Binding of anti-Trypanosoma natural products from African flora against selected drug targets: a docking study

Article *in* Medicinal Chemistry Research · December 2016

DOI: 10.1007/s00044-016-1764-y

CITATIONS
6

READS
282

5 authors, including:

Akachukwu Ibezim
University of Nigeria
29 PUBLICATIONS **84** CITATIONS
SEE PROFILE

Biplab Debnath
Birla Institute of Technology, Mesra
15 PUBLICATIONS **21** CITATIONS
SEE PROFILE

Fidele Ntie-Kang
Martin Luther University Halle-Wittenberg
140 PUBLICATIONS **724** CITATIONS
SEE PROFILE

Chika Mbah
University of Nigeria
37 PUBLICATIONS **109** CITATIONS
SEE PROFILE

*Binding of anti-*Trypanosoma *natural products from African flora against selected drug targets: a docking study*

Akachukwu Ibezim, Biplab Debnath, Fidele Ntie-Kang, Chika John Mbah & Ngozi Justina Nwodo

Medicinal Chemistry Research

ISSN 1054-2523
Volume 26
Number 3

Med Chem Res (2017) 26:562-579
DOI 10.1007/s00044-016-1764-y

Springer

Med Chem Res (2017) 26:562–579
DOI 10.1007/s00044-016-1764-y

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

Binding of anti-*Trypanosoma* natural products from African flora against selected drug targets: a docking study

Akachukwu Ibezim[1] · Biplab Debnath[2] · Fidele Ntie-Kang[3,4] · Chika John Mbah[1] · Ngozi Justina Nwodo[1]

Received: 6 June 2016 / Accepted: 19 December 2016 / Published online: 11 January 2017

Abstract Recent publications have suggested that African medicinal plants and their derived products could be viable source of new and better trypanocidal drugs. Nowadays, in silico methods are often used in drug discovery processes to identify new potential drug leads. In this study, we have developed a small library named Afrotryp, comprising three-dimensional chemical structures of potential trypanocidal compounds derived from medicinal plants in Africa (a total of 321 unique chemical structures). We have predicted the pharmacokinetic properties of the library using Qikprop software and employed the three docking/scoring methods implemented in Molecular Operating Environment Dock Tool to assess the affinity of the library dataset towards the binding site of six selected validated anti-*Trypanosoma* drug targets. It was observed that about 42% of the compounds contained in the Afrotryp dataset were predicted to show a good overall performance in terms of predicted parameters for absorption, distribution, metabolism, elimination and toxicity properties. Docking calculations identified 15 compounds with lowest theoretical binding energies toward the studied proteins, nine of which are suited for the treatment of stage 2 human African trypanosomiasis, due to their low polar surface area. Analysis of their binding modes gave basis for the observed unique molecular interactions which exist between the Afrotryp dataset and the six studied drug targets. The results lay the foundations for the rational development of novel trypanocidal drugs with improved potency.

Keywords African flora · Docking · Drug-likeness · Drug targets · in silico · Trypanosomiasis

Introduction

African trypanosomiasis affects both humans and animals. It is caused by protozoa of the genus *Trypanosoma*. *Trypanosome brucei rhodensiense* and *T. b. gambiense* both cause human African trypanosomiasis (HAT), while *T. b. brucei*, *T. congolense* and *T. vivax* cause nagana in livestock (Maser et al. 2003). The tsetse fly vector (*Glossina* sp.) transmits the parasites from one vertebrate to another (D'Archivio et al. 2011). The disease occurs in 36 sub-Saharan Africa countries. The reported cases of infected persons dropped below 10,000 in 2009 for the first time in 50 years and in 2014 there were 3,796 recorded cases (World Health Organization 2016). Chemotherapy sources of treatment against HAT is poor. Except for Fexinidazole, which is still in phase II/III clinical study in patients with late-stage of the disease, the only four drugs (suramin,

Electronic supplementary material The online version of this article (doi:10.1007/s00044-016-1764-y) contains supplementary material, which is available to authorized users.

✉ Fidele Ntie-Kang
ntiekfidele@gmail.com

✉ Ngozi Justina Nwodo
nwodong@hotmail.com

[1] Department of Pharmaceutical and Medicinal Chemistry, University of Nigeria, Nsukka, Nigeria

[2] Department of Pharmaceutical Sciences and Technology, Birla Institute of Technology, Mesra, Ranchi 835215, India

[3] Department of Pharmaceutical Chemistry, Martin-Luther University Halle-Wittenberg, Wolfgang Langenbeck-Str. 4, 06120 Halle, Germany

[4] Department of Chemistry, University of Buea, Buea, Cameroon

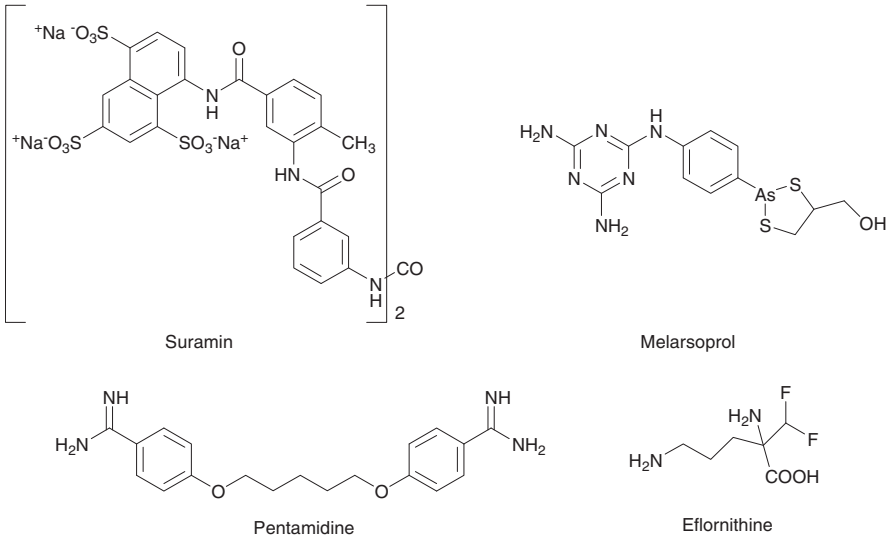

Fig. 1 Drugs approved for first and second stages of HAT treatment

melarsoprol, pentamidine and eflornithine) (Fig. 1) currently used to treat HAT were developed about 30 years ago (Hoet et al. 2004). Besides, each of these drugs has one or more of these challenges; expensive, highly toxic and require parenteral administration (Maser et al. 2003; Amaechi 2001; Kaiser et al. 2011). Therefore, there is an urgent need for new drugs against HAT.

The belief that natural products (NPs) have been optimized over time is the underlying reason why plant metabolites and their derived products are a main focus for new drug development (Koehn and Carter 2005). Recent reviews presented Africa as inhabiting plant materials (and their constituents) with high potentials to treat *Trypanosoma* infections (Ibrahim et al. 2014; Nwodo et al. 2015; Awulu et al. 2013).

Modern drug development procedures using computers and computer software often involve building three dimensional (3-D) virtual compound libraries (databases) for in silico screening, e.g. by docking, binding free energy calculation and pharmacophore-based screening (Paul et al. 2013; Vyas et al. 2013; Fan et al. 2001). Some libraries consist of compounds with activities against several diseases, e.g. the ZINC database (Sterling and Irwin 2015) while others are activity focused libraries, e.g. the naturally occurring plant-based anti-cancer compounds (Mangal et al. 2013). Before any serious time-consuming analysis is performed, the library is usually filtered to eliminate unworthy molecules through a concept referred to as 'rapid elimination of swill' (REOS) (Walters et al. 1996). REOS aids to identify molecules with poor adsorption, distribution, metabolism, elimination and toxicology (ADME/T) properties. REOS is justified by the fact that many drugs fail to make it to the market, after many resources have been invested, because of poor ADME/T performance. Thereafter, virtual screening is carried out by docking the "filtered out" library (or dataset) against validated drug targets, with the view of identifying promising hit compounds, which are then subjected to biological assays.

In this study, we have developed a 3-D virtual library of potential trypanocidal compounds isolated form African medicinal plants (henceforth referred to as Afrotryp). To identify hits, the Afrotryp dataset was analyzed for pharmacokinetic properties and docked towards six selected validated anti-*Trypanosoma* drug targets. Studies of the binding conformations of top scoring compounds, as ranked by each of the three scoring schemes in molecular operating environment (MOE), revealed key molecular interactions formed between the ligands and the amino acid residues at the binding site of the considered anti-*Trypanosoma* drug targets. The binding modes of the ligands could be used to guide further structural modifications in the process of designing novel drugs for HAT with improved activity.

Materials and methods

Data sources

The plants and plant constituents as well as their biological activities were retrieved from literature sources comprising mainly published articles from across the major journals of natural product chemistry, covering the period of 1990–2014. The exact chemical structures of the potential trypanocidal compounds were confirmed from Dictionary of Natural Products (Dictionary of Natural Product 2005). Although some of the compounds in Afrotryp are also found in DNP, the Afrotryp dataset was developed independently of DNP dataset.

Assessment of ADMET profile

3-dimentional molecular structures of the compounds were generated using the graphical user interface of the MOE (Chemical Computing Group 2010) software, energy minimized, using the MMFF94 forcefield (Halgren 1996), until a gradient of 0.001 kcal/mol was reached using MOE software running on a Linux workstation with a 3.5 GHz Intel Core2 Duo processor and saved in .mol2 format. The newly curated Afrotryp dataset was assessed for ADME/T properties using Qikprop software running in normal mode as developed by Jorgensen and Duffy (Duffy and Jorgensen 2000; Jorgensen and Duffy 2000; Schrödinger 2011; Jorgensen and Duffy 2002).

Protein modeling and docking/scoring

Docking procedure with MOE

Protein-ligand docking calculations were performed toward the X-ray crystal structures of the six anti-*Trypanosoma* drug targets discussed in Results and discussion section. Each protein-ligand complex was retrieved from the protein data bank (Berman et al. 2000) and treated thus using MOE software (Chemical Computing Group 2010): the co-crystallized water molecules and non-essential small organic molecules were deleted. The retained protein –ligand complexes were protonated and subjected to energy minimization using the Merck Molecular (MMFF94) Forcefield until a gradient of 0.001 kcal/mol was reached.

Docking and scoring procedure The co-crystallized ligands were removed from the protein and re-docked towards the protein binding sites in three main stages: Conformational Analysis of ligands, Placement of ligands in the protein binding sites and Scoring of the protein-ligand conformations (Edelsbrunner 2007). Conformational analysis of the ligands was performed using the systematic search method to generate conformations from a single 3-D conformation of the respective co-crystallized ligands. Triangle Matcher method was employed in the Placement stage. Triangle Matcher method was preferred because it has been reported to be more systematic in its ligand atoms alignment than other methods like Alpha Triangle. The poses generated from the placement stage were refined using the Forcefield refinement scheme according to its default settings. The scheme makes use of the molecular mechanics set up and the reaction field functional form to perform energy minimization and to treat solvent effect respectively. Scoring was carried out using the three scoring functions implemented in MOE which are Affinity dG, Alpha HB and London dG. Affinity dG scoring function estimates the enthalpic contribution to the free energy (G) of binding by summing up the energies resulting from the following interactions: hydrogen bond donor–acceptor pair (hb), ionic interactions (ion), metal ligation (mlig), hydrophobic interation (hh), and interactions between hydrophobic and polar atoms (hp) and between any two atoms (aa), according to the equation given below:

$$G = C_{hb}f_{hb} + C_{ion}f_{ion} + C_{mlig}f_{mlig} + C_{hh}f_{hh} + C_{hp}f_{hp} + C_{aa}f_{aa}$$

where the f terms fractionally count atomic contacts of specific types and the C's are coefficients representing weighted term contributions to affinity estimation.

