The use of isolated natural products as scaffolds for the generation of chemically diverse screening libraries for drug discovery

Author

Barnes, Emma C, Kumar, Rohitesh, Davis, Rohan A

Published

2016

Journal Title

Natural Product Reports

Version

Accepted Manuscript (AM)

DOI

10.1039/C5NP00121H

The Use of Isolated Natural Products as Scaffolds for the Generation of Chemically Diverse Screening Libraries for Drug Discovery

Journal:	*Natural Product Reports*
Manuscript ID	NP-HIG-09-2015-000121.R1
Article Type:	Highlight
Date Submitted by the Author:	27-Nov-2015
Complete List of Authors:	Barnes, Emma; Griffith University, Eskitis Institute for Drug Discovery; Leibniz Institute for Natural Product Research and Infection Biology HKI, Kumar, Rohitesh; Griffith University, Eskitis Institute for Drug Discovery Davis, Rohan; Griffith University, Eskitis Institute for Drug Discovery

Natural Product Reports

HIGHLIGHT

The Use of Isolated Natural Products as Scaffolds for the Generation of Chemically Diverse Screening Libraries for Drug Discovery

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Emma C. Barnes,[a,b] Rohitesh Kumar,[a] Rohan A. Davis[*a]

Covering: 1980 to end 2014

A diverse range of strategies leading to natural product derived or inspired screening libraries aims to increase the number of new chemical entities emerging per year. However, the use of isolated natural products as scaffolds for the semi-synthesis of larger biological screening libraries remains rare. This particular method avoids the time-consuming and resource intensive *de novo* synthetic strategy for scaffold production, and has become more feasible through improvements to synthetic and isolation methodologies. This Highlight examines the increasing popularity of small- to large-sized screening libraries generated directly from isolated natural products. Several of the examples detailed herein show how this strategy can lead to improvements in not only potency but also other important (and often forgotten) drug discovery parameters such as toxicity, selectivity, lipophilicity and bioavailability. However, there are still improvements to be made to this method, particularly in the choice of the natural product scaffold and the derivatising reagents used. Avoidance of known nuisance compounds or structural alert motifs (e.g. PAINS) that interfere with bioactivity screens, and impact downstream drug development will play a significant role in the future success of this methodology. Incorporation of rational design strategies that take into account the physicochemical parameters (e.g. log P, MW, HBA, HBD) of the final semi-synthetic library analogues will also facilitate the discovery and development of leads and drugs. A multi-pronged approach to drug discovery that incorporates the use of isolated natural product scaffolds for library generation will surely be beneficial.

1. Introduction

"Structurally diverse", "drug-like", "lead-like", "privileged structures," "biologically validated" – these are just some of the terms used to describe natural products (NPs).[1-6] Computational studies have shown that NPs occupy larger, complementary areas of chemical space compared to synthetic compounds, and typically possess a higher number of stereogenic centres, different proportions of heteroatoms, and more diverse core ring scaffolds.[3, 5] This suggests that by mimicking the distribution properties of NPs we can increase the chemical complexity of discovery libraries, which may in turn result in a higher number of biological hits.[1, 4, 5]

Despite the overall number of new chemical entities (NCEs) having fallen per annum in recent years, it is interesting to note that the number of NP and NP-derived NCEs has remained relatively high. Indeed, NPs either directly or indirectly afforded around 39% of all small-molecule drugs approved between 1981-2010.[2, 7]

Contemporary strategies for using NPs to generate new drugs and/or leads has moved beyond *de novo* syntheses followed by analogue generation, to include techniques such as: the use of simplified core motifs found in NPs, diverted total synthesis and diversity-orientated synthesis to explore a greater a number of potentially useful scaffolds inspired by Nature, and the development of NP-like libraries.[1, 5, 8]

But what about the use of isolated NPs for the generation of libraries? For decades chemists have utilised isolated NPs to synthesise structurally complex bioactive compounds and small sets of analogues. However, employing traditional notions of semi-synthesis to construct larger libraries for biological screening is still rare. Well-established NP and biota collections, such as Griffith University's Nature Bank,[9-11] have the potential to provide useful scaffolds, so that the *de novo* synthetic strategy for scaffold production can be by-passed, thus saving considerable time and resources. Such an approach would complement, and no doubt expedite, current NP drug discovery methods that typically involve the high-throughput screening of extract or fraction libraries.

A major focus of our research is the design and synthesis of drug discovery libraries based on unique NP scaffolds.[9-11] This prompted us to investigate the popularity of semi-synthetic

[a] *Eskitis Institute for Drug Discovery, Griffith University, Brisbane, QLD, 4111, Australia. Tel: +61-7-3735-6043. Fax: +61-7-3735-6001. E-mail: r.davis@griffith.edu.au.*
[b] *Leibniz Institute for Natural Product Research and Infection Biology HKI, Beutenbergstr. 11a, Jena, Germany.*

compound libraries, and particularly how this strategy has been used to generate small- to large-sized libraries of drug- or lead-like molecules.

The fact that many secondary metabolites possess favourable physicochemical properties and are biologically active does not mean they automatically qualify as the best leads or drugs.[6, 12] The generation of analogues to improve such factors as the potency, toxicity, and metabolism of any NP and to access structure-activity relationship (SAR) information remains important for the selection of leads for progression.[6, 13] Increasingly, physicochemical parameters such as Lipinski's "Rule of Five" (MW ≤ 500; HBA ≤ 10; HBD ≤ 5; Log P ≤ 5) for drug-like molecules are being calculated and carefully analysed in an effort to triage and prioritise scaffolds and derivatives that have a greater propensity for oral bioavailability.[7, 12, 14]

Drug-hunters are also becoming more skilled at recognising chemical motifs that interfere with certain assay technologies (e.g. 2-aminothiazoles), or have the potential to compromise downstream drug development (e.g. Michael acceptors). Medicinal chemists are becoming more familiar with pan-assay interference compounds (PAINS),[15] since the pursuit of these sorts of molecules increases both the time and expense of not only screening but also lead optimisation campaigns, due to increased failure rates. Their removal from screening decks facilitates detection of more viable hits.[15]

It may also be beneficial to reassess how large a library of compounds needs to be in order to add value to a drug discovery screening program.[13, 16] For this publication, a library was defined as a group of ten or more analogues arising from the chemical elaboration of one or more chemical handles present on a NP scaffold. Such small libraries are more accessible when taking into account the often low isolation yields of NPs, and can provide both new bioactive entities and important structure-activity information, as the examples in this Highlight will testify.

For this Highlight it was not our intention to discuss more familiar and well-reviewed examples, nor was it our goal to perform an exhaustive literature review on every report of an analogue library that has been synthesised using an isolated NP scaffold. Instead, we have focussed on less well-known examples used for drug discovery purposes and the concepts behind such work.

