RESEARCH ARTICLE

View Article Online
View Journal | View Issue Cite this: *Med. Chem. Commun.*,
2017, 8, 1152Identification of steroid-like natural products as
antiplasmodial agents by 2D and 3D similarity-
based virtual screening†‡Elumalai Pavadai,§¶*^a Gurminder Kaur,§^a Sergio Wittlin^{bc} and Kelly Chibale^{†d} ^{*ade}

The emergence of drug resistance in *Plasmodium falciparum* to available antimalarial drugs has challenged current antimalarial treatments. New antimalarials, particularly those with novel mechanisms of action and no cross resistance to current drugs, are therefore urgently needed. To identify new growth inhibitors of *Plasmodium falciparum*, 2D and 3D similarity-based virtual screening methods were employed in parallel with an in-house database of steroid-type natural products using fusidic acid as a search query. The resulting hit compounds were further filtered based on the predicted partition coefficient, log *P*. The virtual screening strategy resulted in the identification of nine new compounds that inhibited parasite growth with IC₅₀ values of <20 μM. Four compounds exhibited IC₅₀ values in the range of 1.39–3.45 μM and three of which showed a promising selectivity index. Further, the predicted ADME properties of the four most active compounds were found to be comparable to fusidic acid. These compounds can be further explored using structural modifications in the identification and development of more potent parasite growth inhibitors with improved selectivity.

Received 4th February 2017,
Accepted 17th March 2017

DOI: 10.1039/c7md00063d

rsc.li/medchemcomm

Introduction

Malaria, a life-threatening parasitic disease, remains a global public health concern and is responsible for millions of infections each year. The World Health Organization recently reported that in 2015 alone, malaria caused 214 million infections and resulted in over half a million deaths.¹ The burden of the disease is heaviest in the African region where an estimated 88% of all malaria deaths occur and children under 5 are the most at risk.¹ Of the five *Plasmodium* parasites that cause malaria, which spreads through the bites of infected *Anopheles* mosquitoes, *Plasmodium falciparum* (*P. falciparum*) is responsible for most malaria infections and deaths. Since no effective malaria vaccines are available to this date, antima-

larial control programs and chemotherapy are used for prevention and treatment of this disease, respectively. However, *P. falciparum* has developed resistance to antimalarial drugs such as chloroquine resulting in their withdrawal in many regions.² In addition, the parasite shows resistance to nearly all antimalarial drugs in current use, including the artemisinin combination therapy (ACT) regimen, a new treatment option to combat antimalarial drug resistance.^{3–5} Thus the discovery and development of novel antimalarial drugs is urgently needed.

Fusidic acid (Fig. 1), a steroid-type and narrow-spectrum antibiotic, has been used since the 1960s for the clinical treatment of infections caused by the Gram-positive bacterium, *Staphylococcus aureus*.^{6–8} This compound inhibits bacterial protein synthesis by locking elongation factor G (EF-G) on the ribosome with guanosine 5'-diphosphate (GDP) in a post-translational state.^{9,10} Besides being an effective anti-bacterial agent, fusidic acid has also shown antiplasmodial activity *in vitro* against the

Fig. 1 Fusidic acid.

^a Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa. E-mail: elumalai.pavadai.a@gmail.com, Kelly.Chibale@uct.ac.za; Fax: +27 21 650 5195; Tel: +1 (305) 348 6661, +27 21 650 2553

^b Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland

^c University of Basel, 4002 Basel, Switzerland

^d Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Rondebosch 7701, South Africa

^e South African Medical Research Council Drug Discovery and Development Research Unit, University of Cape Town, Rondebosch 7701, South Africa

† The authors declare no competing interests.

‡ Electronic supplementary information (ESI) available. See DOI: 10.1039/c7md00063d

§ Equally contributed.

¶ Current address: Department of Physics, Florida International University, Miami 33199, FL, USA.