Alpha HB method scores a ligand by measuring the geometric fit of the ligand to the binding site and the H-bond effect. The free energy of binding of a given pose is estimated by the London dG scoring function from the sum of each of these terms: average gain/loss of rotational/translational entropy, energy due to loss of flexibility of the ligand, H-bond energies, energy due to metal-ligation and desolvation energy of atoms. Docking validation was done in an attempt to identify the best docking parameters which reproduce the experimentally determined ligand conformations within the binding pocket. Five independent docking scores were carried out and the average taken. The accurate poses were assessed using the lowest root-mean-square-deviation (RMSD) values relative to the crystal structure binding poses. Grid dimensions producing ligand conformations closest to the experimental poses (Table 5) were subsequently employed in docking the Afrotryp dataset against the binding sites of the six targets by performing five independent runs and taking the average.

Results and discussion

Current thinking in drug development involves building of a 3-D virtual library. This library serves as a database for the search of viable new drug molecules. Sometimes, a library can be made up of compounds of various activities. At another time, it could be activity focused library (Sterling and Irwin 2015; Mangal et al. 2013). We have developed the later for compounds with activities against *Trypanosoma* infections. It comprises of 321 natural and derived products which have been isolated from medicinal plants harvested from the African continent. The number of plants studied spread across 22 plant families. Most of the plants belong to Lamiaceae (15.31%), followed by Dioncophyllaceae (12.19%), Clusiaceae and Bignoniaceae (7.81%) (Fig. 2a). Figure 2b shows that the plant part most studied and from which the potential anti-*Trypanosoma* compounds were isolated was leaves (36.76%) followed by roots and stems (26.79 and 23.05%, respectively). Examination of the

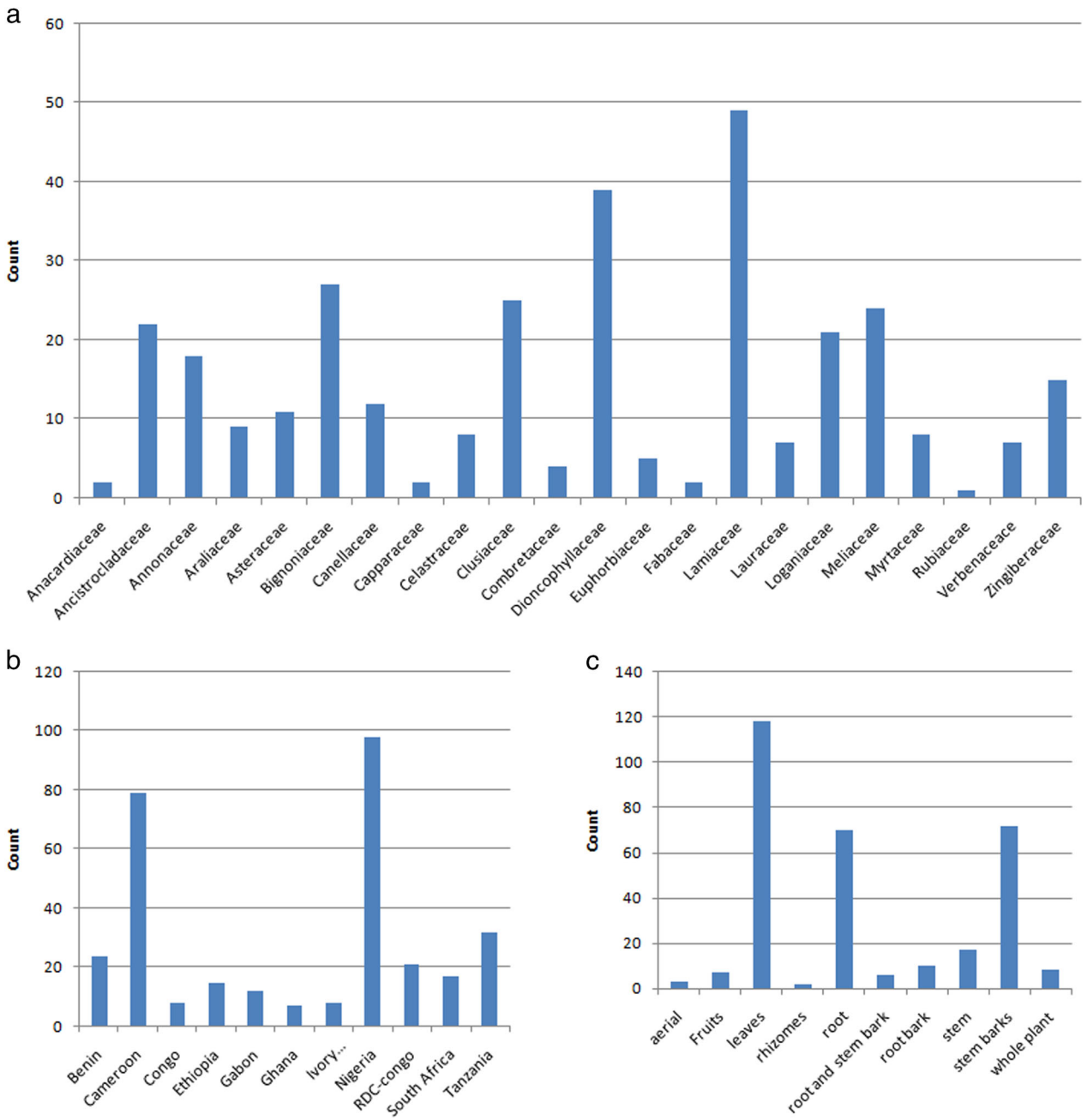

Fig. 2 Bar chart showing **a** plant families of origin **b** distribution of plants within their source countries **c** count of plant part yielding the compounds in Afrotryp (color figure online)

database revealed that West Africa, especially Nigeria and Cameroon (30.53 and 24.62%, respectively) harbor most of these plants (Fig. 2c). This finding confirms the views of Ibrahim and co-workers and Nwodo and team, which that reported that; plants obtained from these countries demonstrated high trypanocidal activities (Ibrahim et al. 2014; Nwodo et al. 2015). In fact, 800 mg/kg leaf extract of one of the plants (*Dissotis rotundifolia*) harvested from Nigeria, killed *T. b. brucei* in vitro within 45 s of incubation.

Predicted absorption property of Afrotryp dataset

Most drugs are administered orally because it is easier to do so. The major challenge faced in designing oral dosage forms lies with poor bioavailability because for orally administered drug to be efficiently absorbed from the gastrointestinal tract, it has to possess appropriate physicochemical properties. Nowadays, it is a common practice to assess how orally bioavailable drug molecule is in silico

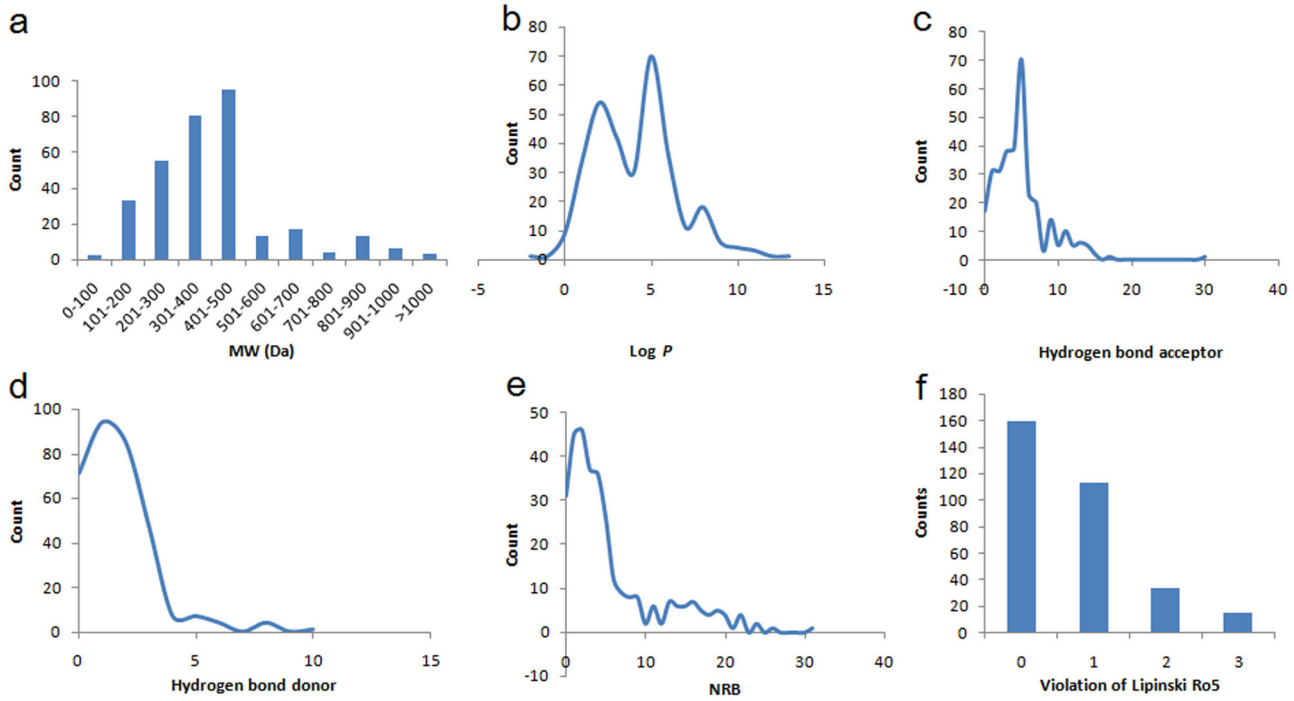

Fig. 3 Distribution of parameters used to assess "drug-like" property and Ro5 violation of Afrotryp compounds: **a** MW, **b** Log *P*, **c** HBA, **d** hydrogen bond donor, **e** NRB, **f** violation of Ro5 (color figure online)

early enough in the of drug discovery process, so as not to waste resources on molecules which would not be bioavailable, after all, in the biological system. The assessment is done according to Lipinski's rule of five (Ro5). As a rule of thumb, orally absorbed (or drug-like) molecules tend to obey this rule (Lipinski et al. 1997). Although, Lipinski had initially postulated that the Ro5 was not respected by NPs, the work of Feher and Schmidt showed they do (Feher and Schmidt 2003). Lipinski analyzed compounds from the World Drugs Index database to arrive at the Ro5 which states that a drug candidate will be orally bioavailable if it the following features: molecular weight (MW) < 500 Da, hydrophobicity (log *P*) < 5, hydrogen bond acceptor and donor (HBA and HBD), respectively ≤10 and ≤5. According to Lipinski, molecules which satisfy above criteria are "drug-like". The importance of molecular flexibility, which depends on the number of rotatable bond (NRB), to oral bioavailability of drug molecule, was reported by Verber and co-workers (Veber et al. 2002). They proposed that NRB of <5 is necessary for a molecule to be drug-like.