2. Examples of analogue libraries based on isolated NP scaffolds

Inula britannica, a traditional Chinese medicinal plant, was extracted by Dong *et al.* in order to access large quantities of 1-*O*-acetylbritannilactone (**1**) for cytotoxicity studies.[17, 18] This compound had previously been found to possess several biological activities such as anti-cancer, anti-inflammatory, anti-bacterial, anti-hepatitic, anti-diabetes, and anti-tumour.[17] 1,6-*O,O*-diacetylbritannilactone (**2**), also isolated from *I. britannica*, was found to be ten times more potent than **1** towards HL-60 and MCF-7 cells. It was hypothesised that the enhanced lipophilicity of **2**, due to the acetyl group at position C-6, was responsible for the increased potency. It was also suggested that the electrophilic α-methylene-γ-lactone motif, which can bind thiols of proteins and induce DNA-fragmentation and apoptosis, was of importance. Dong *et al.* therefore generated 19 ester derivatives at the C-6 position as well as five C-13-modified arylation analogues (Fig. 1). Compound **1** and its derivatives were all tested for their cytotoxicity towards three cancer cell lines (HCT116, HEp-2 and HeLa). Introduction of lipophilic aliphatic chains at C-6 generally led to an increase in activity, while arylation at C-13 resulted in decreased potency. The most potent analogue, **3**, bore a lauroyl group (IC_{50} 2.91–6.78 μM), which was more active than the NP scaffold **1** (IC_{50} 19.3–36.1 μM), and comparable to the known anti-cancer agent, etoposide (IC_{50} 2.13–4.79 μM).[17]

Bu-2313 A (**4**) and B (**5**) are produced by an oligosporic actinomycete strain and have been shown to be active against a number of anaerobic and aerobic bacteria.[19, 20] During the isolation of **4** and **5** it was found that a periodate oxidation reaction yielded the dienoic acid moiety.[19, 20] This chemistry was exploited during a second study in which the tricyclic dienoic acid was reacted with a series of tetramic acid and cyclic 1,3-dicarbonyl moieties to generate 13 analogues of **4** that were screened for their anti-bacterial activity using the strains *Bacteroides fragilis* A20926, *Propionibacterium acnes* A21933, and *Streptococcus pyogenes* A9604 (Fig. 1).[19] The derivatives displayed anti-bacterial activity with the MIC values ranging from 0.1–25 μg/mL, 0.4–6.3 μg/mL, and 1.6–100 μg/mL against *B. fragilis, P. acnes,* and *S. pyogenes*, respectively (MIC 0.05–6.3 μg/mL for **4** and **5**).[19]

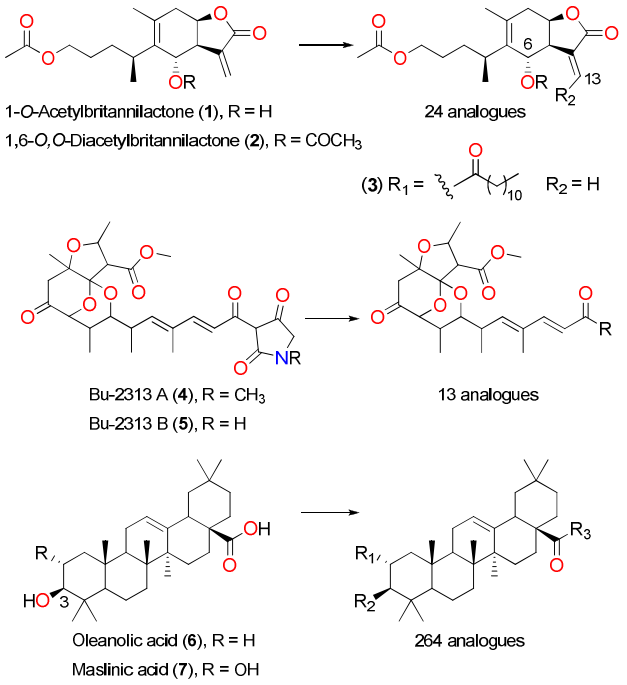

Fig. 1 Overview of the 1-*O*-acetylbritannilactone (**1**), Bu-2313 A (**4**) and B (**5**), oleanolic acid (**6**) and maslinic acid (**7**) libraries.

Oleanolic acid (**6**) and maslinic acid (**7**) are found in high concentrations in olive-pomace oil and can be obtained in large amounts from the solid waste resulting from olive-oil production.[21] Triterpenoids such as **6** and **7** possess a range of medicinal benefits including anti-inflammatory, hepatoprotective, analgesic, anti-microbial, virostatic, anti-cancer, and anti-HIV activities.[21] Parra *et al.* generated 264 derivatives of these two scaffolds (Fig. 1) using combinatorial solid-phase procedures with the objective to obtain analogues with greater cytotoxic and anti-tumour activities.[21] For example, combinatorial attachment of six different amino acids to the carboxylic side chains of **6** and **7** followed by acylation of the hydroxyl group at C-3 using 10 different acid anhydrides led to an extended set of bi-functional analogues. The cytotoxic effects of 90 of the derivatives towards three cancer cell lines (HT29, Hep G2, and B16−F10) were examined. Close to 90% of the compounds displayed cytotoxicity towards the B16−F10 and HT29 cell lines, while 67% showed growth inhibition towards the Hep G2 cell line. Several derivatives of **6**, the parent acid, had low IC_{50} values (< 2 μM), around 300-fold more effective than **6**. SAR studies found that long-chain ω-amino acids on the carboxylic group and small acyl groups on the hydroxyl moieties of **6** and **7** increased the potency of the analogues.[21]

Fredericamycin A (**8**), originally isolated from a culture of *Streptomyces griseus* (FCRC-48),[22] possesses *in vitro* anti-bacterial, anti-fungal, and cytotoxic activities and has also been shown to inhibit topoisomerases I and II and DNA polymerase.[23] This compound also demonstrated *in vivo* activity against P388 leukemic cells, and reduced the median weight of CD8F mammary tumours in mice. Rather than go down the route of total synthesis, as many others had done, Abel *et al.* utilised an optimised fermentation, isolation and purification process to obtain multi-gram quantities of **8** from the culture broth of *Streptomyces griseus* ATCC49344 mutants.[23] NP **8** was subsequently altered via the halogenation of ring B and modification of the methoxy group attached to ring F (Fig. 2). The pentadiene sidechain was also degraded to an aldehyde in order to lower the MW of **8** and provide a chemical handle for further synthetic transformations. A number of the halogenated derivatives, in particular the iodinated derivative (**9**), were found to be more potent (IC_{70} 8–20 nM) compared to **8** (IC_{70} 517 nM) and have higher selectivity when tested against a panel of ten tumour cell lines. Replacement of the pentadiene side chain resulted in less potent molecules.[23]

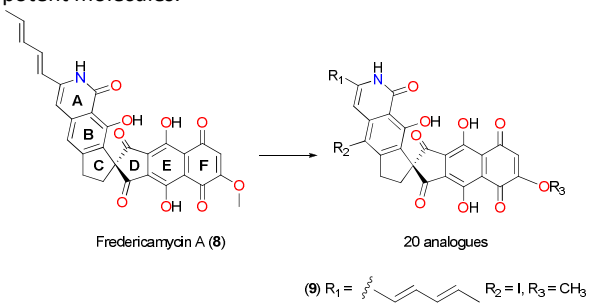

Fig. 2 Overview of the fredericamycin A (**8**) library.

