

See discussions, stats, and author profiles for this publication at: <http://www.researchgate.net/publication/264383421>

Molecular Modeling of Potential Anticancer Agents from African Medicinal Plants

ARTICLE *in* JOURNAL OF CHEMICAL INFORMATION AND MODELING · AUGUST 2014

Impact Factor: 3.74 · DOI: 10.1021/ci5003697

CITATIONS

13

READS

279

9 AUTHORS, INCLUDING:



[Fidele Ntie-Kang](#)

Martin Luther University Halle-Wittenberg

80 PUBLICATIONS 367 CITATIONS

[SEE PROFILE](#)



[Conrad VERANSO Simoben](#)

University of Buea

13 PUBLICATIONS 49 CITATIONS

[SEE PROFILE](#)



[Wolfgang Sippel](#)

Martin Luther University Halle-Wittenberg

195 PUBLICATIONS 3,144 CITATIONS

[SEE PROFILE](#)



[Luc Mbaze Meva'a](#)

University of Douala

47 PUBLICATIONS 421 CITATIONS

[SEE PROFILE](#)

Molecular Modeling of Potential Anticancer Agents from African Medicinal Plants

Fidele Ntie-Kang,^{*,†,‡,§} Justina Ngozi Nwodo,^{§,‡} Akachukwu Ibezim,[§] Conrad Veranso Simoben,[†] Berin Karaman,[‡] Valery Fuh Ngwa,^{||} Wolfgang Sippl,[‡] Michael Umale Adikwu,[‡] and Luc Meva'a Mbaze^{*,⊗}

[†]Department of Chemistry, Chemical and Bioactivity Information Centre, Faculty of Science, University of Buea, P.O. Box 63, Buea, Cameroon

[‡]Department of Pharmaceutical Chemistry, Martin-Luther University of Halle-Wittenberg, Wolfgang-Langenbeck Strasse 4, 06120, Halle Saale, Germany

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

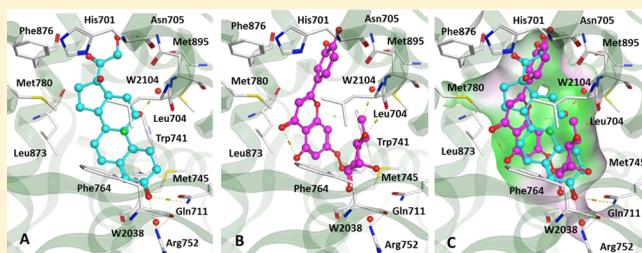
^{||}Department of Biochemistry, Faculty of Science, University of Buea, P.O. Box 63, Buea, Cameroon

[‡]Department of Pharmaceutics, University of Nigeria, Nsukka, Nigeria

[⊗]Department of Chemistry, Faculty of Science, University of Douala, P.O. Box 24157, Douala, Cameroon

S Supporting Information

ABSTRACT: Naturally occurring anticancer compounds represent about half of the chemotherapeutic drugs which have been put in the market against cancer until date. Computer-based or *in silico* virtual screening methods are often used in lead/hit discovery protocols. In this study, the “drug-likeness” of ~400 compounds from African medicinal plants that have shown *in vitro* and/or *in vivo* anticancer, cytotoxic, and antiproliferative activities has been explored. To verify potential binding to anticancer drug targets, the interactions between the compounds and 14 selected targets have been analyzed by *in silico* modeling. Docking and binding affinity calculations were carried out, in comparison with known anticancer agents comprising ~1 500 published naturally occurring plant-based compounds from around the world. The results reveal that African medicinal plants could represent a good starting point for the discovery of anticancer drugs. The small data set generated (named AfroCancer) has been made available for research groups working on virtual screening.



1. INTRODUCTION

The term cancer is often used to refer to all malignant tumors. A neoplasm or tumor is an abnormal mass of tissue, the growth of which exceeds and is uncoordinated with that of the normal tissue and continues in the same manner after cessation of the stimuli which had initiated it.¹ Cancer is the second most common disease-related cause of death within the human population.² It is known to be responsible for ~7.6 million deaths (~13% of all deaths) annually, according to the World Health Organization (WHO).³ In spite of the enormous efforts and progress in the field of cancer research, the WHO estimates that the threat of cancer disease will worsen.⁴ Radiation therapy and surgery as means of cancer treatment are only successful at the early localized stages of the disease. Chemotherapy, in contrast, is the mainstay in the treatment of malignancies because of its ability to cure widespread cancer. The endeavor of anticancer drug discovery from floral matter seems to be promising since it is believed that natural products (NPs) have been optimized during evolution. The above claim partly explains why plant-derived molecules have been highly valued by biomedical

researchers and pharmaceutical companies for the development of drugs against cancer.⁵ Plant materials (and their derived phytochemicals) have been extensively used to treat various forms of cancer and several reviews on medicinal plants used in the treatment of cancer and tumor have been published.^{6–10} According to a recent review on new chemical entities by Newman and Cragg, ~49% of anticancer drugs were either natural products or natural product-related synthetic compounds.^{11–14} These data are coherent with previous report by the same authors, which showed that about 47% of a total of 155 anticancer drugs approved up to 2006 were either NPs or directly derived from NP lead compounds by semisynthesis.¹⁵ The continent of Africa is very rich in floral biodiversity and its plant materials are endowed with NPs having intriguing chemical structures and promising biological activities.¹⁶ Therefore, the next generation of anticancer compounds or the scaffolds

Received: June 24, 2014

Published: August 13, 2014



necessary for their synthesis may be found in plants currently used in African Traditional Medicine (ATM).¹⁷

The African flora is known to contain promising anticancer agents,^{18–21} and the employment of computer-aided drug design (CADD) methods, which often incorporate virtual screening (VS) by docking of large compound databases against validated drug targets, followed by careful selection of virtual hit compounds to be screened by biological assays, has become a promising part of the drug discovery process nowadays.^{22–27} An example of a recent computational study of potential anticancer lead compounds has been carried out on isolates from the medicinal plant, *Graviola*.²⁶ This approach narrows down the number of compounds to be assayed and hence cuts down the costs of discovery of a drug.^{28–31} The effectiveness of such methods could be evaluated by the ability of a scoring method to clearly discriminate active compounds from inactive ones. In a previous study, the focus has been on evaluating how naturally occurring anticancer compounds are “drug-like”.³² Selected molecular descriptors related to drug absorption, distribution, metabolism, elimination, and toxicity (ADMET) were used to assess the pharmacokinetic profiles of about 1 500 NPs exhibiting *in vitro* and/or *in vivo* anticancer properties, which have been derived from the Naturally Occurring Plant-based Anticancer Compound-Activity-Target (NPACT) database.³³ It was observed that 42% of the compounds studied had physicochemical properties within the accepted range for 95% of known drug, while, respectively, 63% and 69% of the corresponding “drug-like” and “lead-like” subsets showed compliance.²⁰ The aim of the study was to bring out the need to return to natural sources in searching for anticancer lead compounds, as opposed to combinatorial synthesis. In the present study, the focus has been laid on forming a 3D structural library of potential anticancer compounds from the African flora (currently containing ~400 structures), which can be easily accessible to medicinal chemists and pharmaceutical researchers. The “drug-likeness” of this small compound library has been evaluated by analyzing computed molecular descriptors commonly used to compare “drug-likeness” properties of virtual screening compound libraries.^{34,35} A comparison has been made with the NPACT³³ and Dictionary of Natural Products (DNP).³⁶ Lastly both the NPACT and the newly developed African natural product anticancer database were virtually screened by docking both libraries toward the binding sites of a set of 14 selected validated anticancer drug targets, which have been cocrystallized with known inhibitors. The binding affinities of the identified compounds from the African flora and those of the NPACT library have been compared with those of the native (cocrystallized) inhibitors for each of the studied drug targets.

2. RESULTS AND DISCUSSION

2.1. Brief Geographical, Chemical, and Phytochemical Classification of Promising Anticancer Agents from African Medicinal Plants. Data from the literature showed that majority of compounds with potential anticancer activities have been isolated from the forests of the Central African countries, which converge toward the Congo Basin (Figure 1). Analysis of collected data also showed that the majority of the 390 identified compounds belongs to the terpenoid class (31%, Figure 2A), while the plant family with a majority of cytotoxic, antiproliferative and anticancer agents is the Rutaceae family (14.1%), followed by the Moraceae, Leguminosae, and Rubiaceae families, respectively, representing 8.8%, 8.7%, and 4.5% (Figure 2B). Other significantly populated classes included

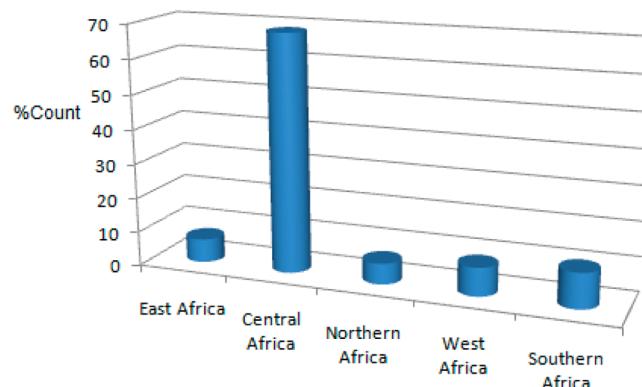


Figure 1. Bar chart showing the distribution of potential anticancer agents from Africa, according to the region of isolation.

the flavonoids, alkaloids, and steroids. These results are comparable with a similar study recently carried out for compounds possessing significant antiproliferative or pro-apoptotic effects, isolated from Cameroonian medicinal plants.³⁷ Moreover, previous studies on plant classifications by compound types in Africa showed that the Rutaceae and Rubiaceae families are noted to contain abundant quantities of alkaloids,^{38–40} some of which have shown remarkable cytotoxicities, anticancer, and antiproliferative properties.⁴¹ Vinca alkaloids are known to exhibit potent anticancer properties and some have been derived from the African flora. Vinca alkaloids are a remarkable class of antimicrotubule agents, discovered by chance in a species of the herbaceous perennial plant *Catharanthus* sp. (Madagascar Periwinkle). The sap from the species *C. roseus* has been used historically in folk medicine to treat various conditions, including diabetes and high blood pressure.⁴² More than 70 compounds from this subclass of alkaloids have been identified from the sap of this plant, two (vincristine and vinblastine) of which have shown remarkable antitumor properties.⁴³ On the contrary, plant-derived terpenoid ingredients are known to possess the ability to suppress nuclear factor-kappaB (NF-κB) signaling, the major regulator in the pathogenesis of inflammatory diseases and cancer.⁴⁴

2.2. “Drug-Likeness” Assessment by Lipinski’s Criteria.

The African flora, as a source of potential anticancer drugs, could be evaluated by analyzing data generated for a number of physicochemical parameters or molecular descriptors and comparing these with the range for known drugs. Lipinski’s “Rule of Five” (Ro5) is one of such criteria.³⁴ Lipinski et al. initially carried out experimental and computational approaches to estimate solubility and permeability of drugs in discovery and development settings. The study was based on a large computerized database of >50 000 drugs from the Derwent Co World Drug Index (WDI), with the view of selecting a subset of compounds from this database that are likely to have superior physicochemical properties suitable for oral bioavailability and hence “drug-likeness”. Initially, NPs were not included in this study, based on the fact that Lipinski et al. initially postulated that NPs do not respect the Ro5. The criteria [the Ro5 predicts that poor absorption or permeation is more likely when there are more than 5 hydrogen bond donors (HBD), 10 hydrogen bond acceptors (HBA), the molecular weight (MW) is greater than 500 Da and the calculated log P (Clog P) is greater than 5 or Moriguchi Log P (Mlog P > 4.15)] have often been used to access the “drug-likeness” of NP libraries to be employed in drug discovery programs.⁴⁵ An additional criterion, which is the

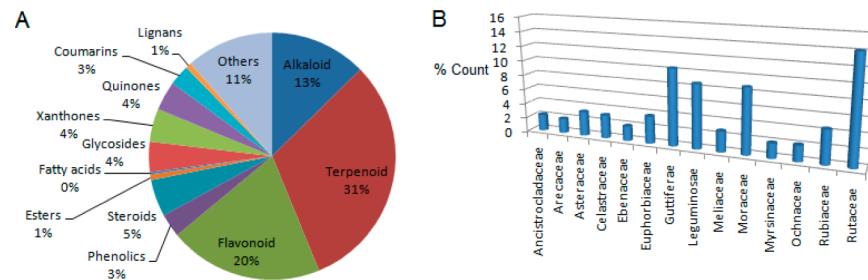


Figure 2. (A) Pie chart showing the distribution of potential anticancer drugs from the African flora by compound types. (B) Bar chart showing the distribution by plant families of origin.

