An experimental toolbox for structure-based hit discovery for *P. aeruginosa* FabF, a promising target for antibiotics

Ludvik Olai Espeland,[$[a, b]] Dr. Charis Georgiou, [$[a]] Dr. Raphael Klein, [$[a,c]] Hemalatha Bhukya,[[d]] Prof. Dr. Bengt Erik Haug,[[b]] Assoc. Prof. Dr. Jarl Underhaug,[[b]] Dr. Prathama S. Mainkar,[[d]] and Prof. Dr. Ruth Brenk*[[a]]

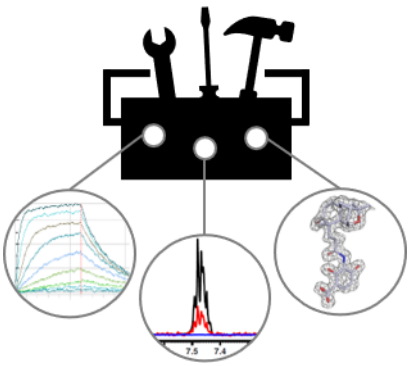
[a] L. O. Espeland, Dr. C. Georgiou, Dr. R. Klein, Prof. Dr. R. Brenk
Department of Biomedicine
University of Bergen
Jonas Lies Vei 91 5020 Bergen (Norway)
E-mail: ruth.brenk@uib.no

[b] L. O. Espeland, Prof. Dr. B. E. Haug, Assoc. Prof. Dr. J. Underhaug
Department of Chemistry
University of Bergen
Allégaten 41 5007 Bergen (Norway)

[c] Dr. R. Klein
Institute of Pharmacy and Biochemistry
Johannes Gutenberg University
Staudingerweg 5, 55128 Mainz (Germany)

[d] Dr. Prathama S. Mainkar, Hemalatha Bhukya
Department of Organic Synthesis & Process Chemistry
CSIR-Indian Institute of Chemical Technology
Tarnaka, Hyderabad 500007, India

$ These authors have contributed equally to the work.



Abstract: FabF (3-oxoacyl-[acyl-carrier-protein] synthase 2), which catalyses the rate limiting condensation reaction in the fatty acid synthesis II pathway, is an attractive target for new antibiotics. Here, we focus on FabF from *P. aeruginosa* (*Pa*FabF) as antibiotics against this pathogen are urgently needed. To facilitate exploration of this target we have set up an experimental toolbox consisting of binding assays using bio-layer interferometry as well as saturation transfer difference (STD) and WaterLOGSY NMR in addition to robust conditions for structure determination. The suitability of the toolbox to support structure-based design of FabF inhibitors was demonstrated through the validation of hits obtained from virtual screening. Screening of our in-house library of almost 5 million compounds resulted in 6 compounds for which binding into the malonyl-binding site of FabF was shown. For one of the hits, the crystal structure in complex with *Pa*FabF was determined. Based on the obtained binding mode, analogues were designed and synthesised, but affinity could not be improved. This work has laid the foundation for structure-based exploration of *Pa*FabF.

Introduction

New antibiotics are urgently needed to maintain the high standard of living that we have got accustomed to.[1] In response to this, the World Health Organization has recently published a list of bacteria for which antibiotics are in high demand and urged that research and drug discovery efforts are directed toward these bacteria.[2] However, most of the compounds in clinical development are not fit for purpose as they are derivatives of already known antibiotics and therefore have cross-resistance with existing agents. Accordingly, there is a high need for new classes of antibiotics without pre-existing cross-resistance. To achieve this, we need more knowledge on new targets and new chemical scaffolds.[3]

A possible source for new targets for antibiotics is the fatty acid synthesis (FAS II) pathway (Figure 1a).[4] In this pathway, fatty acid synthesis is carried out by a series of monofunctional enzymes which are highly conserved among microbial pathogens. In contrast, mammals use the FAS I pathway consisting of the large multifunctional enzyme fatty acid synthase.[5] The fatty acid synthase and the enzymes of the FAS II pathway differ significantly enabling the development of selective inhibitors. While the FASII pathway is also present in mitochondria of human cells, its role is not entirely clear and short-term inhibition by antibiotics is likely to be tolerated.[6] Genes coding for enzymes in the FAS II pathway have been found to be essential in several genetic screens.[7–11] Further validation of this pathway as a target for antibiotics comes from the drug isoniazid which is in clinical use against tuberculosis. The main target for isoniazid is the enoyl-ACP reductase FabI, which is part of the FAS II pathway (Figure 1a).[4] Multiple enzymes in this pathway were predicted to contain druggable binding sites rendering them attractive targets for antibiotic drug discovery.[12]

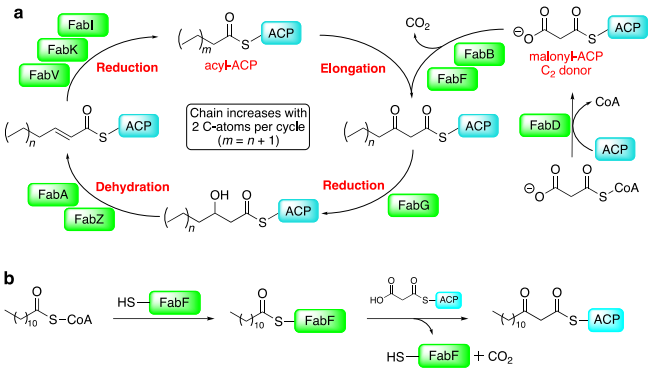


Figure 1. Fatty acid synthesis in bacteria. a) Schematic overview of the elongation part of the FAS II pathway. b) Condensation reaction catalysed by FabF. (ACP: acetyl carrier protein).

Here, we focus on 3-oxoacyl-[acyl-carrier-protein] synthase 2 (also called beta-ketoacyl-ACP synthase 2 with FabF, Kas II, and KasB as commonly used short names). FabF catalyses the rate limiting condensation of malonyl-ACP (acyl carrier protein) with acyl-ACP (Figure 1b),[4] facilitated the by the catalytic Cys-His-His triad. FabB catalyses the same reaction (Figure 1a) but differs in substrate specificity for the fatty acid chain. FabF has been validated in Gram-positive bacteria as the target for the natural products platensimycin, platencin and fasamycin A and B (Figure 2).[13–15] Platensimycin, although inactive against wild-type Gram-negative bacteria, is active against efflux-negative *E. coli*.[13] Similarly, the fasamycins are active against membrane-permeabilized *E. coli* but not the unaltered strain.[14] These observations suggest that the lack of activity of these compounds against Gram-negative bacteria is not caused by a lack of essentiality of the target but owing to the inability of the ligands to access it. Further target validation for FabF comes from the natural products cerulenin and thiolactomycin (Figure 2). Cerulenin is a covalent inhibitor forming a bond with the catalytic cysteine.[16–19] Both natural products are effective against bacteria and thiolactomycin is also efficacious in an animal model, however both lack selectivity and potency for clinical use.[4]

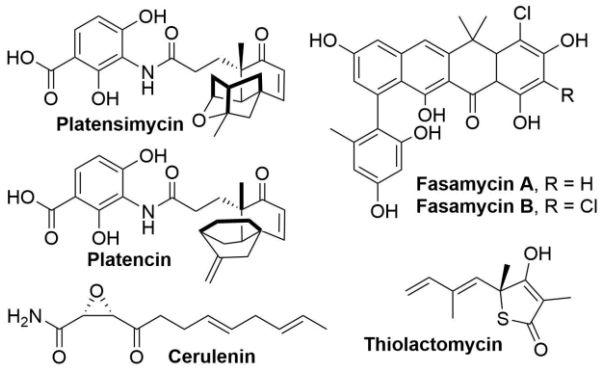


Figure 2. Natural product FabF inhibitors.

The crystal structures of FabF and the related FabB from various bacteria have been determined revealing the architecture of the binding site. The active site contains two sub-pockets: one for accommodating the growing fatty acid chain and one for binding the C_2-donor malonyl-ACP. The inhibitors platensimycin, platencin, and thiolactomycin all bind into the malonyl-ACP pocket whereas cerulenin binds into the fatty acid channel.[13,20–22] However, platensimycin and platencin do not bind strongly to the w. t. enzyme, but only to the lauroyl-CoA FabF intermediate (Figure 1b) and to intermediate-mimicking FabF variants in which the active site Cys has been changed to either Gln or Ala.[20,23] The lack of binding to the w. t. enzyme is probably caused by repulsion between the carboxylate group of these compounds and the active site nucleophilic Cys residue (Figure 3). The benzoic acid moiety of platensimycin is deeply bound in the malonyl pocket and forms hydrogen bonds with the catalytic histidines and edge-face interactions with Phe400/400 (*E. coli* FabF (*Ec*FabF)/*P. aeruginosa* FabF (*Pa*FabF) numbering). The rather hydrophobic aliphatic ring system extends towards the opening of the pockets and is held in place by hydrogen bonds with Thr270/271 and Ala309/310. In addition, the amide linker forms hydrogen bonds to Thr307/308.

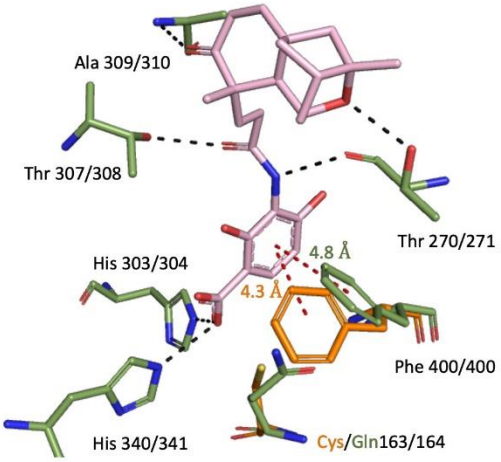


Figure 3. Alignment of apo w. t. *Ec*FabF (orange sticks - PDB ID: 2GFW, for clarity only Phe400 is shown) and *Ec*FabF C163Q (green sticks - PDB ID: 2GFX) in complex with platensimycin (pink sticks). Hydrogen bonds are indicated as black dashed lines and aromatic interactions as red dashed lines. For easier comparison to *Pa*FabF all residues are labelled as *Ec*FabF/*Pa*FabF. Compared to the apo structure, Phe400 is rotated in the holo structure to create space for the ligand to bind.

Platensimycin is the most potent FabF inhibitor known to date. Using the radioactively labelled compound in displacement assays, it was determined that platensimycin binds with an IC_{50} of 19 nM to the lauroyl-CoA FabF intermediate while binding to the w. t. enzyme is much weaker.[23] Total syntheses of platencin[24,25] and platensimycin[26,27] have been carried out, and efforts have been undertaken to increase the antibiotic activity through synthesis of analogues. For platensimycin this includes the modifications of the benzoic acid head group[28,29] and the tetracyclic cage.[30–34] While some of these compounds showed promising activity against Gram-positive bacteria, for most compounds the potency activity against Gram-negative bacteria was not reported. In addition, poor pharmacokinetic properties were also a concern.

To further assess FabF as a target, there is a need for more readily modifiable ligands, ideally also based on diverse scaffolds. Only limited efforts towards obtaining such compounds have been published. Using the crystal structure of *Ec*FabB, Zheng et al.

have conducted a virtual screening which resulted in the identification of a series of benzoxazolinones where the most potent compound had an IC_{50} of 235 µM against *Ec*FabB and 387 µM against *S. aureus* FabF.[35] Further, N-acetylated sulfonamides have been reported as the result of a phenotypic screen and found to be highly specific for *C. trachomatis* FabF Albeit, their binding mode remains elusive.[35]

Here, we report on our efforts toward establishing an experimental toolbox for structure-based drug design for FabF from *P. aeruginosa* to drive forward the exploration of FabF as a drug target for Gram-negative bacteria. The target organism was chosen as there is a particular high need for new antibiotics to fight infections caused by carbapenem-resistant *P. aeruginosa*.[2] Starting from our previous structural experience with *Pa*FabF[36] we have established robust crystallization conditions that allow for soaking with ligands. Further, we have developed a binding assay using bio-layer interferometry (BLI). In addition, we have set up saturation transfer difference (STD) and WaterLOGSY NMR assays for binding and competition experiments to confirm active site binding. We demonstrate the suitability of this toolbox for structure-based ligand discovery through the validation of hits obtained from virtual screening.

Results and Discussion

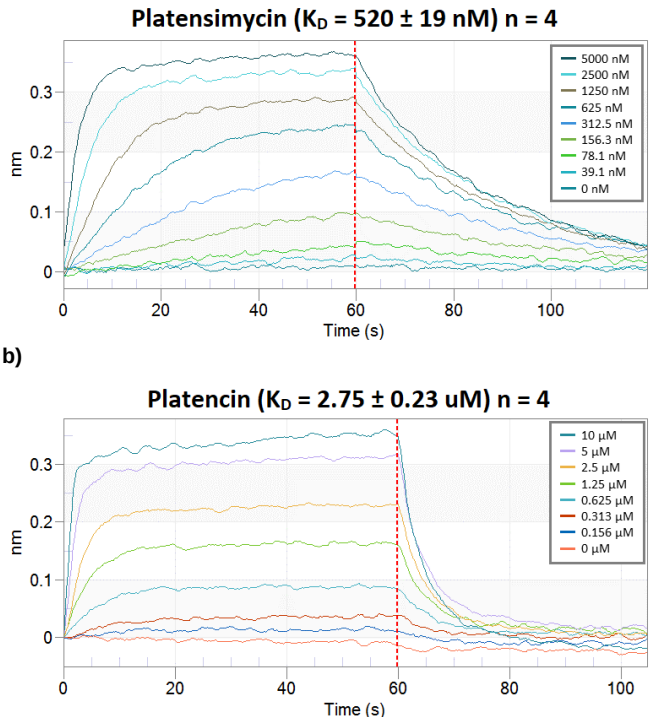
As *Pa*FabF is a promising target for new antibiotics, we aimed at establishing an experimental toolbox that would enable the structure-based design of new ligands. This toolbox should contain binding assays as well as robust conditions for soaking crystals to allow for a thorough assessment of potential ligands.

*Pa*FabF binding assay based on BLI

In order to determine ligand binding to *Pa*FabF and the intermediate-mimicking FabF variants C164Q and C164A, a BLI assay was developed. For this purpose, protein constructs containing an AviTag[37] that can be biotinylated using the biotin ligase BirA[38] were designed and purified as described earlier.[36] Avi-tagged *Pa*FabF C164Q was prone to precipitation and thus not accessible for BLI experiments The remaining biotinylated *Pa*FabF variants were immobilized on Super Streptavidin (SSA) biosensors. Binding of ligands at different concentrations was recorded as wavelength shift of the interference pattern of white light reflected from a layer of immobilized protein on the biosensor tip, and an internal reference layer.[39] At concentrations up to 5 and 50 µM respectively, no binding for platensimycin and platencin to the w. t. enzyme was observed. In contrast, binding constants of 0.52 and 2.75 µM, respectively were determined for *Pa*FabFC164A (Figure 4 and Figure S5 in supplementary material).

The binding data obtained for the BLI assay is overall in agreement with what has been reported previously when measured against *Ec*FabF. In an assay based on displacement of a radioactive platensimycin derivative from the lauroyl-CoA *Ec*FabF intermediate, IC_{50} values of 19 nM and 113 nM were reported for platensimycin and platencin, respectively.[15,23] Using the same assay, no IC_{50} values could be determined for platensimycin binding to w.t. *Ec*FabF.[23] While the binding constants for platensimycin and platencin binding to *Pa*FabF C164A were higher than those for the *Ec*FabF construct, in both cases they differed about 5-fold. Using an enzymatic elongation assay with *Ec*FabF, the IC_{50} values for platensimycin and platencin were determined as 0.29 and 4.9 µM, respectively.[15] The latter values are close to the K_D values for *Pa*FabF C164A obtained using BLI.

a)



b)

Figure 4. BLI sensorgram for a) platensimycin and b) platencin binding to *Pa*FabFC164A. The dashed rad line indicates the start of the dissociation interval.

NMR assay to determine ligand binding

To assess the feasibility of using nuclear Overhauser effect (NOE) based NMR experiments to detect binding between FabF and ligands, saturation transfer difference (STD) spectra were obtained for platensimycin and platencin in the presence of *Pa*FabF C164A in a 10:1 ratio. Attenuation of the compound signals was observed for platencin, but not for the stronger binding platensimycin (Figure 5). The absence of signals for platensimycin is probably due to its slow off-rate causing the high binding affinity (0.52 µM) and resulting in a too slow exchange between free and bound ligand to be detectable.[40] Nevertheless, the data demonstrates that STD NMR measurements can be used to detect ligands binding with medium affinity to *Pa*FabFC164A.

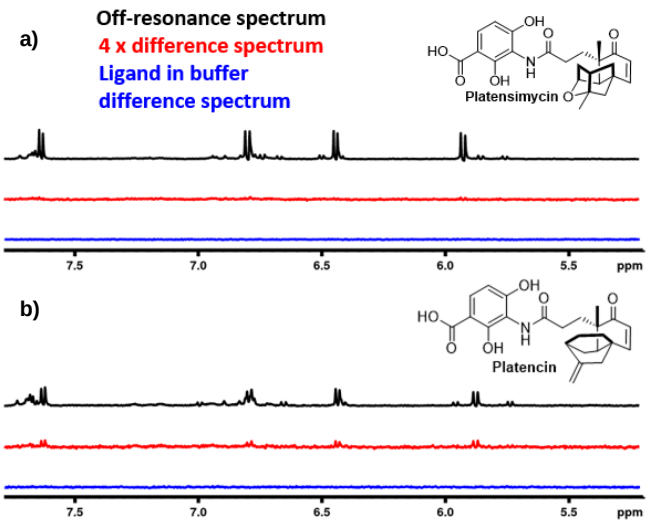


Figure 5. Stacked STD NMR spectra displaying the signals from the aromatic and enone protons. Off-resonance (black), difference spectrum scaled x 4 (red) and difference spectrum for ligand in pure buffer (blue) obtained with a) platensimycin and b) platencin in 10:1 ratio with *Pa*FabFC164A.

Robust soaking conditions for *Pa*FabF

Starting from our previously established crystallization conditions for *Pa*FabF,[36] protein and precipitant concentration as well as the drop-to-buffer ratio were optimized to allow for reliable crystallization and soaking of ligands. As a proof of concept, *Pa*FabF C164Q was soaked with platensimycin and the structure of the complex was determined at a resolution of 1.78 Å (Table S1). As expected, the ligand adopted the same binding modes as observed before for *Ec*FabF (Figure S1a). Compared to the apo structure, Phe400 was rotated to create space to bind the ligands (Figure S1b), as was also observed earlier for *Ec*FabF (Figure 3).

The experimental toolbox put into action

Having established binding assays for *Pa*FabF and robust soaking conditions, the suitability of the experimental toolbox to support structure-based design of FabF inhibitors was subsequently evaluated. To this end, a virtual screening for FabF inhibitors was conducted and binding of the shortlisted compounds was determined using the assays of the toolbox.

Virtual Screening for FabF ligands binding into the malonyl pocket

For virtual screening, the malonyl binding site of FabF was targeted, as most known inhibitors bind into this pocket. As a first step, the binding sites of FabF orthologs were analysed. Alignment of the available crystal structures in the Protein Data Bank (PDB) revealed a rigid binding site with only Phe400 adopting two different rotamers (Figure 3). As mentioned above, in the apo wild type structure, Phe400 blocks the malonyl binding site while upon binding of a fatty acid or inhibitors like cerulenin or platensimycin it rotates into the open conformation. The open conformation is also present in the *Pa*FabF C164Q variant mimicking the acyl-enzyme intermediate (PDB-ID: 4JB6) which was therefore used as template for virtual screening.

A pharmacophore hypothesis for malonyl binding site ligands was generated based on the binding modes of known inhibitors. Several structures of FabF in complex with platensimycin or derivatives thereof, cerulein and a fatty acid were available in the public domain. To increase the diversity of the ligands, the structure of thiolactomycin in complex with the homologous protein FabB was also taken into account. Analysis of these structures suggested that hydrogen-bond interactions to His304 and His341 are crucial for binding as these were formed by all ligands studied. Thus, these interactions were included as mandatory in a pharmacophore model (Figure 6). Further, the backbone amide of Thr271 and the side chain hydroxyl group of Thr308 were found to form hydrogen bonds with platensimycin and its derivatives, and Phe400 to form edge-to-face interactions with the phenyl rings of these ligands (Figure 3). These interactions were included as optional features in the pharmacophore hypothesis but only two of them had to be fulfilled to satisfy the pharmacophore requirements.

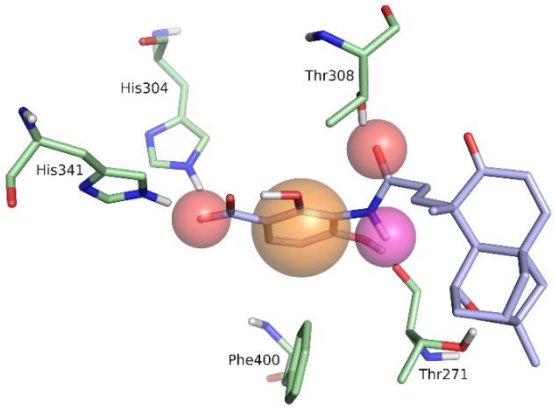


Figure 6. Pharmacophore model for FabF inhibitors binding into the malonlyl-pocket. The interactions with the His residues were set to mandatory. In addition, two of the interactions to the Thr and Phe residues was required.

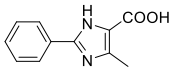
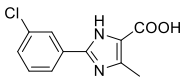
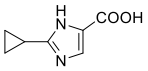
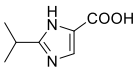
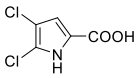
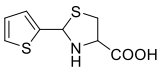
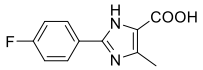
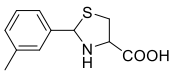
A hierarchical approach was adopted for virtual screening. First, an updated version of our in-house database of around 5 million commercially available compounds[41] was filtered for lead-like molecules fulfilling the above described pharmacophore. The resulting 66053 compounds were subsequently docked into the FabF binding site leading to binding modes for 48064 compounds. Filtering the rigid binding poses again with the pharmacophore resulted in 13244 compounds. Those were ranked by their calculated ligand efficiencies and visually inspected. Finally, 8 compounds were selected for hit validation (Table 1).

Testing of virtual screening hits

The 8 purchased virtual screening hits were tested as singletons at a concentration of 500 µM with STD and WaterLOGSY NMR in the presence of PaFabF C164A. Six compounds (**1**, **2**, and **5**-**8**, Table 1, Figure 7, Figure S2, and Figure S3) showed binding and were progressed to competition experiments. For all of them, binding to *Pa*FabF C164A could be outcompeted by addition of platensimycin (Figure 7b and Figure S3). This indicates that these compounds bind like platensimycin into the malonyl binding site as intended in the virtual screening.

All compounds that showed binding in the NMR experiments were advanced to BLI experiments. The compounds were tested at concentrations ranging from 16 µM to 2 mM. At lower concentrations, no binding and at higher concentrations binding to both protein-loaded and biocytin-blocked reference sensors was observed, prohibiting the determination of binding constants (data not shown).

Table 1. Compounds shortlisted for testing after virtual screening. All compounds were assessed for binding to *Pa*FabF C164A. Competition experiments were carried out with platensimycin as competing ligand. ✓ indicates that binding or out-competing of binding with platensimycin was observed.

#	Structure	STD	WaterLOGSY	Competition
1		✓	✓	✓
2		✓	✓	✓
3		✗	✗	
4		✗	✗	
5		✓	✓	✓
6		✓	✓	✓
7		✓	✓	✓
8		✓	✓	✓

Next, we attempted to determine the binding modes of the compounds and this was successful for one of the hits (Figure 8). The structure was obtained with re-synthesised material of **6** (compound **6a** in Table 2) as we did not have enough of the purchased material left over from the binding studies. The purchased material was a mixture of all four possible stereoisomers. However, in the virtual, only the (4*R*)-isomers scored well. Therefore, this stereo centre was fixed during synthesis while for the other stereo centre both configurations were obtained (Scheme 1). The structure of the complex *Pa*FabF C164Q was determined at 1.78 Å resolution (Table S1) with the ligand clearly visible in the difference electron density map (Figure 8a). Based on the electron density it was not possible to assign the stereochemistry of the compound as density around the undefined stereo centre was lacking. However, it appeared that the (2*S*,4*R*) diastereoisomer was fitting the density better than the (2*R*,4*R*) diastereoisomer. In the determined binding mode, the ligand forms hydrogen bonds to the catalytic histidines and the backbone of Phe399 (Figure 8b). While the interactions with the histidines are close to those predicted by docking, the thiazolidine moiety is clearly placed in a different position (Figure 8c). Accordingly, the bound ligand does not fulfil the postulated pharmacophore hypothesis.

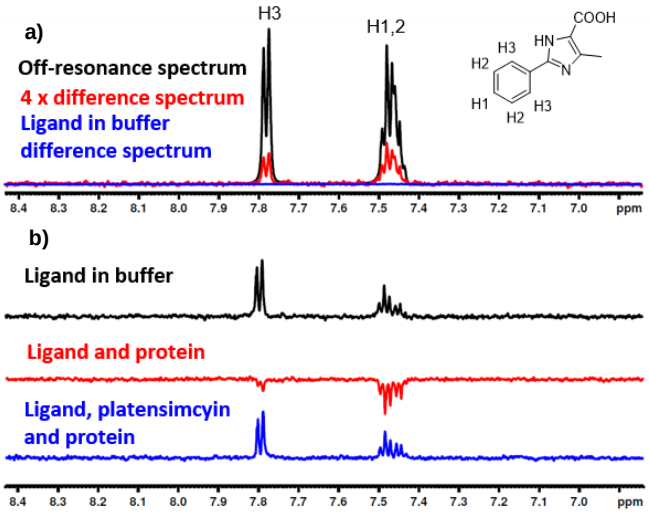


Figure 7. Example from NMR binding assays with a virtual screening hit. a) Overlayed spectra of **1** from STD experiments. b) Stacked WaterLOGSY spectra (non-interacting spectra pointing upwards). Binding of **1** can be outcompeted with platensimycin.

Hit expansion

Two thiazolidines were among the screening hits (**6** and **8**, Table 1) and for one of them the binding mode could be determined (Figure 8). Therefore, we chose this compound class for hit expansion. In an effort to increase the affinities of the compounds, analogues were designed to explore further interactions in the pocket occupied by the thiophene ring and to fill available space in the binding site adjacent to this ring (Figure 8d, Table 2). As docking had suggested that only the (*4R*) enantiomer could bind, this centre was fixed during the design.

A general methodology for condensation of L-cysteine hydrochloride with aromatic aldehydes in presence of sodium hydroxide in ethanol gave the desired thiazolidines **6a-h** (Scheme 1, Table 2). In addition to providing diastereoisomers of hit molecules **6a** and **6f** with a fixed R-configuration at C-4, six more analogues of the hits were synthesized. The compounds were obtained in 91-97% yields. The synthesis yielded a mixture of diastereomers where aldehydes with electron-rich rings favoured formation of the (2*S*,4*R*) isomer over the (2*R*,4*R*) isomer (10:1 ratio for **6d**, Table 2) based on comparison of chemical shifts with those published for similar structures.[42] In most of the cases the diastereomeric ratio was found to be between 1:1 to 4:1. Separation of the individual diastereomers was not carried out at this stage and all compounds were tested as mixtures.

