pubs.acs.org/acschemicalbiology

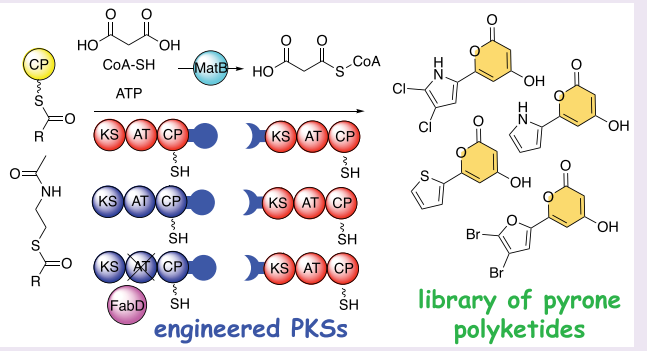
Letters

Biosynthesis-Guided Discovery and Engineering of α-Pyrone Natural Products from Type I Polyketide Synthases

Dongqi Yi and Vinayak Agarwal*

Cite This: https://doi.org/10.1021/acschembio.3c00081

ABSTRACT: Natural products containing the α-pyrone moiety are produced by polyketide synthases (PKSs) in bacteria, fungi, and plants. The conserved biosynthetic logic for the production of the α-pyrone moiety involves the cyclization of a triketide intermediate which also off-loads the polyketide from the activating thioester. In this study, we show that truncating a tetraketide natural product producing PKS assembly line allows for a thioesterase-independent off-loading of an α-pyrone polyketide natural product, one which we find to be natively present in the extracts of the bacterium that otherwise furnishes the tetraketide natural product. By engineering the truncated PKS *in vitro*, we demonstrate that a ketosynthase (KS) domain with relaxed substrate selectivity when coupled with in trans acylation of polyketide extender units can expand the chemical space of α-pyrone polyketide natural products. Findings from this study point toward heterologous intermolecular protein−protein interactions being detrimental to the efficiency of engineered PKS assembly lines.



Polyketide natural products containing the α-pyrone moiety are biosynthesized by modular type I polyketide synthases (PKSs), iterative type II and type III PKSs, and fungal nonreducing PKSs (NRPKSs).[1−5] Examples of natural products furnished by the above-mentioned three PKS types include venemycin (**1**, Figure 1A), triacetic acid lactone (**2**), and pyrophen (**3**), respectively. In each case, progressing from a carboxylic acid thioesterified either to coenzyme A (CoA-SH) or to the phosphopantetheine thiol of a carrier protein (CP), two polyketide extension reactions furnish a triketide intermediate that is then off-loaded via intramolecular annulation to furnish the pyrone natural product (Figure 1B). Reductive tailoring of the triketone prior to off-loading produces lactones rather than pyrones.

We recently described the *in vitro* reconstitution of the pyoluteorin (**4**, Figure 1C) biosynthetic pathway.[6,7] Here, three modules of the type I PKSs PltB and PltC afford a tetraketide that is reduced by the PltC ketoreductase (KR) domain. Off-loading by the standalone thioesterase (TE) PltG furnished **5** that was dehydrated and aromatized to **4**. Curiously, organization of the first two modules of the Plt PKSs is similar to the Vem PKSs that produce **1** (Figure 1C).[1,8] Despite the differences in the starter units (4,5-dichloropyrrole carboxylic acid for **4**; 3,5-dihydroxybenzoic acid for **1**) and the mechanism of the delivery of these starter units to the initiating ketosynthase (KS) domains of the respective PKSs, no reductive tailoring of the triketide intermediate occurs in either assembly line. For the Plt assembly line, the triketide intermediate is further extended to a tetraketide followed by a Dieckmann cyclization likely catalyzed by the PltG TE to afford **5**, while for the Vem assembly line, the terminal TE domain embedded in the VemH PKS off-loads the triketide via an intramolecular esterification to furnish **1**.

Motivated by the similarities in module organization of Plt and Vem PKSs, we asked if the Plt assembly line could be engineered to deliver α-pyrone polyketide products akin to the Vem pathway. For the Vem pathway, selectivity for the starter unit is generated by the adenylation (A) domain of the VemG loading module wherein this A-domain adenylates 3,5-dihydroxybenzoic acid and thioesterifies it to the phosphopantetheine arm of the VemG loading module CP (Figure 1C).[1] For the Plt pathway, the PltB module-1 KS screens and selects for the starter unit that is delivered by the upstream pyrrole maturation and halogenating enzymes. In light of the differential mechanism for starter unit delivery to the PltB PKS as compared to the VemG PKS, we also explored if PKS

Received: February 6, 2023
Accepted: April 17, 2023

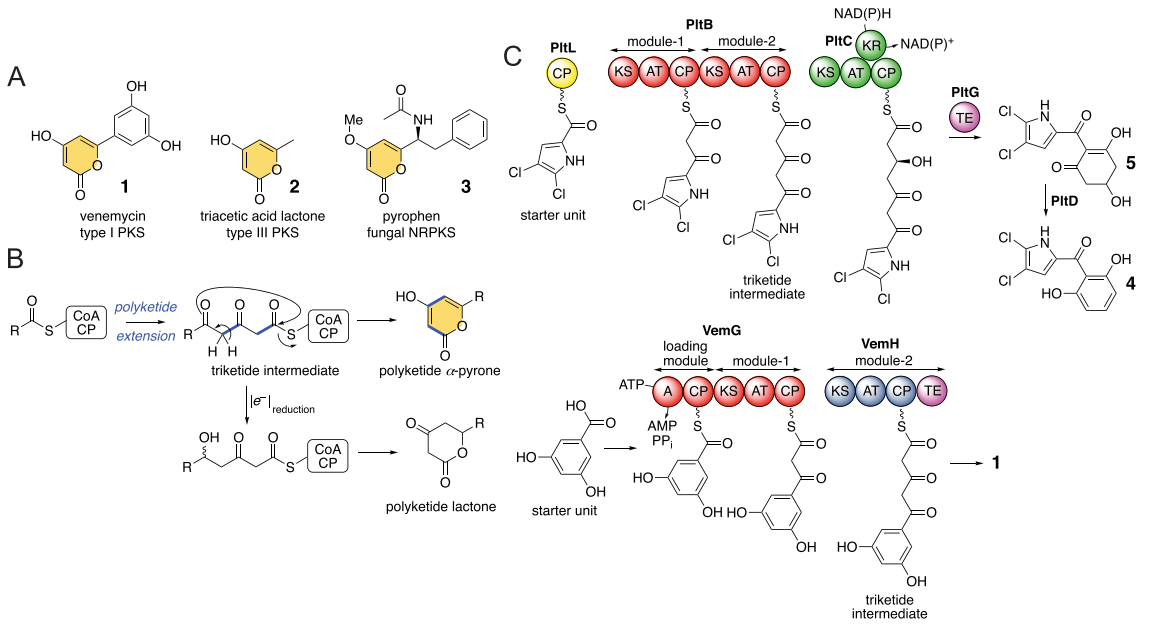


Figure 1. (A) Examples of polyketide natural products containing the α-pyrone moiety (shaded). (B) Formation of the α-pyrones involves triketide cyclization and off-loading; reduction of the triketide leads to lactone production. (C) The Plt and Vem assembly lines.

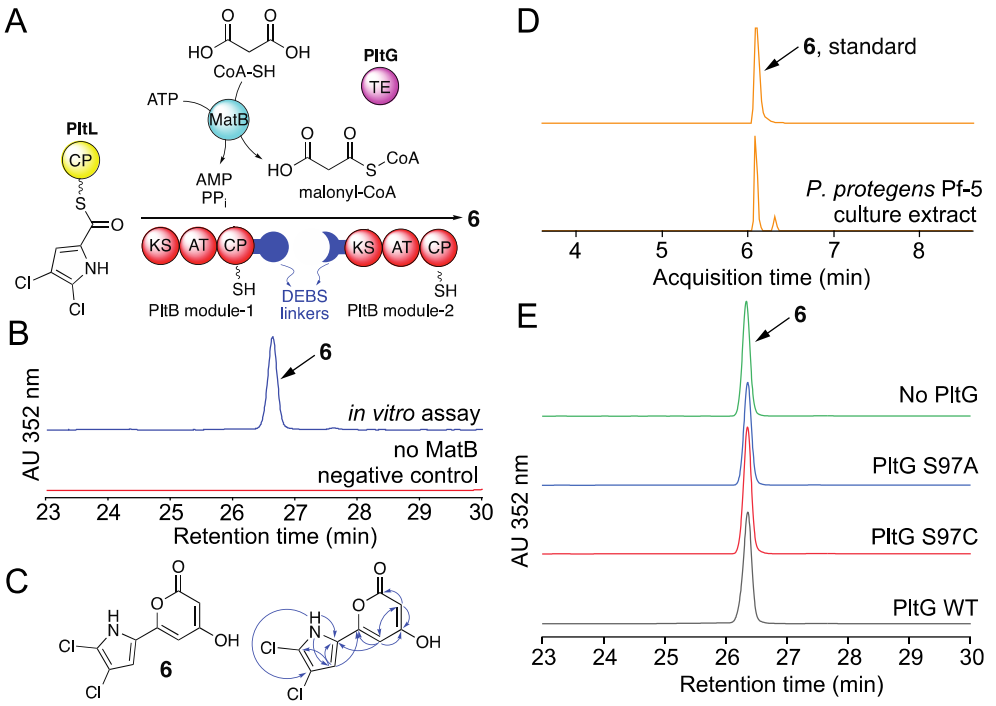


Figure 2. (A) Assay design for production of **6**. (B) UV-absorbance chromatograms demonstrating the production of **6**. A negative control reaction in which MatB was omitted did not produce **6**. (C) Structure of **6** with key HMBC correlations shown as blue arrows. (D) Extracted ion chromatograms demonstrating the presence of **6** in *P. protegens* Pf-5 culture extracts, as compared to a purified standard of **6**. (E) UV-absorbance chromatograms for end point assays demonstrating similar abundance of **6** in reactions carried out as illustrated in panel A, with PltG variants S97A and S97C and when PltG was omitted from the reaction.

engineering could expand the chemical space of the delivered α-pyrone products.

With the observation that both the Vem and Plt PKSs produce a nonreduced triketide intermediate, we explored, *in vitro*, whether the first two modules of the Plt PKS could mimic the activity of the Vem PKS (Figure 1C). The two Plt modules were separated by appending docking domains from the 6-deoxyerythronolide B synthase (DEBS) assembly line to the C-terminus of PltB module-1 and to the N-terminus of PltB module-2 and purified along with the TE PltG (Figure 2A).[6] The starter unit, 4,5-dichloropyrrole carboxylic acid thioesterified to the CP PltL, 4,5-dichloropyrrolyl-*S*-PltL, was chemoenzymatically synthesized (Figure S1).[9] The enzyme MatB was employed to generate malonyl-CoA *in situ* using malonate, CoA-SH, and ATP;[10] malonyl-CoA is used as the substrate by the acyltransferase (AT) domains of the PltB PKS

https://doi.org/10.1021/acschembio.3c00081
ACS Chem. Biol. XXXX, XXX, XXX−XXX

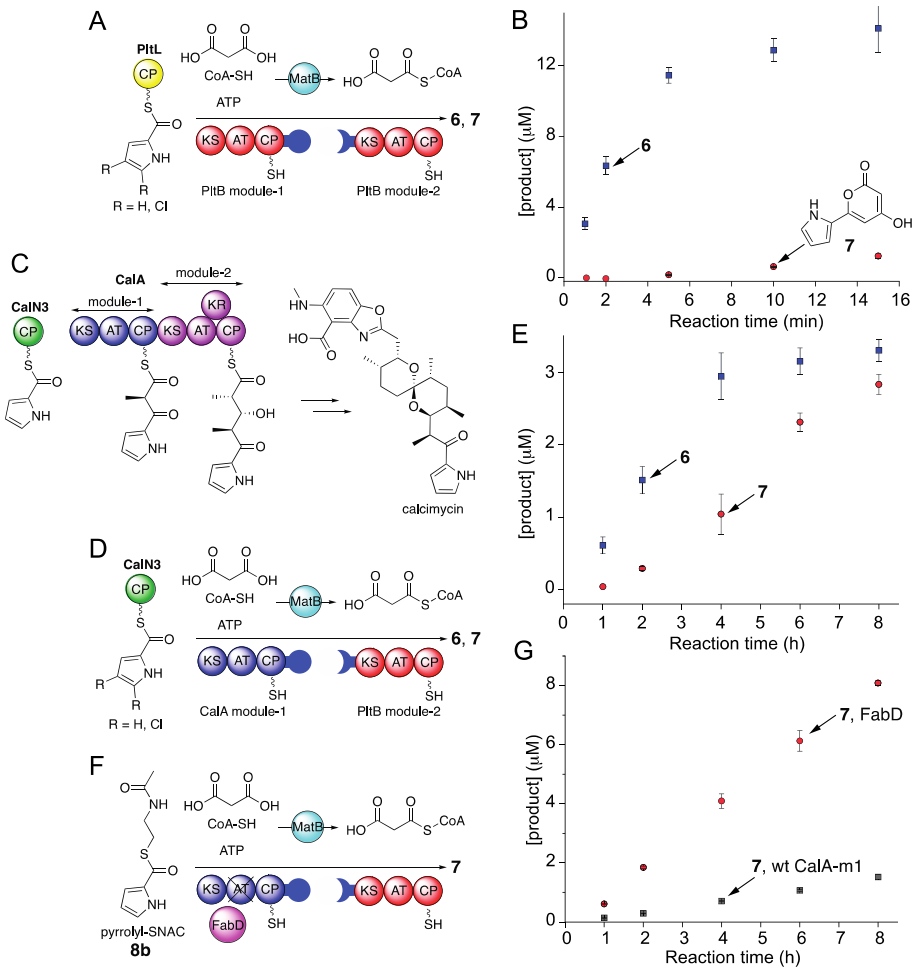


Figure 3. (A) Assay design to test substrate preference of the Plt PKS assembly line. The dichloropyrrolyl- and the pyrrolyl-*S*-PltL substrates lead to **6** and **7**, respectively. (B) Time-dependent accumulation of **6** and **7** produced in the assay illustrated in panel A. (C) The Cal PKS; only the first two modules of the Cal PKS are shown. The pyrrole carboxylic acid starter unit is delivered via the CP CalN3. (D) Assay design in which the PltB module-1 is replaced with the CalA module-1. (E) Time-dependent accumulation of **6** and **7** produced in the assay illustrated in panel D. (F) Assay design in which the CalA module-1 AT domain is inactivated and the *E. coli* FabD enzyme is provided to acylate the CalA module-1 CP in trans. (G) Time-dependent accumulation of **7** produced in the assay illustrated in panel F and with wild-type CalA module-1 with a functional AT domain.

modules. Upon incubation of the assay components, we observed the accumulation of a dichlorinated product (Figures 2B and S2−S4). The production of **4** or **5** was not detected using liquid chromatography/mass spectrometry (LC/MS).

The abundance of the conceivable α-pyrone produced in this assay was limited by the amount of 4,5-dichloropyrrolyl-*S*-PltL substrate that we could provide in an *in vitro* assay. To circumvent this challenge, we coexpressed genes encoding the PltB PKS modules in *E. coli* together with genes encoding MatB and the phosphopantetheinyl transferase Sfp.[11] The cell lysate of this *E. coli* strain was then directly used as a catalyst in an *in vitro* assay wherein we replaced 4,5-dichloropyrrolyl-*S*-PltL with 4,5-dichloropyrrolyl-*S*,*N*-acetylcysteamine (4,5-dichloropyrrolyl-SNAC, **8a**, Scheme S1, Figure S5−S6). These modifications allowed for preparative isolation and structural characterization of the product as the α-pyrone **6** (Figures 2C and S7−S11, Table S1). These data demonstrate that truncating the Plt assembly line to two modules is sufficient to change its output from the dihydrophloroglucinol **5** to an α-pyrone **6**.

Components added to the assay illustrated in Figure 2A are all present *in situ* in the *Pseudomonas protegens* Pf-5 bacterium that produces **4**. However, no pyrone natural products have been reported from *P. protegens* Pf-5. Using LC/MS, we could indeed detect the presence of **6** in the extracts of *P. protegens* Pf-5 (Figure 2D). This result establishes the *in vitro* enzymatically synthesized **6** as a physiologically produced natural product that had evaded prior detection and characterization.

The production of triketide pyrones by the Vem assembly line involves a TE domain that is embedded within the VemH PKS (Figure 1C).[1,8] The Plt assembly line presents a different scenario; here the standalone thioesterase PltG is involved in the production of **5**. In complete contrast, no embedded or standalone TEs are partnered with either the fungal NRPKSs that produce **3** or the bacterial type II PKS that produces α-pyrone polyketides enterocins and wailupemycins.[3,4] In light of these differences, we sought to evaluate the role of PltG in the production of **6**. Variants of PltG were prepared in which the serine residue in the PltG active site was replaced with alanine (PltG S97A) and cysteine (PltG S97C). While the S97A mutation would abolish the thioesterase activity, the S97C mutation could preserve or even enhance[12] the thioesterase activity of PltG. The PltG variants were trialed in the above-

https://doi.org/10.1021/acschembio.3c00081
ACS Chem. Biol. XXXX, XXX, XXX−XXX

mentioned assay with the 4,5-dichloropyrrolyl-*S*-PltL initiator substrate and purified PltB PKS modules. Chromatographic detection of **6** at similar abundance in all four assays demonstrates that the formation of the α-pyrone product did not require the TE participation (Figure 2E). This observation allows us to posit that off-loading of a triketide as an α-pyrone is noncatalytic, likely driven by the thermodynamic stability of the aromatic product.[3,4] This then calls into question the role of the terminal thioesterase domain embedded in the VemH PKS module in the production of **1**.

We have reported previously that the substrate selectivity of the PltB module-1 KS domain constrains the diversity of starter units that can progress along Plt PKS assembly line.[13] In the experimental setup illustrated in Figure 3A, which did not include the PltG TE, we provided the dichloropyrrolyl- and the pyrrolyl-*S*-PltL initiator substrates to the PltB PKS modules. Progressing from the physiological substrate dichloropyrrolyl-*S*-PltL, we observed a time-dependent accumulation of **6**. In contrast, starting from pyrrolyl-*S*-PltL, the level of production of the deschloro derivative of **6**, molecule **7**, was much lower (Figures 3B and S12−S14).

In addition to the substrate selectivity of the PltB module-1 KS domain, in our previous study, we had characterized the substrate promiscuity of the CalA module-1 KS domain.[13] The Cal PKS assembly line uses the pyrrolyl-*S*-CalN3 initiator substrate to furnish the pyrrolic antibiotic calcimycin (Figure 3C).[14] Unlike the Plt PKSs which use malonyl extender units, the Cal PKS module-1 and module-2, which are both present within the CalA polypeptide, use methylmalonyl extender units. The CalA module-1 KS domain was found to accept both pyrrolyl- (its physiological substrate) and the dichloropyrrolyl starter units.[13] In light of these observations, we asked if the CalA module-1 KS domain could be used to circumvent the substrate selectivity of the PltB module-1 KS domain. With this motivation, we swapped the PltB module-1 with the CalA module-1 (Figure 3D). To maintain physiological intermolecular protein−protein contacts, in this assay, the initiator substrates were delivered by the Cal CP, CalN3. Upon incubation of reaction components, we observed a higher level of production of **7** relative to **6** (Figure 3E), as compared to the pyrone production assay with the PltB module-1 (Figure 3B). Thus, the relaxed substrate selectivity of the CalA module-1 KS domain indeed allows for expansion of the α-pyrone product chemical space.

While the abundance of **7** relative to **6** was enhanced by swapping out the PltB module-1 and replacing it with CalA module-1, the product abundances and rates of product formation were diminished (Figure 3B,E). We have previously demonstrated that the delivery of the dichloropyrrolyl molecular cargo to the initiating PltB module-1 and CalA module-1 KSs from their cognate CPs (PltL and CalN3, respectively) was equally efficient and was thus not a likely contributing factor to the reduction in abundance and rate of product formation, at the very least for **6**.[13] Progressing along the PKS assembly lines, there could then be three possible reasons for the decrease in abundance and rate of product formation. First, in the assay set up illustrated in Figure 3D, despite the presence of the DEBS linker domains, the intermolecular protein−protein interactions between the CalA module-1 and the PltB module-2 are non-native (Figure 3D). This mismatch could compromise the efficiency of the transthioesterification of the diketide intermediate from CalA module-1 CP to the PltB module-2 KS.[15,16] Second, the downstream PltB module-2 KS domain could gatekeep against the extension of a nonphysiological diketide intermediate furnished by the upstream PKS module-1.[17] However, the potential gatekeeping activity of PltB module-2 KS is unlikely to affect biosynthesis of **6**, which we *also* observed to be negatively impacted (rate and abundance of **6** in Figure 3B,E). Third, the physiological extender unit incorporated by the CalA module-1 AT domain is methylmalonyl-CoA and not malonyl-CoA. The forced incorporation of the non-native malonyl extender unit by the CalA module-1 AT could compromise the efficiency of the entire engineered bimodular PKS.

To test if the non-native extender unit incorporation by the CalA module-1 AT domain compromised the efficiency of the assay illustrated in Figure 3D, we inactivated the CalA module-1 AT domain by a serine to alanine mutation in the AT active site. To then acylate the CalA module-1 CP in trans, we added purified *E. coli* fatty acid malonyl-CoA:CP transacylase (MAT) FabD to the assay. The physiological substrate for FabD is malonyl-CoA, same as the PltB module-2 AT domain.[18] In line with previous reports where trans-acting ATs have substituted for inactivated cis-ATs,[19] the pyrrole carboxylic acid starter unit was thioesterified to SNAC (**8b**), rather than the CP CalN3 (Figures 3F and S15, Scheme S2). This change also potentially ameliorates substrate degradation due to the transacylating activity of FabD.

By observing an enhancement in the abundance of product **7** when FabD substituted in trans for the inactivated cis-acting CalA module-1 AT domain (Figure 3G), we could confirm that the mismatch in the extender units in the assay illustrated in Figure 3D was indeed one of the contributing factors that compromised rate and abundance of product formation. Engineering of the PKS AT domains is of intense contemporary interest and AT domains in collinear PKS modules have been replaced with cis-acting substrate promiscuous MATs to facilitate the delivery of engineered polyketides.[20−22] Here, we have used a trans-acting MAT to substitute for an inactivated cis-acting AT domain to partially circumvent the regularly observed reduction in yields for engineered PKSs.

With an engineered system in hand to increase the α-pyrone product formation upon replacement of the CalA module-1 AT with FabD, we explored if we could expand the acylpyrone product profile. With this motivation, a panel of thioesterified SNAC starter units was developed (**8a**−**8s**; Figures 4A and S16−S44; Scheme S3−S19). In addition to derivatized pyrroles, the panel of SNAC-thioesters included pentacyclic thiophenes, furans, thiazole, oxazole, hexacyclic pyridines, a phenyl, and branched short chain alkanes. These starter units were provided to the *native* bimodular Plt PKSs and to the *engineered* CalA/PltB/FabD hybrid PKS assembly lines (Figure 4B). Product formation was monitored by mass spectrometry, including the characteristic MS/MS fragmentation of the acyl pyrones (Figures S45−S61). The product abundances were normalized to the abundance of **6**, produced starting from **8a**, as the dichloropyrrolyl starter unit was demonstrated to be competent substrate for both the PltB module-1 KS and the CalA module-1 KS (Figure 3).

The normalized product abundances, illustrated as a heat map in Figure 4C, reveal trends for starter unit specificities (Table S2). In general, as compared to the native PKSs, the engineered PKSs/FabD indeed demonstrated an expanded product profile, with pyrone products detected in modest

https://doi.org/10.1021/acschembio.3c00081
ACS Chem. Biol. XXXX, XXX, XXX−XXX

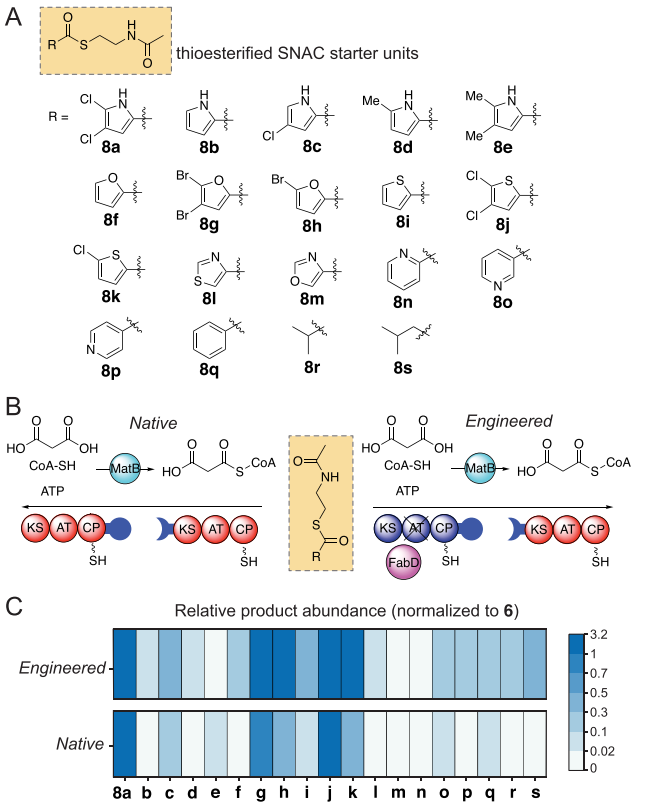


Figure 4. (A) Panel of synthetic thioesterified SNAC starter units prepared in this study. (B) The *Native* and *Engineered* PKS systems evaluated for α-pyrone polyketide product formation. (C) Heat map representing the abundance of α-pyrone products produced from the *Engineered* (top) and *Native* (bottom) PKS systems. The abundance of each product is normalized to **6**, which is produced starting from **8a**.

abundances starting from **8o**−**8s** which were poor substrates to the native PKS system. Interestingly, as for pyrroles, halogenated thiophenes and furans were competent substrates for both systems, with the engineered PKS system demonstrating greater relative product abundance as compared to the native PKS system for these substrates as well. The oxazole and thiazole were poor substrates in both cases.