The marine sesquiterpenoid hydroquinone avarol (**10**), isolated from the sponge *Dysidea avara*,[24] has been found to possess anti-oxidant, anti-platelet, anti-inflammatory, anti-viral, anti-tumour, anti-HIV, and anti-psoriatic activities. A number of studies involving the generation of avarol derivatives have been reported, a few of which are summarised here. Prompted by avarol's anti-oxidant activity and potential use as an anti-psoriatic agent, Amigo *et al.* produced 14 avarol derivatives and examined their potential as antioxidants and inhibitors of cell proliferation and PGE_2 generation in human keratinocytes (Fig. 3).[25]

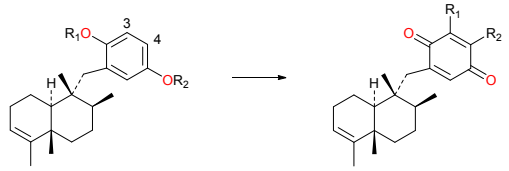

Fig. 3 Overview of the avarol (**10**) library.

Three ester analogues were generated from avarol (**10**), while Ag_2O oxidation of **10** gave avarone, which was subsequently used to generate one thio and ten amino derivatives. The thiosalicylic derivative of **10** demonstrated interesting anti-inflammatory properties, being a potent inhibitor of ROS and PGE2 production. 3'-Methylaminoavarone, which possessed a strong anti-proliferative profile, was identified as a possible anti-psoriatic drug.[25] Avarol (**10**) and its derivatives were further examined for their influences on primary keratinocyte cell growth.[26] 4'-Benzylamino avarone was found to inhibit keratinocyte cell growth and TNFα and COX-2 expression, making it an interesting anti-psoriatic drug candidate.[26] Amigo *et al.* generated a further 14 analogues of **10** to explore its photoprotective potential.[27] The derivatives were evaluated against a spontaneously immortalised human keratinocyte cell line (HaCaT) for cellular viability, TNF-α production, and NF-κβ activation. Two monophenyl thio-avarol derivatives which did not exhibit any cytotoxicity were considered to be promising UVB photoprotective agents due to their potent inhibition of NF-κβ activation and mild antioxidant pharmacological profile.[27] Pejin *et al.* generated 11 additional thio-avarone analogues and tested these along with 12 of the derivatives constructed by Amigo *et al*. in anti-microbial, brine shrimp lethality, free-radical scavenging, and acetylcholinesterase inhibition assays. The activity of the analogues varied in each test, allowing for SAR delineation. For example, in the free-radical scavenging assay it was found that 3' substituted derivatives were more active than those with substitution at 4'.[28]

Sarcophine (**11**) was first isolated from the Red Sea soft coral *Sarcophyton glaucum* in 1974; this major compound constituted ~ 3% of the wet weight of the marine organism.[29, 30] Biological investigations during the 1990's identified **11** and 7α-hydroxy-$\Delta^{8(19)}$-deepoxysarcophine as effective suppressors

of TPA-induced anchorage-independent JB6 cell transformation, and noted the significance of functional groups at C-7 and C-8 for anti-cancer activity.[31] Building on this and other synthetic studies that reported promising biological data and important SAR observations, in 2011 Hassan *et al.* described the production of 12 sarcophine analogues (Fig. 4).[32] The compounds were assessed for their ability to inhibit the proliferation and migration of the human metastatic prostate cancer PC-3 and breast cancer MDA-MB-231 cell lines. Only three of the compounds showed anti-proliferative activity against MDA-MB-231 cells (IC_{50} 20–45 μM), indicating that most lacked cytotoxicity. Anti-migratory activity was assessed in a wound-healing assay and a number of the cembranoid derivatives had higher anti-migratory activity against both cell lines (PC-3 IC_{50} 15.5–24.3 μM; MDA-MB-231 IC_{50} 4.8–30.0 μM) compared to the parent compound and control (4-hydroxyphenylmethylene hydantoin). SAR observations indicated that modification at C-7 enhanced potency, with butyloxy, butyl, or hexylcarbamate functionalities affording the most significant biological effects.[32]

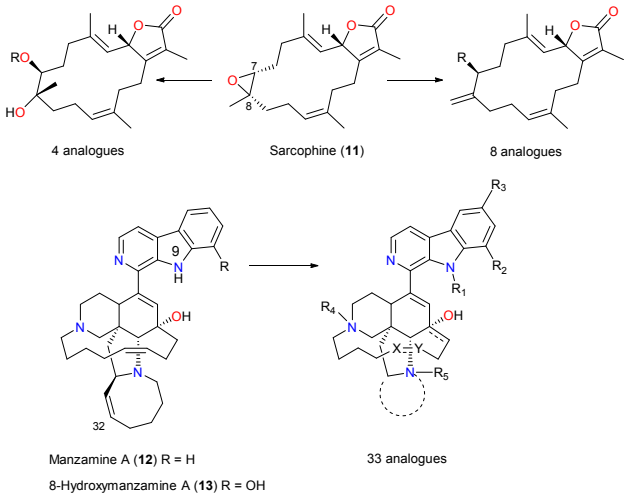

Fig. 4 Overview of the sarcophine (**11**) and manzamine libraries.

