RESEARCH ARTICLE

A Fragment-based Competitive ^{19}F LB-NMR Platform For Hotspot-directed Ligand Profiling

William J. McCarthy,[a] Sherine E. Thomas,[b] Tayo Olaleye,[a] Jennifer A. Boland,[a] R. Andres Floto,[c][d] Glyn Williams,[a] Tom L. Blundell,*[b][d] Anthony G. Coyne,*[a] Chris Abell†[a]

[a] Dr W. J. McCarthy, Dr T. Olaleye, Dr J. A. Boland, Dr G. Williams, Dr A. G. Coyne, Prof C. Abell
Yusuf Hamied Department of Chemistry
University of Cambridge
Lensfield Road, Cambridge, UK, CB2 1EW
E-mail: agc40@cam.ac.uk

[b] Dr S. E. Thomas, Prof T. L. Blundell
Department of Biochemistry
University of Cambridge
Tennis Court Road, Cambridge, UK, CB2 1GA
E-mail: tlb20@cam.ac.uk

[c] Prof R. A. Floto
University of Cambridge Molecular Immunity Unit
MRC Laboratory of Molecular Biology
Cambridge, UK, CB2 0QH

[d] Prof T. L. Blundell, Prof R. A. Floto
VPD Heart Lung Research Institute, Department of Medicine
University of Cambridge
Cambridge, UK, CB2 0BB

Present Addresses:
Dr W. J. McCarthy
Molecular Structure of Cell Signalling Laboratory, The Francis Crick Institute, 1 Midland Road, London, UK, NW1 1AT
Dr S. E. Thomas
Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, UK, CB2 1PD

[†] Prof. C. Abell passed away in October 2020

Supporting information for this article is given via a link at the end of the document. Crystal structures of the *Mabs* PPAT-ligand complexes are deposited in the Protein Data Bank (http://www.rcsb.org/pdb/) under the following accession codes: 7YY0 (PPSH), 7YY3 (**1**), 7YY4 (**2**), 5O0C (**4**), 5O0A (**7**), 8QID (**8**), 8QIY (**9**), 7YY9 (**10**), 8QJ8 (**11**), 5O0D (**12**), 8QIX (**17**).

Abstract: Ligand binding hotspots are regions of protein surfaces that form particularly favourable interactions with small molecule pharmacophores. Targeting interactions with these hotspots maximises the efficiency of ligand binding. Existing methods are capable of identifying hotspots but often lack assays to quantify ligand binding and direct elaboration at these sites. Herein, we describe a fragment-based competitive ^{19}F Ligand Based-NMR (LB-NMR) screening platform that enables routine, quantitative ligand profiling focused at ligand-binding hotspots. As a proof of concept, the method was applied to 4'-phosphopantetheine adenylyltransferase (PPAT) from *Mycobacterium abscessus* (*Mabs*). X-ray crystallographic characterisation of the hits from a 960-member fragment screen identified three ligand-binding hotspots across the PPAT active site. From the fragment hits a collection of ^{19}F reporter candidates were designed and synthesised. By rigorous prioritisation and use of optimisation workflows, a single ^{19}F reporter molecule was generated for each hotspot. Profiling the binding of a set of structurally characterised ligands by competitive ^{19}F LB-NMR with this suite of ^{19}F reporters recapitulated the binding affinity and site ID assignments made by ITC and X-ray crystallography. This quantitative mapping of ligand binding events at hotspot level resolution establishes the utility of the fragment-based competitive ^{19}F LB-NMR screening platform for hotspot-directed ligand profiling

Introduction

Fragment-based drug discovery (FBDD) is a powerful ligand discovery methodology that has found widespread use in both academia and industry.[1–3] As a drug discovery tool FBDD is well-established, with seven FDA approved drugs originating from a fragment based approach.[4–10] As the FBDD field has matured, more specific fragment-based applications have emerged to study focused questions about biological targets. One prominent example is the use of fragment screening for mapping of ligand binding hotspots on protein surfaces.[11]

Hotspots are regions of a protein's surface that form particularly favourable interactions with pharmacophores present in small molecules.[12] Efficient fragment elaboration strategies combine pharmacophores that interact with multiple, spatially proximal

hotspots into a single ligand. To achieve this, discovery of hotspots and identification of chemical matter engaging them is necessary. Fragment screening has proven to be a particularly effective method of empirically mapping ligand binding hotspots. This is underpinned by the fragments' ability to comprehensively sample the protein's molecular recognition capacity, owing to their minimal molecular complexity. The most recent iterations of this approach map hotspots by screening targets against libraries of small fragments specifically designed for detection by X-ray crystallography.[13,14] These methods provide structural characterisation of the hotspots and suggest the types of interactions that define them. However, while methods capable of directing ligand elaboration exist, systematic approaches to directing elaboration at discovered hotspots are lacking.

Herein, we present a methodology that re-purposes fragment hits found to engage hotspots, defined by X-ray crystallography, as ^{19}F reporter molecules that are integrated into a competitive ^{19}F-observed ligand-based NMR (LB-NMR) screening platform. The ^{19}F reporters enable hotspot-directed profiling of other ligand binding events, connecting the identification of ligand binding hotspots to the production of high-quality small molecules that engage them.

The properties of fragment-sized molecules and the features of the LB-NMR methodology complement each other particularly well to detect and quantify ligand binding events in a hotspot mapping platform. LB-NMR can detect binding of the reporter at very low fractional occupancy (typically ≈ 1% ligand bound), owing to accumulation of the bound-state signal in the free-state population as a result of rapid exchange between the bound and unbound fractions. Combined with slower NMR relaxation in the free-state, this allows the bound state signal to be transferred to the free population and amplified during the NMR experiment.[15,16] Fragments are favourable as competitive reporters since their inherently weak affinities allow them to be displaced readily from their preferred binding site.

Although great progress has been made recently in the development of saturation transfer difference (STD)[17] and Carr-Purcell-Meiboom-Gill (CPMG)[18,19] LB-NMR methods for the direct quantitation of binding affinities in the fragment affinity range (low μM – low mM),[20–23] direct approaches are limited in throughput and information content relative to indirect methods. Indirect, competitive binding experiments have the advantage that optimisation of experimental conditions and NMR parameters, as well as characterisation of the ligand binding site, are only required for the relatively small set of reporter molecules.

^{19}F-observed LB-NMR is a well established method for the qualitative identification of ligand:protein interactions.[24–26] To quantify the affinities of ligand:protein interactions, indirect, competitive ^{19}F LB-NMR approaches have also been described.[16,27–30] Several excellent reviews of the theory and practical considerations that underpin these ^{19}F LB-NMR applications are available,[20,31–33] For the purposes of this work, the most important conclusion is that the transverse relaxation of the ^{19}F nucleus is particularly sensitive to changes in its chemical environment. As such, changes in ligand ^{19}F relaxation times on protein binding are typically far greater than changes in their ^{1}H relaxation times. For these reasons, we have established a hotspot-directed ligand profiling platform using the competitive ^{19}F-observed CPMG LB-NMR method.

To establish the method, 4'-phosphopantetheine adenylyltransferase (PPAT) from Mycobacterium abscessus (Mabs) was selected as a model target. Bacterial CoA biosynthesis is an attractive antibacterial target.[34,35] Recently, chemogenomic methods have demonstrated mycobacterial PPAT is a ligandable target with potential as a point of intervention in the CoA biosynthetic pathway.[36] The PPAT active site contains two distinct substrate sites, binding adenosine triphosphate (ATP) and 4'-phosphopantetheine (PPSH), respectively. These substrate sites each comprise two binding sub-sites (**Figure 1A**).[32] The PPSH binding site contains the mercaptamide sub-site and a henceforth termed cryptic pocket. The ATP binding site contains the ribose and adenine sub-sites.