Given in Fig. 3, is the distribution of various properties, which determine drug-likeness of molecules. Counts verses property was plotted to represent the actual number of compounds involved. Figure 3a showed that 265 out of 321 of the Afrotryp dataset fall within the recommended region for MW according to the Ro5 with highest population of the compounds having MW in the range of 401–500 Da. A total of 171 compounds (about 54.0%) respected the log *P* criterion. The distribution of the log *P* values showed a rough Gaussian curve with a peak centered at 5.0 log *P* units (Fig. 3b), which differs from the performance of other libraries consisting of natural and derived products from African medicinal plants (Ntie-Kang et al. 2013a, 2013b, 2013c). The nature of Afrotryp library (which is activity focused while others are not) might account for the discrepancy in log *P* performance. However, the library datasets generally have good log *P* property with 13.72 as the highest recorded value. Although we discuss log *P* under absorption, its effects cut across all ADME/T properties. The total number of the datasets which possess acceptable HBA, HBD and NRB according to Lipinski, are 291, 311 and 221, respectively. On the overall, about 152 of Afrotryp compounds respected all the five criteria that Lipinski used to predict drug-likeness (Fig. 3c, d and e). Furthermore, the result of Lipinski violation presented in Fig. 3f demonstrates that; about 50% of the compounds showed no Lipinski violations, while 85% showed <2 violations. This suggests that the library is enriched with natural and derived products which are drug-like. Therefore, hits/leads obtained from Afrotryp database would likely be bioavailable in systemic circulation. Figure 4a–d are scatter plots showing the relationship between MW vs. log *P*, HBA, HBD and NRB, respectively. The correlation between MW and other relevant properties further validates the fact that Afrotryp library contains high number of compounds which are

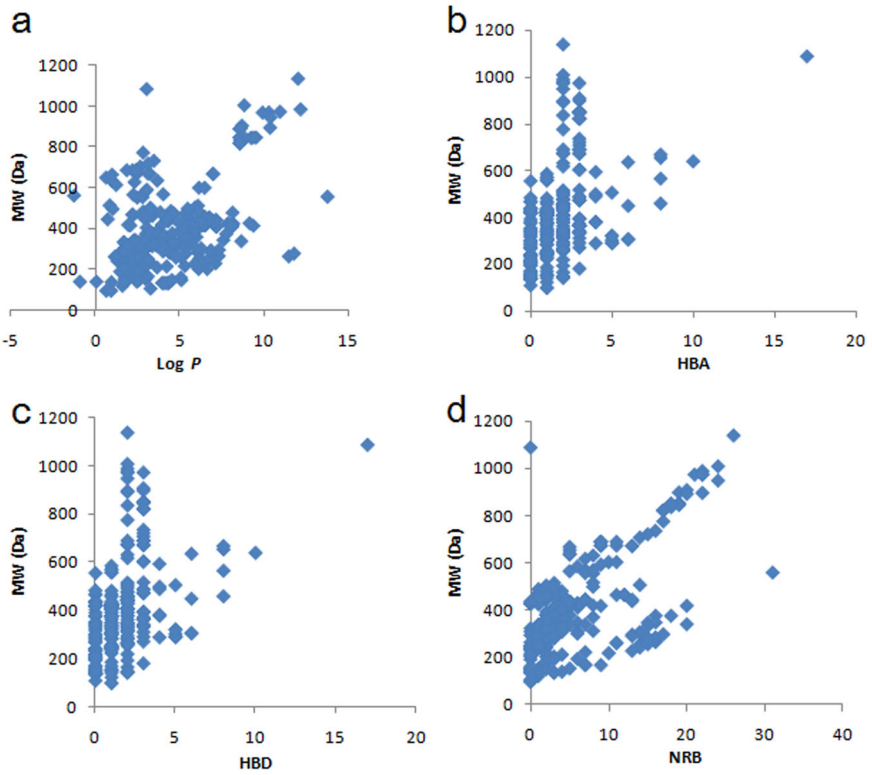

Fig. 4 Scatter plot showing distribution of **a** MW against log *P*, **b** MW against HBA, **c** MW against HBD and **d** MW against NRB. The rectangular box in each plot represents Lipinski regions of compliance

predicted to be drug-like. In each of the plots, it was observed that; the region of highest population density lie within Lipinski's compliance region. In the light of the forgoing results, trypanocidal molecules obtained from the library would most probably be orally bioavailable.

A bar chart representation of more physical properties is shown in Fig. 5a. The Caco-2 monolayer ($BIP_{caco\text{-}2}$) is widely used across the pharmaceutical industry to model: human small internal mucosa and variety of trans-cellular pathways to predict the absorption of orally administered drugs as well as metabolic transformation of test substances (Artursson and Karlsson 1991). The importance of drug solubility (log S_{wat}) is underlined by the fact that drugs are absorbed in the form of solution at the site of absorption and the ones with desired solubility property are usually in good concentration in systemic circulation (Myrdal and Yalkowsky 2007; Vemula et al. 2010). The Caco-2 monolayer and solubility properties, alongside some number of primary metabolites, are physicochemical parameters Jorgenson used to propose the "rule of three" (Ro3). Ro3 considers that a molecule will be orally bioavailable when it respects at least two of the following criteria: log $S_{wat} > -5.7$, $BIP_{caco\text{-}2} > 22$ and number of primary metabolites < 7 (Schrödinger Press 2011). From the result of analysis given in Table 4, 258 (80.37%) and 316 (98.44%) of the Afrotryp compounds possess the general acceptable log S_{wat} and $BIP_{caco\text{-}2}$ properties, respectively required forgood drug absorption. While according to Jorgenson's test, only 117 (38.11%) and 292 (95.11%), respectively respected Ro3 for oral bioavailability. The excellent performance of a good number of Afrotryp dataset with regards to $BIP_{caco\text{-}2}$ property underscores the benefits of the library because since the acknowledgement of the significance of $BIP_{caco\text{-}2}$ in drug development process, percentage failure of drug development on account of poor pharmacokinetic properties has drastically dropped (Artursson 1990). Like $BIP_{caco\text{-}2}$, Madin–Darby canine kidney (MDCK) with a compliance zone of 25 to −500 nm/s,also serves as a model for predicting oral absorption of drug because MDCK cells express more of transporter proteins but very low metabolizing enzymes (Irvine et al. 1999). 296 of the compounds, amounting to 92.21% of the Afrotryp dataset, complied with MDCK criterion. Dermal penetration (log K_p) was computed though there is no HAT drug applied on the skin at the moment. It was observed that 280 (87.23%) of Afrotryp compounds fall within the recommended range for log K_p(−8 to −1) (Potts and Guy 1995). The predicted human oral absorption result displayed in Fig. 5b showed that 161 (52.44%) compounds of Afrotryp database will have high oral absorption in human on a scale of 1, 2 and 3 for low,

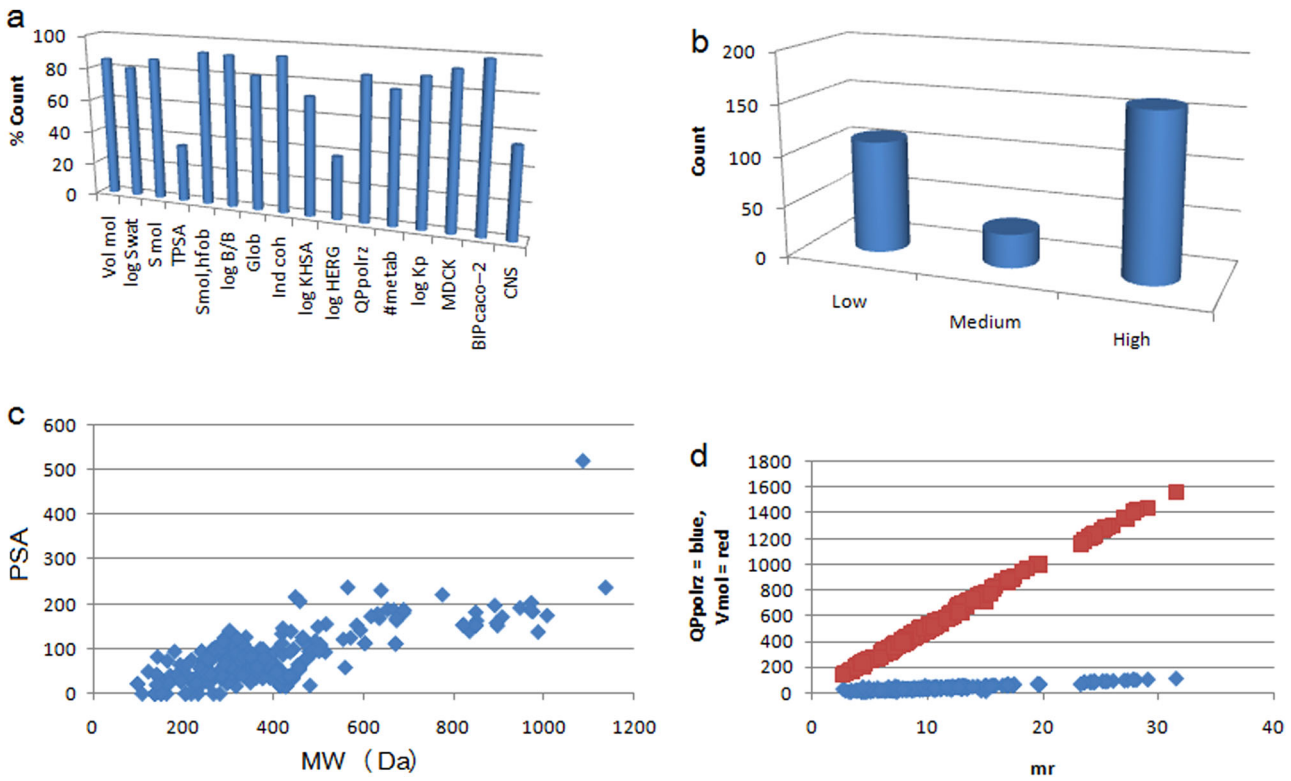

Fig. 5 Plots of Physicochemical properties **a** percentage count vs. ADMET parameters **b** count vs. human oral absorption **c** PSAvs MW and **d** relationship between polarizability and volume against mr (color figure online)

medium and high human oral absorption, respectively. Molar refractivity, which is largely affected by London dispersion forces, of a compound is an essential factor in drug-receptor interaction processes. A strong linear relationship is observed to exist between this property, volume and molecular polarizability (Ghose et al. 1999). Figure 5d demonstrates that Afrotryp dataset complied with this suggested relationship. It was observed that 85.05 and 85.67 percentages of the compounds in the database fall within the acceptable limit for molecular volume (500–2000 $Å^2$) and polarizability (13.00–70.00) properties.

Comparison of drug-likeness of Afrotryp and the dictionary of natural products libraries

It is preferred that the drug-like properties of new compound libraries are enhanced over the existing ones. This is because there is higher probability of obtaining hits with good oral bioavailability profile from such libraries with greater number of drug-like compounds.