The first member of the complex manzamine alkaloid structure class was identified in 1986.[33] Numerous manzamine secondary metabolites have since been reported with a variety of activities, including anti-microbial, anti-neuroinflammatory, anti-malarial, and cytotoxicity.[34, 35] These compounds were also identified as glycogen synthase kinase-3*β* (GSK-3*β*) inhibitors, indicating they could be used to treat neurological disorders like Alzheimer's disease.[34, 35] The Indonesian marine sponge *Acanthostrongylophora* sp. has been a plentiful source of various manzamines, and in 2007 Hamann *et al.* isolated large quantities of several analogues, including manzamine A (**12**) and 8-hydroxymanzamine A (**13**), which allowed for derivatisation studies (Fig. 4).[34] Based on these two metabolites (**12** and **13**), 13 analogues were generated and screened for GSK-3*β* inhibition. Several derivatives were as potent (IC_{50} 4.8–10.4 μM) as the lead compounds (IC_{50} 10.2 [**12**] and 4.8 [**13**] μM). The SAR data suggested that the manzamine structure class contains a promising scaffold from which more potent and selective GSK-3 inhibitors could be designed.[34] In a more recent paper Peng *et al.* produced additional screening libraries using manzamine A (**12**), 8-hydroxymanzamine A (**13**), manzamine F (not shown), and the related ircinal (not shown) with 16, 4, 11 and 8 analogues being generated from each respective NP scaffold.[35] The analogues were evaluated for their anti-malarial, anti-microbial, and anti-neuroinflammatory activities. None possessed higher anti-malarial activity than **12** and **13** (IC_{50} 4.5–8.0 ng/mL), and it was found that leaving 9-NH unsubstituted and the double bond at C-32 intact was important for the potency. Modification of the carbonyl group of manzamine F with hydrazine or alkylation did result in a significant improvement to the anti-malarial activity (IC_{50} 780 ng/mL for manzamine F, compared to IC_{50} 29–86 ng/mL for other derivatives). In the anti-microbial screens, a number of the analogues demonstrated improved potency against *Cryptococcus neoformans* (MIC 0.8–6.5 μg/mL) and *Mycobacterium intracellulare* (MIC 0.02–0.45 μg/mL) compared to the parent NPs. Manzamine A (**12**) and some analogues also demonstrated anti-neuroinflammatory activity alongside minimal cytotoxicity when tested. On the basis of the results of these studies, both groups concluded that it will be possible to undertake rational drug design around the manzamine NP scaffold for the treatment of several different diseases.[34, 35]

The enzyme MraY has been identified as an anti-bacterial target as it is required for the biosynthesis of the polymer peptidoglycan, an essential element in bacterial cell walls.[36] The NPs muraymycins A1 and A3, isolated from a *Streptomyces* sp., were found to possess significant activity against MraY and also displayed good *in vitro* and *in vivo* activity (ED_{50} 1–2 mg/kg in mice) against Gram-positive bacteria.[36, 37] However, they only possessed a moderate therapeutic index of around 4 (measured by the ratio of LD_{50} to ED_{50}). Meanwhile, muraymycin C1 (**14**) did not demonstrate the anti-microbial activity of muraymycins A1 and A3, but was found to have good enzyme inhibitory activity and no toxicity.[36, 37] It was suggested that the lipophilic 12-guanidino or 12-hydroxyguanidino lauroyl group attached to the hydroxyleucyl moiety of muraymycins A1 and A3 may allow for their transportation to MraY in the membrane, but may also be responsible for the toxicity of these compounds.[36] Therefore, Lin *et al.*[36] undertook the modification of muraymycin C1 (**14**) through the introduction of lipophilic groups in order to improve the molecules potency and therapeutic index. Sixteen muraymycin derivatives were synthesised using selective reactions of the primary and/or secondary amino groups of **14** with isocyanates and aldehydes (Fig. 5). Di-substitution resulted in a loss of activity against MraY or MurG, while analogues derived from addition at the secondary amino group demonstrated good inhibitory activity (6.25–100 μg/mL), with two possessing comparable inhibition (6.25 μg/mL) with muraymycin C1 (**14**) when tested in a MurG biochemical assay.

This activity correlated with the lipophilicity of the substituents introduced. The toxicity of the derivatives was not evaluated.[36]

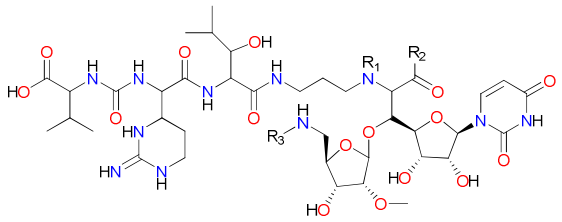

Fig. 5 Overview of the muraymycin C1 (**14**)

The lipopeptide FR901379 (**15**) was isolated from the culture broth of *Coleophoma empetri* F-11899.[38, 39] This NP is highly water-soluble and demonstrates strong anti-fungal activity against *Candida* species (IC_{50} 0.008–0.025 μg/mL).[38] Compound **15** was found to inhibit 1,3-β-glucan synthase (IC_{50} 0.7 μg/mL), which only exists in fungal cell walls.[38] While having promising potency and solubility properties, **15** was found to haemolyse mouse red blood cells *in vitro* at 62 μg/mL and was not effective against another clinically important fungus, *Aspergillus fumigatus*.[38] Thus there was a need to chemically modify **15** to improve the anti-fungal potency and reduce the haemolytic activity of the compound, which was associated with the lipophilic palmitoyl side chain. Enzymatic deacylation and re-acylation products displayed comparable anti-fungal activity to FR901379, but lacked the haemolytic properties associated with the NP.[40] Subsequently a library of acylated analogues of **15** was generated (Fig. 6)[40]

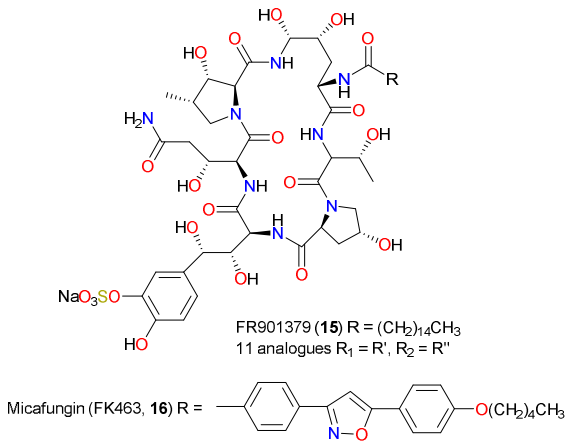

Fig. 6 Overview of the FR901379 (**15**) library.

Initially, derivatives varying at the acyl side chain with improved lipophilicity were designed and constructed. The anti-fungal activity against *C. albicans* was found to be at its maximum when the clog P of the side-chain was ~ 6. Further optimisation of the side-chain showed that rigid moieties gave better biological activities. This resulted in the identification of a naphthyl analogue which displayed approximately 5.5- to 8-fold superior activity (ED_{50} 0.46–0.79 mg/kg) compared to **15** (ED_{50} 3.71–4.31 mg/kg) in *in vivo* models of *C. albicans* and *A. fumigatus* infections.[40] Further optimisation led to the discovery of **16**, whose *in vivo* activity was found to be excellent when administered intravenously.[41] Compound **16**, known as micafungin or FK463,[41] is now marketed as the intravenous anti-fungal drug mycamine®.