Taken together, data described herein demonstrate that truncated type I PKSs can produce α-pyrones without the involvement of an off-loading thioesterase and that engineering efforts directed at circumventing the specificities of the KS and AT domains can expand the scope and yield of α-pyrone products. Several pyrone natural products possess validated pharmaceutical potential and also serve as precursors in synthetic schemes.[23] The above-mentioned efforts thus could serve to generate a biocatalyst toolbox for the targeted delivery of substituted pyrones starting from simple thioesterified substrates with *in situ* enzymatic production of polyketide extender units, as has been achieved for other commodity chemicals.[24] The promise of polyketide engineering, facile as it may seem given the vectorial nature of collinear assembly lines, is yet to be fully realized.[25,26] Even for the simple bimodular extension necessary for the construction of α-pyrones, our study reveals that the selectivity of the KSs and the ATs and the heterologous intermodular interactions can constrain product yield. In addition to chemical delivery to satisfy medicinal, synthetic, or other commercial needs, the simple bimodular PKS systems described in this study could serve as testing systems to evaluate and evolve PKS engineering efforts in the future.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acschembio.3c00081.

Experimental details for recombinant protein production, synthetic procedures, and analytical procedures for compound characterization, Figures S1−S61, Schemes S1−S19, Tables S1−S2, and Supplementary References (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Vinayak Agarwal − *School of Chemistry and Biochemistry and School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* orcid.org/0000-0002-2517-589X; Phone: (+1)404-385-378; Email: vagarwal@gatech.edu

Author

Dongqi Yi − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acschembio.3c00081

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge J. Wysocki for technical assistance, support from the National Science Foundation (CHE-2004030), and J.M. Deutsch, H. Albataineh, and N. Garg for mass spectrometry data.

■ REFERENCES

(1) Thanapipatsiri, A.; Gomez-Escribano, J. P.; Song, L.; Bibb, M. J.; Al-Bassam, M.; Chandra, G.; Thamchaipenet, A.; Challis, G. L.; Bibb, M. J. Discovery of unusual biaryl polyketides by activation of a silent *Streptomyces venezuelae* biosynthetic gene cluster. *ChemBioChem.* **2016**, *17* (22), 2189−2198.

(2) Zhou, Y.; Mirts, E. N.; Yook, S.; Waugh, M.; Martini, R.; Jin, Y.-S.; Lu, Y. Reshaping the 2-pyrone synthase active site for chemoselective biosynthesis of polyketides. *Angew. Chem., Int. Ed.* **2022**, *62* (5), e202212440.

(3) Cheng, Q.; Xiang, L.; Izumikawa, M.; Meluzzi, D.; Moore, B. S. Enzymatic total synthesis of enterocin polyketides. *Nat. Chem. Biol.* **2007**, *3* (9), 557−558.

(4) Hai, Y.; Huang, A.; Tang, Y. Biosynthesis of amino acid derived α-pyrones by an NRPS−NRPKS hybrid megasynthetase in fungi. *J. Nat. Prod.* **2020**, *83* (3), 593−600.

(5) Schäberle, T. F. Biosynthesis of α-pyrones. *Beilstein J. Org. Chem.* **2016**, *12*, 571−588.

(6) Yi, D.; Niroula, D.; Gutekunst, W. R.; Loper, J. E.; Yan, Q.; Agarwal, V. A nonfunctional halogenase masquerades as an aromatizing dehydratase in biosynthesis of pyrrolic polyketides by type I polyketide synthases. *ACS Chem. Biol.* **2022**, *17* (6), 1351−1356.

(7) Nowak-Thompson, B.; Chaney, N.; Wing, J. S.; Gould, S. J.; Loper, J. E. Characterization of the pyoluteorin biosynthetic gene cluster of *Pseudomonas fluorescens* Pf-5. *J. Bacteriol.* **1999**, *181* (7), 2166−74.

(8) Miyazawa, T.; Hirsch, M.; Zhang, Z.; Keatinge-Clay, A. T. An *in vitro* platform for engineering and harnessing modular polyketide synthases. *Nat. Commun.* **2020**, *11* (1), 80.
(9) Agarwal, V.; Diethelm, S.; Ray, L.; Garg, N.; Awakawa, T.; Dorrestein, P. C.; Moore, B. S. Chemoenzymatic synthesis of acyl coenzyme A substrates enables *in situ* labeling of small molecules and proteins. *Org. Lett.* **2015**, *17* (18), 4452−5.
(10) Hughes, A. J.; Keatinge-Clay, A. Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* **2011**, *18* (2), 165−76.
(11) Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **1998**, *37* (6), 1585−95.
(12) Koch, A. A.; Hansen, D. A.; Shende, V. V.; Furan, L. R.; Houk, K. N.; Jiménez-Osés, G.; Sherman, D. H. A single active site mutation in the pikromycin thioesterase generates a more effective macrocyclization catalyst. *J. Am. Chem. Soc.* **2017**, *139* (38), 13456−13465.
(13) Yi, D.; Acharya, A.; Gumbart, J. C.; Gutekunst, W. R.; Agarwal, V. Gatekeeping ketosynthases dictate initiation of assembly line biosynthesis of pyrrolic polyketides. *J. Am. Chem. Soc.* **2021**, *143* (20), 7617−7622.
(14) Wu, Q.; Liang, J.; Lin, S.; Zhou, X.; Bai, L.; Deng, Z.; Wang, Z. Characterization of the biosynthesis gene cluster for the pyrrole polyether antibiotic calcimycin (A23187) in *Streptomyces chartreusis* NRRL 3882. *Antimicrob. Agents Chemother.* **2011**, *55* (3), 974−82.
(15) Klaus, M.; Buyachuihan, L.; Grininger, M. Ketosynthase domain constrains the design of polyketide synthases. *ACS Chem. Biol.* **2020**, *15* (9), 2422−2432.
(16) Klaus, M.; Ostrowski, M. P.; Austerjost, J.; Robbins, T.; Lowry, B.; Cane, D. E.; Khosla, C. Protein-protein interactions, not substrate recognition, dominate the turnover of chimeric assembly line polyketide synthases. *J. Biol. Chem.* **2016**, *291* (31), 16404−16415.
(17) Hirsch, M.; Fitzgerald, B. J.; Keatinge-Clay, A. T. How cis-acyltransferase assembly-line ketosynthases gatekeep for processed polyketide intermediates. *ACS Chem. Biol.* **2021**, *16* (11), 2515−2526.
(18) Serre, L.; Verbree, E. C.; Dauter, Z.; Stuitje, A. R.; Derewenda, Z. S. The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5-Å resolution.: crystal structure of a fatty acd synthase component. *J. Biol. Chem.* **1995**, *270* (22), 12961−12964.
(19) Walker, M. C.; Thuronyi, B. W.; Charkoudian, L. K.; Lowry, B.; Khosla, C.; Chang, M. C. Y. Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* **2013**, *341* (6150), 1089−1094.
(20) Kalkreuter, E.; Bingham, K. S.; Keeler, A. M.; Lowell, A. N.; Schmidt, J. J.; Sherman, D. H.; Williams, G. J. Computationally-guided exchange of substrate selectivity motifs in a modular polyketide synthase acyltransferase. *Nat. Commun.* **2021**, *12* (1), 2193.
(21) Kalkreuter, E.; CroweTipton, J. M.; Lowell, A. N.; Sherman, D. H.; Williams, G. J. Engineering the substrate specificity of a modular polyketide synthase for installation of consecutive non-natural extender units. *J. Am. Chem. Soc.* **2019**, *141* (5), 1961−1969.
(22) Rittner, A.; Joppe, M.; Schmidt, J. J.; Mayer, L. M.; Reiners, S.; Heid, E.; Herzberg, D.; Sherman, D. H.; Grininger, M. Chemoenzymatic synthesis of fluorinated polyketides. *Nat. Chem.* **2022**, *14* (9), 1000−1006.
(23) McGlacken, G. P.; Fairlamb, I. J. S. 2-Pyrone natural products and mimetics: isolation, characterisation and biological activity. *Nat. Prod. Rep.* **2005**, *22* (3), 369−385.
(24) Hagen, A.; Poust, S.; Rond, T. d.; Fortman, J. L.; Katz, L.; Petzold, C. J.; Keasling, J. D. Engineering a polyketide synthase for *in vitro* production of adipic acid. *ACS Synth. Biol.* **2016**, *5* (1), 21−27.
(25) Poust, S.; Hagen, A.; Katz, L.; Keasling, J. D. Narrowing the gap between the promise and reality of polyketide synthases as a synthetic biology platform. *Curr. Opin. Biotechnol.* **2014**, *30*, 32−39.
(26) Klaus, M.; Grininger, M. Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* **2018**, *35* (10), 1070−1081.

https://doi.org/10.1021/acschembio.3c00081
ACS Chem. Biol. XXXX, XXX, XXX−XXX

SUPPLEMENTARY INFORMATION FOR:

Biosynthesis-guided discovery and engineering of α-pyrone natural products from type I polyketide synthases

Dongqi Yi[1] and Vinayak Agarwal[1,2,*]

[1]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA

[2]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA

*Correspondence: vagarwal@gatech.edu; Ph: (+1)404-385-378

Supplementary Information document contains:

SUPPLEMENTARY MATERIALS AND METHODS

General materials and instrumentation

All chemicals, solvents, and media components were obtained commercially from Sigma-Aldrich, Fisher Scientific, and VWR, and used without further purification. Phusion high-fidelity DNA polymerase and Gibson assembly Master Mix were purchased from New England Biolabs. PrimeSTAR DNA polymerase Master Mix was purchased from Takaro Bio. Synthetic DNA fragments were ordered from Twist Biosciences. Reactions were monitored by thin layer chromatography (TLC) carried out on Supelco silica gel (60 F_{254}) glass plates visualized by UV light. Silica gel (SiliaFlash GE60, 60–200 μm) was used for flash chromatography. Nuclear magnetic resonance (NMR) spectra were recorded on Bruker Avance III HD 400, 500, or 700 MHz instruments and calibrated using residual undeuterated solvent as the internal reference ($CDCl_3$ δ_H 7.26 and δ_C 77.16, MeOD δ_H 3.31 and δ_C 49.00, DMSO-d_6 δ_H 2.50 and δ_C 39.52). The splitting patterns were reported as s=singlet, d=doublet, t=triplet, q=quartet, m=multiplet, br=broad. Mass spectra were recorded on a Bruker Impact II high resolution time of flight (ToF) mass spectrometer with an electrospray ionization (ESI) source coupled to an Agilent 1290 ultra high-performance liquid chromatography system equipped with a diode array detector.

Vector construction

Standard molecular biology techniques were used to carry out plasmid construction. Amplification of target DNA fragments were carried out with either Phusion or PrimeSTAR high-fidelity DNA polymerases. Gibson assembly was used to subclone genes encoding for proteins of interest into the target vectors. The sequences of recombinant plasmids were confirmed by Sanger sequencing at Eton Biosciences.

The construction of PltB module1 with DEBS2 C-terminal docking domain (dd) sequences in pET28(+), PltB module2 with DEBS3 N-terminal docking domain sequences in pET28(+), PltG in pET24(+), Sfp in the first multiple cloning site (MCS) of pCDFDuet-1, MatB in pET28-MBP expression vectors has been described previously.[1-2] To construct the expression vector for CalA module1 fused with DEBS2 dd, *calA* module1 gene was amplified from previously synthesized pET28(+)-CalA module1 vector,[1] while the pET28(+) vector backbone with DEBS2dd sequences ligated was amplified using pET28(+)-PltB module1-DEBS2dd as the template. The amplified DNA fragment encoding for CalA module1 was then inserted into pET28(+) vector containing C-terminal DEBS2dd sequences using Gibson assembly. For the construction of FabD overexpression vector, the DNA fragment encoding *fabD* gene was first amplified from the *Escherichia coli* BL21Gold(DE3) genomic DNA and then assembled into pET28(+)

vector. Plasmids for PltG and CalA module1 mutants were generated by site-directed mutagenesis using standard procedures.

For production of compound **6**, a four-gene, three-plasmid system was designed for co-overexpression of PltB module1-DEBS2dd, DEBS3dd-PltB module2, Sfp, and MatB. The DNA fragments encoding for MatB, PltB module1 fused with DEBS2dd and PltB module2 linked to DEBS3dd were amplified using vectors constructed above as the template. The gene encoding for PltB module1-DEBS2dd was first inserted into the second MCS (between NdeI and XhoI restriction sites) of pETDuet-1 vector, followed by the introduction of the *matB* gene into the first MCS (between NcoI and HindIII restriction sites) to construct the recombinant pETDuet-1 vector expressing both MatB and PltB module1 DEBS2dd. The DNA fragment containing PltB module2 DEBS3dd was inserted into the MCS2 between NdeI and XhoI restriction sites of pACYCDuet-1 vector.

Recombinant protein expression and purification

For overexpression of holo-PKSs, the pET28(+) vector carrying PKS modules and the pCDFDuet-1 vector carrying Sfp were co-transformed into *E. coli* BL21Gold(DE3). Overnight cultures were inoculated into 2–4 L terrific both media supplemented with appropriate antibiotics. The cells were grown at 30 ˚C until OD_{600} reached 0.4–0.5 at which time growth temperature was reduced to 18 ˚C. When OD_{600} reached 0.7–0.8, protein expression was induced by the addition of 0.05–0.1 mM IPTG together with 0.25 mM calcium pantothenate. Bacterial cells were cultured at 18 ˚C for an addition of 18 h before harvested by centrifugation at 2,000×gfor 25 min. Cell pellets were stored at −80 ˚C until purification. For MatB, PltG (wild type and mutants), and FabD, expression vectors encoding for proteins of interest were transformed into *E. coli* BL21Gold(DE3). Similar growth condition was used as described above, except that 1 L terrific broth medium was used for cell culture and protein expression was induced with 0.2 mM IPTG.

All steps for protein purifications were performed at 4 ˚C or on ice. Cell pellets were resuspended in binding buffer (20 mM Tris-HCl pH=8.0, 500 mM NaCl, 10% glycerol) and lysed by sonication. The lysate was clarified by centrifugation at 25,000×g for 45 min, and then applied to a 5 mL HisTrap HP column. The column was washed extensively with wash buffer (20 mM Tris-HCl pH=8.0, 30 mM imidazole, 500 mM NaCl, 10% glycerol) till a stable UV-absorbance base line was observed, and then eluted with a linear gradient to 100 % of elution buffer (20 mM Tris-HCl pH=8.0, 250 mM imidazole, 500 mM NaCl, 10% glycerol) over 8 column volumes using ÄKTAprime plus FPLC system. Purity of eluent protein fractions were checked by SDS-PAGE, and fractions containing desired purified proteins were combined. For holo-PKSs, pooled protein solutions were concentrated using 50 kDa Amicon centrifugal filters and

desalted into binding buffer with PD-10 columns. Combined FabD protein solution was dialyzed overnight in 2 L binding buffer before storage. For PltG wild type and mutants, combined protein fractions were dialyzed in 2 L buffer composed of 20 mM Tris-HCl pH=8.9, 50 mM KCl and 10% glycerol overnight. The protein samples after dialysis were loaded to a 5 mL Hi-Trap Q column, washed with 5 column volumes of buffer A (20 mM Tris-HCl pH=8.9, 50 mM KCl), and then eluted with buffer B (20 mM Tris-HCl pH=8.9, 1 M KCl). Eluted protein fractions were checked by SDS-PAGE. Fractions containing protein of interest were combined, glycerol was added to a final concentration of 10% v/v and stored in small aliquots at −80˚C. Fresh aliquots were used each time for enzyme assays.

Enzyme assay for pyrone formation

In PKS assays with CPs, PltB module1 was incubated with its cognate CP partner-PltL, while CalA module1 was paired with CalN3. The preparation of pyrrolyl-*S*-PltL, dichloropyrrolyl-*S*-PltL, pyrrolyl-*S*-CalN3, and dichloropyrrolyl-*S*-CalN3 has been described previously (Fig. S1).[3] Briefly, the assay was composed of 1.2 mM synthesized (dichloro)pyrrolyl acyl-*S*-pantetheines, 10 mM $MgCl_2$, 50 mM HEPES-Na (pH=7.9), 400 μM CPs, 3 μM Sfp, 2 μM CoaA, 2 μM CoaD, and 2 μM CoaE in a total volume of 2.5 mL. The enzymatic reaction was initiated by the addition of 10 mM ATP after incubation at 30 °C for 5 min. The assay was incubated at 30 ˚C for 4.5 h, followed by 4 ˚C overnight, and then desalted into binding buffer (20 mM Tris-HCl pH=8.0, 500 mM NaCl, 10% glycerol) with PD-10 columns.

The PKS assay was performed in a total volume of 200 μL containing 50 μM CP substrates, 5 mM ATP, 10 mM $MgCl_2$, 1.2 mM Na-malonate, 0.6 mM coenzyme A, 5 mM TCEP, 5 μM MatB, 5 μM holo-PltB module1 or 10 μM holo-CalA module1 with equimolar holo-PltB module2, 0 or 5 μM PltG (wild type or mutants), and 400 mM potassium phosphate buffer (pH=7.5). The enzyme assays were incubated at 30 ˚C for 1 h before being quenched with 25 μL 3 M HCl. The quenched reaction mixture was extracted using EtOAc (250 μL, 3×). The combined organic layers were concentrated under vacuum, reconstituted in 100 μL MeOH, and then analyzed by HPLC or LC-MS after clarification by centrifugation. For time-course competitive assay, the assay was performed in a total volume of 1.2 mL supplied with equimolar dichloropyrrolyl- and unsubstituted pyrrolyl-*S*-CPs (25 μM each). For PltB, 200 μL samples were taken from the reaction mixture and quenched with 25 μL 3 M HCl at 1, 2, 5, 10, 15 min. 200 μL samples were taken and quenched at 1, 2, 4, 6, 8 h for CalA reactions. The quenched reaction mixture was extracted with EtOAc (250 μL, 3×), concentrated, and reconstituted in 100 μL MeOH before HPLC analysis. The data points were collected from experiments performed in triplicate.

When acyl-SNACs were used as substrates, the assays were performed in a total volume of 150 µL containing 0.15 mM acyl-SNACs, 5 mM ATP, 10 mM $MgCl_2$, 1.2 mM Na-malonate, 0.6 mM coenzyme A, 5 mM TCEP, 5 µM MatB, 7.5 µM holo-PltB module1 or holo-CalA module1, 7.5 µM holo-PltB module2, and 400 mM potassium phosphate buffer (pH=7.5). For *trans*-AT system, 7.5 µM holo-CalA module1 S633A together with equimolar FabD were added to the enzyme reaction instead of wild-type holo-CalA. The reactions were incubated at 30 ˚C for 4.5 h before being quenched with 75 µL MeOH containing 10% v/v formic acid and 0.03 mM internal standard (4-hydroxy-6-methyl-2-pyrone). For time-course assay, the reactions were performed in a total volume of 400 µL; 70 µL aliquots were withdrawn at 1, 2, 4, 6, 8 h and added to 30 µL MeOH with 10% v/v formic acid. Each experiment was conduct in triplicate. The quenched reactions were analyzed by LC-MS or HPLC.

Organic extraction of *Pseudomonas protegens* Pf-5

The bacterium *P. protegens* Pf-5 was inoculated into 50 mL LB media from glycerol stock and grown at 30 ˚C for 24 h. Cell pellets and liquid media were separated by centrifugation at 2,000×g for 20 min. The supernatant was acidified with 3 M HCl to pH = 2.0, and then extracted with EtOAc (50 mL, 3×). The organic layers were combined, dried with anhydrous Na_2SO_4, and concentrated under vacuum. The concentrated residue was dissolved in 1 mL MeOH and analyzed by HPLC-HRMS in the negative ionization mode.

LC-MS and HPLC analysis

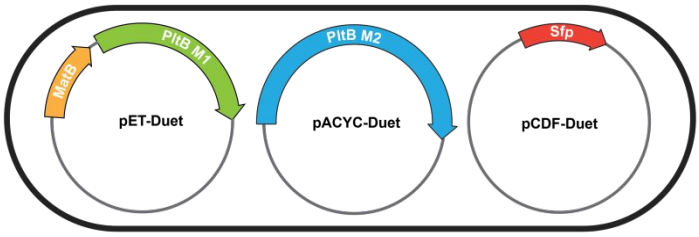
High resolution mass spectrometry (HRMS) data for synthesized acyl-SNACs were collected on an Agilent 1290 Infinity II UHPLC system coupled to a Bruker Impact II ToF mass spectrometer operating at room temperature with Kinetex 1.7 µm C18 100 Å column (50×2.1 mm). Mixture of water and acetonitrile with 0.1% formic acid was used as the mobile phase at a flow rate of 0.5 $mL \cdot min^{-1}$. All MS data for acyl-SNACs were collected in positive mode.

Production of pyrone analogs were monitored by either LC-MS or HPLC. HPLC analyses to monitor the production of **6** and **7** were carried out using Phenomenex Luna 5 µm C8(2) 100 Å LC column (250×4.6 mm) using Agilent 1260 Infinity II HPLC system. Water (solvent A) and MeCN (solvent B) with 0.1 % TFA were used as the mobile phase. A flow rate of 0.5 $mL \cdot min^{-1}$ was used with the following gradient: 0–5 min: 5% B, 5–30 min: linear gradient to 100% B, 30–34 min: 100% B, 34–35 min: linear gradient to 5% B, 35–36 min: 5 % B, 36–37 min: linear gradient to 100% B, 37–38 min: 100% B, 38–39 min: linear

gradient to 5% B. UV absorbance was monitored at 352 nm. LC-MS analyses were carried out on Agilent Poroshell 120 EC-C18 column (2.7 µm, 4.6 × 100 mm) using Agilent 1260 Infinity HPLC coupled to a Bruker amaZon SL operating in the negative ionization mode. Water (solvent A) and MeCN (solvent B) with 0.1 % formic acid were used as the mobile phase. A flow rate of 0.5 mL·min^{-1} was used with the following gradient: 0–3 min: 5% B, 3–16 min: linear gradient to 100% B, 16–20 min: 100% B, 20–21 min: linear gradient to 5% B, 21–22 min: 5% B, 22–23 min: linear gradient to 100% B, 23–24 min: 100% B, 24–25 min: linear gradient to 5% B, 25–27 min: 5% B. Ions corresponding to pyrones and internal standard were extracted, and peak areas for extracted ion counts were integrated. Ion count ratio between pyrone product and internal standard was calculated, normalized based on monoisotopic abundance, and averaged between three triplicate experiments. To represent relative activity, activity of each Plt PKS system (native or engineered) towards dichloropyrrole-2-carbonyl-SNAC is set as 1. Production level of other pyrone analogs by native or engineered PKSs was normalized to the production level of **6** by each PKS system. The same UHPLC-HRMS system was used to collect HRMS data for pyrone analogs except that the MS data for pyrone analogs were collected in the negative mode.

Organic extracts from *P. protegens* Pf-5 were analyzed by UHPLC-HRMS system operating in negative mode on Thermo Scientific™ Accucore™ Phenyl-X column (2.6 µm, 2.1×50 mm) with the following gradient: 0–2 min: 5% B, 2–12 min: linear gradient to 100% B, 12–16 min: 100% B, 16–17 min: linear gradient to 5% B, 17–18 min: 5% B, 18–19 min: linear gradient to 100% B, 19–20 min: 100% B, 20–21 min: linear gradient to 5% B, 21–22 min: 5% B.

Large scale production of dichloropyrrolyl pyrone



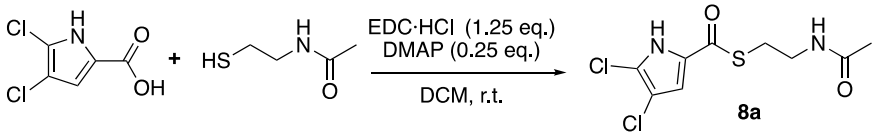
Expression vectors, pETDuet-1 carrying MatB and PltB module1 DEBS2dd, pACYCDuet-DEBS3dd PltB module2 and pCDFDuet-Sfp, were co-transformed into *E. coli* BL21Gold(DE3). BL21 cells with expression plasmids were cultured under the same conditions as described for protein overexpression. After 24 h induction, *E. coli* cells were harvested by centrifugation at 2,000×g for 25 min. Cell pellets collected from 2 L culture were combined, resuspended in 60 mL buffer containing 400 mM potassium phosphate pH=7.5 and 10% glycerol, and sonicated for 40 min. The lysate was clarified by centrifugation

and used directly for large scale enzyme assays which was composed of 10 mM ATP, 10 mM $MgCl_2$, 10 mM Na-malonate, 0.2 mM coenzyme A, 0.4 mM dichloropyrrole-2-carbonyl-SNAC, 5 mM TCEP, 5 mM calcium pantothenate, 400 mM potassium phosphate pH=7.5 and 10% glycerol in a total volume of 100 mL. ATP, Na-malonate and dichloropyrrole-2-carbonyl-SNAC were divided into 4 batches and added every 2 h. The reaction mixture was shaken at 60 rpm at 30 ˚C for 24 h before quenched with TFA to pH = 2.0. The acidified mixture with precipitate was extracted with EtOAC (100 mL, 3×). The organic and aqueous layers were separated by centrifugation at 1,000×g for 5 min. The organic layers were combined, dried with anhydrous Na_2SO_4, and concentrated under vacuum. The crude extract was first purified by silica flash column (0–100% EtOAc in hexane). Fractions contain desired products (checked by LC-MS) were pooled, concentrated, and further purified by two rounds of preparative HPLC. For the first round of preparative HPLC purification, it was carried out on Luna 5 µm C18(2) 100 Å LC column (250×10 mm) using Agilent 1260 Infinity II HPLC system, using mixture of water (solvent A) and acetonitrile (solvent B) with 0.1 % TFA as the mobile phase. A flow rate of 2 $mL \cdot min^{-1}$ was used with the following gradient: 0–5 min: 5% B, 5–30 min: linear gradient to 50% B, 30–36 min: 50% B, 36–42 min: linear gradient to 100% B, 42–46 min: 100% B, 46–47 min: linear gradient to 5% B, 47–48 min: 5 % B, 48–49 min: linear gradient to 100% B, 49–50 min: 100% B, 50–51 min: linear gradient to 5% B. The second round of purification was achieved on Luna 5 µm C8(2) 100 Å LC column (250×4.6 mm) with isocratic elution at 50% B using the same solvent system at a flow rate of 0.5 $mL \cdot min^{-1}$. The product was monitored at UV 352 and 254 nm. Purified molecule was analyzed by HRMS and NMR spectra.