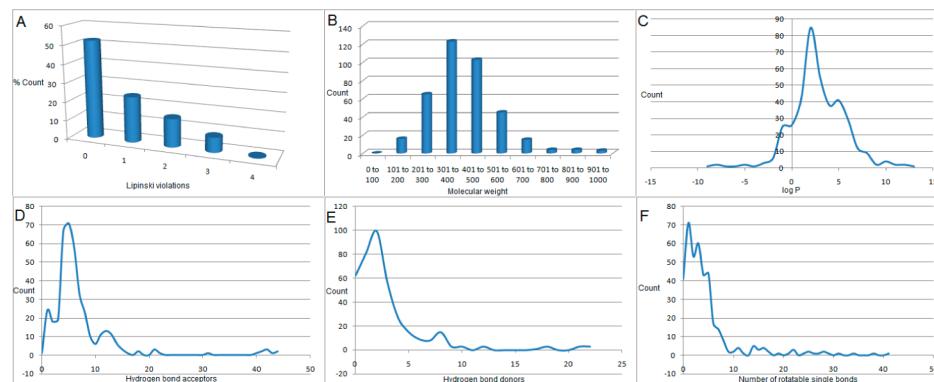


Figure 3. Distribution of Lipinski parameters used to access the “drug-likeness” of the potential anticancer compounds from African flora: (A) bar chart showing the percentage distribution of Lipinski violations, (B) bar chart showing the distributions of molecular weights, (C) Plot of $\log P$ distribution, (D) plot of number of hydrogen bond acceptors, (E) plot of hydrogen bond donors, and (F) plot of number of rotatable single bonds.

number of rotatable bonds (NRB) < 5 for drug-like compounds, is often included in the assessment. This is because the NRB property has been proven experimentally to influence bioavailability in rats.⁴⁶

In our present study, Lipinski parameters (MW, Clog P or $\log P$, HBA, and HBD) were computed for the ~400 potential anticancer, cytotoxic, and antiproliferative agents derived from the African flora. Plots of the distributions of the various parameters are shown in Figure 3. It was observed that about 50% of the compounds showed no Lipinski violations (Figure 3A), meanwhile about 75% showed <2 violations. The peak of the MW distribution was in the interval 301–400 (Figure 3B), corresponding to 31.5% of the compounds. Moreover, 307 out of the 390 compounds (i.e., 78.7%) respected the Lipinski criterion for the MW parameter for drug-like compounds ($MW < 500$) while 288 compounds (i.e., 73.8%) respected the $\log P$ criterion ($\log P < 5$). The distribution curve for the $\log P$ parameter gave a rugged Gaussian shape with a peak centered at $\log P = 1.5$ units (Figure 3B), a situation similar to previous studies for data sets from African medicinal plants,^{20,38,47,48} with only nine (9) compounds having $\log P > 10$ units. A total of 319 compounds (representing about 82%) had $HBA \leq 10$ and a corresponding 331 compounds (~85%) had $HBD \leq 5$. Interestingly, only one-fifth of the compounds showed $NRB > 5$. The distribution curves for the HBA and HBD parameters, respectively, rose to peak values of five acceptors and two donors, then quickly fell to give a total of 44 acceptors and 22 donors for the bulkiest NPs (Figure 3D,E). Similarly the distribution curve for the NRB rose to a peak value for 1 RB (corresponding to 18.2% of the compounds). The minima, maxima, and mean values of each of the considered physicochemical parameters have been shown in Table 1. From the data, it comes out that the mean molecular weight for

Table 1. Summary of Physicochemical Properties (Often Used to Predict “Drug-Likeness”) of the 390 Potential Anticancer Agents Derived from the African Flora

property	maximum ^f	minimum ^g	avg ^h
MW ^a (Da)	1999.16	120.17	437.87
$\log P^b$	13.60	-8.09	3.27
HBA ^c	44	0	7.05
HBD ^d	22	0	2.97
NRB ^e	41	0	4.57

^aMolecular weight. ^bLogarithm of the calculated *n*-octanol/water partition coefficient. ^cNumber of hydrogen bond acceptors. ^dNumber of hydrogen bond donors. ^eNumber of rotatable single bonds. ^fMaximum number. ^gMinimum number. ^hMean value.

potential anticancer agents from African flora is about 437 Da, which is much higher than that for the average drug (310 Da).⁴⁹ However, the mean values for the other parameters; $\log P$ (3.27), HBA (7.05), HBD (2.97), and NRB (4.57) all fall within the acceptable limits for “drug-like” compounds, according to the Ro5. In addition, the pairwise scatter plots (MW against $\log P$, HBA, HBD, and NRB), respectively, show that the highest density of points lie in the Lipinski compliance regions (Figure S1 in the Supporting Information).

The challenge for lead identification is a major obstacle to overcome in drug discovery. This is particularly the case in cancer chemotherapy. The identification of specific leads compounds that bind to only particular types of receptors is constantly needed.^{50,51} Following Oprea’s criteria for “lead-like” compounds,^{52–54} 165 out of the 390 compounds from African flora (corresponding to 42.3%) could be regarded as “lead-like”, having physicochemical parameters falling within the intervals ($150 \leq MW \leq 350$; $\log P \leq 4$; $HBD \leq 3$; $HBA \leq 6$), often

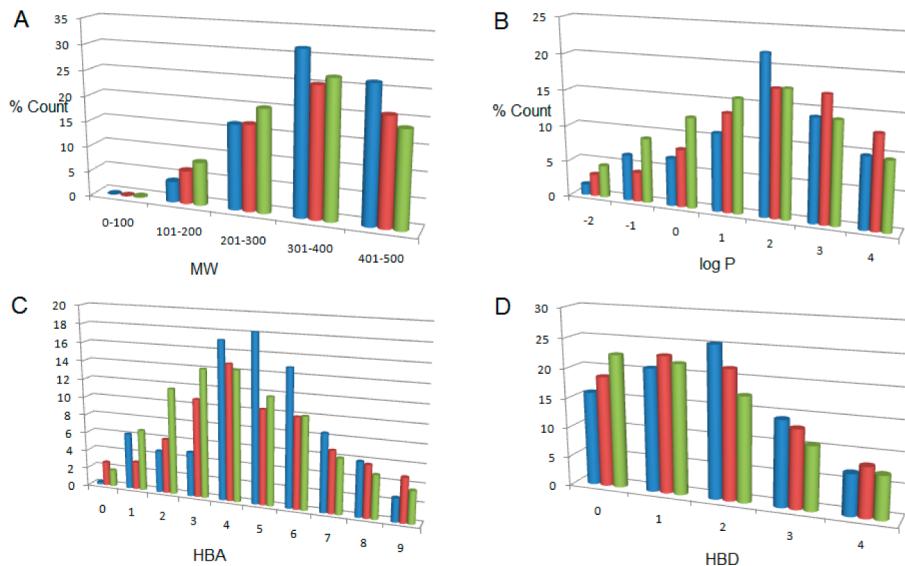


Figure 4. Bar charts showing a comparison of the percentage performances of the three data sets: AfroCancer, NPACT, and the DNP for the four Lipinski parameters used to access the “drug-likeness”: (A) molecular weights, (B) $\log P$, (C) hydrogen bond acceptors, and (D) hydrogen bond donors. In all cases, AfroCancer is shown in blue, NPACT in red, and the DNP in green.

Table 2. Selected Computed ADMET-Related Descriptors and Their Recommended Ranges for 95% of Known Drugs, Along with Percentage Compliances and Mean Values for the AfroCancer and NPACT Data Sets

descriptor	definition	required range	percentage compliances		mean values	
			AfroCancer	NPACT ^a	AfroCancer	NPACT ^a
S_{mol}	the total solvent-accessible molecular surface, in \AA^2 (probe radius 1.4 \AA)	300 to 1000 \AA^2	94.0	87.7	683.1	706.32
$S_{\text{mol,hfob}}$	the hydrophobic portion of the solvent-accessible molecular surface, in \AA^2 (probe radius 1.4 \AA)	0 to 750 \AA^2	95.0	89.7	402.3	417.44
V_{mol}	the total volume of molecule enclosed by solvent-accessible molecular surface, in \AA^3 (probe radius 1.4 \AA)	500 to 2000 \AA^3	95.0	86.5	1296.6	1346.7
$\log S_{\text{wat}}$	the logarithm of aqueous solubility ^{55,56}	-6.0 to 0.5	80.3	80.6	-4.95	-4.62
$\log K_{\text{HSA}}$	the logarithm of predicted binding constant to human serum albumin ⁵⁷	-1.5 to 1.2	84.2	90.7	0.46	0.25
$\log B/B$	the logarithm of predicted blood/brain barrier partition coefficient ⁵⁸⁻⁶⁰	-3.0 to 1.0	95.1	83.6	-1.02	-1.50
BIP _{caco-2}	the predicted apparent Caco-2 cell membrane permeability, in nm s^{-1} (in Boehringer–Ingelheim scale) ⁶¹⁻⁶³	<5 low, >100 high	81.4	41.0	1450.7	1168.9
MDCK	the predicted apparent Madin-Darby canine kidney cell permeability in nm s^{-1} ⁶²	<25 poor, >500 great	45.3	47.6	800.5	661.7
Ind _{coh}	the index of cohesion interaction in solids ⁶⁴	0.0 to 0.05	91.3	91.4	0.013	0.018
Glob	the globularity descriptor	0.75 to 0.95	94.0	87.5	0.85	0.84
QP _{polarz}	the predicted polarizability	13.0 to 70.0	96.2	92.2	43.09	43.60
$\log \text{HERG}$	the predicted IC ₅₀ value for blockage of HERG K ⁺ channels ^{65,66}	concern < -5	58.5	62.6	-4.61	-4.56
$\log K_p$	the predicted skin permeability ^{67,68}	-8.0 to -1.0	91.8	94.3	-2.78	-3.25
#metab	the number of likely metabolic reactions	1 to 8	83.6	76.8	5.60	6.29

^aAs reported in ref 32.

referred to as the “Rule of 3.5”. Thus, from this study, an analysis of the physicochemical parameters indicates that over two-fifths of the potential anticancer compounds isolated from the African flora are predicted to be potential “lead-like” compounds.

2.3. Comparison with the NPACT Library and the Dictionary of Natural Products. In a previous study,³² an attempt was made to compare the “drug-likeness” features of naturally occurring, plant-derived anticancer agents from the NPACT database, containing 1 574 compounds³³ with a relatively large data set of NPs, which is the Dictionary of Natural Products (DNP), containing 126 140 compounds.³⁶ In this study, we further carried out a comparison of the three data sets, this time including the data set from the African flora, henceforth referred to as AfroCancer. Histograms showing the four Lipinski physicochemical parameters are represented in

Figure 4. For each subfigure, only the Lipinski regions of interest are shown ($\text{MW} < 500 \text{ Da}$, $-2 \leq \log P < 5$, $\text{HBA} < 10$, and $\text{HBD} < 5$) and percentage comparisons are carried out and shown in histograms (Figure 4). It was observed that, overall, these histograms show an enhancement of the distributions of the AfroCancer data set over the NPACT and the DNP for Lipinski properties. As an example, the MW distribution histograms (Figure 4A) show that the AfroCancer data set has the highest percentage abundance for the interval 301–500 Da, even though all three data sets peak in the interval 301–400 Da. It was also observed that the AfroCancer data set had an enhancement of 11.8 over the NPACT data set and a corresponding 12.1% over the DNP for the region $301 \text{ Da} < \text{MW} \leq 500 \text{ Da}$. Below this range, the percentages were reduced for the AfroCancer and NPACT data sets, when compared to the DNP, with the first two

data sets being equal for the region 201–300 Da. As per the $\log P$ parameter, the AfroCancer data set was conspicuously higher than the other two data sets for the interval from 2.0 to 2.9 $\log P$ units (Figure 4B), which was the peak of the curve for the AfroCancer library. On the contrary, this parameter was lowest for the other intervals, when compared with the NPACT and the DNP libraries. For the HBA and HBD parameters, these were respectively enhanced for the intervals $3 < \text{HBA} \leq 8$ and $1 < \text{HBD} \leq 3$ for the AfroCancer over the NPACT and the DNP (Figure 4C,D).

2.4. Comparison of Selected ADMET Predicted Properties for NPACT and AfroCancer. A number of computed parameters are related to the absorption, distribution, metabolism, elimination, and toxicity (ADMET) of drugs. In this study, the selected parameters are shown in Table 2, which is a comparison between the NPACT and AfroCancer data sets. Each parameter was computed and the percentage of compounds for each data set that falls within the recommended range for 95% of known drugs was recorded. Among the parameters described in Table 2, related to the prediction of drug metabolism and pharmacokinetics (DMPK)/ADMET properties, the aqueous solubility ($\log S_{\text{wat}}$), brain/blood partition coefficient ($\log B/B$), Caco-2 cell permeability ($\text{BIP}_{\text{caco-2}}$), binding affinity to human serum albumin ($\log K_{\text{HSA}}$), number of likely metabolic reactions (#metab), solvent accessible surface area (S_{mol}), blockage of human-ether-a-go-go potassium ion (HERG K^+) channels and the total volume of each molecule enclosed by the solvent accessible surface area are discussed. The bioavailability of a compound depends on the processes of absorption and liver first-pass metabolism.⁶⁹ Absorption in turn depends on the solubility and permeability of the compound as well as interactions with transporters and metabolizing enzymes in the gut wall. The computed parameters used to assess oral absorption are the predicted aqueous solubility ($\log S_{\text{wat}}$), the conformation-independent predicted aqueous solubility, the predicted qualitative human oral absorption, the predicted % human oral absorption and compliance to the RoS.³⁴ It should be noted that the size of a molecule as well as its capacity to make hydrogen bonds, its overall lipophilicity, and its shape and flexibility are important properties to consider when determining permeability. Molecular flexibility has been seen as a parameter which is dependent on the NRB parameter, which is known to influence bioavailability.⁴⁶ The blood/brain partition coefficients ($\log B/B$) were computed and used as a predictor for access to the central nervous system (CNS). Madin-Darby canine kidney (MDCK) monolayers are widely used to make oral absorption estimates, the reason being that these cells also express transporter proteins but only express very low levels of metabolizing enzymes.⁴⁶ They are also used as an additional criterion to predict blood/brain barrier (BBB) penetration. Thus, the calculated apparent MDCK cell permeability could be considered to be a good mimic for the BBB (for nonactive transport). The efficiency of a drug may be affected by the degree to which it binds to the proteins within blood plasma. Binding of drugs to plasma proteins (like human serum albumin, lipoprotein, glycoprotein, and α , β , and γ globulins) greatly reduces the quantity of the drug in general blood circulation, and hence the less bound a drug is, the more efficiently it can traverse or diffuse cell membranes. The predicted plasma-protein binding has been estimated by the prediction of binding to human serum albumin; the $\log K_{\text{HSA}}$ parameter (the recommended range is -1.5 to 1.5 for 95% of known drugs). Human ether-a-go-go related gene (HERG) encodes a potassium ion (K^+) channel that

is implicated in the fatal arrhythmia known as torsade de pointes or the long QT syndrome.⁷⁰ The HERG K^+ channel, which is best known for its contribution to the electrical activity of the heart that coordinates the heart's beating, appears to be the molecular target responsible for the cardiac toxicity of a wide range of therapeutic drugs.^{71,72} HERG has also been associated with modulating the functions of some cells of the nervous system and with establishing and maintaining cancer-like features in leukemic cells.⁷² Thus, HERG K^+ channel blockers are potentially toxic and the predicted IC_{50} values often provide reasonable predictions for cardiac toxicity of drugs in the early stages of drug discovery.⁷³ Additionally, the index of cohesion interaction in solids was calculated from the HBA, HBD, and the surface area accessible to the solvent (S_{mol}) using the relation:

$$\text{Ind}_{\text{coh}} = \text{HBA} \times \sqrt{\text{HBD}} / S_{\text{mol}} \quad (1)$$

This is an indication of the sum of cohesive forces acting in a molecule in the solid state, hence the ability or inability to dissociate and enter into solutions (blood and plasma).