The flavanone glycoside naringin (**17**) is a widely abundant plant NP that was first isolated from *Eucalyptus globulus*. Flavonoids such as **17** are abundant in the human diet, and have been shown to exhibit a variety of biological properties such as radical scavenging, anti-cancer, anti-bacterial, anti-inflammatory and anti-oxidant activities.[42] In 2005, Hanessian *et al.* reported the preparation of 28 analogues of naringin using solution-phase parallel synthesis (Fig. 7). The 6-hydroxyl group of the *β*-D-glucopyranosyl subunit was transformed into urethanes, amides, sulphonamides, and secondary and tertiary amines via a 6-amino naringin intermediate.[42] The synthesised analogues constituted a large diversity of aromatic and heterocylic *N*-derivatives based on the naringin scaffold. While no bioactivity was reported in this particular paper, the authors did make the comment that the library *"may find application in biological screens searching for new activities associated with flavonoids."*[42]

The poly-prenylated acyl phloroglucinol guttiferone A (**18**) was first isolated from the roots of the plant *Symphonia globulifera,*[43] but has also been purified from other species belonging to the Clusiaceae family, such as *Garcinia livingstonei* and *G. macrophylla*.[44] Guttiferone A (**18**) has been evaluated in a number of biological screens and shown to display anti-bacterial, anti-HIV, trypanocidal, anti-malarial, anti-cancer, and leishmanicidal activities.[44] In 2013 Fromentin *et al.* reported the synthesis of 20 ester and ether derivatives of guttiferone A that probed the catechol pharmacomodulation of this plant metabolite (Fig. 7).[44] All the derivatives were screened *in vitro* for their anti-parasitic properties. Six of the compounds demonstrated similar or slightly improved potency (IC_{50} 1.42–3.23 μM) compared to **18** (IC_{50} 2.95 μM) toward *Trypanosoma brucei*. Just one analogue possessed activity (IC_{50} 2.00 μM) less than **18** (IC_{50} 3.32 μM), with five other molecules demonstrating similar potencies (IC_{50} 3.63–5.06 μM) against *Plasmodium falciparum*. In regards to anti-leishmanial activity, none of the compounds (IC_{50} 5.04–100 μM) showed higher activity than **18** (IC_{50} 5.83 μM).[44] Cytotoxicity evaluation of the derivatives (VERO cells) identified that some catechol modulations decreased toxicity while maintaining the anti-parasitic activity of **18**.[44]

The triterpenoid NP sipholenol A (**19**) was first purified from the organic extract of the Red Sea sponge *Callyspongia siphonella*.[45] Bioactivity testing of **19** and some related analogues subsequently demonstrated their ability to reverse P-glycoprotein (P-gp) induced multi-drug resistance (MDR) in human epidermoid cancer cells, and members of this structure class have been reported as inhibitors of breast cancer cell migration and invasion.[46, 47] In an initial study in 2013 Foudah *et al.* described the generation of a series of ester, ether, oxime, and carbamate analogues of **19** (Fig. 7).[46] The anti-

migratory activities of sipholenol A and its analogues were tested in a wound-healing assay against the highly metastatic human breast cancer cell line MDA-MB-231. All but two of the analogues possessed improved activities (IC_{50} 5.3−52.7 μM) compared to **19**, with two 4-chlorobenzoate ester derivatives demonstrating the highest potencies (IC_{50} 5.3 and 5.9 μM). Protein tyrosine kinase 6 (PTK6), which promotes growth factor signalling and migration in breast tumour cells, was identified as a potential target using a KINOMEscan analysis. Sipholenol A 4β-4-chlorobenzoate and 19,20-anhydrosipholenol A 4β-4-chlorobenzoate esters were found to possess the most potent inhibition of PTK6 phosphorylation when tested.[46] Foudah *et al*. went on to report a second study involving the production of 15 new sipholenol A esters in an attempt to further improve the breast cancer anti-migratory and anti-proliferative activities.[47] The 4β-4',5'-dichlorobenzoate ester of sipholenol A, **20**, was shown to be the most potent compound, with an IC_{50} value of 1.3 μM in the migration assay. Furthermore, pharmacophore modelling and 3D-QSAR analyses identified and correlated important pharmacophoric features, providing data for the future design of novel anti-migratory compounds based on a simplified sipholane structure.[47]

The styryl-lactone goniothalamin (**21**) is abundant in several *Goniothalamus* species,[48, 49] and has demonstrated good activity against a number of tumour cell lines. In an attempt to improve the potency and lipophilicity of **21** Zhou *et al.* produced 19 derivatives by *ortho*-substitution of the aromatic ring with amine and amino acid moieties (Fig. 7).[49] Goniothalamin (**21**) together with its derivatives were then evaluated for their *in vitro* anti-tumour activity using a SRB (sulforhodamine B) assay against human promyelcytic leukemia (HL-60), human hepatoma (BEL-7402), human lung carcinoma (A549), and human stomach cancer (SGC-7901) cell lines. Most of the compounds showed an inhibitory effect against HL-60 cells (IC_{50} 2.8−10.2 μM, control etoposide IC_{50} 0.22 μM), and two simple nitro and amine derivatives also demonstrated good activities against SGC-7901 cells (IC_{50} 4.5 and 5.3 μM, respectively, control etoposide IC_{50} 10.3 μM). None of the compounds were found to have better activity profiles than the parent molecule **21**.[49] The log P values of the analogues were also calculated and it was shown that the amino acid derivatives possessing higher log P values showed poor anti-tumour activities against HL-60 cells while compounds with log P < 5 exhibited better activity.[49]

A number of macrocyclic diterpene NPs isolated from plants of the genus *Euphorbia*, including jolkinol D (**22**),[50] have been found to be strong modulators of P-gp, which is implicated in the MDR of tumours.[51] In order to identify more potent P-gp modulators, Reis *et al.* carried out a phytochemical study of *Euphorbia piscatoria* in which a large amount of jolkinol D (**22**) was isolated.[50, 51] Based on this NP scaffold, 13 analogues were generated by esterification of the hydroxyl group at C-3 and by hydrolysis of the ester at C-15 (Fig. 7). All derivatives possessed higher anti-proliferative activity than jolkinol D (**22**), with their selectivity indices indicating that the compounds were equally active against L5178Y mouse lymphoma cells (PAR cells, chemo-sensitive) and in human MDR1-gene transfected L5178Y mouse lymphoma cells (MDR cells, chemoresistant by overexpression of P-gp). The potential P-gp-mediated MDR reversing activity of the derivatives were assessed using a rhodamine-123 exclusion assay in which the fluorescence activity ratio (FAR) was measured. Five of the library members possessed FAR values greater than 10 (10.55 − 28.09) at 2 μM, indicating they were strong MDR modulators. The control compound verapamil had a FAR value of 20 at 22 μM. The combined effect of the analogues when administered with the anti-neoplastic drug doxorubicin was analysed, with all but one of the compounds being found to synergistically enhance its cytotoxicity. A number of physicochemical parameters were determined and the compounds clustered according to similarities. This analysis revealed significant correlations between potency and parameters such as log P, solvent accessible area, and MW, and also indicated the importance of aromatic moieties for P-gp efflux modulation. Altogether these results indicated that the derivatives had potential as MDR reversal agents.[51]