In this study, we set out a workflow for implementing the hotspot-directed ligand profiling methodology that follows the general structure outlined in **Figure 1B**: (1) ligand binding hotspots are identified by screening of a fragment library, with X-ray crystallography applied to structurally understand the molecular interactions defining the hotspots, (2i) fragment hits that engage these hot spots are modified to introduce fluorine atoms in appropriate positions, (2ii) within individual hotspots the highest quality ^{19}F reporter candidates are prioritised, (2iii) optimal conditions for application of each ^{19}F reporter in competitive ^{19}F LB-NMR experiments are established, and (3) the established ^{19}F reporters are applied to quantitatively characterise ligand binding events at specific hotspots. In this competitive ^{19}F LB-NMR platform, site-ID information, provided by X-ray crystallography, is encoded into the fluorinated reporters. This allows them to both quantify the binding affinity and identify the site of binding of competitive ligands with hotspot level resolution.

Results and Discussion

^{19}F Reporter Generation and Selection

Recently, we conducted a fragment screen against *Mabs* PPAT.[37] A 960-member fragment library was screened by differential scanning fluorimetry (DSF), identifying 28 hits. 16 of these were subsequently structurally characterised by X-ray crystallography (**Figure S1**). Analysis of the binding mode of the fragment hits revealed the ribose and adenine sub-sites as well as the cryptic pocket to be ligand binding hotspots (**Figure 1A**). Discovery of the cryptic pocket as a hotspot highlights the advantage of experimental hotspot mapping (**Figure S2**), which is capable of identifying highly ligandable sub-sites that may have otherwise been overlooked by computational methods.[38] As well as identifying the ligand binding hotspots of *Mabs* PPAT, the fragment screen discovered chemical matter engaging these sites. From the original fragment screen, validated hits **1** and **2** were identified as promising fluorinated reporter candidates for the cryptic hotspot (**Figure S1**). Both contain a fluorinated motif and their binding modes were characterised by X-ray crystallography upfront. **3**, a hit identified in the DSF primary screen but not validated by X-ray crystallography, was also included as a

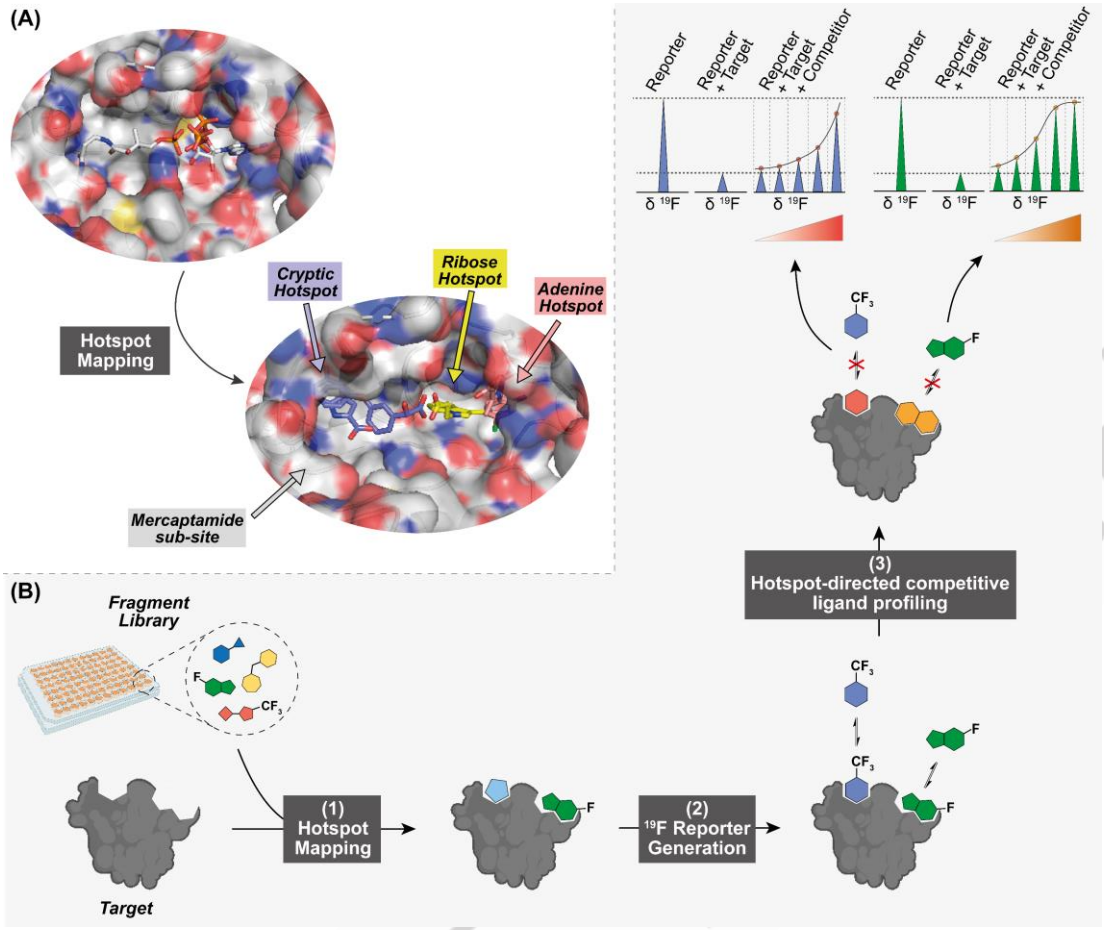


Figure 1. (A) X-ray crystal structure highlighting the active site of an *Mabs* PPAT monomer overlaid with substrates PPSH and ATP. Ligand binding hotspots are highlighted with bound fragment hits; cryptic (purple ligands), ribose (yellow ligand), adenine (pink ligands) hotspots. PDB entry: 7YY0 (PPSH), 5O0D (**12**). (B) Schematic representation of the workflow for identifying ligandable hotspots and focused profiling of ligand binding events at these hotspots by competitive ^{19}F LB-NMR.

fluorinated reporter candidate (**Figure S1**). Although its binding site was unassigned at this point, its aryl trifluoromethyl moiety was an attractive feature for ^{19}F LB-NMR. For the ribose and adenine hotspots, no fluorinated fragment hits were identified in the fragment screen (**Figure S1**). Introduction of fluorinated motifs into structurally characterized fragment hits by structure-based design appeared an efficient means of generating fluorinated reporter candidates for these hotspots. In the ribose hotspot, each hit contained a carboxylic acid, and the indole core was found to be a privileged scaffold (**Figure S1**). As such, fragment **4** was selected for modification and two fluorinated ribose candidates, **5** and **6**, were designed (**Figure S3A**). In the adenine hotspot, only two fragment hits were validated by X-ray crystallography (**Figure S1**). For the design of fluorinated reporter candidates, **7** was selected as scaffold. Two fluorinated analogues, **8** and **9**, were designed (**Figure S3B**).

Once produced, the seven fluorinated reporter candidates were passed through a biophysical filtering cascade (**Figure 2A**). This workflow was designed to select the highest quality reporter candidate within each hotspot for progression to ^{19}F CPMG experiments. The cascade consisted of four tiers. The first tier tested the candidate's stability, solubility and tendency to aggregate under the conditions envisioned for the CPMG experiments. The second tier assessed candidate binding to PPAT by ^{1}H- and ^{19}F-observed CPMG LB-NMR, establishing whether reporter binding was responsive to detection by NMR methods. The third tier attempted to assign binding sites to the reporter candidates, where not previously assigned, directly by X-ray crystallography or indirectly by competitive ^{1}H LB-NMR methods. The fourth and final tier attempted to quantify the binding of reporter candidates. Isothermal titration calorimetry (ITC) was used to determine affinities but, where this was not possible, differential scanning fluorimetry (DSF) was used as a comparative metric. The filtering cascade workflow is described in greater detail in **Section S1** and the outcomes summarised **Table S1**. This cascade selected three candidates for progression to characterisation of their ^{19}F LB-NMR reporter capabilities (**Figure 2A**, **Table S2**). **1** was chosen for the cryptic hotspot, **9** for the adenine hotspot and **3** was putatively selected for the ribose hotspot.