D10 strain of *P. falciparum* with an IC_{50} of 52.8 μM .¹¹ We recently synthesized a series of fusidic acid derivatives by replacing the carboxylic acid group with various bioisosteres, which resulted in active inhibitors against the chloroquine-sensitive *P. falciparum* NF54 strain (PfNF54).¹² Though the mode of action of fusidic acid against *P. falciparum* is unclear, a few studies have demonstrated that it targets EF-Gs of *P. falciparum* (PfEF-Gs) located in the apicoplast and mitochondria.^{11,13}

Ligand and structure-based approaches are well-established computational methods, which are employed to identify new chemotypes in the modern era of drug discovery.¹⁴ Ligand-based methods use known active ligands, while structure-based methods use the 3D structure of a biological target. The former has become the approach of choice when the 3D-structure of a biological target is unknown. Similarity-search is one such ligand-based approach that searches a database for compounds that most closely resemble the 2D structure or 3D shape of a given query molecule.¹⁵ 2D and 3D similarity-search methods have been successfully employed in parallel or in combination to identify novel compounds for various targets.^{16–18}

This communication reports the findings of a study in which 2D and 3D similarity-based virtual screening approaches were employed to identify fusidic acid-like structurally diverse compounds from an in-house database using fusidic acid as

a search query. Selected hit compounds were evaluated *in vitro* for antiparasmodial activity against the *P. falciparum* NF54 strain to reveal new active antiparasmodial compounds.

Results and discussion

Two different similarity-search virtual screening approaches, 2D and 3D similarity-searches, were performed in parallel, using fusidic acid as a search query molecule (Fig. 1), with the aim of identifying fusidic acid-like structurally diverse compounds with improved biological activity from our in-house database containing 708 steroid-type natural products and their semi-synthetic derivatives. Although fusidic acid and most compounds in the database have a common tetracyclic ring system, they are structurally diverse with varying substituents and substitution patterns. In addition, one of the aliphatic rings of the tetracyclic ring system in some of the database compounds is replaced by an aromatic ring. A schematic representation of the screening cascade is shown in Fig. 2.

2D fingerprint methods were first adopted for a 2D similarity-based virtual screening campaign. The fingerprint is the characterization of 2D molecular structure in a binary bit format.^{19,20} These methods are frequently used as a measure for structural similarity between molecules, which is computed

Fig. 2 Schematic representation of virtual screening strategy adopted in this study.

by overlapping between the fingerprints of molecules.²¹ The 2D fingerprint methods have become the approach of choice for 2D similarity-based virtual screening because of its computational efficiency and effectiveness, and do not need the generation of the 3D conformation of molecules for screening. The usefulness of these fingerprint methods has also been suggested for scaffold hopping studies.²²

At the time of commencement of the current study, there were no fusidic acid derivatives known to be active against *Pf*NF54 available for the validation of 2D and 3D similarity methods. Therefore, a total of 13 fingerprints of Discovery Studio 4.0 (Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, San Diego, USA) including MDLPublicKeys and 12 extended connectivity fingerprints (ESI† S1) were investigated with the in-house database to choose the most suitable 2D fingerprint methods for the virtual screening. Fusidic acid was used as a search query molecule and the widely used Tanimoto coefficient (Tc) (ESI† S1) was employed as the evaluation criterion for selecting fingerprint methods and ranking the hit compounds for all 2D similarity-search methods. We considered that compounds with a Tc value cut-off of ≥ 0.6 are active and therefore the compounds with a Tc value of < 0.6 are inactive. Hence, we selected the fingerprints that were able to filter the database compounds with the Tc value of ≥ 0.6 . Of thirteen fingerprints, FCFP_2, ECFC_4 and FCFC_4 were capable of identifying compounds with Tc values of > 0.6 , and thus they were selected and subsequently employed in parallel for the fingerprint-based similarity search with the in-house database. The top 25-ranked compounds from each of the FCFP_2, ECFC_4 and FCFC_4 were selected. The Tc values for FCFP_2, ECFC_4 and FCFC_4 were > 0.66 , > 0.6 and > 0.92 , respectively (Table 1).

The second similarity method that was used to find compounds with the desired biological profile was the 3D shape-based similarity search. Shape similarity is an alignment of the 3D shape of molecules, which is based on the principle that molecules can be similar if their shapes overlap well. The shape-based screening is effectively used to optimize leads and to understand structure–activity relationships by superposition of similar compounds. Additionally, many studies have shown its successful applications in virtual screening experiments.^{23,24} In this study, the phase-shape screening tool of Schrodinger (Small-Molecule Drug Discovery Suite Phase, Schrödinger, LLC, New York, USA) was employed for 3D shape-based virtual screening. This tool uses geometric overlapping between the conformations of a given molecule and the search query to compute shape similarity. Finally, the alignment returns with the best similarity score for a given molecule.