Our group has also explored the use of isolated NPs as scaffolds, with a major focus on the generation of orally bioavailable leads or drugs. The Eskitis Institute's Nature Bank collection, which includes ~63,000 plant and marine biota samples and 2738 pure NPs, is a unique resource for the procurement of NP scaffolds. The pure compound library was analysed for molecules fitting a simple set of structural parameters that were identified as positive attributes for potential scaffolds (MW < 300, quantity available > 50 mg, log P < 5, chemical handles > 1 and stereogenic centres > 1).[10] This screening process identified the sesquiterpene 14-hydroxy-6,12-muuroloadien-15-oic acid (**23**), whose low MW (250), favourable log P (2.42), multiple stereogenic centres (4) and potential chemical handles (i.e. the carboxylic acid and allylic hydroxyl group) for synthetic elaboration made it a particularly attractive NP scaffold (Fig. 7).[10] Compound **23** has been isolated from several species belonging to the Australian endemic plant genus *Eremophila*.[10, 52] Prior to analogue generation of **23**, a virtual library of potential products was constructed using lists of commercially available amines and isocyanates. The virtual compounds were compared using Lipinski's drug-like "Rule of Five" (MW < 500; HBA < 10; HBD < 5; log P < 5) criteria in order to select those with the most drug-like properties for construction. NP **23** was then used to generate 12 amide and carbamate analogues, which were evaluated for their anti-malarial activity. Several of the compounds displayed low to moderate potency with IC_{50} values ranging from 14 to 33 μM, as opposed to the scaffold itself (**23**) which displayed no activity.[10] This library has subsequently been added to an open-access compound repository at the Eskitis Institute (Compounds Australia), where it has been screened by both local and international collaborators in a variety of assays. Several library members are currently being further evaluated for their anti-infective potential.

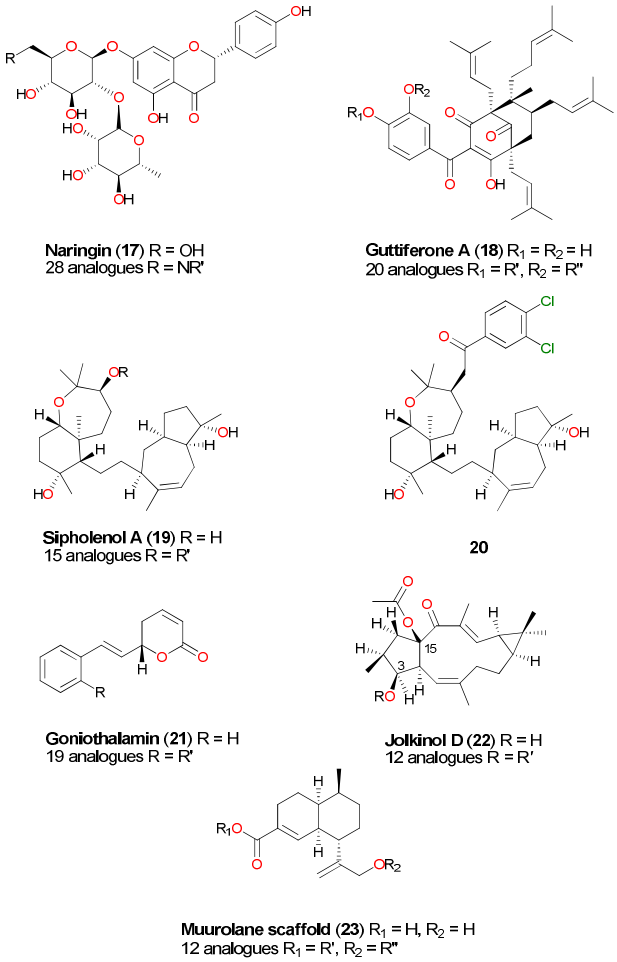

Fig. 7 Overview of libraries derived from NP scaffolds **17**-**19** and **21**-**23**.

3. Discussion

The use of isolated NPs as scaffolds in library production is by no means a well-established method; however, there has been a steady increase in the number of publications using this strategy over the years. The earliest study given in this review is from 1980, but the remaining examples were published during 2002–2014. While there are some examples of large-scale combinatorial libraries having been generated from isolated NP scaffolds [*e.g.* oleanolic (**6**) and maslinic acids (**7**)], the majority of the library examples described here are smaller in size (12–28 membered libraries). These smaller libraries are certainly worthwhile since they give a preliminary SAR overview, that can be used to assist in the prioritisation of a project, whilst minimising time and effort, along with consumption of resources. Although there are some exceptions [*e.g.* Bu-2313 A (**4**) and muraymycin C1 (**14**)], the NP-derived library examples given here tended to be successful in terms of achieving an improvement in either potency or other factors such as toxicity or lipophilicity compared to their NP starting scaffold. The majority of papers also discuss SAR trends, with some of the data supporting additional library generation on particularly encouraging scaffolds. Despite a number of the NP scaffolds mentioned here possessing a range of previously identified bioactivities [*e.g.* fredericamycin A (**8**), 1,6-*O,O*-diacetylbritannilactone (**1**), avarol (**10**)], typically their analogues are only reported to have been tested in one bioactivity screen. Bioprofiling, the testing of compounds in a range of diverse assays, will no doubt add further value to these unique libraries, and we hope that open access compound repositories like the one currently operating at Griffith University (Compounds Australia), will assist in making such compounds more accessible in the future. Easier access to such libraries will lead to more screening (i.e. bioprofiling), which will translate into future successes for small-molecule lead/drug discovery.

The quality of a compound library is determined by its molecular diversity and biological properties. In order to maximise these qualities it can be useful to utilise physicochemical parameters in order to select suitable scaffolds and/or analogues for construction. For example, Lipinski's "Rule of Five" for drug-like molecules can be employed to bias library members towards orally bioavailable leads and/or drugs, while other property combinations can be pursued to facilitate the development of injectables.[7] Indeed, a statistical analysis of drugs derived from a natural progenitor showed that physicochemical properties can vary markedly across different therapy areas.[7] Computational design using physicochemical properties prior to derivative synthesis has been increasingly used to produce more desirable libraries.[14] Virtual libraries can be created in order to examine the properties of candidate analogues, bypassing the need to generate large sets of potentially redundant molecules. It is likely that using even a simple set of indicators such as MW, log P, and the number of stereocentres to select scaffolds and/or potential products can increase the chances of obtaining biologically active hits.[14] If information is available about the target of a bioactive NP, then the synthesis of specific derivatives to induce a desired biological effect, such as the use of sulfonamidation to encourage interactions with certain enzymes, is also a useful strategy.