The globularity descriptor is also defined by

$$\text{Glob} = (4\pi r^2) / S_{\text{mol}} \quad (2)$$

where r is the radius of the sphere whose volume is equal to the molecular volume.

From the percentage compliances of the two data sets compared (Table 2), it was observed that eight (S_{mol} , $S_{\text{mol,hfob}}$, V_{mol} , $\log B/B$, $\text{BIP}_{\text{caco-2}}$, Glob , QP_{polz} , and #metab) out of the 14 selected parameters showed an enhancement of the AfroCancer data set over the NPACT data set. Two of the parameters ($\log S_{\text{wat}}$ and Ind_{coh}) showed almost equal compliances, while only four showed an enhancement of the NPACT over the AfroCancer data set. Additionally, in terms of the overall "drug-likeness" descriptor implemented in QikProp, the "#star" parameter,^{74,75} 57.9% of the AfroCancer compounds showed #star = 0, as opposed to only a corresponding 42.0% for the NPACT data set.³² It is noteworthy that the #star parameter is often employed as an assessment criterion by using a set of 24 essential property descriptors computed by QikProp.^{74,75} For each of the 24 essential descriptors, if a descriptor computed for an input molecule falls outside the recommended range for 95% of known drugs (Table 2), a #star of 1 is given. Hence, the most promising "drug-like" molecules show a compliance score (#star = 0). The human oral absorption and percentage human oral absorption parameters (not shown in Table 2) were also accessed. The predicted human oral absorption showed that 62.8% of the compounds within the AfroCancer library were likely to have high oral absorption in humans (scale 3), as opposed to only 53.4% for the NPACT library.³² Furthermore, the percent human oral absorption showed that 52.5% of the AfroCancer compounds (as opposed to only 33.9% for the NPACT) are likely to be absorbed orally at 100%, while a corresponding 63.9% (as opposed to only 47.1% for the NPACT) were likely to be absorbed at >90%. An analysis of the predicted central nervous system (CNS) activity (in the scale of $-2 = \text{inactive}$ to $+2 = \text{active}$), however, showed that only about 19% of the AfroCancer compounds could likely show from a fairly high (scale = 1) to high (scale = 2) activity in the CNS, i.e., have a propensity for brain tumors. The higher percentage of compounds within the AfroCancer data set complying with the recommended range of the BBB parameter for 95% of known drugs rather indicates that a greater proportion of the compounds from African flora are likely to penetrate the BBB

Table 3. Description of Selected Human Anticancer Drug Targets in This Study

PDB code	drug target class or role	co-crystallized ligand	resolution of X-ray crystal structure (Å)	ref
1GS4	androgen receptor	9 α -fluorocortisol	1.95	76
2X9E	mitotic regulator for chromosomal alignment and segregation	NMS-P715	3.10	77
3PE2	human protein kinase	CX-5011	1.90	78
SP21	oncogene protein	phosphoaminophosphonic acid guanylate ester OR GPPNHP (GPPNHP)	1.35	79
4M8H	retinoid X nuclear receptor	(R)-4-methyl-9C4AB30	2.20	80
3PE1	human protein kinase	CX-4945	1.60	78
3E37	protein farnesyltransferase	tert-butyl 4-((2-(4-cyanophenyl)[(1-methyl-1 <i>H</i> -imidazol-5-yl)methyl]amino)ethyl)[(2-methylphenyl)sulfonyl]amino)methyl)piperidine-1-carboxylate	1.80	81
3PP1	RAS-RAF-mitogen-activated protein kinase/extracellular signal-regulated kinase	3-[(2 <i>R</i>)-2,3-dihydroxypropyl]-6-fluoro-5-[(2-fluoro-4-iodophenyl)amino]-8-methylpyrido[2,3- <i>d</i>]pyrimidine-4,7(3 <i>H</i> ,8 <i>H</i>)-dione	2.70	82
4BBG	cell cycle regulator, critical for the assembly of the mitotic spindle.	3-[(<i>R</i>)-azanylethylsulfanyl-(3-ethylphenyl)-phenyl-methyl]phenol	2.50	83
3KKP	signaling protein	GPPNHP (GPPNHP)	1.35	84
2XMY	cyclin-dependent kinase (CDK) responsible for regulating transcription	4-[4-(3,4-dimethyl-2-oxo-2,3-dihydro-thiazol-5-yl)-pyrimidin-2-ylamino]-N-(2-methoxy-ethyl)-benzenesulfonamide	1.90	85
4ACM	glycogen synthase kinase	3-amino-6-(4-[(2-(dimethylamino)ethyl)sulfamoyl]-phenyl)- <i>N</i> -pyridin-3-ylpyrazine-2-carboxamide	1.63	86
1IEP	tyrosine kinase	PD173955 and imatinib (STI-571)	2.10	87
4BKX	class I histone deacetylase	MTA1	3.00	88
4O2B	tubulin	BAL27862	2.30	89

and hence access the CNS. Whether they will be active in the CNS or not is another aspect to be examined, since only 1.6% of the AfroCancer were predicted to have high CNS activity.

2.5. Docking Studies on Selected Anticancer Targets.

2.5.1. Selected Targets for Docking Studies. This study also involved evaluating the drug–target interactions between the two data sets (AfroCancer and NPACT) and 14 selected validated human anticancer drug targets. A summary of the descriptions of the selected targets is shown in Table 3. These include the androgen receptor (AR), cocrystallized with the agonist 9 α -fluorocortisol in the ligand-binding domain of the protein, PDB code 1GS4.⁷⁶ This drug target is significant in that it has previously been reported as a high-affinity cortisol/cortisone responsive AR (AR^{CCR}) isolated from androgen-independent human prostate cancer cell lines.⁹⁰

The drug target described by Colombo et al.⁷⁷ is responsible for mitotic regulation for chromosomal alignment and segregation and has been cocrystallized with a selective and orally bioavailable small-molecule inhibitor (NMS-P715), which selectively reduces cancer cell proliferation, leaving normal cells almost unaffected. Battistutta et al.⁷⁸ also describe 5-(3-chlorophenylamino)benzo[*c*][2,6]naphthyridine-8-carboxylic acid (CX-4945), which is a protein kinase CK2 inhibitor at the low nanomolar range with interesting selectivity versus other kinases, cocrystallized within the binding pocket of the human protein kinase CK2. Two other CK2 or type I inhibitors have been cocrystallized with this protein, which mark the highest selectivity ever reported for CK2 inhibitors. Another drug target involved in this study is the H-ras oncogene protein p21, which has been cocrystallized with the slowly hydrolyzing GTP analogue GppNp,⁷⁹ as well as the retinoid X receptor (RXR), cocrystallized with the selective rexinoid agonists, which are 4-methyl analogues of (2*E*,4*E*,6*Z*,8*Z*)-8-(3',4'-dihydro-1'(2*H*)-naphthalen-1'-ylidene)-3,7-dimethyl-2,3,6-octatrienoic acid or 9cUAB30.⁸⁰

Another target employed in this study is the protein farnesyltransferase (FTase) of PDB code 3E37, which catalyzes an essential posttranslational lipid modification of more than 60

proteins involved in intracellular signal transduction networks.⁸¹ Recently, FTase inhibitors have emerged as significant targets for development of anticancer therapeutics, which has been cocrystallized with five inhibitors based on an ethylenediamine scaffold. Moreover, it has also been verified that in mixed lymphocyte cultures, using tumor cells as antigenic stimulators, addition of recombinant galectin-1 dose-dependently inhibited lymphocyte production. The mitotic kinesin spindle protein (Eg5) plays a critical role in the assembly of the mitotic spindle and is thus a promising chemotherapy target.⁹¹ In an attempt to identify optimized ligands for this protein target, Good et al.⁸³ have cocrystallized this protein with *S*-trityl-*L*-cysteine-based inhibitors exhibiting moderate *in vivo* antitumor activity in lung cancer xenograft models.

The GTP-bound Ras protein is another target in this study.⁸² Ras is the name given to a family of related proteins which is ubiquitously expressed in cells.⁹² All Ras protein family members belong to a class of protein called small GTPase and is made up of proteins involved in transmitting signals within cells (cellular signal transduction). The cyclin-dependent kinases (CDK7 and especially CDK9) are responsible for regulating transcription.⁹³ Transcriptional CDK target, cocrystallized with the 2-anilino-4-(thiazol-5-yl)pyrimidine inhibitors, have also been employed in this study.⁸⁵ The glycogen synthase kinase-3 β (GSK3 β), cocrystallized with its novel potent and highly selective inhibitor (PDB code 4ACM), was also employed in this study.⁸⁶ Also known as tau phosphorylating kinase, this protein is a proline-directed serine/threonine kinase which was originally identified due to its role in glycogen metabolism.⁹⁴ It is noteworthy that glycogen synthase kinase-3 (GSK-3) is a serine/threonine protein kinase activity, which is inhibited by a variety of extracellular stimuli including insulin, growth factors, cell specification factors, and cell adhesion. Consequently, inhibition of GSK-3 activity has been proposed to play a role in the regulation of numerous signaling pathways that elicit pleiotropic cellular responses.⁹⁴

The tyrosine kinase (Bcr-Abl) that transforms cells and causes chronic myelogenous leukemia (PDB code 1IEP)⁸⁷ has also been

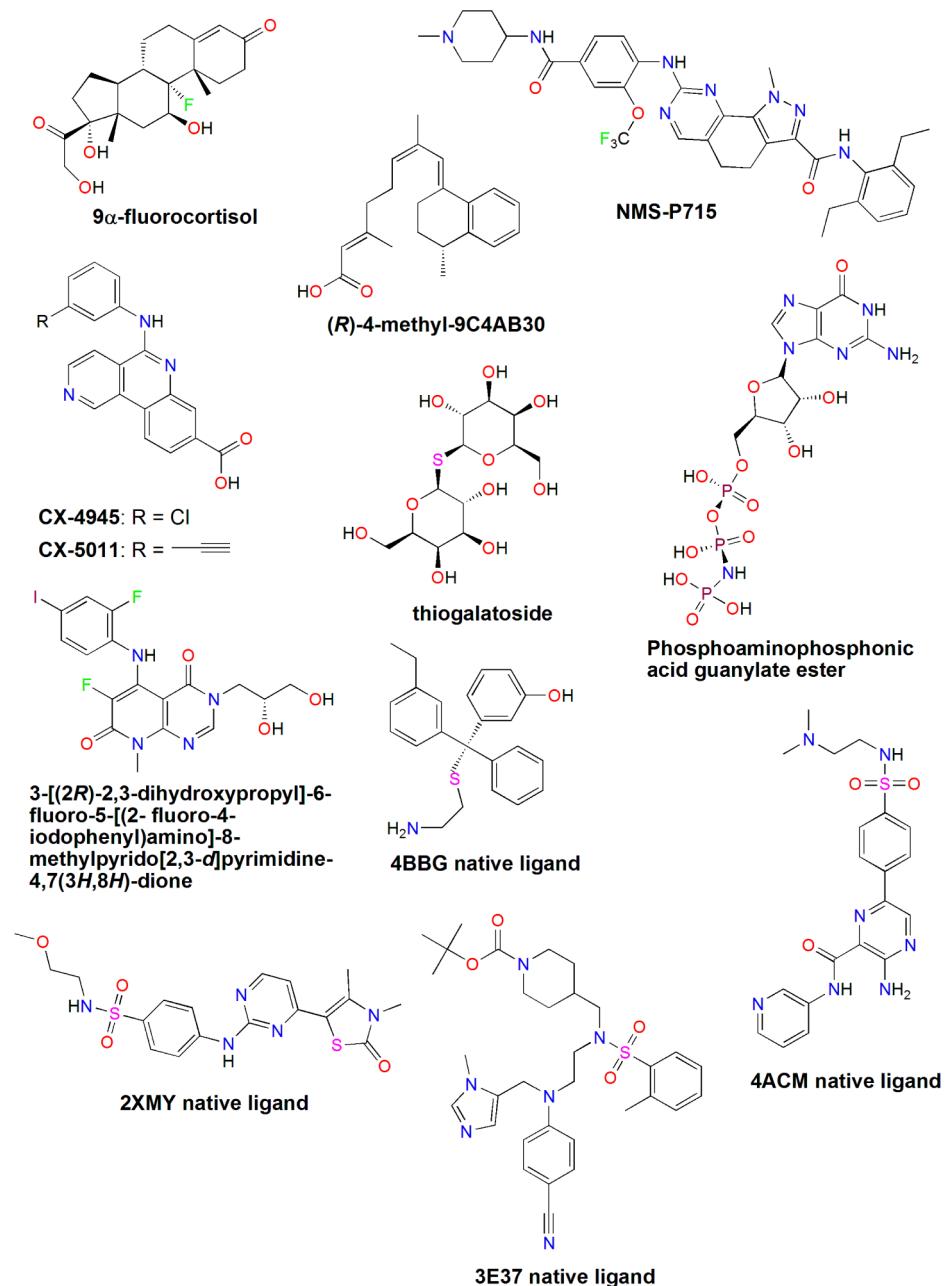


Figure 5. Chemical structures of some native ligands in the protein–ligand complexes studied.

included in this study. This is because small molecule inhibitors of Bcr-Abl that bind to the kinase domain can be used to treat chronic myelogenous leukemia. Class I histone deacetylases (HDAC1, HDAC2, and HDAC3) are recruited by cognate corepressor proteins into specific transcriptional repression complexes that target HDAC activity to chromatin, resulting in chromatin condensation and transcriptional silencing. This epigenetic target (PDB code 4BKX),⁸⁸ bound to the inhibitor MTA1, is one of the drug targets under study. The last of the studied targets is tubulin. This is because microtubule-targeting agents are widely used for the treatment of cancer and as tool compounds to study the microtubule cytoskeleton.⁸⁹