The X-ray crystal structure of **1** in complex with PPAT demonstrates its occupation of the cryptic hotspot (**Figure 2B, S5A**). The K_d of **1** at the cryptic hotspot was found to be 242 μM, using ITC. Attempts to characterise the binding mode and affinity of **3** by X-ray crystallography and ITC proved unsuccessful. By ITC, **3** produced very low enthalpies, thus no binding isotherm could be resolved (**Figure S4**). However, competitive ^{1}H LB-NMR experiments conducted in the third tier of the biophysical filtering.

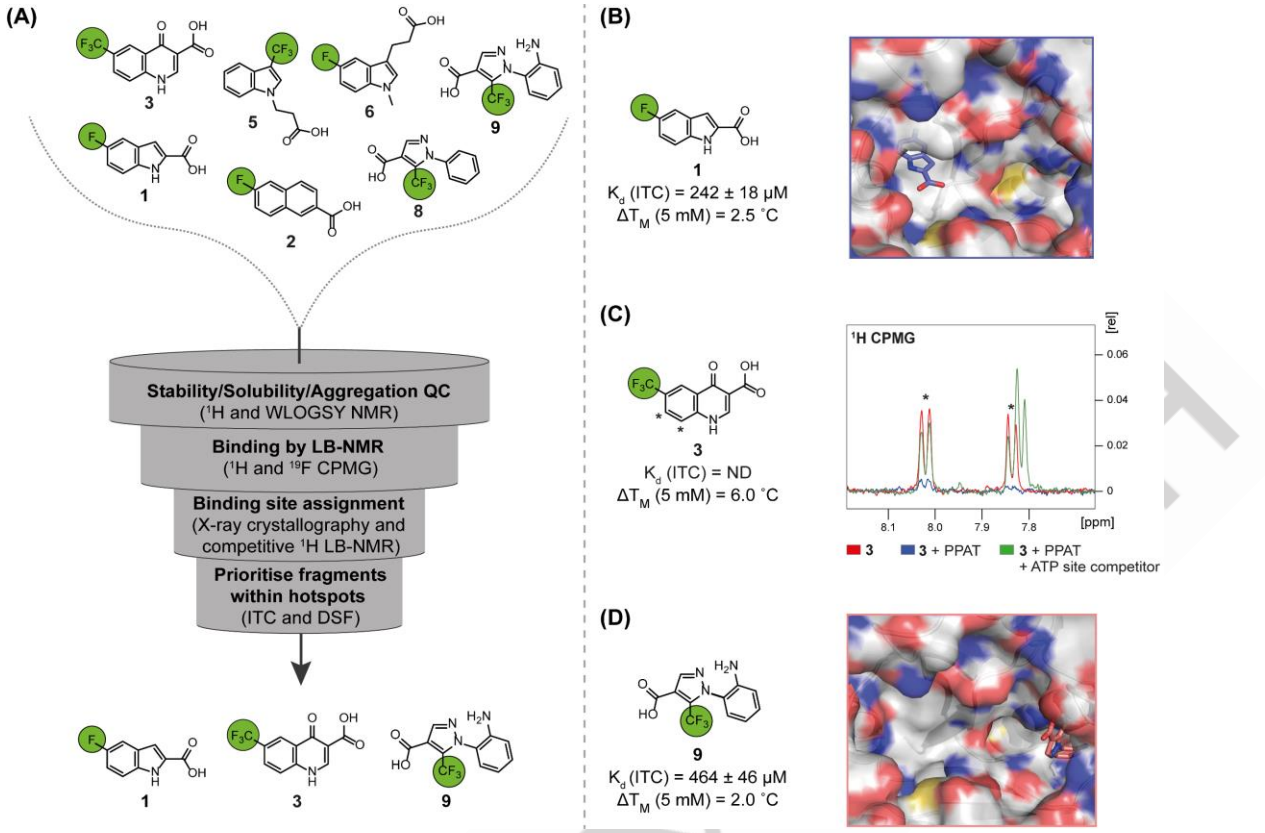


Figure 2. (A) Biophysical filtering cascade applied to select the highest quality fluorinated reporter candidate in each hotspot. **1, 3** and **9** were selected for the cryptic, ribose and adenine hotspots, respectively. (B) Cryptic hotspot candidate **1**. X-ray crystal structure of PPAT with **1** (purple) occupying the cryptic pocket. (C) Ribose hotspot candidate **3**. Competitive ^{1}H LB-NMR of **3** (showing protons of **3** indicated by *) demonstrating occupation of the ATP site by competition with an ATP site tool compound. (D) Adenine hotspot candidate **9**. X-ray crystal structure of PPAT with **9** (pink) occupying the adenine sub-site. PDB entries: 7YY3 (**1**), 8QIY (**9**).

cascade revealed **3** occupied the ATP site (**Figure 2C**), and later, competitive ^{19}F CPMG experiments showed **3** engages the ribose hotspot (*vide infra*). The X-ray crystal structure of **9** in complex with PPAT was solved and showed that it engaged the adenine hotspot as designed (**Figure 2D**), recapitulating the binding site of parent fragment **7** and engaging all key interactions defining the adenine hotspot (**Figure S5B**). The K_d of **9** at the adenine hotspot was found to be 464 µM, using ITC.

Optimisation of ^{19}F CPMG Experimental Conditions

Conversion of the selected fluorinated fragments into effective ^{19}F reporters first required optimisation of the ^{19}F CPMG parameters and experimental conditions. The target concentration and length of the CPMG delay were simultaneously optimised by determination of the T_2 relaxation rate (R_2) of the reporter ^{19}F nucleus across a range of PPAT concentrations. Furthermore, the reporter concentration was adjusted during the optimisation process to improve the sensitivity to competition experiments. The optimisation workflow implemented is described in detail in **Section S2**.

Firstly, the R_2 of the reporter candidates' ^{19}F nuclei were determined using the CPMG pulse sequence. The intensity of the ^{19}F signal decays as the implemented CPMG delay time increases (**Figure 3A**) and is defined by R_2. Fitting the normalised integration of the ^{19}F signal at each CPMG delay time to **Equation S1** determines the R_2 of the reporter ^{19}F nucleus. Experimental conditions were chosen such that all measured ^{19}F signals had a S/N ratio ≥ 3. The trifluoromethyl group of **9** was found to relax with an R_2 = 0.73 s^{-1} (**Figure 3B**), while the R_2 of the fluorine and trifluoromethyl moieties of **1** and **3** were found to be 0.57 s^{-1} and 1.67 s^{-1}, respectively (**Figures S6B, S7B**).

The difference in the reporter ^{19}F nucleus' R_2 in the presence and absence of protein defines the size of the assay window for competition experiments.[16] In practise, the increase in R_2 can be tuned by controlling the concentration of protein present. Calculation of the relaxation contrast value (c_2) guides selection of the protein concentration that produces the best resolution between the free and bound reporter states (**Equation S2**).[39] The reporters' ^{19}F R_2 values were recorded at several protein concentrations, and using the reporter R_2 values measured for the free (R_2^{free}) and protein-bound (R_2^{obs}) states, a series of c_2 values were calculated for the range of protein concentrations tested (**Figure 3C, S6C, S7C**). Generally, c_2 values greater than 70% produce satisfactorily large assay windows for competition experiments.[39] From this titration a single optimal protein concentration was chosen and used in all future experiments. For **9**, 25 µM was selected as an optimal PPAT concentration, producing an ^{19}F R_2 = 3.28 s^{-1}, and c_2 = 78% (**Figure 3D**). Similarly, 25 µM was selected as the optimal PPAT concentration for **3**, yielding an ^{19}F R_2 = 7.19 s^{-1}, and c_2 = 77% (**Figure S7D**). The ^{19}F R_2 of the fluorine nucleus in **1** was found to be much more responsive to protein concentration than the trifluoromethyl groups in **9** and **3**. As little as 0.75 µM PPAT produced an ^{19}F R_2

