Non-invasive 19F NMR analysis of a protein-templated N-acylhydrazone dynamic combinatorial library.

Journal Name

ARTICLE

Non-invasive ^{19}F NMR analysis of a protein-templated N-acylhydrazone dynamic combinatorial library

Alexander G. Ekström, Jue Theresa Wang, Juraj Bella and Dominic J. Campopiano

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

Dynamic combinatorial chemistry (DCC) is a powerful tool to identify ligands for biological targets. We used ^{19}F NMR as an *in situ*, non-invasive technique for measuring the composition of a dynamic combinatorial library (DCL) of *N*-acylhydrazones (NAHs). An NAH DCL, constructed from a fluoro-aromatic aldehyde and a small set of hydrazides, was targetted at *ec*FabH, an essential enzyme in bacterial fatty acid biosynthesis. Our NMR analysis identified a tert-butyl NAH as the best binder which was confirmed by enzymatic assay.

Introduction

Dynamic combinatorial chemistry (DCC) is a technique derived from fragment-based screening that exploits chemically-reversible reactions to generate a thermodynamically equilibrated library from a pool of building blocks[1]. The reversibility of the dynamic combinatorial library (DCL) allows perturbation of the equilibrium through the addition of a template. Since its inception, DCC has been applied to a variety of fields including self-assembly systems, dynamic polymers and host-guest chemistry, however an exciting application is in the field of drug discovery, whereby a protein is used to self-select a ligand through amplification of the best binding species at the expense of other combinations (see fig. 1).

The constantly expanding toolbox of DCC-compatible reversible reactions[2,3] has helped facilitate the wide range of applications of the technique as a whole, however the number of bio-compatible reactions is limited by the requirement for the reaction to proceed on a reasonable timescale under physiological conditions (temperature, solvent, pH). From this panel of suitable reactions, the chemistry of *N*-acylhydrazone (NAH) exchange has been well documented and successfully used to identify ligands for a number of different protein targets[4–8]. The reaction between an aldehyde and a hydrazide occurs rapidly at pH 4 with an equilibrium constant and stability that strongly favours product formation[9]. In order to allow the reaction to proceed at a physiologically relevant pH, Bhat *et al.* applied early work from Jencks on oxime exchange catalysis[10]. By using aniline as a nucleophilic catalyst they showed that their 10-member *N*-acylhydrazone library equilibrated within a few hours at pH 6.2[5,10,11].

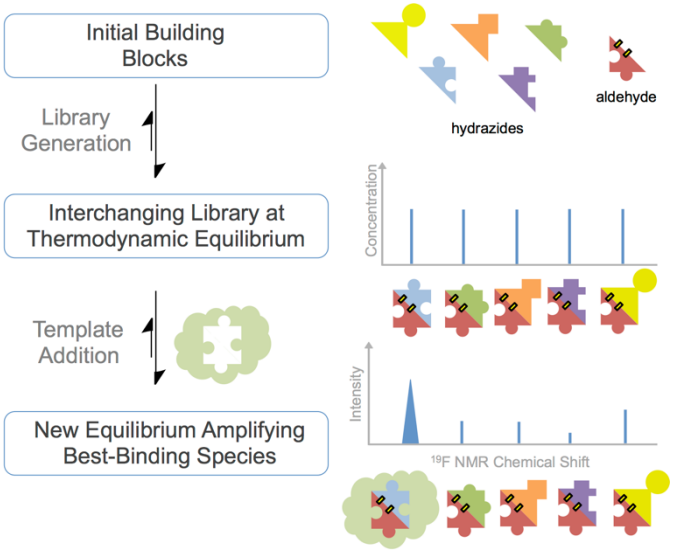

Figure 1. Formation of a DCL and analysis of the blank and protein-templated DCL product distribution by ^{19}F NMR.

A DCC experiment typically involves 3 main steps: firstly, the DCL is equilibrated under conditions promoting a thermodynamic equilibrium. Secondly, the protein template is added and the DCL is allowed to re-equilibrate establish a new product distribution. Thirdly, the final and arguably most vital step, is the analysis of the DCL in the presence and absence of the template. Various techniques to deconvolute the library mixture have been demonstrated[12], including HPLC[5,13–15], mass spectrometry[8,16–18] and ^{1}H NMR and Saturation Transfer Difference (STD) NMR[7,19,20], dynamic deconvolution[21] and X-ray crystallography[22]. Recent reviews of this area have highlighted both experimental and theoretical approaches used to analyse DCLs[12,23].