2.5.2. Docking Validation with MOE. For each of the selected protein–ligand complexes (see Figure 5, the native ligand was retrieved and redocked toward the binding site of each protein, following several predetermined docking and scoring methods

(as described in the Experimental Section). The various parameters for the centroids and dimensions of the retained grid boxes used for docking, after several docking trials, are shown in Table S1 in the Supporting Information. The results for docking validation using three methods for each complex are shown in Table 4. For all three docking/scoring methods (London dG, Affinity dG, and Alpha HB), the RMSD values were $<3.00 \text{ \AA}$, with only 6 of the 36 RMSD values exceeding 2.00 \AA . The most interesting docking protocols were observed for the native ligand of the 3PE1 complex (RMSD = 0.56 \AA) for the Affinity dG method and native ligand of the 3PE2 complex using the Alpha HB method (RMSD = 0.64 \AA). The mean values of RMSDs showed that the London dG method gave $\text{RMSD}^{\text{mean}} = 1.75 \text{ \AA}$ for all 12 complexes, while those of Affinity dG and Alpha HB both gave $\text{RMSD}^{\text{mean}} = 1.42 \text{ \AA}$. These docking/scoring methods could therefore be considered efficient, being capable of

Table 4. Lowest RMSD Values and Corresponding Docking Scores for Docking to the Co-Crystallized Ligands into the Binding Sites of the Selected Anticancer Drug Targets in This Study

protein PDB code	scoring methods					
	London dG		Affinity dG		Alpha HB	
	score	RMSD	score	RMSD	score	RMSD
1GS4	-26.92	1.07	-23.48	1.03	-26.81	0.83
2X9E	-12.78	2.57	-7.95	2.00	-13.10	2.80
3PE2	-7.25	2.83	-6.99	1.18	-7.29	0.64
SP21	-12.28	2.22	-12.41	1.69	-14.95	1.23
4M8H	-20.73	1.18	-19.93	1.31	-14.10	0.88
3E37	-4.58	1.98	-4.68	2.01	-4.79	1.68
3PE1	-15.26	1.52	-17.76	0.56	-12.81	1.12
3PP1	-15.95	1.47	-16.43	1.37	-16.09	1.51
4BBG	-13.39	1.69	-12.72	1.88	-13.83	2.77
3KKP	-5.68	1.67	-5.37	1.82	-6.74	1.50
2XMY	-14.10	1.52	-12.47	1.05	-11.57	1.05
4ACM	-15.68	1.25	-13.47	1.09	-12.77	1.00

reproducing the experimental binding orientations of the native ligands up to an average of <1.50 RMSD for the Affinity dG and Alpha HB scoring functions.

2.5.3. MOE Docking Results for the NPACT and AfroCancer Data Sets against the Selected Anticancer Targets. This involved 12 of the selected protein targets, the compounds

contained in the two data sets (AfroCancer and NPACT) were docked toward the binding site of the cocrystallized native ligand, using the methods, grid box centroids and dimensions obtained in the docking validation phase. A potential wall was set up as an additional criterion to prevent docking poses with the ligands outside the defined boundaries. For each compound, the five top scoring (lowest binding affinity) poses were selected and the distribution of docking scores are shown in Figure 6 (for the targets 1GS4, 2X9E, 2XMY, and 3KKP), while those of the other targets are given in the Supporting Information (Figures S1–S3). In each case, only percentage comparisons were carried out for the distributions of docking scores, comparing the two data sets within the same set of axes. It was observed that the docking scores for the both data sets fell within the same range (-40 to -4 kcal/mol) for the 1GS4 target, using the London dG scoring scheme. However, while the distribution of the London dG scores for the AfroCancer data set showed two peaks (at -25 kcal/mol and at -20 kcal/mol) that for the NPACT data set rather showed a rugged asymmetrical Gaussian distribution with a peak value at -23 kcal/mol. It was also noted that 11.56% of the compounds within the AfroCancer library had London dG scores (binding affinities) lower than that of the native ligand in the 1GS4 complex (which is -26.92 kcal/mol). The corresponding percentage for the NPACT library was slightly higher (13.20%). A similar analysis of the score distributions using the other scoring schemes for this target showed that up to 34.24% of the compounds within the AfroCancer library had lower Affinity dG scores than the bound 9 α -fluorocortisol (-23.48 kcal/mol), with

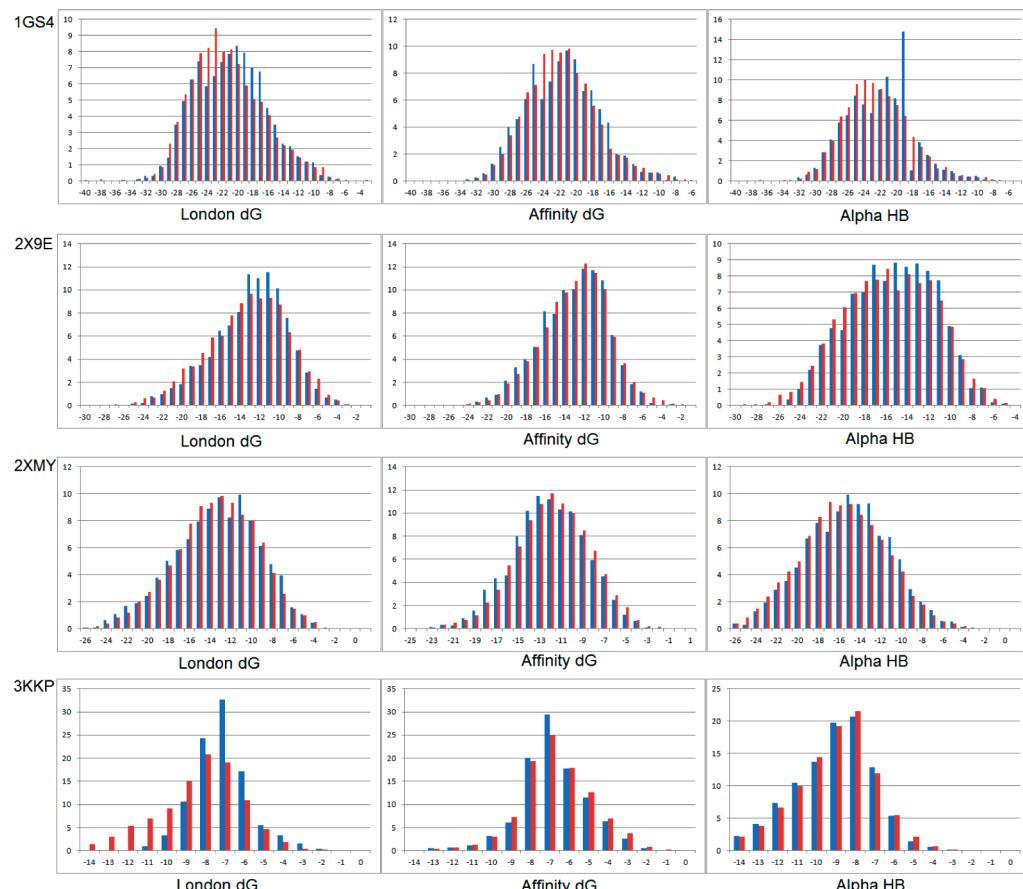


Figure 6. Distributions of docking scores for the two data sets (AfroCancer and NPACT) toward the binding sites of four of the studied targets (1GS4, 2X9E, 2XMY, and 3KKP) using the three scoring methods. In each case the blue = AfroCancer and red = NPACT.

Table 5. Percentage of Docking Scores Lower than the Reference Ligand for All 12 Targets under Study, Using the Three Scoring Schemes

protein PDB code	scoring method/data set					
	London dG		Affinity dG		Alpha HB	
	AfroCancer	NPACT	AfroCancer	NPACT	AfroCancer	NPACT
1GS4	11.56	13.20	34.24	35.91	15.21	16.35
2X9E	49.33	54.67	96.51	95.55	64.61	67.21
3EP2	98.67	97.08	95.69	95.50	99.43	98.64
SP21	78.21	78.48	54.36	53.71	66.36	33.16
4M8H	11.37	15.97	16.67	17.02	60.07	71.25
3PE1	47.43	49.10	33.42	32.11	73.59	86.71
3E37	1.92	4.47	4.98	3.51	5.56	9.33
3PP1	28.49	21.33	9.12	13.58	33.73	38.63
4BBG	34.56	31.92	32.82	28.48	56.38	60.67
3KKP	89.08	92.88	78.92	75.57	92.51	91.56
2XMY	37.13	38.79	45.33	41.63	80.46	84.02
4ACM	11.92	40.86	6.78	38.58	53.41	88.96

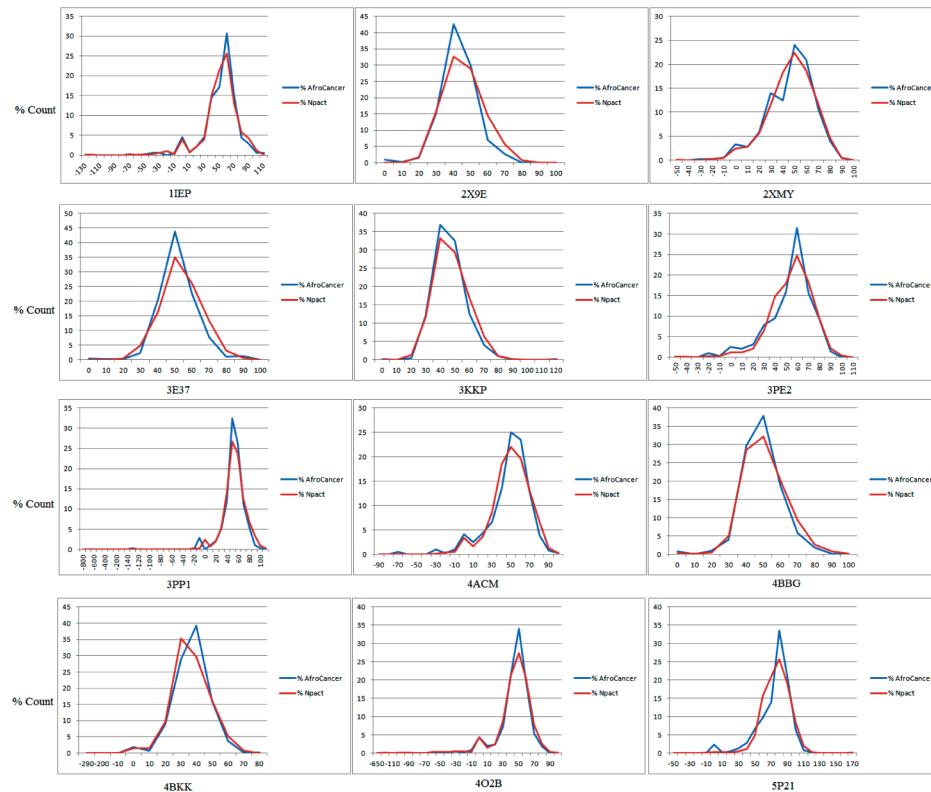


Figure 7. Distributions of docking scores for the two data sets (AfroCancer and NPACT) toward the binding sites of 12 of the studied targets (except 1GS4 and 4M8H) using the GoldScore fitness function. In each case the blue = AfroCancer and red = NPACT.

a corresponding 35.91% for the NPACT data set. The Alpha HB scoring scheme showed that 15.21% of the AfroCancer (and a corresponding 16.35% of the NPACT) had lower scores than the reference ligand (-26.81 kcal/mol).

A summary of the percentages of compounds having docking scores lower than the reference (bound) ligands within each of the 12 drug targets in this study are given in Table 5 for the two data sets. The results showed that the highest percentages of compounds having lower docking scores than the reference (bound) ligands for all scoring schemes were experienced for the 3EP2 and 3KKP targets. This could be explained by the fact that the bound ligands (CX-5011 and GPNP, respectively) are both highly functionalized, with several polar groups. When compared

with the naturally occurring anticancer data sets in our study (a majority made of less functionalized terpenoids), the nature of binding interactions could be quite different, when compared with the native ligands within these target proteins. The same explanation could be offered for the 3PE1 target, even though this is most pronounced for the Alpha HB scoring method, the others being, respectively, about half and one-third of the both data sets for the London dG and Affinity dG methods. It was generally observed that the proportions of compounds within the two data sets with docking scores lower than those of the reference ligands were within the same range. Exceptions to this rule were observed for the SP21 target, for which the AfroCancer data set was 2-fold, when compared with the NPACT for the

Table 6. Summary of the Percentages of Docking Poses in the Two Data Sets with Lower Scores than the Native Ligands for Gold and Glide Docking, Along with the Docking Scores of the Native Ligands

target PDB code	GoldScore			GlideScore (XP)		
	G_s^{GS} (ref ligand, kcal/mol)	AfroCancer	NPACT	G_s^{XP} (ref ligand, kcal/mol)	AfroCancer	NPACT
1GS4	79.03	0.00	0.00	-8.23	2.76	1.42
1IEP	117.40	0.26	0.00	-11.80	0.28	0.07
2X9E	69.10	3.07	7.12 ^a	-8.87	7.36	5.54
2XMY	81.79	3.58	3.81 ^a	-12.30	0.00	0.54 ^a
3E37	65.20	19.95	27.38 ^a	-7.24	29.62	39.65 ^a
3KKP	127.76	0.00	0.00	-9.82	1.10	2.82 ^a
3PE2	95.21	0.26	1.21 ^a	-11.50	0.83	0.57
3PP1	76.68	10.23	14.87 ^a	-8.77	10.58	12.52 ^a
4ACM	88.50	1.79	1.91 ^a	-8.28	10.20	8.08
4BBG	90.33	0.00	1.01 ^a	-10.14	0.58	0.55
4BKK	60.47	3.70	5.83 ^a	-6.53	8.12	10.63 ^a
4M8H	87.56	0.00	0.00	-13.41	1.16	0.00
4O2B	69.00	8.95	11.75 ^a	-7.66	26.67	22.08
SP21	173.33	0.00	0.00	-10.18	3.03	3.77 ^a

^aNPACT dominates over AfroCancer in terms of percentages.