SUPPLEMENTARY SCHEMES

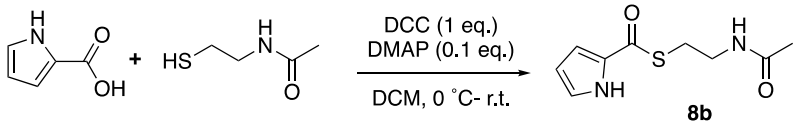
Synthesis of acyl-SNACS

Scheme S1 synthesis of 4,5-dichloro-1H-pyrrole-2-carbonyl-SNAC (**8a**)



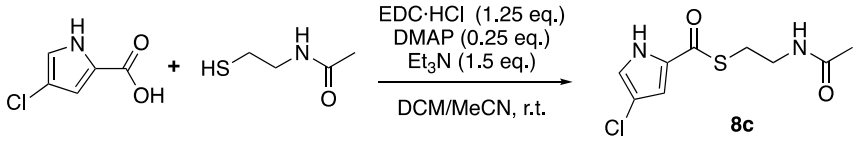
The compound 4,5-dichloro-1H-pyrrole-2-carboxylic acid was synthesized as previous described.[1, 4] Dichloropyrrolyl-2-carboxylic acid (72.7 mg, 0.61 mmol, 1 eq.), N-acetylcysteamine (109.7 mg, 0.61 mmol, 1 eq.), EDC·HCl (146.1 mg, 0.76 mmol, 1.25 eq.), and DMAP (22.3 mg, 0.18 mmol, 0.3 eq.) were dissolved in 3 mL DCM at room temperature. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by silica flash column (50% EtOAc in hexane to 5% MeOH in EtOAc) to give **8a** as yellow solid (27.5 mg, 16 %). ^{1}H NMR (500 MHz, MeOD) δ 8.26 (br s, 1H), 6.93 (s, 1H), 3.39 (td, J = 6.7, 5.0 Hz, 2H), 3.15 (t, J = 6.6 Hz, 2H), 1.93 (s, 3H). ^{13}C NMR (126 MHz, MeOD) δ 180.91, 173.53, 128.98, 120.91, 115.16, 111.75, 40.41, 28.47, 22.48. HRMS (ESI) *m/z* calculated for $C_9H_{11}Cl_2N_2O_2S$ ($[M+H]^+$) 280.9913, found 280.9915.

Scheme S2 synthesis of 1H-pyrrole-2-carbonyl-SNAC (**8b**)



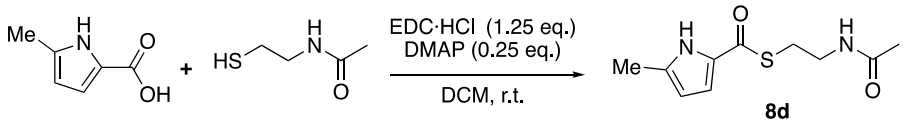
The molecule pyrrolyl-2-carbonyl-SNAC was prepared as described in literature and the ^{1}H NMR spectra matched the previous report.[5] Pyrrole-2-carboxylic acid (300 mg, 2.7 mmol, 1 eq.), N-acetylcysteamine (321.9 mg, 2.7 mmol, 1 eq.), and DMAP (33.0 mg, 0.27 mmol, 0.1 eq.) were dissolved in 8.2 mL dry DCM at 0 ˚C. A solution of DCC (557.2 mg, 2.7 mmol, 1 eq.) dissolved in 5.4 mL dry DCM was then added dropwise at 0 ˚C. The reaction mixture was stirred overnight from 0 ˚C to room temperature before concentration under vacuum. The crude mixture was purified by silica flash column (50% EtOAc in hexane to EtOAc) to give **8b** as white solid (272.3 mg, 48 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.34 (br s, 1H), 7.03 (tdd, J = 3.9, 2.6, 1.3 Hz, 2H), 6.29 (dt, J = 3.9, 2.5 Hz, 1H), 6.13 (br s, 1H), 3.52 (dt, J = 6.5, 5.6 Hz, 2H), 3.19 (dd, J = 6.9, 5.7 Hz, 2H), 1.98 (s, 3H). HRMS (ESI) *m/z* calculated for $C_9H_{13}N_2O_2S$ ($[M+H]^+$) 213.0692, found 213.0694.

Scheme S3 synthesis of 4-chloro-1H-pyrrole-2-carbonyl-SNAC (**8c**)



4-Chloro-1H-pyrrole-2-carboxylic acid (60 mg, 0.41 mmol, 1 eq.), N-acetylcysteamine (49.1 mg, 0.41 mmol, 1 eq.), EDC·HCl(98.8 mg, 0.52 mmol, 1.25 eq.), and DMAP (12.6 mg, 0.10 mmol, 0.25 eq.) were dissolved in 4.2 mL 2:1 DCM/acetonitrile at room temperature, followed by the addition of Et_3N (62.6 mg, 0.62 mmol, 1.5 eq.). The reaction was stirred at room temperature overnight before concentrated under vacuum. The reaction mixture was first purified by silica flash chromatography (20% hexane in EtOAc to EtOAc), followed by preparative HPLC purification to give **8c** as white solid (4.1 mg, 4 %). Preparative HPLC purification was carried out on Luna 5 µm C18(2) 100 Å LC column (250×10 mm) using Agilent 1260 Infinity II HPLC system using mixture of water (buffer A) and acetonitrile (buffer A) containing 0.1% TFA as the mobile phase. A flow rate of 2 mL·min^{-1} was used with the following gradient: 0-5 min: 5% B, 5-30 min: linear gradient to 60% B, 30-36 min: 60% B, 36-38 min: linear gradient to 100% B, 38-43 min: 100% B, 43-44 min: linear gradient to 5% B, 44-45 min: 5 % B, 45-46 min: linear gradient to 100% B, 46-47 min: 100% B, 47-48 min: linear gradient to 5% B. ^{1}H NMR (500 MHz, MeOD) δ 7.02 (d, *J* = 1.6 Hz, 1H), 6.87 (d, *J* = 1.6 Hz, 1H), 3.39 (t, *J* = 6.6 Hz, 2H), 3.14 (t, *J* = 6.6 Hz, 2H), 1.93 (s, 3H). ^{13}C NMR (126 MHz, MeOD) δ 181.63, 173.52, 130.55, 123.02, 114.69, 114.60, 40.50, 28.40, 22.48. HRMS (ESI) *m/z* calculated for $C_9H_{12}ClN_2O_2S([M+H]^+)$ 247.0303, found 247.0306

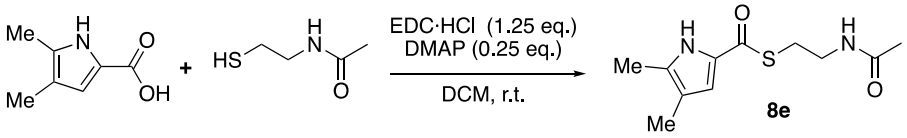
Scheme S4 synthesis of 5-methyl-1H-pyrrole-2-carbonyl-SNAC (**8d**)



5-Methyl-1H-pyrrole-2-carboxylic acid was prepared as previously described.[1, 6] 5-Methyl-pyrrole-2-carboxylic acid (100 mg, 0.80 mmol, 1 eq.), N-acetylcysteamine (95.3 mg, 0.80 mmol, 1 eq.), EDC·HCl (191.5 mg, 1.0 mmol, 1.25 eq.), and DMAP (24.4 mg, 0.20 mmol, 0.25 eq.) were dissolved in 4 mL DCM at room temperature. The reaction was stirred at room temperature overnight before concentrated under vacuum. The crude mixture was purified by silica flash chromatography (DCM to 5% MeOH in DCM) to give **8d** as white solid (115.9 mg, 64 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.95 (br s, 1H), 6.94 (dd, *J* = 3.9, 2.5 Hz, 1H), 6.30 (br s, 1H), 5.99 (ddd, *J* = 3.7, 2.7, 0.8 Hz, 1H), 3.52 (q, *J* = 5.8 Hz, 2H), 3.17 (dd, *J* = 6.9, 5.6 Hz, 2H), 2.32 (s, 3H), 2.01 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 180.83, 170.77, 135.85, 128.68,

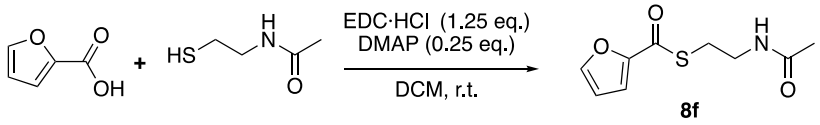
117.27, 110.01, 40.49, 27.72, 23.26, 13.38. HRMS (ESI) *m/z* calculated for $C_{10}H_{15}N_2O_2S$ ($[M+H]^+$) 227.0849, found 227.0852.

Scheme S5 synthesis of 4,5-dimethyl-1H-pyrrole-2-carbonyl-SNAC (**8e**)



4, 5-Dimethyl-1H-pyrrole-2-carboxylic acid (60 mg, 0.43 mmol, 1 eq.), N-acetylcysteamine (51.4 mg, 0.43 mmol, 1 eq.), EDC·HCl (103.3 mg, 0.54 mmol, 1.25 eq.), and DMAP (13.2 mg, 0.11 mmol, 0.25 eq.) were dissolved in 2.2 mL DCM at room temperature. The reaction was stirred at room temperature overnight before concentrated under vacuum. The crude mixture was purified by silica flash chromatography (20% hexane in EtOAc to 2% MeOH in EtOAc) to give **8e** as pink solid (76.7 mg, 74 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.04 (br s, 1H), 6.80 (d, *J* = 2.7 Hz, 1H), 6.23 (br s, 1H), 3.50 (q, *J* = 5.8 Hz, 2H), 3.15 (dd, *J* = 6.8, 5.6 Hz, 2H), 2.22 (s, 3H), 2.01 (s, 3H), 1.97 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 180.37, 170.72, 133.15, 126.95, 118.66, 117.91, 40.58, 27.66, 23.27, 11.57, 10.95. HRMS (ESI) *m/z* calculated for $C_{11}H_{17}N_2O_2S$ ($[M+H]^+$) 241.1005, found 241.1008

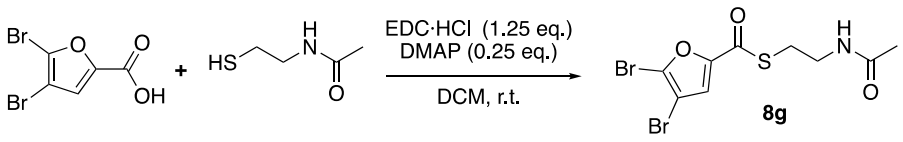
Scheme S6 synthesis of furan-2-carbonyl-SNAC (**8f**)



2-Furoic acid (200 mg, 1.78 mmol, 1 eq.), N-acetylcysteamine (255.2 mg, 2.14 mmol, 1.2 eq.), EDC·HCl (427.6 mg, 2.23 mmol, 1.25 eq.), and DMAP (54.5 mg, 0.45 mmol, 0.25 eq.) were dissolved in 8.9 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (90% EtOAc in hexane to 5% MeOH in EtOAc) to give **8f** as white solid (312.2 mg, 82 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.60 (dd, *J* = 1.7, 0.8 Hz, 1H), 7.22 (dd, *J* = 3.6, 0.8 Hz, 1H), 6.56 (dd, *J* = 3.6, 1.7 Hz, 1H), 6.07 (br s, 1H), 3.57 – 3.48 (m, 2H), 3.21 (dd, *J* = 6.8, 5.9 Hz, 2H), 1.99 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ

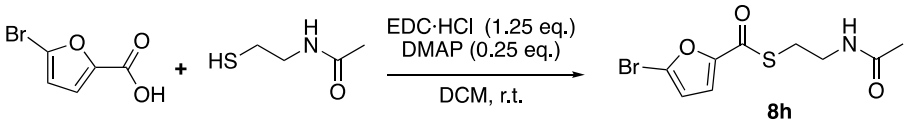
180.81, 170.71, 150.65, 146.70, 116.32, 112.56, 39.93, 27.84, 23.29. HRMS (ESI) *m/z* calculated for $C_9H_{12}NO_3S$ ([M+H]$^+$) 214.0532, found 214.0534

Scheme S7 synthesis of 4,5-dibromofuran-2-carbonyl-SNAC (**8g**)



4,5-Dibromofuran-2-carboxylic acid (300 mg, 1.11 mmol, 1 eq.), N-acetylcysteamine (132.5 mg, 1.11 mmol, 1 eq.), EDC·HCl (266.4 mg, 1.39 mmol, 1.25 eq.), and DMAP (34.0 mg, 0.28 mmol, 0.25 eq.) were dissolved in 5.6 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (75% EtOAc in hexane to 5% MeOH in EtOAc) to give **8g** as white solid (249.9 mg, 61 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.20 (s, 1H), 6.01 (br s, 1H), 3.52 (q, *J* = 6.1 Hz, 2H), 3.22 (t, *J* = 6.4 Hz, 2H), 1.99 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 179.21, 170.81, 151.62, 129.16, 119.86, 104.86, 39.65, 28.12, 23.25. HRMS (ESI) *m/z* calculated for $C_9H_{10}Br_2NO_3S$ ([M+H]$^+$) 369.8743, found 369.8739.

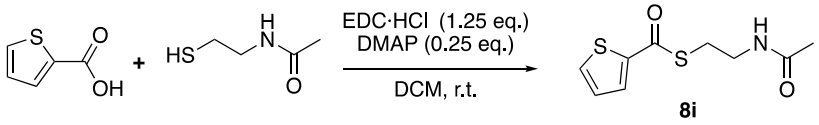
Scheme S8 synthesis of 5-bromofuran-2-carbonyl-SNAC (**8h**)



5-Bromofuran-2-carboxylic acid (180 mg, 0.94 mmol, 1 eq.), N-acetylcysteamine (112.3 mg, 0.94 mmol, 1 eq.), EDC·HCl (225.9 mg, 1.18 mmol, 1.25 eq.), and DMAP (28.8 mg, 0.24 mmol, 0.25 eq.) were dissolved in 4.7 mL DCM at room temperature. The reaction was stirred at room temperature for 2 days before quenched with saturated NH_4Cl (10 mL). The organic phase was then washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 4×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (80% EtOAc in hexane to 90% EtOAc in hexane) to give **8h** as white solid (207.8 mg, 75 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.15 (d, *J* = 3.6 Hz, 1H), 6.50 (d, *J* = 3.6 Hz, 1H), 6.05 (br s, 1H), 3.52 (q, *J* = 6.2 Hz, 2H), 3.21 (dd, *J* = 6.8, 5.9 Hz, 2H), 1.99 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 179.54,

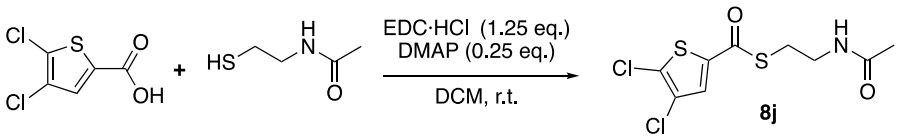
170.78, 152.14, 128.62, 118.27, 114.64, 39.80, 27.96, 23.28. HRMS (ESI) *m/z* calculated for $C_9H_{11}BrNO_3S$ ($[M+H]^+$) 291.9638, found 291.9636.

Scheme S9 synthesis of thiophene-2-carbonyl-SNAC (**8i**)



Thiophene-2-carboxylic acid (100 mg, 0.78 mmol, 1 eq.), N-acetylcysteamine (93.0 mg, 0.78 mmol, 1 eq.), EDC·HCl (187.0 mg, 0.98 mmol, 1.25 eq.), and DMAP (23.8 mg, 0.20 mmol, 0.25 eq.) were dissolved in 3.9 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (90% EtOAc in hexane to 5% MeOH in EtOAc) to give **8i** as white solid (159.5 mg, 89 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 (d, J = 3.9 Hz, 1H), 7.65 (d, J = 4.9 Hz, 1H), 7.13 (t, J = 4.4 Hz, 1H), 6.08 (br s, 1H), 3.53 (q, J = 6.1 Hz, 2H), 3.22 (t, J = 6.4 Hz, 2H), 1.98 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 184.40, 170.67, 141.69, 133.36, 131.75, 128.21, 39.96, 28.86, 23.35. HRMS (ESI) *m/z* calculated for $C_9H_{12}NO_2S_2$ ($[M+H]^+$) 230.0304, found 230.0301

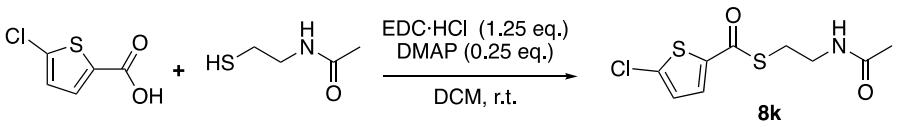
Scheme S10 synthesis of 4,5-dichlorothiophene-2-carbonyl-SNAC (**8j**)



4,5-Dichlorothiophene-2-carboxylic acid (180 mg, 0.91 mmol, 1 eq.), N-acetylcysteamine (108.9 mg, 0.91 mmol, 1 eq.), EDC·HCl (218.9 mg, 1.14 mmol, 1.25 eq.), and DMAP (27.9 mg, 0.23 mmol, 0.25 eq.) were dissolved in 4.6 mL DCM at room temperature. The reaction was stirred at room temperature for 2 days before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (80% EtOAc in hexane to 90% EtOAc in hexane) to give **8j** as white solid (209.6 mg, 77 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.58 (s, 1H), 5.97 (br s, 1H), 3.52 (q, J = 6.2 Hz, 2H), 3.23 (t, J = 6.4

Hz, 2H), 1.99 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 183.01, 170.67, 137.37, 133.60, 130.70, 125.55, 39.67, 29.02, 23.33. HRMS (ESI) *m/z* calculated for $C_9H_{10}Cl_2NO_2S_2$ ($[M+H]^+$) 297.9525, found 297.9522.

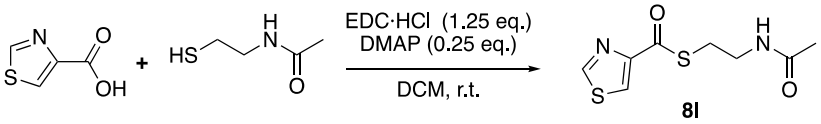
Scheme S11 synthesis of 5-chlorothiophene-2-carbonyl-SNAC (**8k**)



8k

5-Chlorothiophene-2-carboxylic acid (200 mg, 1.23 mmol, 1 eq.), N-acetylcysteamine (146.6 mg, 1.23 mmol, 1 eq.), EDC·HCl (294.8 mg, 1.54 mmol, 1.25 eq.), and DMAP (37.6 mg, 0.31 mmol, 0.25 eq.) were dissolved in 6.2 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (80% DCM in hexane to 5% MeOH in DCM) to give **8k** as white solid (279.0 mg, 86 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.59 (d, *J* = 4.1 Hz, 1H), 6.95 (d, *J* = 4.1 Hz, 1H), 6.20 (br s, 1H), 3.53 – 3.49 (m, 2H), 3.21 (dd, *J* = 6.8, 6.1 Hz, 2H), 1.98 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 183.39, 170.74, 139.88, 139.30, 131.20, 127.66, 40.43, 28.80, 23.25. HRMS (ESI) *m/z* calculated for $C_9H_{11}ClNO_2S_2$ ($[M+H]^+$) 263.9914, found 263.9911.

Scheme S12 synthesis of thiazole-4-carbonyl-SNAC (**8l**)

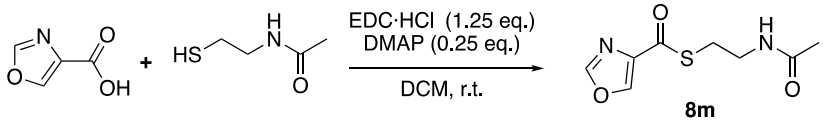


8l

Thiazole-4-carbonyl-SNAC was synthesized as previously described and the spectra matched the previous report.[7] Thiazole-4-carboxylic acid (100 mg, 0.77 mmol, 1 eq.), N-acetylcysteamine (92.3 mg, 0.77 mmol, 1 eq.), EDC·HCl (185.6 mg, 0.97 mmol, 1.25 eq.), and DMAP (23.7 mg, 0.19 mmol, 0.25 eq.) were dissolved in 3.9 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (75% EtOAc in hexane to 10% MeOH in EtOAc) to give **8l** as white solid (126.9 mg, 71 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.87 (d, *J* = 2.1 Hz, 1H), 8.19 (d, *J* = 2.1 Hz, 1H), 5.99 (br s, 1H),

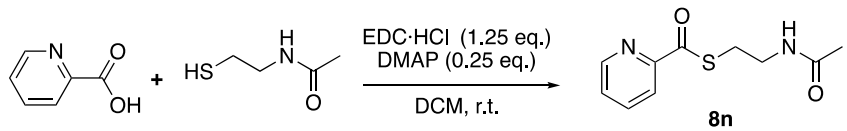
3.59 – 3.52 (m, 2H), 3.24 (dd, *J* = 6.8, 5.7 Hz, 2H), 1.99 (s, 5H). HRMS (ESI) *m/z* calculated for $C_8H_{11}N_2O_2S_2$ ([M+H]$^+$) 231.0256, found 231.0258

Scheme S13 synthesis of oxazole-4-carbonyl-SNAC (**8m**)



Oxazole-4-carbonyl-SNAC was synthesized as previously described and the spectra matched the previous report.[7] Oxazole-4-carboxylic acid (150 mg, 1.33 mmol, 1 eq.), N-acetylcysteamine (158.1 mg, 1.33 mmol, 1 eq.), EDC·HCl (317.9 mg, 1.66 mmol, 1.25 eq.), and DMAP (40.5 mg, 0.33 mmol, 0.25 eq.) were dissolved in 6.6 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (DCM to 5% MeOH in DCM) to give **8m** as white solid (217.9 mg, 77 %).[1]H NMR (400 MHz, $CDCl_3$) δ 8.29 (d, *J* = 1.0 Hz, 1H), 7.96 (d, *J* = 1.0 Hz, 1H), 6.12 (br s, 1H), 3.59 – 3.53 (m, 2H), 3.25 (dd, *J* = 6.8, 5.8 Hz, 2H), 2.03 (s, 3H). HRMS (ESI) *m/z* calculated for $C_8H_{11}N_2O_3S$ ([M+H]$^+$) 215.0485, found 215.0486

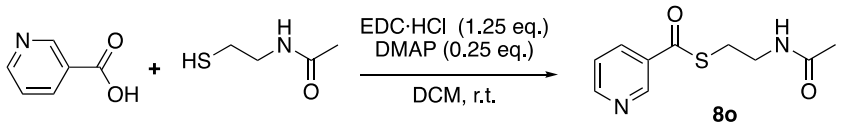
Scheme S14 synthesis of picolinoyl-SNAC (**8n**)



Picolinoyl-SNAC was synthesized as previously described and the spectra matched the previous report.[7] Picolinic acid (154.9 mg, 1.26 mmol, 1 eq.), N-acetylcysteamine (150 mg, 1.26 mmol, 1 eq.), EDC·HCl (301.6 mg, 1.57 mmol, 1.25 eq.), and DMAP (38.4 mg, 0.32 mmol, 0.25 eq.) were dissolved in 6.3 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (90% EtOAc in hexane to 10% MeOH in EtOAc) to give **8n** as white solid (226.0 mg, 80 %). [1]H NMR (400 MHz, $CDCl_3$) δ 8.71 (ddd, *J* = 4.8, 1.7, 0.9 Hz, 1H), 7.97 (dt, *J* = 7.8, 1.1 Hz, 1H), 7.88 (td, *J* = 7.7, 1.7 Hz,

1H), 7.55 (ddd, *J* = 7.5, 4.7, 1.3 Hz, 1H), 5.98 (br s, 1H), 3.58 – 3.51 (m, 2H), 3.22 (dd, *J* = 6.9, 5.7 Hz, 2H), 1.98 (s, 3H). HRMS (ESI) *m/z* calculated for $C_{10}H_{13}N_2O_2S$ ($[M+H]^+$) 225.0692, found 225.0694

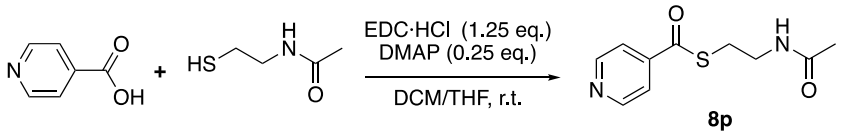
Scheme S15 synthesis of nicotinoyl-SNAC (**8o**)



8o

Nicotinoyl-SNAC was synthesized as previously described and the spectra matched the previous report.[7] Nicotinic acid (150 mg, 1.22 mmol, 1 eq.), N-acetylcysteamine (145.2 mg, 1.22 mmol, 1 eq.), EDC·HCl (292.0 mg, 1.52 mmol, 1.25 eq.), and DMAP (37.2 mg, 0.31 mmol, 0.25 eq.) were dissolved in 6.1 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (90% EtOAc in hexane to 15% MeOH in EtOAc) to give **8o** as white solid (231.8 mg, 85 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.21 (d, *J* = 2.1 Hz, 1H), 8.87 – 8.82 (m, 1H), 8.36 (dt, *J* = 8.0, 1.9 Hz, 1H), 7.61 – 7.54 (m, 1H), 5.98 (br s, 1H), 3.56 (q, *J* = 6.3 Hz, 2H), 3.30 (t, *J* = 6.5 Hz, 2H), 1.99 (s, 3H). HRMS (ESI) *m/z* calculated for $C_{10}H_{13}N_2O_2S$ ($[M+H]^+$) 225.0692, found 225.0695.