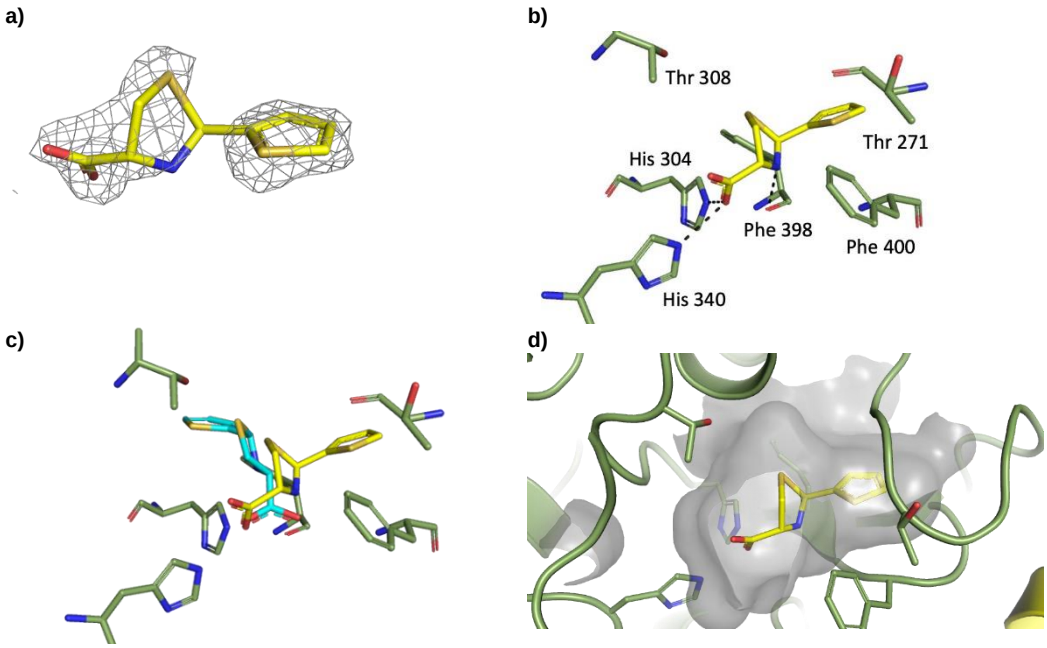
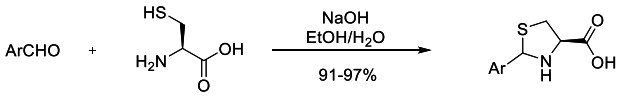


Figure 8 Binding of compound **6a** to *Pa*FabF C164Q (PDB ID 7OC0). a) Fo-Fc map of **6a** binding to *Pa*FabF C164Q, contoured at 3.0 sigma. b) Interactions between **6a** (yellow) and *Pa*FabF C164Q (green). Hydrogen-bonds are indicated as dashed lines. c) Alignment between the docked pose (cyan sticks) and the crystallographically determined binding mode of **6a** (yellow sticks). d) Solvent accessible surface of the *Pa*FabF C164Q binding site indicating available space next to the thiophen group of the ligand.



Scheme 1. Synthesis of substituted thiazolidines through a cyclodehydration reaction between L-cysteine and aryl aldehydes.

The analogues were tested for binding to *Pa*FabF C164A using binding assays and *Pa*FabF C164Q using X-ray crystallography. Two of the compounds were not soluble under the assay conditions (**6d** and **e**) and **6h** decomposed during shipping, thus, these were not further considered. For the remaining compounds, some attenuation of the WaterLOGSY signal was already observed for ligand in pure buffer indicating compound aggregation (Figure 9 and Figure S4). All tested compounds showed binding to *Pa*FabF C164A using WaterLOGSY NMR (Table 2). However, competition experiments did not lead to a full reduction in WaterLOGSY signal, suggesting that some of the signal stemmed from non-active site interactions. To further elucidate the binding mode of these compounds, experiments with w. t. *Pa*FabF were also carried out. For acidic compounds binding close to the catalytic histidines, one would expect no binding to the w. t. enzyme due to repulsion caused by the cysteine as observed for platencin and platensimycin. For compounds **6a** and **c**, binding to w. t. *Pa*FabF was reduced compared to *Pa*FabF C164A while for compounds **6f** and **g** the binding levels for both enzyme variants were comparable (Figure 9 and Figure S4). This suggests that the compounds **6b** and **c** adopt a similar binding mode as **6a** closely interacting with the catalytic histidines while **6f** and **g** still bind into the malonyl binding site as their binding competes with platensimycin, however in a different orientation that is not affected by the presence or absence of the Cys164 side chain. Binding of **6b** was inconclusive, as some signals were more intense in presence than in the absence of the competing ligand without a clear explanation (Figure S4b).

For none of these ligands, binding constants could be determined using BLI due to binding to both protein and reference sensors at higher concentrations. Likewise, determination of binding modes with soaking experiments was unsuccessful.

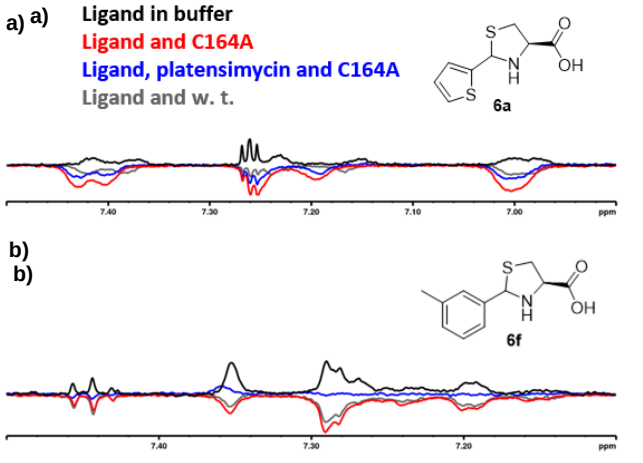
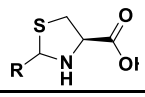
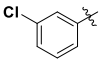
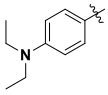
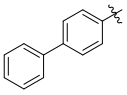


Figure 9. Overlay of WaterLOGSY spectra showing ligand in pure buffer (black), ligand with *Pa*FabF C164A (red), ligand with *Pa*FabF C164A and platensimycin (blue) and ligand with w. t. *Pa*FabF (gray). a) Compound **6a** showed a stronger signal (here pointing downwards) with only *Pa*FabF C164A than in competition and w. t. experiments. b) Compound **6f** showed a similar signal with *Pa*FabF C164A and the w. t. enzyme while the signal in the competition experiment was weaker.

Table 2. Analogues of virtual screening hit **6**. All compounds were assessed for binding to *Pa*FabF C164A. Competition experiments were carried out with platensimycin as competing ligand. ✓ indicates binding (C164A) or out-competing of binding with platensimycin (Competition). For WaterLOGSY no w. t., ✓ indicates the signals for the competition experiment and with w. t. enzyme were similar (low binding to the w. t. enzyme), while ✗ indicates no difference in the signal for w. t. enzyme and *Pa*FabF C164A (binding to both enzymes).

# 6	Structure	Dia-stereomer ratio	WaterLOGSY		
			C164A	Competition	no w. t.
a		3:2	✓	✓	✓
b		3:2	✓	a)	a)
c		1:1	✓	✓	✓
d		10:1	-b)		
e		3:2	-b)		
f		1.2:1	✓	✓	✗
g		4:1	✓	✓	✗
h*		3:2	-c)		

a) inconclusive experiment. b) compound was not soluble in 20 mM in DMSO. c) compound decomposed during shipment.

Conclusion

FabF is an attractive target for antibiotics that was shown to be essential in several genetic screens.[7–11] Ligands with diverse scaffolds are needed to explore this target. Here, we have established an experimental toolbox to support structure-based drug discovery efforts for *Pa*FabF. We went on to use this toolbox to validate hits obtained from virtual screening.

The BLI and NMR assays allow for the determination of binding affinities and binding sites. In most studies aiming at deriving SAR for platensimycin analogues no affinity or inhibition data was published. This is probably due to the complexity of inhibition and binding assays used so far requiring either coupling enzymes in addition to other helper enzymes to synthesize the substrates and get a read-out or radiolabelled ligands. [23,43] In this respect, the BLI assay is a clear improvement making the determination of binding constants for FabF straightforward. However, solubility of the compounds and aggregate formation is clearly a limiting factor, preventing the determination of binding constants of the virtual screening hits and the synthesised analogues. The NMR assays we have set up are well suited to detect binding of ligands with medium to low affinity. In combination with a competing ligand or the use of different enzyme variants they can also deliver information about the binding sites of the compounds.

Further, we have successfully optimized our previous crystallization conditions for *Pa*FabF to allow for soaking of small molecules in DMSO containing solutions.[36] We routinely obtain crystals diffracting to 1.7 Å allowing for detailed binding information for ligands. Nevertheless, for only one of the virtual screening hits a complex structure could be obtained. The frustratingly low rate of obtaining structural information for fragment hits has been reported before.[44] The reasons for this are not entirely clear, but it is possible that the binding affinity of the fragment hits in the assay buffer is different from the affinity in the soaking buffer causing the low success rate.

Virtual screening delivered 6 hits for which binding was confirmed using STD and WaterLOGSY experiments. However, the exact affinity could not be determined due to a combination of low solubility and low affinity. This is common for fragments and some rounds of optimization based on structural information might be needed to obtain measurable affinity.[45] Our attempts to optimize one of the compound has failed. Further efforts should be directed toward increasing the solubility of the compound series and exploring different parts of the binding site.

This work has laid the foundation to support structure-based design efforts for *Pa*FabF ligands. The assays are also suitable for experimental fragment screening using NMR, BLI or X-ray crystallography and such efforts are ongoing to explore this promising target further.

Experimental Section

Expression and purification of *Pa*FabF variants

All genes except of w.t. *Pa*FabF and *Pa*FabF C164Q were synthesized and cloned into pET-28a-TEV vectors by Genscript (New Jersey). The protein constructs for FabF containing an AviTag[37] for BLI experiments (here called *Pa*AviFabF) were designed as follows: N-terminal His6-tag, TEV cleavage site, Avi-tag, followed by linker residues (GS). All FabF variants were expressed as described earlier,[36] with only minor deviations. The plasmids were heat-shock transformed (45 s, 315 K) into Rosetta (DE3) pLysS competent *E. coli* cells (Novagen) and cultured in autoinduction medium (1L LB with 50 mL 20X NPS, 20 mL 50X 5052 and 1 mL MgSO4) supplemented with chloramphenicol (1 mL 30 mg/mL) and kanamycin (1 mL 50 mg/mL) for 48 h at 293 K with shaking at 250 rpm. Cells were harvested by centrifugation (45 min, 5000g, 277 K), and each pellet resuspended in 30 mL lysis buffer [50 mM Na2HPO4 pH 7.8, 300 mM NaCl, 20 mM imidazole, 10 % (v/v) glycerol, half a tablet of protease-inhibitor EDTA-free cocktail (Roche) and 0.05 mg DNase I (Sigma Aldrich)] and lysed on ice by sonication. The protein was purified as described earlier, but with a slight variation in gel filtration buffer [50 mM Na2HPO4, 150 mM NaCl, pH = 7.6, 10% (v/v) glycerol]. A high level of protein purity was confirmed by SDS–PAGE (Mini-PROTEAN TGX Stain-Free Precast

Gel; Bio-Rad) and the yield, typically in excess of 10 mg/L, determined using Nanodrop 2000 (ThermoFisher).

Expression and purification of GST-BirA

The GST-BirA (*E. coli* biotin ligase fused to glutathione-S-transferase) pGEX-6P-1 vector was a kind donation from Petra Hänzelmann (University of Würzburg). The vector was transformed into Rosetta(DE3)pLysS competent *E. coli* cells (Novagen) and GST-BirA was expressed and purified as described earlier.[38]

Biotinylation of *Pa*AviFabF variants

*Pa*AviFabF variants were biotinylated in a similar manner as described earlier.[38] In short, a 1 mL mixture of 100 µM *Pa*AviFabF, 1 uM GST-BirA, 5 mM $MgCl_2$, 150 uM D-biotin and 2 mM ATP in 1 x PBS was incubated at 303 K for 1 h shaking at 60 rpm. 3 µL of 50 mM D-biotin and 50 µL of 20 uM GST-BirA in 1 x PBS were then added, before the mixture was allowed to shake for another 1 h. 0.1 mL of Glutathione Sepharose (50% slurry in 1 x PBS; Cytiva) was added and the resulting suspension incubated on ice for 1 h. The mixture was then centrifuged (2.000 x g, 2 min) and the supernatant was brought through a prewashed (2 mL 1x PBS at 2.000 x g, 2 min, twice) Zeba Spin desalting column (7k Molecular weight cut-off; ThermoFischer) to remove the cofactors and the flow-through collected.

BLI measurements

BLI measurements were performed using an Octet RED96 instrument (FòrteBio) in polypropylene 96-well F-Bottom microplates (Greiner Bio-One) at a constant temperature of 298 K. Wells contained a total volume of 200 µL liquid. The biotinylated proteins at a concentration of 500 µg/mL in assay buffer [1 x PBS, 0.0001 % (v/v) TritonX100, 5 mM DTT, 5% (v/v) DMSO] were immobilized onto Super Streptavidin (SSA) biosensors (FòrteBio). The loading was carried out in a series of sequential steps. Typically, first an initial baseline step of 300 s was used, followed by immobilization of the protein for 900 s until a response of 12-13 nm was reached. Unoccupied streptavidin sites were then quenched with biocytin (10 µg/mL) for 120 s, followed by a final washing step of 60 s in the assay buffer. Binding of compounds to *Pa*AviFabF (both w. t. and C164A), and biocytin blocked reference sensors was measured in three steps: 30 s baseline, 60 s association and 30-60 s disassociation. The obtained curves were analysed using the FortéBio software. The curves were aligned to their baselines and double-referenced against buffer wells and biocytin reference sensors. Binding constants were obtained by using a steady state model fitted to the averaged response obtained between 50 and 55 s. The reported K_D values are averages of four independent measurements, using two different protein batches for platencin, but the same protein batch for platensimycin.

X-ray crystallography

Purified *Pa*FabF C164Q (14.5 - 15.6 mg/mL) was mixed with well buffer (0.20 - 0.24M ammonium formate and 26% - 31.2% PEG3350) at 0.8:0.9 µL drop ratio and let to equilibrate at 25 C using the vapor diffusion, sitting drop technique. Crystals were observed as large plates growing from the same centre within 24 hours and crystals growth continued for up to 3 days. For soaking, 300nL of platensimycin and **6a** in solutions (2 and 50mM respectively in 20% DMSO, 0.20mM ammonium formate and 26% PEG3350) were added in the crystal drop and let to equilibrate for a maximum of 2.5 hours. Crystals were then fished and frozen in liquid nitrogen. X-ray data were collected at 100K at the Biomax beamline, MaxIV synchrotron radiation facility in Lund Sweden. Dimple[46] from the CCP4i2 suite[47] was used for molecular replacement using the PDB ID: 4JB6 as input model. Modelled structures were manually inspected and corrected using Coot[48] while Refmac5[49] was used for the refinement of the crystal structures. Data collection, processing and refinement statistics are listed in Table S1. The complex *Pa*FabF-platensimycin has obtained the PDB ID 7OC1 and the complex *Pa*FabF-**6a** the PDB ID 7OC0.

Binding measurements using NMR

All NMR spectra were acquired in 3 mm NMR tubes at 298 K using a Bruker AVANCE NEO 600 MHz spectrometer (Bruker BioSpin, Zürich, Switzerland) equipped with a SampleJet and a QCI-P CryoProbe. Initial STD (stddiffesgp.3) and WaterLOGSY (ephogsygpno.2) screening of commercial compounds was carried out as singletons containing a final concentration of 500 µM ligand, 2.5 µM *Pa*FabF C164A, and follow up WaterLOGSY experiments were carried out with 400 µM ligand, 10 µM platensimycin and 2.5 µM *Pa*FabF C164A. All samples were dissolved in PBS buffer (pH = 7.4) containing 10% D2O and 2.5% d6-DMSO. Synthesized compounds were screened at 1 mM with 10 µM *Pa*FabF C164A, w. t. *Pa*FabF and 20 µM platensimycin for competition experiments in a similar buffer to above, but with 5% d6-DMSO. STD spectra of platensimycin and platensin measurements were obtained using 100 µM ligand and 10 µM hisFabF C164A with 5% d6-DMSO. Buffer controls were recorded in all cases. D-glucose was used as negative control (Figure S2i and S3e).

STD spectra were acquired using an on-resonance irradiation at 0.7 ppm and off-resonance irradiation at −40 ppm. A saturation time of 1 s and relaxation delay of 2 s were used. The number of scans was 256. WaterLOGSY experiments were conducted using a saturation time of 1 s and a mixing time of 1.7 s. The number of scans was 512.

Virtual Screening

Modelling tasks were carried out using the Molecular Operating Environment (MOE; Chemical Computing Group, Montreal, QC, Canada) if not stated otherwise. The apo *Pa*FabF C164Q structure (PDB id 4JB6) was used as template for docking. However, the side chain of Glu164 was modelled as cysteine to obtain an open malonyl-competitive binding site with a native catalytic triade. Polar hydrogen atoms were added, their positions energy minimized, and partial charges were assigned based on AMBER force field parameters. All water molecules were deleted. The polar aromatic head group of platensimycin was used to define a sphere set for ligand placement. Grid-based excluded volume, van-der-Waals potential, electrostatic potential and solvent occlusion maps were calculated as described earlier.[50,51] The partial charges of His304 and His341 were adjusted to favour hydrogen bonds to these residues. For that purpose, the partial charge of NE2 was reduced by 0.3 and the partial charge of HNE was increased by 0.3, keeping the net charge constant. The setup was validated by predicting the binding mode of platensimycin which resulted in a pose with all key interactions preserved.

A pharmacophore for FabF inhibitors was derived from the binding modes of known binders (platensimycin and its derivatives, thiolactomycin and cerulenin). Also structures of the homologue protein FabB were taken into account to increase the diversity of ligands. The final pharmacophore setup contained an essential hydrogen bond acceptor feature to His304 and His341 and three optional features of which at least 2 needed to be fulfilled: a hydrogen-bond acceptor to Thr308, a hydrogen-bond donor to Thr271 and an aromatic feature for the interaction with Phe400 (Figure 6). The radii for all hydrogen-bond features were set to 1 Å and the radius for the hydrophobic interaction was set to 1.5 Å.

Our in-house MySQL database of around 5 million commercially available compounds from various suppliers was filtered for lead-like molecules fulfilling the following criteria: 10-30 heavy atoms, 1-7 hydrogen-bond acceptors, 1-3 hydrogen-bond donors , $-3 \leq \text{clogP} < 3$, < 8 rotatable bonds, 1-3 fused ring systems, < 2 nitril groups. $< 8 < 8 < 7$.[41] The filtered compounds were charged, tautomerized and stereoisomerized using in-house python scripts based on OpenEye's OEChem toolkit (OEChem, version 2016.6.1, OpenEye Scientific Software, Inc., Santa Fe,NM, USA) (scripts are available at https://github.com/ruthbrenk/compound_preparation). Conformers were generated using OpenEye's OMEGA toolkit. Subsequently, filtered and prepared molecules were screened with the above described pharmacophore. Compounds that passed the pharmacophore filter were transformed into a format suitable for docking as described previously.[52]

The prepared compounds were docked into the receptor using DOCK 3.6.[51,53,54] Docking parameters were set as follows: ligand and receptor bins 0.4 Å, overlap bins 0.2 Å and distance tolerance for matching ligand atoms to receptor matching sites 1.5 Å. For each compound, the best scored representative was stored in the docking hit list. Docking poses were analysed using MOE. First, rigid docking poses were filtered with the pharmacophore described above. Subsequently, the compounds passing this filter were ranked by their predicted ligand efficiency which was calculated as docking score / number of heavy atoms[55] and the 1000 highest-ranked poses were inspected by eye. The poses were rated based on the quality of the putative hydrogen bonds and the ligand conformations. Fragment-sized ligands forming the desired hydrogen bond to His304 and His341 were prioritized for purchase.

Design of analogues of compound 6

Compounds with substitutions in the 2-position of the thiazolidine ring were drawn in Maestro (Schrödinger Release 2018-1: Maestro, Schrödinger, LLC, New York, NY, 2021) and minimized in the *Pa*FabF C164Q binding site. Compounds that explored additional interactions in the pocket in which the thiophene ring was placed and/or that filled available space in the binding site adjacent to this ring (Figure 8d) were shortlisted for synthesis.

Chemical compounds

Platensimycin was purchased from Sigma Aldrich and platencin was purchased from Cayman Chemicals. The virtual screening hits **1**-**7** were purchased from Enamine Ltd., virtual screening hit **8** from Vitas-M Laboratory Ltd. Identities were confirmed using ^{1}H NMR and HRMS (see supplementary material).

Synthesis

Synthesis of compounds **6a**-**h** (Table 2) is described in the supplementary material.

Acknowledgements

The work was supported by the Research Council of Norway (grant number 273588) and the Indian Council of Medical Research (grant number AMR/IN/111/2017 – ECD – II). LOE obtained a PhD studentship through the Global Challenges initiative at the University of Bergen and further support from the Melzer Research Fund. We made use of the Facility for Biophysics, Structural Biology and Screening at the University of Bergen (BiSS), which has received funding from the Research Council of Norway (RCN) through the NORCRYST (grant number 245828) and NOR–OPENSCREEN (grant number 245922) consortia. Further, this work was partly supported by the Bergen Research Foundation, Sparebankstiftinga Sogn og Fjordane, and the Research Council of Norway through The Norwegian NMR Platform, NNP (226244/F50). Diffraction data were collected at MaxIV (Lund-Sweden), ESRF (France), and DLS (Oxford-UK). We thank Khan Kim Dao for excellent support with protein purification and Openeye for free software licenses.

Keywords: antibiotics, bio-layer interferometry, ligand-based NMR, structure-based design, virtual screening

[1] E. D. Brown, G. D. Wright, *Nature* **2016**, *529*, 336–343.

[2] E. Tacconelli, E. Carrara, A. Savoldi, S. Harbarth, M. Mendelson, D. L. Monnet, C. Pulcini, G. Kahlmeter, J. Kluytmans, Y. Carmeli, M. Ouellette, K. Outterson, J. Patel, M. Cavaleri, E. M. Cox, C. R. Houchens, M. L. Grayson, P. Hansen, N. Singh, U. Theuretzbacher, N. Magrini, *Lancet Infect. Dis.* **2018**, *18*, 318–327.

[3] U. Theuretzbacher, K. Bush, S. Harbarth, M. Paul, J. H. Rex, E. Tacconelli, G. E. Thwaites, *Nat. Rev. Microbiol.* **2020**, *18*, 286–298.

[4] J. Yao, C. O. Rock, *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* **2017**, *1862*, 1300–1309.

[5] F. J. Asturias, J. Z. Chadick, I. K. Cheung, H. Stark, A. Witkowski, A. K. Joshi, S. Smith, *Nat. Struct. Mol. Biol.* **2005**, *12*, 225–232.

[6] A. J. Kastaniotis, K. J. Autio, J. M. Kerätär, G. Monteuuis, A. M. Mäkelä, R. R. Nair, L. P. Pietikäinen, A. Shvetsova, Z. Chen, J. K. Hiltunen, *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* **2017**, *1862*, 39–48.

[7] B. E. Poulsen, R. Yang, A. E. Clatworthy, T. White, S. J. Osmulski, L. Li, C. Penaranda, E. S. Lander, N. Shoresh, D. T. Hung, *Proc. Natl. Acad. Sci.* **2019**, *116*, 10072–10080.

[8] K. H. Turner, A. K. Wessel, G. C. Palmer, J. L. Murray, M. Whiteley, *Proc. Natl. Acad. Sci.* **2015**, *112*, 4110–4115.

[9] S. A. Lee, L. A. Gallagher, M. Thongdee, B. J. Staudinger, S. Lippman, P. K. Singh, C. Manoil, *Proc. Natl. Acad. Sci.* **2015**, *112*, 5189–5194.

[10] D. Skurnik, D. Roux, H. Aschard, V. Cattoir, D. Yoder-Himes, S. Lory, G. B. Pier, *PLoS Pathog.* **2013**, *9*, e1003582.

[11] N. T. Liberati, J. M. Urbach, S. Miyata, D. G. Lee, E. Drenkard, G. Wu, J. Villanueva, T. Wei, F. M. Ausubel, *Proc. Natl. Acad. Sci.* **2006**, *103*, 2833–2838.

[12] A. Sarkar, R. Brenk, *PloS One* **2015**, *10*, e0137279.

[13] 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, S. B. Singh, *Nature* **2006**, *441*, 358–361.

[14] Z. Feng, D. Chakraborty, S. B. Dewell, B. V. Reddy, S. F. Brady, *J. Am. Chem. Soc.* **2012**, *134*, 2981–2987.

[15] 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, S. B. Singh, *Proc. Natl. Acad. Sci.* **2007**, *104*, 7612–7616.

[16] S. Kauppinen, M. Siggaard-Andersen, P. von Wettstein-Knowles, *Carlsberg Res. Commun.* **1988**, *53*, 357–370.

[17] G. D'Agnolo, I. S. Rosenfeld, J. Awaya, S. Omura, P. R. Vagelos, *Biochim. Biophys. Acta* **1973**, *326*, 155–156.

[18] I. Nishida, A. Kawaguchi, M. Yamada, *J. Biochem. (Tokyo)* **1986**, *99*, 1447–1454.

[19] S. Jackowski, C. M. Murphy, J. E. Cronan, C. O. Rock, *J. Biol. Chem.* **1989**, *264*, 7624–7629.

[20] S. B. Singh, J. G. Ondeyka, K. B. Herath, C. Zhang, H. Jayasuriya, D. L. Zink, G. Parthasarathy, J. W. Becker, J. Wang, S. M. Soisson, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4756–4759.

[21] A. C. Price, K.-H. Choi, R. J. Heath, Z. Li, S. W. White, C. O. Rock, *J. Biol. Chem.* **2001**, *276*, 6551–6559.

[22] M. Moche, G. Schneider, P. Edwards, K. Dehesh, Y. Lindqvist, *J. Biol. Chem.* **1999**, *274*, 6031–6034.

[23] 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, S. B. Singh, *Nature* **2006**, *441*, 358–361.

[24] K. Tiefenbacher, J. Mulzer, *J. Org. Chem.* **2009**, *74*, 2937–2941.

[25] K. C. Nicolaou, G. S. Tria, D. J. Edmonds, *Angew. Chem.* **2008**, *120*, 1804–1807.

[26] K. C. Nicolaou, A. Li, D. J. Edmonds, *Angew. Chem. Int. Ed.* **2006**, *45*, 7086–7090.

[27] A. K. Ghosh, K. Xi, *J. Org. Chem.* **2009**, *74*, 1163–1170.

[28] K. Tian, Y. Deng, L. Qiu, X. Zhu, B. Shen, Y. Duan, Y. Huang, *ChemistrySelect* **2018**, *3*, 12625–12629.