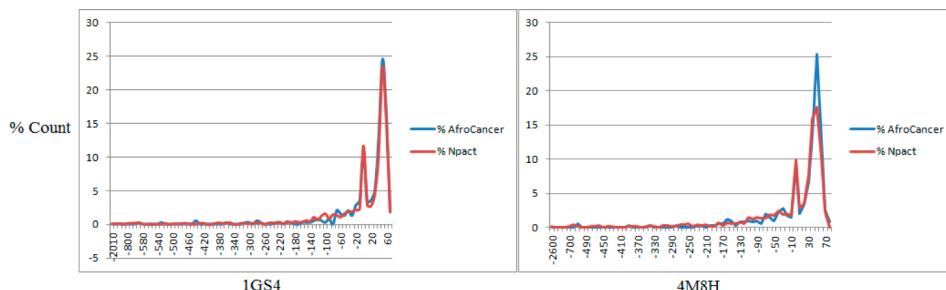


Figure 8. Distributions of docking scores for the two data sets (AfroCancer and NPACT) toward the binding sites of two of the studied targets (1GS4 and 4M8H) using the GoldScore fitness function. In each case the blue = AfroCancer and red = NPACT.

Alpha HB scheme as well as for the 4ACM target with much discrepancy observed for all scoring methods, the NPACT library having significantly higher percentages (40.86% versus 11.92% for London dG, 38.58 versus 6.78% for Affinity dG, and 88.96% versus 53.41% for Alpha HB). For the SP21 target, this is reflected in the disparity of the intervals of the docking scores, those of the AfroCancer data set ranging from -28 kcal/mol to -8 kcal/mol, while those of the NPACT data set are distributed from -24 kcal/mol to -2 kcal/mol (Figure S2 in the Supporting Information).

For three targets (4M8H, 3EP1, and 2XMY), the Alpha HB scoring scheme gave significantly higher proportions of compounds with lower docking scores than the reference ligands, when compared with the other two scoring methods. It may seem as if “fitting” of the bulkier NPs within the two data sets was disfavored by steric clashes within the binding pockets of these target proteins, whose native ligands may be much smaller, thus suggesting “induced fit” docking instead of docking a “rigid” protein. The lowest proportions of compounds within the two data sets that showed docking scores lower than the reference ligand were seen in the 3E37 target. Additionally, the plot of the distributions show that the AfroCancer docking scores are enhanced over those of the NPACT for all scoring methods for the interval -5 kcal/mol to the origin of the coordinate axes, where both curves meet (Figure S2 in the Supporting Information). In the X-ray crystal structure, this target is bound to a ligand having several polar functional groups and aromatic centers. Hence, the binding of the native ligand to its

target site seems to be much specific and such could not be easily mimicked during docking simulation by a majority of the NPs within our data sets.

2.5.4. Gold Docking Results. For each of the 14 selected protein targets, the compounds are contained in the two data sets (AfroCancer and NPACT) have been docked towards the target binding sites as described in the experimental section. The distribution of the GoldScores (G_s^{GS}) of the docking poses have been compared by plotting the histograms for both data sets side by side (Figure 7), while the percentages of docking G_s^{GS} values lower than the reference (cocryallized) ligand have been shown in Table 6 in each case. From Table 6, it appears that a significant proportion of docking poses showed better docking scores than the native ligands for three targets: 3E37 (19.95% AfroCancer and 27.38% NPACT), for 3PP1 (10.23% AfroCancer and 14.87%), and for 4O2B (8.95% AfroCancer and 11.75% NPACT). Additionally, it was generally observed that the distribution of the docking scores showed Gaussian distributions (Figure 7), except for the case of the 1GS4 and 4M8H targets with unusual distributions (Figure 8), with the AfroCancer curve having a higher peak in both cases (particularly distinguished in the case of the 4M8H target). For the other targets, the distributions of the G_s^{GS} fitness function for the docked poses showed that the AfroCancer data set had a higher peak than the NPACT data set in all cases. This dominance of the AfroCancer G_s^{GS} scores (Figure 7) may indicate that a search for potential anticancer drugs from African medicinal plants would be a good endeavor.

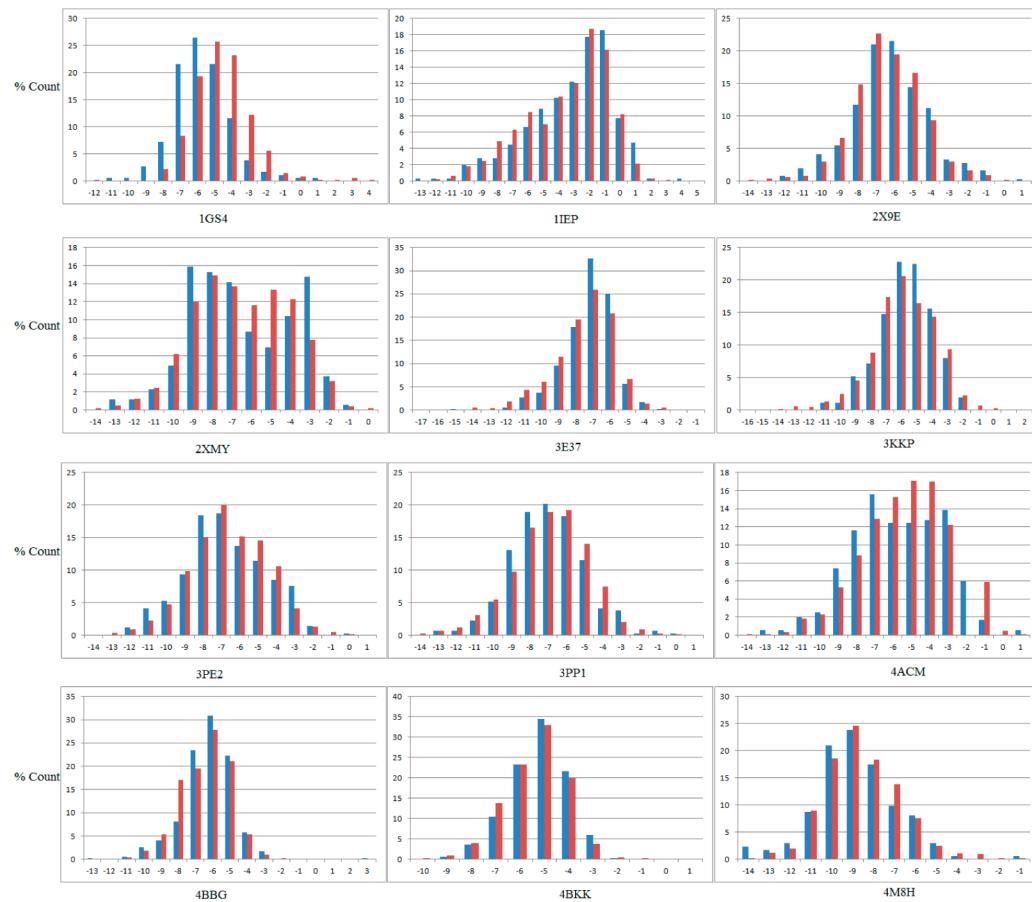


Figure 9. Distributions of docking scores for the two data sets (AfroCancer and NPACT) toward the binding sites of 12 of the studied targets (except 4O2B and 5P21) using the GlideScore XP function. In each case the blue = AfroCancer and red = NPACT.

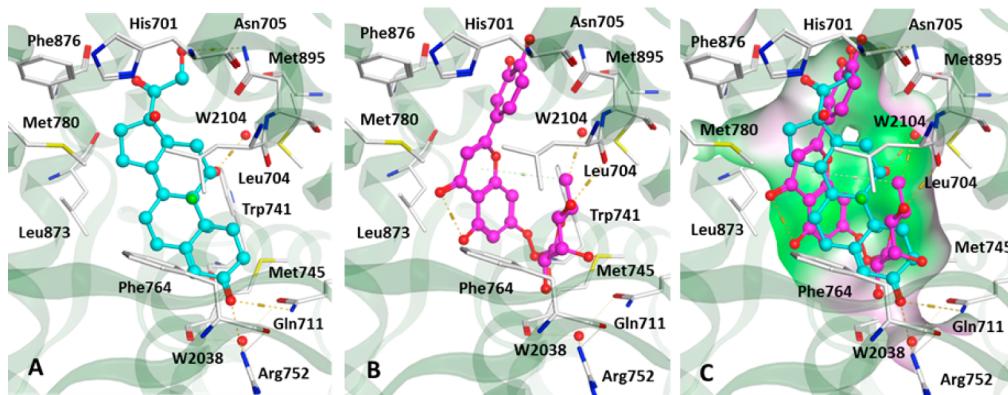


Figure 10. Top scoring pose for Glide docking of AfroCancer: (A) crystal structure of the androgen receptor (1GS4, dark green ribbon and amino acid residues, white sticks) in complex with cocrystallized 9 α -fluorocortisol (carbon atoms colored cyan). Water molecules are shown as red spheres. Only interacting amino acid residues are shown for clarity. Hydrogen bonds are shown as dashed lines. (B) Crystal structure of the androgen receptor (1GS4, dark green ribbon) in complex with docked luteolin-7-O- β -glucopyranoside (carbon atoms colored pink). Water molecules are shown as red. Only interacting amino acid residues are shown for clarity. Hydrogen bonds are shown as dashed lines. (C) Docking pose of the luteolin-7-O- β -glucopyranoside (carbon atoms colored pink) compared to the cocrystallized 9 α -fluorocortisol (carbon atoms colored cyan). The molecular surface of the binding pocket is displayed and colored according to the hydrophobicity. Polar regions are shown in magenta, hydrophobic regions in green.

2.5.4. Glide Docking Results. For each of the 14 selected protein targets, the compounds contained in the two data sets (AfroCancer and NPACT) have been docked towards the target binding sites as described in the experimental section. The GLIDE-XP docking protocol was able to predict the correct binding of all the resolved compounds ($\text{RMSD} < 2.5 \text{ \AA}$) to their corresponding targets except targets 3E37 and 4BKX. Docking

the 1GS4 receptor with both data sets gave Glide-XP scores (G_s^{XP}) with Gaussian-shaped distributions, the AfroCancer ligands clearly majoring over those of the NPACT for lower Glide-XP scores, i.e., in the interval $-12 \text{ kcal/mol} \leq G_s^{\text{XP}} \leq -6 \text{ kcal/mol}$ (Figure 9). This trend of the left handed dominance of the AfroCancer data set (toward lower G_s^{XP} scores) was also experienced in the targets 3PE2, 3PP1, and 4O2B. For the 2XMY

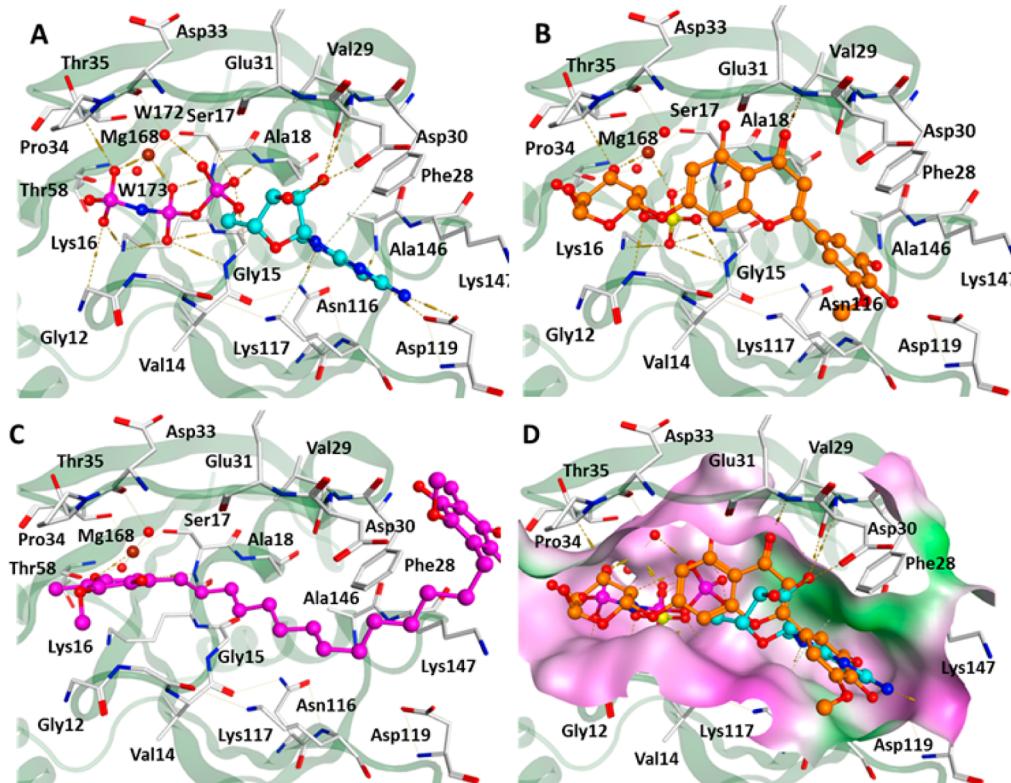


Figure 11. Top scoring pose for Gold docking of AfroCancer: (A) crystal structure of the Oncogene protein (SP21, dark green ribbon and amino acid residues, white sticks) in complex with cocrystallized GPNP (carbon atoms colored cyan). Hydrogen bonds are shown as dashed lines. Water molecules are shown as red spheres and magnesium ion is shown as a brown sphere. (B) Crystal structure of the oncogene protein (SP21, dark green ribbon) in complex with docked tricin-7-O- β -glucopyranoside-2"-sulfate (carbon atoms colored orange). Hydrogen bonds are shown as dashed lines. Water molecules are shown as red spheres (not labeled this time for clarity) and magnesium ion is shown as a brown sphere. (C) Docking result obtained for ardisiaquinone L (carbon atoms colored pink). (D) Docking pose of tricin-7-O- β -glucopyranoside-2"-sulfate (carbon atoms colored orange) compared to the cocrystallized GPNP (carbon atoms colored cyan). The molecular surface of the binding pocket is displayed and colored according to the hydrophobicity. Polar regions are shown in magenta, hydrophobic regions in green.

target, the distribution of G_s^{XP} values for the docking poses showed rugged curves for both data sets, the NPACT majoring within the interval from -7 kcal/mol to -4 kcal/mol, literally where the African data set showed a trough. On the contrary, there was a dominance of the African data set slightly above and below this range, while the curves were superposed at the extremes. A similar situation was seen in the 4ACM distribution. For the 3E37 distribution graphs, both data sets showed Gaussian distributions centered at -7.5 kcal/mol, but there was a clear dominance of the AfroCancer data set, with a peak >10 units higher than the NPACT counterpart. A similar situation is witnessed for the 3KKP and 4BBG targets. Meanwhile, a comparison of the G_s^{XP} distributions for six of the studied targets (1IEP, 2X9E, 4BKX, 4M8H, 5P21, and 5P21) did not show any clear demarcation between the two data sets. In terms of the percentage of compounds in the two data sets having lower docking scores than the native ligands (Table 6), the AfroCancer data set shows a higher compliance than the reference (or native) ligands, when compared with the NPACT data set for 8 out of the 14 targets.