DNP is a comprehensive database of natural products containing 126,140 unique compounds. As in our past study (Ntie-Kang et al. 2014), a comparison of the anti-trypanosomal NP library has been carried out with the much larger DNP as shown in Fig. 6 and Table 1. The oral bioavailability performances of the datasets of the two librariesare presented as percentage count. The MW distribution of Afrotryp dataset peaked at interval (401–500 Da) similar to ConMedNP (Ntie-Kang et al. 2013a), a dataset of natural products from Congo Basin, while that of DNP peaked at 301–400 Da. The percentages of compounds in the two databases that satisfy Lipinski's MW property (<500 Da) were 74.15 for DNP and 82.81 for Afrotryp. Table 1 shows that, in the overall, Afrotryp library have more compounds that are orally bioavailable than the ones in DNP library by 10.99%. Though both datasets peaked at the same log P unit (log $P = 2$ unit), the DNP datasets performed significantly better than Afrotryp dataset in all the ranges examined. 4.99 and 15.03% of Afrotryp dataset are enhanced for HBA and HBD, respectively over DNP's. In the light of the above observation, it can be concluded that the small Afrotryp database contain more compounds, except for log P property, with improved drug-like property than those in DNP.

Lead compounds usually have a smaller number of rings, less NRB, smaller MW and are less lipophilic (Quinn et al.

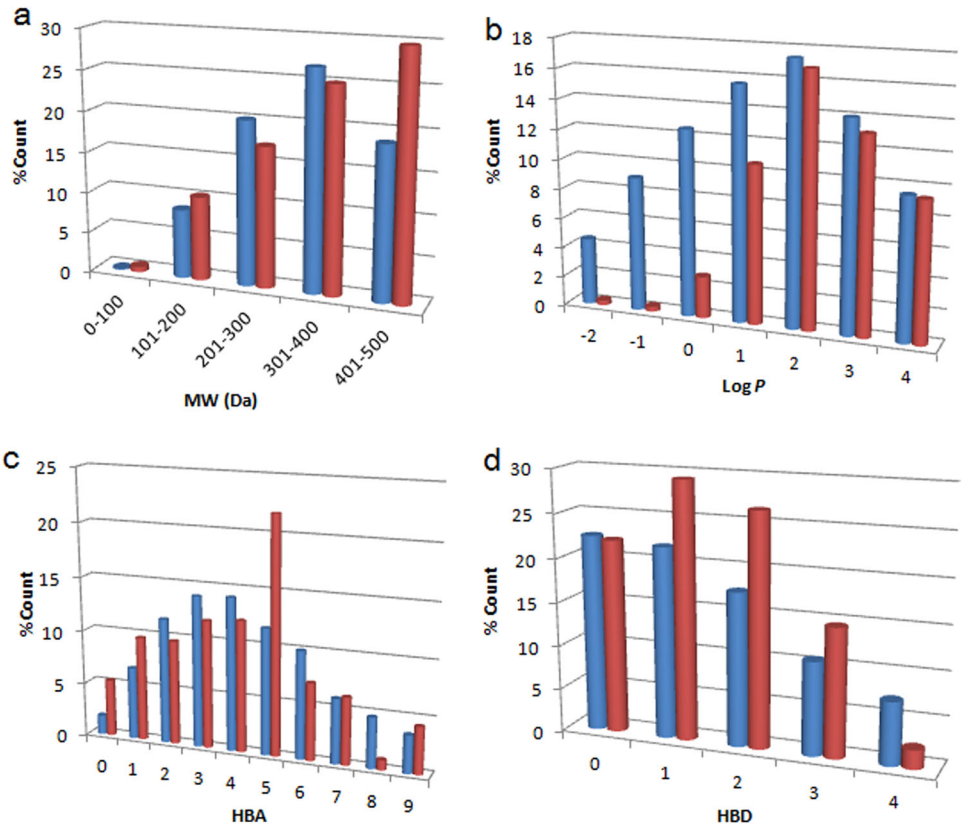

Fig. 6 Plots of percentage performance of Afrotryp and DNP compounds for oral bioavailability parameters according to Ro5. **a** MW **b** log *P* **c** HBA **d** HBD. In all cases, DNP is shown in *blue* and Afrotryp in *red* (color figure online)

Table 1 Percentage performance of Lipinski parameters for the two datasets

Property	DNP	Afrotryp	Difference
MW	74.15	82.81	10.99
Log *P*	82.5	53.27	29.23*
HBA	84.11	89.1	4.99
HBD	79.67	94.7	15.03

The value marked * is for the difference in which DNP is enhanced over Afrotryp dataset

2008). Therefore criteria for "lead-likeness" and "fragment-likeness"; ($150 \leq MW \leq 350$, log $P \leq 4$, $HBD \leq 3$ and $HBA \leq 6$) and ($MW \leq 250$, $-2 \leq \log P \leq 3$, $HBD < 3$, $HBA < 6$ and $NRB < 3$), respectively, have also been proposed, which have slightly more stringent criteria than the Lipinski's Ro5. 138 and 62 compounds of the library have physicochemical parameters falling within the intervals of the two criteria respectively (Table 2) (Teague et al. 1999; Oprea 2002; Schneider 2002). The challenge for lead identification is a major obstacle to overcome in drug discovery. Therefore, the high presence of lead like compounds in Afrotryp library highlights the importance of the library as a resource for the search of new trypanocides.

Drug distribution

Once a drug has been absorbed, it is distributed around the blood supply, then to various tissues and organs. Some physical properties, among other factors, determine the rate and extent of distribution. Second stage of HAT begins when Trypanosome crosses blood-brain-barrier (BBB) and invades central nervous system (CNS). Therefore, anti-trypanosomal drugs intended for the treatment of the second stage of the disease are required to possess properties which will enable them to cross the BBB and act on CNS (Kirchhoff 2000). The criteria for crossing BBB have been defined as: log B/B = −3 to +1 and CNS ≥ 0 (Luco 1999; Ghose et al. 2012; Palm et al. 1997). It was observed that 91.90 and 53.90% of the Afrotryp compounds respected log B/B and CNS criteria. The parameter, total polar surface area (TPSA), has been shown to correlate well with log B/B. Drugs with TPSA values 60–140 $Å^2$ tend to permeate the BBB (Ertl et al. 2000; DesJarlais et al. 1988). Although a significant small percent (34.58%) of the studied compounds possess the TPSA recommended property, their high log B/B performance suggest that they may have the ability to crossed the BBB with the aid of carrier proteins and/or through pinocytosis (Table 3).

Table 2 Distribution of oral bioavailable properties of Afrotryp compounds

Property	Drug-like	Lead-like	Fragment-like	Maximum	Minimum	Mean
MW	265	138	62	1137.33	98.15	403.92
Log *P*	172	142	99	13.72	−1.28	4.62
HBA	291	249	226	30	0	4.94
HBD	311	297	250	17	0	1.72
NRB	221	159	122	31	0	5.69

Table 3 Results of analysis of parameters for ADMET property prediction

Property	Symbol	Number of compliance	Percentage compliance	Max.	Min.	Mean	Acceptable range
The total volume of molecule enclosed by solvent molecular in Å^3	Vol_{mol}	273	85.05	3321.48	437.13	1205.5	500–2000
The logarithm of aqueous solubility	log S_{wat}	258	80.37	−0.11	−15.12	−5.03	−6 to +5
The total solvent accessible molecular surface, in Å^2	S_{mol}	278	86.6	1591.19	293.47	639.03	300–1000
The topological polar surface area in Å^2	TPSA	111	34.58	521.73	0	71.91	60–140
The hydrophobic portion of the solvent-accessible molecular surface, in Å^2	$S_{mol,hfob}$	297	92.52	1229.11	145.96	514.43	0 to 750
The logarithm of predicted blood/brain barrier partition coefficient	log *B*/*B*	295	91.9	1.74	−8.65	−0.77	−3 to +1
The globularity	Glob	260	81	0.98	0.64	0.8	0.75–0.95
The index of cohesion interaction in solids	Ind_{coh}	299	93.15	0.12	0	0.01	0–0.05
The logarithm of predicted binding constant to human	log K_{HSA}	229	71.34	3.72	−1.69	0.65	−1.5 to 1.2
The predicted IC_{50} value for blockage of HERG K+ channels	log HERG	121	37.69	1.1	−8.95	−4.25	< −5.0
The predicted polarizability	QPpolrz	275	85.67	113.06	10.69	39.45	13.0–70
The number of likely metabolic reaction	#metab	253	78.82	25	0	5.61	1 to 8
The predicted skin permeability	log K_p	280	87.23	0.17	−13.30	−2.56	−8 to −1
The predicted apparent Madin-Darby canine kidney cell permeability in nm/s	MDCK	296	92.21	10,000	0.00	1124.73	25 to −500
The predicted apparent Caco-2 cell membrane permeability, in nm/s	BIP_{caco-2}	316	98.44	9906.04	0.00	1958.43	5 to >100
The central nervous system	CNS	173	53.9	2	−2	−0.41	≥0
Compliance to "drug-like" property	#star	129	42.43	17	0.00	2.76	0

Worthy of note is the fact that, compounds with TPSA good property also have the required MW as shown in Fig. 5c.

Molecular shape of a ligand (drug) affects how well it binds to its biological target. This property is described using the descriptor—Molecular Globularity (Glob) (Meyer 1986). 0.75–0.95 Glob value is the acceptable limit for good performance. Based on that, 260 (81.00%) of Afrotryp dataset have Glob values within the acceptable range and so can efficiently interact with the various *Trypanosoma* drug targets.

It was observed that 71.34% of the datasets can comfortably pass through the blood capillaries without binding to the blood plasma proteins. This is because these compounds have the recommended property of logarithm of predicted binding constant to human serum albumin (log K_{HSA}: −1.5 to 2.0) (Colmenarejo et al. 2001). Drugs bound to the proteins will no longer be available to reach and interact with its target. Therefore, the high percentage of the datasets with recommended log K_{HSA}value, again raises the importance of Afrotryp library as a good point to

kick-off the search for molecules to treat *Trypanosoma* infections.

Drug metabolism and toxicity

Metabolic enzymes in the body attack drugs to convert them to metabolites most of which are inactives. Therefore, it is preferred that drugs undergo less number of metabolic reactions (#metab). The performance of Afrotryp dataset with regards to #metab spanned from 0 to 25 with a mean value of about 5. In general, 78.82% of the studied compounds undergo the recommended limit of #metab (1–8) (Ntie-Kang et al. 2014). Drug toxicity is a great concern to the medical world. Chemical substances which block or cause the blockage of human ether-a-go-go-related gene potassium ion (HERG K^+) channel are potential toxic substances, because HERG K^+ serves as a molecular target for cardiac toxicity of many drugs (Cavalli et al. 2002). The recommended range of log HERG is < -5.00 (Aronov 2005). Notwithstanding that less than half of the compounds in the library respected log HERG criterion, their mean performance (log HERG = −4.25) fall within the acceptable limit. This indicates that, on the average, more of the datasets will not block or cause the blockage of human ether-a-go-go-related gene potassium ion.