While early work in the generation of semi-synthetic libraries has been valuable, there are still improvements that can be made. For example, a number of the NP scaffolds contain moieties such as Michael acceptors (*e.g.* fredericamycin A and goniothalamin) and catechol functionalities (*e.g.* naringin) that would be classified as PAINS.[15] It would be advantageous to move this strategy forward by using more stringent medicinal chemistry tactics, including the careful selection of scaffolds so as to avoid these motifs, and to further incorporate rational design strategies using physicochemical properties. The combination of NP chemistry and physicochemical parameter "awareness" during drug/lead/probe discovery library design will prove to be a beneficial partnership. This approach, if adhered to, will yield libraries that have better bioavailability, and will translate into better hit or lead molecules.

Historically, one of the biggest problems associated with the use of isolated natural products for the semi-synthesis of screening libraries was supply. Not all desirable NP scaffolds could be readily purified in quantities that enable traditional medicinal chemistry. Furthermore, structural complexity, stability, the requirement for protection chemistry, and the lack of chemical handles are other issues that can prevent and restrict library synthesis around a NP scaffold. However, advances in technology have meant that the miniaturisation of processes associated with drug discovery programs is now possible. For example, access to high-field NMR spectrometers fitted with cold probes, the increasing usage of 1536-well microtitre plates for HTS, and uPLC systems has meant that library synthesis on scarce NP scaffolds is now possible. Extensive research, including full spectroscopic/spectrometric characterisation and bioprofiling can be performed on libraries containing sub-milligram quantities. Furthermore, synthetic methodologies have also been constantly improving.[5, 6] Careful strategies utilising minimal quantities of the chosen NP scaffold and those leading to site selective modifications are now more common place and facilitate site specific editing of a NP skeleton.[5, 6] Degradation and fragmentation studies involving the simplification of a NP to a core feature or the creation of new, more amendable chemical handles are also valuable methods in the utilisation of NP scaffolds.[53, 54] Site selective and degradation methods have been used in the modification of isolated NPs to make sets of analogues.[55-58]

While we have predominantly focussed on the use of NP analogue libraries for drug discovery, the impact of NPs and their derivatives on chemical biology (through their use as chemical probes) for biological target identification and mode of action studies should also not be underestimated or forgotten.[8]

Conclusions

NPs have been instrumental to drug research, and while they have the potential to continue making significant contributions to drug discovery and development, new and complementary strategies to traditional NP methods need to be developed. The use of isolated NPs as scaffolds for screening library generation may prove to be yet another way that these compounds can be efficiently used in the discovery of new leads and/or drugs. A multi-pronged approach to drug discovery that incorporates traditional approaches in new ways such as that discussed here will surely be beneficial.

Acknowledgements

RAD wishes to acknowledge the National Health and Medical Research Council (APP1024314) and the Australian Research Council (LP120200339) for funding. RK would like to thank Griffith University for a Griffith University International Postgraduate Research Scholarship (GUIPRS). ECB thanks the Australian Government for an Australian Postgraduate Award. We thank David Camp (Griffith University) for helpful discussions during the preparation of this manuscript.

Notes and references

1 K. C. Morrison and P. J. Hergenrother, *Nat. Prod. Rep*., 2014, **31**, 6-14.
2 D. J. Newman and G. M. Cragg, *J. Nat. Prod*., 2012, **75**, 311-335.
3 K. Grabowski, K.-H. Baringhaus and G. Schneider, *Nat. Prod. Rep.*, 2008, **25**, 892-904.
4 F. M. Tajabadi, M. R. Campitelli and R. J. Quinn, *Springer Science Reviews*, 2013, **1**, 141-151.
5 M. E. Maier, *Org. Biomol. Chem*., 2015, **13**, 5302-5343.
6 O. Robles and D. Romo, *Nat. Prod. Rep*., 2014, **31**, 318-334.
7 D. Camp, A. Garavelas and M. Campitelli, *J. Nat. Prod*., 2015, **78**, 1370-1382.
8 S. Rizzo and H. Waldmann, *Chem. Rev*., 2014, **114**, 4621–4639.
9 R. Kumar, M. C. Sadowski, C. Levrier, C. C. Nelson, A. J. Jones, J. P. Holleran, V. M. Avery, P. C. Healy and R. A. Davis, *J. Nat. Prod.*, 2015, **78**, 914–918.
10 E. C. Barnes, V. Choomuenwai, K. T. Andrews, R. J. Quinn and R. A. Davis, *Org. Biomol. Chem*., 2012, **10**, 4015-4023.
11 V. Choomuenwai, K. T. Andrews and R. A. Davis, *Bioorg. Med. Chem*., 2012, **20**, 7167-7174.
12 D. Camp, R. A. Davis, E. A. Evans-Illidge and R. J. Quinn, *Future Med. Chem*., 2012, **4**, 1067-1084.
13 R. Breinbauer, I. R. Vetter and H. Waldmann, *Angew. Chem. Int. Ed*., 2002, **41**, 2878-2890.
14 M. Pascolutti and R. J. Quinn, *Drug Discov. Today*, 2014, **19**, 215-221.
15 J. Baell and M. A. Walters, *Nature*, 2014, **513**, 481-483.
16 Y. C. Martin, J. L. Kofron and L. M. Traphagen, *J. Med. Chem*., 2002, **45**, 4350-4358.
17 S. Dong, J.-J. Tang, C.-C. Zhang, J.-M. Tian, J.-T. Guo, Q. Zhang, H. Li and J.-M. Gao, *Eur. J. Med. Chem*., 2014, **80**, 71-82.
18 Z. Bing-Nan, B. Nai-Sheng, L. Long-Ze and G. A. Cordell, *Phytochemistry*, 1993, **34**, 249-252.
19 S. Toda, S. Nakagawa, T. Naito and H. Kawaguchi, *J. Antibiot*., 1980, **33**, 173-181.
20 M. Tsunakawa, S. Toda, T. Okita, M. Hanada, S. Nakagawa, H. Tsukiura, T. Naito and H. Kawaguchi, *J. Antibiot*., 1980, **33**, 166-172.
21 A. Parra, S. Martin-Fonseca, F. Rivas, F. J. Reyes-Zurita, M. Medina-O'Donnell, E. E. Rufino-Palomares, A. Martinez, A. Garcia-Granados, J. A. Lupianez and F. Albericio, *ACS Comb. Sci.*, 2014, **16**, 428–447.
22 R. C. Pandey, M. W. Toussaint, R. M. Stroshane, C. C. Kalita, A. A. Aszalos, A. L. Garretson, T. T. Wei, K. M. Byrne, R. F. Geoghegan, Jr. and R. J. White, *J. Antibiot*., 1981, **34**, 1389-1401.
23 U. Abel, W. Simon, P. Eckard and F. G. Hansske*, Bioorg. Med. Chem. Lett*., 2006, **16**, 3292-3297.
24 L. Minale, R. Riccio and G. Sodano, *Tetrahedron Lett*., 1974, **15**, 3401-3404.
25 M. Amigo, M. C. Terencio, M. Mitova, C. Iodice, M. Paya and S. De Rosa, *J. Nat. Prod*., 2004, **67**, 1459-1463.
26 M. Amigo, J. Schalkwijk, D. Olthuis, S. De Rosa, M. Paya, M. C. Terencio and E. Lamme, *Life Sci*., 2006, **79**, 2395-2404.