The main difficulty lies in not perturbing the DCL during the analysis. This has previously been achieved by either chemically or kinetically freezing the exchange reaction. For example, NAH

[a] *EaStCHEM School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, EH9 3FJ, UK*

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

exchange is effectively stopped by increasing the pH to 9. The second hurdle is ensuring that the complete product population is analysed. This problem is exemplified in HPLC analysis of protein-templated DCL, where the protein must be removed from the mixture prior to analysis to prevent column fouling[5]. In the absence of denaturation agents which may perturb the finely balanced protein-templated DCL equilibrium, target binders may remain bound to the protein and are excluded from the DCL analysis. This technique of "ligand fishing" has in fact been demonstrated in the DCC context to find competitive inhibitors of lysozyme[24].

Analysis by non-denaturing electrospray ionisation mass spectrometry (ESI-MS) has been demonstrated in a number of elegant DCC experiments[8,20,25], however the technique relies on the protein ionising efficiently under non-denaturing conditions and is therefore not universal. STD-NMR has also been used to identify binders from a DCL, however the technique requires upwards of a 20-fold molar excess of ligand. At such stoichiometry, STD experiments do not leverage the dynamic self-selection qualities central to the DCC concept. Instead, they exploit NAH formation as a facile, one-pot synthetic route to a library of compounds for binding assay. With the exception of STD-NMR, the aforementioned techniques are all destructive in a sense that a portion of the DCL is consumed in the analysis.

Our preliminary HPLC analyses of a protein-templated NAH DCL delivered striking differences in the product distribution depending on the method of protein removal (see fig. S1). To resolve this discrepancy, we set out to employ a non-destructive technique that would allow the DCL composition to be monitored in pseudo real-time.

Fluorine has become a central part of drug discovery, with approximately 25% of all marketed drugs containing at least one fluorine atom[26]. As a bioisostere of hydrogen, its inclusion can help medicinal chemists to modulate the pharmacokinetic, pharmacodynamic and physical binding properties of a compound[26]. Subsequently ^{19}F NMR has developed into a popular technique for screening fluorine-based fragment libraries, with the ability to screen up to 20 fragments in one experiment. The high gyromagnetic ratio and chemical shift anisotropy of the ^{19}F nucleus allow for well-resolved signals that relay information on protein binding through signal broadening or chemical shift perturbation. The scarce biological abundance of ^{19}F allows for background-free spectra, unaffected by protonated buffers and solvents, to be recorded in minutes on a low-field NMR spectrometer. The caveat of ^{19}F screening is the requirement of individual well-resolved signals. This can be achieved though intelligent library design facilitated by increasingly accurate chemical shift prediction software (e.g. MNova, TopSpin). More complex NMR experiments have been used to resolve overlapping signals, both by 2D homonuclear correlation-ordered experiments (^{1}H-^{19}F COSY) and pseudo-2D diffusion-ordered (^{19}F DOSY) experiments, where the signals are separated on second axis by diffusion coefficient[26–28]. The elegance of ^{19}F NMR analysis has already been demonstrated in a number of abiotic DCC examples[29–31]. This method was also used to monitor binding of a 4-component imine library to a domain of a human β-catenin target[32]. Herein, we describe our design and analysis of a protein-templated NAH DCL to further demonstrate the additional advantages of ^{19}F NMR in DCC for drug discovery.

Results and Discussion

We chose to design a library targeting β-ketoacyl-ACP synthase III (FabH) from *E. coli*, the initial condensing enzyme in bacterial fatty acid biosynthesis (FAS II). Although the active site residues and primary sequence remain highly conserved across both Gram-positive and Gram-negative bacteria, small differences in the binding pocket architecture determine the substrate specificity across bacterial species. This makes FabH a plausible target for novel narrow-spectrum antibiotics[33]. There are currently few known FabH inhibitors, and none that report good *in vivo* efficacy[34–36]. Zhang *et al.* have recently published a number of reports on the discovery of NAH inhibitors targeting FabH where each compound was prepared individually by organic synthesis. Since each molecule was composed of A-B ring systems joined by an NAH linker it provided the basis around which our dynamic library was designed[37–41].

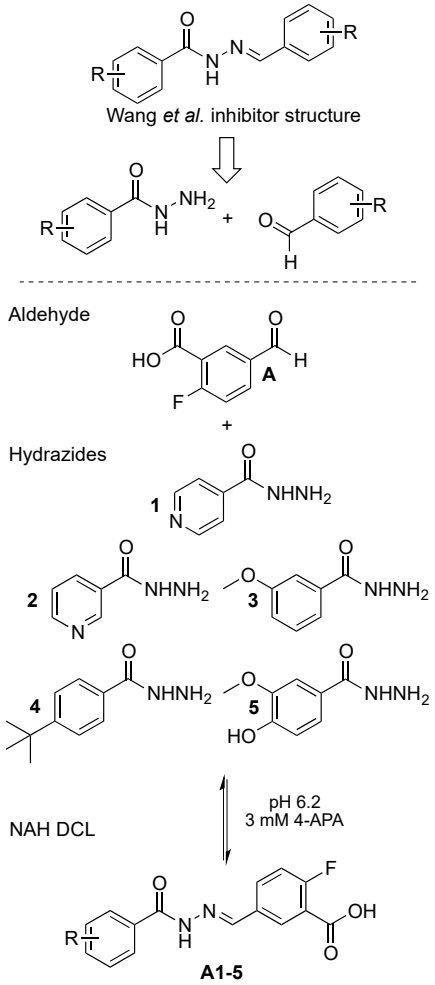

Figure 2. Rationale for the DCL and composition of the 5-membered proof-of-concept ^{19}F-labelled DCL. Hydrazide **1** (isoniazid) is a front-line TB therapy with another FAS II target.