Scheme S16 synthesis of isonicotinoyl SNAC (**8p**)

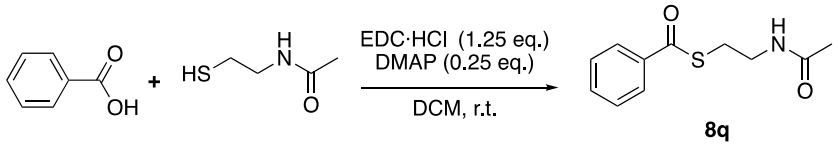


8p

Isonicotinic acid (150 mg, 1.22 mmol, 1 eq.), N-acetylcysteamine (145.2 mg, 1.22 mmol, 1 eq.), EDC·HCl (292.0 mg, 1.52 mmol, 1.25 eq.), and DMAP (37.2 mg, 0.31 mmol, 0.25 eq.) were dissolved in 6.1 mL DCM and 1.5 mL THF at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (EtOAc to 10% MeOH in EtOAc) to give **8p** as white solid (231.8 mg, 85 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.84 – 8.79 (m, 2H), 7.82 – 7.77 (m, 2H), 5.94 (br s, 1H), 3.54 (q, *J* = 6.3 Hz, 2H),

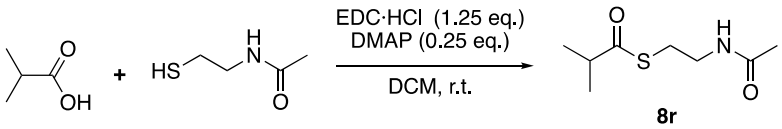
3.28 (t, J = 6.5 Hz, 2H), 1.98 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.24, 170.54, 150.18, 143.58, 120.82, 39.30, 29.12, 23.38. HRMS (ESI) *m/z* calculated for $C_{10}H_{13}N_2O_2S$ ($[M+H]^+$) 225.0692, found 225.0693.

Scheme S17 synthesis of benzoyl-SNAC (**8q**)



Benzoyl-SNAC was synthesized as previously described and the spectra matched the previous report.[7] Benzoic acid (150 mg, 1.23 mmol, 1 eq.), N-acetylcysteamine (146.4 mg, 1.23 mmol, 1 eq.), EDC·HCl (294.3 mg, 1.54 mmol, 1.25 eq.), and DMAP (37.5 mg, 0.31 mmol, 0.25 eq.) were dissolved in 6.1 mL DCM at room temperature. The reaction was stirred at room temperature overnight before the addition of saturated NH_4Cl (10 mL). The organic phase was separated and washed with saturated $NaHCO_3$ (10 mL). The aqueous layer was extracted with EtOAc (10 mL, 3×). The combined organic layers were dried with anhydrous Na_2SO_4, concentrated under vacuum and the purified by silica flash chromatography (75% EtOAc in hexane to 5% MeOH in EtOAc) to give **8q** as white solid (260.5 mg, 95 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 – 7.94 (m, 2H), 7.62 – 7.57 (m, 1H), 7.50 – 7.44 (m, 2H), 5.93 (br s, 1H), 3.57 – 3.51 (m, 2H), 3.24 (dd, J = 6.9, 5.9 Hz, 2H), 1.98 (s, 3H). HRMS (ESI) *m/z* calculated for $C_{11}H_{14}NO_2S$ ($[M+H]^+$) 224.0740, found 224.0744

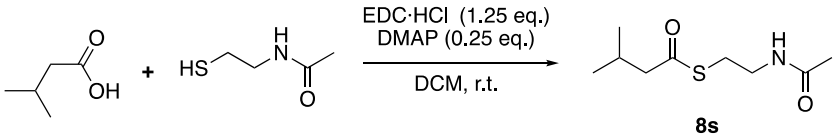
Scheme S18 synthesis of isobutyryl-SNAC (**8r**)



Isobutyric acid (120 mg, 1.36 mmol, 1 eq.), N-acetylcysteamine (162.3 mg, 1.36 mmol, 1 eq.), EDC·HCl (326.4 mg, 1.70 mmol, 1.25 eq.), and DMAP (41.6 mg, 0.34 mmol, 0.25 eq.) were dissolved in 6.8 mL DCM at room temperature. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by silica flash column (80% EtOAc in hexane to 90% EtOAc in hexane) to give **8r** as colorless oil (207.4 mg, 80 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.97 (br s, 1H), 3.47 – 3.39 (m, 2H), 3.01 (dd, J = 6.9, 6.0 Hz, 2H), 2.76 (m, 1H), 1.97 (s, 3H), 1.19 (d, J = 6.9 Hz, 6H). ^{13}C

NMR (126 MHz, $CDCl_3$) δ 204.81, 170.61, 43.25, 39.85, 28.19, 23.19, 19.46. HRMS (ESI) *m/z* calculated for $C_8H_{16}NO_2S$ ($[M+H]^+$) 190.0896, found 190.0896.

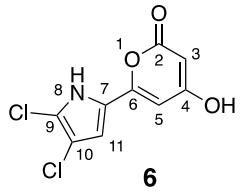
Scheme S19 synthesis of isovaleryl-SNAC (**8s**)



Isovaleric acid (150 mg, 1.47 mmol, 1 eq.), N-acetylcysteamine (175.2 mg, 1.47 mmol, 1 eq.), EDC·HCl (352.2 mg, 1.84 mmol, 1.25 eq.), and DMAP (44.8 mg, 0.37 mmol, 0.25 eq.) were dissolved in 7.3 mL DCM at room temperature. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by silica flash column (80% EtOAc in hexane to 90% EtOAc in hexane) to give **8s** as colorless oil (293.3 mg, 98 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.14 (br s, 1H), 3.40 (q, J = 6.2 Hz, 2H), 3.00 (t, J = 6.4 Hz, 2H), 2.47 – 2.40 (m, 2H), 2.13 (m, 1H), 1.94 (s, 3H), 0.93 (dd, J = 6.8, 2.0 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 199.72, 170.57, 52.95, 39.90, 28.45, 26.59, 23.21, 22.31. HRMS (ESI) *m/z* calculated for $C_9H_{18}NO_2S$ ($[M+H]^+$) 204.1053, found 204.1055.

SUPPLEMENTARY TABLES

Table S1: NMR assignments of **6** in DMSO-d_6



6

position	δ_H (*J* in Hz)	δ_C
2	-	162.5
3	5.26 (d, 2.0)	88.4
4	-	170.7
5	6.47 (d, 2.0)	95.0
6	-	153.4
7	-	122.8
9	-	116.0
10	-	109.7
11	6.86 (d, 2.7)	110.1
NH	12.99 – 12.90 (m)	-
OH	11.93 (s)	-

Table S2: Relative activity of native and engineered system towards a library of synthesized acyl-SNACs

Substrate	Engineered	Native
8a	1 ± 0.13	1 ± 0.04
8b	0.06 ± 0.01	0.00 ± 0.00
8c	0.39 ± 0.05	0.14 ± 0.04
8d	0.03 ± 0.00	0.00 ± 0.00
8e	0.01 ± 0.00	0.04 ± 0.01
8f	0.14 ± 0.02	0.02 ± 0.00
8g	3.16 ± 0.36	0.89 ± 0.23
8h	1.32 ± 0.17	0.36 ± 0.04
8i	0.45 ± 0.05	0.03 ± 0.01
8j	2.42 ± 0.60	1.03 ± 0.17
8k	2.96 ± 0.45	0.34 ± 0.08
8l	0.06 ± 0.00	0.01 ± 0.00
8m	0.02 ± 0.00	0.00 ± 0.00
8n	0.00 ± 0.00	0.00 ± 0.00
8o	0.23 ± 0.04	0.02 ± 0.00
8p	0.23 ± 0.00	0.02 ± 0.00
8q	0.28 ± 0.03	0.03 ± 0.00
8r	0.12 ± 0.00	Not detected
8s	0.38 ± 0.02	0.00 ± 0.00

SUPPLEMENTARY FIGURES

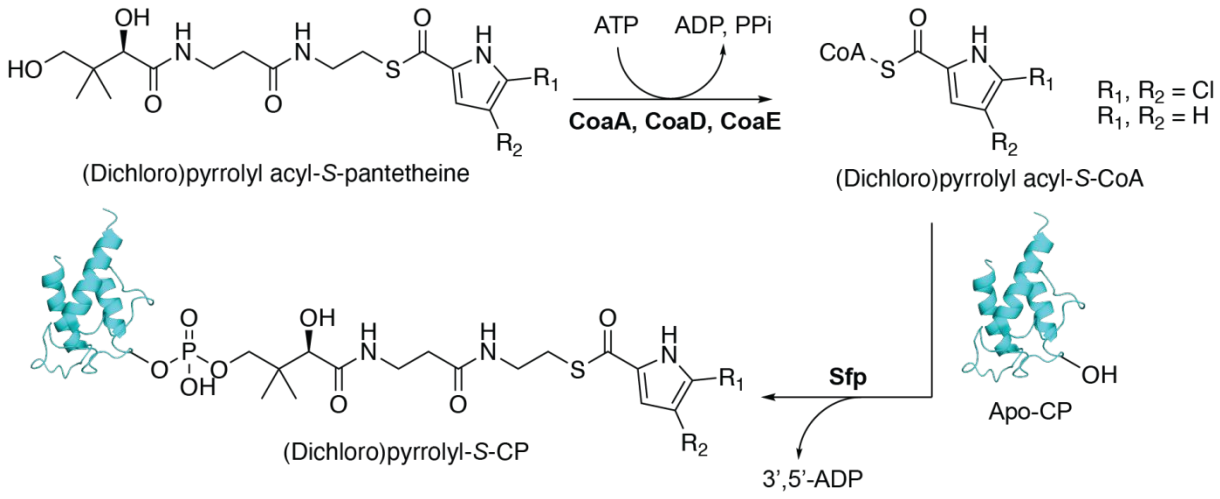


Fig. S1: Scheme for the one-pot chemoenzymatic synthesis of (dichloro)pyrrolyl-*S*-CPs starting from acyl-*S*-pantetheines using CoaA, CoaD, CoaE, and Sfp enzymes.

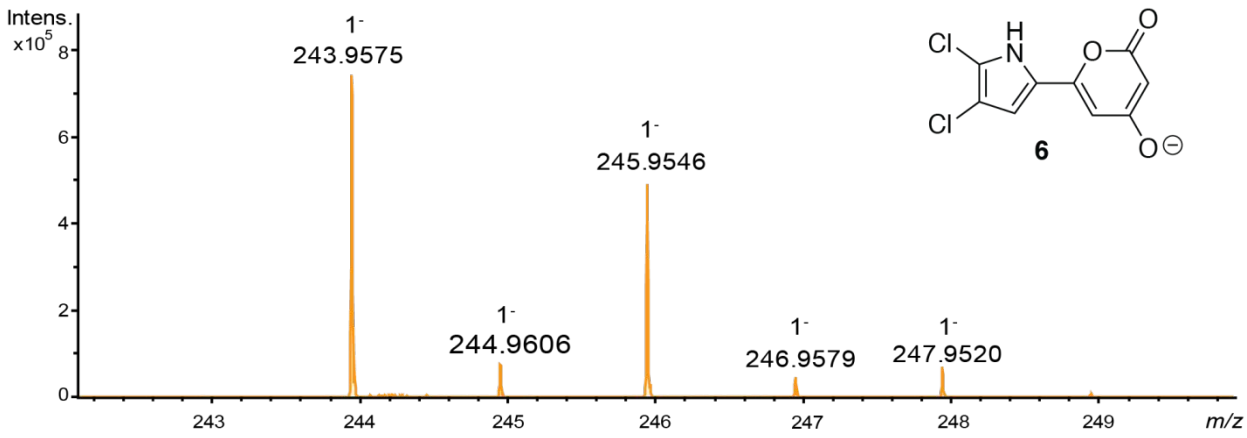


Fig. S2: $[M-H]^{1-}$ MS^1 spectra of compound **6**. HRMS (ESI) *m/z* calculated for $C_9H_4Cl_2NO_3$ ($[M-H]^-$) 243.9574, found 243.9575. Isotopic distribution pattern indicates the presence of two chlorine atoms.

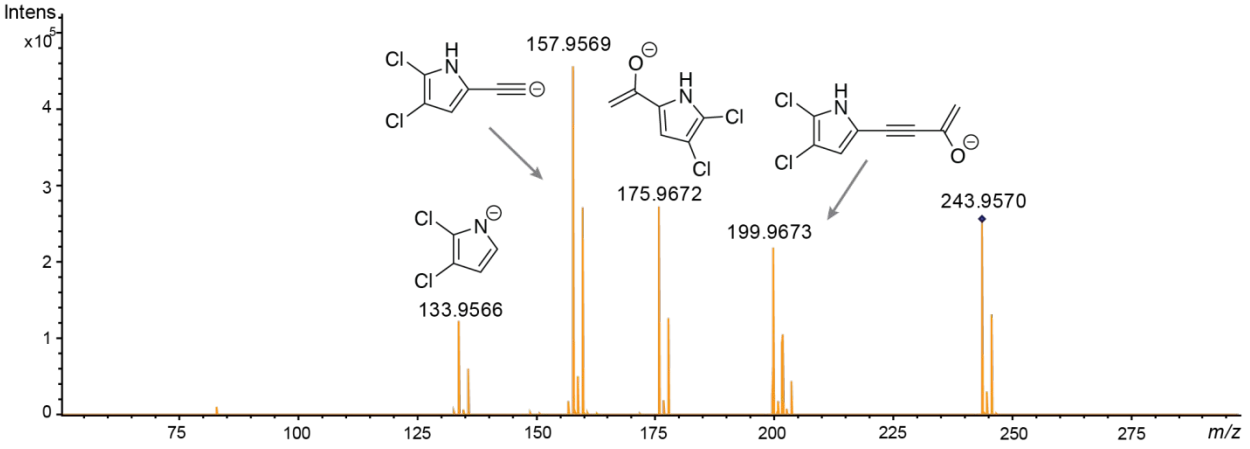


Fig. S3: MS^2 spectra of compound **6** with rationalized structural annotations of fragment ions.

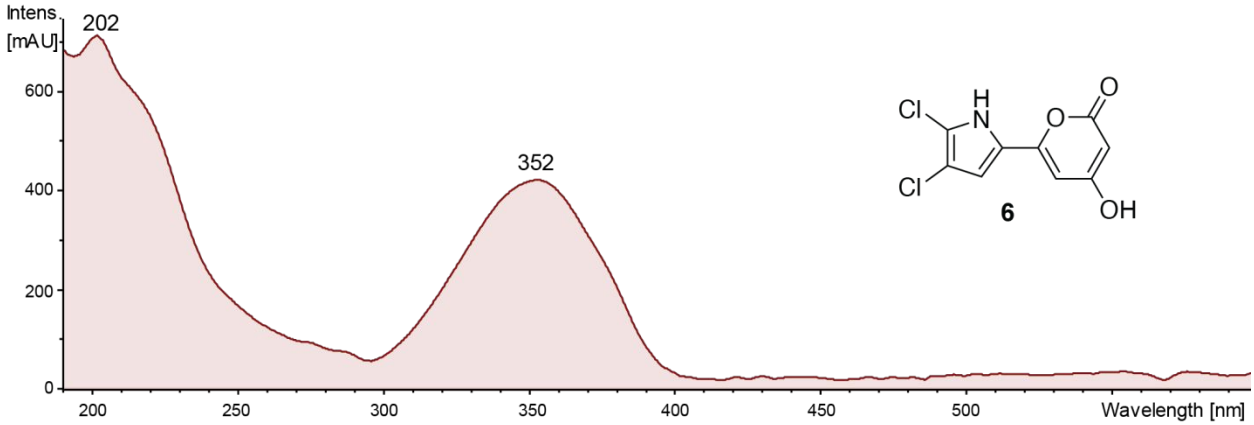


Fig. S4: UV-Vis absorbance profile of compound **6**.

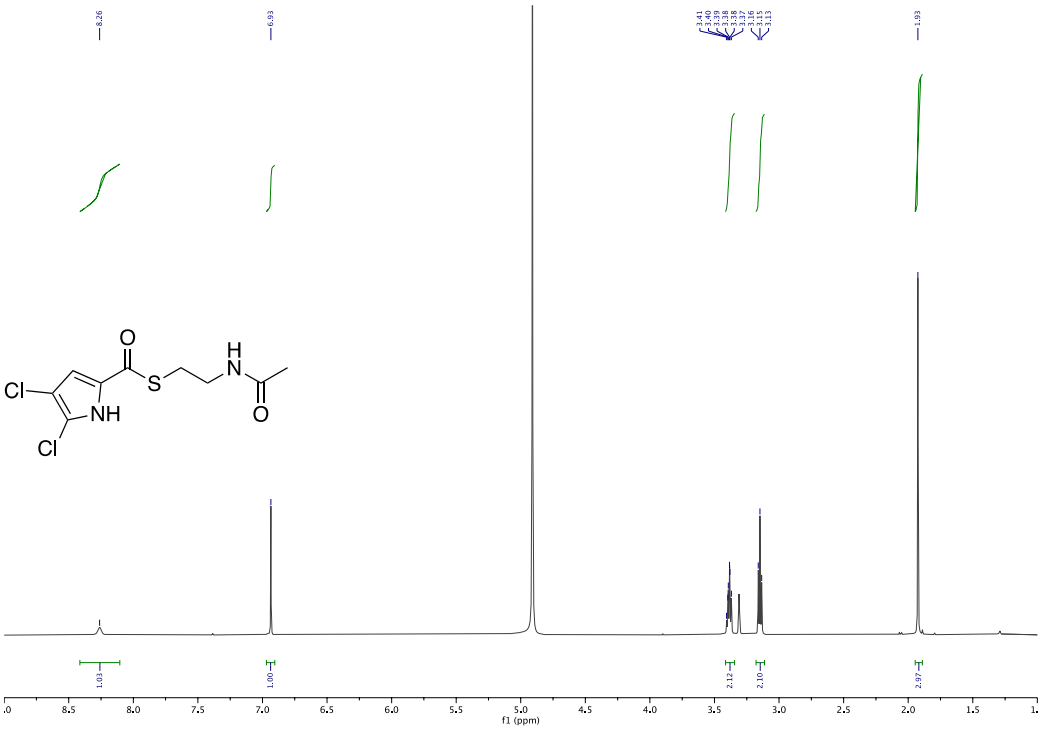


Fig. S5: [1]H NMR (500 MHz, MeOD) of compound **8a.**

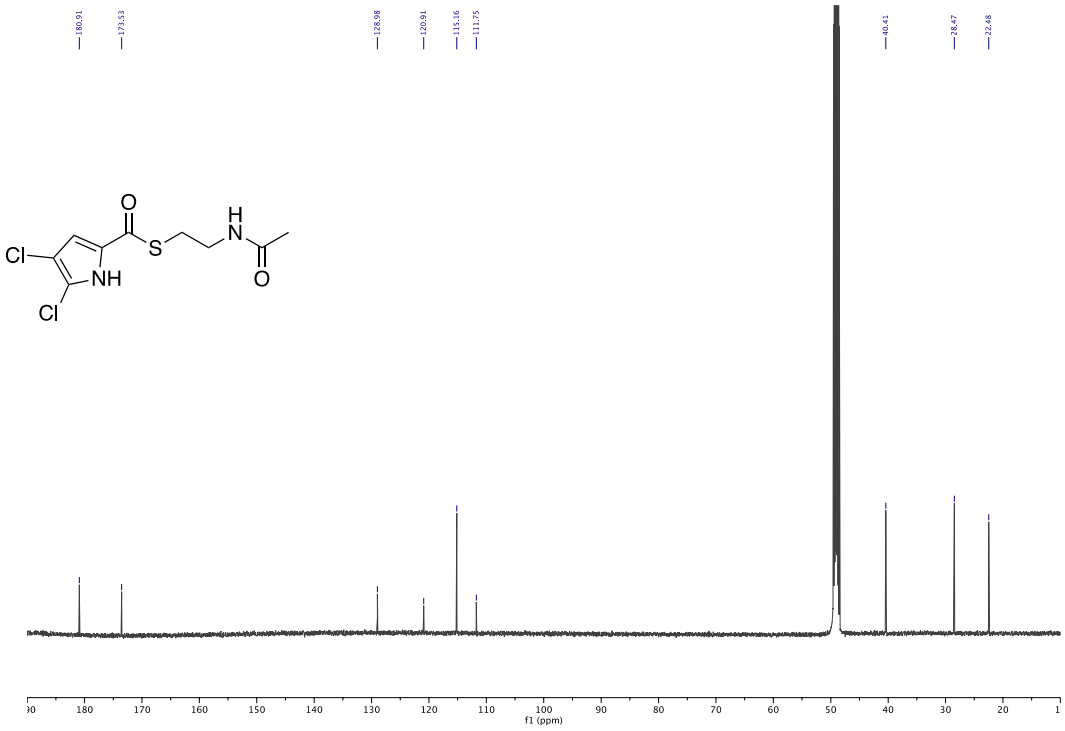


Fig. S6: [13]C NMR (126 MHz, MeOD) of compound **8a**.

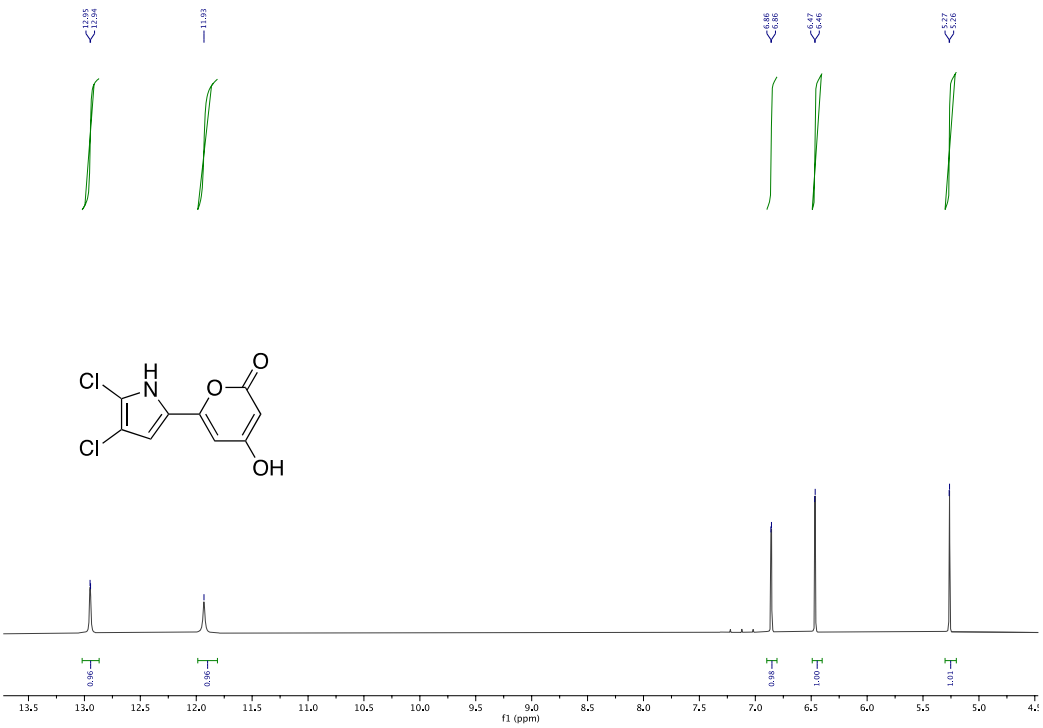


Fig. S7: ^{1}H NMR spectrum (500 MHz, DMSO-d_6) of compound **6**.

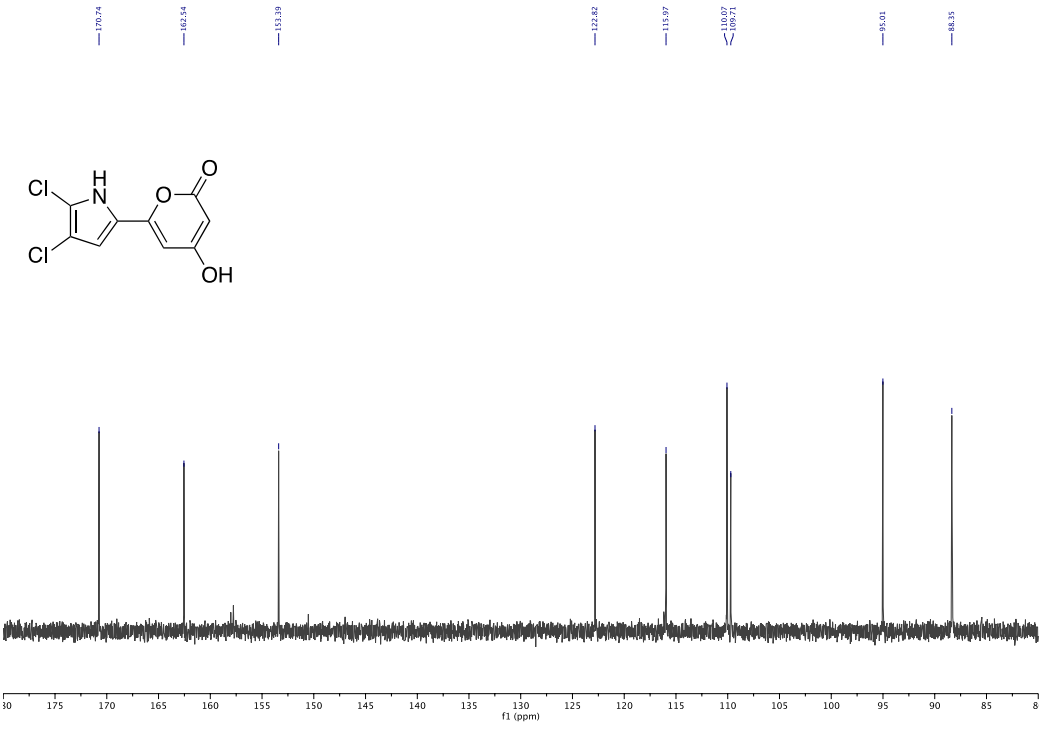


Fig. S8: ^{13}C NMR spectrum (126 MHz, DMSO-d_6) of compound **6**.