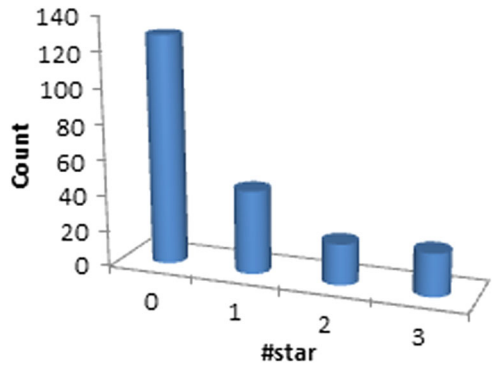

Fig. 7 Bar chart showing #star performance of the Afrotryp dataset. #star means the overall ADME-compliance score—drug-likeness parameters (color figure online)

Afrotryp dataset's overall ADME-compliance performance

Qikprop software evaluates drug-like and pharmacokinetic profile of molecules using a set of 24 descriptors. The overall ADME-compliance-performance of a compound represented as #star shows the number of computed property descriptors which fall outside the recommended rangeof values for 95% of known drugs. Therefore, #star of zero is recommended (Aronov 2005). Thus, as shown in Fig. 7, 42.43% of the compounds in Afrotryp library demonstrated compliance to #star criterion. Meaning that about half of the datasets in the library have the recommended pharmacokinetic property for a good drug molecule.

Docking calculations

Selected targets used in docking studies

This current study considers the protein-ligand interactions of the compounds within the Afrotryp dataset with six selected validated anti-trypanosomal drug targets with available 3D crystal structures in the PDB. Table 4 shows the summary of the studied targets, which include uridylyltransferase (UTF), an enzyme which is vital for modulating non-coding RNA pathways in pathogen systems by catalyzing the synthesis of diphospate and uridylyl-[protein II], has been co-crystallized and characterized with the substrate—Uridine 5'-triphosphate (PDB code 2B56) (Deng et al. 2005; Demir et al. 2014). Many works have reported farnesyldiphosphate synthase (FDS; PDB code 2EWG) as a

Table 4 Description of the 6 anti-*Trypanosoma* drug targets used in this study

PDB code	Protein target	co-crystallized ligand	Resolution of protein X-ray crystal Structure (Å)	Reference
2B56	UTF	uridine-5'-triphosphate	1.97	Deng et al. 2005
2EWG	FDS	α,β-(1-hydroxy-2-imidazo[1,2-α]pyridine-3-ylethine-1,1-diyl)bis(phosphonic acid)	2.48	Mao et al. 2006
2TOD	ODC	α-difluoromethylornithine	2	Grishin et al. 1991
2XTB	adenosine kinase	4-[5-(4-phenoxyphenyl)-1H-pyrazol-3-yl]morpholine	2.8	Kuettel et al. 2011
4H6O	sterol 14α-demethylase	α-1-[3-(4-chloro-3,5-dimethylphenoxy)	2.8	Andriani et al. 2013
4TIM	TIM	2-phosphoglyceric acid	2.4	Noble et al. 1991

Table 5 Retained grid parameters used in the docking protocols, the lowest RMSD values in Å of the docked ligands against the X-ray ligand and the corresponding docking scores of the co-crystallized ligands into thebinding sites of the six (6) studied trypanocidal drug targets

PDB code	Scoring scheme	Origin			Radius			RMSD (Å)	ΔG (Kcal/mol)
		X	Y	Z	X	Y	Z		
2B56	Affinity dG	11.20	36.94	14.89	2.90	6.60	5.20	2.43	−8.08
	Alpha HB	11.20	36.94	14.89	2.90	6.60	5.20	2.91	−10.99
	London dG	11.20	36.94	14.89	2.90	6.60	5.20	2.27	−10.17
2EWG	Affinity dG	26.50	0.18	9.60	3.90	2.10	4.00	2.67	−10.06
	Alpha HB	26.50	0.18	9.60	3.90	2.10	3.80	0.69	−9.95
	London dG	26.50	0.18	9.60	3.90	2.29	3.80	0.79	−16.12
2TOD	Affinity dG	24.98	24.96	−2.13	31.79	41.73	22.13	3.01	−8.31
	Alpha HB	20.98	27.96	−4.13	5.95	3.73	2.13	3.43	−7.52
	London dG	21.98	28.55	−4.23	4.50	3.00	2.00	1.16	−6.16
2XTB	Affinity dG	18.41	−28.30	6.50	6.83	3.90	4.10	2.06	−9.35
	Alpha HB	18.41	−28.30	6.50	6.83	3.90	4.10	2.07	−8.93
	London dG	18.41	−28.30	6.50	6.83	3.90	4.10	0.89	−8.48
4H6O	Affinity dG	2.00	28.00	17.90	4.40	5.00	4.10	2.41	−10.02
	Alpha HB	2.00	28.00	17.90	4.40	5.00	4.10	3.06	−6.49
	London dG	2.00	28.00	17.90	4.40	5.00	4.10	2.95	−6.55
4TIM	Affinity dG	47.00	14.00	−10.00	2.00	5.00	5.00	1.96	−6.58
	Alpha HB	47.00	14.00	−10.00	2.00	5.00	5.00	1.91	−8.04
	London dG	42.44	9.26	−6.04	26.72	23.51	24.12	0.92	−10.54

Table 6 The mean, minimum and maximum docking scores recorded for compounds in Afrotryp library using the three scoring methods implemented in MOE, docking scores of the native ligands (NL) and the number of Afrotryp compounds which scored better than the NL

PDB code	Scoring methods	Min. ΔG	Max. ΔG	Mean ΔG	ΔG of native ligand (NL)	No compounds with ΔG > NL
2B56	Affinity dG	−20.76	−3.92	−12.69	−8.08	298
	Alpha HB	−22.31	−7.96	−16.22	−10.99	303
	London dG	23.94	−6.79	−15.72	−10.17	299
2EWG	Affinity dG	−18.33	0.03	−6.99	−10.06	231
	Alpha HB	−20.18	−6.68	−7.86	−9.95	309
	London dG	−21.02	−5.97	−8.87	−16.12	76
2TOD	Affinity dG	−22.20	−5.86	−11.28	−8.31	288
	Alpha HB	−17.64	−3.62	−14.25	−7.52	278
	London dG	−14.34	−1.80	−14.02	−6.16	286
2XTB	Affinity dG	−15.08	−2.44	−11.32	−9.35	139
	Alpha HB	−14.90	−5.52	−10.98	−8.93	237
	London dG	−14.36	−5.64	−16.22	−8.48	239
4H60	Affinity dG	−18.76	−3.30	−8.99	−10.01	220
	Alpha HB	−22.24	−1.39	−10.08	−6.49	283
	London dG	−30.29	−6.11	−9.69	−6.55	319
4TIM	Affinity dG	−15.70	0.38	−8.30	−6.58	215
	Alpha HB	−16.83	−6.71	−12.54	−8.04	319
	London dG	−20.64	−4.86	−14.29	−10.54	294

therapeutic target in *T. brucei* (Montalvetti et al. 2003; Linares et al. 2006). This enzyme catalyzes the formation of farnesyldiphosphate in the isoprenoid biosynthetic pathway, an important precursor for the prenylation of certain proteins and synthesis of unbiquinones, dolichols, sterols and heme a (Mao et al. 2006). Grishin et al. (1999) described the X-ray

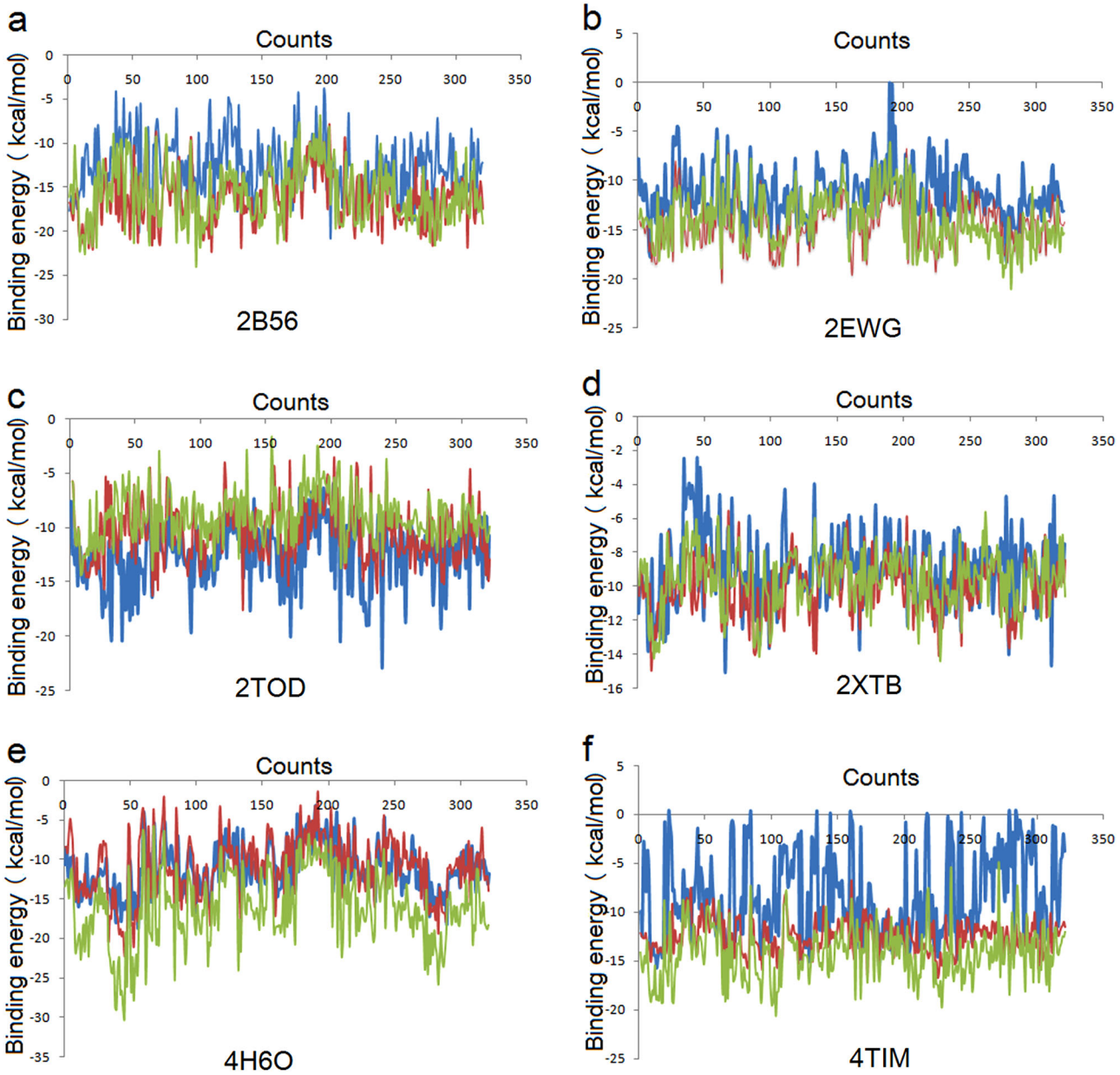

Fig. 8 Docking Scores of the compounds in Afrotryp library using the three scoring schemes in MOE toward the binding sites of the six studied targets. In each graph, the color *blue* = Affinity dG, *red* = Alpha HB and *green* = London dG (color figure online)

structure of ornithine decarboxylase (ODC) from *T. brucei* in complex with alpha-difluoromethylornithine (PDB code 2TOD). ODC catalyzes the decarboxylation of ornithine to produce putrescine, an essential precursor for the synthesis of higher polyamines (spermidines and spermines), which are vital for the normal and neoplastic cell growth (Deng et al. 2008; Pegg 1988). Adenosine kinase is considered as a drug target for chemoprevention of trypanosomiasis because the causative organism of the disease, trypanosomes, lacks de novo purine synthesis and solely depends on salvage from their host (Kuettel et al. 2011). Adenosine kinase (PDB code 2XTB) is a key enzyme in the purine salvage pathway, catalyzing the transfer of gamma-phosphate from adenosine triphosphate to adenosine leading to the formation of adenosine monophosphate (Kuettel et al. 2009). Sterols, which are essential lipid components of eukaryotic membranes, are biosynthesized by the catalytic action of sterol 14α-demethylase (PDB code 4H6O). This enzyme has been considered as a target for anti-fungal and anti-protozoan chemotherapy (Andriani et al. 2013; Lepesheva et al. 2006). The last in the list of enzymes included in this study is triose phosphate isomerase (TIM; PDB code 4TIM) that catalyzes