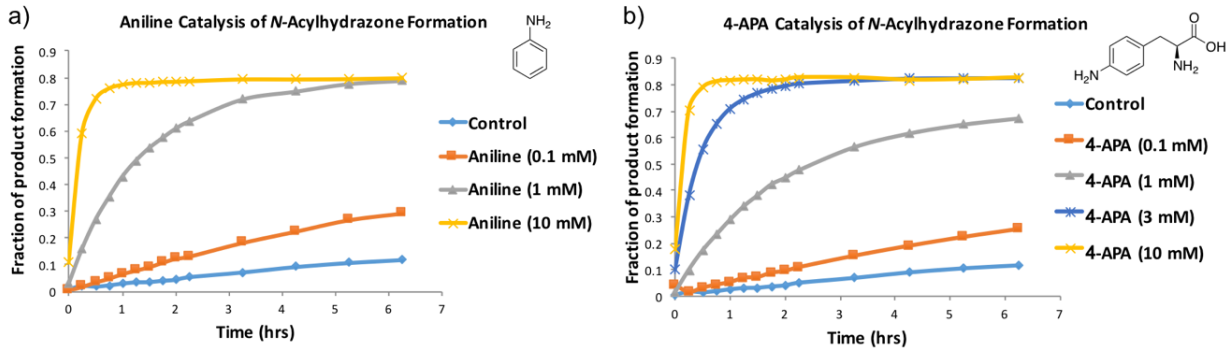

Figure 3. Validation of 4-APA as an alternative to aniline. Fig. 3a shows the formation of NAH **A4** from aldehyde **A** and hydrazide **4** over time in the presence of different concentrations of aniline. Figure 3b shows the same experiment, catalysed by different concentrations of 4-APA. The data shown is a compilation of the NAH **A4** HPLC peaks recorded at given time points (see fig. S3).

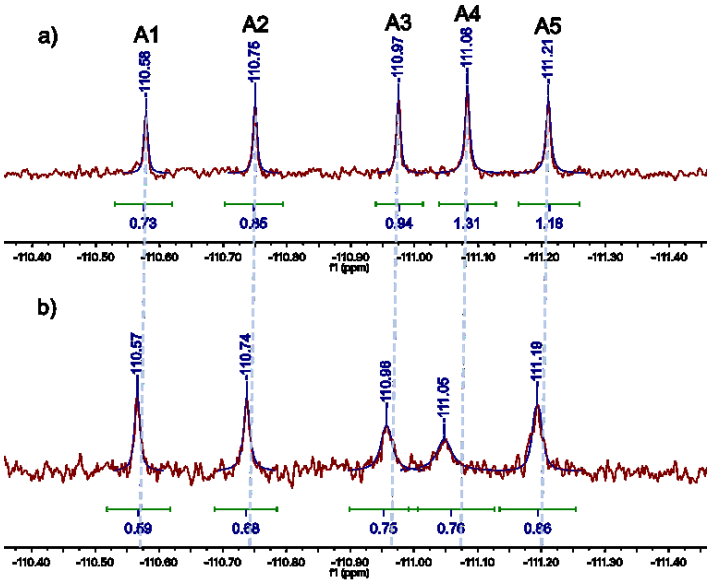

Figure 4. ^{19}F NMR spectra of the ^{19}F-labelled DCL: a) the DCL equilibrated in the absence of the protein target, b) the protein-templated DCL equilibrated in the presence of FabH.

The 5-membered proof-of-concept library was based around commercially available fluoro-aromatic aldehyde **A** as the central core along with five commercially available aromatic hydrazides **1-5** with differing chemical and physical properties (Fig.2). The high degree of conjugation through the system allowed us to observe sufficiently resolved ^{19}F chemical shifts of individual NAHs despite the differences in chemistry being up to 13 bonds away[42] (see Fig. S2).