[29] K. C. Nicolaou, A. F. Stepan, T. Lister, A. Li, A. Montero, G. S. Tria, C. I. Turner, Y. Tang, J. Wang, R. M. Denton, D. J. Edmonds, *J. Am. Chem. Soc.* **2008**, *130*, 13110–13119.
[30] K. C. Nicolaou, T. Lister, R. M. Denton, A. Montero, D. J. Edmonds, *Angew. Chem.* **2007**, *119*, 4796–4798.
[31] J. Wang, V. Lee, H. O. Sintim, *Chem. – Eur. J.* **2009**, *15*, 2747–2750.
[32] J. Krauss, V. Knorr, V. Manhardt, S. Scheffels, F. Bracher, *Arch. Pharm. (Weinheim)* **2008**, *341*, 386–392.
[33] M. Fisher, R. Basak, A. P. Kalverda, C. W. G. Fishwick, W. B. Turnbull, A. Nelson, *Org. Biomol. Chem.* **2013**, *12*, 486–494.
[34] Y. Deng, X. Weng, Y. Li, M. Su, Z. Wen, X. Ji, N. Ren, B. Shen, Y. Duan, Y. Huang, *J. Med. Chem.* **2019**, *62*, 6682–6693.
[35] Z. Zheng, J. B. Parsons, R. Tangallapally, W. Zhang, C. O. Rock, R. E. Lee, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2585–2588.
[36] B. Baum, L. S. M. Lecker, M. Zoltner, E. Jaenicke, R. Schnell, W. N. Hunter, R. Brenk, *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2015**, *71*, 1020–6.
[37] D. Beckett, E. Kovaleva, P. J. Schatz, *Protein Sci.* **1999**, *8*, 921–929.
[38] M. Fairhead, M. Howarth, *Methods Mol. Biol. Clifton NJ* **2015**, *1266*, 171–184.
[39] E. Martin, J. Wang, I. Zaror, J. Yu, K. Yan, M. Doyle, P. Feucht, K. Shoemaker, B. Warne, M. Chin, B. Sy, L. Leder, M. Meyerhofer, C. Wartchow, D. Yao, in *Label-Free Technol. Drug Discov.* (Eds.: M. Cooper, L.M. Mayr), John Wiley & Sons, Ltd, **2011**, pp. 223–240.
[40] J. Angulo, P. M. Nieto, *Eur. Biophys. J.* **2011**, *40*, 1357–1369.
[41] R. Brenk, A. Schipani, D. James, A. Krasowski, I. H. Gilbert, J. Frearson, P. G. Wyatt, *ChemMedChem* **2008**, *3*, 435–444.
[42] A. Yu. Ershov, D. G. Nasledov, I. V. Lagoda, V. V. Shamanin, *Chem. Heterocycl. Compd.* **2014**, *50*, 1032–1038.
[43] J. G. Borgaro, A. Chang, C. A. Machutta, X. Zhang, P. J. Tonge, *Biochemistry* **2011**, *50*, 10678–10686.
[44] E. Hassaan, P.-O. Eriksson, S. Geschwindner, A. Heine, G. Klebe, *ChemMedChem* **2020**, *15*, 324–337.
[45] A. C. Gibbs, M. C. Abad, X. Zhang, B. A. Tounge, F. A. Lewandowski, G. T. Struble, W. Sun, Z. Sui, L. C. Kuo, *J. Med. Chem.* **2010**, *53*, 7979–7991.
[46] M. Wojdyr, R. Keegan, G. Winter, A. Ashton, *Acta Crystallogr. Sect. A* **2013**, *69*, s299.
[47] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. von Delft, D. Waterman, K. Wilson, M. Winn, M. Wojdyr, *Acta Crystallogr. Sect. Struct. Biol.* **2018**, *74*, 68–84.
[48] P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
[49] 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.
[50] R. Brenk, J. J. Irwin, B. K. Shoichet, *J. Biomol. Screen. Off. J. Soc. Biomol. Screen.* **2005**, *10*, 667–674.
[51] M. M. Mysinger, B. K. Shoichet, *J. Chem. Inf. Model.* **2010**, *50*, 1561–1573.
[52] C. P. Mpamhanga, D. Spinks, L. B. Tulloch, E. J. Shanks, D. A. Robinson, I. T. Collie, A. H. Fairlamb, P. G. Wyatt, J. A. Frearson, W. N. Hunter, I. H. Gilbert, R. Brenk, *J. Med. Chem.* **2009**, *52*, 4454–4465.
[53] D. M. Lorber, B. K. Shoichet, *Protein Sci.* **1998**, *7*, 938–50.
[54] B. Q. Wei, W. A. Baase, L. H. Weaver, B. W. Matthews, B. K. Shoichet, *J. Mol. Biol.* **2002**, *322*, 339–355.
[55] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discov. Today* **2004**, *9*, 430–431.

Supporting information

An experimental toolbox for structure-based hit discovery for *P. aeruginosa* FabF, a promising target for antibiotics

Ludvik Olai Espeland, Dr. Charis Georgiou, Dr. Raphael Klein, Hemalatha Bhukya, Prof. Dr. Bengt Erik Haug, Assoc. Prof. Dr. Jarl Underhaug, Dr. Prathama S Mainkar, and Prof. Dr. Ruth Brenk

Table of contents

1 X-ray crystallography 4
Figure S1: Binding of platensimycin to *Pa*FabF C164Q 4
Table S1: Data-collection and refinement statistics 5
2 Ligand-observed spectra in presence of *Pa*FabF 6
2.1 Saturation transfer difference (STD) NMR of virtual screening hits 6
Figure S2a: Stacked STD spectra of compound **1** 6
Figure S2b: Stacked STD spectra of compound **2** 7
Figure S2c: Stacked STD spectra of compound **3** 8
Figure S2d: Stacked STD spectra of compound **4** 9
Figure S2e: Stacked STD spectra of compound **5** 10
Figure S2f: Stacked STD spectra of compound **6** 11
Figure S2g: Stacked STD spectra of compound **7** 12
Figure S2h: Stacked STD spectra of compound **8** 13
Figure S2i: Stacked STD spectra of negative control D-glucose 14
2.2 Competitive WaterLOGSY with purchased hits 15
Figure S3a: Stacked WaterLOGSY spectra of compound **1** 15
Figure S3b: Stacked WaterLOGSY spectra of compound **2** 16
Figure S3c: Stacked WaterLOGSY spectra of compound **5** 17
Figure S3d: Stacked WaterLOGSY spectra of compound **6** 18
Figure S3e: Stacked WaterLOGSY spectra of compound **7** 19
Figure S3f: Stacked WaterLOGSY spectra of compound **8** 20
Figure S3e: Stacked WaterLOGSY spectra of negative control D-glucose 21
2.3 WaterLOGSY for synthesized compounds 22
Figure S4a: Stacked WaterLOGSY spectra of compound **6a** 22
Figure S4b: Stacked WaterLOGSY spectra of compound **6b** 23
Figure S4c: Stacked WaterLOGSY spectra of compound **6c** 24
Figure S4d: Stacked WaterLOGSY spectra of compound **6f** 25
Figure S4e: Stacked WaterLOGSY spectra of compound **6g** 26

3 Steady state analysis 27
Figure S5: Steady state analysis 27
4 Structure validation of hits and synthesis of analogues of hit **6** 28
4.1 General Considerations 28
4.2 Structure validation of hit compounds 28
4-Methyl-2-phenyl-1*H*-imidazole-5-carboxylic acid (**1**) 28
2-(3-Chlorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (**2**) 28
4,5-Dichloro-1*H*-pyrrole-2-carboxylic acid (**5**) 28
(2*RS*, 4*RS*)-2-(Thiophen-2-yl)thiazolidine-4-carboxylic acid (**6**) 28
2-(4-Fluorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (**7**) 28
(2*RS*, 4*RS*)-2-(*m*-Tolyl)thiazolidine-4-carboxylic acid (**8**) 28
4.3 General procedure for synthesis of substituted thiazolidine-4-carboxylic acids 29
(2*RS*, 4*R*)-2-(Thiophen-2-yl)thiazolidine-4-carboxylic acid (**6a**) 29
(2*RS*, 4*R*)-2-(3-Chlorophenyl)thiazolidine-4-carboxylic acid (**6b**) 29
(2*RS*, 4*R*)-2-(*p*-Tolyl)thiazolidine-4-carboxylic acid (**6c**) 29
(2*RS*, 4*R*)-2-(4-(Diethylamino)phenyl)thiazolidine-4-carboxylic acid (**6d**) 29
(2*RS*, 4*R*)-2-([1,1'-Biphenyl]-4-yl)thiazolidine-4-carboxylic acid (**6e**) 30
(2*RS*, 4*R*)-2-(m-Tolyl)thiazolidine-4-carboxylic acid (**6f**) 30
(2*RS*, 4*R*)-2-(Pyridin-4-yl)thiazolidine-4-carboxylic acid (**6g**) 30
(2*RS*, 4*R*)-2-(1H-Imidazol-4-yl)thiazolidine-4-carboxylic acid (**6h**) 30
4.4 ^{1}H NMR spectra of purchased hit compounds 31
Figure S6a: ^{1}H NMR spectrum of 4-methyl-2-phenyl-1*H*-imidazole-5-carboxylic acid (**1**) 31
Figure S6b: ^{1}H NMR spectrum of 2-(3-chlorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (**2**) 32
Figure S6c: ^{1}H NMR spectrum of 4,5-dichloro-1H-pyrrole-2-carboxylic acid (**5**) 32
Figure S6d: ^{1}H NMR spectrum of (2*RS*,4*RS*)-2-(thiophen-2-yl)thiazolidine-4-carboxylic acid (**6**) 33
Figure S6e: ^{1}H NMR spectrum 2-(4-fluorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (**7**) 33
Figure S6f: ^{1}H NMR spectrum of (2RS, 4RS)-2-(*m*-tolyl)thiazolidine-4-carboxylic acid (**8**) 34
4.5 ^{1}H and ^{13}C NMR spectra of synthesized compound 35
Figure S7a: 1H NMR spectrum of (2RS, 4R)-2-(thiophen-2-yl)thiazolidine-4-carboxylic acid (**6a**) 35
Figure S7b: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(thiophen-2-yl)thiazolidine-4-carboxylic acid (6**a)** 35
Figure S7c: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(3-chlorophenyl)thiazolidine-4-carboxylic acid (**6b**) 36
Figure S7d: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(3-chlorophenyl)thiazolidine-4-carboxylic acid (**6b**) 36
Figure S7e: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(*p*-tolyl)thiazolidine-4-carboxylic acid (**6c**) 37
Figure S7f: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(*p*-tolyl)thiazolidine-4-carboxylic acid (**6c**) 37
Figure S7g: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(4-(diethylamino)phenyl)thiazolidine-4-carboxylic acid (**6d**) 38
Figure S7d: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(4-(diethylamino)phenyl)thiazolidine-4-carboxylic acid (**6d**) 38
Figure S7e: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-([1,1'-biphenyl]-4-yl)thiazolidine-4-carboxylic acid (**6e**) 39
Figure S7f: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-([1,1'-biphenyl]-4-yl)thiazolidine-4-carboxylic acid (**6e**) 39
Figure S7g: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(*m*-tolyl)thiazolidine-4-carboxylic acid (**6f**) 40
Figure S7h: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(m-tolyl)thiazolidine-4-carboxylic acid (**6f**) 40
Figure S7i: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(m-tolyl)thiazolidine-4-carboxylic acid (**6f**) 41
Figure S7j: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(pyridin-4-yl)thiazolidine-4-carboxylic acid (**6g**) 41

Figure S7k: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(1H-imidazol-4-yl)thiazolidine-4-carboxylic acid (**6h**). 42
Figure S7l: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(1H-imidazol-4-yl)thiazolidine-4-carboxylic acid (**6h**). 42

1 X-ray crystallography

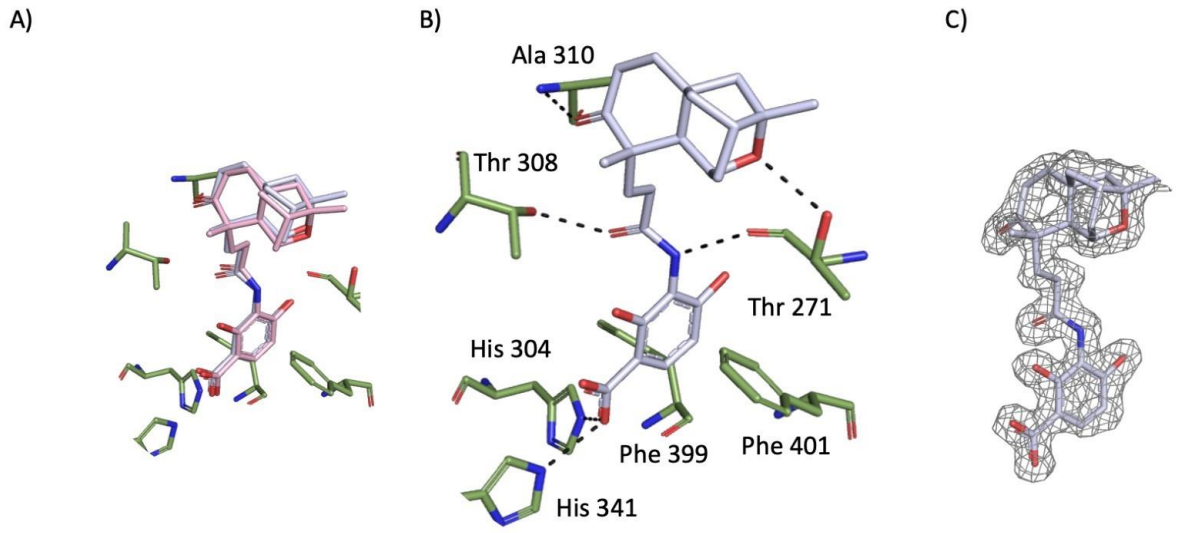


Figure S1: Binding of platensimycin to *Pa*FabF C164Q. **A)** Overlay of platensimycin binding to *Ec*FabF C163Q (pink sticks - PDB ID: 2GFX) and *Pa*FabF C164Q (light purple sticks for ligand and green sticks for protein). **B)** Interactions between platensimycin and *Pa*FabF C164Q. **C**) Fo-Fc map of platensimycin binding to *Pa*FabF C164Q, contoured at 3.0 sigma.

Table S1: Data-collection and refinement statistics
Values in parentheses are for the highest resolution shell.

Structure	*Pa*FabF-C164Q - platensimycin	*Pa*FabF-C164Q - **6a**
PDB code	7OC1	7OC0
Data collection and processing		
Space group	$P\ 2_12_12_1$	$P\ 2_12_12_1$
***a*, *b*, *c* (Å)**	84.02 137.19 65.69	65.71, 84.05, 137.72
α, β, γ (°)	90.00 90.00 90.00	90.00 90.00 90.00
Solvent content (%)	42	44
Diffraction data		
Resolution range (Å)	68.56 (1.8)	48.46 (1.78)
Unique reflections	71222 (5212)	63035 (2473)
Multiplicity	6.7 (6.7)	5.8 (1.9)
R_{merge}	0.10 (0.29)	0.05 (0.36)
Completeness (%)	100.0	85.36
I/sigI	11.05 (4.98)	18.8 (2.2)
Refinement		
R_{work}/R_{free}	0.16/0.19	0.16/0.19
Quaternary structure	dimer	dimer
Protein residues (in a dimer)	412	411
Cations	$1K^+$	$1Mg^{2+}$
Ligands	2 * PMN	2 * **6a**
Other molecules	12	6
Water molecules	432	499
R.m.s.d.s		
Bonds (Å)	0.011	0.006
Angles (Å)	1.49	1.20
Ramachandran plot, residues in (%)		
Favoured regions	98	98
Allowed regions	2	2
Outlier regions	0	0
Mean *B* factors (Å^2)		
Overall	10	24
Protein atoms	8	21
Ligands	6	35

2 Ligand-observed spectra in presence of *Pa*FabF

2.1 Saturation transfer difference (STD) NMR of virtual screening hits

The DMSO peak can be seen in all spectra at 2.60 ppm, while the multiples at 3.45-3.57 and 3.66-3.70 are due to glycerol impurity.

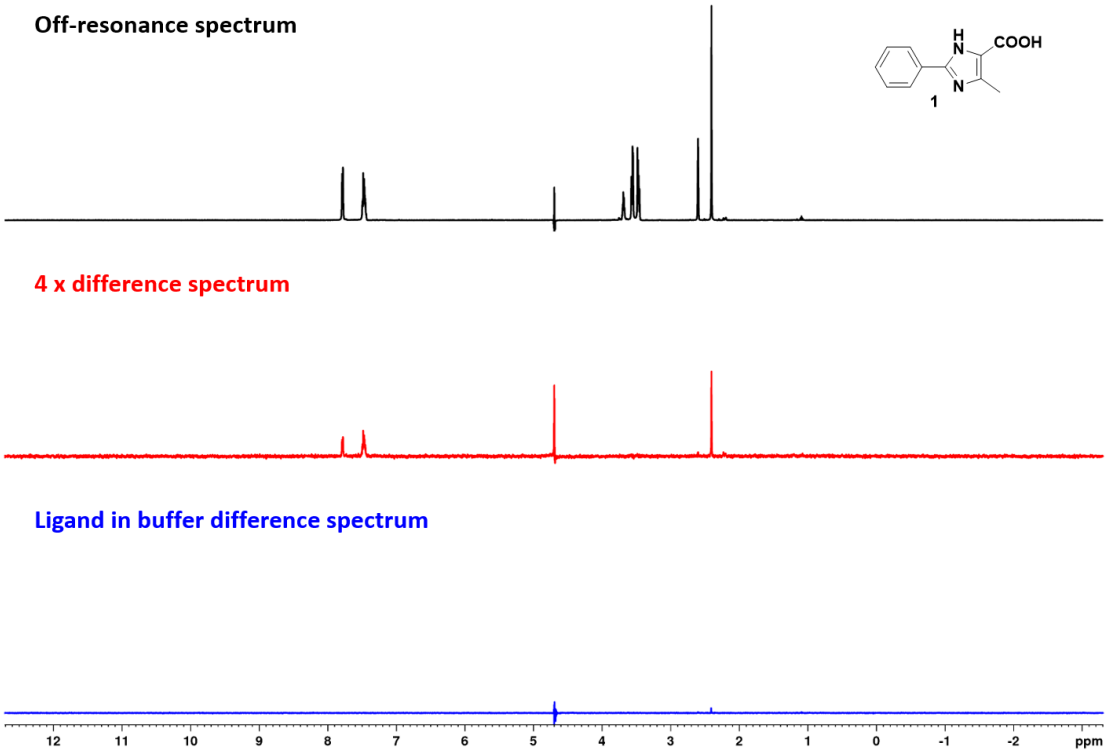


Figure S2a: Stacked STD spectra of compound **1** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 4 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

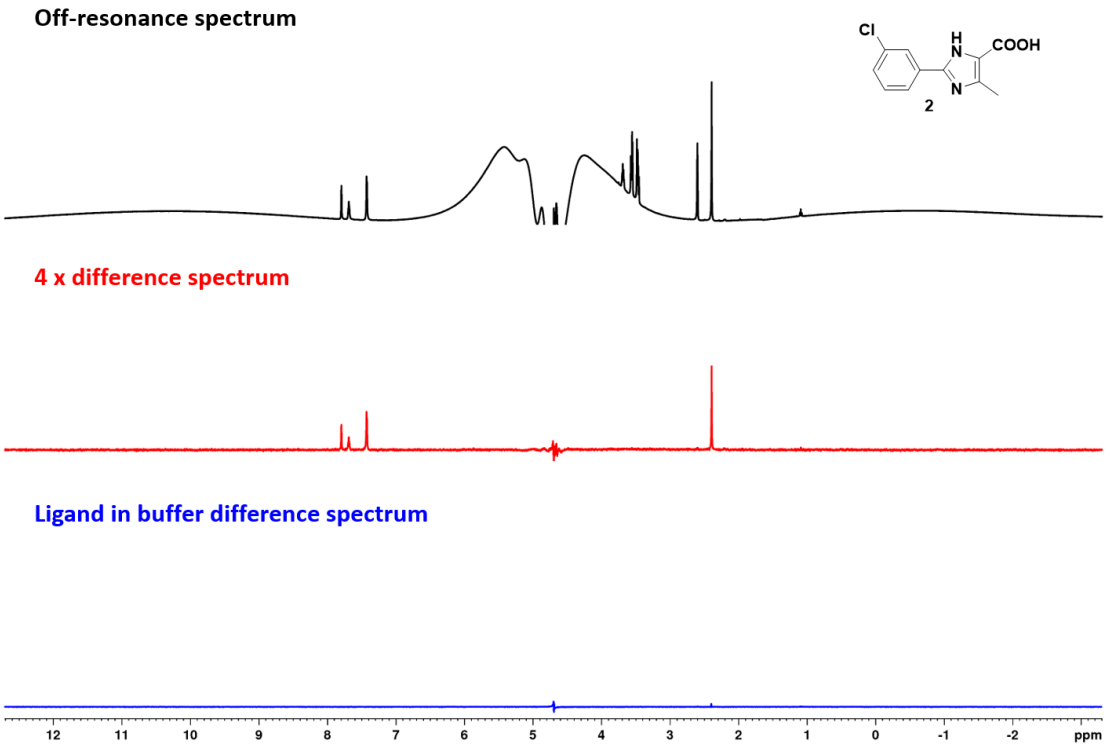


Figure S2b: Stacked STD spectra of compound **2** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 4 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

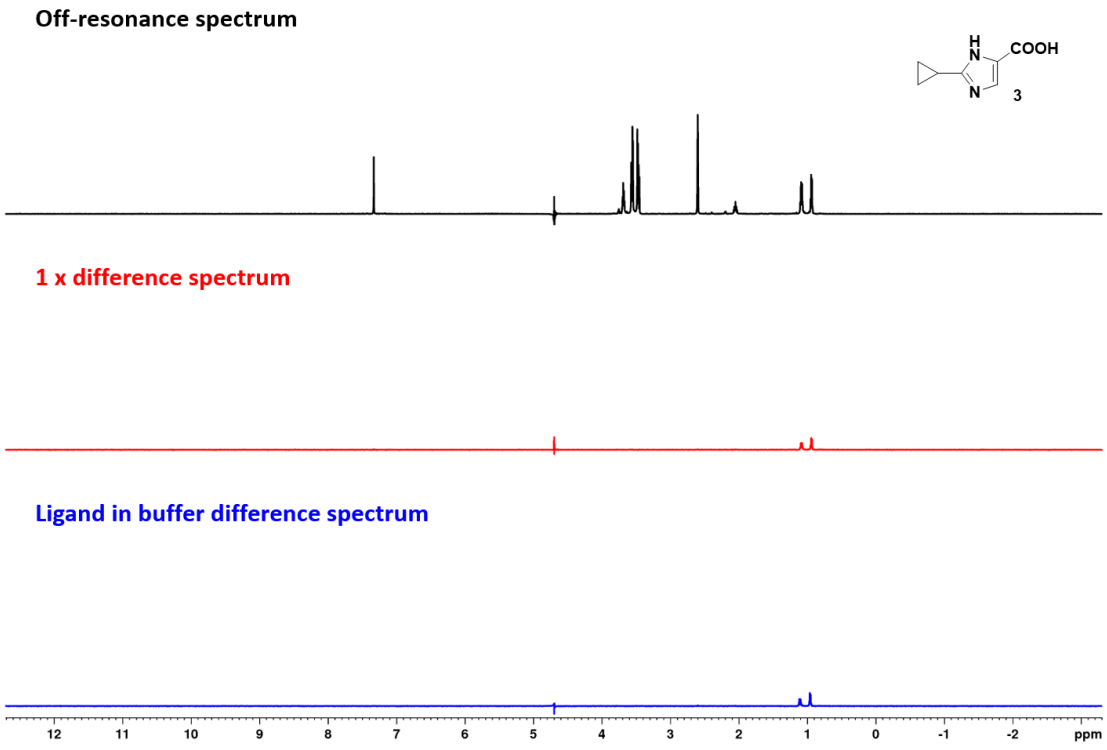


Figure S2c: Stacked STD spectra of compound **3** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 1 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

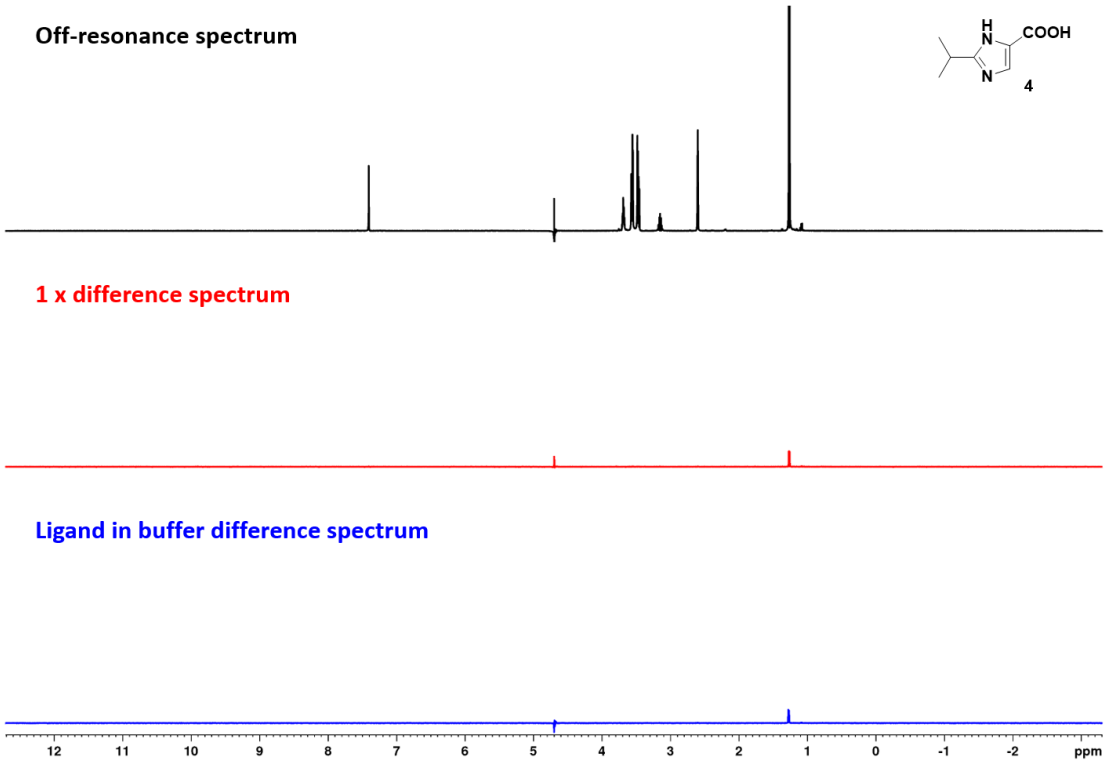


Figure S2d: Stacked STD spectra of compound **4** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 1 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