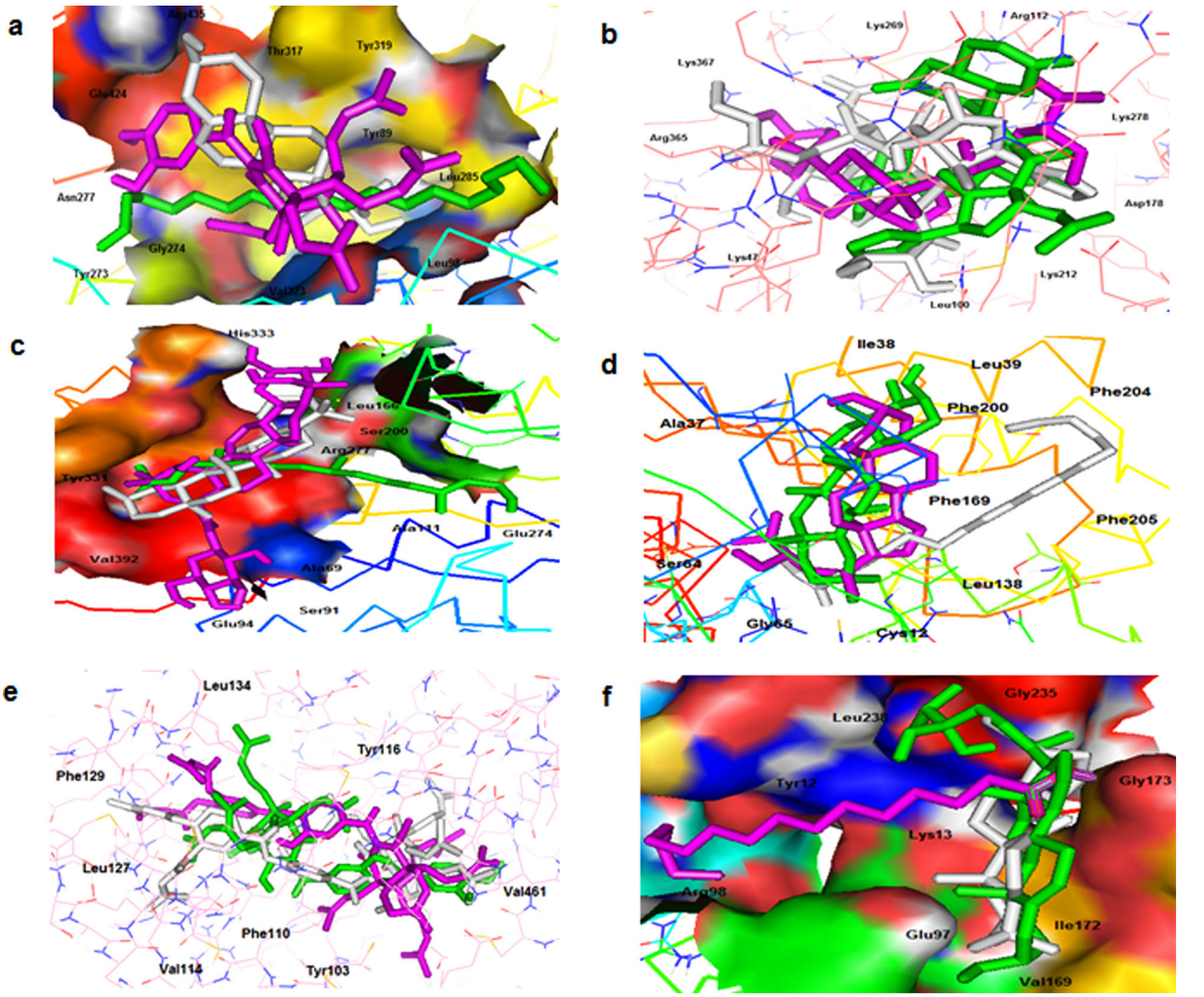

Fig. 9 Binding interactions of the highest ranked compounds from Afrotryp library towards the six studied proteins according the three scoring methods implemented in MOE. The top ranked candidates according to Affinity dG, Alpha HB and London dG scoring methods are coloredin *green*, *magenta* and *white*, respectively. The proteins with PDB codes 2B56 = A, 2EWG = B, 2TOD = C, 2XTB = D, 4H6O = E and 4TIM = F (color figure online)

the interconversion between glyceraldehyde 3-phosphate and dihydroxyacetone phosphate in the glycolytic pathway is another target employed in this study (Noble et al. 1991; Bakker et al. 1999). The work of Olivares-Illana et al., which reported that there are structure differences in mammalian and trypanosomal TIM and the dependency of bloodstream form of *T. brucei* on glycolytic pathway for energy metabolism, prove that TIM is a target of choice for the treatment of trypanosomiasis (Olivares-Illana et al. 2006).

Validation of docking protocol using MOE

The main results have been shown in Table 5. Docking protocols used in this study for each of the scoring methods implemented in MOE (Affinity dG, Alpha HB and London dG) were validated as described in the experimental section. The recorded RMSD values between the X-ray crystallized and the docked ligands ranged from 0.69 to 3.43 Å with only three out of eighteen determined RMSD values exceeding 3.00 Å. The parameters for the centroids and dimensions of the retained grid boxes which gave RMSD values > 3.00 Å (that is the RMSD for 2TOD Affinity dG = 3.01 Å, Alpha HB = 3.43 Å and 4H6O Alpha HB = 3.06 Å) were used nonetheless, because they can be taken as 3.00 Å on approximation and The London dG and Alpha HB docking methods gave the most interesting docking protocols for the native ligands of the 4TIM complex (RMSD = 0.69 and 0.70 Å, respectively). The mean RMSD values observed for each of the three docking methods for all the six complexes are 2.42, 2.35 and 1.50 Å for Affinity dG,

Table 7 Some physicochemical properties of the top scoring compounds (toward 2B56 target as separately recorded by Alpha HB, Affinity dG and London dG scoring functions) toward the six studied proteins

Codes of top scoring compounds	MW	NRB	HBA	HBD	LV	LogP	TPSA
BEN-008	426.73	0	1	1	1	7.82	35.90
BEN-010	414.72	6	1	1	1	8.07	51.14
BEN-013	428.70	6	2	2	1	6.99	105.14
CMR-046	971.24	22	11	2	3	9.88	245.12
CMR-056	821.06	17	9	3	2	8.53	259.90
CMR-066	428.75	0	1	1	1	9.11	31.84
CMR-103	348.53	4	3	0	1	4.51	106.16
CMR-111	602.81	11	6	3	2	6.45	202.63
ETH-132	652.82	5	11	8	3	0.61	324.69
NIG-172	256.43	14	2	1	1	6.14	100.68
NIG-203	268.49	16	1	0	1	7.28	61.28
NIG-227	284.48	16	2	0	1	6.85	74.42
NIG-239	296.54	13	1	1	1	6.82	75.01
SAF-278	672.72	11	13	2	2	2.45	273.55
SAF-281	734.84	16	14	3	2	3.44	247.99

Alpha HB and London dG, respectively. Unlike our earlier work (Ntie-Kang et al. 2014), London dG scoring function appears to be the most efficient docking method for this current study since it gave the least RMSD values in all docking protocols for the studied protein targets except for (PDB codes: 2EWG and 4H6O).

MOE docking calculations

Each of the three docking methods/scoring functions implemented in MOE software suite was used to predict the binding free energy of interaction between the compounds in the Afrotryp dataset with the eight validated anti-*Trypanosoma* protein targets. The lowest binding affinity for each compound was selected from the recorded five top scoring poses and used to plot the distributions shown in Fig. 7 (also included as a supplementary file). The docking scores of the three docking methods showed that the compounds of the Afrotryp dataset take unique binding modes within the binding site cavities of the proteins characterized mainly by hydrophobic or lipophilic interactions.

The results shown in Table 6, indicate that the docking scores of the datasets toward each of the studied protein targets fall within the same range of <3 kcal/mol difference except for the targets with PBD codes 2XTB and 4TIM which have differences of about the double of −3; −5.24 and −5.99 Kcal/mol, respectively.

An analysis of the plots of the binding affinity distributions of the Afrotryp dataset toward the targets (Fig. 8) show, on the general, that the three docking schemes implemented in MOE software suite scored the compounds in a similar pattern but in varying energy grades. It was observed that the Affinity dG docking method assigned the lowest binding affinities to the greater proportion of the compounds docked toward the targets with PDB codes 2B56, 2EWG, 2M9H, 2XTB and 4TIM. On the other hand, as we have observed and reported in our work (Ntie-Kang et al. 2014), *Alpha HB* scoring function gave higher proportion of the compounds with lower docking scores than the reference ligands for five targets (2B56, 2EWG, 4TIM and 2XTB) when compared with the other two scoring functions.

Binding mode of top scoring compounds

Docking calculation results revealed that the Afrotryp dataset adopted preferential conformations within the binding cavity of the studied proteins. The three scoring methods have different compounds as the top scoring candidate for a protein target in all the cases except for sterol 14α-demethylase (PDB code: 4H6O) where *Alpha HB* and *London dG* have the same compound as the top scoring candidate.

The best docked conformations of the three top scored compounds (octadecanal (NIG-203): $\Delta G = -20.76$ Kcal/mol, xanthochymol (CMR-111): $\Delta G = -22.31$ Kcal/mol, β-sitosterol (BEN-010): $\Delta G = -23.94$ Kcal/mol) according to Affinity dG, Alpha HB and London dG scoring methods respectively suggest that molecules with lipophilic moieties at one end and hydrophobic moieties at another end is necessary for maximal interaction with the binding site residues of the UTF (PDB code = 2B56) (Fig. 9a). The compounds made major hydrogen bonding with Tyr274, Gly274, Asn277, Glu424 and Arg435 residues and significant hydrophobic interactions with the alkyl groups of Leu98, 285, Tyr89 and Val323.

It appeared that the structure of sitosterol (BEN-010) allowed it to better accommodate itself within the inner side of the hydrophobic pocket which could have accounted for its highest affinity for the UTF as scored by London dG. Rohituka-7(SAF-278) ($\Delta G = -18.34$ Kcal/mol), β-sitosterol (BEN-010)($\Delta G = -20.18$ Kcal/mol) and trichilia A (NIG-227) ($\Delta G = -21.02$ Kcal/mol) were ranked topmost by each of the respective scoring functions implemented in MOE Dock Tool. The poses of these compounds (Fig. 9b) showed that all the three communicated with FDS by making three hydrogen bonds with the NH of Lys47, 367 and Arg365. This binding mode indicates that interactions with these three residues are important for the inhibition of this enzyme activity.