To minimise the time required to reach equilibrium, the use of aniline as a nucleophilic catalyst for NAH exchange was introduced by Dawson[11], and has since been successfully used in many protein-templated DCC experiments[5,6]. From preliminary ^{1}H-STD-NMR experiments we observed that aniline was binding to FabH and would potentially interfere with the DCL. Issues with aniline were noted by Blanden *et al.* in their attempts to optimise hydrazone ligation for biomolecular labelling. They identified 4-amino-L-phenylalanine (4-APA) as a suitable replacement for aniline so we investigated if 4-APA could be substituted for aniline in a DCL[43]. We validated that 4-APA has comparable catalytic efficiency to aniline with our NAH DCL, both in the forward reaction (Fig. 3 and S4) and the reverse reaction (Fig. S5). When a library is prepared from product **A5** and hydrazides **3** and **4** (see Fig. 2), the same product distribution is observed as when a library is prepared from hydrazides **3**, **4** and **5** and aldehyde **A** (Fig. S4 and S5). Binding of 4-APA to FabH was not observed by ^{1}H-STD-NMR.

The DCL was assembled with each of the hydrazides **1-5** and aldehyde **A** at a final concentration of 200 µM, buffered at pH 6.2 using 50 mM sodium phosphate buffer with 10% DMSO and 3 mM 4-APA. The DCL cocktail was spiked with 200 µM 5-fluorouracil as an internal reference and either 200 µM *ec*FabH or an equivalent volume of buffer. Each library was transferred to an NMR tube and a ^{19}F spectrum was acquired every two hours over a twelve-hour period. In agreement with the 4-APA characterisation experiment (Fig. 3), the DCL reached equilibrium within 2 hours. Fig. 4a shows the blank library 8

hours after mixing. Notably not all products are at an equal concentration, suggesting the electronic substituents of the ring have an effect on the intrinsic thermodynamic stability of each NAH. Fig. 4b shows the ^{19}F spectrum of the library templated by the target, *ec*FabH. Slight signal broadening is observable for all DCL members, suggesting a slow exchange process is occurring between bound and unbound ligand states. Most impressively, the signal broadening and upfield chemical shift perturbation of compound **A4** is indicative of the compound being present in its bound state at a relatively higher proportion to the other library members[27].

It must be noted that there are two opposing forces at play when considering signal integral: on one hand Le Chatelier's principle will strive toward a product distribution proportional to the depth of the energy well of each species, although this is not always the case[2]. Therefore, if product **A4** is selected by the protein its relative concentration and corresponding signal integral should increase. Concomitantly, the signal broadening resulting from exchange between chemical shifts corresponding to bound and unbound states of the ligand may present itself as an apparent integral decrease of the bound ligand. Due to the breadth and low intensity of the bound state signal we have as yet been unable to determine the chemical shift of the bound state by ^{19}F COSY-NMR. This method is therefore not strictly quantitative, but acts as a qualitative indicator of potential binders whose properties can be further characterised by quantitative techniques.

Using individually synthesised DCL members an *in vitro* assay validated the results from the ^{19}F NMR DCC experiment. All 5 NAHs showed inhibitory activity at concentrations of 3 mM. Gratifyingly, compound **A4** showed the strongest inhibition, causing a 50% decrease in activity at 3 mM compared to the DMSO vehicle control and the FabH inhibitor HR45 as a negative control[44,45] (Fig. S6). That such a weak inhibitor could have been picked out of the NMR analysis is encouraging. It also leads to the question of what effect would a tight nM binder have? Broadening of the signal such that it disappears altogether into the baseline may also be useful in identifying hits by comparison with protein-free controls. We do not seek to claim that the NAH molecules described here may be of therapeutic value, rather we present this as a proof of concept that ^{19}F NMR analysis can be used to interrogate an NAH DCL derived from a an appropriately F-labelled building block.

Conclusions

In summary, we have demonstrated a non-invasive analysis of a protein-templated DCL by ^{19}F NMR, using 4-APA as a biologically benign alternative to aniline-catalysed NAH exchange. The results from the DCL agreed well with preliminary inhibition data from an *in vitro* FabH assay. Screening of much larger compound libraries will no doubt require multiple methods to identify, then validate, hit molecules. A combination of ^{19}F fragment screening and MS-based methods are complementary analytical tools suitable for protein-templated DCC. Development of such rapid, cost-effective, and universal methods should help DCC become more widely used in hit discovery.

Experimental

Nuclear magnetic resonance (NMR) spectra were recorded at 298 K on Bruker PRO500, AVA400 or AVA500 spectrometers running at 500 MHz (^{1}H spectra), 126 MHz (^{12}C spectra) or 94 MHz (^{19}F spectra). Chemical shift values (δ) are reported in parts per million (ppm) relative to tetramethylsilane (δ_{TMS} = 0 ppm) and are referenced to the residual solvent peak, or to the signal of internal standard 5-fluorouracil (δ_{5FU} = -169.19 ppm) in the case of the ^{19}F-labelled DCL. ^{1}H NMR data are reported in the format: chemical shift, relative intensity, multiplicity (s = singlet, d = doublet, t = triplet, m = multiplet, br = broad), coupling constant (*J* value, Hz), and assignment. ^{13}C NMR data are reported in the format: chemical shift and assignment.