2.5.5. Selected Top Scoring Ligands from the AfroCancer Data Set. To further highlight the potential of the AfroCancer data set, selected high scoring docking poses for two targets are shown in Figures 10 and 11, from Glide (1GS4 target) and Gold (5P21 target) docking, respectively. The comparative binding interactions of the native 9α -fluorocortisol are explored and compared with those of the docked pose of luteolin-7-O- β -

glucopyranoside, isolated from *Livistona australis* from Egypt.⁹⁵ This plant is used locally in treating various tumors, and the isolated metabolite showed *in vitro* cytotoxicities when assayed against liver carcinoma HEPG2, breast carcinoma MCF7 and colon carcinoma HCT116 cell lines. The American National Cancer Institute assigns a significant cytotoxic effect of extract for future bioguided studies if it exerts an IC_{50} value $\leq 30 \mu\text{g/mL}$.⁹⁶ Among the compounds isolated from this plant, luteolin-7-O- β -glucopyranoside revealed the highest antiproliferative activity with IC_{50} values of 13.5, 15.2, and 16.5 $\mu\text{g/mL}$ against HEPG2, MCF7, and HCT116, respectively, while the methanol extract exhibited less activity against HEPG2 and MCF7 cell lines with IC_{50} values of 21.9 and 22 $\mu\text{g/mL}$, respectively, and the weakest against colon carcinoma HCT116 cell line with IC_{50} values of 45.8 $\mu\text{g/mL}$. Furthermore, the target interactions formed by the top scoring docking pose of this ligand were similar to those of the native ligand for the 1GS4 target (Figure 10).

A similar examination of the best docking pose of the 5P21 target using the Gold docking tool revealed the flavone glycoside tricin-7-O- β -glucopyranoside-2"-sulfate and the anthraquinone ardisiaquinone L, respectively, isolated from *Livistona australis* (harvested in Egypt)⁹⁵ and from the leaves of *Ardisia kivuensis* (harvested from Cameroon).⁹⁷ In addition to its activities against a wide of microbial strains, this compound had shown a cytotoxicity against *Artemia salina* at 38%.⁹⁷ Similar studies have revealed that anthraquinones from this plant species show interesting anticancer properties.^{97,98} The top scoring binding

pose of ardisiaquinone L showed a similar binding mode with that of the native ligand in target 5P21 (Figure 11).

3. EXPERIMENTAL SECTION

3.1. Data Collection and Analysis. The plant sources, geographical collection sites, chemical structures of pure compounds, etc. were retrieved from literature sources comprising of data collected from Ph.D. theses and articles from peer reviewed journals, spanning the period 1971–2014. A full list of journals consulted is given in the Supporting Information (Table S2). The keyword queries in major natural product and medicinal chemistry journals were searched with keywords like cancer, anticancer, tumor, cytotoxic, antiproliferative, etc. Article hits were further hand-picked on the basis of the geographical region of collection of the plant material. Only articles corresponding to plant samples collected from the African region were finally selected. The data collected from the retained articles and available Ph.D. theses include plant sources, uses of plant material in traditional medicine, plant families, region of collection of plant material, isolated metabolites and type (e.g., flavonoid, terpenoid, etc.), and measured biological activities of isolated compounds. This data was compiled on an Excel sheet and analyzed. The range of description of potential anticancer activities of the identified compounds include, among others, induction of apoptosis and suppression of extracellular signal-regulated kinases/mitogen activated protein kinase NF- κ B pathways in lung tumor cells; cytotoxicity against sensitive and multidrug-resistant cancer cell lines; cytotoxic activity toward multifactorial drug resistant cancer cells; antiproliferative activity against BGC-823 and Hela cells; antiproliferative activity against the leukemia cell line TPH-1; *in vivo* antiprostate tumor activity against xenograft tumors; cytotoxic activity against a panel of breast cancer cell lines using the MTT assay; cytotoxicity against lung A549 adenocarcinoma, breast carcinoma MCF-7, prostate carcinoma PC-3, cervical carcinoma HeLa, and the acute monocytic leukemia cell line THP-1 and PF-382 and melanoma colo-38; cytotoxicity against A549 lung carcinoma cells; cytotoxic activity against the human Caucasian prostate adenocarcinoma cell PC-3 line, NCI-H226 (lung), CRL1579 (melanoma) *in vitro*; against the HT-29 and HCT 116 human colon cancer cell lines; against human promyelocytic leukemia (HL60) cell line by using the MTT method; antiproliferative and antioxidant properties on MB-MDA435 cell lines; enhancement of cAMP-regulated chloride conductance of cells expressing cystic fibrosis transmembrane conductance regulator (CFTR) gene CFTR Δ F508; antiproliferative activity against Drosophila S2 cells and the HCT116 cell line; antiproliferative activity against Ishikawa and A431 cell lines; induction of apoptosis in human tumor cells derived from B cell lymphoma and multiple myeloma; preferential necrotic cell death of PANC-1 and PSN-1 cells under nutrient-deprived and serum-sensitive conditions; apoptotic and antiproliferative activities against human leukemic B lymphocytes, such as the hairy cell leukemia-derived ESKOL cell line and cells from B-CLL (B-cell chronic lymphocytic leukemia) patients; against human DU-145 and hepatocarcinoma Hep G2 cells in the potato disk tumor induction and XTT assays. The cutoffs for biological activities were taken from “very active” to “weakly active”, as defined in the respective journal articles and theses. The anticancer properties of some of the compounds identified in the current study have been already described in some of our previous papers.^{20,39,47,99}

3.2. Preparation of 3D Models of Compounds and Calculation of Physicochemical Parameters. On the basis

of the known chemical structures of the NPs, all 3D molecular structures were generated using the graphical user interface (GUI) of the MOE software¹⁰⁰ running on a Linux workstation with a 3.5 GHz Intel Core2 Duo processor. The 3D structures were generated using the builder module of MOE and energy minimization was subsequently carried out using the MMFF94 force field,¹⁰¹ until a gradient of 0.01 kcal/mol was reached. The low energy 3D structures of the compounds were then saved as .mol2 files, subsequently included into a MOE database (.mdb) file and converted to other file formats (.sdf, .mol, .mol2, and .ldb), which are suitable for use in several virtual screening workflow protocols. The MW, NRB, log *P*, HBA, HBD, and number of Lipinski violations were calculated using the molecular descriptor calculator included in the QuSAR module of the MOE package.¹⁰⁰

The low energy 3D chemical structures in the AfroCancer library, which had been saved in .mol2 format were initially treated with LigPrep.¹⁰² This implementation was carried out with the graphical user interface (GUI) of the Maestro software package,¹⁰³ using the Optimized Potentials for Liquid Simulations (OPLS) force field.^{104–106} Protonation states at biologically relevant pH were correctly assigned (group I metals in simple salts were disconnected, strong acids were deprotonated, strong bases protonated, while topological duplicates and explicit hydrogens were added). A set of ADMET-related properties (a total of 46 molecular descriptors) were calculated by using the QikProp program^{74,75} running in normal mode. QikProp generates physically relevant descriptors and uses them to perform ADMET predictions. An overall ADME-compliance score—drug-likeness parameter (indicated by #stars)—was used to assess the pharmacokinetic profiles of the compounds within the AfroCancer library. The #stars parameter indicates the number of property descriptors computed by QikProp that fall outside the optimum range of values for 95% of known drugs. The methods implemented were developed by Jorgensen and Duffy.^{55,56,64} Some of the computed ADMET descriptors are shown in Table 2, along with their recommended ranges for 95% of known drugs. The NPACT database had already been prepared, following a similar procedure, as described in our previous study.³²

3.3. Protein Modeling and Docking Procedure with MOE. The X-ray structures of the protein–ligand complexes for the anticancer targets discussed in the Results and Discussion were retrieved from the protein databank.¹⁰⁷ Each protein–ligand complex was treated as follows; all the complexes were treated using MOE software.¹⁰⁰ For each complex, the cocrystallized water molecules and small molecules were deleted. The retained protein–ligand complexes were protonated using the protonate 3D procedure implemented in MOE.¹⁰⁰ The protonated complexes were then energy minimized in order to remove atomic clashes, using the Merck Molecular (MMFF94) Force field¹⁰¹ until a gradient of 0.001 kcal/mol was attained. For dimers, all the complexes were aligned and superposed using MOE software and the chains with missing amino acid residues in the middle of the chain were deleted, together with the bound ligands, while the remaining chain was treated as previously described.

The docking of all compounds toward the binding site of the 12 anticancer targets was carried out using the MOE Dock tool. Three main stages were involved in the docking process: Conformational Analysis of ligands, Placement, and Scoring. In the ligand Conformational Analysis stage, conformations from a single 3D conformation input ligand were generated by

conducting a systematic search methodology. All combinations of angles were created for each ligand. During the Placement stage, a collection of poses was generated from the pool of ligand conformations using the Triangle Matcher placement method. Several poses were generated by superposition of ligand atom triplets and triplet points in the receptor binding site. The receptor site points are composed of alpha sphere centers, representing locations of tight packing. At each iteration, a random conformation was selected, a random triplet of ligand atoms and a random triplet of alpha sphere centers were used to determine the pose. At the Scoring stage, the poses generated during the Placement stage were scored using three scoring methods: London dG, Affinity dG, and Alpha HB, all implemented in MOE. The London dG scoring function estimates the free energy of binding of the ligand from a given pose, as a combination of several terms, including the average rotational and translational entropy terms, energy lost as a result of the flexibility of the ligand, hydrogen bonding, metal contacts, and a desolvation term due to the volumes of the atoms of the protein and ligand in contact with the solvent. The Affinity dG scoring estimates the enthalpic contribution to the free energy of binding and takes into consideration the atomic contributions to the enthalpic term, including H-bond donor/acceptor pairs, Coulomb interactions between ions, metal ligation, hydrophobic contributions, etc. The Alpha HB scoring methods takes into consideration the geometric “fit” of the ligand within the binding site and H-bonding. The top 5 poses for each ligand were output in a MOE database for both AfroCancer and NPACT data sets. Docking validation was an attempt to identify the best docking parameters which reproduces the ligand conformation (docking poses) within the binding pocket, i.e., having the lowest RMSD values, with respect to the experimental binding mode (X-ray crystal structure). During the docking validation procedure, the native ligands present within the binding pocket of each of the protein complexes were docked toward their respective receptor sites using different grid parameters. The parameters retained with lowest RMSD values (Table S1 in the Supporting Information) were then used to carry out docking for the two data sets (AfroCancer and NPACT) toward the binding sites of the 12 selected anticancer drugs.

3.3. Protein Modeling and Docking Procedure with Gold. The X-ray structures of the protein–ligand complexes for the anticancer targets discussed in the Results and Discussion were prepared as previously described in the MOE docking procedure. Docking studies of the AfroCancer and the NPACT data sets were performed using the docking program GOLD version 5.2, distributed by the Cambridge Crystallographic Data Centre, Cambridge, U.K.^{108–114} The potential binding pocket was defined using the ligand in the corresponding crystal structure as a center and all the atoms within a radius of 10 Å to the ligand included in the binding site. Scoring was carried out using the GoldScore (GS) fitness function^{109,113} implemented within the GOLD program. Water molecules retained in the binding pocket were handled by turning on the toggling option in GOLD setup.

3.4. Protein Modeling and Docking Procedure with Glide. For Glide docking, protein preparation was carried out using the Protein Preparation Wizard workflow included in the Maestro suite.¹⁰³ The crystal structures were protonated, unwanted bound molecules were removed, and then minimization was carried out with the OPLS2005 force field,^{104–106} with the aim of relaxing the atomic coordinates until the geometric convergence (0.01 Å RMSD). The Glide program,^{115–118}

version 5.8, implemented in the Maestro Schrodinger suite was used in the current work. The receptor docking was generated using the receptor grid generator workflow in the region which was defined as 10 Å around the cocrystallized ligands. The position of the inhibitor in the corresponding crystal structure was used to define the binding site. Water molecules retained in the corresponding structures were treated as rigid molecules, and all other parameters for the grid generation were kept as default. The dockings were performed using the GLIDE extra precision (XP) mode, treating the ligand as flexible. In all docking studies, 10 docking poses were calculated for each ligand and all other options were left at their default values. The best-ranked pose from each docking run was included in the analysis. During the docking procedure using both GOLD and GLIDE, the bound ligand to each target was also included in the AfroCancer data set for the sake of comparison.

4. CONCLUSION

In this study, we have been able to identify ~400 compounds from the literature which have been isolated from medicinal plants growing on the African continent, having the potential for the development of anticancer drugs. The drug-likeness properties of these compounds have been evaluated by computed molecular descriptors, in comparison with two known natural product libraries, the NPACT and the DNP, and it was observed that, although the AfroCancer data set is much smaller than the other two data sets, its “drug-likeness” properties were enhanced over the latter, in terms of percentages. This library has been virtualized and made available for research groups involved in virtual screening of chemical libraries for the identification of potential hits, which could be further evaluated by biological assays. A successful docking attempt has been carried out on 14 selected known anticancer drug targets and our results show that, for five scoring methods, a significant number of the identified compounds within our data set have docking scores comparable with those of the bound inhibitors within the X-ray structures of the drug targets, implying the presence of potential binders within this data set. The virtual library derived from this work (containing low energy 3D atomic coordinates of the compounds and some relevant molecular descriptors) can be freely downloaded from the Supporting Information (data set S1), meanwhile enquiries about sample availability for bioassays can be addressed to the p-ANAPL consortium.⁴⁷ It is noteworthy that some compounds often exert interesting biological activities while also presenting limitations to drug discovery programs as a result of the presence of substructures which may be responsible for triggering toxicological profiles like carcinogenicity, mutagenicity, chromosome damage, genotoxicity, hepatotoxicity, irritation, ocular toxicity, and other toxicity end points. We intend to examine the AfroCancer data set for the presence of structural alerts or toxicophores common to known toxic agents to humans as well as generate pharmacophore models which may be useful in virtual screening programs for anticancer agents from plant sources.