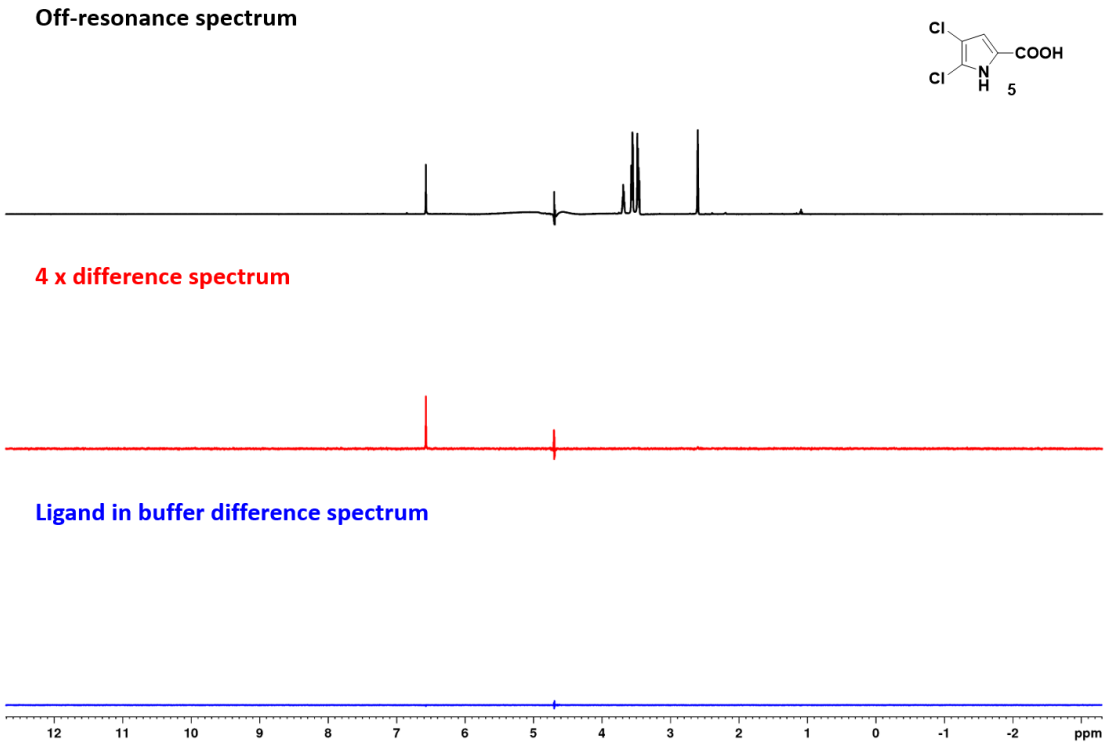


Figure S2e: Stacked STD spectra of compound **5** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 4 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

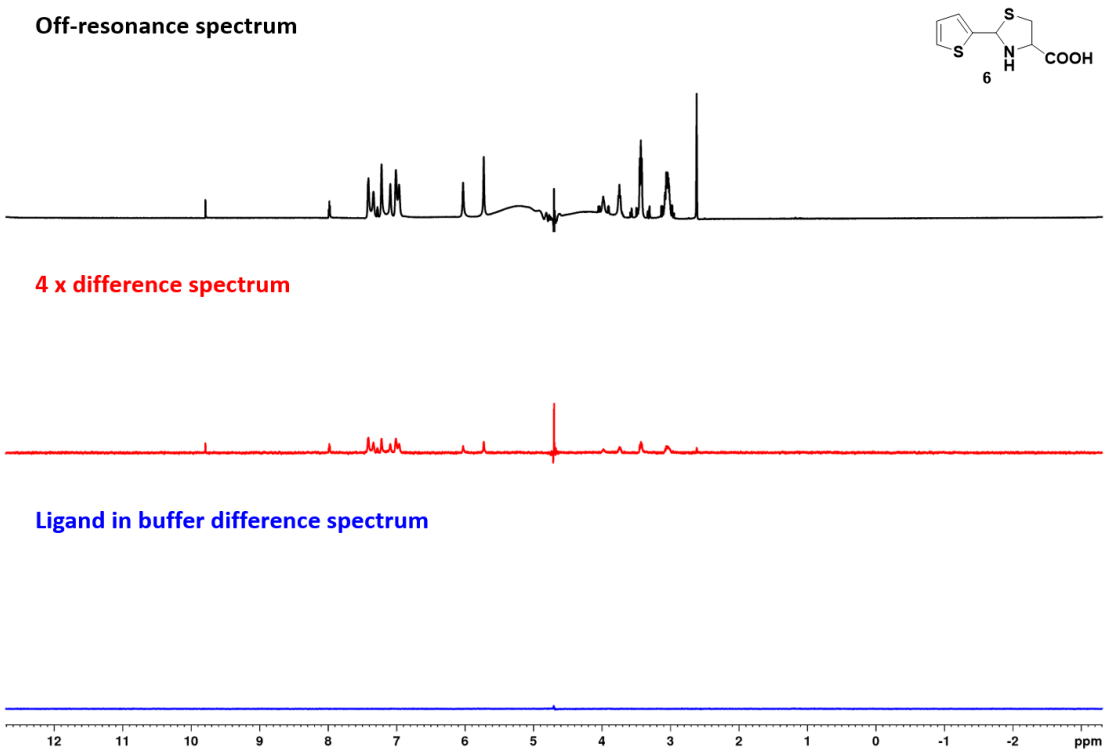


Figure S2f: Stacked STD spectra of compound **6** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 4 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

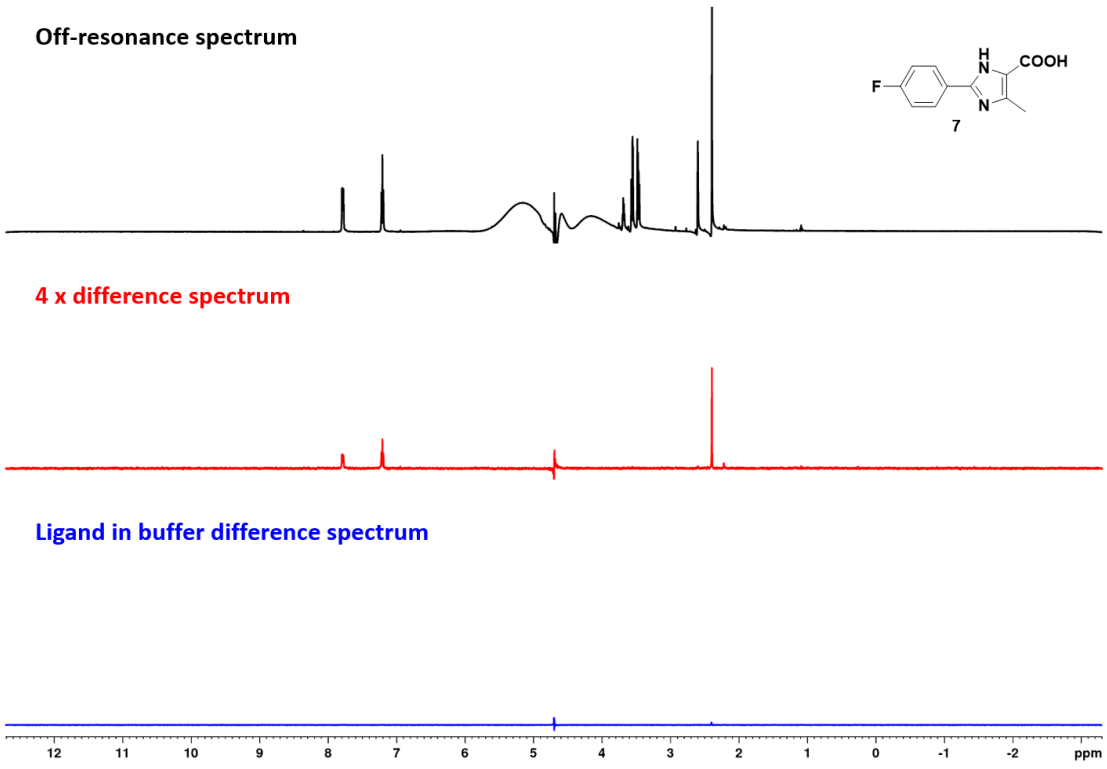


Figure S2g: Stacked STD spectra of compound **7** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 4 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

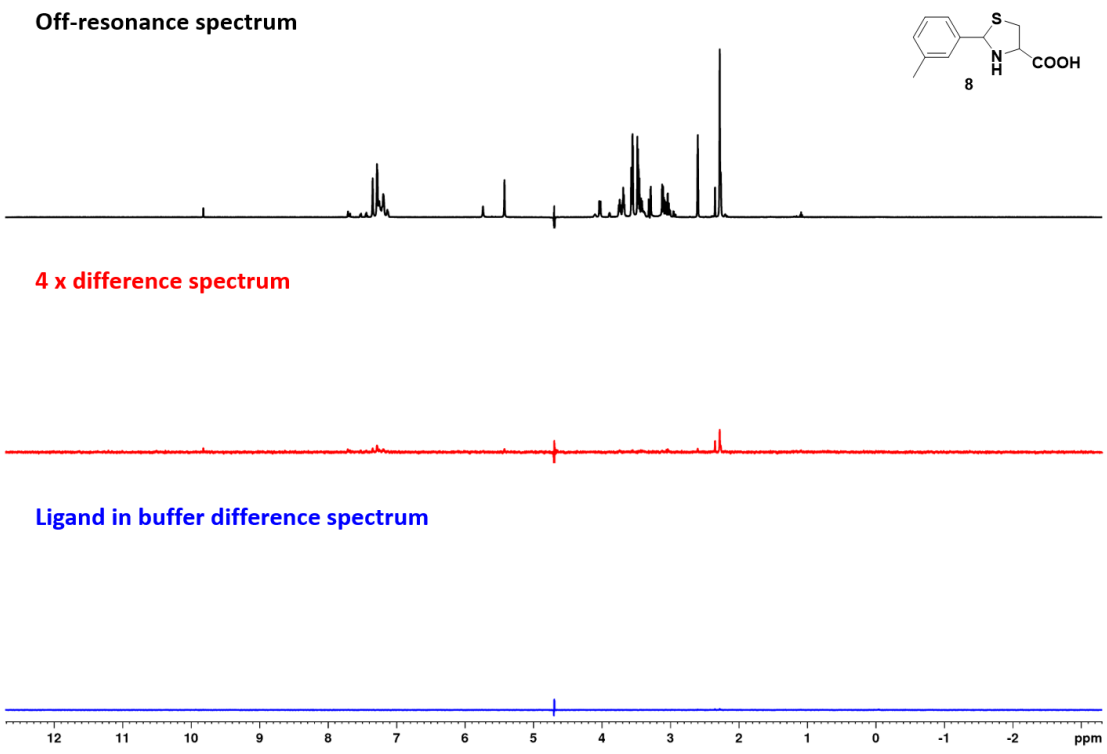


Figure S2h: Stacked STD spectra of compound **8** off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 4 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

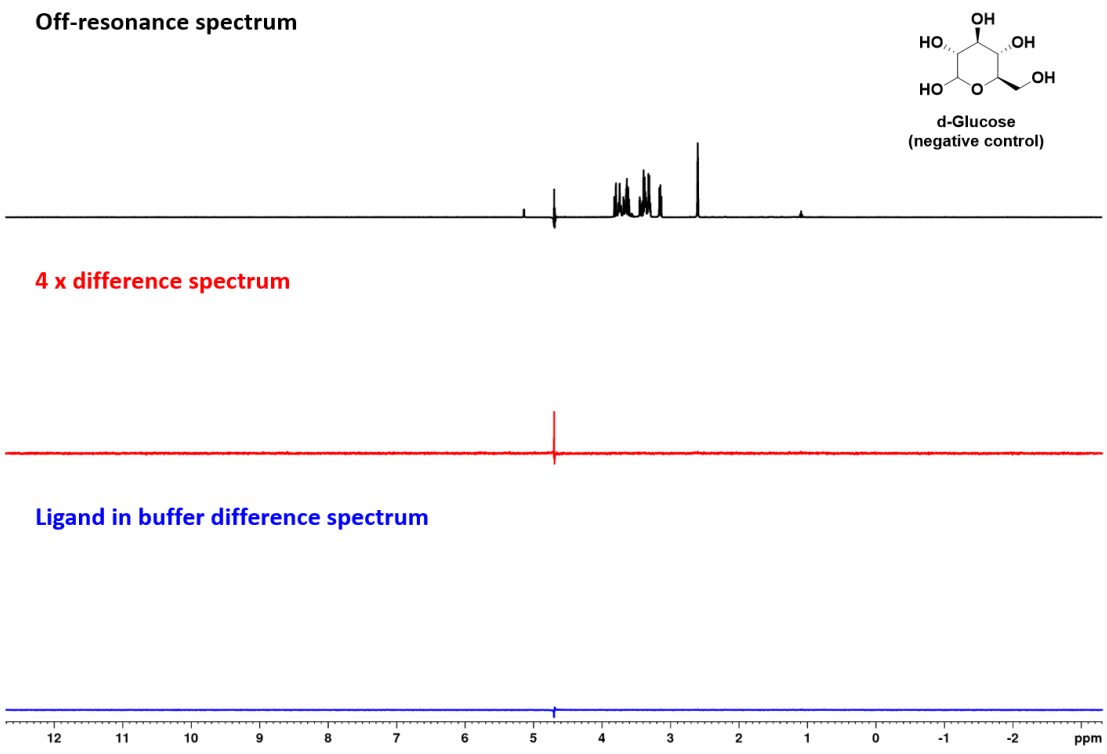


Figure S2i: Stacked STD spectra of compound d-glucose off-resonance (-40 ppm) spectrum (black, top), difference spectrum scaled 4 x (red, middle, on-resonance 0.7 ppm), and buffer control difference spectrum (blue, bottom).

2.2 Competitive WaterLOGSY with purchased hits

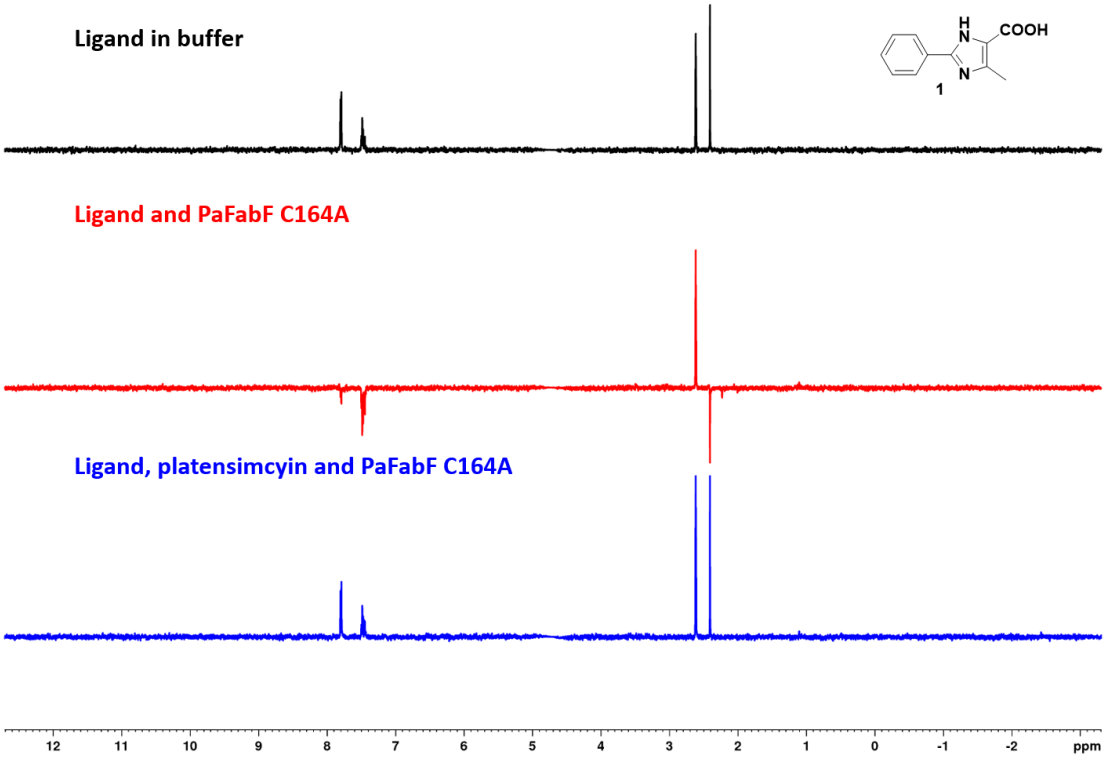


Figure S3a: Stacked WaterLOGSY spectra of compound **1** in buffer (black, top), with *Pa*FabF C164A (red, middle), and with both *Pa*FabF C164A and platensimycin (blue, bottom).

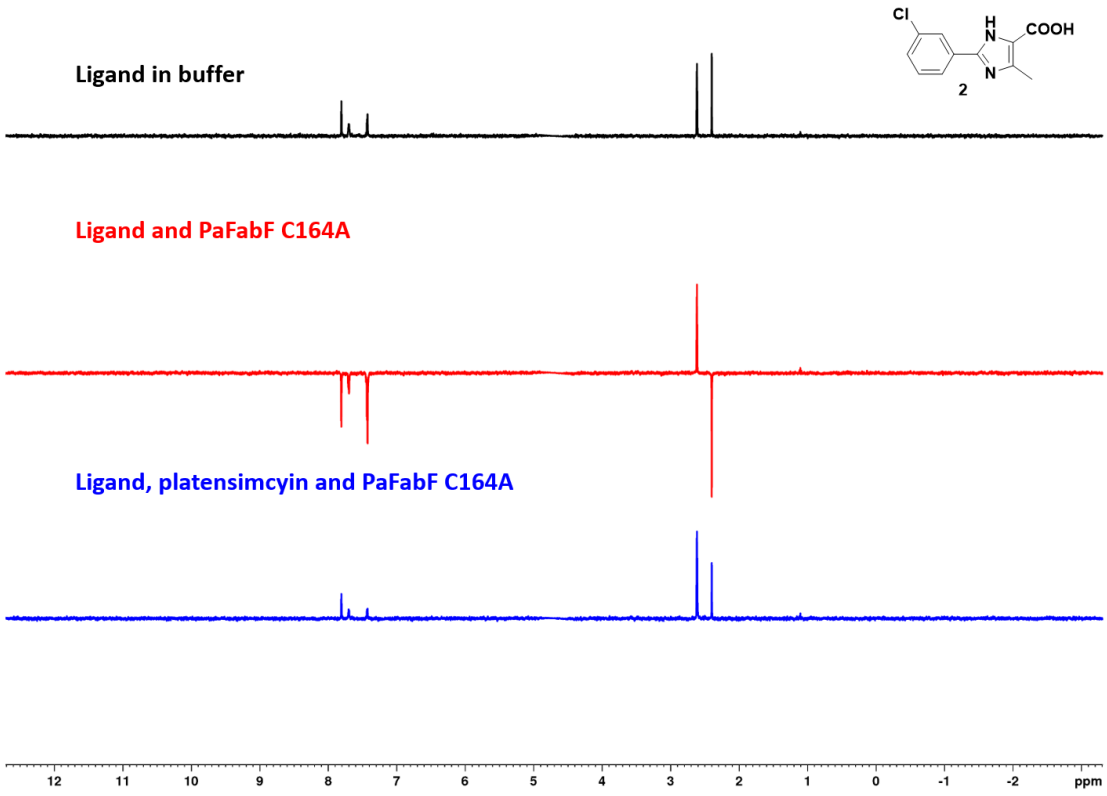


Figure S3b: Stacked WaterLOGSY spectra of compound **2** in buffer (black, top), with *Pa*FabF C164A (red, middle), and with both *Pa*FabF C164A and platensimycin (blue, bottom).

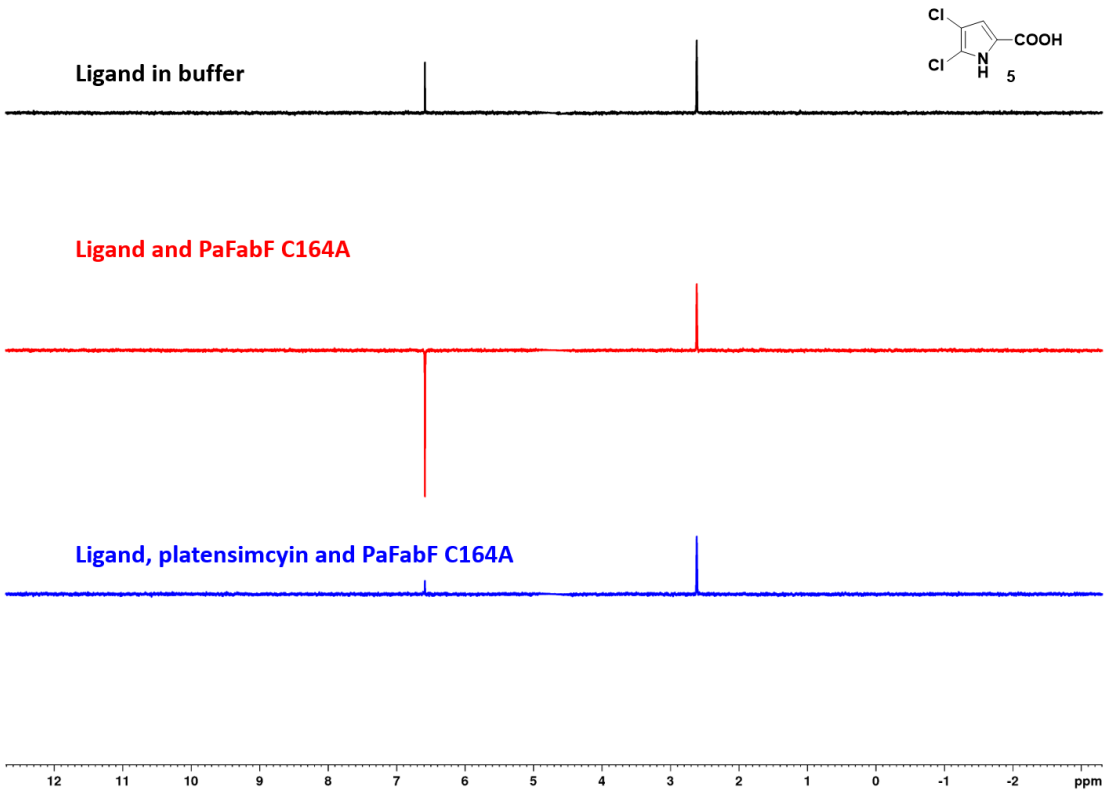


Figure S3c: Stacked WaterLOGSY spectra of compound **5** in buffer (black, top), with *Pa*FabF C164A (red, middle), and with both *Pa*FabF C164A and platensimycin (blue, bottom).

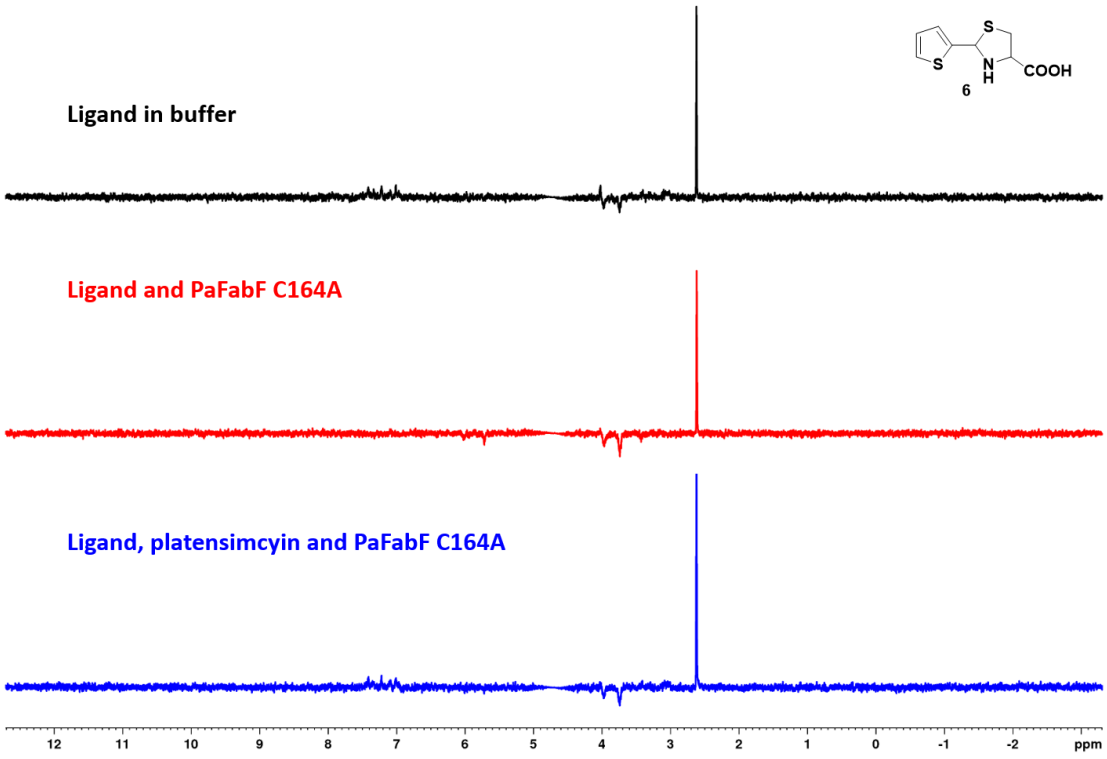


Figure S3d: Stacked WaterLOGSY spectra of compound **6** in buffer (black, top), with *Pa*FabF C164A (red, middle), and with both *Pa*FabF C164A and platensimycin (blue, bottom).

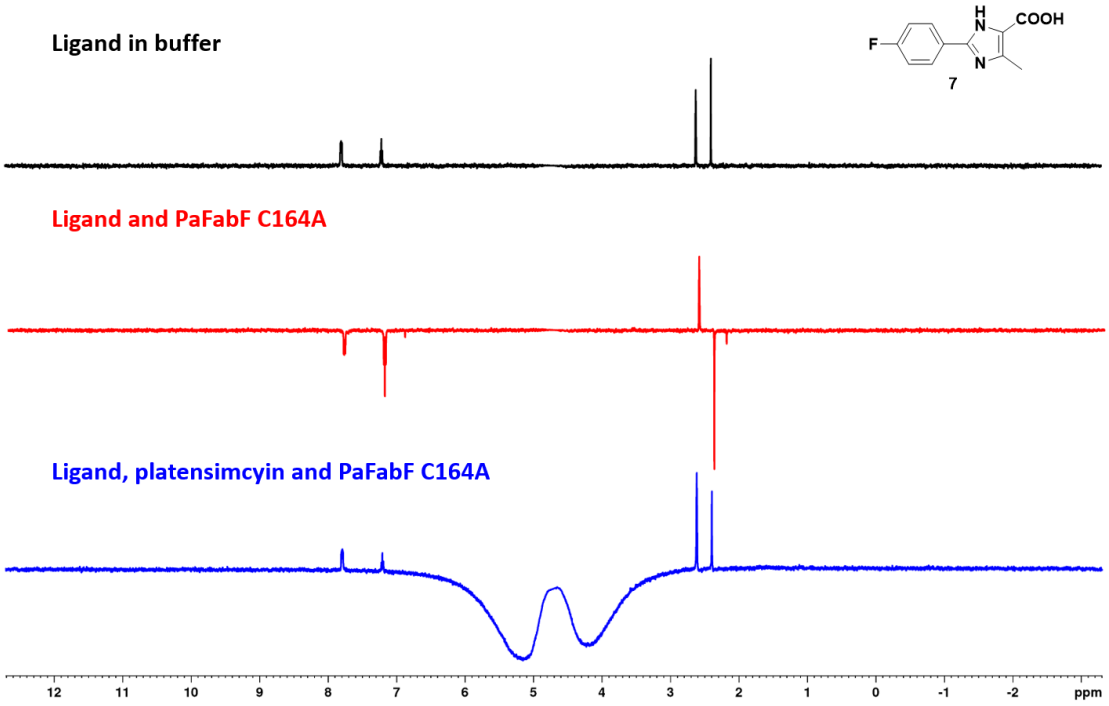


Figure S3e: Stacked WaterLOGSY spectra of compound **7** in buffer (black, top), with *Pa*FabF C164A (red, middle), and with both *Pa*FabF C164A and platensimycin (blue, bottom).