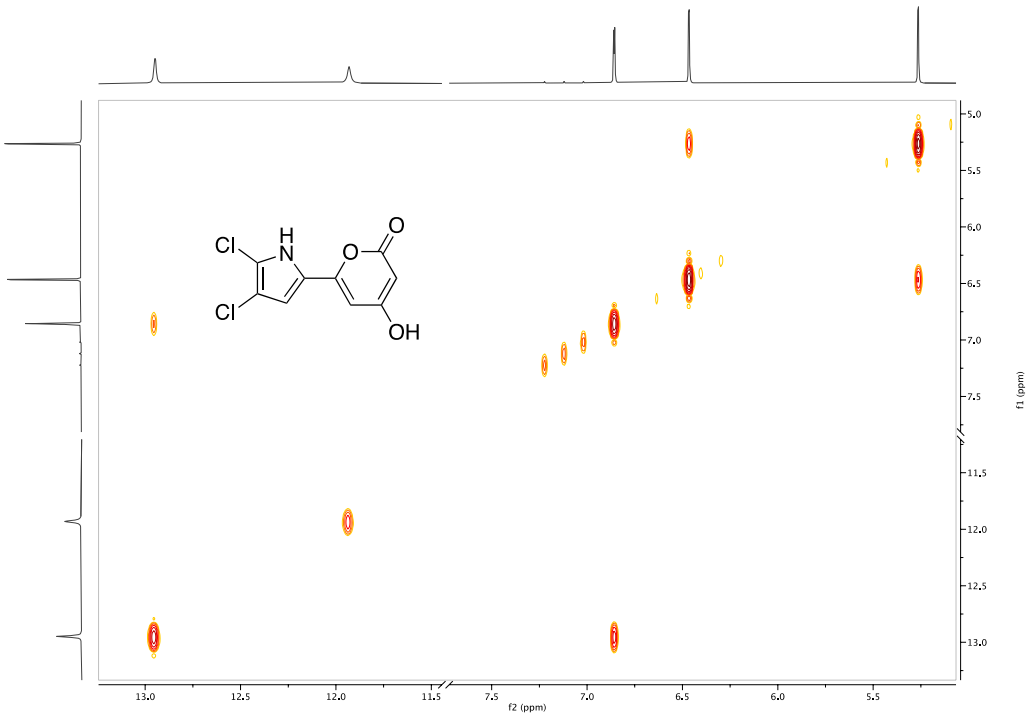


Fig. S9: ^{1}H -^{1}H COSY spectrum compound **6** in DMSO-d_6.

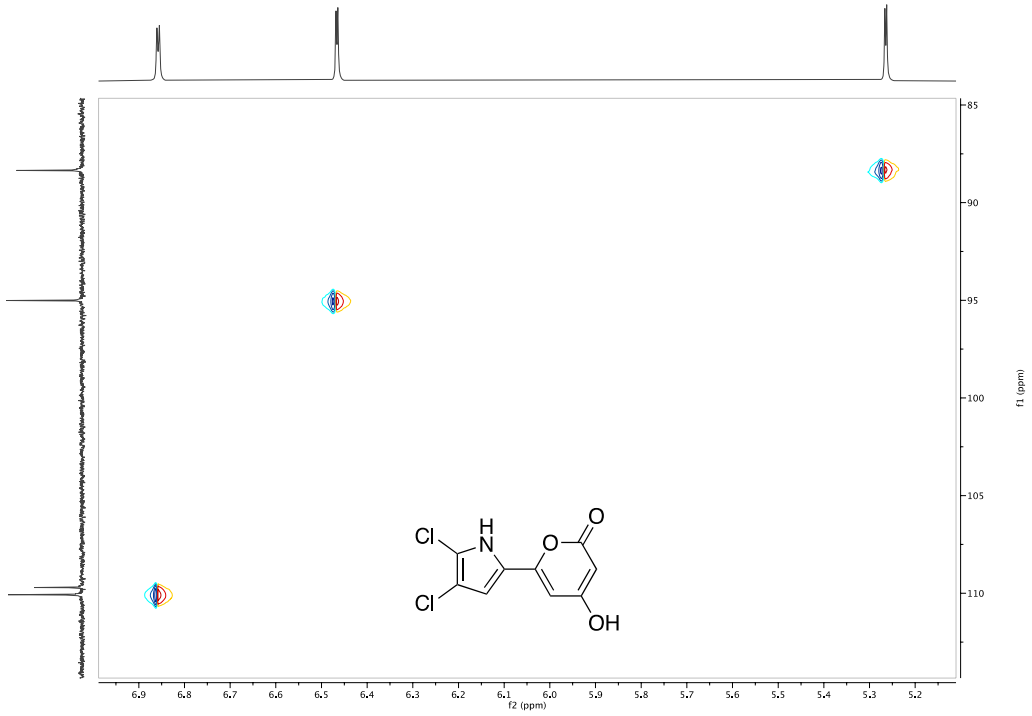


Fig. S10: ^{1}H -^{13}C HSQC spectrum compound **6** in DMSO-d_6.

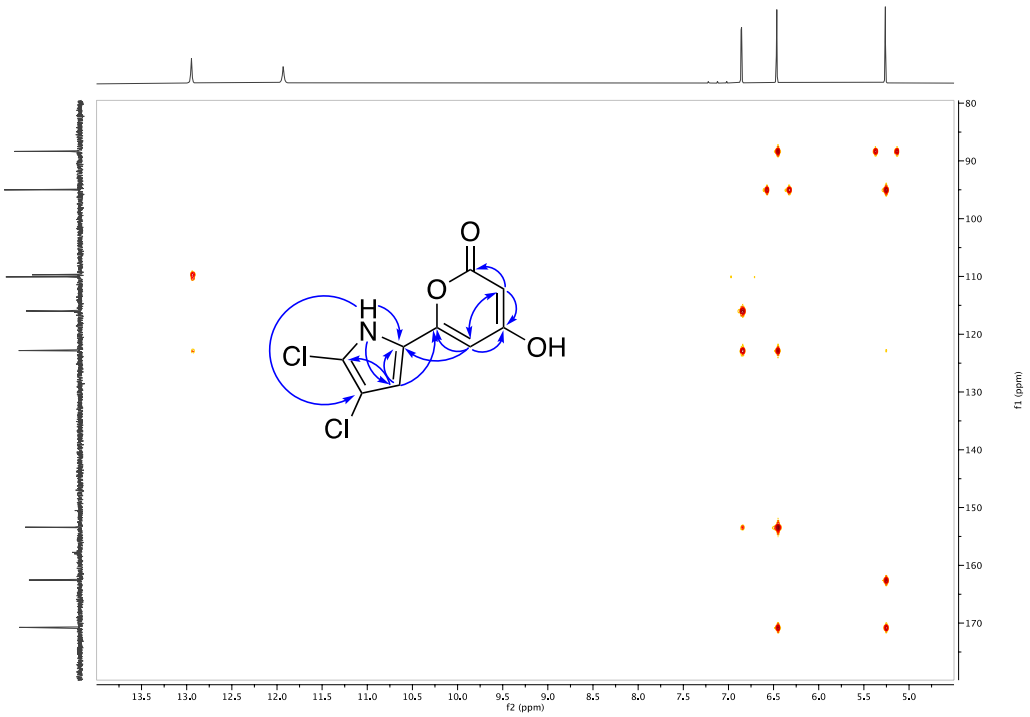


Fig. S11: ^{1}H -^{13}C HMBC spectrum compound **6** in DMSO-d_6.

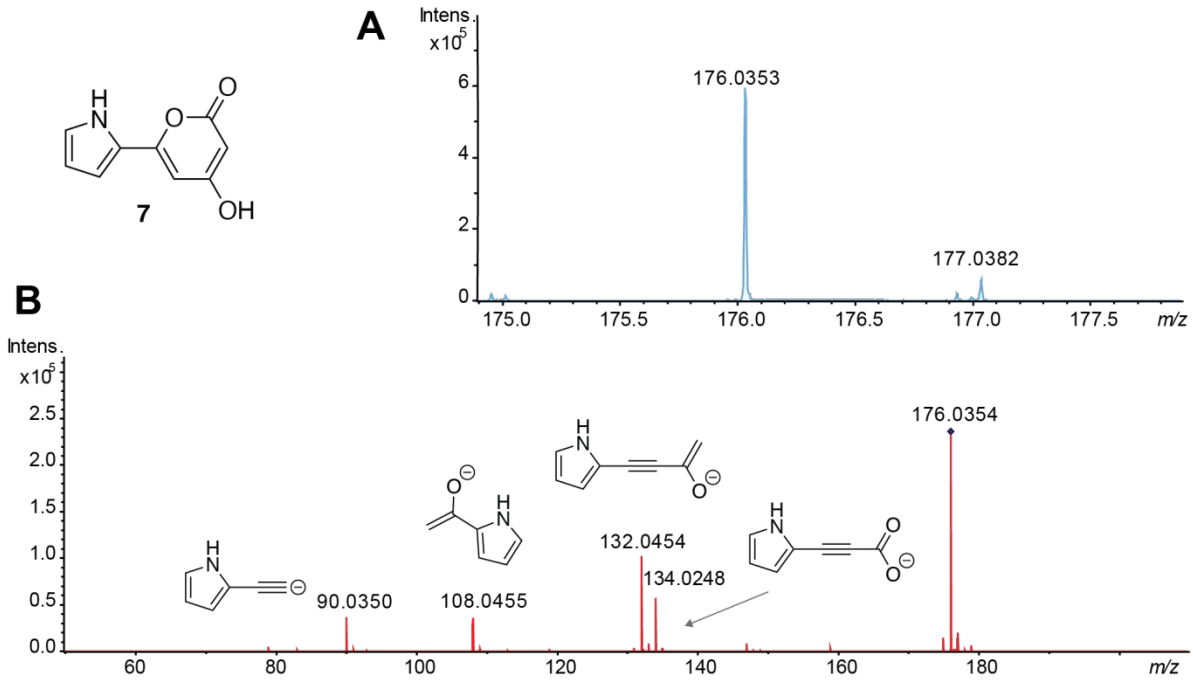


Fig. S12: MS^1 and MS^2 spectra of compound **7**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **7** (*m/z* calculated for $C_9H_6NO_3$ 176.0353, found 176.0353). (B) MS^2 spectra of compound **7** with rationalized structural annotations of fragment ions. Note the similarity in fragmentation pattern for **7** and **6** (Fig. S3).

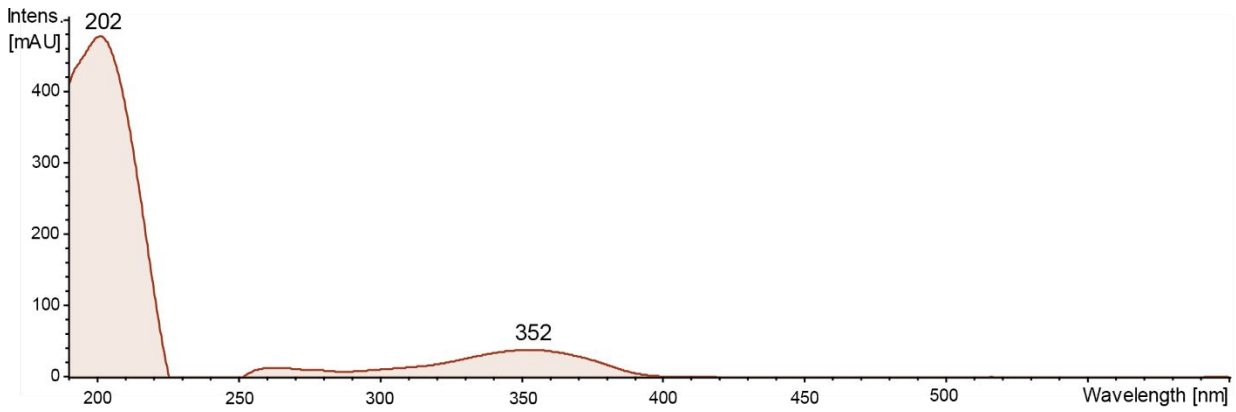


Fig. S13: UV absorbance profile of compound **7**. Note the similarity in UV-Vis absorbance spectra for **7** and **6** (Fig. S4).

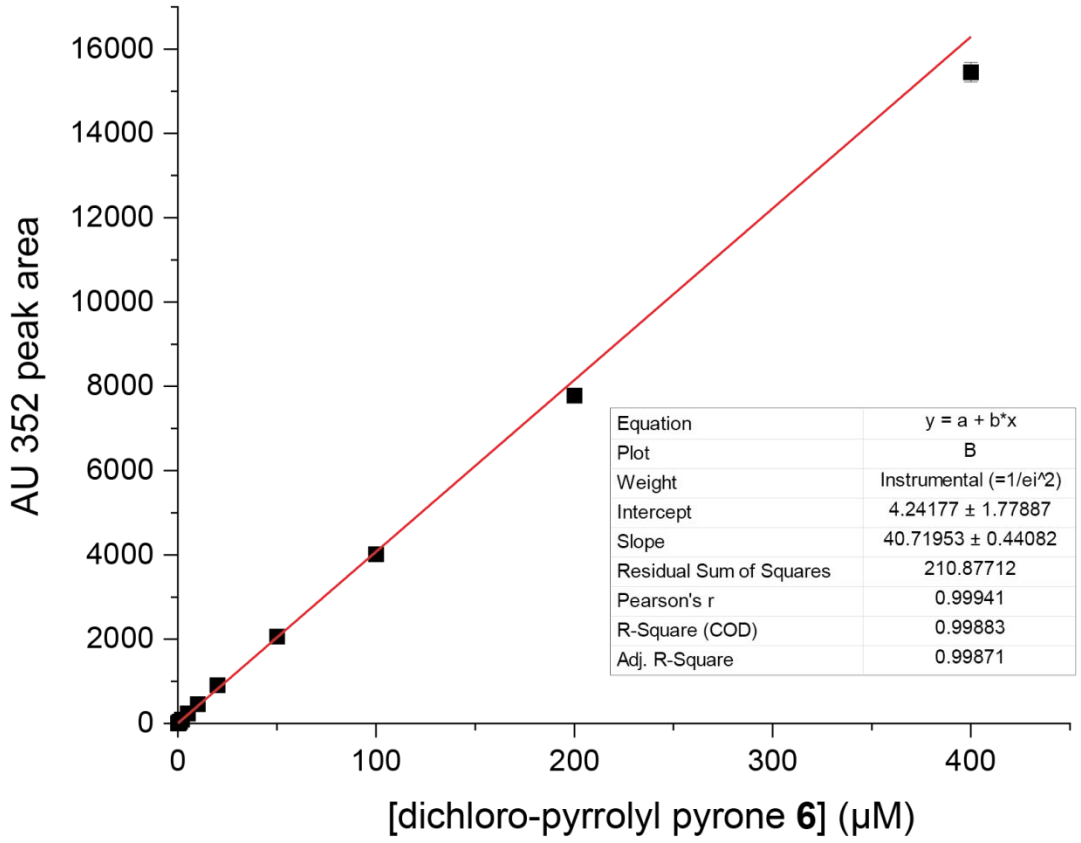


Fig. S14: Standard curve between area under the UV-Vis absorbance chromatogram recorded at 352 nm and concentration of compound **6**. This standard curve was used to quantify the concentration of **6** and **7** produced in assays with PKSs. Note that molecule **7** also possesses an absorbance maximum at 352 nm (Fig. S13).

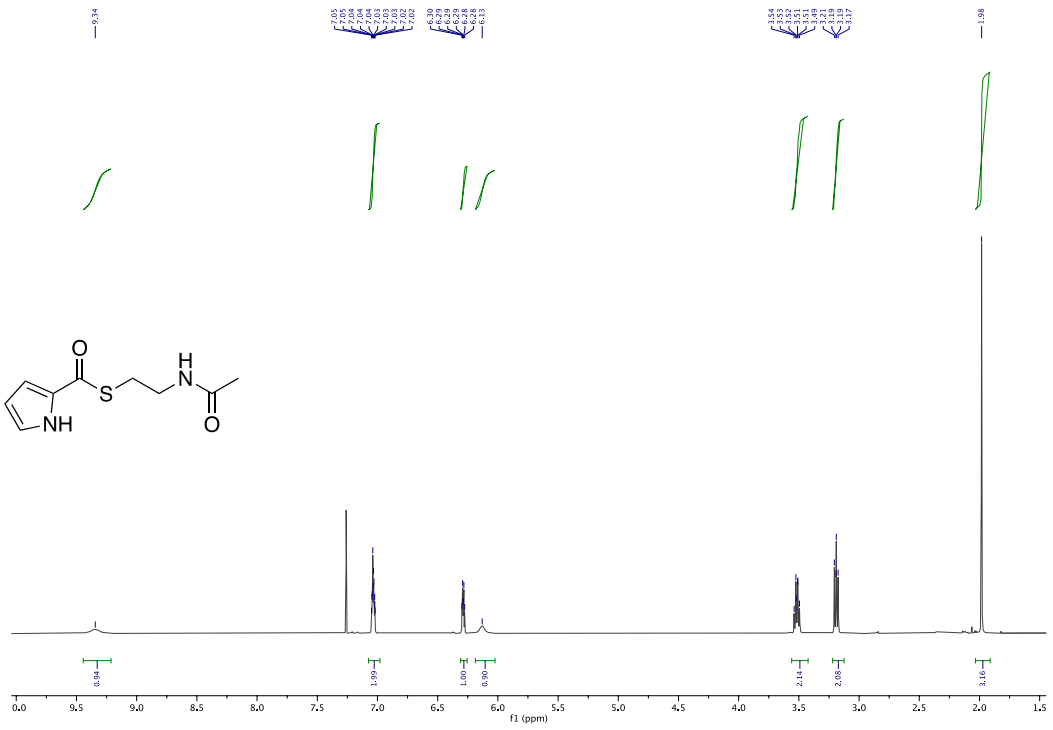


Fig. S15: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8b**.

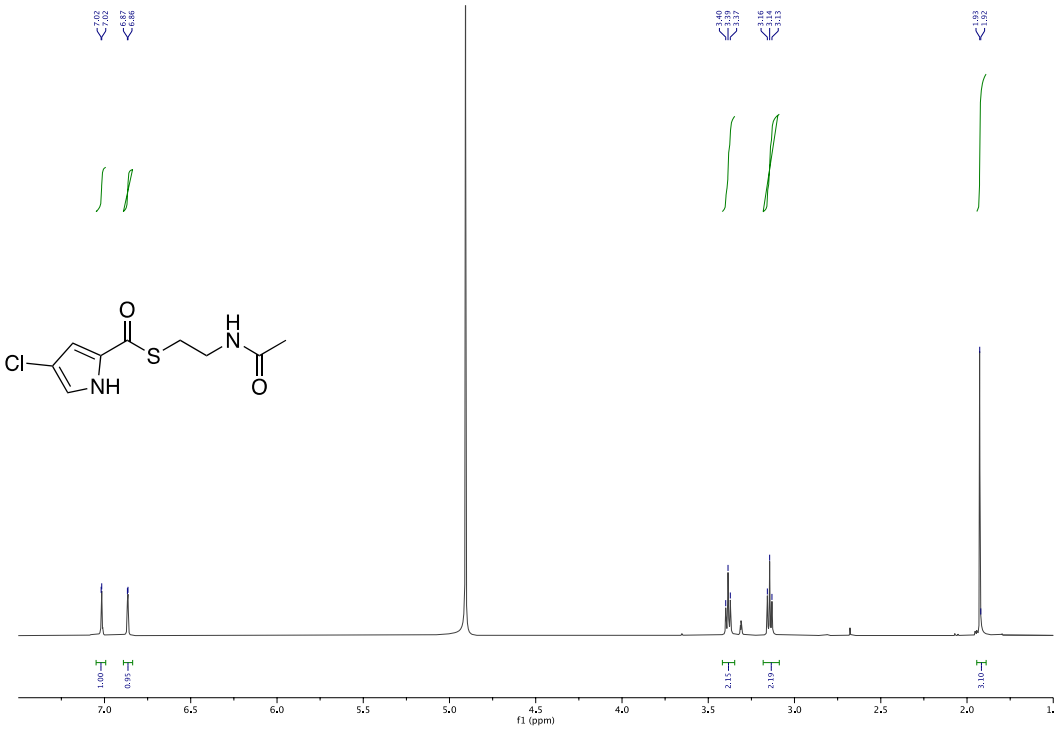


Fig. S16: [1]H NMR spectrum (500 MHz, MeOD) of compound **8c**.

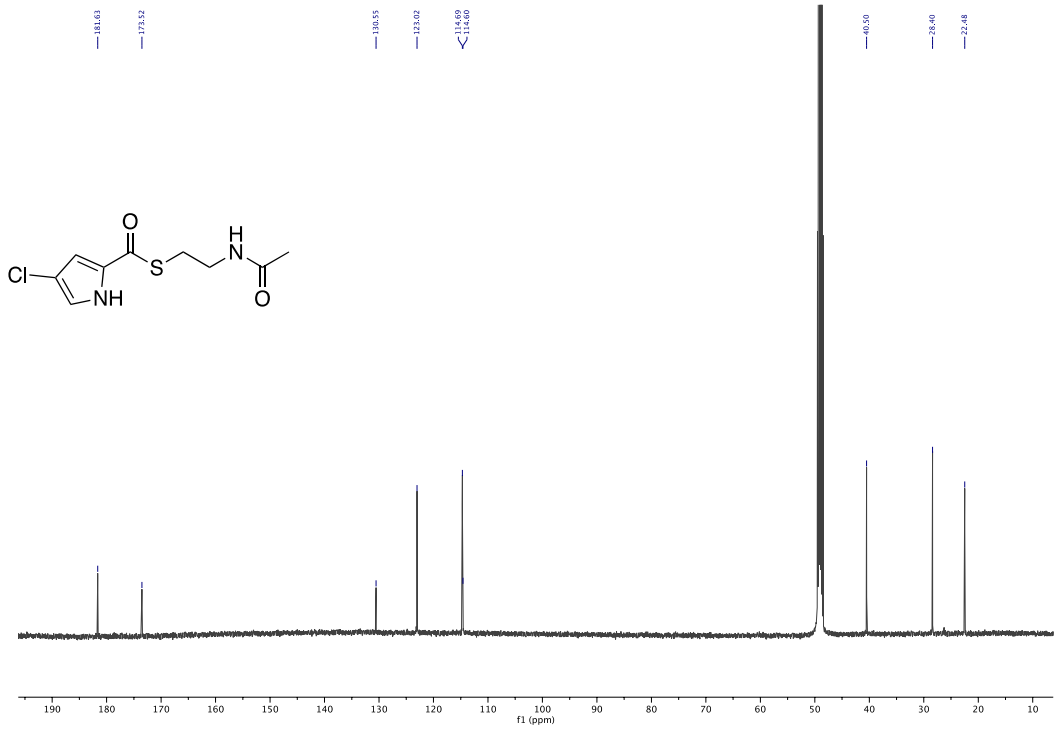


Fig. S17: [13]C NMR spectrum (126 MHz, MeOD) of compound **8c**.

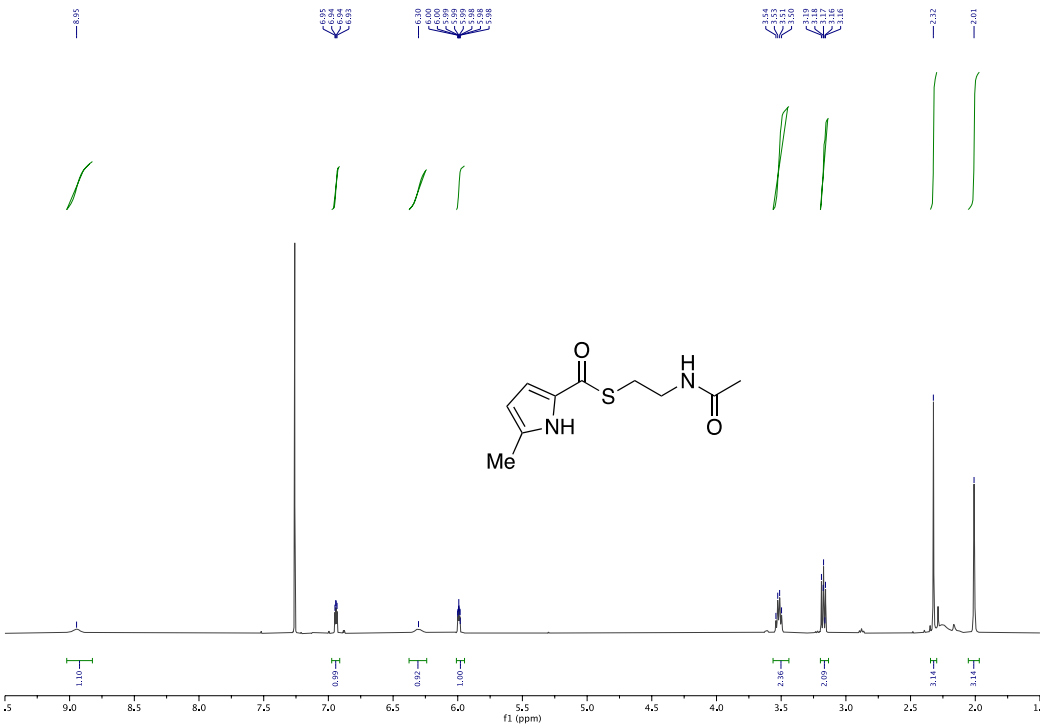


Fig. S18: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8d**.