Expression and purification of *ec*FabH. The *ecFabH*/pET-28a construct (4 µL) was transformed into an aliquot (50 µL) of BL21(DE3) cells and set on ice for 25 minutes. The cells were heat shocked at 42 °C for 40 seconds and set back on ice for a further 2 minutes. SOC media (100 µL) was added and the mixture was agitated at 37 °C for 1 hour. The mixture was spread on LB agar (30 µg/mL kanamycin) and incubated overnight at 37 °C. A single transformant was used to inoculate two seed cultures of sterile LB broth (2 x 250 mL, 30 µg/mL kanamycin) and agitated overnight at 37 °C. One of the overnight seed cultures was used to sub-culture sterile LB broth (5 x 500 mL, 30 µg/mL kanamycin) to an OD_{600} of 0.1. The cultures were agitated at 37 °C until the OD_{600} reached 0.6, at which point expression was induced by addition of IPTG (final conc. 0.1 mM). Cells were harvested after a further 3 hours at 30 °C and subsequently stored at -20 °C.

N-terminal histidine-tagged *ec*FabH was purified at 4 °C by Ni-affinity chromatography followed by size exclusion chromatography. The BL21 (DE3) cell pellet expressing *ec*FabH was resuspended in lysis buffer (30 mL, 20 mM Tris-HCl pH 7.6, 300 mM NaCl, 5 mM imidazole) and lysed for 15 minutes with rounds of 30 second of sonication followed by 30 seconds of rest. The cell lysate was clarified by centrifugation (*18,000 g*, 30 minutes, 4 °C) and the cell-free extract was injected onto a HisTrap 5 mL (GE Healthcare) Ni^{2+}-affinity chromatography column pre-equilibrated in lysis buffer. The column was washed with lysis buffer (5 CV) before the histidine-tagged protein was eluted using a gradient (0-100%) of lysis buffer to elution buffer (20 mM Tris-HCl pH 7.6, 300 mM NaCl, 400 mM imidazole) over 20 CV. Each elution fraction was analysed by SDS-PAGE, and the fractions containing His-tagged *ec*FabH were pooled, and concentrated to a volume less than 5 mL. *ec*FabH was further purified by size exclusion chromatography (HiLoad Superdex 200 16/60, GE Healthcare) with an isocratic elution of mobile phase buffer (20 mM Tris-HCl pH 7.6, 100 mM NaCl, 10% glycerol) at 1 mL/min over 120 minutes. *ec*FabH eluted at approximately 70 minutes and the most concentrated fractions were pooled and flash frozen and stored at -80 °C.

DCL conditions. The library was assembled with each of the hydrazides 1-5 and aldehyde A at a final concentration of 200 µM, buffered at pH 6.2 using 50 mM sodium phosphate with 50% D_2O, 10% DMSO, 50 mM NaCl and 3 mM 4-APA. The library cocktail also contained 200 µM 5-fluorouracil as a non-competing internal reference and either 200 µM *ec*FabH or an equivalent volume of enzyme purification buffer. Each library was transferred to an NMR tube and a ^{19}F spectrum (94 MHz, 512 scans, T_1 = 1 second) was acquired every two hours over a twelve-hour period.

Length of relaxation time in ^{19}F NMR pulse sequence. A three-member library was assembled from *N*-acylhydrazones A1, A3 and A5 (200 µM each) in sodium phosphate buffer (50 mM, D_2O, pH 6.2), 5-fluorouracil (internal standard, 200 µM) and a total of 10% DMSO. A series of ^{19}F NMR experiments were conducted with 512 scans and relaxation times of 1, 2, 3 *or* 4 seconds. The integral of the signals corresponding each *N*-acylhydrazone relative to the internal standard were compared at different relaxation times to determine the required relaxation time for the DCL ^{19}F NMR experiment. No difference in relative signal integral was apparent between all T_1 intervals, therefore a T_1 of 1 second was used in the DCL experiments.

FabH assay. The activity of FabH was quantified by 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB) assay with the following procedure. An assay cocktail was prepared from DTNB (1 mL, 4 mM in 50 mM NH_4OAc), HEPES buffer (1 mL, 1 M, pH 8.0) and ddH_2O (8 mL). The assay was run in 96-well plate format with the following protocol (the following bracketed figures refer to final concentrations): 100 µL DTNB assay cocktail, 25 µL *ec*FabH (0.63 µM), 25 µL Acetyl-CoA (375 µM), 25 µL Malonyl-CoA (375 µM), 5 µL ddH_2O, 20 µL NAH A1-A5 (3 mM) *or* HR45 (200 µM) in DMSO *or* DMSO control. Following 20 mins incubation at 37 °C, the reaction was initiated by the addition of malonyl-CoA and followed by monitoring absorbance change at 412 nm over 20 mins at 37 °C. The DMSO concentration was maintained at 10% to aid the solubility of the NAH compounds.