A total of four phase-shape methods of Schrodinger (ESI† S2) were initially investigated with the in-house database to choose the most suitable 3D shape methods for the virtual screening. For this, the lowest energy conformation of fusidic acid, a reasonable approximation of the bioactive geometry generated by ConfGen, was used as a 3D search query. The shape similarity score was used as the evaluation criterion for selecting phase-shape methods and ranking the hit compounds. Shape_element and Shape_pharm were employed in

parallel to screen the 3D conformers of the in-house database compounds generated by ConfGen (ESI† S2). The top 20-ranked compounds from each of the Shape_element and Shape_pharm were selected, which had shape similarity scores of > 0.69 and > 0.39 for Shape_element and Shape_pharm, respectively.

The 2D and 3D similarity-based virtual screening yielded a total of 115 hit compounds and the removal of duplicates resulted in 79 compounds. To obtain compounds with the desired lipophilic properties, these hits were subjected to log *P* based screening. The log *P*, known as the octanol/water partition coefficient, is a measure of the lipophilicity of the drug molecule, and plays a key role in biological processes such as metabolism and penetration across the cell membranes. The log *P* values for 79 hit compounds were predicted using StarDrop software (Optibrium, Ltd, Cambridge, UK). Any compound having log *P* ≥ 5.5 was discarded. Based on this, 41 hit compounds were selected for further analysis. These 41 compounds were clustered using ECFP_4 and each of them was visually inspected. Based on the structural diversity and sample availability, 27 compounds were finally selected and tested for *in vitro* activity against *P. falciparum* and for cytotoxicity against the mammalian Chinese Hamster Ovarian (CHO) cell line. The purity of the tested compounds was determined to be $\geq 95\%$.

The antiplasmodial activity of all selected compounds was evaluated against a chloroquine-sensitive (NF54) strain of *P. falciparum* (*Pf*NF54) using the modified [³H]hypoxanthine incorporation assay (ESI† S3). Chloroquine, Artesunate, and fusidic acid were used as positive controls, which showed IC₅₀ values of 0.016, 0.004 and 59 μ M, respectively. Nine out of 27 compounds showed IC₅₀ values of ≤ 20 μ M (Table 1). However, none of the compounds, except fusidic acid, was tested beyond this concentration. The most active compounds were 2, 3, 4, 5, 6 and 7, which exhibited IC₅₀ values of 1.39, 1.76, 2.92, 3.45, 6 and 7 μ M, respectively (Table 1). In addition, three compounds (8, 9 and 10) showed IC₅₀ values ranging from 13 to 20 μ M and other compounds showed IC₅₀ values of > 20 μ M (Table S1†). The antiplasmodial activity of identified hit compounds is superior to that of the search query fusidic acid. This indicates that the employed protocol of 2D and 3D similarity-search approaches is capable of identifying fusidic acid-like structurally diverse compounds with improved biological activity. In addition, it is noteworthy that 3D shape similarity search methods identified 7 active compounds while 2D fingerprint search methods identified 5 active compounds. Three of these 12 active compounds were common in both types of searches.

The *in vitro* cytotoxicity test, which uses specific cell lines, is carried out in the early phases of drug discovery in order to ensure that cytotoxic compounds are not further progressed. For this purpose, the most active compounds 2, 3, 4 and 5 were evaluated for *in vitro* cytotoxicity against a mammalian cell-line, CHO (ESI† S4). Emetine was used as a positive control, which showed an IC₅₀ value of 0.05 μ M. Compounds 2, 3, 4 and 5 displayed cytotoxicity with IC₅₀

Table 1 Nine hit compounds (**2–10**) that showed $IC_{50} < 20 \mu M$ against *PfNF54* and their inhibitory activities, cytotoxicity, log *P* and similarity score

Compd ID	Structure	<i>PfNF54</i> IC_{50}^a (μM)	CHO IC_{50} (μM)	log <i>P</i>	Similarity score
1	Fusidic acid	59	194	4.45	ND
2^b		1.39	1.99	2.87	0.67
3^c		1.76	8.51	2.67	0.62
4^d		2.92	55.2	2.26	0.92
5^e		3.45	141	4.63	0.63
6^e		6.00	ND	4.66	0.75
7^e		7.00	ND	4.13	0.7
8^b		13.06	ND	5.37	0.68
9^f		15.80	ND	3.95	0.40
10^f		19.76	ND	2.92	0.39

Table 1 (continued)

Compd ID	Structure	<i>Pf</i> NF54 IC ₅₀ ^a (μM)	CHO IC ₅₀ (μM)	log <i>P</i>	Similarity score
	Chloroquine	0.016	ND	ND	ND
	Artesunate	0.004	ND	ND	ND
	Emetine	ND	0.05	ND	ND

ND: not determined. ^a Mean from at least 2 independent experiments with chloroquine-sensitive strain (NF54) of *P. falciparum*. ^b Compound screened by FCFP_2. ^c Compound screened by ECFC_4. ^d Compound screened by FCFC_4. ^e Compound screened by Shape_ele. ^f Compound screened by Shape_pharm.