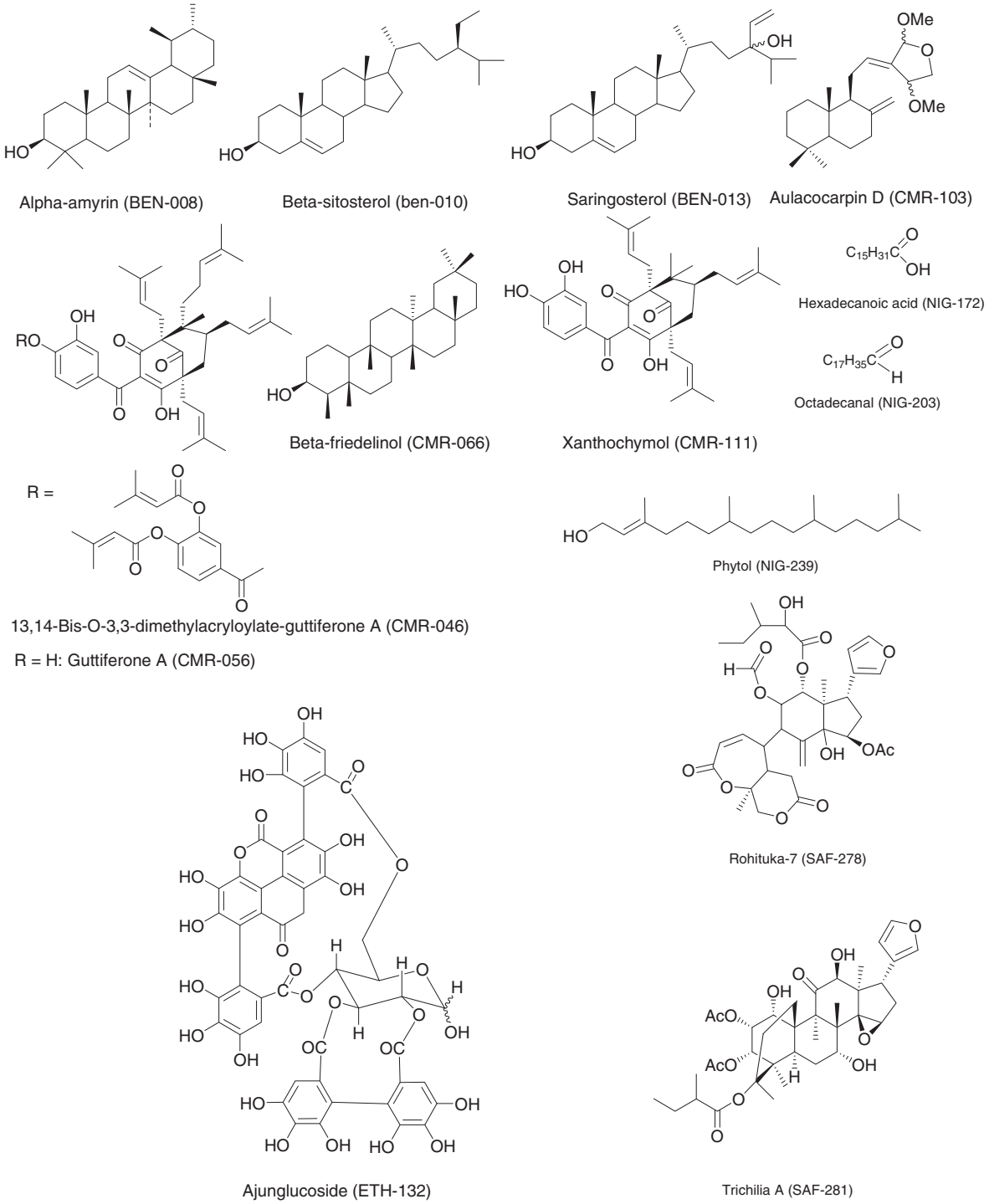

Fig. 10 2D chemical structures of the top scoring compounds from Afrotryp dataset

Furthermore, while the lipophilic moieties of the three compounds made contact with the polar ends of Leu100, Lys212 and 278, the hydrophobic part of the compounds especially took advantage of the non-lipophilic portions of the residues to establish interaction with the protein. Each of the compounds with the lowest binding energy as scored by the respective docking methods in MOE (NIG-239: $\Delta G = -22.92$ Kcal/mol, ETH-132: $\Delta G = -17.65$ Kcal/mol, BEN-008: $\Delta G = -14.34$ Kcal/mol), adopted unique poses within the binding site cavity of ODC (Fig. 9c). Nonetheless, hydrophobic contacts are found between all the three compounds and the aromatic rings of Tyr331, 389 and the

alkyl group of Val392. It appears that the nature of Alpha HB calculations (i.e. geometrically fitting the ligands into the binding site of protein target) favored the bulky glucoside moiety of arjunglucoside (ETH-132) which got accommodated in Ala69, Glu94 and Ser91 surrounded pocket. The activity of adenosine kinase was most inhibited by β-friedelinol (CMR-066) ($\Delta G = -15.08$ Kcal/mol), β-sitosterol (BEN-010) ($\Delta G = -14.89$ Kcal/mol) and methyl-hexadecanoate (NIG-227) ($\Delta G = -14.36$ Kcal/mol) as recorded by the respective scoring functions in MOE (Fig. 9d).

It appears hydrophobic interactions with Phe169, 200 aromatic rings and the alkyl side chain of Leu138 are pertinent for optimal interaction with the protein target because all the top scored compounds made contacts with them. The pose of *London dG* top scored ligand allowed its ester moiety to make hydrogen bonds with Ser64, Gly63, 65 and Cys12 amino acid. On the other hand, the Affinity dG and Alpha HB top scored compounds for the adenosine kinas protein used their OH groups to have a lipophilic intercourse with the polar ends of Ala37, Ile38 and Leu39 while still maintaining a hydrophobic relationship with the non-lipophilic parts of these residues. Affinity dG docking method identified guttiferone A(CMR-056) as the top scoring candidate while both Alpha HB and London dG ranked its derivative, 13,14-bis-O-3,3-dimethylacryloylate-guttiferone A (CMR-046), as the best docked compounds toward sterol 14α-demethylase (Fig. 9e). The observed π–π interactions between the aromatic rings of Phe110, Tyr103, 116 and the central benzene ring of the compounds suggest its importance for affinity towards the protein target. Except for the conspicuous hydrogen bonds with Tyr116 and Val114 residues, the best poses revealed that hydrophobic forces mainly account for the fitting of the compounds into the large non-lipophilic binding site groove of sterol 14α-demethylase.

All the top scoring compounds identified by each of the MOE Dock Tool, made hydrogen bonding with the Gly173 carboxylic group and Lys13 NHmoiety at the binding site of TIM (Fig. 9f). The best scored compound by Affinity dG docking scheme (BEN-013), had two significant lipophilic contacts with the NH groups of Glu97 and Gly235 and further proceed to have hydrophobic communication with alkyl side chain of Leu238 in order to strengthen its association with the protein. Alpha HB and London dG top ranking compounds (NIG-172 and CMR-103, respectively) had non-lipophilic contacts with the hydrophobic portions of Arg98, Tyr12 and Val169, Ile169, Glu97, respectively (Table 7, Fig. 10).

Physicochemical analysis of drug-like property, according to Lipinski's Ro5, of the individual top scoring compounds shows that all of them would be orally bioavailable in the biological system except the compounds CMR-046 and ETH-132 which violated more than two of the rules. In addition, nine of the fifteen compounds with top ranking scores can be developed into drugs for the treatment of the second stage of HAT. This is because of their low polar surface area, a crucial feature needed to permeate the BBB to treat second stage of the disease.

Conclusion

We have developed 3D virtual library of potential anti-trypanosomal compounds isolated from medicinal plants grown in Africa. The low energy .mol2 files for the compounds within the newly cureted dataset are available for virtual screening (as a supplementary file), along with additional tabular information describing the origins (plant sources) of the compounds, compound classes, full compound names, major authors and literature references. ADME/T profile of Afrotryp library was evaluated and its drug-likeness compared to that of DNP library. Appreciable percentage of the compounds in Afrotryp library possesses good ADME/T properties and even demonstrated an enhanced drug-likeness over the DNP datasets. This study, therefore, suggests that hits/leads discovered from Afrotryp library will likely have acceptable ADME/T profile. Docking calculations revealed 15 compounds with lowest theoretical binding energies and the analysis of their binding modes could lay the foundations for the rational development of novel trypanocidal drugs. We hope to check the stability of the complexes of the top scoring compounds through molecular dynamic simulation, before moving further to obtain/synthesize live compounds of successful candidates in order to carry out biological testings to confirm the in silico results. We expect that the aftermath of this work will be a novel anti-*Trypanosoma* molecule with improved activity that is ready to go through clinical trial.

Acknowledgements The authors thank Chemical and Bioactivity Information Centre, Cameroon for access to the natural products database and technical support. FNK currently holds a Georg Forster fellowship for postdoctoral researchers, obtained from the Alexander von Humboldt Foundation, Germany.

Compliance with Ethical Standards

Competing financial interests The authors declare no competing financial interests.

References

Amaechi N (2001) Toxicity of antiprotozoan drug diminazeneaceturate in rats. J Sustain Agric Environ 3:365–370

Andriani G, Amata E, Beatty J, Clements Z, Coffey BJ, Courtemanche G, Devine W, Erath J, Juda CE, Wawozak Z, Wood JT,

Lepesheva GI, Rodriguez A, Pollastri MP (2013) Antitrpanosomal lead discovery: identification of a ligand-efficient inhibitor of Trypanosomacruzi CYP51 and parasite growth. J Med Chem 56:2556–2567

Aronov AM (2005) Predictive in silico modeling for hERG channel blockers. Drug Discov Today 10:149–155

Artursson P (1990) Epithelial transport of drug in cell culture I: a model for studying the passive diffusion of drugs over intestinal absortive (Caco-2) cells. J Pharm Sci 79(6):476–482

Artursson P, Karlsson J (1991) Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco-2) cells. Biochem Biophys Rec Commun 175(3):880–885

Awulu EA, Oniye SJ, Adelanwa MA (2013) Phytochemical screening and *in vivo* antitrypanosomal activity of methanol extracts of *Peristrophe bicalyculata* in mice infected with *Trypanosoma evansi*. Int J Sci Technol 3:34–39

Bakker BM, Michelsi PAM, Opperdoesi FR, Hans V, Westerhoff HV (1999) Metabolic control analysis of glycolysis in trypanosomes as an approach to improve selectivity and effectiveness of drugs. J Biol Chem 274:14551–14559

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–342

Cavalli A, Poluzzi E, De Ponti F, Recanatini M (2002) Toward a pharmacophore for drugs inducing the long QT syndrome: insights from a CoMFA study of HERG K^+ channel blockers. J Med Chem 45:3844–3853

Chemical Computing Group (2010) Molecular Operating Envrinment (MOE) software

Colmenarejo G, Alvarez-Pedraglio A, Lavandera JL (2001) Cheminformatic models to predict binding affinities to human serum albumin. J Med Chem 44:4370–4378

D'Archivio S, Medina M, Cosson A, Chamond N, Rotureau B, Minoprio P (2011) Genetic engineering of *Trypanosoma* (*Dutonella*) *vivax*and*in vitro* differentiation under axenic conditions. PloS Negl Trop Dis 5:e1461