■ ASSOCIATED CONTENT

§ Supporting Information

Pair-wise scatter plots for the distribution of Lipinski parameters among the 390 potential anticancer compounds from the African flora, distributions of docking scores for AfroCancer and NPACT towards the binding sites of several of the studied targets, distributions of Glide XP docking scores for AfroCancer and NPACT towards the binding sites of two of the studied targets,

Dataset S1 (separate .mol file), retained grid parameters for the docking sites of the selected targets, and list of journals consulted in constructing AfroCancer. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: lmbazze@yahoo.fr.

*E-mail: ntiekfidele@gmail.com.

Author Contributions

[#]F.N.-K. and J.N.N. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Computational resources were made available through the Molecular Simulations Laboratory, Department of Chemistry, University of Buea, Cameroon. Schrodinger Inc. is acknowledged for the academic license to use QikProp and LigPrep.

REFERENCES

- (1) Ashutosh, K. *Medicinal Chemistry*; New Age International Ltd.: New Delhi, India, 2007; p 794.
- (2) Gibbs, J. B. Mechanism-based target identification and drug discovery in cancer research. *Science* **2000**, *287*, 1969–1973.
- (3) WHO Media Centre. Fact Sheet No. 297, 2013, <http://www.who.int/mediacentre/factsheets/fs297/en/index.html> (accessed February 4, 2014).
- (4) Globocan Fact Stats, 2008, <http://globocan.iarc.fr/factsheets/populations/asp?uno=900> (accessed August 4, 2013).
- (5) Koehn, F. E.; Carter, G. T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discovery* **2005**, *4*, 206–220.
- (6) Hartwell, J. Plant used against cancer. A survey. *Lloydia* **1970**, *33*, 97–425.
- (7) Ashidi, J. S.; Houghton, P. J.; Hylands, P. J.; Efferth, T. Ethnobotanical survey and cytotoxicity testing of plants South-western Nigeria used to treat cancer, with isolation cytotoxic constituents from *Cajanus cajan* Millsp. leaves. *J. Ethnopharmacol.* **2010**, *128*, 501–512.
- (8) Graham, J. G.; Quinn, M. L.; Fabricant, D. S.; Farnsworth, N. R. Plants used against cancer- an extension of the work of Jonathan Hartwell. *J. Ethnopharmacol.* **2010**, *73*, 347–377.
- (9) Rahman, M. M.; Khan, M. A. Anti-cancer potential of South Asian plants. *Nat. Prod. Bioprospect.* **2013**, *3*, 74–88.
- (10) Cragg, G. M.; Newman, D. J. Plants as a source of anti-cancer and anti-HIV agents. *Ann. Appl. Biol.* **2003**, *143*, 127–133.
- (11) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **2010**, *75*, 331–335.
- (12) Li, J. W. H.; Vedera, J. C. Drug discovery and natural products: end to an era or an endless frontier? *Science* **2009**, *325*, 161–165.
- (13) Chin, Y. W.; Balunas, M. J.; Chai, H. B.; Renghorn, A. D. Drug discovery from natural sources. *APPS J.* **2006**, *8*, 239–253.
- (14) Potterat, O.; Hamburger, M. Progress in drug research: natural compounds as drug. In *Drug Discovery and Development with Plant-Derived Compounds*; Peterson, F., Amstutz, R., Eds.; Birkhäuser Verlag: Basel, Switzerland, 2008; pp 45–118.
- (15) Newman, D. J.; Cragg, G. M. Natural products as source of new drugs over the last 25 years. *J. Nat. Prod.* **2007**, *70*, 461–477.
- (16) Hostettmann, K.; Marston, A.; Ndjoko, K.; Wolfender, J. L. The potential of African plants as a source of drugs. *Curr. Org. Chem.* **2000**, *4*, 973–1010.
- (17) Efange, S. M. N. Natural products: a continuing source of inspiration for the medicinal chemist. In *Ethnomedicine and Drug Discovery*; Advances in Phytomedicine, Vol. 1; Iwu, M. M., Wootton, J. C., Eds.; Elsevier Science: Amsterdam, The Netherlands, 2002; pp 61–69.
- (18) Steenkamp, V. Phytomedicines for the prostate. *Fitoterapia* **2003**, *74*, 545–552.
- (19) Fadeyi, S. A.; Fadeyi, O. O.; Adejumo, A. A.; Okoro, C.; Myles, E. L. *In vitro* anticancer screening of 24 locally used Nigerian medicinal plants. *BMC Complement. Altern. Med.* **2013**, *13*, 79.
- (20) Ntie-Kang, F.; Zofou, D.; Babiaka, S. B.; Meudom, R.; Scharfe, M.; Lifongo, L. L.; Mbah, J. A.; Mbaze, L. M.; Sippl, W.; Efange, S. M. N. AfroDb: A select highly potent and diverse natural product library from African medicinal plants. *PLoS One* **2013**, *8*, e78085.
- (21) Sawadogo, W. R.; Schumacher, M.; Teiten, M. H.; Dicato, M.; Diederich, M. Traditional West African pharmacopeia, plants and derived compounds for cancer therapy. *Biochem. Pharmacol.* **2012**, *84*, 1225–1240.
- (22) Ferdous, S.; Mirza, M. U.; Saeed, U. Docking studies reveal phytochemicals as the long searched anticancer drugs for breast cancer. *Int. J. Comput. Appl.* **2013**, *67*, 1–5.
- (23) Meshram, R. J.; Bhagiade, N. H.; Gacche, R. N.; Jangle, S. N. Virtual screening and docking exploration on oestrogen receptor: an *in silico* approach to decipher novel anticancer agents. *Ind. J. Biotechnol.* **2012**, *11*, 389–395.
- (24) Fei, J.; Zhou, L.; Liu, Y.; Tang, X. Y. Pharmacophore modeling, virtual screening, and molecular docking studies for discovery of novel Akt2 inhibitors. *Int. J. Med. Sci.* **2013**, *10*, 265–275.
- (25) Chen, H.; Yao, K.; Nadas, J.; Bode, A. M.; Malakhova, M.; Oi, M.; Li, H.; Lubet, R. A.; Dong, Z. Prediction of molecular targets of cancer preventing flavonoid compounds using computational methods. *PLoS One* **2012**, *7*, e38261.
- (26) Paul, J.; Gnanam, R.; Jayadeepa, R. M.; Arul, L. Anti cancer activity on *Graviola*, an exciting medicinal plant extract *vs* various cancer cell lines and a detailed computational study on its potent anti-cancerous leads. *Curr. Top. Med. Chem.* **2013**, *13*, 1666–1673.
- (27) Vyas, V. K.; Ghate, M.; Goel, A. Pharmacophore modeling, virtual screening, docking and *in silico* ADMET analysis of protein kinase B (PKB β) inhibitors. *J. Mol. Graph. Model.* **2013**, *42*, 17–25.
- (28) Harvey, A. L. Natural products in drug discovery. *Drug Discovery Today* **2008**, *13*, 894–901.
- (29) Cragg, G. M.; Newman, D. J.; Snader, K. M. Natural products in drug discovery and development. *J. Nat. Prod.* **1997**, *60*, 52–60.
- (30) Klebe, G. Virtual ligand screening strategies; perspective and limitations. *Drug Discovery Today* **2006**, *11*, 580–594.
- (31) Kubinyi, H. Structure-based design of enzyme inhibitors and receptors ligands. *Curr. Opin. Drug Discovery Dev.* **1998**, *1*, 4–15.
- (32) Ntie-Kang, F.; Lifongo, L. L.; Judson, P. N.; Sippl, W.; Efange, S. M. N. How “drug-like” are naturally occurring anti-cancer compounds? *J. Mol. Model.* **2014**, *20*, 1–13.
- (33) Mangal, M.; Sagar, P.; Singh, H.; Raghava, G. P. S.; Agarwal, S. M. NPACT: naturally occurring plant-based anti-cancer compound-activity-target database. *Nucleic Acids Res.* **2013**, *41*, D1124–D1129.
- (34) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- (35) Good, A. C.; Cho, S. J.; Mason, J. S. Descriptors you can count on? Normalized and filtered pharmacophore descriptors for virtual screening. *J. Comput.-Aided Mol. Des.* **2004**, *18*, 523–527.
- (36) *Dictionary of Natural Products on CD-Rom*; Chapman and Hall/CRC Press: London, 2005.
- (37) Orang-Ojong, B. B.; Munyangaju, J. E.; Wei, M. S.; Lin, M.; Wei, F. G.; Fokounang, C.; Zhu, Y. Impact of natural resources and research on cancer treatment and prevention: A perspective from Cameroon (Review). *Mol. Clin. Oncol.* **2013**, *1*, 610–620.
- (38) Ntie-Kang, F.; Mbah, J. A.; Mbaze, L. M.; Lifongo, L. L.; Scharfe, M.; Ngo Hanna, J.; Cho-Ngwa, F.; Amoa Onguéné, P.; Owono Owono, L. C.; Megnassan, E.; Sippl, W.; Efange, S. M. N. CamMedNP: Building the Cameroonian 3D structural natural products database for virtual screening. *BMC Complement. Altern. Med.* **2013**, *13*, 88.
- (39) Ntie-Kang, F.; Lifongo, L. L.; Mbaze, L. M.; Ekwelle, N.; Owono Owono, L. C.; Megnassan, E.; Judson, P. N.; Sippl, W.; Efange, S. M. N. Cameroonian medicinal plants: a bioactivity versus ethnobotanical