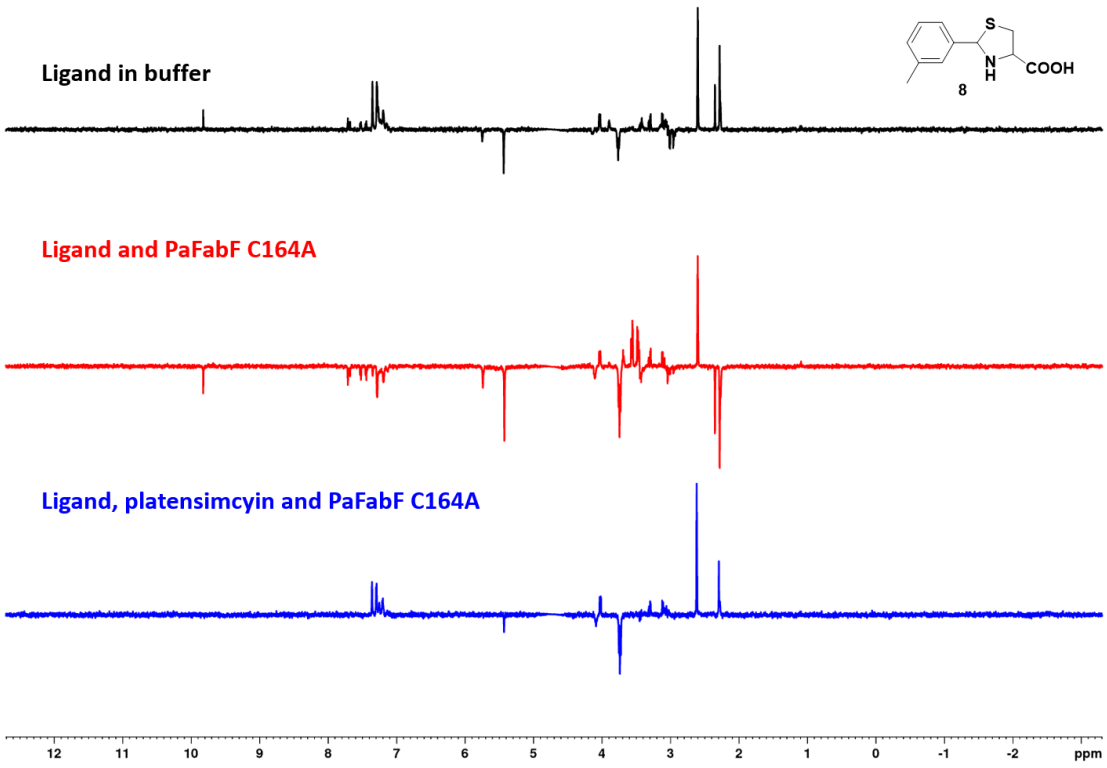


Figure S3f: Stacked WaterLOGSY spectra of compound **8** in buffer (black, top), with *Pa*FabF C164A (red, middle), and with both *Pa*FabF C164A and platensimycin (blue, bottom).

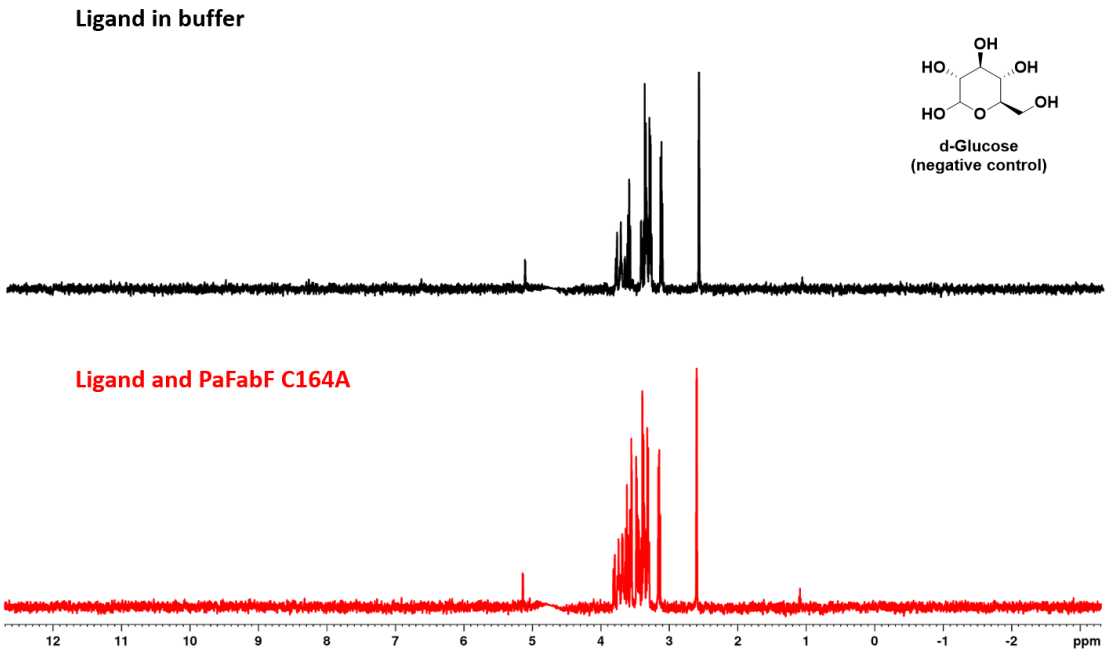


Figure S3e: Stacked WaterLOGSY spectra of negative control D-glucose in buffer (black, top) and with *Pa*FabF C164A (red, bottom).

2.3 WaterLOGSY for synthesized compounds

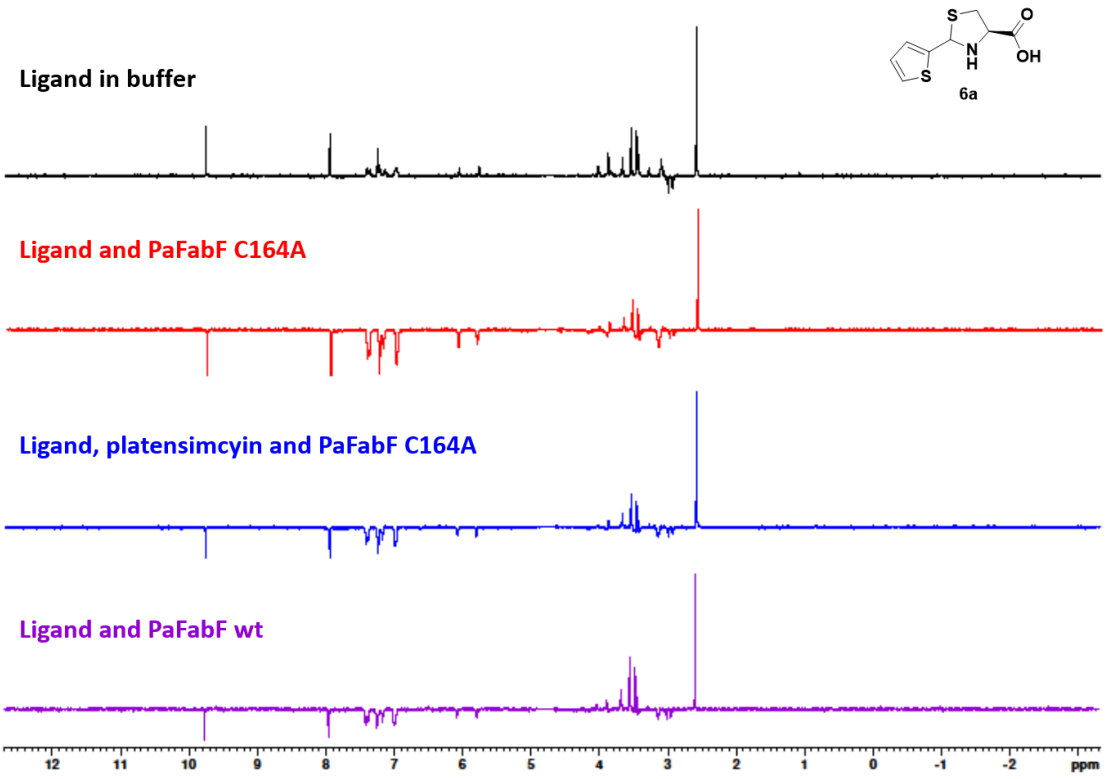


Figure S4a: Stacked WaterLOGSY spectra of compound **6a** in buffer (black, top), with *Pa*FabF C164A (red, upper middle), with both *Pa*FabF C164A and platensimycin (blue, lower middle) and with w. t. *Pa*FabF (purple, bottom).

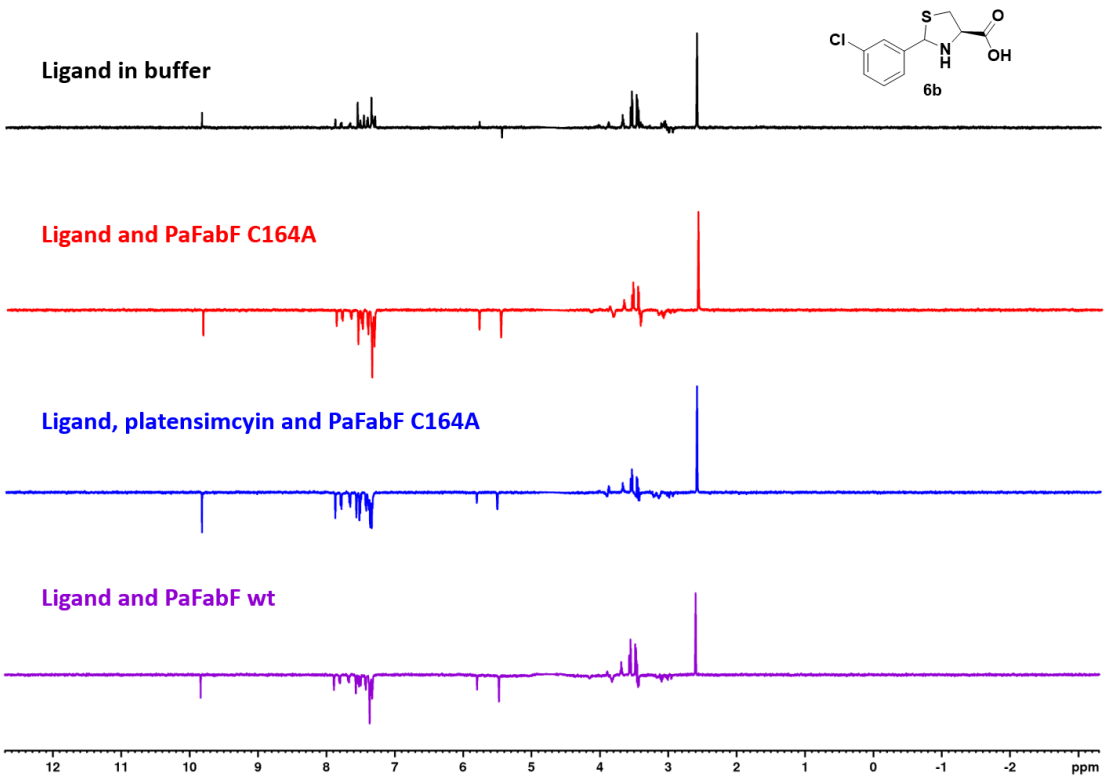


Figure S4b: Stacked WaterLOGSY spectra of compound **6b** in buffer (black, top), with *Pa*FabF C164A (red, upper middle), with both *Pa*FabF C164A and platensimycin (blue, lower middle) and with w. t. *Pa*FabF(purple, bottom).

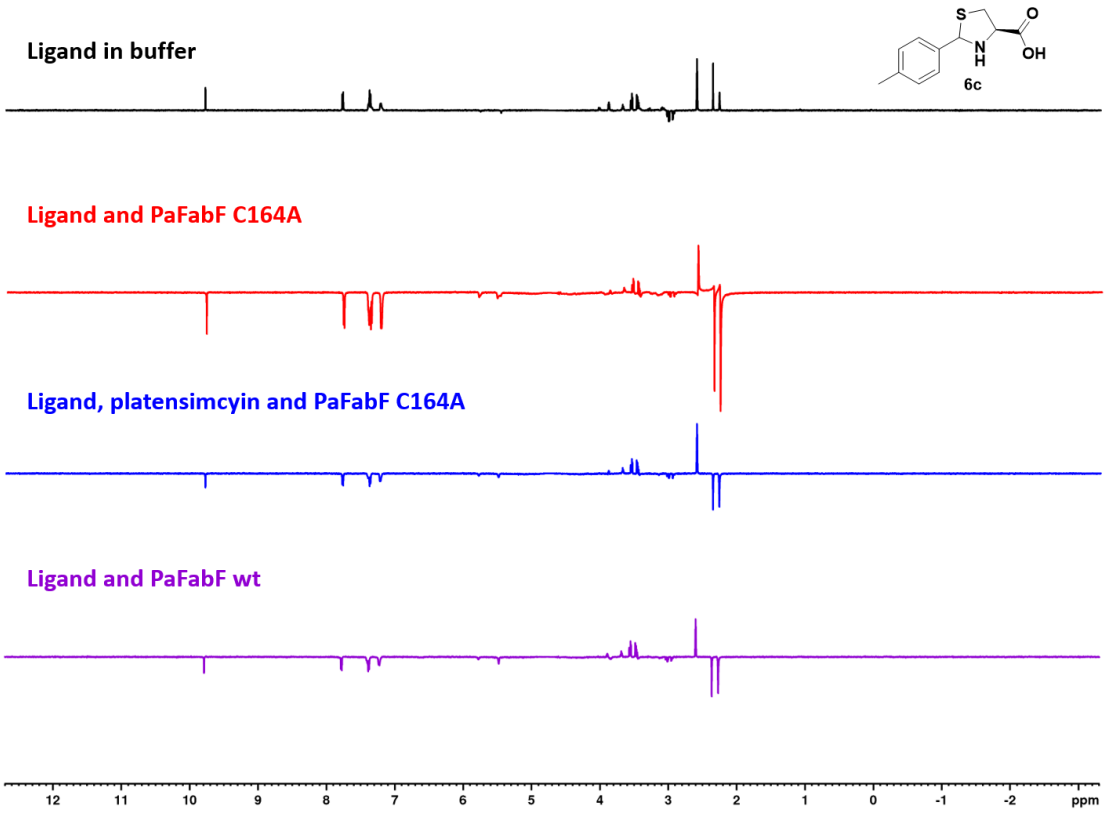


Figure S4c: Stacked WaterLOGSY spectra of compound **6c** in buffer (black, top), with *Pa*FabF C164A (red, upper middle), with both *Pa*FabF C164A and platensimycin (blue, lower middle) and with w. t. *Pa*FabF(purple, bottom).

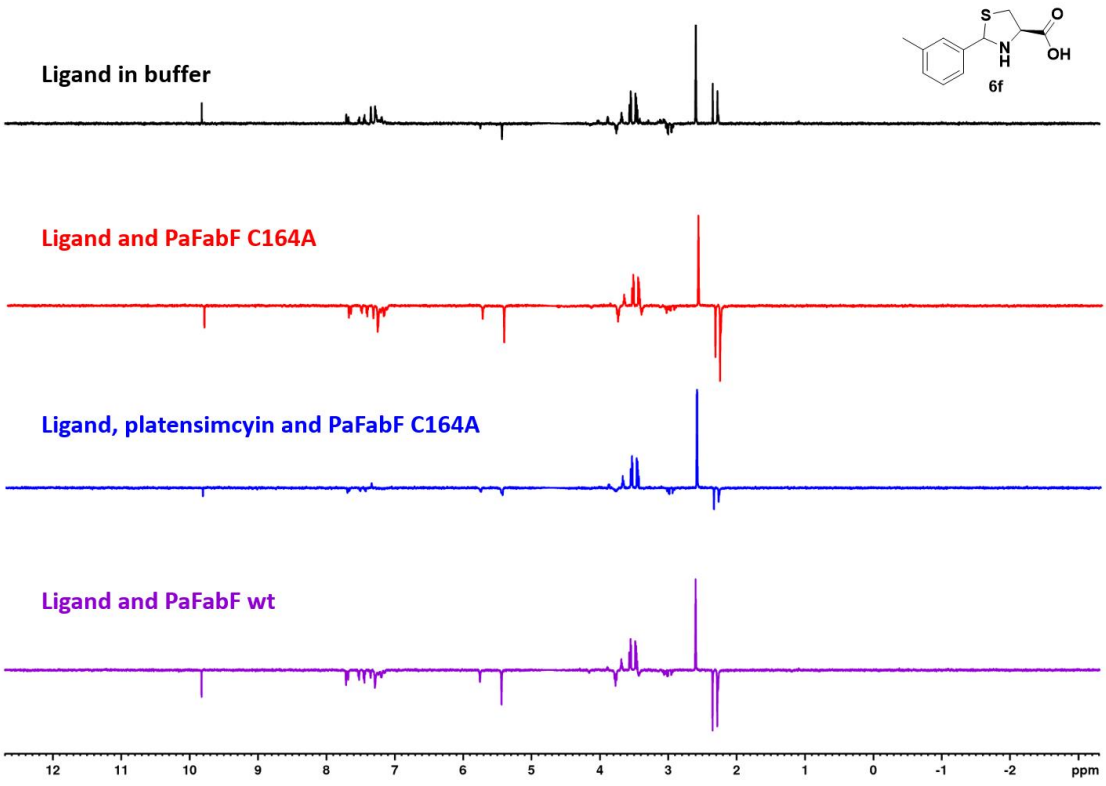


Figure S4d: Stacked WaterLOGSY spectra of compound **6f** in buffer (black, top), with *Pa*FabF C164A (red, upper middle), with both *Pa*FabF C164A and platensimycin (blue, lower middle) and with w. t. *Pa*FabF (purple, bottom).

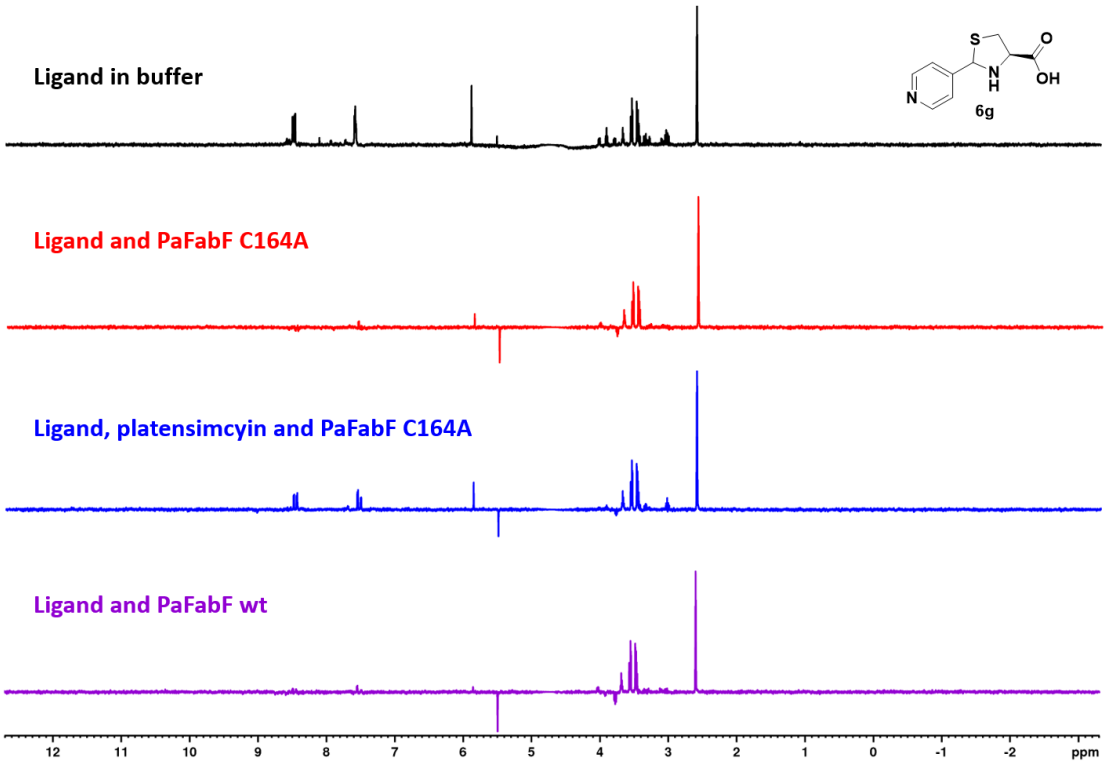


Figure S4e: Stacked WaterLOGSY spectra of compound **6g** in buffer (black, top), with *Pa*FabF C164A (red, upper middle), with both *Pa*FabF C164A and platensimycin (blue, lower middle) and with w. t. *Pa*FabF (purple, bottom).

3 Steady state analysis

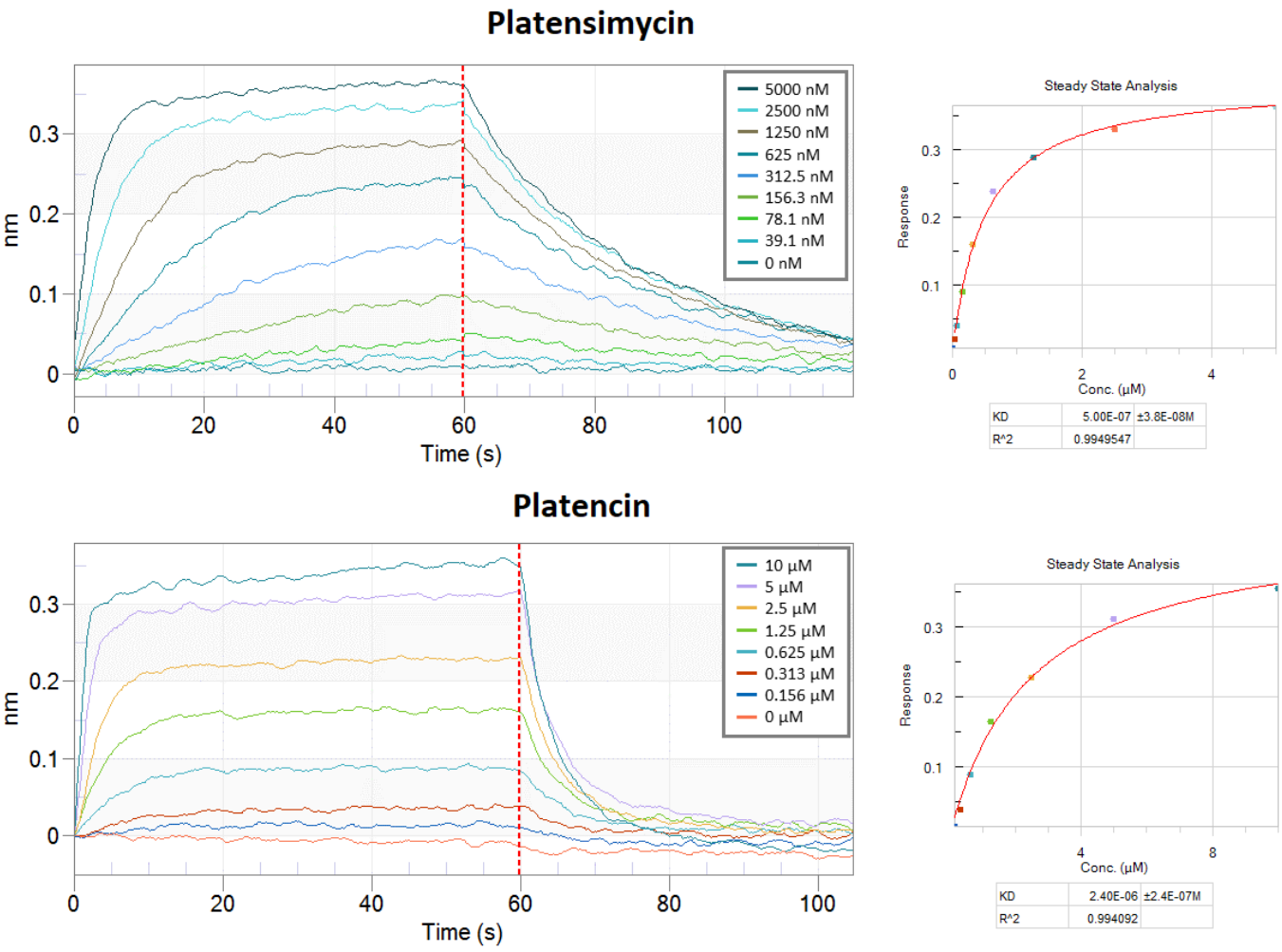


Figure S5: Steady state analysis of sensograms for platensimycin (top) and platencin (bottom) binding to PaFabF C164A together with calculated KD and R2 values.

4 Structure validation of hits and synthesis of analogues of hit 6

4.1 General Considerations

All chemicals have been purchased from commercial sources and were used without further purification unless otherwise noted. All solvents are reagent grade or HPLC grade. Yields refer to spectroscopically pure compounds after isolation. Most ^{1}H- and ^{13}C-NMR spectra were recorded in DMSO-d_6 using 400, 500 or 600 MHz (^{1}H) and 100 or 125 MHz (^{13}C). Chemical shifts (δ-values) are reported in ppm, spectra were calibrated related to solvents' residual proton chemical shifts (DMSO-d_6 δ = 2.5, D_2O δ = 4.79) and solvents' residual carbon chemical shifts (DMSO-d_6 δ = 49.0), multiplicity is reported as follows: s = singlet, brs = broad singlet, d = doublet, dd = doublet of a doublet, t = triplet, q = quartet, m = multiplet or unresolved and coupling constants *J* in Hz. Melting points (mp) were determined in open capillaries and are uncorrected. Infrared spectra (IR) were recorded on a 0.1 mm KBr demountable cell. High-resolution mass spectra (HRMS) were obtained by electrospray ionization (ESI) using a Q-TOF-Waters mass spectrometer (Xevo GS-XS model) in positive ion mode (M + H or M + Na) or by direct infusion on a JEOL AccuTOF T100GC with ESI in negative mode (M – H) as indicated.

4.2 Structure validation of hit compounds

4-Methyl-2-phenyl-1*H*-imidazole-5-carboxylic acid (1)
^{1}H-NMR (600 MHz, DMSO-d6): δ 7.98 (d, *J* = 7.6, 2H), 7.45 (t, *J* = 7.7, 2H), 7.38 (t, *J* = 7.5, 1H), 2.47 (s, 3H). HRMS (ESI): *m*/*z* calculated for $[M-H]^-$ $C_{11}H_9N_2O_2$ 201.0664, found 201.0660.

2-(3-Chlorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (2)
^{1}H NMR (600 MHz, DMSO-d6): δ 8.09 (t, *J* = 1.5, 2H), 7.97 (dt, *J* = 1.5, 7.6, 1H), 7.53-7.48 (m, 2H), 2.47 (s, 3H). HRMS (ESI): *m*/*z* calculated for $[M-H]^-$ $C_{11}H_8Cl_1N_2O_2$ 235.0274, found 235.0270.

4,5-Dichloro-1*H*-pyrrole-2-carboxylic acid (5)
^{1}H-NMR (600 MHz, DMSO-d6): δ 6.80 (s,1H). HRMS (ESI): *m*/*z* calculated for $[M-H]^-$ $C_5H_2Cl_2N_1O_2$ 177.9462, found 177.9458.