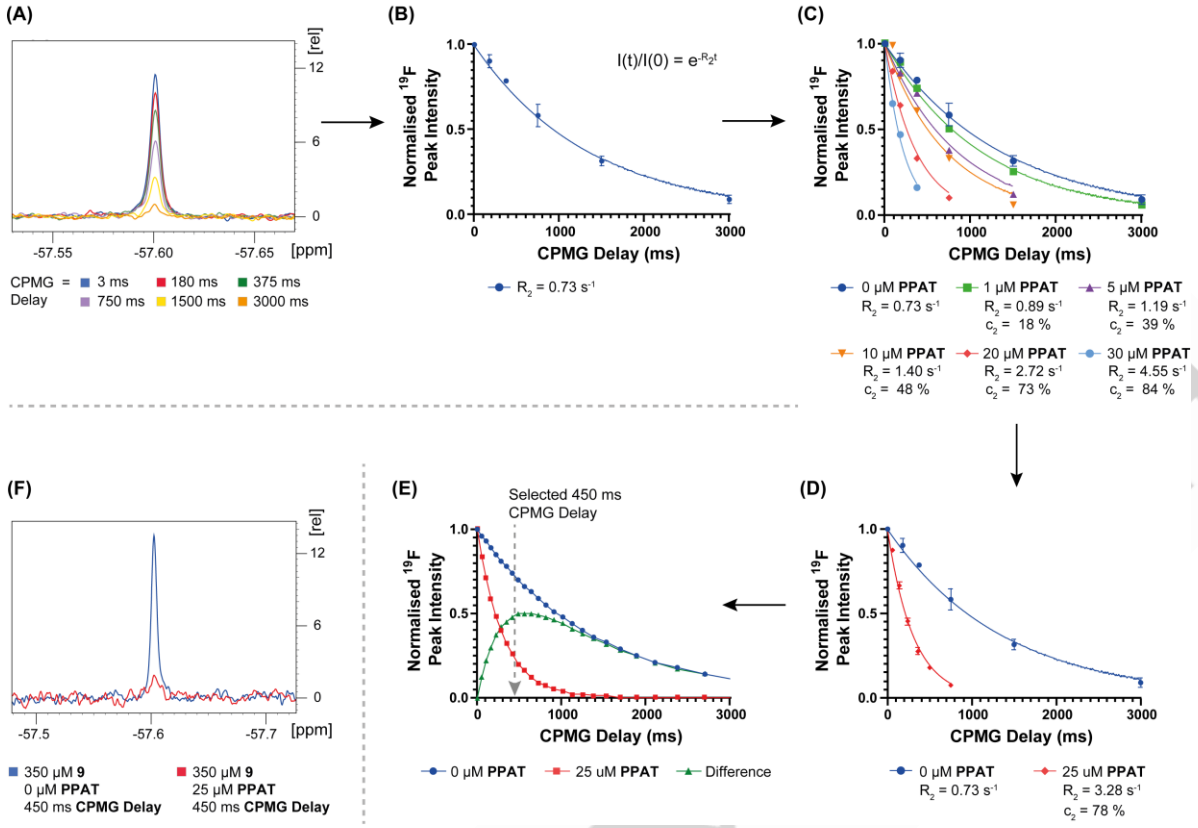


Figure 3. Optimisation of experimental conditions for use of **9** as an ^{19}F reporter. (A) ^{19}F NMR spectra of the trifluoromethyl signal of **9** at several CPMG delay lengths. (B) Determination of the trifluoromethyl signal R_2 from (A) using the exponential decay described in **Equation S1** (inset) (n=3). (C) Repetition of the R_2 measurement protocol from (B) in the presence of several PPAT concentrations and calculation of c_2 values for each condition according to **Equation S2** (n=1). (D) Measurement of the trifluoromethyl signal R_2 in the absence (blue) and presence (red) of the optimal target concentration (n≥2). (E) Simulation of the trifluoromethyl signal R_2 in the absence (blue) and presence (red) of the optimal target concentration and calculation of the difference curve (green). (F) ^{19}F NMR spectra of the trifluoromethyl signal of **9** at the optimised reporter concentration and CPMG delay in the absence (blue) and presence (red) of the optimised target concentration. Data in panels (A)-(E) was generated with 500 µM reporter **9**. Data in panel (F) was generated with 350 µM reporter **9**.

= 2.40 s^{-1}, and c_2 = 76% (**Figure S6D**). This ≈ 30-fold greater sensitivity of the ^{19}F nucleus of **1**, in comparison with **9** and **3**, is likely the result of a greater ^{19}F chemical shift anisotropy (CSA) between the free and bound reporters states for the mono-fluorinated reporter. The aryl fluoride of **1** is conjugated with the aromatic system and also coupled to local protons, and so is likely more sensitive to changes in the electron distribution of the indole ring on binding compared to the trifluoromethyl fluorines of **9** and **3**.

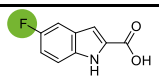
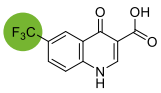
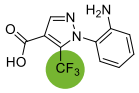
Next, a single optimum CPMG delay time for each reporter was sought. This optimum corresponds to the delay time that produces the largest difference in intensity of the reporter ^{19}F signal in the presence and absence of protein. The optimum CPMG delay time can be simply calculated using the measured R_2^{free} and R_2^{obs} values according to **Equation S3**. However, the consequences of any deviation from this optimum value on the size of the assay window can be understood quantitatively by simulation of the difference in ^{19}F signal intensities between the free and bound reproter states, across the CPMG delay timescale, according to **Equation S1**. Plotting the difference between the free and bound states as a curve (R_2^{free} - R_2^{obs}) pinpoints the optimum CPMG delay as the maximum of this simulated curve (**Figure 3E, S6E, S7E**).

In subsequent competition experiments, each ^{19}F reporter was used at a lower concentration than in these optimization experiments, facilitating easier reporter displacement. To maintain a robust ^{19}F signal, the lower ^{19}F reporter concentration was offset by applying a shorter CPMG delay than the theoretical optimum (**Table 1**). The simulation curves guided selection of a shorter CPMG delay that did not compromise significantly on the size of the assay window, ensuring a S/N ratio ≥ 3 was achieved for all measured ^{19}F signal intensities in the fewest scan numbers. The ^{19}F CPMG parameter optimisation workflow outlined here allowed three-dimensional optimisation of the (i) target concentration, (ii) CPMG delay length, and (iii) reporter concentration. Application of these conditions together established the assay window for subsequent competition experiments (**Figure 3F**). Optimal concentrations of reporter and PPAT as well as the optimal CPMG delay were identified for all three reporters and are summarised in **Table 1**.

Hotspot-directed Competitive ^{19}F LB-NMR Ligand Profiling

To validate the reporters' hotspot specificity, a single concentration displacement experiment with high concentration of competing ligands was performed. For this purpose, a set of

Table 1. Summary of the optimised experimental conditions for each of the ^{19}F reporter fragments.[a]

Reporter	Structure	K_d (ITC)	[Reporter]	[PPAT]	c_2	Optimal CPMG Delay	CPMG Delay Used
1		242 ± 18 µM	125 µM	0.75 µM	74% (±8%)	614 ms	360 ms
3		[b]	250 µM	25 µM	77% (±5%)	255 ms	192 ms
9		464 ± 46 µM	350 µM	25 µM	74% (±6%)	536 ms	450 ms

[a] Errors in K_d values represent the fitting error (n=2). Errors in c_2 values are propagated from the R_2 measurement error (n≥2). [b] A K_d value is not reported for **3** as the ITC binding isotherm could not be reliably fit.

ligands with well-characterised binding modes was assembled (**Figure 4A**). **11** is a ligand developed in our group that occupies the ATP site completely with a K_d = 3.4 µM (manuscript in preparation). The selected fragments together occupy the three hotspots targeted by the suite of fluorinated reporters. **12**, **4** and **7** occupy the cryptic, ribose and adenine hotspots, respectively. Also, as the hotspot occupied by **3** had not been determined directly, it was included in competition experiments with the other ^{19}F reporters to explore the hotspot(s) that it engages. Biophysical data for the validation set ligands is summarised in **Table S4**.