General procedure for the synthesis of *N*-acylhydrazones A1-A5. 2-Fluoro-5-formylbenzoic acid (50 mg, 0.30 mmol, 1.0 eq.) and the hydrazide (0.33 mmol, 1.1 eq.) were dissolved in ethanol (2 mL). A few drops of glacial acetic acid were added and the mixture was stirred overnight at room temperature. The solid formed was collected by vacuum filtration and washed in diethyl ether (5 mL) and water (5 mL) to yield the target *N*-acylhydrazone as a white solid.

A1 (54 mg, 0.188 mmol, 63%): m.p. 263-264°C; IR (nujol, cm^{-1}) 3497, 3402, 3196, 3073, 3028, 1657, 1612, 1564; ^1H NMR (500 MHz, DMSO) δ 12.17 (1H, s, OH-1), 8.83–8.78 (2H, m, CH-15,17), 8.50 (1H, s, CH-9), 8.27 (1H, m, CH-5), 8.02–7.99 (1H, m, CH-8), 7.86–7.81 (2H, m, CH-14,18), 7.44 (1H, dd, *J* = 10.6, 8.6 Hz, CH-6); ^{13}C NMR (126 MHz, DMSO) δ 165.1, 162.3, 150.8, 147.5, 140.8, 133.8, 131.1, 130.9, 112.0, 118.4, 118.2; ^{19}F NMR (470 MHz, DMSO) δ -110.58; HRMS m/z (ESI^+) calcd. $C_{14}H_{10}FN_3O_3$ $[M+Na]^+$ requires 310.0598, found 310.0599.

A2 (72 mg, 0.251 mmol, 84%): m.p. 265-266°C; IR (nujol, cm^{-1}) 3196, 3065, 2446, 2426, 1647, 1618, 1601, 1553; ^1H NMR (500 MHz, DMSO) δ 12.12 (1H, s, OH-1), 9.10–9.05 (1H, m, CH-18), 8.78 (1H, d, *J* = 3.7 Hz, CH-16), 8.48 (1H, s, CH-9), 8.28–8.26 (2H, m, CH-5,14), 7.59 (1H, dd, *J* = 7.8, 4.9 Hz, CH-15), 7.44 (1H, dd, *J* = 10.5, 8.6 Hz, CH-6); ^{13}C NMR (126 MHz, DMSO) δ 165.6, 165.1, 152.9, 149.1, 146.9, 136.0, 133.7, 131.2, 130.8, 129.6, 124.1, 120.5, 118.4, 118.4, 118.2, 31.2; ^{19}F NMR (470 MHz, DMSO) δ -110.75; HRMS m/z (ESI^+) calcd. $C_{14}H_{10}FN_3O_3$ $[M]^+$ requires 288.0779, found 288.0786.

A3 (82 mg, 0.259 mmol, 87%): m.p. 223-224°C; IR (nujol, cm^{-1}) 3401, 3225, 2995, 1701, 1659, 1618, 1589, 1545; ^1H NMR (500 MHz, DMSO) δ 11.91 (1H, s, OH-1), 8.49 (1H, s, CH-9), 8.27–8.21 (1H, m, CH-5), 7.98 (1H, s, CH-8), 7.51–7.38 (4H, m, CH-14,15,16,18), 7.21–7.15 (1H, m, 6), 3.85 (3H, s, CH_3-19); ^{13}C NMR (126 MHz, DMSO) δ 165.1, 163.47, 161.3, 159.7, 146.3, 135.2, 133.6, 131.4, 130.7, 130.2, 120.5, 120.3, 118.3, 118.1, 118.1, 113.4, 55.9, 31.2; ^{19}F NMR (470 MHz, DMSO) δ -110.97; HRMS m/z (ESI^+) calcd. $C_{16}H_{13}FN_2O_4$ $[M+Na]^+$ requires 339.0752, found 339.0755.

A4 (33 mg, 0.096 mmol, 32%): m.p. 230-232°C; IR (nujol, cm^{-1}) 3441, 3289, 2953, 2619, 2596, 1701, 1612, 1557, 1503; ^1H NMR (500 MHz, DMSO) δ 11.88 (1H, s, OH-1), 8.48 (1H, s, CH-9), 8.24 (1H, d, *J* = 6.9 Hz, CH-5), 7.99–7.96 (1H, m, CH-8), 7.86 (2H, d, *J* = 8.1 Hz, CH-14,18), 7.56 (2H, d, *J* = 8.3 Hz, CH-15,17), 7.42 (1H, m, CH-6), 1.33 (9H, s, CH_3-19,20,21); ^{13}C NMR (126 MHz, DMSO) δ 172.4, 165.1, 164.0, 146.0, 144.2, 130.6, 128.0, 125.8, 120.4, 118.1, 35.2, 31.4, 31.2, 21.5; ^{19}F NMR (470 MHz, DMSO) δ -111.08; HRMS m/z (ESI^+) calcd. $C_{19}H_{19}FN_2O_3$ $[M]^+$ requires 343.1453, found 343.1455.