(2*RS*, 4*RS*)-2-(Thiophen-2-yl)thiazolidine-4-carboxylic acid (6)
^{1}H-NMR (600 MHz, DMSO-d6): δ 7.50 (dd, *J* = 5.2, 1.2, 0.4H), 7.42 (dd, *J* = 5.3, 1.3, 0.6H), 7.20 (d, *J* = 3.3, 0.4H), 7.06 (d, *J* = 3.6, 0.6H), 7.00 (dd, *J* = 5.0, 3.6, 0.4H), 6.95 (dd, *J* = 5.1, 3.6, 0.6H), 5.92 (s, 0.6H), 5.74 (s, 0.4H), 4.01 (t, J = 6.4, 0.6 H), 3.90 (dd, *J* = 7.1, 0.4H), 3.38-3.31 (m, 1H), 3.10-3.03 (m, 1H). HRMS (ESI): *m*/*z* calculated for $[M-H]^-$ $C_8H_8N_1O_2S_2$ 213.9996, found 213.9990.

2-(4-Fluorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (7)
^{1}H-NMR (600 MHz, H_2O/D_2O, 9:1): δ 7.78 (dd, *J* = 8.9, 5.3, 2H) 7.21 (t, *J* = 8.9, 2H), 2.40 (s, 1H). HRMS (ESI): *m*/*z* calculated for $[M-H]^-$ $C_{11}H_8F_1N_2O_2$ 219.0569, found 219.0564.

(2*RS*, 4*RS*)-2-(*m*-Tolyl)thiazolidine-4-carboxylic acid (8)
^{1}H-NMR (600 MHz, DMSO-d6): δ 7.32 (s, 0.5H), 7.28-7.26 (m, 1H), 7.25-7.24 (m, 1H), 7.21-7.20 (m, 1.5 H), 7.14 (d, *J* = 7.2, 0.5 H), 7.08 (3.6, 0.5H), 5.61 (s, 0.5H), 5.45 (s, 0.5H), 4.24 (dd, *J* = 7.1, 4.5, 0.5H), 3.90 (dd, *J* = 8.7, 7.2, 0.5H), 3.36 (dd, *J* = 10.1, 7.1, 0.5H), 3.28 (dd, *J* = 10.2, 7.1, 0.5H), 3.13 (dd, *J* = 10.3, 4.4, 0.5H), 3.05 (dd, *J* = 9.9, 8.9, 0.5H), 2.30 (s, 1.5H), 2.29 (s, 1.5H). HRMS (ESI): *m*/*z* calculated for $[M-H]^-$ $C_{11}H_{12}N_1O_2S_1$ 222.0582, found 222.0588.

4.3 General procedure for synthesis of substituted thiazolidine-4-carboxylic acids

To a stirred suspension of L-cysteine monohydrate hydrochloride (1.0 equiv.) in water (4 times vol.), was added NaOH (1 equiv.) at room temperature. To this mixture ethanol (5 times vol.) was added and after stirring for 15 min to make the mixture homogenous, aldehyde (1.0 equiv.) was added. The reaction mixture was stirred at room temperature overnight. After completion of reaction the solvent was removed under reduced pressure and the residue was triturated with hexane or pentane to get desired product (91-97%).

(2*RS*, 4*R*)-2-(Thiophen-2-yl)thiazolidine-4-carboxylic acid (6a)

Thiazolidine **2e** was obtained by using L-cysteine monohydrate.HCl (0.78 g, 4.5 mmol) in water (3.1 mL), NaOH (0.18 g, 4.5 mmol), ethanol (3.7 mL) and thiophene-2-carbaldehyde (0.5 g, 4.5 mmol) with a diastereomeric ratio 3:2 (0.5 g, 96%) as a light brown solid. mp = 177-179 °C. **^{1}H NMR** (300 MHz, DMSO-d_6): δ 7.52 (d, J = 5.0, 0.4H), 7.48 – 7.38 (m, 0.6H), 7.21 (d, J = 3.1, 0.4H), 7.07 (d, J = 3.2, 0.6H), 7.00 (dd, J = 5.0, 3.6, 0.4H), 6.95 (dd, J = 5.0, 3.6, 0.6H), 5.93 (s, 0.6H), 5.76 (s, 0.4H), 4.07 (t, J = 6.3, 0.6H), 3.96 – 3.87 (m, 0.4H), 3.36 – 3.30 (m, 1H), 3.15 – 2.98 (m, 1H). **^{13}C NMR** (75 MHz, DMSO-d_6): δ 172.7, 172.0, 147.0, 142.7, 127.9, 126.8, 126.2, 126.0, 125.5, 125.8, 66.5, 66.1, 65.2, 64.4, 39.5, 38.3, 38.0 ppm. IR (thin film): ν_{max} 1406, 1582, 2590, 2830, 3114, 3276, 3506 cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_8H_{10}O_2NS_2$ 216.0153, found 216.0147.

(2*RS*, 4*R*)-2-(3-Chlorophenyl)thiazolidine-4-carboxylic acid (6b)

Thiazolidine compound **6b** was obtained by using L-cysteine monohydrate.HCl (1.24 g, 7.1 mmol) in water (5.0 mL), NaOH (0.28 g, 7.1 mmol), ethanol (6.0 mL) and 3-chlorobenzaldehyde (1.0 g, 7.1 mmol) with a diastereomeric ratio 3:2 (1.0 g, 95.4%) as a white solid. mp = 195-198 °C. **^{1}H NMR** (300 MHz, DMSO-d_6): δ 7.65 (s, 0.4H), 7.50 (s, 0.6H), 7.48 – 7.44 (m, 0.4H), 7.39 (d, J = 5.3, 0.8H), 7.36 (s, 0.8H), 7.34 – 7.30 (m, 1.2H), 5.69 (s, 0.6H), 5.51 (s, 0.4H), 4.18 (dd, J = 6.7, 5.0, 0.6H), 3.90 (dd, J = 8.9, 7.0, 0.4H), 3.30 – 3.29 (m, 0.4H), 3.27 (s, 0.4H), 3.13 (d, J = 4.7, 0.6H), 3.07 (d, J = 8.8, 0.6H) ppm. **^{13}C NMR** (125 MHz, DMSO-d_6): δ 172.8, 144.4, 132.9, 130.1, 127.5, 126.6, 125.7, 69.9, 64.8, 38.0 ppm. IR (thin film): ν_{max} 1413, 1594, 3115, 3226, 3279, 3447, 3511 cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_{10}H_{11}O_2NClS$ 244.0199, found 244.0193.

(2*RS*, 4*R*)-2-(*p*-Tolyl)thiazolidine-4-carboxylic acid (6c)

Thiazolidine compound **6c** was obtained by using L-cysteine monohydrate.HCl (1.46 g, 8.3 mmol) in water (5.0 mL), NaOH (0.33 g, 8.3 mmol), ethanol (6.0 mL) and 4-methylbenzaldehyde (1.0 g, 8.3 mmol) with a diastereomeric ratio 1:1 (1.0 g, 94.8%) as a white solid. mp = 198.5-200 °C. **^{1}H NMR** (300 MHz, DMSO-d_6): δ 7.39 (d, J = 7.9, 1H), 7.32 (d, J = 7.9, 1H), 7.16 (dd, J = 12.6, 8.0, 2H), 5.61 (s, 0.5H), 5.46 (s, 0.5H), 4.24 (dd, J = 6.9, 4.3, 0.5H), 3.94 – 3.82 (m, 0.5H), 3.28 (m, J = 10.2, 7.2, 1H), 3.18 – 3.02 (m, 1H), 2.30 (s, 1.5H), 2.28 (s, 1.5H) ppm. **^{13}C NMR** (125 MHz, DMSO-d_6): δ 173.0, 138.0, 136.8, 128.7, 126.9, 71.0, 64.8, 37.9, 20.7 ppm. IR (thin film): ν_{max} 1412, 1490, 1583, 2962, 2996, 3052 cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_{11}H_{14}O_2NS$ 224.0741, found 224.0739.

(2*RS*, 4*R*)-2-(4-(Diethylamino)phenyl)thiazolidine-4-carboxylic acid (6d)

Thiazolidine compound **6d** was obtained by using L-cysteine monohydrate.HCl (0.34 g, 2.8 mmol) in water (2 mL), NaOH (0.11 g, 2.8 mmol), ethanol (2.3 mL) and 4-(diethylamino)benzaldehyde (0.5 g, 2.8 mmol) with a diastereomeric ratio 10:1 (0.5 g, 95.4%) as a white solid. mp = 192-194 °C. **^{1}H NMR** (400 MHz, D_2O) δ 7.64 (d, J = 8.7, 2H), 7.44 (d, J = 8.7, 2H), 5.23 (s, 1H), 4.07 (dd, J = 6.6, 5.0, 1H), 3.95 (dd, J = 6.7, 4.5, 1H), 3.51 (q, J = 7.3, 4H), 3.10 (dd, J = 7.0, 3.7, 2H), 0.96 (t, J = 7.2, 6H). **^{13}C NMR** (100 MHz, D2O) δ 170.8, 170.6, 140.8, 136.8, 132.1, 130.1, 123.1, 53.8, 52.7, 52.7, 51.9, 36.4, 32.2, 32.1, 9.6 ppm. IR (thin film): ν_{max} 773, 1412, 1487, 1609, 2589, 2800, 2655, 3118, 3445, 3496, 3513 cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_{14}H_{20}N_2O_2S$ 281.0968, found 281.0960.

(2*RS*, 4*R*)-2-([1,1'-Biphenyl]-4-yl)thiazolidine-4-carboxylic acid (6e)

Thiazolidine compound **6e** was obtained by using L-cysteine monohydrate.HCl (0.96 g, 5.5 mmol) in water (4 mL), NaOH (0.22 g, 5.5 mmol), ethanol (4.6 mL) and [1,1'-biphenyl]-4-carbaldehyde (1.0 g, 5.5 mmol) with a diastereomeric ratio 3:2 (1.0 g, 91.0%) as a white solid. mp = 199-201 °C. **^{1}H NMR** (500 MHz, DMSO-*d6*): *δ* 7.67 (dd, *J* = 7.6, 5.0, 3H), 7.62 (dd, *J* = 12.1, 8.3, 2H), 7.53 (d, *J* = 8.3, 1H), 7.47 (dt, *J* = 13.0, 6.5, 2H), 7.37 (q, *J* = 7.3, 1H), 5.74 (s, 0.6H), 5.56 (s, 0.4H), 4.25 (dd, *J* = 7.0, 4.7, 0.6H), 3.93 (dd, *J* = 8.6, 7.2, 0.4H), 3.40 (dd, *J* = 10.1, 7.1, 0.6H), 3.33 (dd, *J* = 10.2, 7.1, 0.6H), 3.18 – 3.06 (m, 0.4H), 2.92 (m, *J* = 20.3, 14.2, 5.1, 0.6H) ppm. **^{13}C NMR** (125 MHz, DMSO-*d6*): *δ* 172.9, 140.5, 139.8, 138.1, 128.9, 127.9, 127.5, 126.8, 126.7, 126.6, 126.5, 70.8, 64.9, 38.0 ppm. IR (thin film): ν_{max} 1411, 1591, 2997, 3107, 3176, 3204 cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_{16}H_{15}NO_2S$ 286.0488, found 286.0492.

(2*RS*, 4*R*)-2-(m-Tolyl)thiazolidine-4-carboxylic acid (6f)

Thiazolidine compound **2f** was obtained by using L-cysteine monohydrate.HCl (0.29 g, 1.7 mmol) in water (1.2 mL), NaOH (0.07 g, 1.7 mmol), ethanol (1.3 mL) and 3-methylbenzaldehyde (0.2 g, 1.7 mmol) with diastereomeric ratio 1.2:1 (0.2 g, 96.5%) as a white solid. mp = 183-185 °C. **^{1}H NMR** (500 MHz, DMSO-*d6*): *δ* 7.34 (s, 0.5H), 7.30 – 7.26 (m, 1H), 7.25 – 7.20 (m, 1.5H), 7.14 (d, *J* = 7.2, 0.6H), 7.08 (d, *J* = 3.7, 0.4H), 5.62 (s, 0.5H), 5.46 (s, 0.5H), 4.24 (dd, *J* = 7.1, 4.5, 0.5H), 3.89 (dd, *J* = 8.7, 7.2, 0.4H), 3.37 (dd, *J* = 10.1, 7.2, 0.5H), 3.29 (dd, *J* = 10.2, 7.1, 0.5H), 3.14 (dd, *J* = 10.2, 4.4, 0.5H), 3.07 (dd, *J* = 10.0, 8.9, 0.5H), 2.32 (s, 1.5H), 2.31 (s, 1.5H) ppm. **^{13}C NMR** (100 MHz, DMSO-*d$_6$*): *δ* 173.0, 141.0, 137.4, 128.9, 128.1, 127.7, 124.0, 71.1, 64.9, 38.0, 21.0 ppm. IR (thin film): ν_{max} 1410, 1484, 1592, 2587, 2787, 2876, 2979 3115, 3382, 3461cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_{11}H_{14}O_2NS$ 224.0740, found 224.0739.

(2*RS*, 4*R*)-2-(Pyridin-4-yl)thiazolidine-4-carboxylic acid (6g)

Thiazolidine compound **2h** was obtained by using L-cysteine monohydrate.HCl (0.49 g, 2.8 mmol) in water (2 mL), NaOH (0.11 g, 2.8 mmol), ethanol (2.8 mL) and isonicotinaldehyde (0.3 g, 2.8 mmol) with a diastereomeric ratio 4:1 (0.3 g, 92.3%) as a brown solid. mp = 203-205 °C. **^{1}H NMR** (400 MHz, DMSO) δ 8.55 (dd, *J* = 4.4, 1.6, 0.4H), 8.49 (d, *J* = 6.1, 1.6H), 7.50 (d, *J* = 6.1, 0.4H), 7.38 (d, *J* = 6.0, 1.6H), 5.77 (s, 0.8H), 5.52 (s, 0.2H), 4.03 (t, *J* = 6.4, 1.2H), 3.89 (dd, *J* = 8.7, 7.0, 0.6H), 3.35 (dd, *J* = 10.0, 7.0, 0.4H), 3.28 (dd, *J* = 10.2, 6.8, 1H), 3.02 (dd, *J* = 10.2, 6.1, 1H) ppm. **^{13}C NMR** (100 MHz, DMSO) *δ* 172.5, 149.6, 121.6, 69.2, 65.0, 38.2, 35.3 ppm. IR (thin film): ν_{max} 1411, 1492, 1593, 2870, 2913, 2994, 3118, 3175, 3493, 3631 cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_{27}H_{27}N_2O_6$ 211.0541, found 211.0535.

(2*RS*, 4*R*)-2-(1H-Imidazol-4-yl)thiazolidine-4-carboxylic acid (6h)

Thiazolidine compound **6h** was obtained by using L-cysteine monohydrate.HCl (0.73 g, 4.2 mmol) in water (2.9 mL), NaOH (0.17 g, 4.2 mmol), ethanol (3.4 mL) and 1H-imidazole-4-carbaldehyde (0.4 g, 4.2 mmol) with a diastereomeric ratio 3:2 (0.4 g, 97.2%) as a light brown solid. mp = 119-201 °C. **^{1}H NMR** (400 MHz, DMSO-*d$_6$*): *δ* 7.61 (t, *J* = 8.6, 0.4H), 7.56 (d, *J* = 0.9, 0.6H), 7.10 (d, *J* = 0.8, 0.4H), 6.99 (s, 0.6H), 5.69 (s, 0.4H), 5.52 (s, 0.6H), 4.21 (dd, *J* = 6.8, 5.1, 0.6H), 3.68 (dd, *J* = 8.8, 7.1, 0.4H), 3.29 (dd, *J* = 9.9, 7.0, 0.4H), 3.23 (dd, *J* = 10.1, 7.1, 0.6H), 3.03 (dd, *J* = 10.1, 4.9, 0.4H), 2.87 (t, *J* = 9.4, 0.4H) ppm. **^{13}C NMR** (100 MHz, DMSO-*d$_6$*): *δ* 172.7, 149.6, 121.6, 69.2, 65.2, 38.3 ppm. IR (thin film): ν_{max} 1410, 1484, 1592, 2587, 2787, 2979, 3115, 3214, 3382 cm^{-1}. HRMS (ESI): *m*/*z* calculated for $[M+H]^+$ $C_7H_{10}O_2N_3S$ 200.0492, found 200.0488.

4.4 ^{1}H NMR spectra of purchased hit compounds

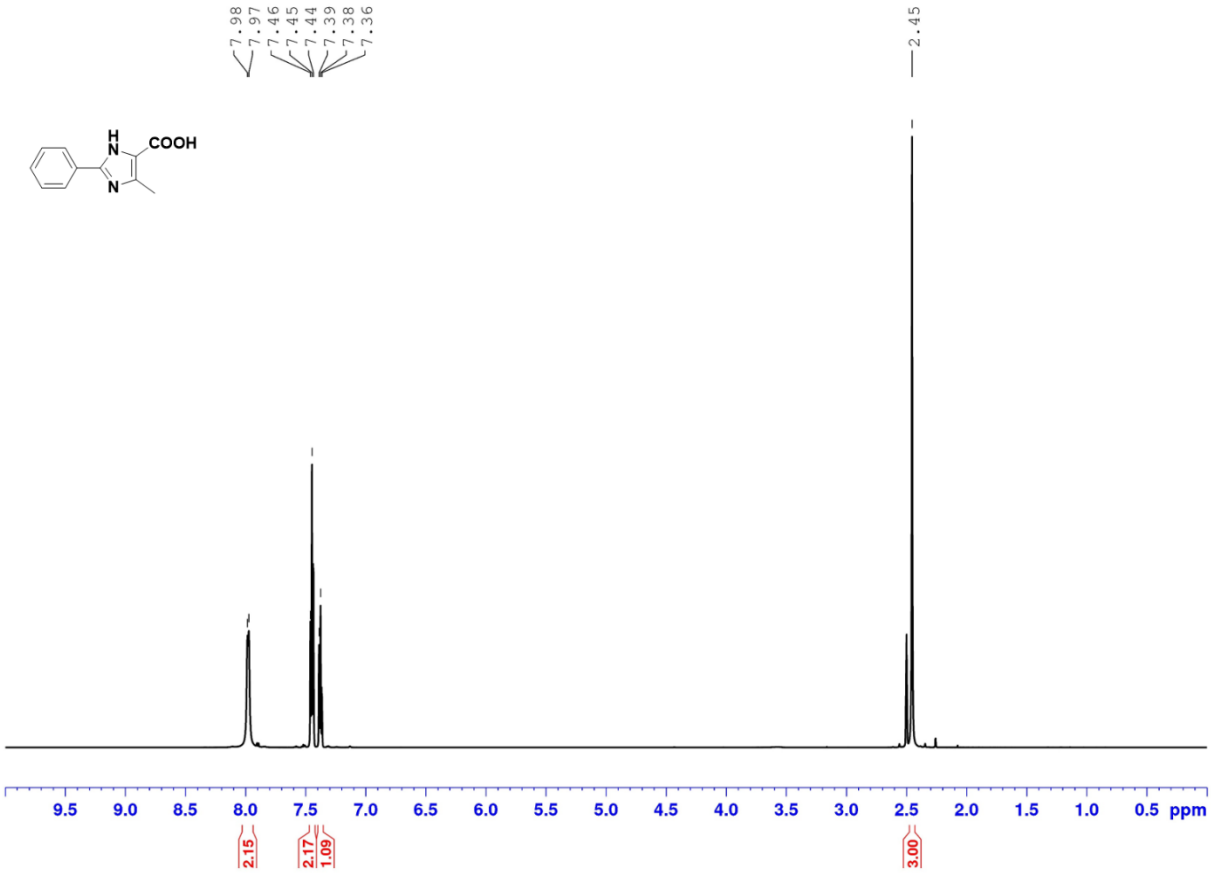


Figure S6a: ^{1}H NMR spectrum of 4-methyl-2-phenyl-1*H*-imidazole-5-carboxylic acid (**1**).

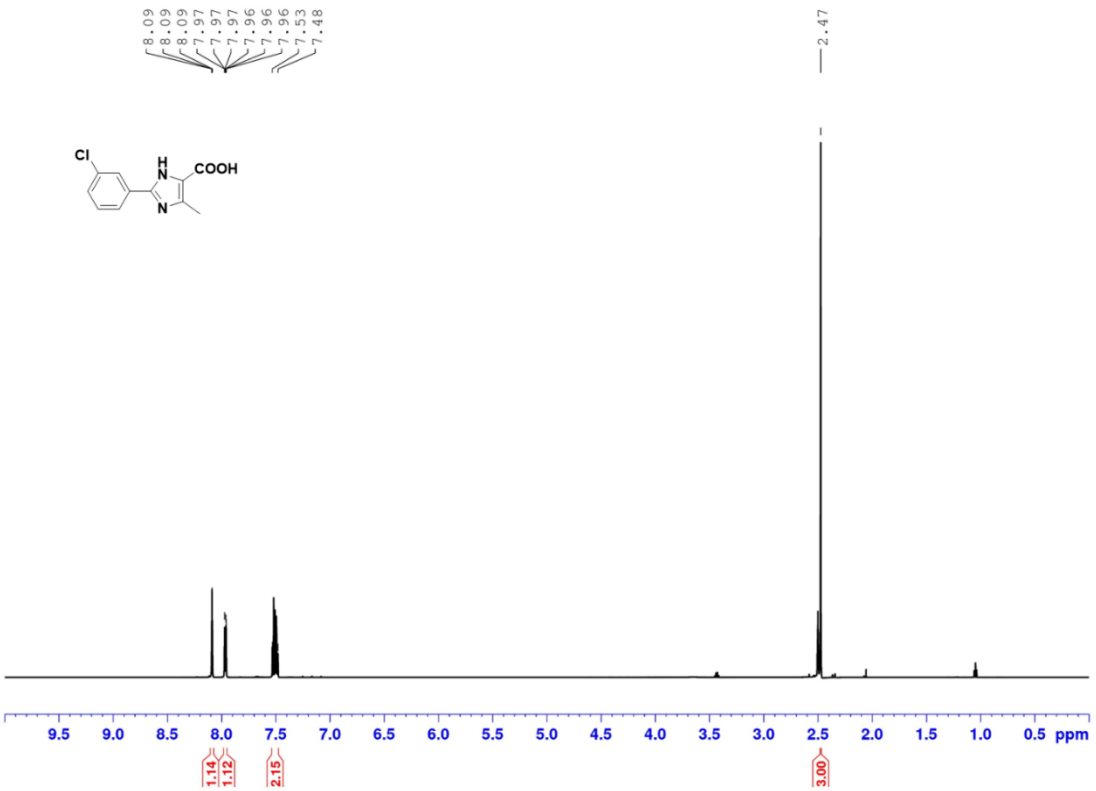


Figure S6b: [1]H NMR spectrum of 2-(3-chlorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (**2**).

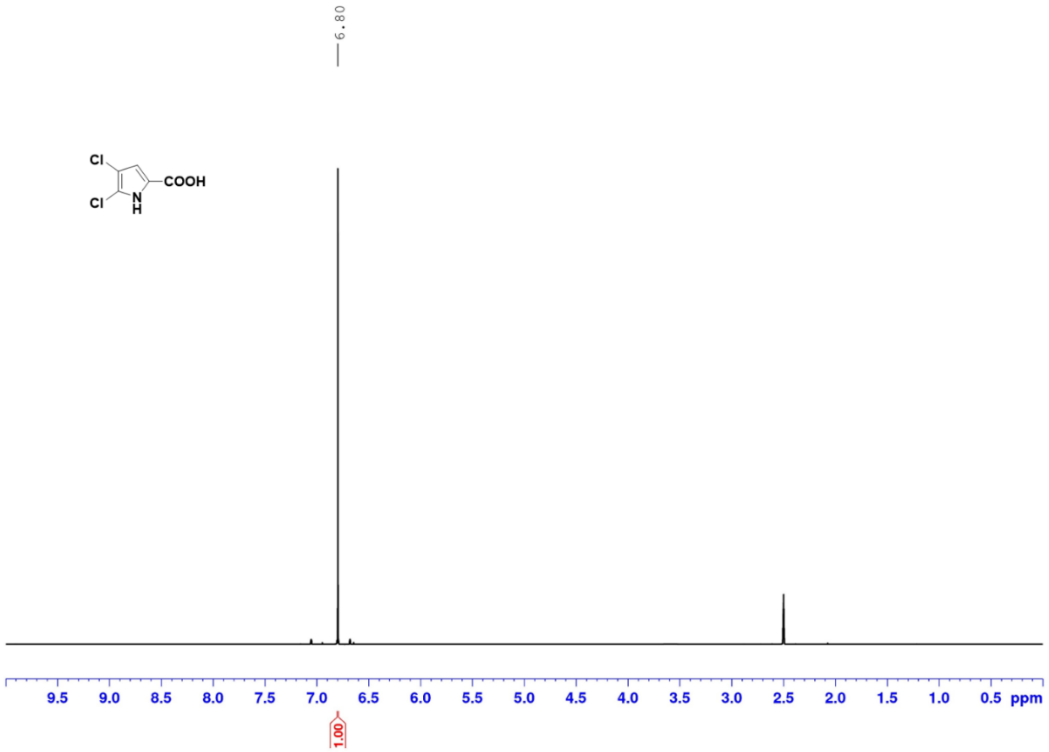


Figure S6c: [1]H NMR spectrum of 4,5-dichloro-1H-pyrrole-2-carboxylic acid (**5**).

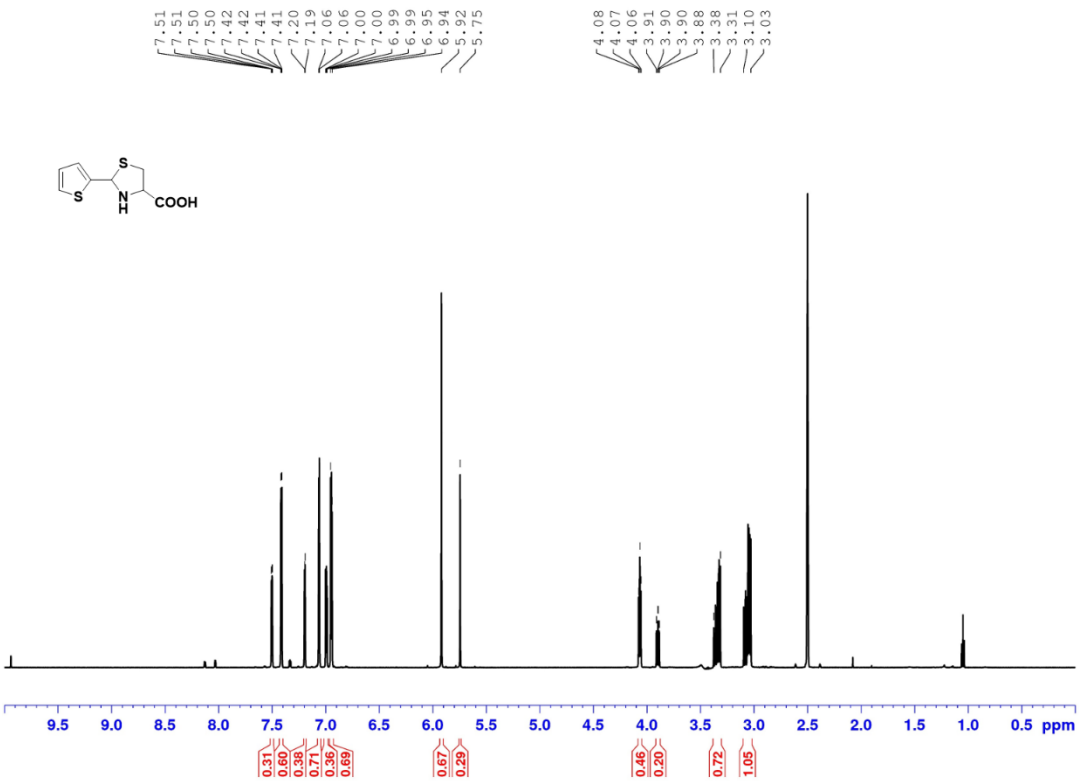


Figure S6d: [1]H NMR spectrum of (2*RS*,4*RS*)-2-(thiophen-2-yl)thiazolidine-4-carboxylic acid (**6**).