27 M. Amigo, M. C. Terencio, M. Paya, C. Iodice and S. De Rosa, *Bioorg. Med. Chem. Lett*., 2007, **17**, 2561-2565.
28 B. Pejin, C. Iodice, G. Tommonaro and S. De Rosa, *J. Nat. Prod*., 2008, **71**, 1850-1853.
29 J. Bernstein, U. Shmeuli, E. Zadock, Y. Kashman and I. Neeman, *Tetrahedron*, 1974, **30**, 2817-2824.
30 I. Ne'eman, L. Fishelson and Y. Kashman, *Toxicon*, 1974, **12**, 593-594.
31 K. A. El Sayed, M. T. Hamann, C. A. Waddling, C. Jensen, S. K. Lee, C. A. Dunstan and J. M. Pezzuto, *J. Org. Chem*., 1998, **63**, 7449-7455.
32 H. M. Hassan, A. A. Sallam, R. Mohammed, M. S. Hifnawy, D. T. A. Youssef and K. A. El Sayed, *Bioorg. Med. Chem*., 2011, **19**, 4928-4934.
33 R. Sakai, T. Higa, C. W. Jefford and G. Bernardinelli, *J. Am. Chem. Soc*., 1986, **108**, 6404-6405.
34 M. Hamann, D. Alonso, E. Martin-Aparicio, A. Fuertes, M. J. Perez-Puerto, A. Castro, S. Morales, M. L. Navarro, M. del Monte-Millan, M. Medina, H. Pennaka, A. Balaiah, J. Peng, J. Cook, S. Wahyuono and A. Martínez, *J. Nat. Prod*., 2007, **70**, 1397-1405.
35 J. Peng, S. Kudrimoti, S. Prasanna, S. Odde, R. J. Doerksen, H. K. Pennaka, Y.-M. Choo, K. V. Rao, B. L. Tekwani, V. Madgula, S. I. Khan, B. Wang, A. M. S. Mayer, M. R. Jacob, L. C. Tu, J. Gertsch and M. T. Hamann, *J. Med. Chem*., 2010, **53**, 61-76.
36 Y.-I. Lin, Z. Li, G. D. Francisco, L. A. McDonald, R. A. Davis, G. Singh, Y. Yang and T. S. Mansour, *Bioorg. Med. Chem. Lett*., 2002, **12**, 2341-2344.
37 L. A. McDonald, L. R. Barbieri, G. T. Carter, E. Lenoy, J. Lotvin, P. J. Petersen, M. M. Siegel, G. Singh and R. T. Williamson, *J. Am. Chem. Soc*., 2002, **124**, 10260-10261.
38 T. Iwamoto, A. Fujie, K. Nitta, S. Hashimoto, M. Okuhara and M. Kohsaka, *J. Antibiot*., 1994, **47**, 1092-1097.
39 T. Iwamoto, A. Fujie, K. Sakamoto, Y. Tsurumi, N. Shigematsu, M. Yamashita, S. Hashimoto, M. Okuhara and M. Kohsaka, *J. Antibiot*., 1994, **47**, 1084-1091.
40 M. Tomishima, H. Ohki, A. Yamada, K. Maki and F. Ikeda, *Bioorg. Med. Chem. Lett*., 2008, **18**, 1474-1477.
41 M. Tomishima, H. Ohki, A. Yamada, K. Maki and F. Ikeda, *Bioorg. Med. Chem. Lett*., 2008, **18**, 2886-2890.
42 S. Hanessian and K. K. Kothakonda, *J. Comb. Chem*., 2005, **7**, 837-842.
43 K. R. Gustafson, J. W. Blunt, M. H. G. Munro, R. W. Fuller, T. C. McKee, J. H. Cardellina, J. B. McMahon, G. M. Cragg and M. R. Boyd, *Tetrahedron*, 1992, **48**, 10093-10102.
44 Y. Fromentin, N. Gaboriaud-Kolar, B. N. Lenta, J. D. Wansi, D. Buisson, E. Mouray, P. Grellier, P. M. Loiseau, M.-C. Lallemand and S. Michel, *Eur. J. Med. Chem*., 2013, **65**, 284-294.
45 S. Carmely and Y. Kashman, *J. Org. Chem*., 1983, **48**, 3517-3525.
46 A. I. Foudah, S. Jain, B. A. Busnena and K. A. El Sayed, *ChemMedChem*, 2013, **8**, 497-510.
47 A. I. Foudah, A. A. Sallam, M. R. Akl and K. A. El Sayed, *Eur. J. Med. Chem*., 2014, **73**, 310-324.
48 J. R. Hlubucek and A. V. Robertson, *Aust. J. Chem*., 1967, **20**, 2199-2206.
49 F. S. Zhou, W. D. Tang, Q. Mu, G. X. Yang, Y. Wang, G. L. Liang and L. G. Lou, *Chem. Pharm. Bull*., 2005, **53**, 1387-1391.
50 D. Uemura, K. Nobuhara, Y. Nakayama, Y. Shizuri and Y. Hirata, *Tetrahedron Lett*., 1976, **17**, 4593-4596.
51 M. Reis, R. J. Ferreira, M. M. M. Santos, D. J. V. A. dos Santos, J. Molnar and M.-J. U. Ferreira, *J. Med. Chem*., 2013, **56**, 748-760.
52 E. L. Ghisalberti, P. R. Jefferies, B. W. Skelton, A. H. White and R. S. F. Williams, Tetrahedron, 1989, **45**, 6297-6308.
53 J. Niggemann, K. Michaelis, R. Frank, N. Zander and G. Höfle, *J. Chem. Soc., Perkin Trans. 1*, 2002, 2490–2503.
54 J.-Y. Ortholand and A. Ganesan, *Curr. Opin. Chem. Biol., 2004, **8**, 271–280.*
55 P. A. Wender, O. D. Jankowski, E. A. Tabet and H. Seto, *Org. Lett.*, 2003, **5**, 487-490.
56 C. A. Lewis, K. E. Longcore, S. J. Miller and P. A. Wender, *J. Nat. Prod*., 2009, **72**, 1864-1869.
57 R. A. Davis, A. R. Carroll and R. J. Quinn, *Aust. J. Chem*., 2001, **54**, 355–359.
58 S. D. Dong, K. Sundermann, K. M. J. Smith, J. Petryka, F. Liu and D. C. Myles, *Tetrahedron Lett*., 2004, **45**, 1945–1947.