Table 2 Predicted ADMET properties of fusidic acid and identified four active hit compounds that showed *Pf*NF54 < 5 μM

Properties	Fusidic acid	2	3	4	5
log <i>S</i> @pH 7.4	1.992	1.576	1.457	1.457	0.7515
Human intestinal absorption	+	+	+	+	–
P-glycoprotein substrate	Yes	Yes	Yes	Yes	No
Plasma protein binding	High	Low	Low	Low	Low
BBB log ([brain]:[blood])	–0.7931	–1.051	–0.9395	–0.8719	–0.3384
Blood–brain barrier penetration category	–	–	–	–	–
Caco-2 cell permeability in nm s ^{–1}	52.648	59.698	152.892	84.888	278.188
Cytochrome P450 CYP2C9 p <i>K</i> _i	4.883	5.034	5.114	5.09	4.867
Cytochrome P450 CYP2D6 affinity	Very high	High	High	High	Very high
hERG pIC ₅₀	4.357	4.16	4.447	4.442	4.42
Ames mutagenicity test	Non-mutagen	Non-mutagen	Non-mutagen	Non-mutagen	Non-mutagen

log *S* (μM) = solubility; p*K*_i = –log₁₀ *K*_i (μM); hERG pIC₅₀ = –log₁₀ IC₅₀ (μM); 25 nm s^{–1} (poor) < Caco-2 permeability >500 nm s^{–1}, (great); Ames mutagenicity was computed by Discovery Studio 4.0; Caco-2 cell permeability was computed by QikProp of Schrodinger.

values of 1.99, 8.51, 55.2, and 141 μM, respectively (Table 1). The selectivity index (SI), calculated as the ratio of the CHO IC₅₀ to the *Pf*NF54 IC₅₀, was found to be 1.43 for 2, 4.83 for 3, 18.90 for 4, and 40.87 for compound 5. A SI > 10 is desirable as it suggests selective antiparasmodial activity and the potential for a desirable safety profile. In contrast, a low SI suggests that the antiparasmodial activity may be due to compound cytotoxicity rather than specific activity against the parasite itself. Compounds 4 and 5, in addition to superior antiparasmodial activity compared to fusidic acid, also displayed a greatly improved selectivity. It is noteworthy that compounds 2–4 are cucurbitacins, which were first identified in plants of the family Cucurbitaceae. To date, no fewer than twenty distinct cucurbitacins have been reported in the literature.²⁵

In addition to early detection of cytotoxicity, early optimization of ADME (Absorption, Distribution, Metabolism, Excretion) properties of a drug candidate is also an important step in drug discovery process as they have a major impact on the likelihood of success of a drug. Most drug candidates fail in the clinical trials due to a poor ADME profile. Therefore, the ADME properties of compounds 2, 3, 4, and 5 were predicted using StarDrop 5.5 (Optibrium, Ltd, Cambridge, UK), Discovery Studio 4.0 and QikProp of Schrodinger, and were compared with those of the fusidic acid (Table 2). Most of the predicted ADME properties of these compounds are comparable to fusidic acid, indicating that the hit compounds have acceptable ADME properties.

In conclusion, 2D fingerprint- and 3D-shape-based virtual screening was performed with the in-house database using

fusidic acid as a search query to identify fusidic acid-like structurally diverse growth inhibitors of *P. falciparum*. The virtual screening strategy successfully identified nine hit compounds that inhibited growth of the *Pf*NF54 parasite with IC₅₀ values of <20 μM. Four of these compounds (2, 3, 4 and 5) displayed IC₅₀ values between 1 and 4 μM, which were better than fusidic acid (59 μM). Compounds 4, and 5 also displayed promising selective antiparasmodial activity. In addition, the predicted ADME properties of these four compounds were comparable to fusidic acid. It could be expected that the systematic SAR analysis of identified hit compounds may be useful for the identification and development of more potent malaria parasite growth inhibitors with better SI values.

Acknowledgements

The University of Cape Town, South African Medical Research Council, and South African Research Chairs initiative of the Department of Science and Technology administered through the South African National Research Foundation are gratefully acknowledged for support (KC). The authors would like to acknowledge the Centre of High Performance Computing, Cape Town (<http://www.chpc.ac.za>) for Discovery Studio 4.0 license.