Whereas in the previous section, a reporters' ^{19}F R_2 was characterised by measuring the percentage decrease in intensity of its ^{19}F signal over a range of CPMG delays, displacement of the fluorinated reporter by competing ligands is measured by the percentage recovery in intensity of the ^{19}F signal that results from the decreased fractional occupancy of the reporter. The extent to which the reporter is displaced was quantified by integration of the ^{19}F peak and converted into a percentage displacement according to **Equation S4** (**Figure 4B**). The methodology for quantifying displacement of reporters by competing ligands in single concentration or dose-response experiments is detailed in **Section S3**.

Displacement of a reporter was considered to be significant if the ^{19}F signal integral in the presence of competitor exceeded the standard deviation (n≥3) of the ^{19}F signal integral in the absence of competitor by > 3-fold. **1**, as expected, was not competed by **11**, nor significantly by the ribose and adenine hotspot fragments **3**, **4** and **7**. **1** was only substantially displaced by **12**, the other cryptic hotspot binder in the set, demonstrating that **1** specifically reports on binding events at the cryptic hotspot. **9**, as expected, was competed almost quantitatively by **11**, confirming its occupation of the ATP site. **9** was not competed by the cryptic hotspot fragment, **12**, and only modestly by **3** and the ribose hotspot fragment **4**. **9** was substantially displaced by the other adenine hotspot binder, **7**, demonstrating its capacity to specifically map the adenine hotspot, within the ATP pocket. **3** was competed almost quantitatively by **11**. In combination with its lack of displacement by the cryptic hotspot fragment, **12**, this data confirmed **3** binds within the ATP site. Gratifyingly, **3** was substantially competed by the ribose hotspot fragment **4**, establishing it can report on ligand binding to the ribose hotspot**.**

To establish the capacity of the ^{19}F LB-NMR platform for quantitative mapping of ligand binding events, a series of competition experiments with ^{19}F reporters **1**, **3** and **9** was performed, using a set of structurally characterised *Mabs* PPAT ligands (**Figure 5A-C**). This set consisted of 11 compounds, 5 fragments and 6 more elaborate compounds developed in our group. The fragments included were all hits from the original fragment screen. **7** and **10** were the only adenine hotspot hits. **4** was representative of the indole containing fragments privileged in the ribose hotspot. **2** and **12** were selected as representative ligands for the cryptic hotspot. Two series of elaborated ligands were included, as each simultaneously occupied two distinct hotspots. The details of the fragment elaboration campaign that yielded these series is detailed elsewhere (manuscript in preparation).

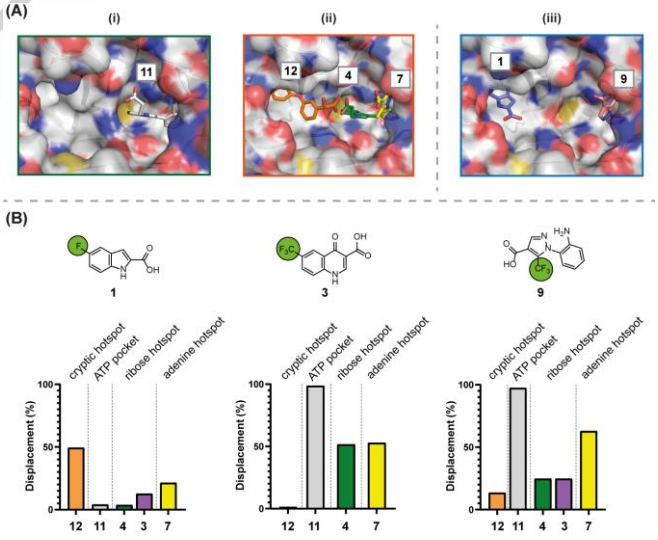


Figure 4. Validation of hotspot specificity of the ^{19}F reporter fragments using a set of structurally characterised ligands. (A) X-ray crystal structures showing the binding sites occupied by validation set ligands. (B) Percentage displacement bar charts from single concentration competitive ^{19}F CPMG experiments with the validation set ligands (n=1). Key experimental conditions for the competitive ^{19}F CPMG experiments with each reporter are as reported in **Table 1**. Ligands were used at 1 mM (**11**) or 2.5 mM (**12**, **4**, **7** and **3**) concentrations. PDB entries: 8QJ8 (**11**), 5O0D (**12**), 5O0C (**4**), 5O0A (**7**) 7YY3 (**1**), 8QIY (**9**).

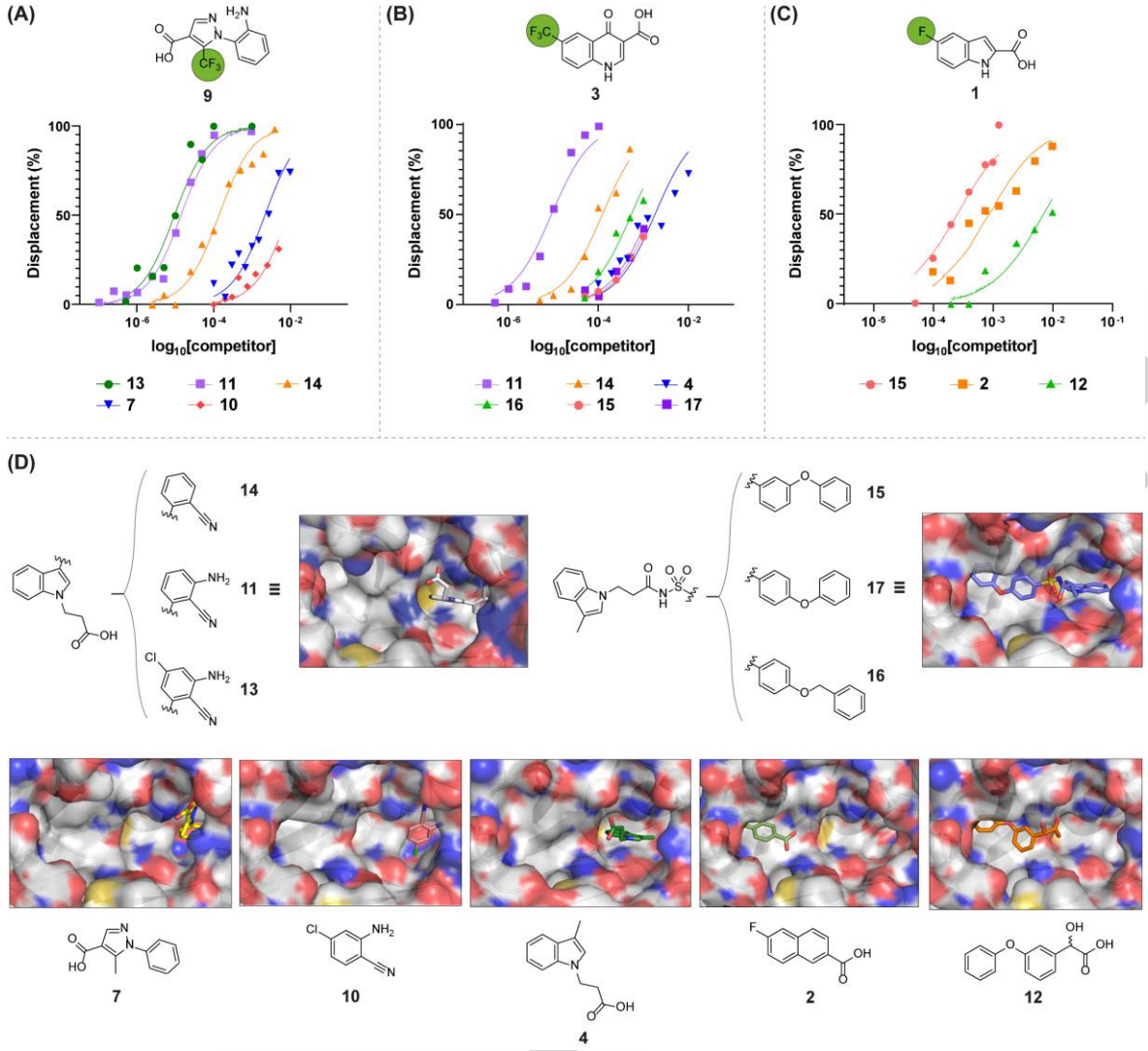


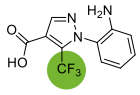
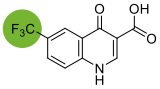
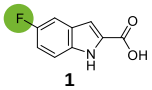
Figure 5. Hotspot directed ligand profiling of the adenine, ribose and cryptic hotspots by competitive ^{19}F LB-NMR. (A-C) Overlay of competitive ^{19}F LB-NMR experiments with **9** (A), **3** (B) and **1** (C) against a panel of ligands engaging the adenine, ribose and cryptic hotspots, respectively (n=1). (D) Chemical structures of the competing ligands used in (A-C) with associated X-ray crystal structures in complex with PPAT. Key experimental conditions for the competitive ^{19}F LB-NMR experiments with each reporter are as reported in **Table 1**. PDB entries: 8QJ8 (**11**), 8QIX (**17**), 5O0A (**7**), 7YY9 (**10**), 5O0C (**4**), 7YY4 (**2**), 5O0D (**12**).