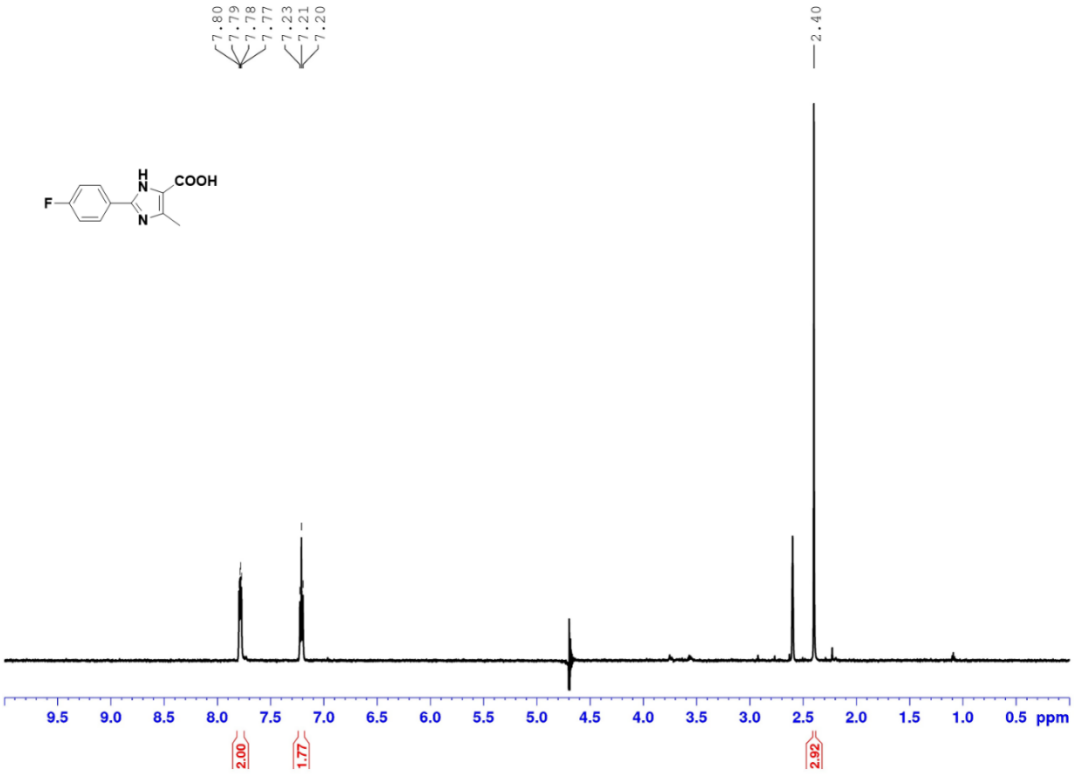


Figure **S6e:** [1]H NMR spectrum 2-(4-fluorophenyl)-4-methyl-1*H*-imidazole-5-carboxylic acid (**7**).

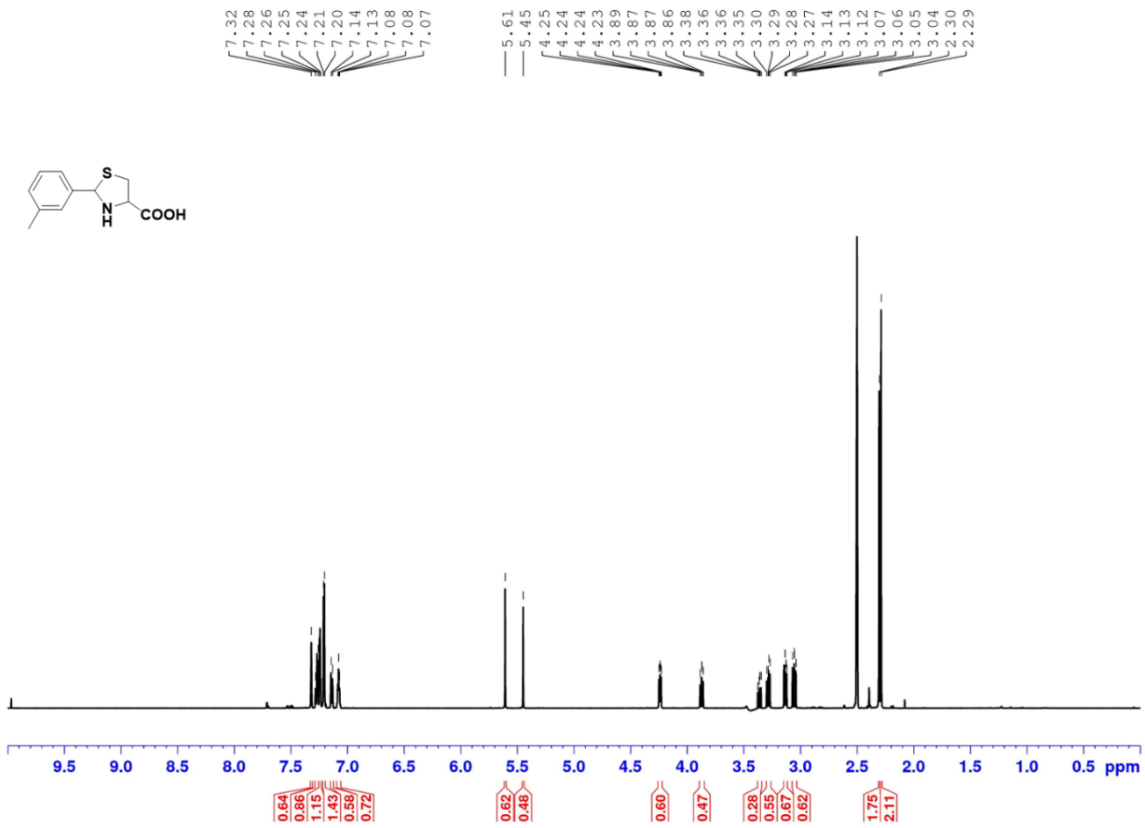


Figure S6f: ^{1}H NMR spectrum of (2RS, 4RS)-2-(*m*-tolyl)thiazolidine-4-carboxylic acid (**8**).

4.5 ^{1}H and ^{13}C NMR spectra of synthesized compound

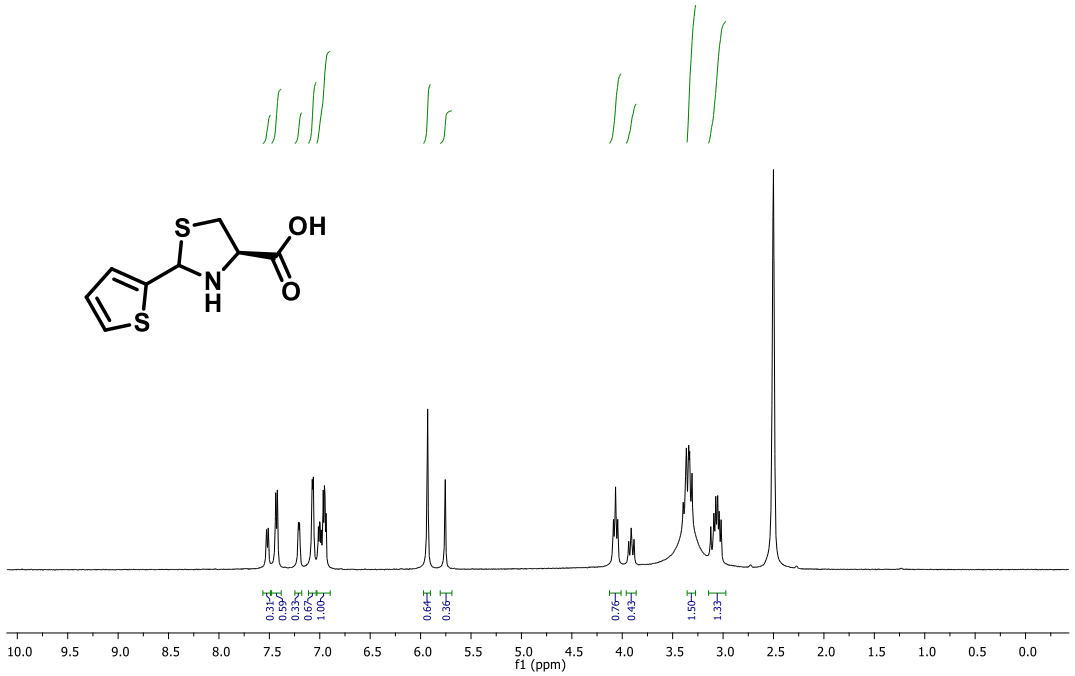


Figure S7a: 1H NMR spectrum of (2RS, 4R)-2-(thiophen-2-yl)thiazolidine-4-carboxylic acid (**6a).**

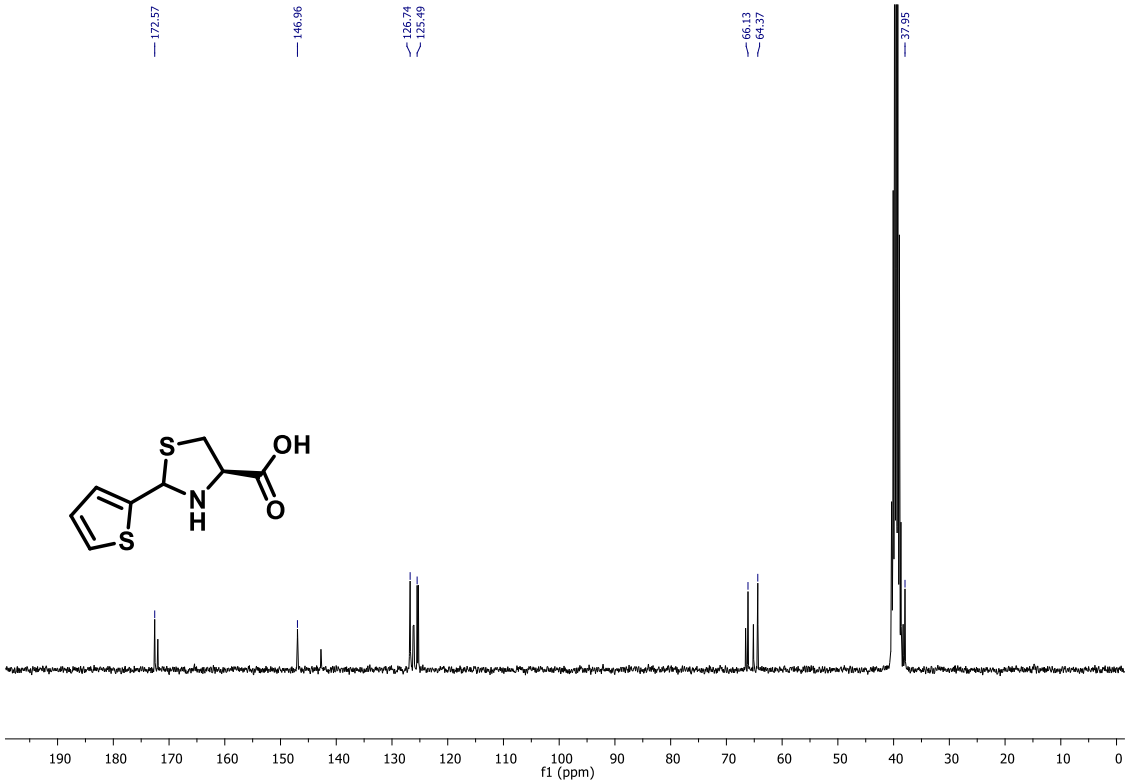


Figure S7b: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(thiophen-2-yl)thiazolidine-4-carboxylic acid (**6a**).

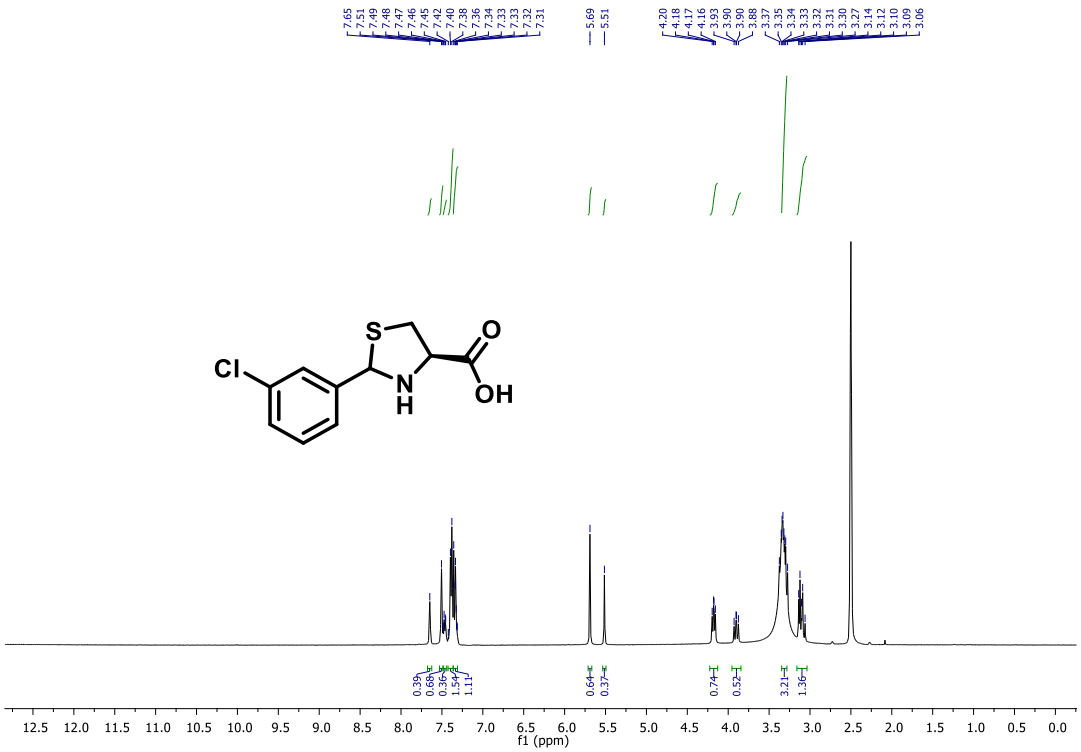


Figure S7c: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(3-chlorophenyl)thiazolidine-4-carboxylic acid (**6b**).

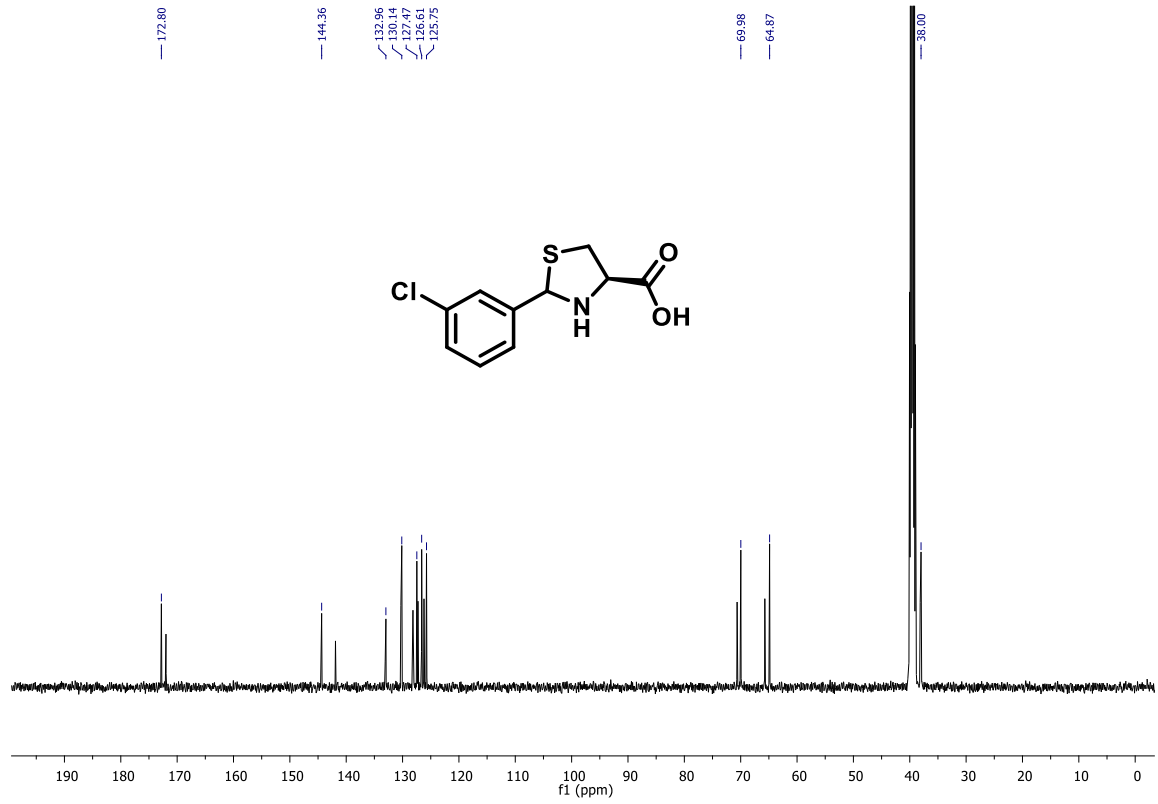


Figure S7d: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(3-chlorophenyl)thiazolidine-4-carboxylic acid (**6b**).

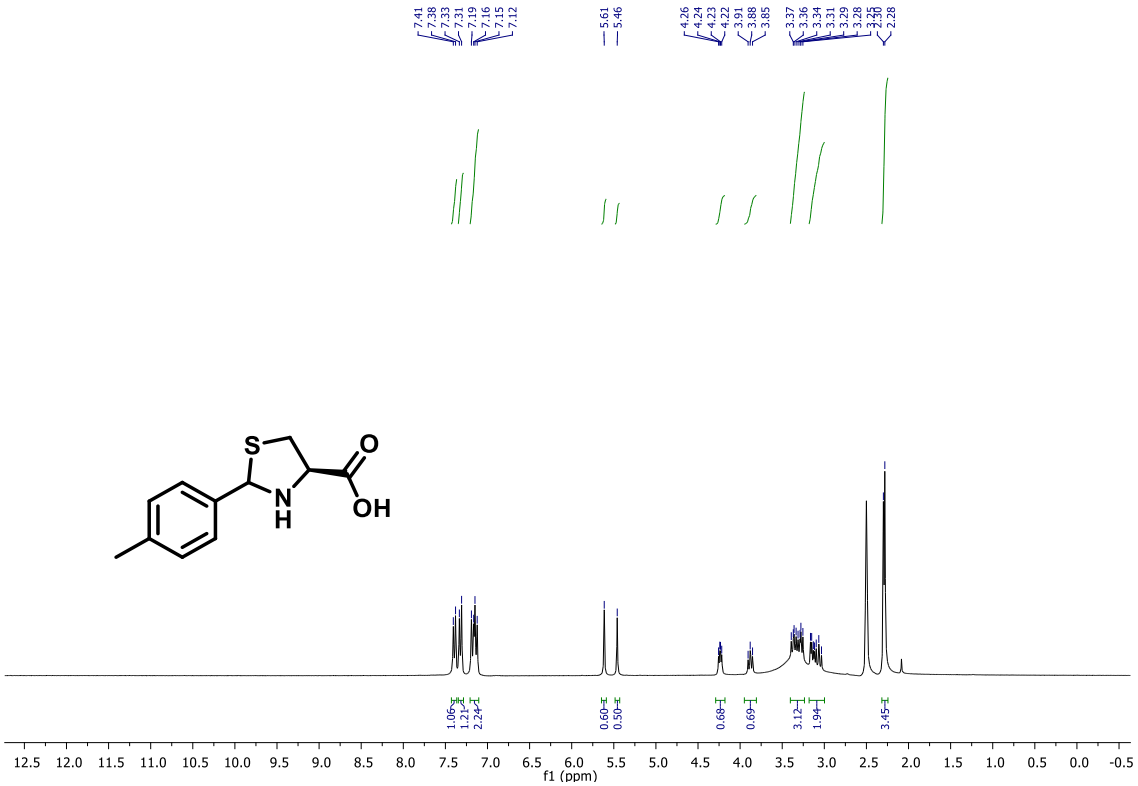


Figure S7e: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(*p*-tolyl)thiazolidine-4-carboxylic acid (**6c**).

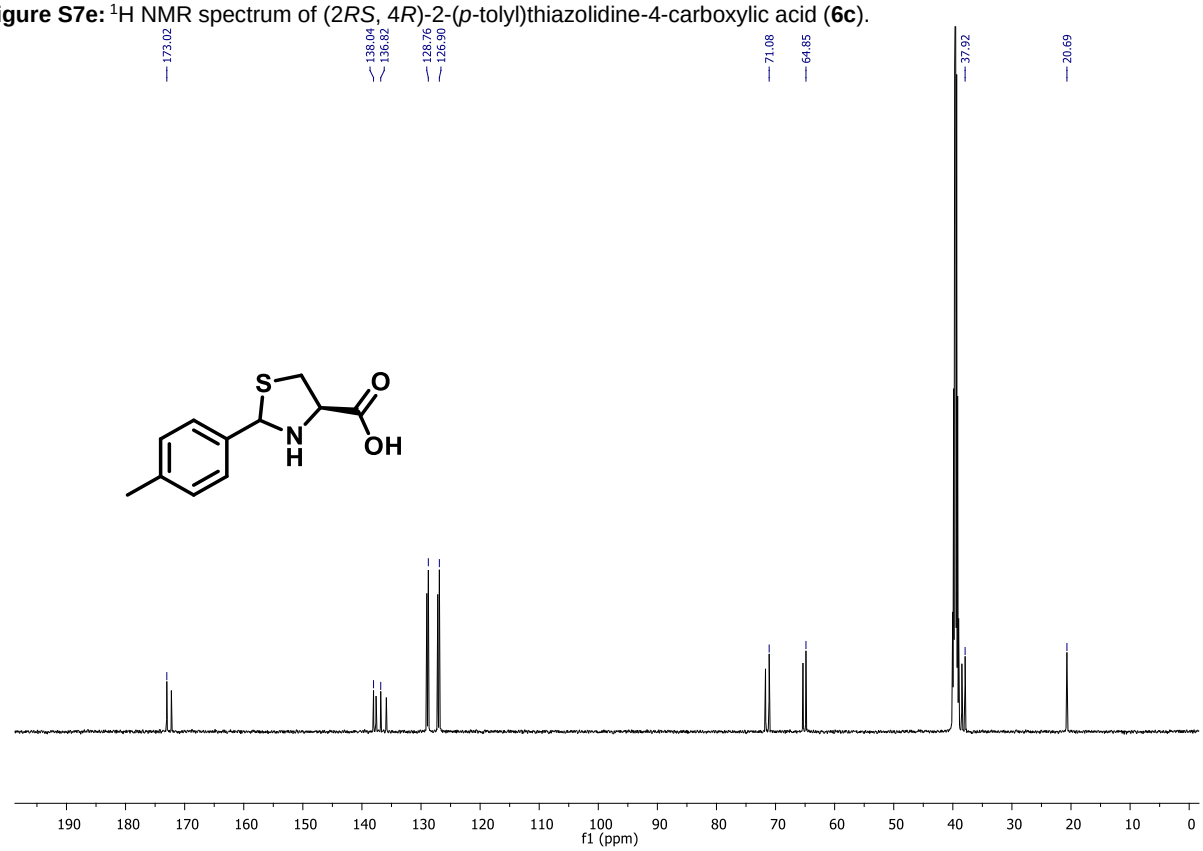


Figure S7f: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(*p*-tolyl)thiazolidine-4-carboxylic acid (**6c**).

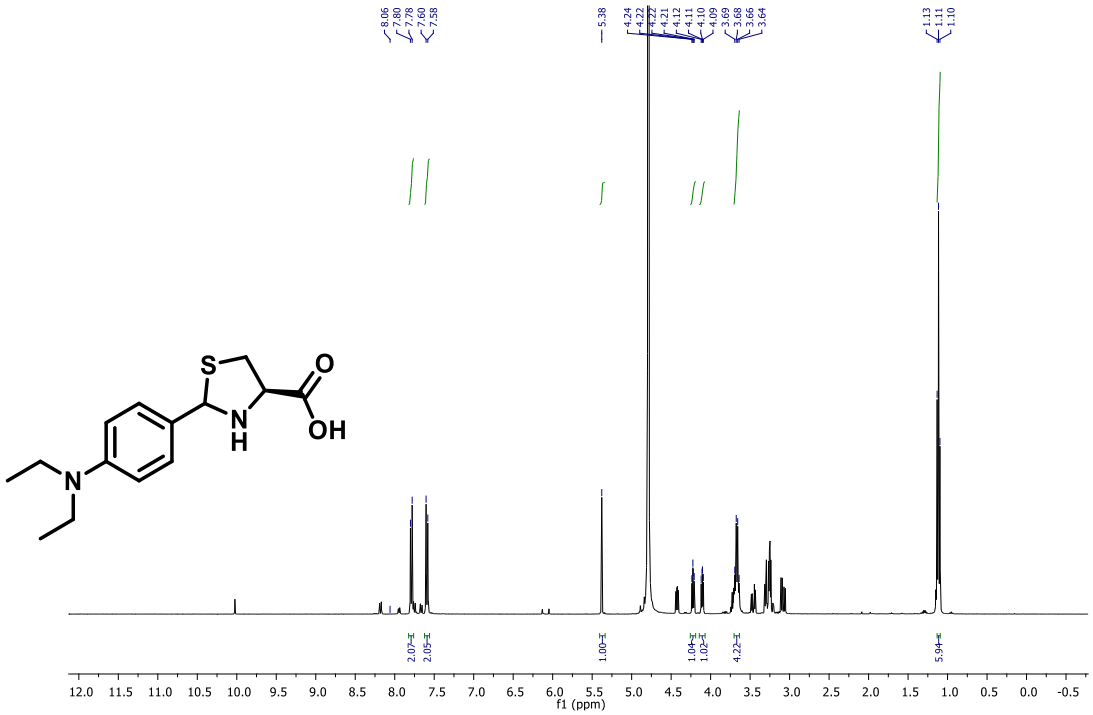


Figure S7g: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(4-(diethylamino)phenyl)thiazolidine-4-carboxylic acid (**6d**).