A5 (72 mg, 0.226 mmol, 73%): m.p. 260-263°C; IR (nujol, cm^{-1}) 3292, 3092, 1695, 1651, 1618, 1589, 1557, 1520; ^1H NMR (500 MHz, DMSO) δ 11.71 (1H, s, OH-1), 9.71 (1H, s, OH-20), 8.47 (1H, s, CH-5), 8.23 (1H, dd, *J* = 7.3, 2.3 Hz, CH-8), 7.95 (1H, d, *J* = 6.7 Hz, CH-6), 7.48–7.44 (1H, m, CH-18), 7.44–7.38 (1H, m, CH-14), 6.89 (1H, d, *J* = 8.2 Hz, CH-15), 3.86 (3H, s, CH_3-19); ^{13}C NMR (126 MHz, DMSO) δ 165.1, 163.2, 161.1, 150.7, 147.8, 133.4, 133.3, 131.6, 130.5, 124.5, 121.9, 120.4, 120.3, 118.3, 118.1, 115.5, 112.2, 56.3, 19.0; ^{19}F NMR (470 MHz, DMSO) δ -111.21; HRMS m/z (ESI^+) calcd. $C_{16}H_{13}FN_2O_5$ $[M+Na]^+$ requires 355.0701, found 355.0705.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

We would like to thank the Scottish Universities Life Sciences Alliance (SULSA) for funding a PhD studentship (AGE).

Notes and references

1 O. Ramström and J.-M. Lehn, *Nat. Rev. Drug Discov.*, 2002, **1**, 26–36.

2 A. Herrmann, *Chem. Soc. Rev.*, 2014, **43**, 1899–1933.

3 M. Mondal and A. K. H. Hirsch, *Chem. Soc. Rev.*, 2015, **44**, 2455–2488.

4 S. R. Beeren and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2011, **133**, 3804–3807.

5 V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano and M. F. Greaney, *Nat. Chem.*, 2010, **2**, 490–497.

6 A. J. Clipson, V. T. Bhat, I. McNae, A. M. Caniard, D. J. Campopiano and M. F. Greaney, *Chem. A Eur. J.*, 2012, **18**, 10562–10570.

7 M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe and A. K. H. Hirsch, *Angew. Chemie Int. Ed.*, 2014, **53**, 3259–3263.

8 S.-A. Poulsen, *J. Am. Soc. Mass Spectrom.*, 2006, **17**, 1074–1080.

9 J. Kalia and R. T. Raines, *Angew. Chemie Int. Ed.*, 2008, **47**, 7523–7526.

10 E. H. Cordes and W. P. Jencks, *J. Am. Chem. Soc.*, 1962, **84**, 826–831.

11 A. Dirksen, S. Dirksen, T. M. Hackeng and P. E. Dawson, *J. Am. Chem. Soc.*, 2006, **128**, 15602–15603.

12 R. Huang and I. Leung, *Molecules*, 2016, **21**, 1–18.

13 M. Hochgürtel, H. Kroth, D. Piecha, M. W. Hofmann, C. Nicolau, S. Krause, O. Schaaf, G. Sonnenmoser and A. V Eliseev, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 3382–3387.

14 I. Huc and J. M. Lehn, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 2106–2110.

15 O. Ramström and J. M. Lehn, *Chembiochem*, 2000, **1**, 41–48.

16 T. Hotchkiss, H. B. Kramer, K. J. Doores, D. P. Gamblin, N. J. Oldham and B. G. Davis, *Chem. Commun.*, 2005, 4264–4266.

17 R. F. Ludlow and S. Otto, *Components*, 2008, **3**, 12218–12219.

18 B. M. R. Liénard, N. Selevsek, N. J. Oldham and C. J. Schofield, *ChemMedChem*, 2007, **2**, 175–179.

19 R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero and O. Ramström, *Angew. Chemie Int. Ed.*, 2010, **49**, 589–593.

20 M. Demetriades, I. K. H. Leung, R. Chowdhury, M. C. Chan, M. A. McDonough, K. K. Yeoh, Y. M. Tian, T. D. W. Claridge, P. J. Ratcliffe, E. C. Y. Woon and C. J. Schofield, *Angew. Chemie Int. Ed.*, 2012, **51**, 6672–6675.

21 T. Bunyapaiboonsri, O. Ramström, S. Lohmann, J. M. Lehn, L. Peng and M. Goeldner, *Chembiochem*, 2001, **2**, 438–444.

22 M. S. Congreve, D. J. Davis, L. Devine, C. Granata, M. O'Reilly, P. G. Wyatt and H. Jhoti, *Angew. Chemie Int. Ed.*, 2003, **42**, 4479–4482.