13, **11** and **14** occupy both the ribose and adenine hotspots, while **15**, **16** and **17** engage both the ribose and cryptic hotspots (**Figure 5D**). For competition experiments, the weak binding limit of affinity quantitation is often set by the solubility limit of the competing ligands. For the experiments described herein, competing ligands were screened at concentrations up to their measured solubility limits. An IC_{50} value quantifying the reporter displacement by a competing ligand was derived from dose-response competition experiments by fitting the data to a standard dose-response model. In cases where the K_d of the ^{19}F reporter is known, the IC_{50} values can be converted into K_i values using the Nikolovska-Coleska relationship (**Equation S5**).[40] Where quantified, the IC_{50} and K_i values derived from the competition experiments with the suite of ^{19}F reporters are reported in **Table 2**.

The competitive profiling experiments performed with **3** (**Figure 5B**, **Table 2**) demonstrates a useful property of the competitive ^{19}F CPMG system. By using the IC_{50} values derived for ligands with known affinities engaging the ribose hotspot (**11** and **14**), the binding affinity of **3** could be estimated by internal calibration. The best agreement between the affinity values was given with $K_d \approx$ 150 μM for **3**. This estimated affinity for **3** was subsequently used to convert IC_{50} values to K_i values for other ligands.

It was found that there was a strong correlation (R^2 = 0.974) between K_i values determined by competitive ^{19}F LB-NMR and K_d values determined by ITC (**Figure S8**). The absolute and relative affinities of ligands in the series simultaneously occupying the adenine and ribose hotspots agreed well with the K_d values derived by ITC (**Figure 5A/B**, **Table 2**). Of the five fragment hits assessed by the competitive ^{19}F LB-NMR method, ITC had only been capable of quantifying the affinity of two of these, **7** and **2**, likely due to a generally low enthalpy output observed with fragments engaging the ATP binding site. IC_{50} and K_i values for both these fragments were measurable with the ^{19}F LB-NMR platform. A K_i = 1.4 mM for **7** was derived from profiling with adenine reporter **9**, ≈ 6-fold higher than the affinity determined by ITC (K_d = 0.24 mM) (**Figure 5A**, **Table 2**).

Table 2. IC_{50} and K_i values for competing ligands derived from the competitive ^{19}F LB-NMR experiments with reporters **9**, **3** and **1** outlined in **Figure 5**.[a]

Reporter	Competitor	K_d (μM) (ITC)	IC_{50} (μM) (^{19}F NMR)	K_i (μM) (^{19}F NMR)
 **9**	**13**	2.0 ± 0.5	9.4 ± 4.4	3.3 ± 2.8
	11	3.4 ± 0.9	13.8 ± 3.5	6.2 ± 2.3
	14	39.5 ± 7.5	138 ± 38	86 ± 24
	7	240 ± 14	2200 ± 680	1410 ± 430
 **3**	**11**	3.4 ± 0.9	9.3 ± 3.8	4.4 ± 1.5
	14	39.5 ± 7.5	124 ± 41	51 ± 17
	4	[b]	1830 ± 1100	820 ± 490
 **1**	**15**	[b]	249 ± 112	172 ± 77
	2	137 ± 10	860 ± 370	610 ± 260

[a] K_d values measured by ITC are included where available for comparison. For ^{19}F LB-NMR data, n=1. for ITC data, n≥1. Errors in IC_{50} values represent the 95% confidence interval (CI) for the fitting. Errors in K_i values represent a range defined by the upper and lower IC_{50} values from the 95% CI. Errors in K_d values represent the fitting error. IC_{50} and K_i values are only reported for ligands that achieved at least 70% displacement of the reporter. [b] K_d values for **4** and **15** are not reported as the ITC binding isotherms could not be reliably fit.

A K_i = 610 μM was measured for **2** by profiling with cryptic reporter **1**. This was ≈ 4-fold higher than the affinity determined by ITC (K_d = 137 μM) (**Figure 5C**, **Table 2**). The errors in K_i values determined by competitive ^{19}F LB-NMR were typically higher than for K_d values determined by ITC (**Table 2**).

Nonetheless, the competitive ^{19}F LB-NMR platform extended the range of quantifiable ligand binding interactions across *Mabs* PPAT hotspots beyond that of ITC. The binding affinity of fragment **4** was measured at the ribose hotspot by competition with **3** (**Figure 5B**, **Table 2**). Although **4** did not completely displace **3** in the concentration range tested, sufficient displacement was achieved to allow an IC_{50} value to be extracted from the dose-response curve, corresponding to a K_i = 820 μM. **15** occupies the ribose and cryptic hotspots simultaneously and was profiled against both **3** and **1**. **15** did not displace **3** to a great enough extent in the concentration range tested to quantify an IC_{50} value. However, its binding affinity at the cryptic hotspot was quantifiable using **1** (**Figure 5C**, **Table 2**). **1** was completely displaced by **15** and the affinity of the interaction was found to have K_i = 172 μM.

An advantage of the competitive ^{19}F LB-NMR platform over many direct methods is that it can provide reliable measures of relative ligand affinities when the absolute affinities cannot be determined or when the ligand concentration range is constrained, for example, by solubility. This is clearly demonstrated by data for the competition of adenine hotspot fragments **7** and **10** with **9** (**Figure 6A**). **7** was found to have a K_i = 1.4 mM, but **10** did not displace **9** to a great enough extent to reliably quantify an IC_{50} value. However, **7** achieved ≈ 50% displacement of **9** at ≈ 2 mM concentration, while **10** had only achieved ≈ 30% displacement of **9** at its top concentration of 5 mM. The X-ray crystal structures of **7** and **10** in complex with PPAT show the phenyl rings of both fragments occupy the same site and each forms a H-bond with the side-chain hydroxyl of Thr14, via the pyrazole nitrogen (**7**) and nitrile (**10**), respectively (**Figure 6B**). The aniline amine of **10** forms H-bonds with the backbone carbonyls of Tyr122 and Val125. The carboxylic acid of **7** forms direct H-bonds with the side-chain hydroxyl and backbone NH of Ser127, as well as water mediated H-bonds with the side-chain imidazole of His17 and side-chain hydroxyl of Ser128. As **7** engages the adenine sub-site with greater affinity than **10**, these polar interactions and additional space occupied by the pyrazole ring must contribute more affinity to **7** than the aniline amine contributes to **10**.

At the cryptic hotspot, the relative affinities of fragments **12** and **2** was determined by profiling with **1** (**Figure 6C**). **2** was found to have a K_i = 610 μM, but **12** did not displace **1** to a great enough extent to reliably quantify an IC_{50} value. While **2** achieved ≈ 50% displacement of **1** at ≈ 800 μM concentration, **12** had only achieved ≈ 50% displacement of **1** at its top concentration of 10 mM. Thus, the affinity of **2** at the cryptic hotspot can be estimated to be ≈ 10-fold greater than that of **12**. These relative affinities can be rationalised by examination of the X-ray crystal structures of both fragments in complex with PPAT (**Figure 6D**). The naphthalene core of **2** more completely occupies the cryptic pocket than the terminal phenyl of **12**, suggesting greater occupation of the cryptic pocket increases binding affinity. Furthermore, the additional interactions gained by the carboxylic acid of **12** are possibly outweighed by unfavourable solvent exposure of the linking phenyl ring, which superimposes with the carboxylic acid of **2**. That even very weak binding interactions (affinity > 5 mM) can be understood semi-quantitatively by comparison with other ligands demonstrates the sensitivity of the ^{19}F LB-NMR platform and its capacity to guide prioritisation of ligands engaging hotspots for subsequent elaboration.