References

- 1 World Health Organization, *World Malaria Report*, 2015.
- 2 A. M. Dondorp, S. Yeung, L. White, C. Nguon, N. P. J. Day, D. Socheat and L. von Seidlein, *Nat. Rev. Microbiol.*, 2010, 8, 272–280.

- 3 A. M. Dondorp, F. Nosten, P. Yi, D. Das, A. P. Phyto, J. Tarning, D. Ph, K. M. Lwin, F. Arie, W. Hanpithakpong, S. J. Lee, P. Ringwald, K. Silamut, T. Herdman, S. S. An, S. Yeung, D. Socheat and N. J. White, *N. Engl. J. Med.*, 2009, **361**, 455–467.
- 4 T. E. Wellems and C. V. Plowe, *J. Infect. Dis.*, 2001, **184**, 770–776.
- 5 I. Petersen, R. Eastman and M. Lanzer, *FEBS Lett.*, 2011, **585**, 1551–1562.
- 6 B. Long, *Acta Derm.-Venereol.*, 2008, **88**, 14–20.
- 7 L. Verbist, *J. Antimicrob. Chemother.*, 1990, **25**, 1–5.
- 8 L. T. W. Godtfredsen and K. Roholt, *Lancet*, 1962, **1**, 928–931.
- 9 W. Bodley and J. Zieve, *J. Biol. Chem.*, 1970, **45**, 5662–5667.
- 10 E. Cundliffe, *Biochem. Biophys. Res. Commun.*, 1972, **46**, 1794–1801.
- 11 R. A. Johnson, G. I. McFadden and C. D. Goodman, *PLoS One*, 2011, **6**, e20633.
- 12 G. Kaur, K. Singh, E. Pavada, M. Njoroge, M. Espinoza-moraga, C. De Kock, P. J. Smith, S. Wittlin and K. Chibale, *MedChemComm*, 2015, **6**, 2023–2028.
- 13 A. Gupta, S. S. Mir, U. Saqib, S. Biswas, S. Vaishya, K. Srivastava, M. I. Siddiqi and S. Habib, *Mol. Biochem. Parasitol.*, 2013, **192**, 39–48.
- 14 R. Braga, V. Alves, A. Silva, M. Nascimento, F. Silva, L. Liao and C. Andrade, *Curr. Top. Med. Chem.*, 2014, **14**, 1899–1912.
- 15 *Concepts and Applications of Molecular Similarity*, ed. M. A. Johnson and G. M. Maggiora, John Wiley & Sons, 1990.
- 16 G. Hu, G. Kuang, W. Xiao, W. Li, G. Liu and Y. Tang, *J. Chem. Inf. Model.*, 2012, **52**, 1103–1113.
- 17 S. Distinto, F. Esposito, J. Kirchmair, M. C. Cardia, M. Gaspari, E. MacCioni, S. Alcaro, P. Markt, G. Wolber, L. Zinzula and E. Tramontano, *Eur. J. Med. Chem.*, 2012, **50**, 216–229.
- 18 K. Dobi, I. Hajdú, B. Flachner, G. Fabó, M. Szaszko, M. Bognár, C. Magyar, I. Simon, D. Szisz, Z. Lőrincz, S. Cseh and G. Dormán, *Molecules*, 2014, **19**, 7008–7039.
- 19 T. E. Johann Gasteiger, *Chemoinformatics: A Textbook*, John Wiley & Sons, 2006.
- 20 A. R. Leach and V. J. Gillet, *An introduction to chemoinformatics*, Springer Science and Business Media, 2007.
- 21 D. Rogers and M. Hahn, *J. Chem. Inf. Model.*, 2010, **50**, 742–754.
- 22 E. J. Gardiner, J. D. Holliday, C. O'Dowd and P. Willett, *Future Med. Chem.*, 2011, **3**, 405–414.
- 23 T. S. Rush, J. A. Grant, L. Mosyak and A. Nicholls, *J. Med. Chem.*, 2005, **48**, 1489–1495.
- 24 D. S. Chekmarev, V. Kholodovych, K. V. Balakin, Y. Ivanenkov, S. Ekins and W. J. Welsh, *Chem. Res. Toxicol.*, 2008, **21**, 1304–1314.
- 25 J. C. Chen, M. H. Chiu, R. L. Nie, G. A. Cordell and S. X. Qiu, *Nat. Prod. Rep.*, 2005, **22**, 386–399.