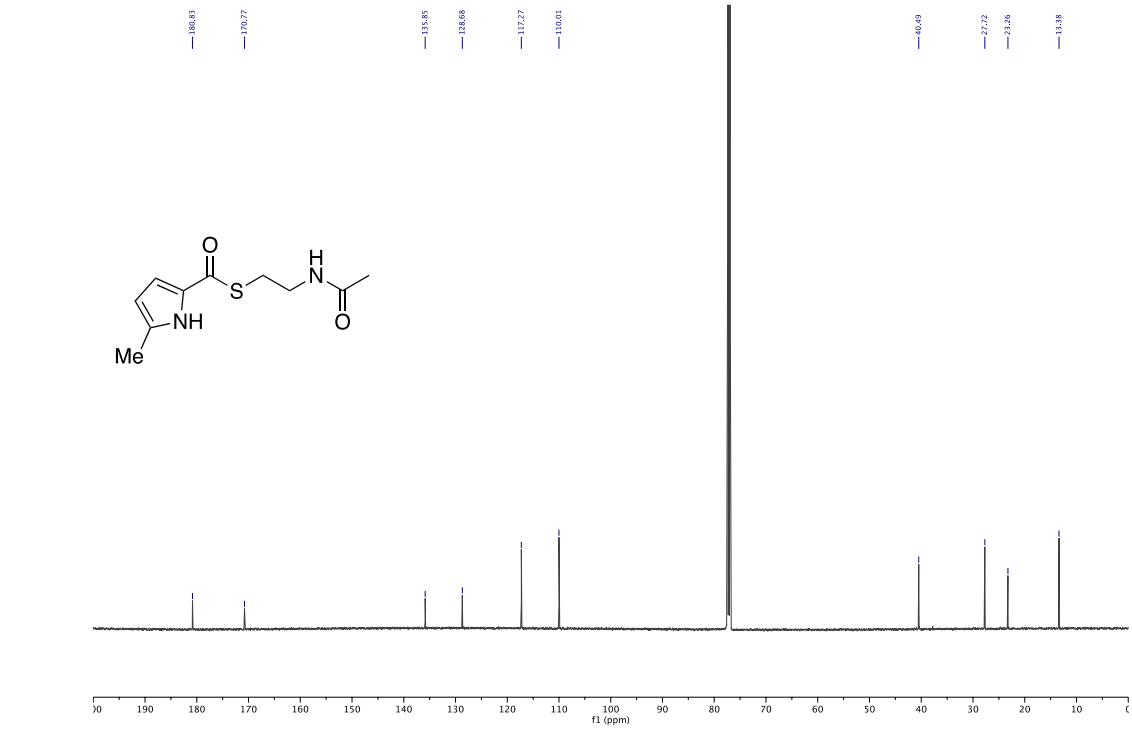


Fig. S19: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8d**.

Fig. S20: ^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **8e**.

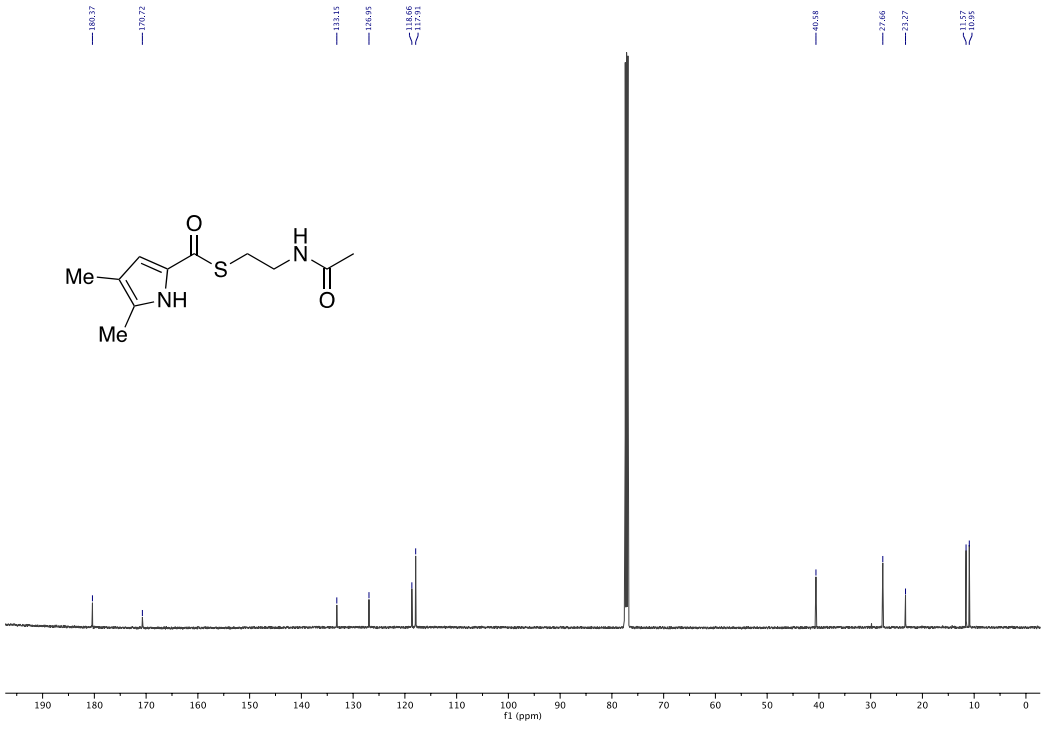


Fig. S21: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8e**.

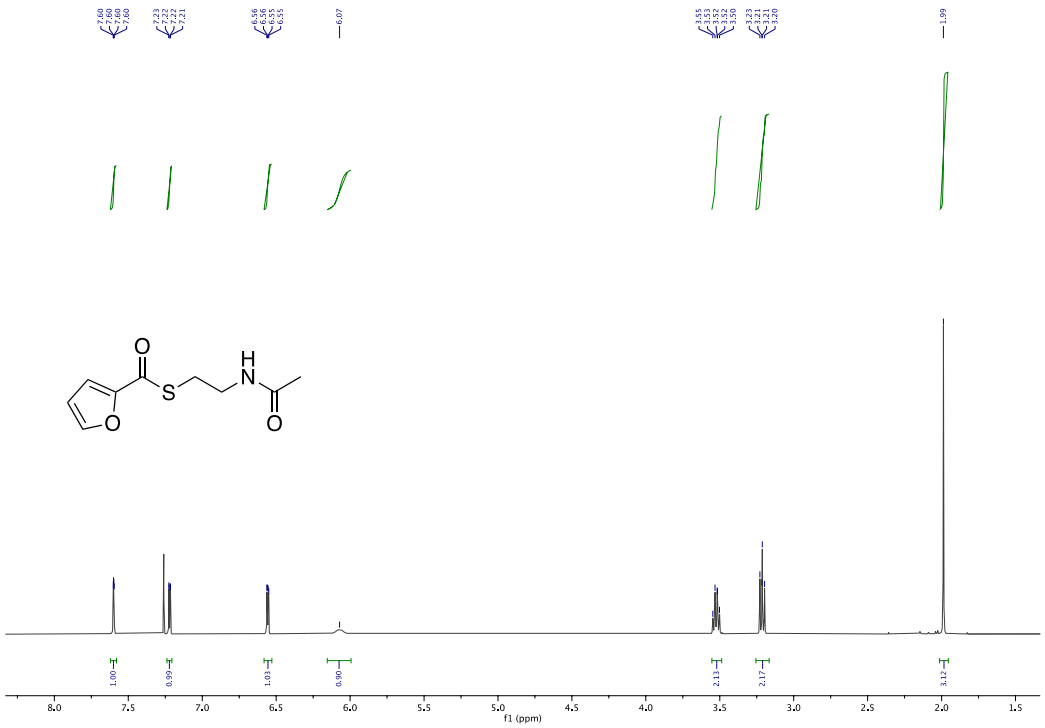


Fig. S22: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8f**.

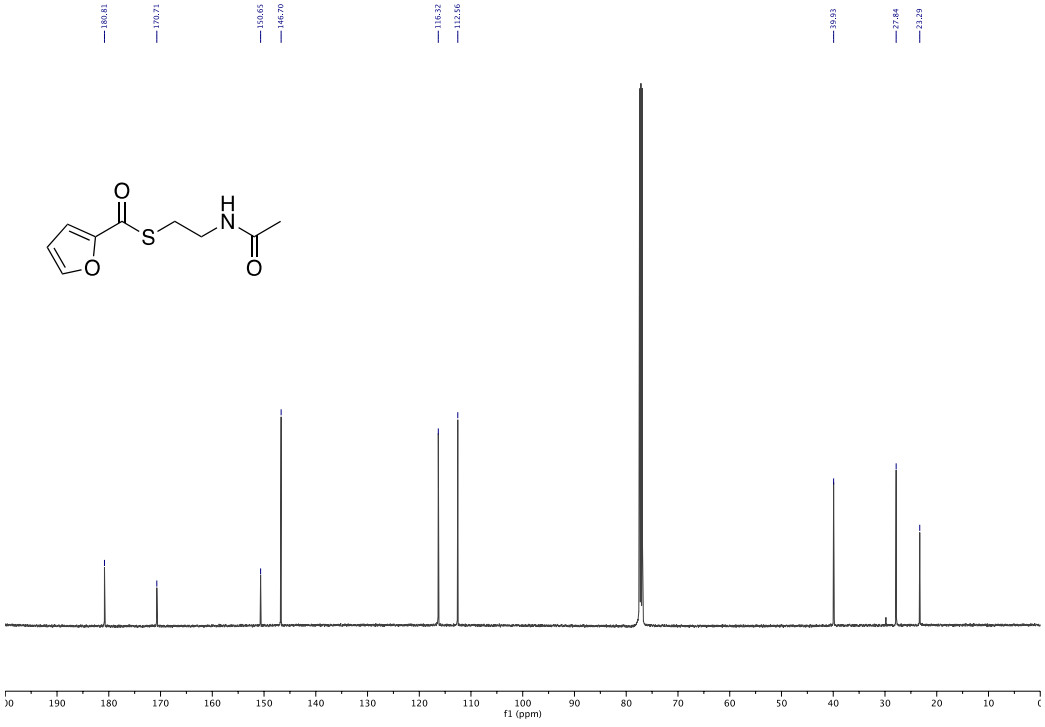


Fig. S23: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8f**.

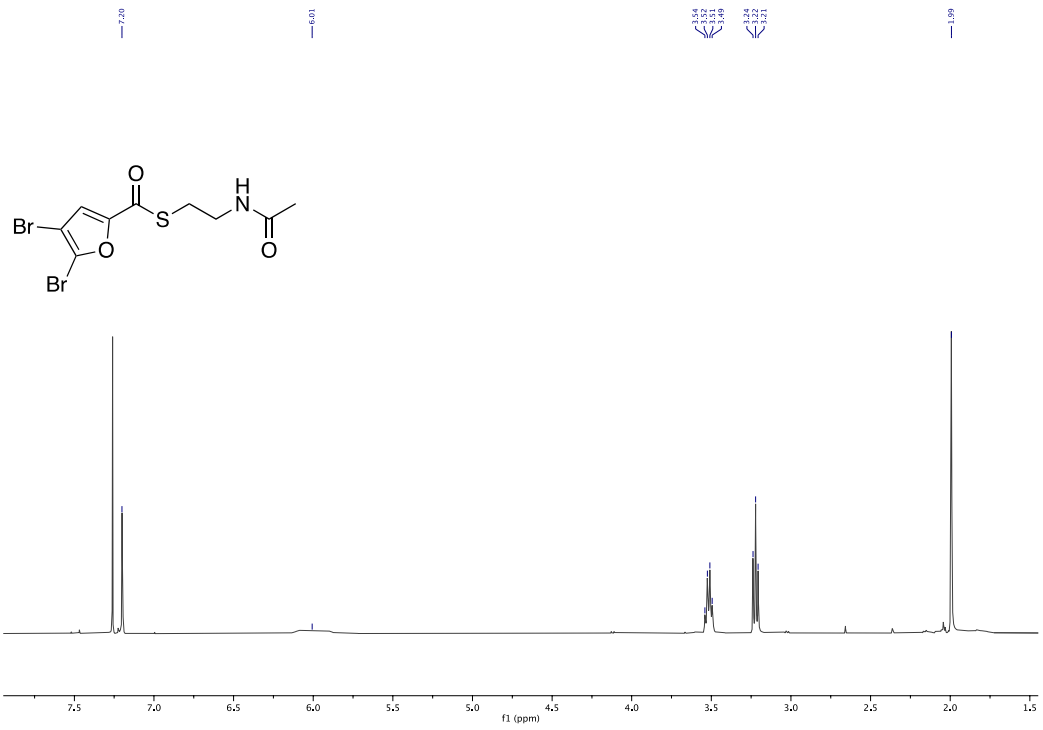


Fig. S24: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8g**.

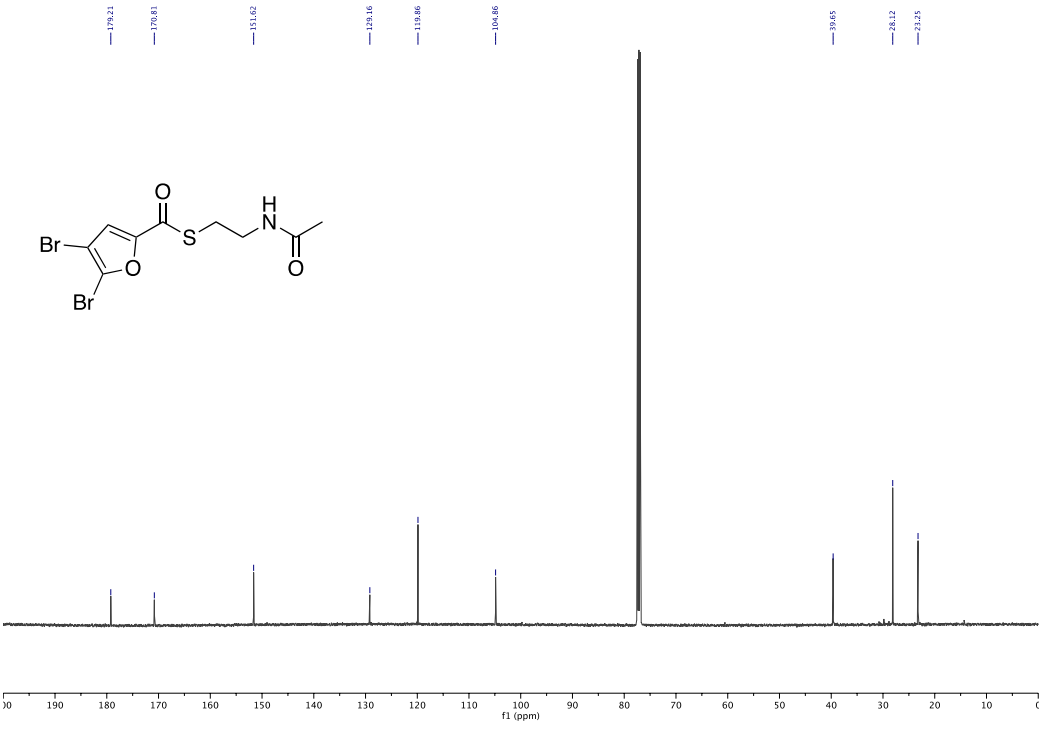


Fig. S25: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8g**.

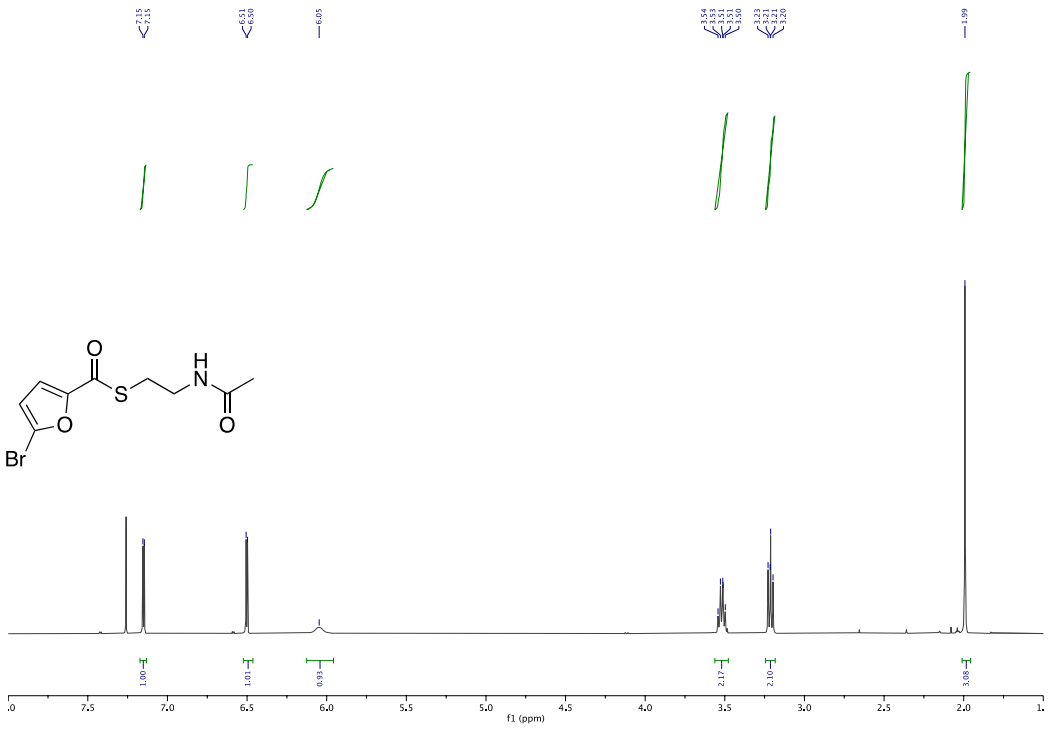


Fig. S26: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8h**.

Fig. S27: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8h**.

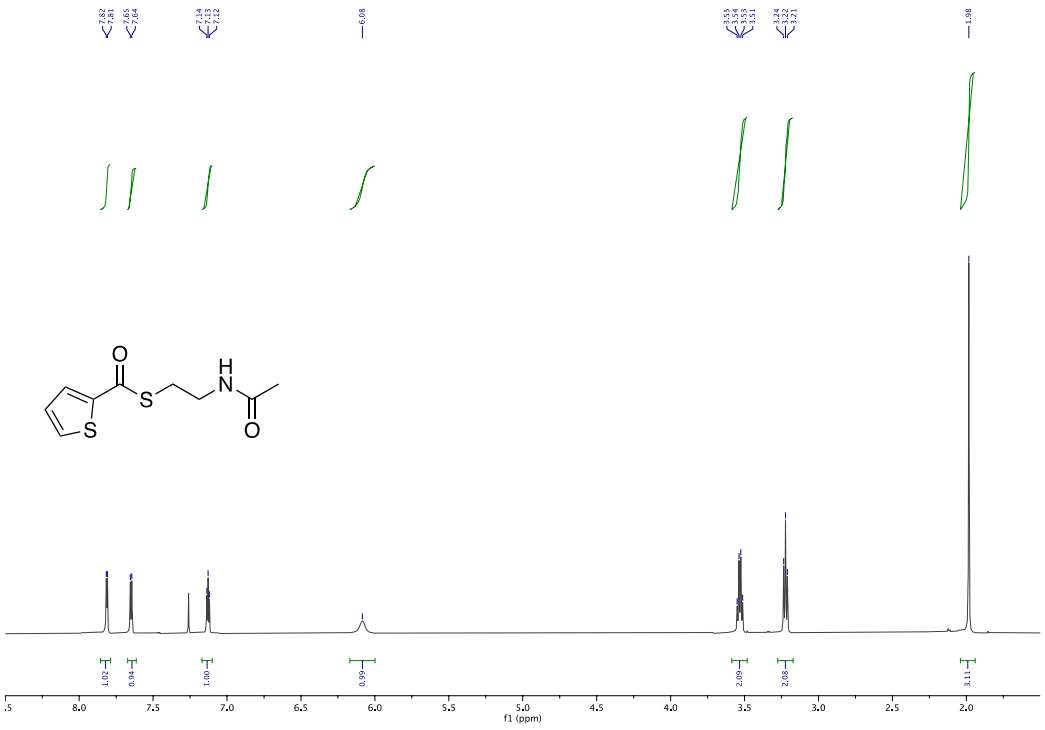


Fig. S28: ^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **8i**.

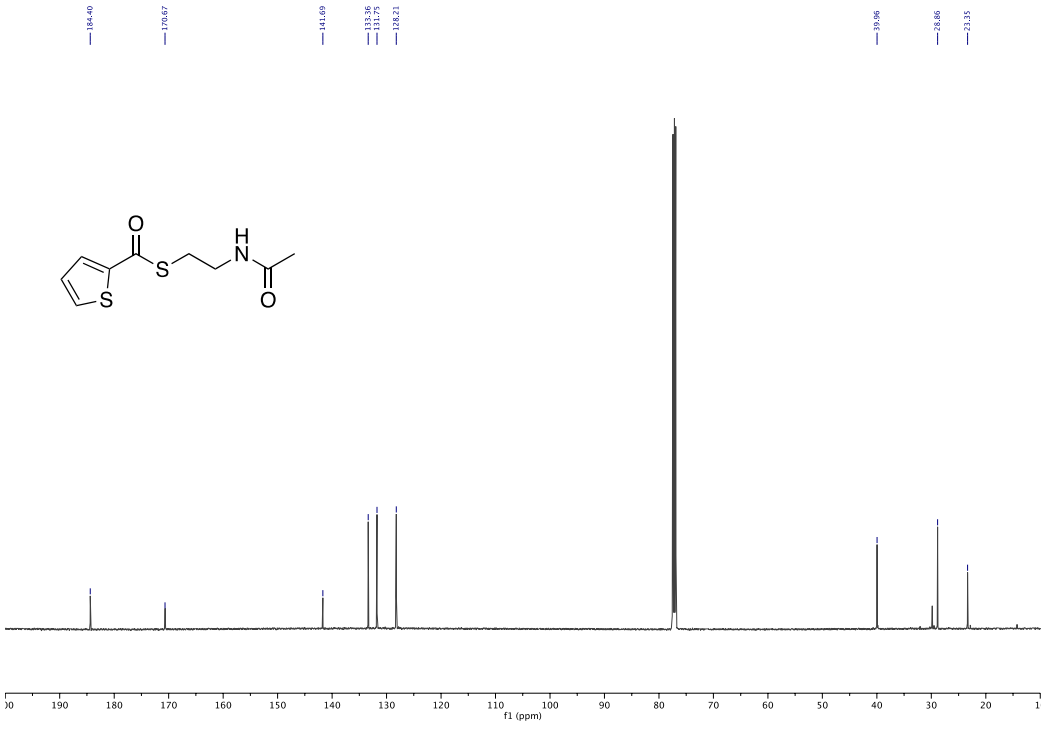


Fig. S29: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8i**.

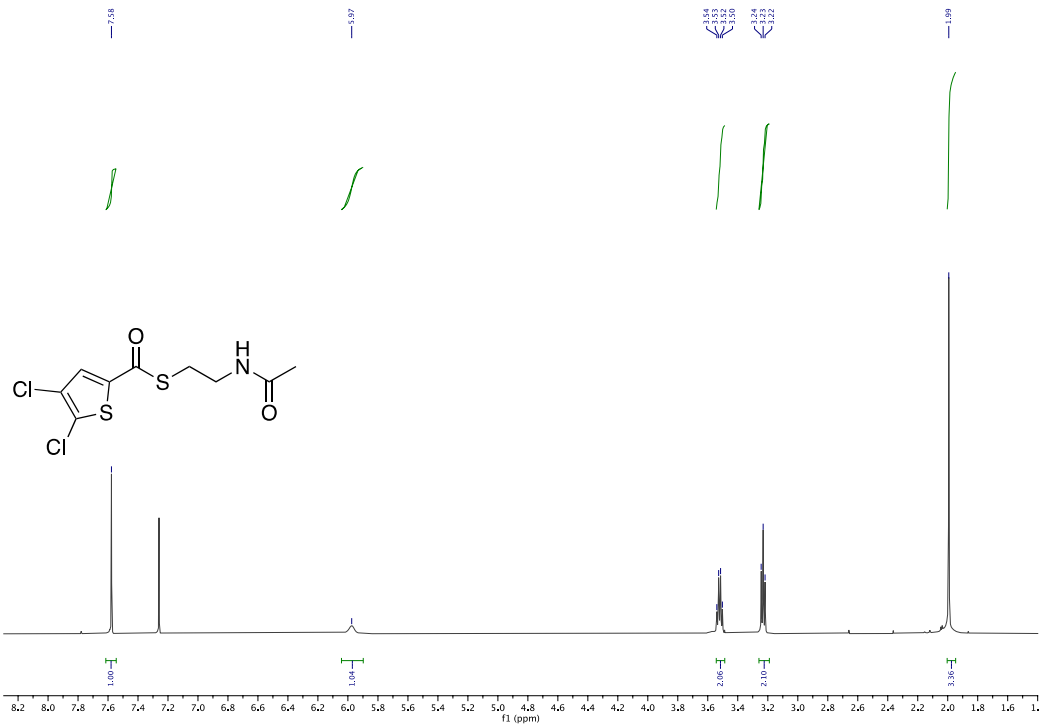


Fig. S30: ^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **8j**.