Demir O, Ladaied M, Merritt C, Stuart K, Amaro RE (2014) Computer-aided discovery of trypanosomabrucei RNA-editing terminal uridylyltransferase 2 inhibitors. Chem Biol Drug Des 84 (2):131–139

Deng J, Ernst NL, Turley S, Stuart KD, Hol WG (2005) Structural basis for UTP specificity of RNA editing TUTases from trypanosomabrucei. Embo J 24:4007–4017

Deng W, Jiang X, Mei Y, Sun J, Ma R, Liu X, Sun H, Tian H, Sun X (2008) Role of ornithine decarboxylase in breast cancer. Acta Biochem Biophys Sin 40:235–243

DesJarlais RL, Sheridan RP, Seibel GL, Dixon JS, Kuntz ID, Venkataraghavan R (1988) Using shape complementary as an initial screen in designing ligands for a receptor binding site of known 3D structure. J Med Chem 31:722–729

Dictionary of natural products on CD-rom (2005) Chapman and Hall/CRC Press, London

Duffy EM, Jorgensen WL (2000) Prediction of properties from simulations: free energies of solvation in hexadecane, octanol, andwater. J Am Chem Soc 122:2878–2888

Edelsbrunner H (2007) Weighted alpha shapes. Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL

Ertl P, Rohde B, Selzer P (2000) Fast calculation of molecular PSA as a sum of fragment based contributions and its application to the prediction of drug transport properties. J Med Chem 43:3714–3717

Fan Y, Shi LM, Kohn KW, Pommier Y, Weinstein JN (2001) Quantitative structure-antitumoractivity relationships of camptothecin analogues: cluster analysis and genetic algorithmbased-studies. J Med Chem 44:3254–3263

Feher M, Schmidt JM (2003) Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. J Chem Inf Comput Sci 43(1):218–227

Ghose AK, Herbertz T, Hudkins RL, Dorsey BD, Mallamo JP (2012) Knowledge-based central nervous system (CNS) lead selection and lead optimization for CNS drug discovery. Neuroscience 3:50–68

Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) A knowledge based approach in designing combinatorial or medicinal chemistry libraries for drug discovery 1. A qualitative and quantitative characterization of known drug databases. J Comb Chem 1 (1):55–68

Grishin NV, Osterman AL, Brooks HB, Phillips MA, Goldsmith EJ (1999) X-ray structure of ornithine decarboxylase from *Trypanosoma brucei*: the native structure and the structure in complex with alpha-difluoromethylornithine. Biochemistry 38: 15174–15184

Halgren TA (1996) Merck molecular forcefield. J Comput Chem 17:490–641

Hoet S, Opperdoes F, Brun R, Quetin-Leclercq J (2004) Natural products active against African trypanosomes: a step towards new drugs. Nat Prod Rep 21:353–364

Ibrahim MA, Mohammed A, Isah MB, Aliyu AB (2014) Antitrypanosomal activity of African medicinal plants: a review update. J Ethnopharmacol 154:26–54

Irvine JD, Takahashi L, Lockhart K, Cheong J, Tolan JW, Selick HE, Grove JR (1999) MDCK (Madin-Darby canine kidney) cells: a tool for membrane permeability screening. J Pharm Sci 88:28–33

Jorgensen WL, Duffy EM (2000) Prediction of drug solubility from Monte Carlo simulations. Bioorg Med Chem Lett 10:1155–1158

Jorgensen WL, Duffy EM (2002) Prediction of drug solubility fromstructure. Adv Drug Deliv Rev 54:355–366

Kaiser M, Bray MA, Cal M, Trunz BB, Torreele E, Brun R (2011) Antitrypanosomal activity of Fexinidazole, a new oral nitroimidazole drug candidate for treatment of sleeping sickness. Antimicrob Agents Chemother 55:5602–5608

Kirchhoff LV (2000) Agents of African Trypanosomiasis (Sleeping Sickness). In: Mandell GL, Benett JE and Dolin R (ed) Mandell, Douglas, and Benett's principles and practice of infectious diseases, Churchill Livingstone, Philadelphia, PA, pp 2853–2858

Koehn FE, Carter GT (2005) The evolving role of natural products in drug discovery. Nat Rev Drug Discov 4:206–220

Kuettel S, Greenwald J, Kostrewa D, Ahmed S, Scapozza L, Perozzo R (2011) Crystal structures of T.b. rhodesiense adenosine kinase complexed with inhibitor and activator: Implications for catalysis and hyperactivation. PloS Negl Trop Dis 5:E1164

Kuettel S, Mosimann M, Maser P, Kaiser M, Brun R, Scapozza L, Perozzo R (2009) Adenosine kinase of T. b. Rhodesiense identified as the putative target of 4-[5-(4-phenoxyphenyl)-2H-pyrazol-3-yl]morpholine using chemical proteomics. PloS Negl Trop Dis 3(8):e506

Lepesheva GI, Zaitseva NG, Nes WD, Zhou W, Arase M, Liu J, Hill GC, Waterman MR (2006) CYP51 from Trypanosomacruzi: a phyla-specific residue in the β' helix defines substrate s preferences of Sterol 14α-demethylase. J Biol Chem 281:3577–3585

Linares GEG, Ravaschino EL, Rodriguez JB (2006) Progresses in the field of drug design to combat tropical protozoan parasitic diseases. Curr Med Chem 13:335–360

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 23:3–25

Luco JM (1999) Prediction of brain–blood distribution of a large set of drugs from structurally derived descriptors using partial least squares (PLS) modelling. J Chem Inf Comput Sci 39:396–404

Mangal M, Sagar P, Singh H, Raghava GPS, Agarwal SM (2013) NPACT: Naturally occurring plant-based anti-cancer compound activity-target database. Nucleic Acids Res 41:D1124–D1129

Mao J, Mukherjee S, Zhang Y, Cao R, Sanders JM, Song Y, Zhang Y, Meints GA, Gao YG, Mukkamala D, Hudock MP, Oldfield E (2006) Solid-state NMR, crystallographic, and computational investigation of bisphosphonates and farnesyldiphosphate synthase-bisphosphate complexes. J Am ChemSoc 128: 14485–14497

Maser P, Luscher A, Kaminsky R (2003) Drug transport and drug resistance in African trypanosomes. Drug Resist Updat 6: 281–290

Meyer AY (1986) Molecular mechanics and molecular shape. III Surface area and cross-sectional areas of organic molecules. J Comput Chem 7:144–152

Chemical Computing Group Inc. (2010) Molecular Operating Environment version 2010: Montreal, Canada

Montalvetti A, Fernandez A, Sanders JM, Ghosh S, Van Brussel E, Oldfield E, Docampo R (2003) Farnesyl pyrophosphate synthase is an essential enzyme in *Trypanosoma brucei*, in vitro RNA interference and *in vivo* inhibition studies. J Biol Chem 278:17075–17083

Myrdal PB, Yalkowsky SH (2007) Solubilization of drugs in aqueous media. In: Swarbrick J (Ed) Encyclopedia of pharmaceutical technology. Informa Health Care, New York, NY

Noble ME, Verlinde CL, Groendijk H, Kalk KH, Wierenga RK, Hol WG (1991) Crystallographic and molecular modeling studies on trypanosomal triosephosphateisomerase: Critical assessment of the predicted and observed structures of the complex with 2-phosphoglycerate. J Med Chem 34:2709–2718

Ntie-Kang F, Nwodo NJ, Ibezim A, Simoben CV, Karaman B, Ngwa FV, Sippl W, Adikwu MU, Mbaze LM (2014) Molecular modeling of potential anticancer agents from African medicinal plants. J Chem Inf Model 54:2433–2450

Ntie-Kang F, Lifongo LL, Judson NP, Sippl W, Efange SMN (2014) How "drug-like" are naturally occurring anti-cancer compounds? J Mol Model 20:1–13

Ntie-Kang F, Lifongo LL, Mbah JA, Owono LCO, Megnassan E, Mbaze LM, Judson PN, Sippl W, Efange SMN (2013a) *In silico* drug metabolism and pharmacokinetic profiles of natural products from medicinal plants in the Congo basin. In Silico Pharmacol 1:12

Ntie-Kang F, Mbah JA, Lifongo LL, OwonoOwono LC, Megnassan E, Mbaze LM, Judson PN, Sippl W, Efange SMN (2013c) Assessing the pharmacokinetic profile of the CamMedNP natural products database: an *in silico* approach. Org Med Chem Lett 3:10

Ntie-Kang F, Zofou D, Babiaka SB, Meudom R, Scharfe M, Lifongo LL, Mbah JA, Mbaze LM, Sippl W, Efange SMN (2013b) AfroDb: a select highly potent and diverse natural product library from African medicinal plants. PLoS One 8(10): e78085

Nwodo NJ, Ibezim A, Ntie-Kang F, Adikwu MU, Mbah CJ (2015) Anti-trypanosomal activity of Nigerian plants and their constituents. Molecules 20:7750–7771

Olivares-Illana V, Pérez-Montfort R, López-Calahorra F, Costas M, Rodríguez-Romero A, Tuena de Gómez-Puyou MT, Gómez Puyou A (2006) Structural differences in triosephosphateisomerase from different species and discovery of a multitrypanosomatid inhibitor. Biochemistry 45:2556–2560

Oprea TI (2002) Current trends in lead discovery: are we looking for the appropriate properties? J Comput Aided Mol Des 16:325–334

Palm K, Stenbery P, Luthman K, Artursson P (1997) Polar molecular surface properties predict the intestinal absorption of drugs in humans. Pharm Res 14:154–571

Paul J, Gnanam R, Jayadeepa RM, Arul L (2013) Anti-cancer activity on *Graviola*, an excitingmedicinal plant extract vs various cancer cell lines and a detailed computational study onits potent anticancerous leads. Curr Top Med Chem 13:1666–1673

Pegg AE (1988) Polyamine metabolism and its importance in neoplastic growth and a target for chemopreventive. Cancer Res 48:759–774

Potts RO, Guy RH (1995) A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bond activity. Pharm Res 12:1628–1633

Quinn RJ, Carroll AR, Pham NB, Baron P, Palframan ME, Suraweera L, Pierens GK, Muresan S (2008) Developing a drug-like natural product library. J Nat Prod 71:464–468

Schneider G (2002) Trends in virtual computational library design. Curr Med Chem 9:2095–2102

Schrödinger Press (2011) QikProp 3.4 user manual. LLC, New York, NY

Sterling T, Irwin JJ (2015) ZINC 15 – ligand discovery for everyone. J Chem Inf Model 55:2324–2337

Teague SJ, Davis AM, Leeson PD, Opea TI (1999) The design of leadlike combinatorial libraries. Angew Chem Int Ed 38:3743–3748

Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD (2002) Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 45:2615–2623

Vemula VR, Lagishetty V, Lingala S (2010) Solubility enhancement techniques. Intl J Pharm Sci Rev Res 5(1):41–51

Vyas VK, Ghate M, Goel A (2013) Pharmacophore modeling, virtual screening, docking and insilico ADMET analysis of protein Kinase B (PKB β) inhibitors. J Mol Graph Model 42:17–25

Walters PW, Stahl MT, Murck MA (1996) Virtual screening – an overview. Res Focus 3(4):160–178

WHO Media Centre (2016) Trypanosomiasis, human African (sleeping sickness). http://www.who.int/mediacentre/factsheets/fs259/en. Accessed 31 May 2016