- survey and chemotaxonomic classification. *BMC Complement. Altern. Med.* **2013**, *13*, 147.
- (40) Zofou, D.; Ntie-Kang, F.; Sippl, W.; Efange, S. M. N. Bioactive natural products derived from the Central African flora against neglected tropical diseases and HIV. *Nat. Prod. Rep.* **2013**, *30*, 1098–1120.
- (41) Noble, R. L.; Beer, C. T.; Cutts, J. H. Role of chance observation in chemotherapy: *Vinca rosea*. *Ann. N. Y. Acad. Sci.* **1958**, *76*, 882–894.
- (42) Noble, R. L. The discovery of the vinca alkaloids-chemotherapeutic agents against cancer. *Biochem. Cell Biol.* **1990**, *68*, 1344–1351.
- (43) Brown, I.; Sangrithi-Wallace, J. N.; Schofield, A. C. Antimicrotubule agents. In *Anticancer Therapeutics*; Missailidis, S., Ed.; Wiley-Blackwell: Chichester, U.K., 2008; pp 80–89.
- (44) Salminen, A.; Lehtonen, M.; Suuronen, T.; Kaarniranta, K.; Huuskonen, J. Terpenoids: natural inhibitors of NF- κ B signaling with anti-inflammatory and anticancer potential. *Cell. Mol. Life Sci.* **2008**, *65*, 2979–2999.
- (45) Quinn, R. J.; Carroll, A. R.; Pham, M. B.; Baron, P.; Palframan, M. E.; Suraweera, L.; Pierens, G. K.; Muresan, S. Developing a drug-like natural product library. *J. Nat. Prod.* **2008**, *71*, 464–468.
- (46) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623.
- (47) Ntie-Kang, F.; Amoa Onguéné, P.; Fotso, G. W.; Andrae-Marobela, K.; Bezabih, M.; Ndom, J. C.; Ngadjui, B. T.; Ogundaini, A. O.; Abegaz, B. M.; Mbaze, L. M. Virtualizing the p-ANAPL library: A step towards drug discovery from African medicinal plants. *PLoS One* **2014**, *9*, e90655.
- (48) Ntie-Kang, F.; Amoa Onguéné, P.; Scharfe, M.; Mbaze, L. M.; Owono Owono, L. C.; Megnassan, E.; Sippl, W.; Efange, S. M. N. ConMedNP:natural product library from medicinal plants in Central Africa. *RSC Adv.* **2014**, *4*, 409–419.
- (49) Khanna, V.; Ranganathan, S. Structural diversity of biologically interesting datasets: a scaffold analysis approach. *J. Cheminform.* **2011**, *3*, 30.
- (50) Mans, D. R. A.; Da Rocha, A. B.; Schwartsmann, G. Anti-cancer drug discovery and development in Brazil: Targeted plant collection as a rational strategy to acquire candidate anti-cancer compounds. *Oncologist* **2000**, *5*, 185–198.
- (51) Simmons, T. L.; Andrianasolo, E.; McPhail, K.; Flatt, P.; Gerwick, W. H. Marine natural products as anticancer drugs. *Mol. Cancer Ther.* **2005**, *4*, 333–342.
- (52) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. I. The design of leadlike combinatorial libraries. *Angew. Chem., Int. Ed.* **1999**, *38*, 3743–3748.
- (53) Hann, M. M.; Oprea, T. I. Pursuing the leadlikeness concept in pharmaceutical research. *Curr. Opin. Chem. Biol.* **2004**, *8*, 255–263.
- (54) Oprea, T. I. Current trends in lead discovery: are we looking for the appropriate properties. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 325–334.
- (55) Jorgensen, W. L.; Duffy, E. M. Prediction of drug solubility from structure. *Adv. Drug Delivery Rev.* **2002**, *54*, 355–366.
- (56) Duffy, E. M.; Jorgensen, W. L. Prediction of properties from simulations: free energies of solvation in hexadecane, octanol, and water. *J. Am. Chem. Soc.* **2000**, *122*, 2878–2888.
- (57) Colmenarejo, G.; Alvarez-Pedraglio, A.; Lavandera, J. L. Cheminformatic models to predict binding affinities to human serum albumin. *J. Med. Chem.* **2001**, *44*, 4370–4378.
- (58) Luco, J. M. 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.* **1999**, *39*, 396–404.
- (59) Kelder, J.; Grootenhuis, P. D.; Bayada, D. M.; Delbresine, L. P.; Ploemen, J. P. Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharm. Res.* **1999**, *16*, 1514–1519.
- (60) Ajay; Bermis, G. W.; Murkko, M. A. Designing libraries with CNS activity. *J. Med. Chem.* **1999**, *42*, 4942–4951.
- (61) Yazdanian, M.; Glynn, S. L.; Wright, J. L.; Hawi, A. Correlating partitioning and caco-2 cell permeability of structurally diverse small molecular weight compounds. *Pharm. Res.* **1998**, *15*, 1490–1494.
- (62) Irvine, J. D.; Takahashi, L.; Lockhart, K.; Cheong, J.; Tolan, J. W.; Selick, H. E.; Grove, J. R. MDCK (Madin-Darby canine kidney) cells: a tool for membrane permeability screening. *J. Pharm. Sci.* **1999**, *88*, 28–33.
- (63) Stenberg, P.; Norinder, U.; Luthman, K.; Artursson, P. Experimental and computational screening models for the prediction of intestinal drug absorption. *J. Med. Chem.* **2001**, *44*, 1927–1937.
- (64) Jorgensen, W. L.; Duffy, E. M. Prediction of drug solubility from Monte Carlo simulations. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1155–1158.
- (65) Cavalli, A.; Poluzzi, E.; De Ponti, F.; Recanatini, M. Toward a pharmacophore for drugs inducing the long QT syndrome: insights from a CoMFA study of HERG K⁺ channel blockers. *J. Med. Chem.* **2002**, *45*, 3844–3853.
- (66) De Ponti, F.; Poluzzi, E.; Montanaro, N. Organising evidence on QT prolongation and occurrence of *Torsades de Pointes* with non-antiarrhythmic drugs: a call for consensus. *Eur. J. Clin. Pharmacol.* **2001**, *57*, 185–209.
- (67) Potts, R. O.; Guy, R. H. Predicting skin permeability. *Pharm. Res.* **1992**, *9*, 663–669.
- (68) Potts, R. O.; Guy, R. H. A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bond activity. *Pharm. Res.* **1995**, *12*, 1628–1633.
- (69) van de Waterbeemd, H.; Gifford, E. ADMET *in silico* modelling: towards prediction paradise? *Nat. Rev. Drug Discovery* **2003**, *2*, 192–204.
- (70) Hedley, P. L.; Jørgensen, P.; Schlamowitz, S.; Wangari, R.; Moolman-Smook, J.; Brink, P. A.; Kanters, J. K.; Corfield, V. A.; Christiansen, M. The genetic basis of long QT and short QT syndromes: a mutation update. *Hum. Mutat.* **2009**, *30*, 1486–1511.
- (71) Vandenberg, J. I.; Walker, B. D.; Campbell, T. J. HERG K⁺ channels: friend or foe. *Trends Pharmacol. Sci.* **2001**, *22*, 240–246.
- (72) Chiesa, N.; Rosati, B.; Arcangeli, A.; Olivotto, M.; Wanke, E. A novel role for HERG K⁺ channels: spike-frequency adaptation. *J. Physiol.* **1997**, *501*, 313–318.
- (73) Aronov, A. M. Predictive *in silico* modeling for hERG channel blockers. *Drug Discovery Today* **2005**, *10*, 149–155.
- (74) QikProp 3.4 User Manual; Schrödinger Press LLC: New York, NY, 2011.
- (75) QikProp, version 3.4; Schrödinger LLC: New York, NY, 2011.
- (76) Matias, P. M.; Carrondo, M. A.; Coelho, R.; Thomaz, M.; Zhao, X. Y.; Wegg, A.; Crusius, K.; Egner, U.; Donner, P. Structural basis for the glucocorticoid response in a mutant human androgen receptor (AR^{ccr}) derived from an androgen-independent prostate cancer. *J. Med. Chem.* **2002**, *45*, 1439–1446.
- (77) Colombo, R.; Caldarelli, M.; Mennecozzi, M.; Giorgini, M. L.; Sola, F.; Cappella, P.; Perrera, C.; Depaolini, S. R.; Rusconi, L.; Cucchi, U.; Avanzi, N.; Bertrand, J. A.; Bossi, R. T.; Pesenti, E.; Galvani, A.; Isacchi, A.; Colotta, F.; Donati, D.; Moll, J. Targeting the mitotic checkpoint for cancer therapy with nms-p715, an inhibitor of mps1 kinase. *J. Cancer Res.* **2010**, *70*, 10255–10264.
- (78) Battistutta, R.; Cozza, G.; Pierre, F.; Papinutto, E.; Lolli, G.; Sarno, S.; O'Brien, S. E.; Siddiqui-Jain, A.; Haddach, M.; Anderes, K.; Ryckman, D. M.; Meggio, F.; Pinna, L. A. Unprecedented selectivity and structural determinants of a new class of protein kinase CK2 inhibitors in clinical trials for the treatment of cancer. *Biochemistry* **2011**, *50*, 8478–8488.
- (79) Pai, E. F.; Krengel, U.; Petsko, G. A.; Goody, R. S.; Kabsch, W.; Wittinghofer, A. Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **1990**, *9*, 2351–2359.
- (80) Deshpande, A.; Xia, G.; Boerma, L. J.; Vines, K. K.; Atigadda, V. R.; Lobo-Ruppert, S.; Grubbs, C. J.; Moeinpour, F. L.; Smith, C. D.; Christov, K.; Brouillette, W. J.; Muccio, D. D. Methyl-substituted conformationally constrained rexinoid agonists for the retinoid X receptors demonstrate improved efficacy for cancer therapy and prevention. *Bioorg. Med. Chem.* **2014**, *22*, 178–185.

- (81) Hast, M. A.; Fletcher, S.; Cummings, C. G.; Pusateri, E. E.; Blaskovich, M. A.; Rivas, K.; Gelb, M. H.; Van Voorhis, W. C.; Sebti, S. M.; Hamilton, A. D.; Beese, L. S. Structural basis for binding and selectivity of antimalarial and anticancer ethylenediamine inhibitors to protein farnesyltransferase. *Chem. Biol.* **2009**, *16*, 181–192.
- (82) Dong, Q.; Dougan, D. R.; Gong, X.; Halkowycz, P.; Jin, B.; Kanouni, T.; O'Connell, S. M.; Scorah, N.; Shi, L.; Wallace, M. B.; Zhou, F. Discovery of TAK-733, a potent and selective MEK allosteric site inhibitor for the treatment of cancer. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1315–1319.
- (83) Good, J. A. D.; Wang, F.; Rath, O.; Kaan, H. Y. K.; Talapatra, S. K.; Podgórski, D.; MacKay, S. P.; Kozielski, F. Optimized S-trityl-L-cysteine-based inhibitors of kinesin spindle protein with potent *in vivo* antitumor activity in lung cancer xenograft models. *J. Med. Chem.* **2013**, *S6*, 1878–1893.
- (84) Shima, F.; Ijiri, Y.; Muraoka, S.; Liao, J.; Ye, M.; Araki, M.; Matsumoto, K.; Yamamoto, N.; Sugimoto, T.; Yoshikawa, Y.; Kumasaki, T.; Yamamoto, M.; Tamura, A.; Kataoka, T. Structural basis for conformational dynamics of GTP-bound Ras protein. *J. Biol. Chem.* **2010**, *285*, 22696–22705.
- (85) Wang, S.; Griffiths, G.; Midgley, C. A.; Barnett, A. L.; Cooper, M. Discovery and characterization of 2-anilino-4-(thiazol-5-yl)pyrimidine transcriptional CDK inhibitors as anticancer agents. *Chem. Biol.* **2010**, *17*, 1111–1121.
- (86) Berg, S.; Bergh, M.; Hellberg, S.; Hogdin, K.; Lo-Alfredsson, Y.; Söderman, P.; von Berg, S.; Weigelt, T.; Ormö, M.; Xue, Y.; Tucker, J.; Neelissen, J.; Jerning, E.; Nilsson, Y.; Bhat, R. Discovery of novel potent and highly selective glycogen synthase kinase-3 β (GSK3 β) inhibitors for Alzheimer's disease: design, synthesis, and characterization of pyrazines. *J. Med. Chem.* **2012**, *55*, 9107–9119.
- (87) Nagar, B.; Bornmann, W.; Pellicena, P.; Schindler, T.; Veach, D. R.; Miller, W. T.; Clarkson, B.; Kuriyan, J. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res.* **2002**, *62*, 4236–4243.
- (88) Millard, C. J.; Watson, P. J.; Celardo, I.; Gordiyenko, Y.; Cowley, S. M.; Robinson, C. V.; Fairall, L.; Schwabe, J. W. Class I HDACs share a common mechanism of regulation by inositol phosphates. *Mol. Cell* **2013**, *51*, 57.
- (89) Prota, A. E.; Danel, F.; Bachmann, F.; Bargsten, K.; Buey, R. M.; Pohlmann, J.; Reinelt, S.; Lane, H.; Steinmetz, M. O. The novel microtubule-destabilizing drug BAL27862 binds to the colchicine site of tubulin with distinct effects on microtubule organization. *J. Mol. Biol.* **2014**, *426*, 1848–1860.
- (90) Zhao, X. Y.; Malloy, P. J.; Krishnan, A. V.; Swami, S.; Navone, N. M.; Peehl, D. M.; Feldman, D. Glucocorticoids can promote androgen-independent growth of prostate cancer cells through a mutated androgen receptor. *Nature Med.* **2000**, *6*, 703–706.
- (91) Marcus, A. I.; Peters, U.; Thomas, S. L.; Garrett, S.; Zelnak, A.; Kapoor, T. M.; Giannakakou, P. Mitotic kinesin inhibitors induce mitotic arrest and cell death in Taxol-resistant and -sensitive cancer cells. *J. Biol. Chem.* **2005**, *280*, 11569–11577.
- (92) Goodsell, D. S. The molecular perspective: The ras Oncogene. *Oncologist* **1999**, *4*, 263–264.
- (93) Malumbres, M.; Barbacid, M. Mammalian cyclin-dependent kinases. *Trends Biochem. Sci.* **2005**, *30*, 630–641.
- (94) Coghlan, M. P.; Culbert, A. A.; Cross, D. A.; Corcoran, S. L.; Yates, J. W.; Pearce, N. J.; Rausch, O. L.; Murphy, G. J.; Carter, P. S.; Roxbee, C. L.; Mills, D.; Brown, M. J.; Haigh, D.; Ward, R. W.; Smith, D. G.; Murray, K. J.; Reith, A. D.; Holder, J. C. Selective small molecule inhibitors of glycogen synthase kinase-3 modulate glycogen metabolism and gene transcription. *Chem. Biol.* **2000**, *7*, 793–803.
- (95) Kassem, M. E. S.; Shoela, S.; Marzouk, M. M.; Sleem, A. A. A sulphated flavone glycoside from *Livistona australis* and its antioxidant and cytotoxic activity. *Nat. Prod. Res.* **2012**, *26*, 1381–1387.
- (96) Stiffness, M.; Pezzuto, J. M. Assays related to cancer drug discovery. In *Assays for Bioactivity; Methods in Plant Biochemistry*, Vol. 6; Hostettmann, K., Ed.; Academic Press: London, 1990; pp 71–153.
- (97) Ndontsa, B. L.; Tala, M. F.; Talontsi, F. M.; Wabo, H. K.; Tene, M.; Laatsch, H.; Tane, P. New cytotoxic alkylbenzoquinone derivatives from leaves and stem of *Ardisia kivuensis* (Myrsinaceae). *Phytochem. Lett.* **2012**, *5*, 463–466.
- (98) Ndontsa, B. L.; Tatximo, J. S. N.; Csopor, D.; Forgo, P.; Berkecz, R.; Berényi, A.; Tene, M.; Molnár, J.; Zupkó, I.; Hohmann, J.; Tane, P. Alkylbenzoquinones with antiproliferative effect against human cancer cell lines from stem of *Ardisia kivuensis*. *Phytochem. Lett.* **2011**, *4*, 227–230.
- (99) Lifongo, L. L.; Simoben, C. V.; Ntie-Kang, F.; Babiaka, S. B.; Judson, P. N. A bioactivity versus ethnobotanical survey of medicinal plants from Nigeria, West Africa. *Nat. Prod. Bioprospect.* **2014**, *4*, 1–19.
- (100) *Molecular Operating Environment*, version 2010; Chemical Computing Group Inc: Montreal, Canada, 2010.
- (101) Halgren, T. A. Merck molecular forcefield. *J. Comput. Chem.* **1996**, *17*, 490–641.
- (102) *LigPrep*, version 2.5; Schrödinger LLC: New York, NY, 2011.
- (103) *Maestro*, version 9.2; Schrödinger LLC: New York, NY, 2011.
- (104) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comput.* **2010**, *6*, 1509–1519.
- (105) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (106) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- (107) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–342.
- (108) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.
- (109) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (110) Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. A new test set for validating predictions of protein-ligand interaction. *Proteins* **2002**, *49*, 457–471.
- (111) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623.
- (112) Cole, J. C.; Nissink, J. W. M.; Taylor, R. Protein-ligand docking and virtual screening with GOLD. In *Virtual Screening in Drug Discovery*; Shoichet, B., Alvarez, J., Eds.; Taylor & Francis CRC Press: Boca Raton, FL, 2005.
- (113) Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. Modeling water molecules in protein-ligand docking using GOLD. *J. Med. Chem.* **2005**, *48*, 6504–6515.
- (114) Hartshorn, M. J.; Verdonk, M. L.; Chessari, G.; Brewerton, S. C.; Mooij, W. T. M.; Mortenson, P. N.; Murray, C. W. Diverse, high-quality test set for the validation of protein-ligand docking performance. *J. Med. Chem.* **2007**, *50*, 726–741.
- (115) *Glide*, version 5.8; Schrödinger LLC: New York, NY, 2012.
- (116) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–619.
- (117) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shaw, D. E.; Shelley, M.; Perry, J. K.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (118) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A new approach for rapid,

accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.