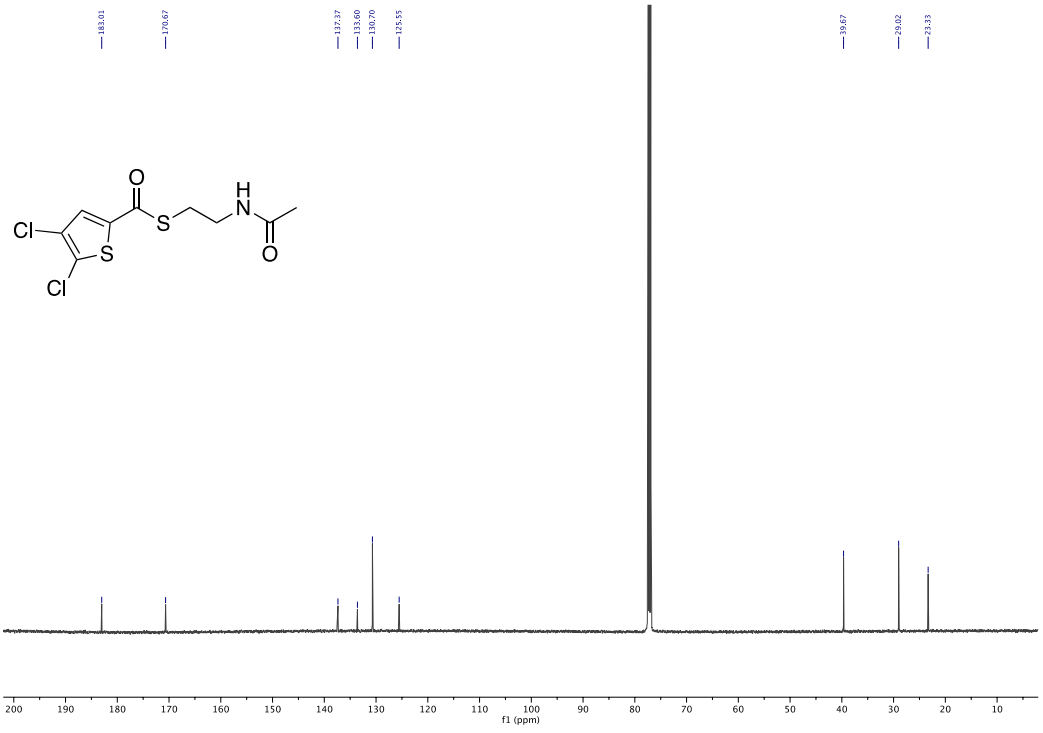


Fig. S31: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8j**.

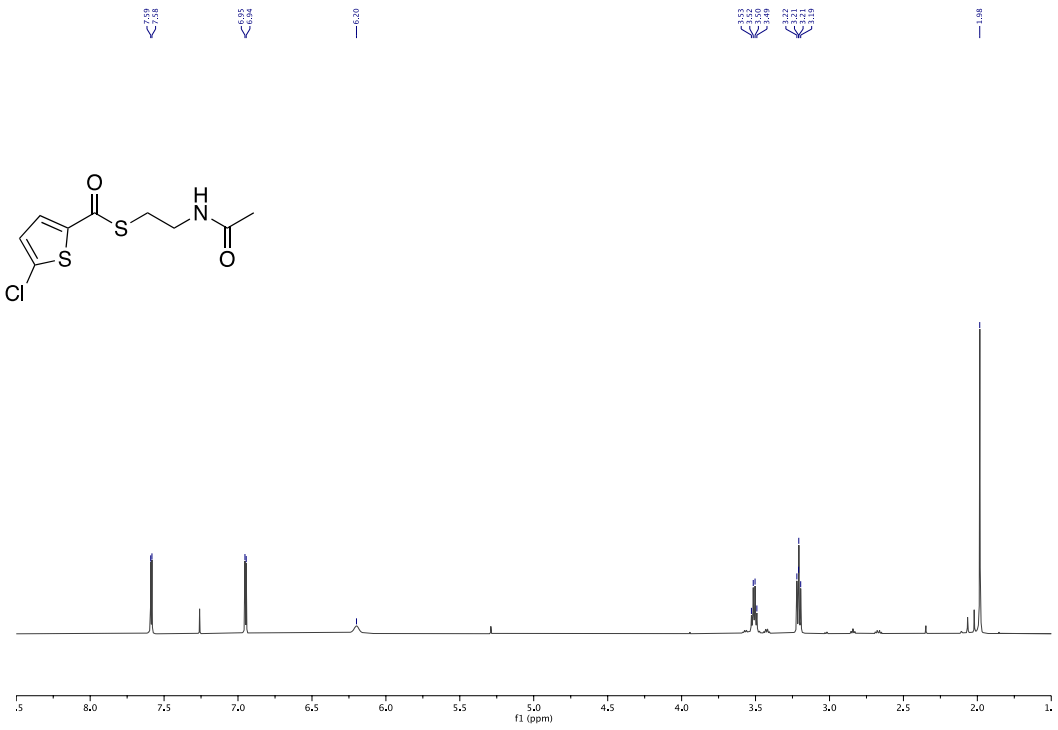


Fig. S32: ^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **8k**.

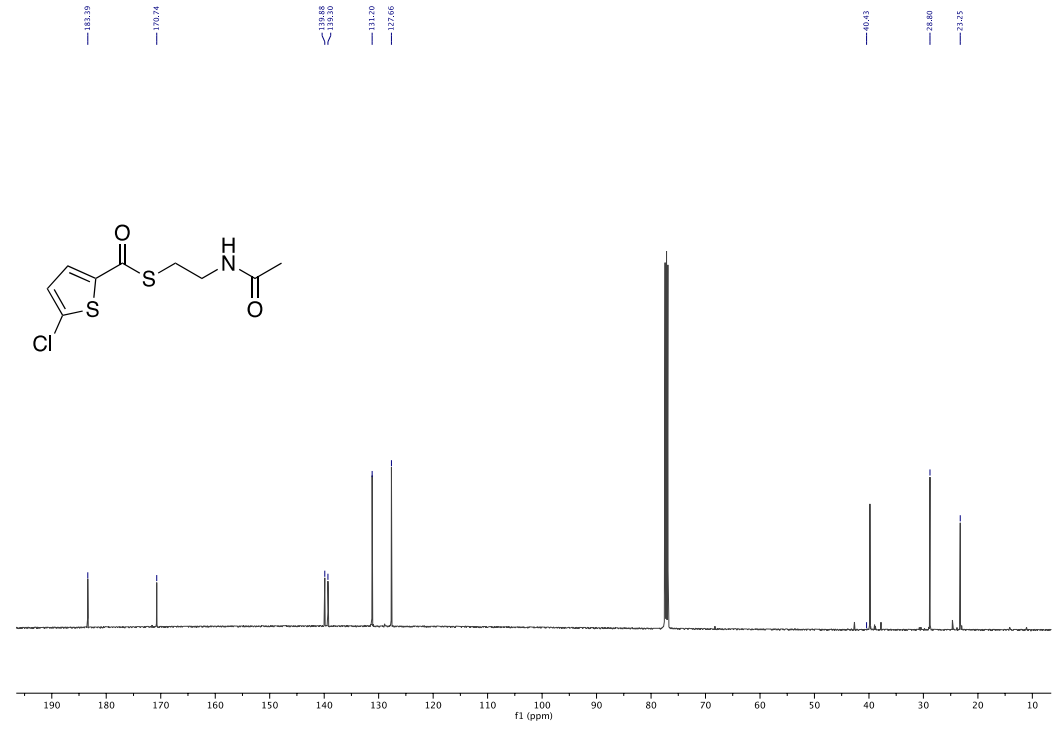


Fig. S33: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8k**.

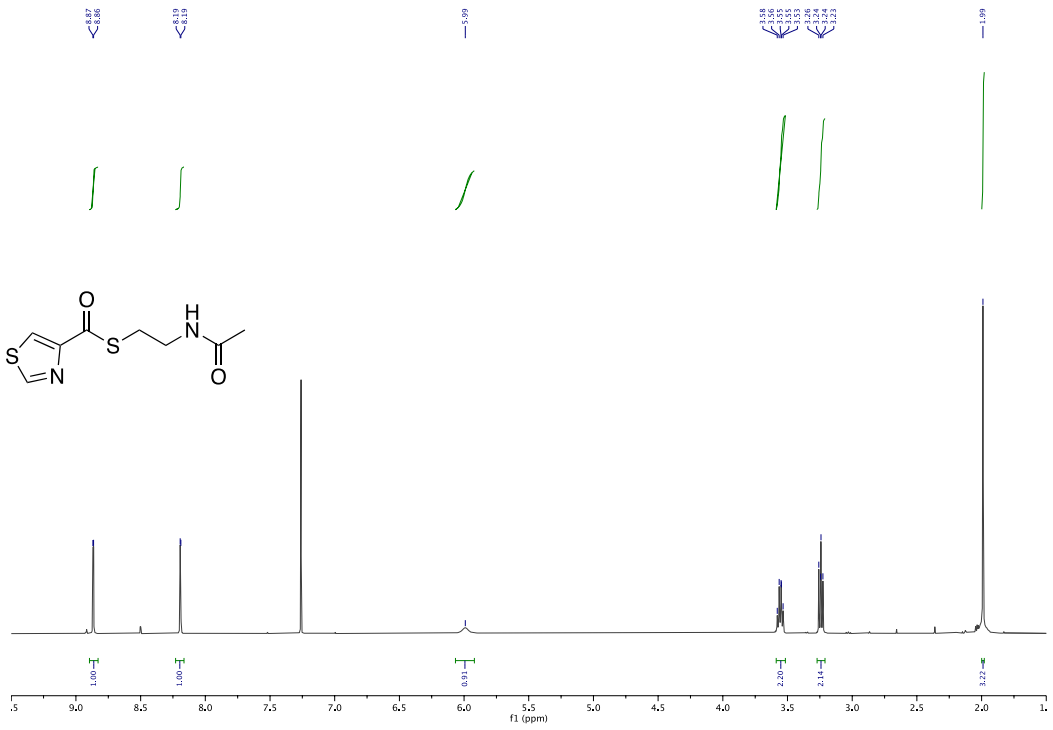


Fig. S34: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8l**.

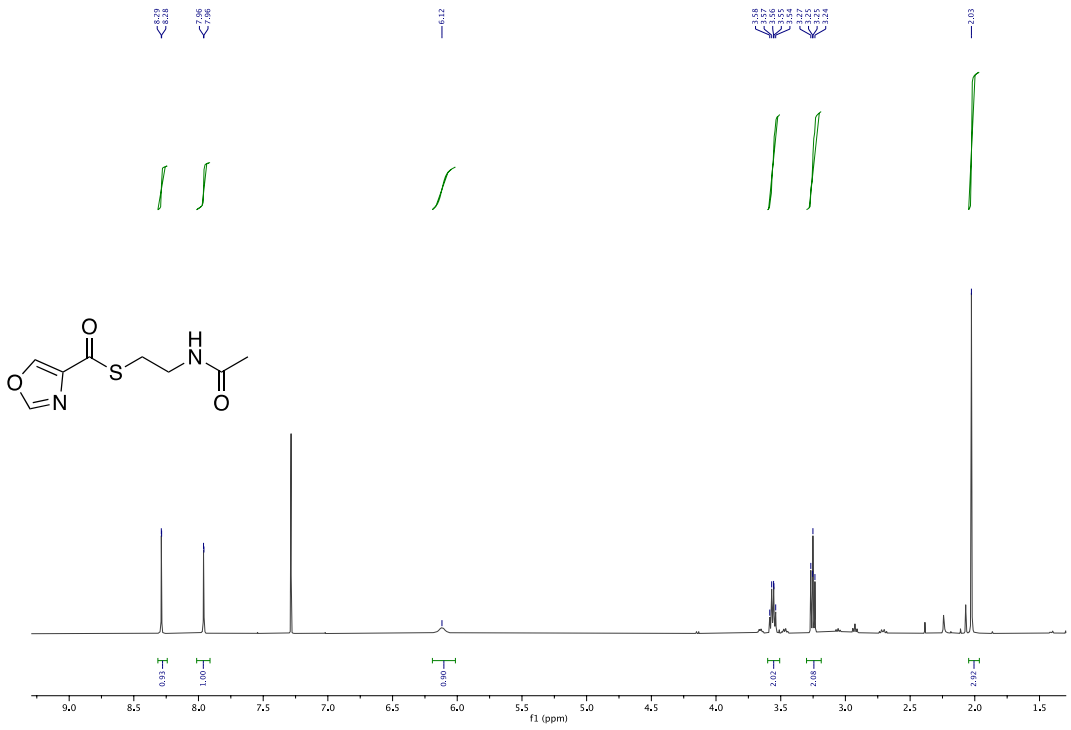


Fig. S35: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8m**.

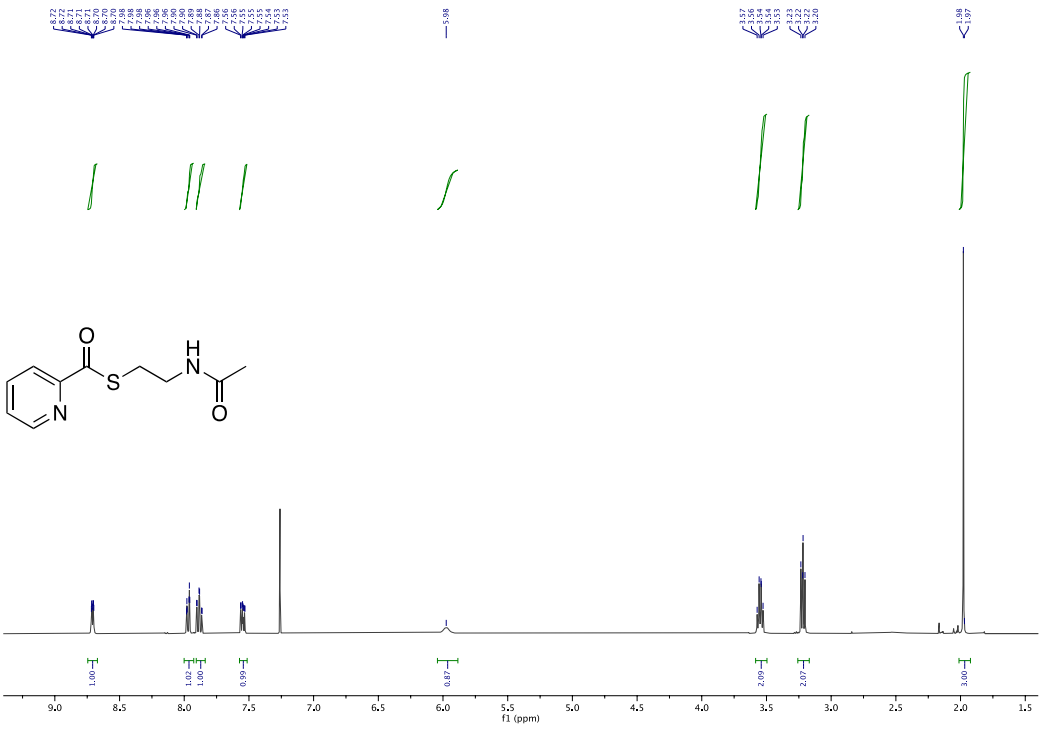


Fig. S36: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8n**.

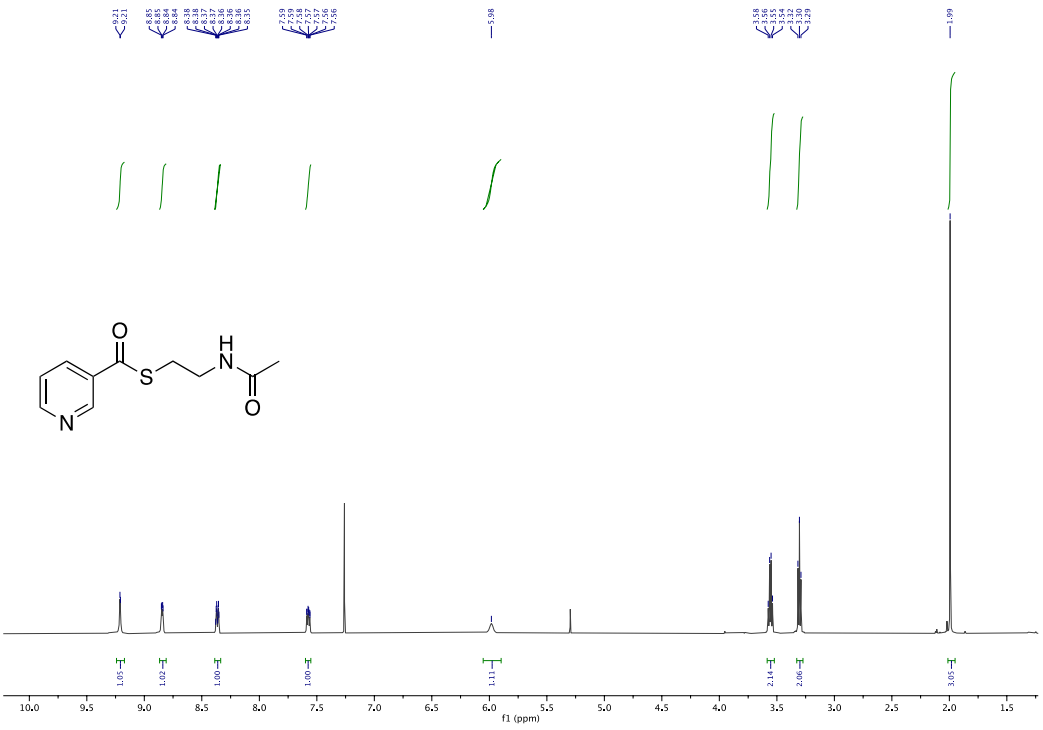


Fig. S37: ^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **8o**.

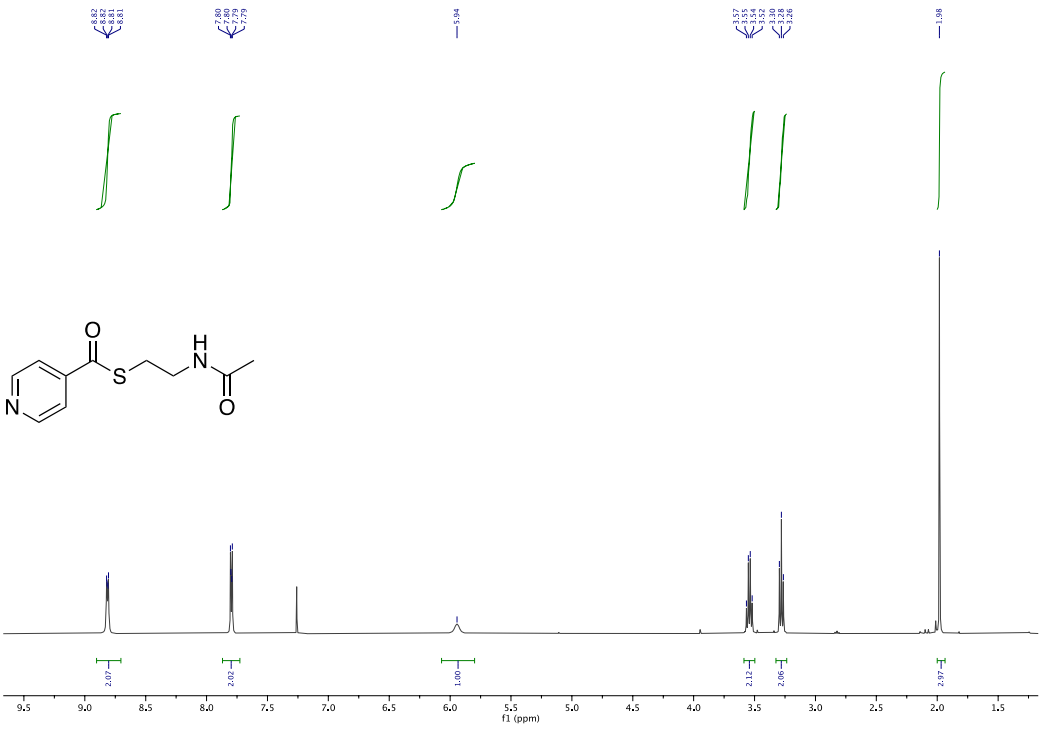


Fig. S38: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8p**.

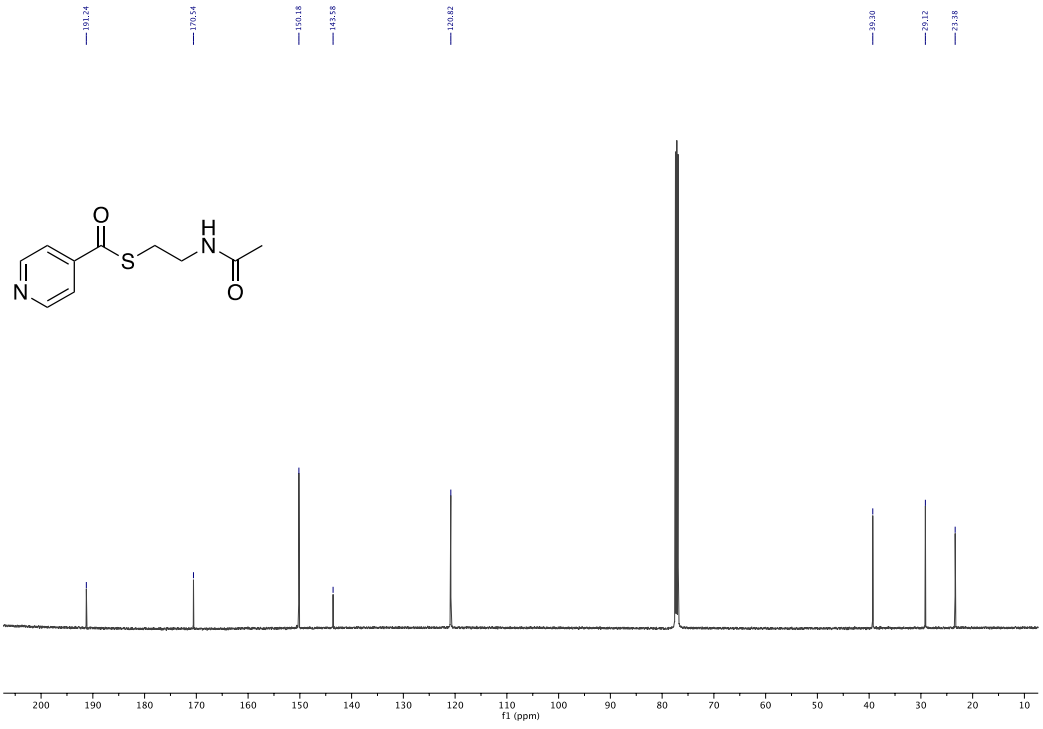


Fig. S39: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8p**.

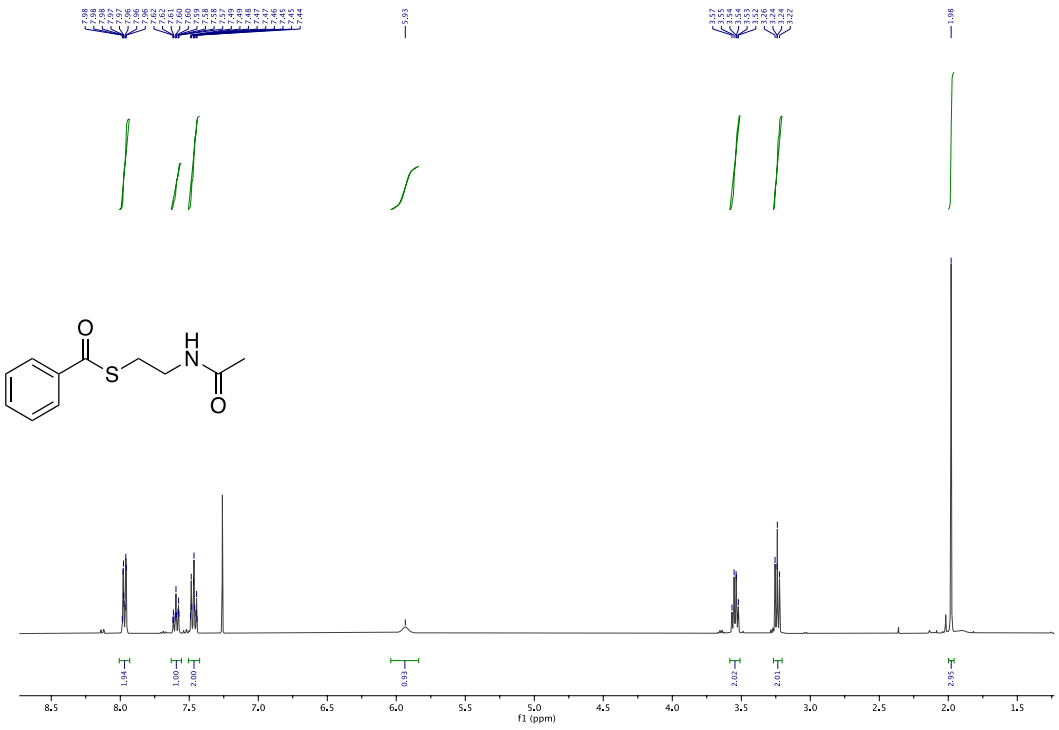


Fig. S40: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8q**.

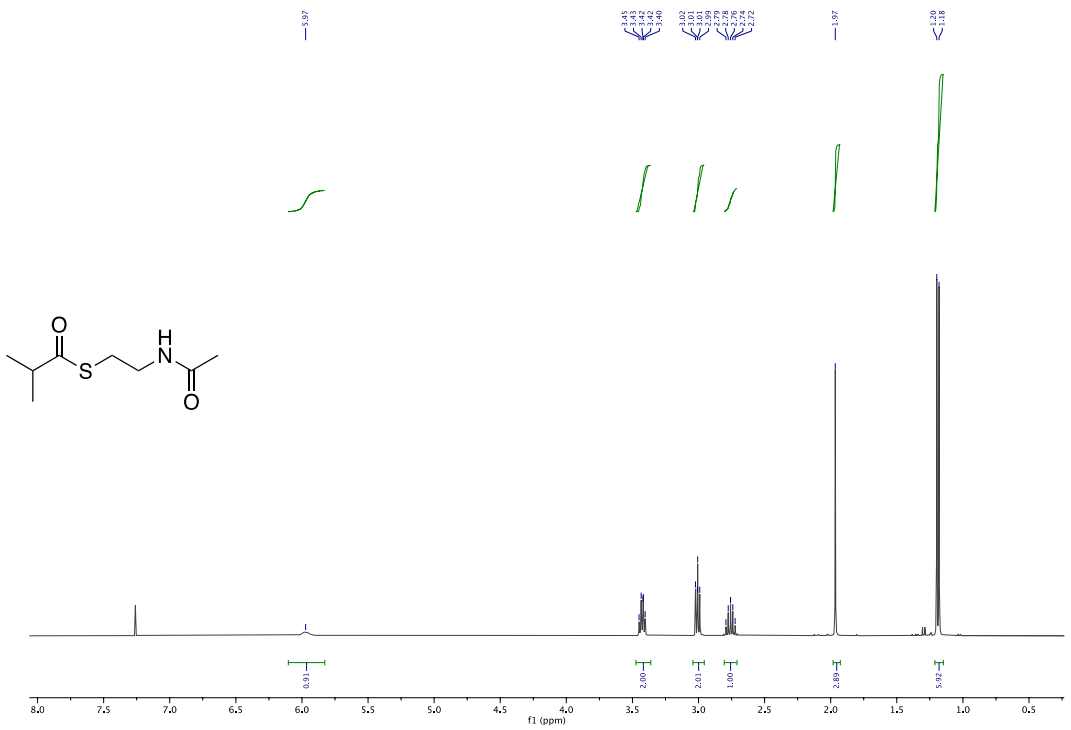


Fig. S41: ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of compound **8r**.