23 M. Misuraca and E. Moulin, *New J. Chem.*, 2014, **38**, 3336–3349.

24 Z. Fang, W. He, X. Li, Z. Li, B. Chen, P. Ouyang and K. Guo, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 5174–5177.

25 E. C. Y. Woon, M. Demetriades, E. A. L. Bagg, W. Aik, S. M. Krylova, J. H. Y. Ma, M. Chan, L. J. Walport, D. W. Wegman, K. N. Dack, M. A. McDonough, S. N. Krylov and C. J. Schofield, *J. Med. Chem.*, 2012, **55**, 2173–2184.

26 A. Vulpetti and C. Dalvit, *ChemMedChem*, 2013, **8**, 2057–2069.

27 C. Dalvit and A. Vulpetti, *Magn. Reson. Chem.*, 2012, **50**, 592–597.

28 C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2007, **51**, 243–271.

29 J. W. Sadownik and D. Philp, *Org. Biomol. Chem.*, 2015, **13**, 10392–10401.

30 T. Kosikova, H. MacKenzie and D. Philp, *Chem. A Eur. J.*, 2016, **22**, 1831–1839.

31 J. W. Sadownik and D. Philp, *Angew. Chemie Int. Ed.*, 2008, **47**, 9965–9970.

32 M. A. Henen, L. Abou-Zeid, A. El-Din Barghash and H. Eisa, *Curr. Top. Pept. Protein Res.*, 2014, **15**, 35–39.

33 Z. Nie, C. Perretta, J. Lu, Y. Su, S. Margosiak, K. S. Gajiwala, J. Cortez, V. Nikulin, K. M. Yager, K. Appelt and S. Chu, *J. Med. Chem.*, 2005, **48**, 1596–1609.

34 J. W. Campbell and J. E. Cronan Jr., *Annu. Rev. Microbiol.*, 2001, **55**, 305–332.

35 J. Wang, S. M. Soisson, K. Young, W. Shoop, S. Kodali, A. Galgoci, R. Painter, G. Parthasarathy, Y. S. Tang, R. Cummings, S. Ha, K. Dorso, M. Motyl, H. Jayasuriya, J. Ondeyka, K. Herath, C. Zhang, L. Hernandez, J. Allocco, A. Basilio, J. R. Tormo, O. Genilloud, F. Vicente, F. Pelaez, L. Colwell, S. H. Lee, B. Michael, T. Felcetto, C. Gill, L. L. Silver, J. D. Hermes, K. Bartizal, J. Barrett, D. Schmatz, J. W. Becker, D. Cully and S. B. Singh, *Nature*, 2006, **441**, 358–361.

36 J. Wang, S. Kodali, S. H. Lee, A. Galgoci, R. Painter, K. Dorso, F. Racine, M. Motyl, L. Hernandez, E. Tinney, S. L. Colletti, K. Herath, R. Cummings, O. Salazar, I. González, A. Basilio, F. Vicente, O. Genilloud, F. Pelaez, H. Jayasuriya, K. Young, D. F. Cully and S. B. Singh, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 7612–7616.

37 F. Zhang, Q. Wen, S.-F. Wang, B. Shahla Karim, Y.-S. Yang, J.-J. Liu, W.-M. Zhang and H.-L. Zhu, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 90–95.

38 Y. Li, C. P. Zhao, H. P. Ma, M. Y. Zhao, Y. R. Xue, X. M. Wang and H. L. Zhu, *Bioorg. Med. Chem.*, 2013, **21**, 3120–3126.

39 L. Shi, R.-Q. Fang, Z.-W. Zhu, Y. Yang, K. Cheng, W.-Q. Zhong and H.-L. Zhu, *Eur. J. Med. Chem.*, 2010, **45**, 4358–4364.

40 X. L. Wang, Y. Bin Zhang, J. F. Tang, Y. S. Yang, R. Q. Chen, F. Zhang and H. L. Zhu, *Eur. J. Med. Chem.*, 2012, **57**, 373–382.

41 X. Song, Y. Yang, J. Zhao and Y. Chen, *Chem. Pharm. Bull.*, 2014, **62**, 1110–1118.

42 J. T. Gerig, *Methods Enzymol.*, 1989, **177**, 3–23.

43 A. R. Blanden, K. Mukherjee, O. Dilek, M. Loew and S. Bane, *Bioconjug. Chem.*, 2011, **22**, 1954–1961.

44 X. He, A. M. Reeve, U. R. Desai, E. Glen, K. A. Reynolds and G. E. Kellogg, *Antimicrob. Agents Chemother.*, 2004, **48**, 3093–3102.

45 A. G. Ekström, V. Kelly, J. Marles-Wright, S. L. Cockroft and D. J. Campopiano, *Org. Biomol. Chem.*, 2017, **15**, 6310–6313.