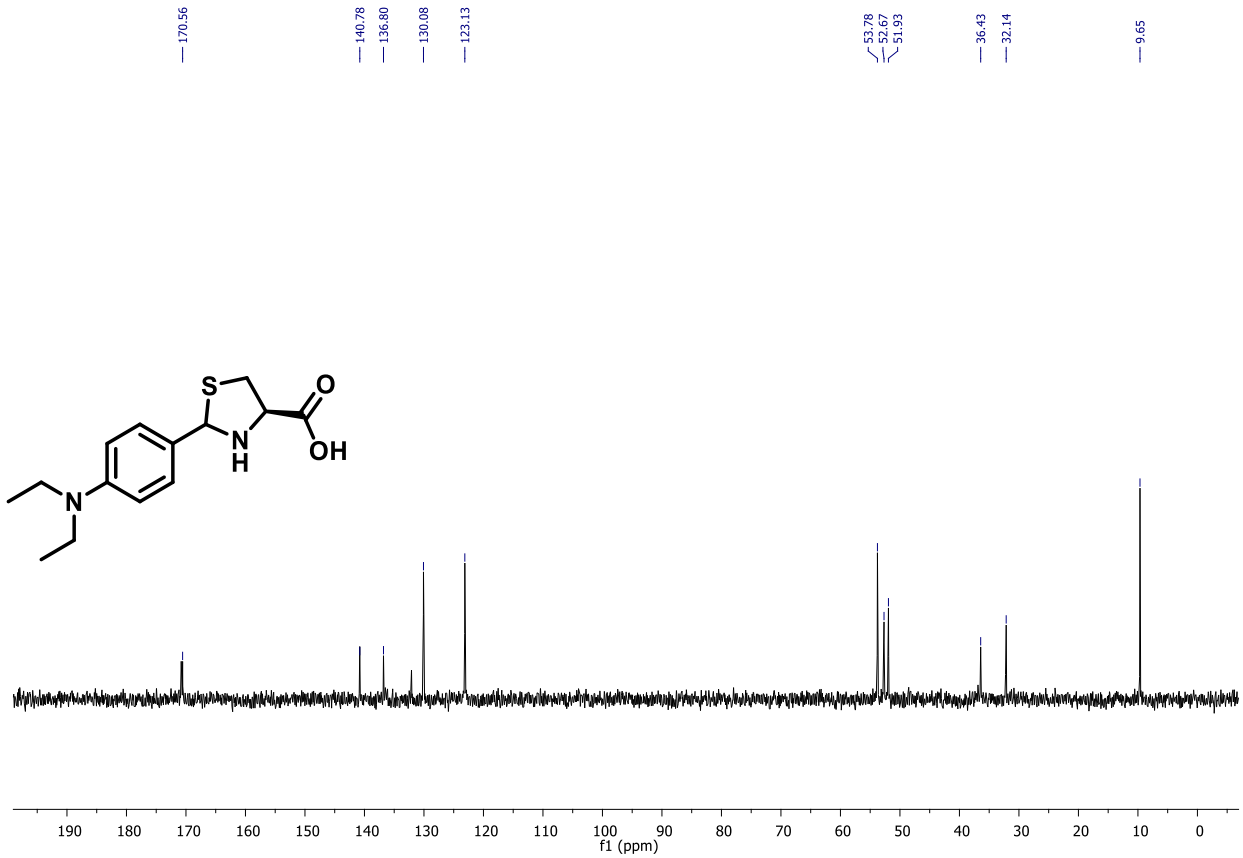


Figure S7d: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(4-(diethylamino)phenyl)thiazolidine-4-carboxylic acid (**6d**).

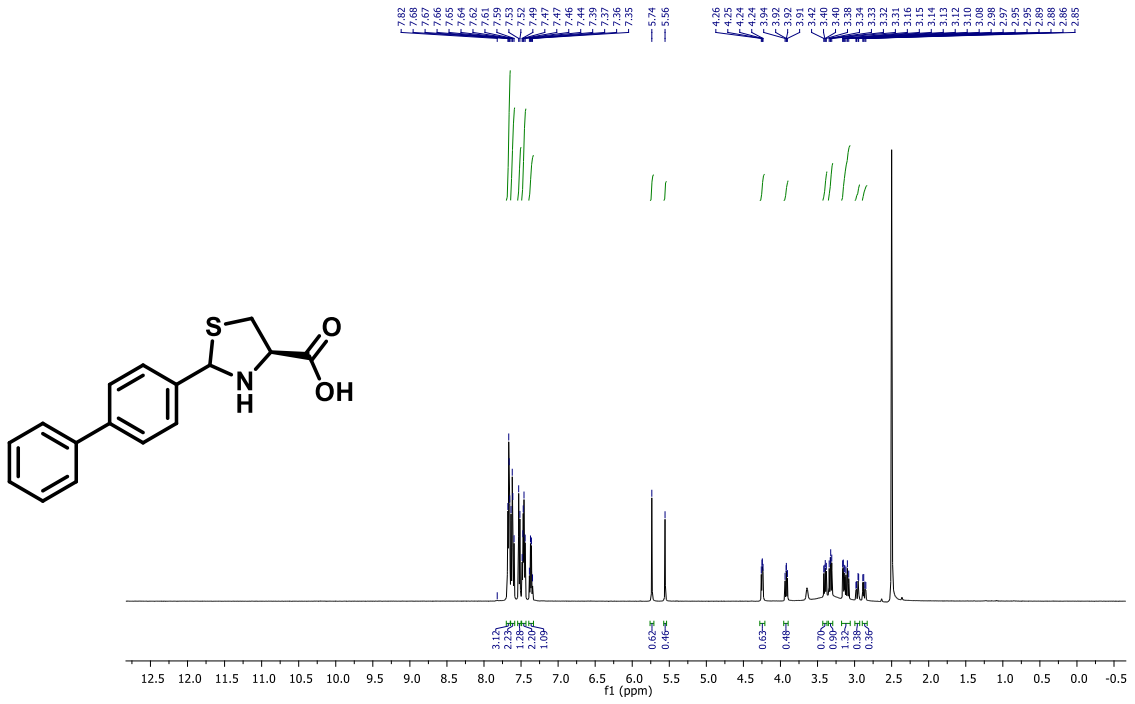


Figure S7e: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-([1,1'-biphenyl]-4-yl)thiazolidine-4-carboxylic acid (**6e**).

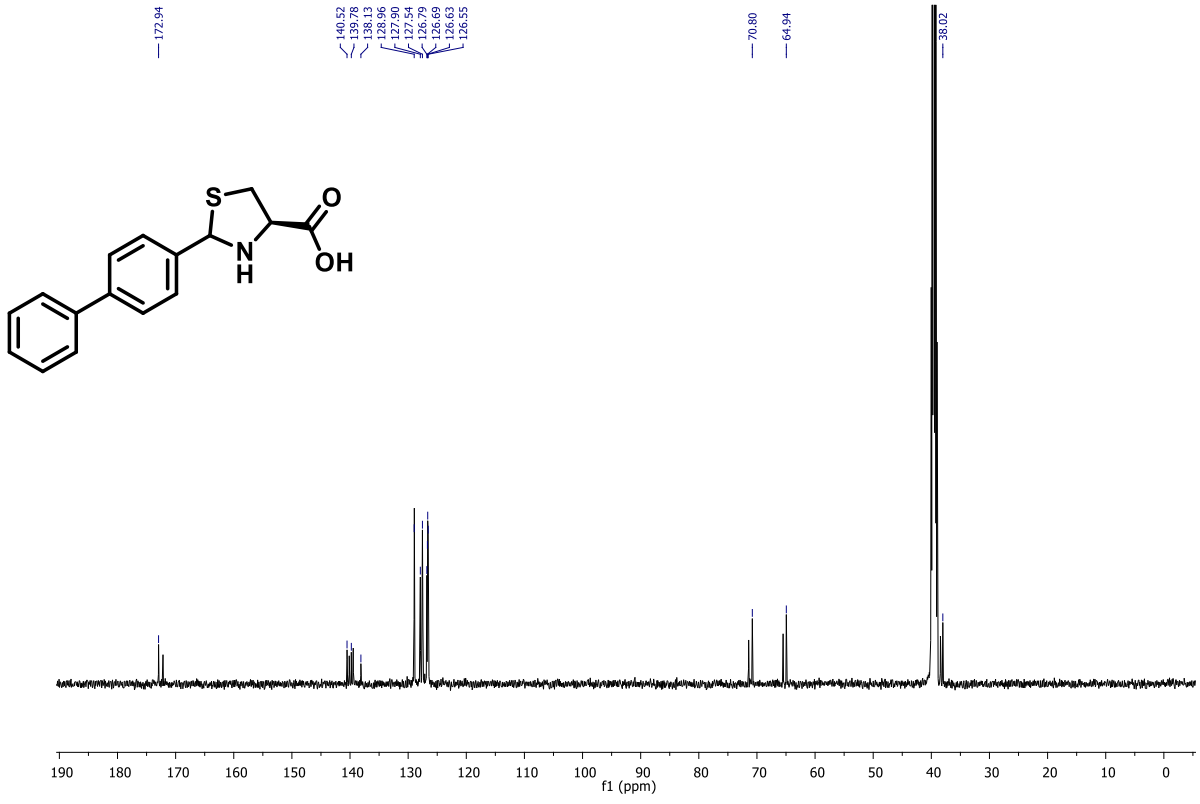


Figure S7f: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-([1,1'-biphenyl]-4-yl)thiazolidine-4-carboxylic acid (**6e**).

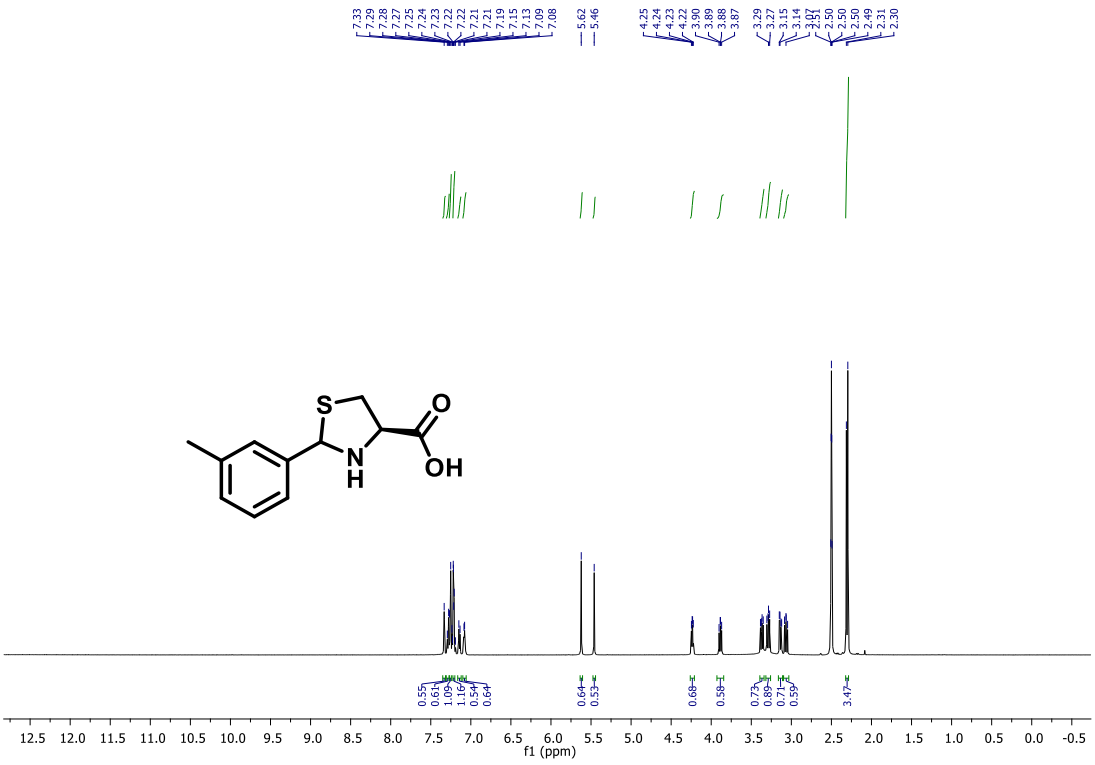


Figure S7g: [1]H NMR spectrum of (2*RS*, 4*R*)-2-(*m*-tolyl)thiazolidine-4-carboxylic acid (**6f**)

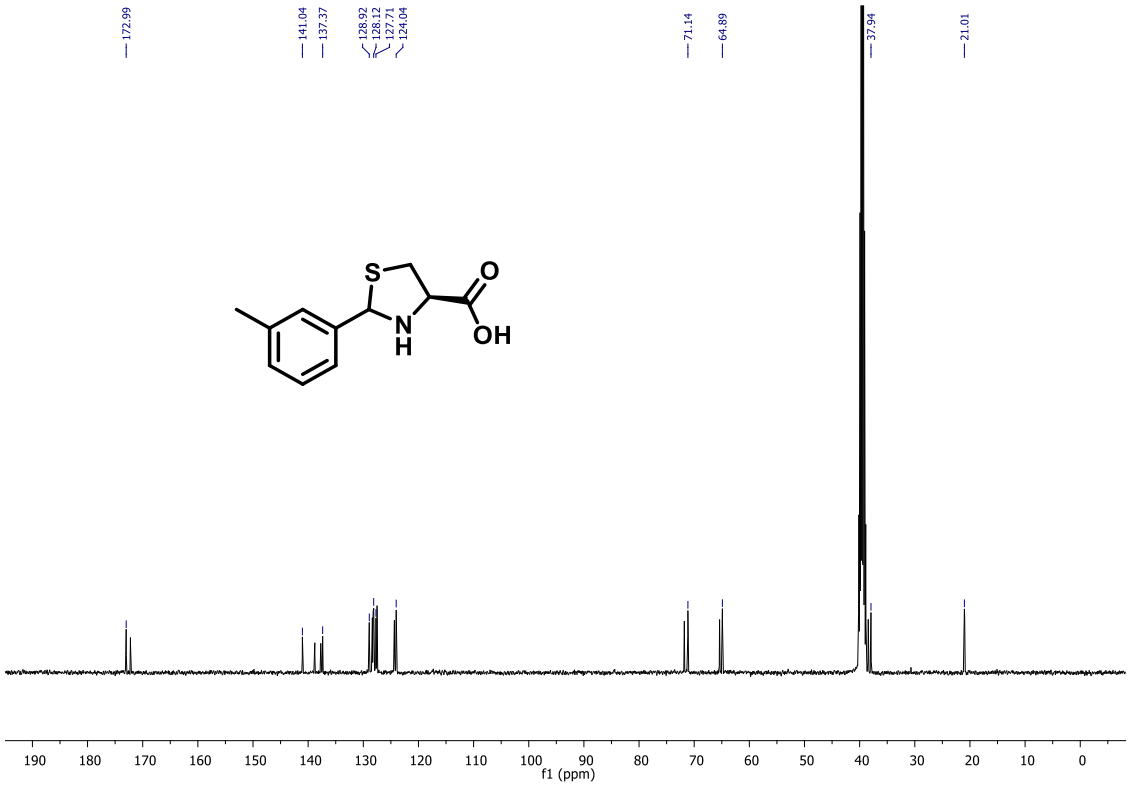


Figure S7h: [1]H NMR spectrum of (2*RS*, 4*R*)-2-(m-tolyl)thiazolidine-4-carboxylic acid (**6f**).

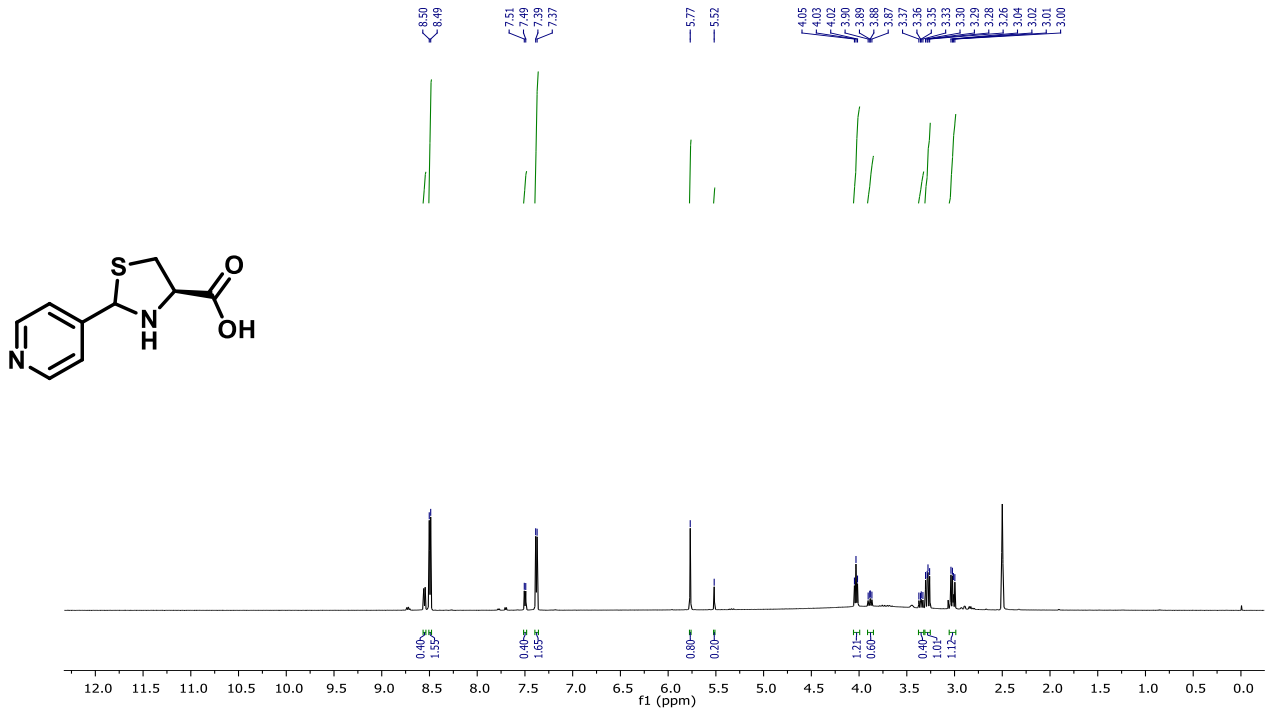


Figure S7i: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(m-tolyl)thiazolidine-4-carboxylic acid (**6f**).

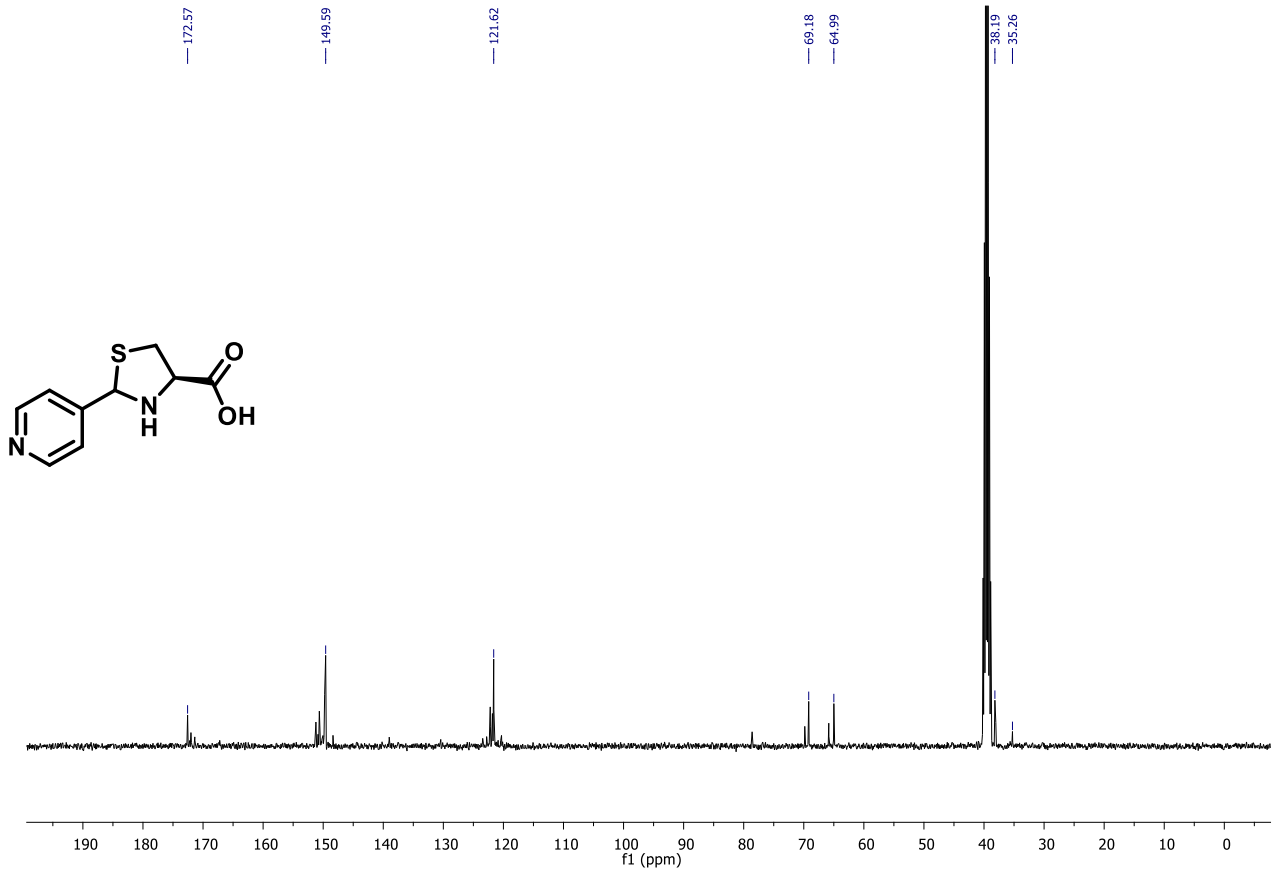


Figure S7j: ^{13}C NMR spectrum of (2*RS*, 4*R*)-2-(pyridin-4-yl)thiazolidine-4-carboxylic acid (**6g**).

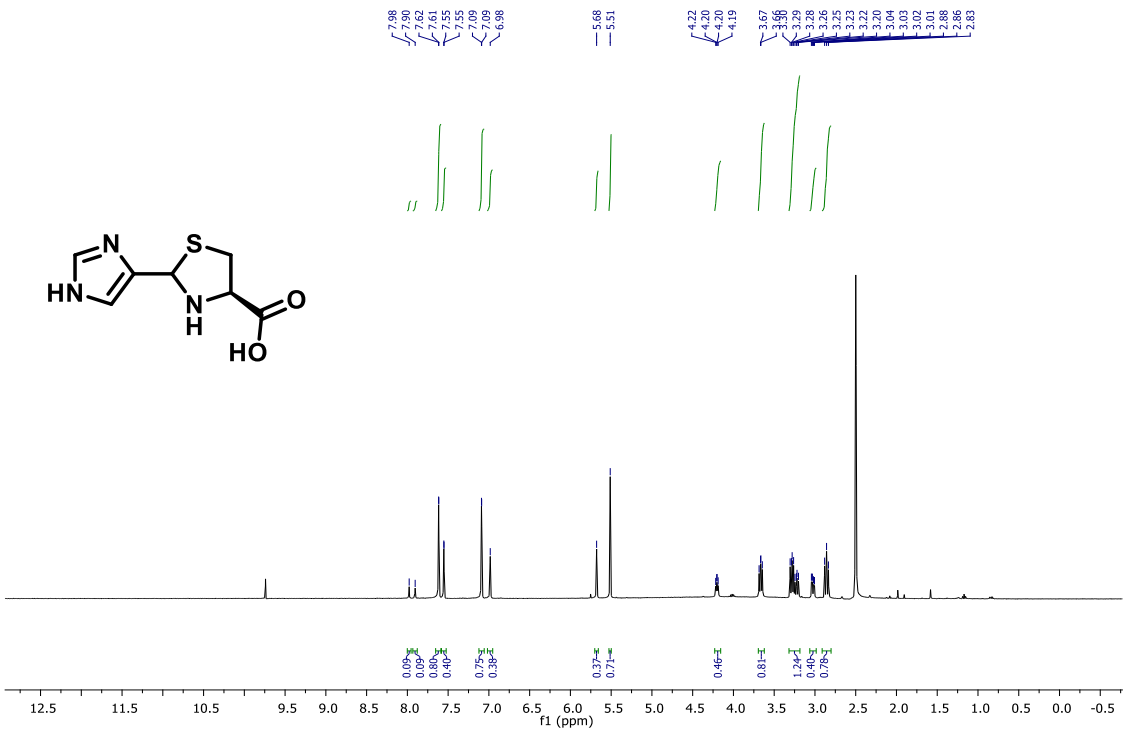


Figure S7k: ^{1}H NMR spectrum of (2*RS*, 4*R*)-2-(1H-imidazol-4-yl)thiazolidine-4-carboxylic acid (**6h**).

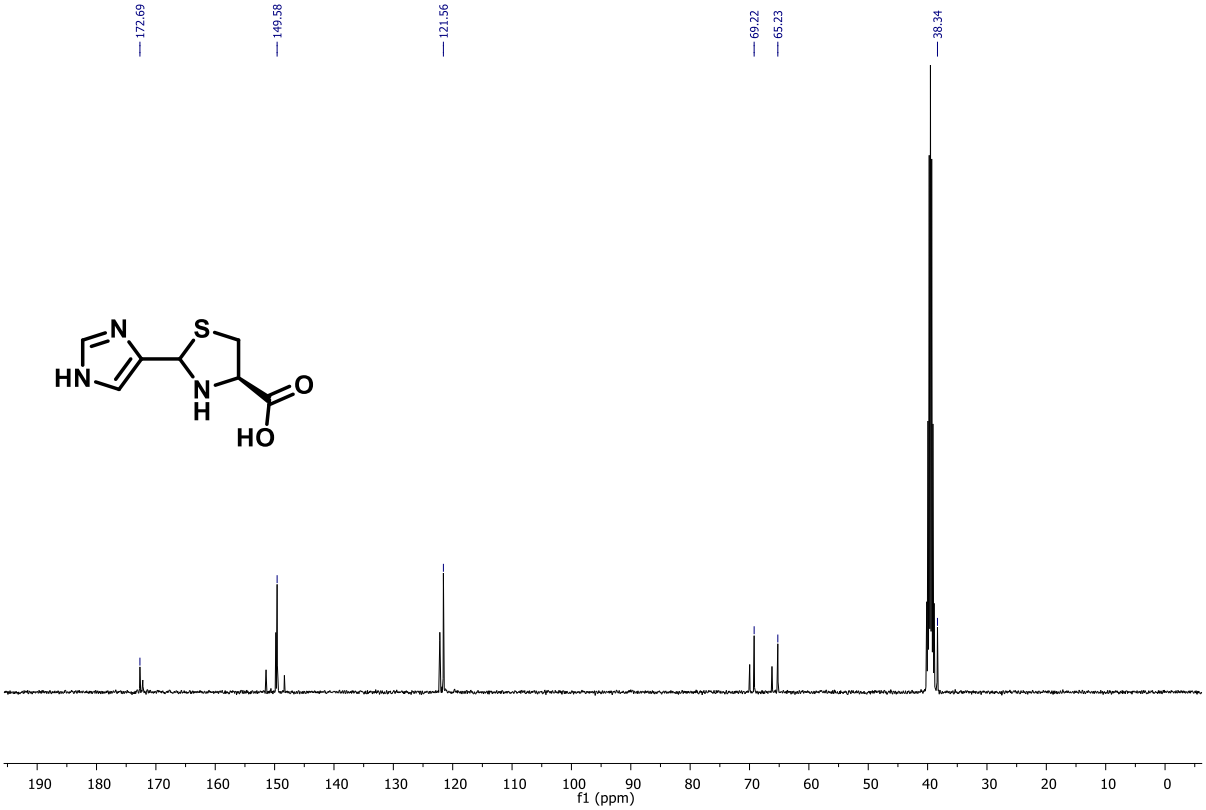


Figure S7l: ^{13}**C** NMR spectrum of (2*RS*, 4*R*)-2-(1H-imidazol-4-yl)thiazolidine-4-carboxylic acid (6h**).**