In summary, where available, the K_i values derived by the competitive ^{19}F LB-NMR method agreed well with K_d values measured by ITC. Furthermore, several ligands whose affinities could not be measured by ITC were successfully quantified by the ^{19}F LB-NMR method. And where fragment affinities escaped absolute quantification by ^{19}F LB-NMR, their relative affinities could be confidently assigned. Successful application of the three ^{19}F reporter fragments, **1**, **3** and **9** in single shot and dose-response competition experiments established the capability of the ^{19}F CPMG platform to quantify and rank the affinities of ligands binding across the PPAT active site with hotspot specific resolution.

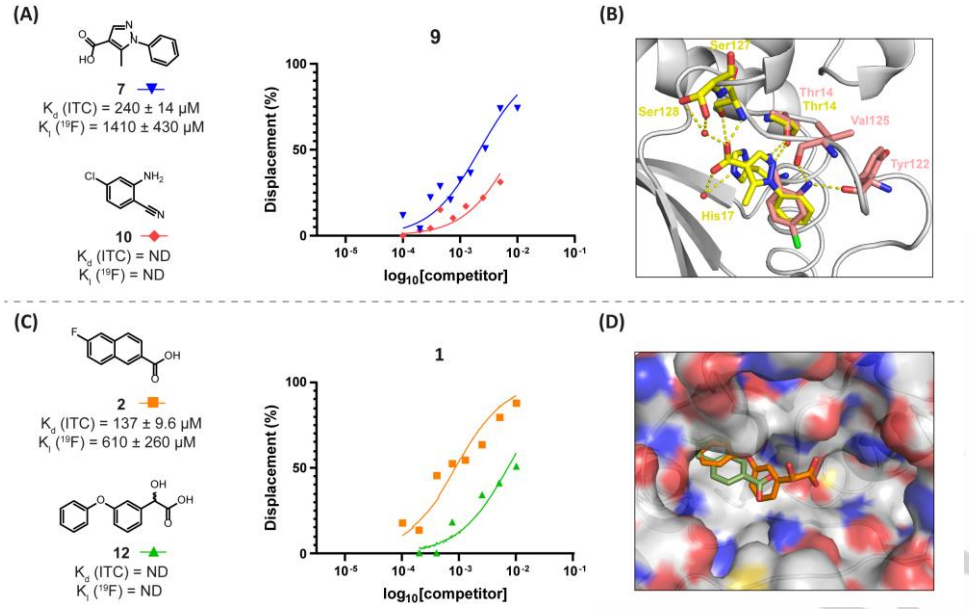


Figure 6. Overlay of competitive ^{19}F LB-NMR experiments with (A) **9** against adenine hotspot fragments **7** and **10** (n=1), and (C) **1** against cryptic hotspot fragments **2** and **12** (n=1). Overlay of X-ray crystal structures of (B) adenine hotspot fragments **7** (yellow) and **10** (pink), and (D) cryptic hotspot fragments **2** (green) and **12** (orange). ND - K_d and K_i were not quantifiable for **10**, and **12** by ITC nor competitive ^{19}F LB-NMR. Key experimental conditions for the competitive ^{19}F LB-NMR experiments with each reporter are as reported in **Table 1.** PDB entries: 5O0A (**7**), 7YY9 (**10**) 7YY4 (**2**), 5O0D (**12**).

Conclusion

Regarding the general applicability of the competitive 19F LB-NMR platform outlined herein, there are two important points to make. Firstly, the fragment hits from which the established *Mabs* PPAT ^{19}F reporters were derived had very different upfront suitabilities as ^{19}F reporters, (i) structurally characterised, fluorinated fragment hits (**1**), (ii) structurally uncharacterised, fluorinated fragment hits (**3**), and (iii) structurally characterised, non-fluorinated fragment hits (**9**). That an ^{19}F reporter was developed from fragments in each of these categories and successfully applied for quantitative hotspot mapping demonstrates that fragment hits from not only directed ^{19}F fragment screens but any format of fragment screen can be developed into good ^{19}F reporters. Secondly, ^{19}F reporters whose binding interactions are characterised to very different degrees can all be usefully applied in the competitive ^{19}F LB-NMR hotspot mapping method. In this sense, the three ^{19}F reporters employed herein for *Mabs* PPAT represented two distinct classes of ^{19}F reporter, (i) binding mode and affinity characterised (**1** and **9**), and (ii) binding mode and affinity uncharacterised (**3**). Due to their high level of characterisation, reporter fragments **1** and **9** could be applied directly for hotspot-specific quantitative profiling. Application of **3** for specific and quantitative hotspot mapping relied on indirect characterisation of its binding site and affinity using well characterised ligands. These examples demonstrate that the level of reporter characterisation necessary depends on the availability of tool compounds.

When designing fluorinated reporter fragments, key considerations are (i) the intensity of the ^{19}F NMR signal, and (ii) its sensitivity to protein binding. While trifluoromethyl motifs produce 3-fold higher intensity ^{19}F signals than mono-fluorinated substituents, their relaxation rates are typically less sensitive to binding. The lower ^{19}F signal intensity of mono-fluorinated fragments can be compenstaed by the acquisition of more scans, while the lower sensitivity of the trifluoromethyl group to binding can be offset by an increase in the protein concentration.

By following the workflows outlined herein, others applying this method in the future will be able to convert fragment hits, with or without a pre-existing fluorinated motif and quantified binding affinities, into high quality reporters of ligand binding hotspots. With the addition of structural information on the distances and relative orientations of these hotspots, from X-ray crystallography for example, synthetic strategies to combine them can be explored. In principle, the fragment-based competitive ^{19}F LB-NMR workflow can be applied to fragment hits against any target that is responsive to LB-NMR. In conclusion, this study has demonstrated that the minimal molecular recognition that is intrinsic to fragments can be exploited for high precision spatial mapping of ligand binding events, with simultaneous quantification of the binding interaction using competitive ^{19}F LB-NMR.

Supporting Information

The authors have cited additional references within the Supporting Information.[28, 37, 40-52]

Acknowledgements

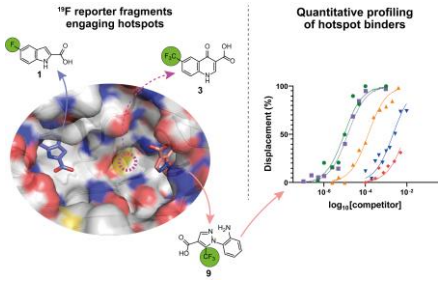
This work was supported by the Cystic Fibrosis Trust (IH001). We are also grateful for funding from the Cambridge Philosophical Society (W.J.M.C). T.O was funded by the Bill and Melinda Gates foundation (OPP1024021 awarded to C.A and T.L.B). We would also like to thank the NMR facility staff at the Yusuf Hamied

Department of Chemistry, Duncan Howe and Andrew Mason for their expert technical assistance.

Keywords: Drug Discovery • Enzymes • Fluorine • Fragment-based discovery • NMR spectroscopy

[1] D. E. Scott, A. G. Coyne, S. A. Hudson, C. Abell, *Biochemistry* **2012**, *51*, 4990–5003.
[2] C. W. Murray, D. C. Rees, *Nat. Chem.* **2009**, *1*, 187–192.
[3] D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discov.* **2016**, *15*, 605–619.
[4] C. for D. E. and Research, *FDA* **2023**.
[5] G. Bollag, J. Tsai, J. Zhang, C. Zhang, P. Ibrahim, K. Nolop, P. Hirth, *Nat. Rev. Drug Discov.* **2012**, *11*, 873–886.
[6] A. Mullard, *Nat. Rev. Drug Discov.* **2016**, *15*, 147–149.
[7] H. Gelderblom, M. van de Sande, *Future Oncol.* **2020**, *16*, 2345–2356.
[8] C. W. Murray, D. R. Newell, P. Angibaud, *MedChemComm* **2019**, *10*, 1509–1511.
[9] F. Skoulidis, B. T. Li, G. K. Dy, T. J. Price, G. S. Falchook, J. Wolf, A. Italiano, M. Schuler, H. Borghaei, F. Barlesi, T. Kato, A. Curioni-Fontecedro, A. Sacher, A. Spira, S. S. Ramalingam, T. Takahashi, B. Besse, A. Anderson, A. Ang, Q. Tran, O. Mather, H. Henary, G. Ngarmchamnanrith, G. Friberg, V. Velcheti, R. Govindan, *N. Engl. J. Med.* **2021**, *384*, 2371–2381.
[10] T. P. Hughes, M. J. Mauro, J. E. Cortes, H. Minami, D. Rea, D. J. DeAngelo, M. Breccia, Y.-T. Goh, M. Talpaz, A. Hochhaus, P. le Coutre, O. Ottmann, M. C. Heinrich, J. L. Steegmann, M. W. N. Deininger, J. J. W. M. Janssen, F.-X. Mahon, Y. Minami, D. Yeung, D. M. Ross, M. S. Tallman, J. H. Park, B. J. Druker, D. Hynds, Y. Duan, C. Meille, F. Hourcade-Potelleret, K. G. Vanasse, F. Lang, D.-W. Kim, *N. Engl. J. Med.* **2019**, *381*, 2315–2326.
[11] C. J. Radoux, T. S. G. Olsson, W. R. Pitt, C. R. Groom, T. L. Blundell, *J. Med. Chem.* **2016**, *59*, 4314–4325.
[12] T. Clackson, J. A. Wells, *Science* **1995**, *267*, 383–386.
[13] M. O'Reilly, A. Cleasby, T. G. Davies, R. J. Hall, R. F. Ludlow, C. W. Murray, D. Tisi, H. Jhoti, *Drug Discov. Today* **2019**, *24*, 1081–1086.
[14] D. J. Wood, J. D. Lopez-Fernandez, L. E. Knight, I. Al-Khawaldeh, C. Gai, S. Lin, M. P. Martin, D. C. Miller, C. Cano, J. A. Endicott, I. R. Hardcastle, M. E. M. Noble, M. J. Waring, *J. Med. Chem.* **2019**, *62*, 3741–3752.
[15] C. Dalvit, A. D. Gossert, J. Coutant, M. Piotto, *Magn. Reson. Chem.* **2011**, *49*, 199–202.
[16] T. B. Almeida, S. Panova, R. Walser, *SLAS Discov.* **2021**, *26*, 1020–1028.
[17] M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784–1788.
[18] H. Y. Carr, E. M. Purcell, *Phys. Rev.* **1954**, *94*, 630–638.
[19] S. Meiboom, D. Gill, *Rev. Sci. Instrum.* **2004**, *29*, 688–691.
[20] C. Dalvit, A. Parent, F. Vallée, M. Mathieu, A. Rak, *ChemMedChem* **2019**, *14*, 1115–1127.
[21] M. Liu, A. Mirza, P. C. McAndrew, A. Thapaliya, O. A. Pierrat, M. Stubbs, T. Hahner, N. E. A. Chessum, P. Innocenti, J. Caldwell, M. D. Cheeseman, B. R. Bellenie, R. L. M. van Montfort, G. K. Newton, R. Burke, I. Collins, S. Hoelder, *J. Med. Chem.* **2023**, *66*, 10617–10627.
[22] J. Angulo, P. M. Enríquez-Navas, P. M. Nieto, *Chem. – Eur. J.* **2010**, *16*, 7803–7812.
[23] S. Monaco, J. Angulo, M. Wallace, *J. Am. Chem. Soc.* **2023**, *145*, 16391–16397.
[24] C. Dalvit, P. E. Fagerness, D. T. A. Hadden, R. W. Sarver, B. J. Stockman, *J. Am. Chem. Soc.* **2003**, *125*, 7696–7703.
[25] A. Vulpetti, C. Dalvit, *ChemMedChem* **2013**, *8*, 2057–2069.
[26] A. Lingel, A. Vulpetti, T. Reinsperger, A. Proudfoot, R. Denay, A. Frommlet, C. Henry, U. Hommel, A. D. Gossert, B. Luy, A. O. Frank, *Angew. Chem. Int. Ed.* **2020**, *59*, 14809–14817.
[27] C. R. Buchholz, W. C. K. Pomerantz, *RSC Chem. Biol.* **2021**, *2*, 1312–1330.
[28] G. V. de Castro, A. Ciulli, *Chem. Commun.* **2019**, *55*, 1482–1485.
[29] G. V. de Castro, A. Ciulli, *RSC Med. Chem.* **2021**, *12*, 1765–1770.
[30] C. Dalvit, I. Gmür, P. Rößler, A. D. Gossert, *Prog. Nucl. Magn. Reson. Spectrosc.* **2023**, *138–139*, 52–69.
[31] C. Dalvit, A. Vulpetti, *J. Med. Chem.* **2019**, *62*, 2218–2244.
[32] C. Dalvit, M. Piotto, *Magn. Reson. Chem.* **2017**, *55*, 106–114.
[33] C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 243–271.
[34] C. Spry, K. Kirk, K. J. Saliba, *FEMS Microbiol. Rev.* **2008**, *32*, 56–106.
[35] H. S. Butman, T. J. Kotzé, C. S. Dowd, E. Strauss, *Front. Cell. Infect. Microbiol.* **2020**, *10*, 605662.
[36] J. El Bakali, M. Blaszczyk, J. C. Evans, J. A. Boland, W. J. McCarthy, I. Fathoni, M. V. B. Dias, E. O. Johnson, A. G. Coyne, V. Mizrahi, T. L. Blundell, C. Abell, C. Spry, *Angew. Chem. Int. Ed.* **2023**, *62*, e202300221.
[37] S. E. Thomas, W. J. McCarthy, J. El Bakali, K. P. Brown, S. Y. Kim, M. Blaszczyk, V. Mendes, C. Abell, R. A. Floto, A. G. Coyne, T. L. Blundell, *Front. Mol. Biosci.* **2022**, *9,* 880432.
[38] T. Krojer, J. S. Fraser, F. von Delft, *Curr. Opin. Struct. Biol.* **2020**, *65*, 209–216.
[39] R. Buratto, D. Mammoli, E. Chiarparin, G. Williams, G. Bodenhausen, *Angew. Chem. Int. Ed.* **2014**, *53*, 11376–11380.
[40] Z. Nikolovska-Coleska, R. Wang, X. Fang, H. Pan, Y. Tomita, P. Li, P. P. Roller, K. Krajewski, N. G. Saito, J. A. Stuckey, S. Wang, *Anal. Biochem.* **2004**, *332*, 261–273.
[41] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 293–302.
[42] W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132.
[43] P. R. Evans, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 282–292.
[44] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. D Biol. Crystallogr.* **2013**, *69*, 1204–1214.
[45] S. French, K. Wilson, *Acta Crystallogr. A* **1978**, *34*, 517–525.
[46] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242.
[47] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
[48] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367.
[49] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221.
[50] P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
[51] L. Dubois, F. C. Acher, I. McCort-Tranchepain, *Synlett* **2012**, *23*, 791–795.
[52] T. W. Greulich, C. G. Daniliuc, A. Studer, *Org. Lett.* **2015**, *17*, 254–257.

Entry for the Table of Contents



Fragment screening identified three ligand binding hotspots across the active site of 4'-phosphopantetheine adenylyltransferase from *Mycobacterium abscessus*. Fragment sized fluorinated reporters were developed for each hotspot and used in competitive ^{19}F LB-NMR experiments to quantify hotspot-specific ligand binding affinities across three orders of magnitude, from low mM to low µM.