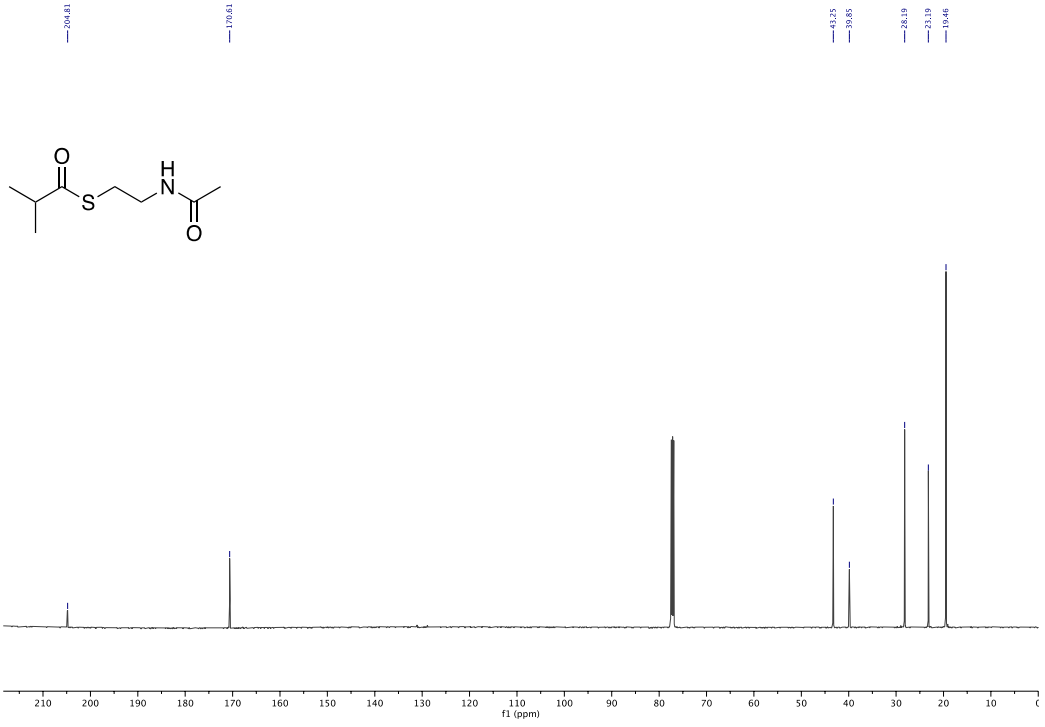


Fig. S42: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8r**.

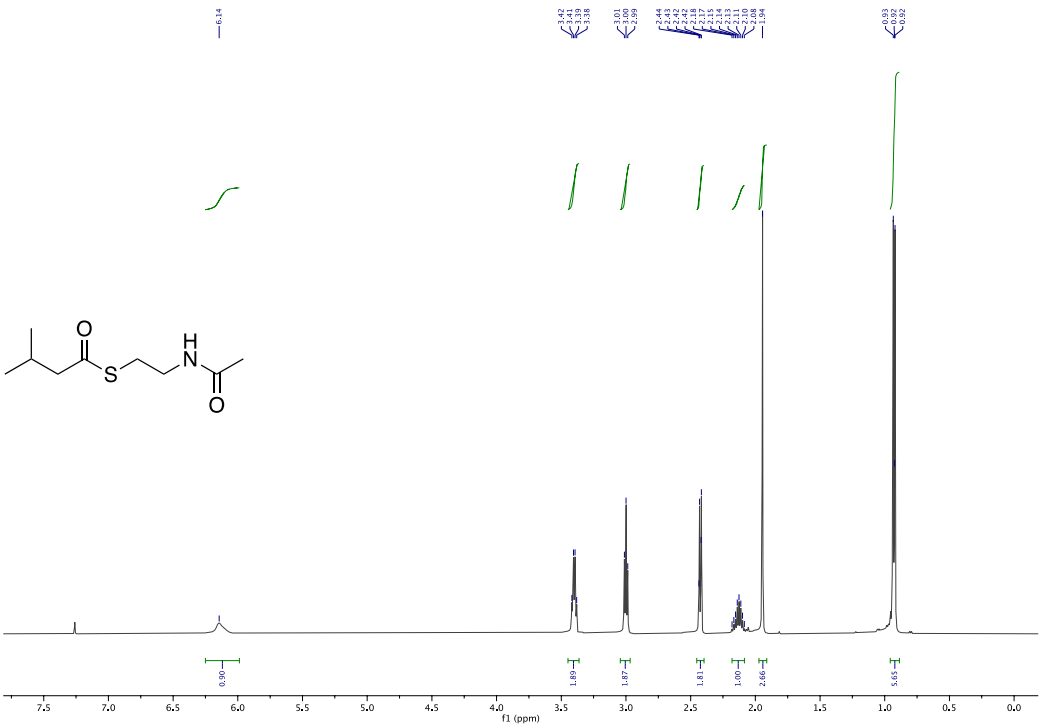


Fig. S43: ^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **8s**.

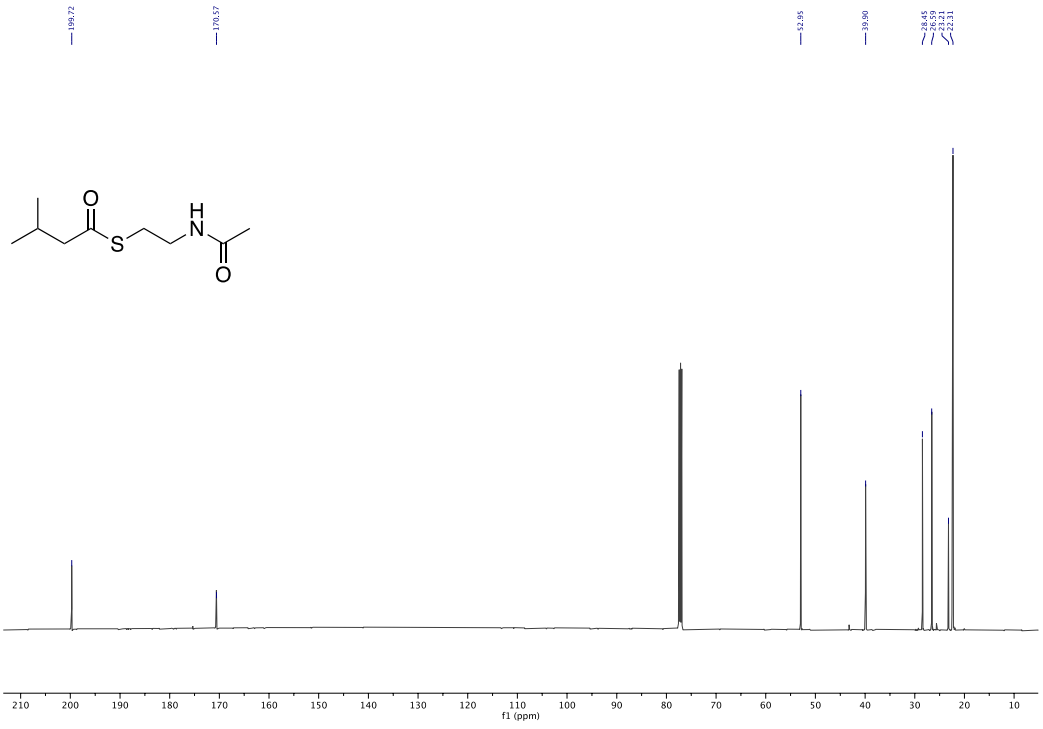


Fig. S44: ^{13}C NMR spectrum (126 MHz, $CDCl_3$) of compound **8s**.

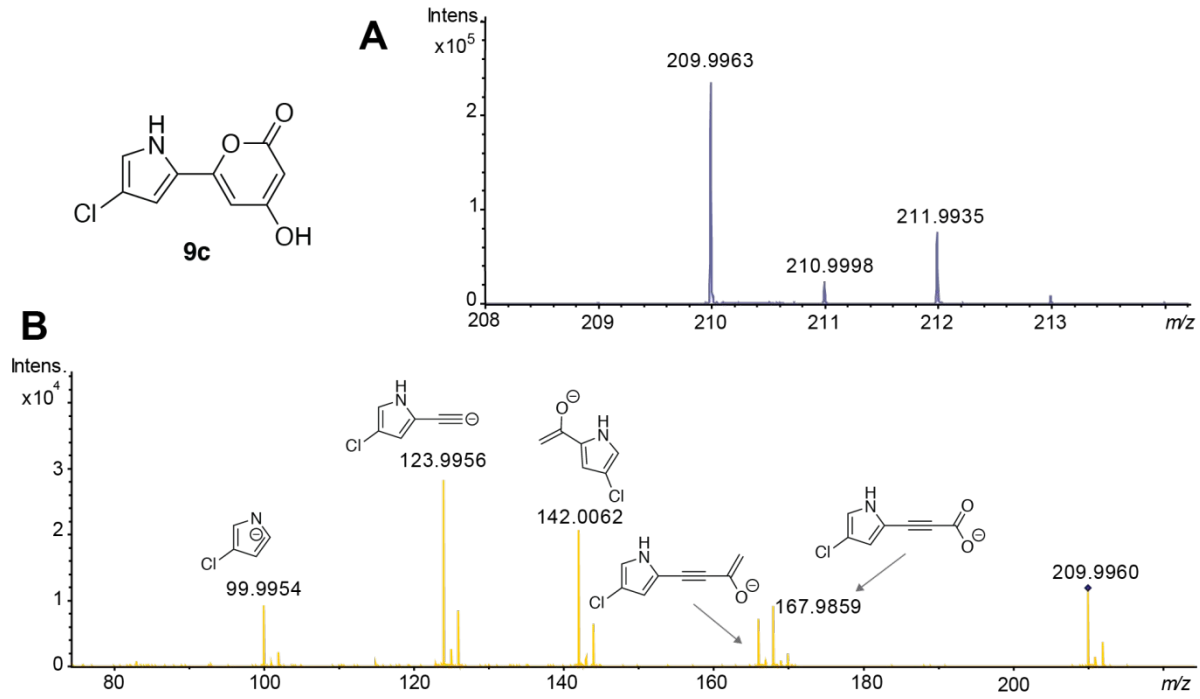


Fig. S45: MS^1 and MS^2 spectra of compound **9c**, product derived from substrate **8c**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9c** (*m/z* calculated for $C_9H_5ClNO_3$ 209.9963, found 209.9963). (B) MS^2 spectra of compound **9c** with rationalized structural annotations of fragment ions.

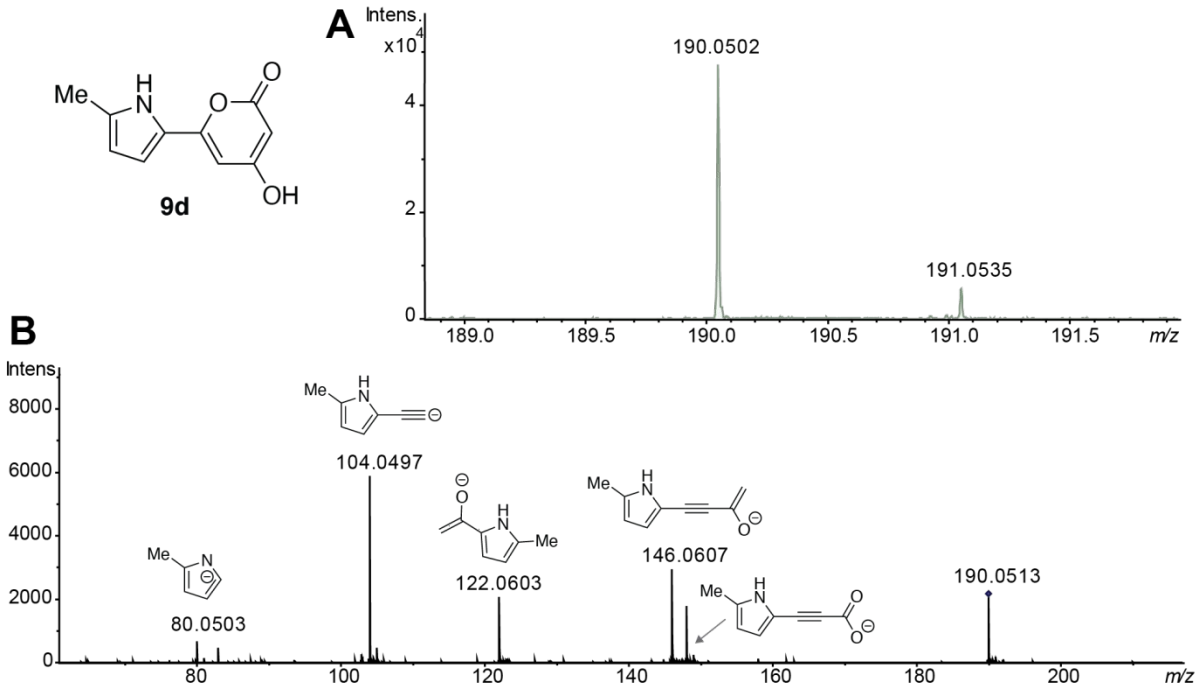


Fig. S46: MS^1 and MS^2 spectra of compound **9d**, product derived from substrate **8d**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9d** (*m/z* calculated for $C_{10}H_8NO_3$ 190.0510, found 190.0502). (B) MS^2 spectra of compound **9d** with rationalized structural annotations of fragment ions.

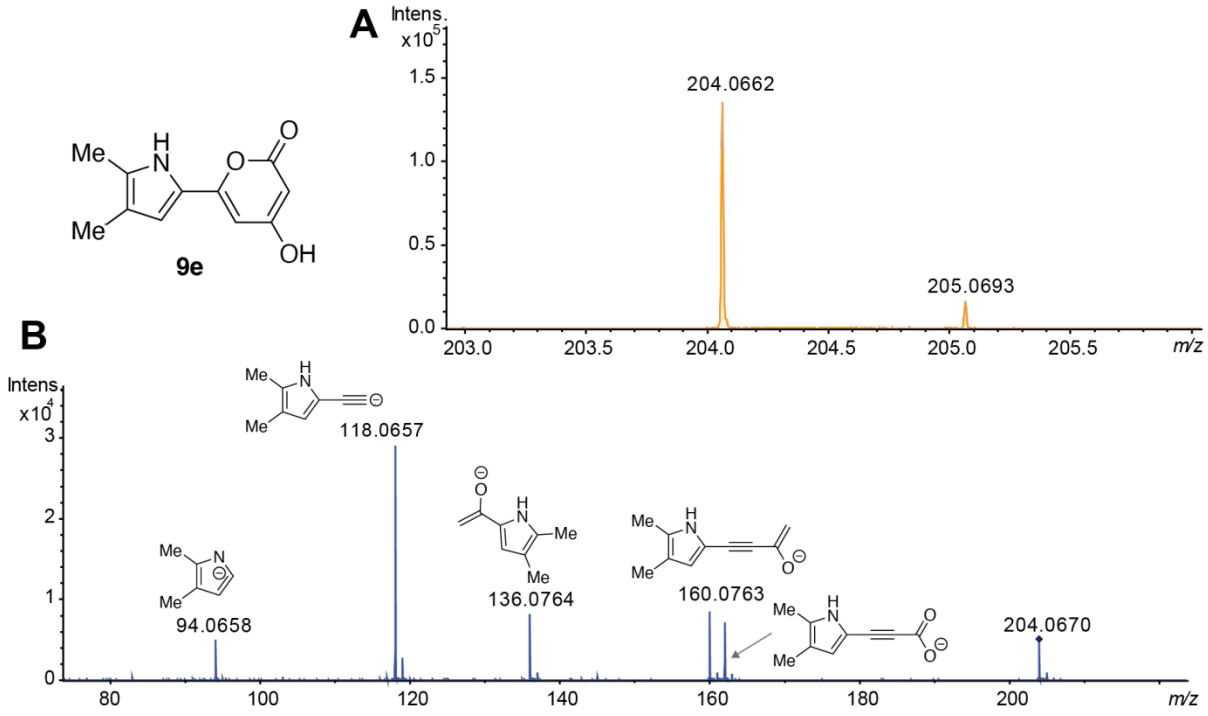


Fig. S47: MS^1 and MS^2 spectra of compound **9e**, product derived from substrate **8e**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9e** (*m/z* calculated for $C_{11}H_{10}NO_3$ 204.0666, found 204.0662). (B) MS^2 spectra of compound **9e** with rationalized structural annotations of fragment ions.

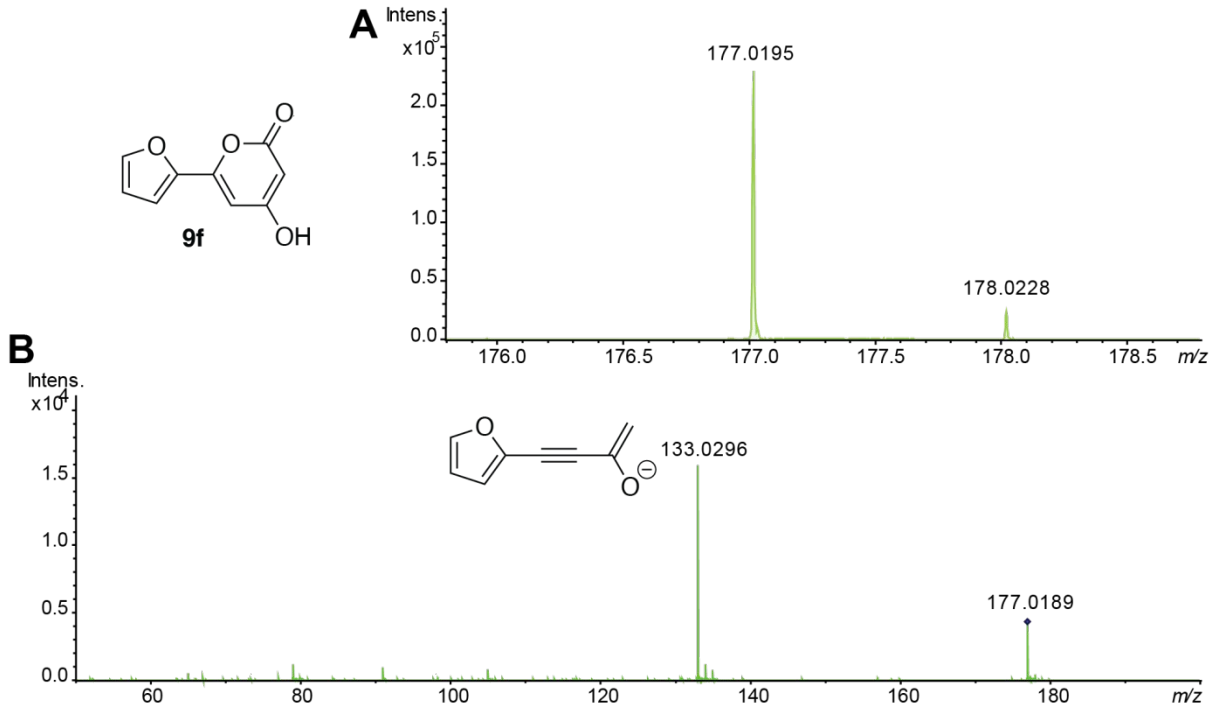


Fig. S48: MS^1 and MS^2 spectra of compound **9f**, product derived from substrate **8f**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9f** (*m/z* calculated for $C_9H_5O_4$ 177.0193, found 177.0195). (B) MS^2 spectra of compound **9f** with rationalized structural annotations of fragment ions.

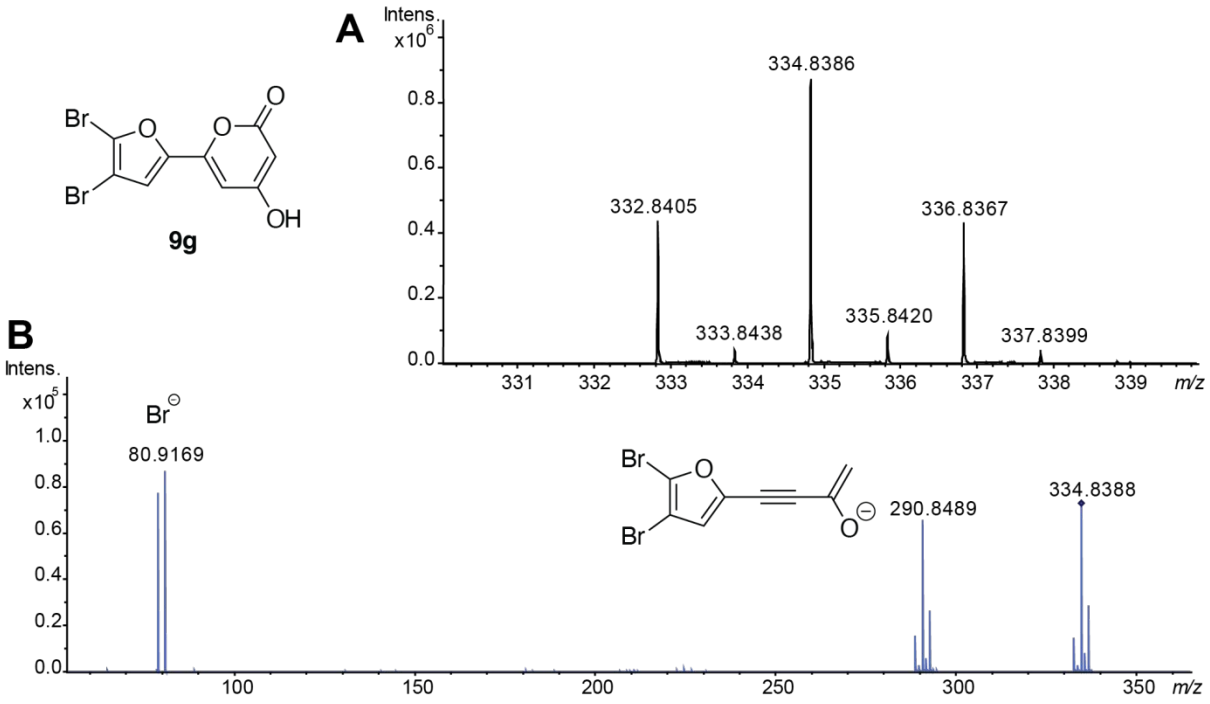


Fig. S49: MS^1 and MS^2 spectra of compound **9g**, product derived from substrate **8g**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9g** (*m/z* calculated for $C_9H_3Br_2O_4$ 332.8404, found 332.8405). (B) MS^2 spectra of compound **9g** with rationalized structural annotations of fragment ions.

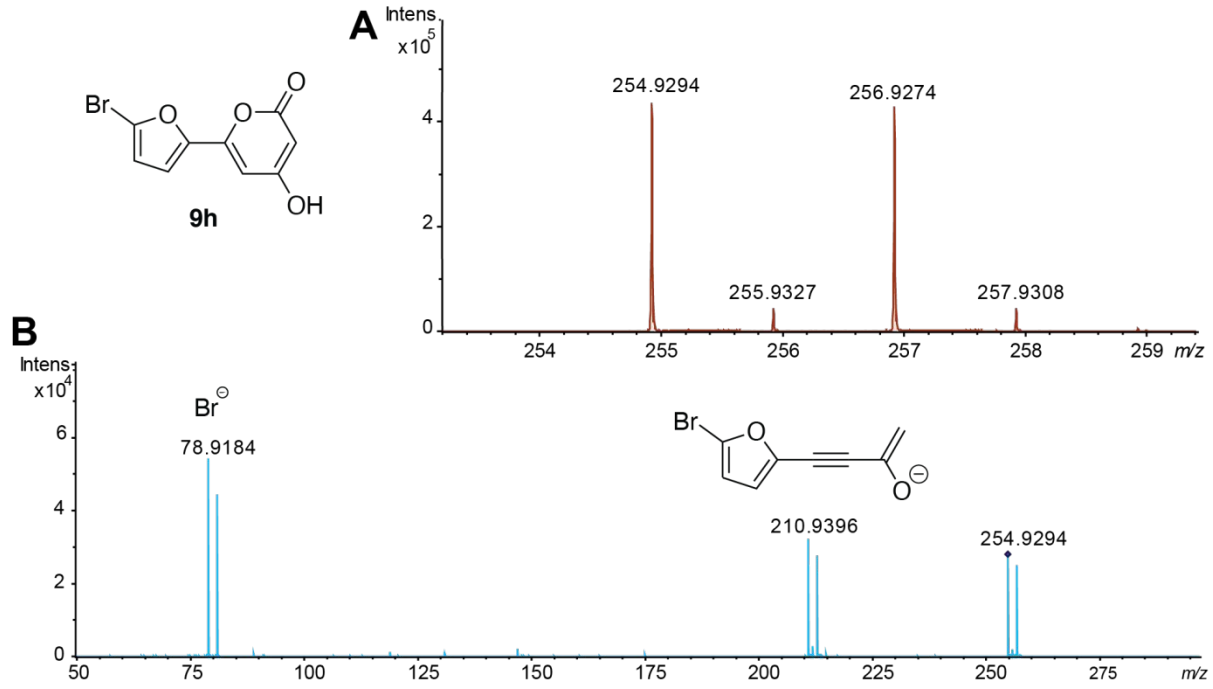


Fig. S50: MS^1 and MS^2 spectra of compound **9h**, product derived from substrate **8h**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9h** (*m/z* calculated for $C_9H_4BrO_4$ 254.9298, found 254.9294). (B) MS^2 spectra of compound **9h** with rationalized structural annotations of fragment ions.

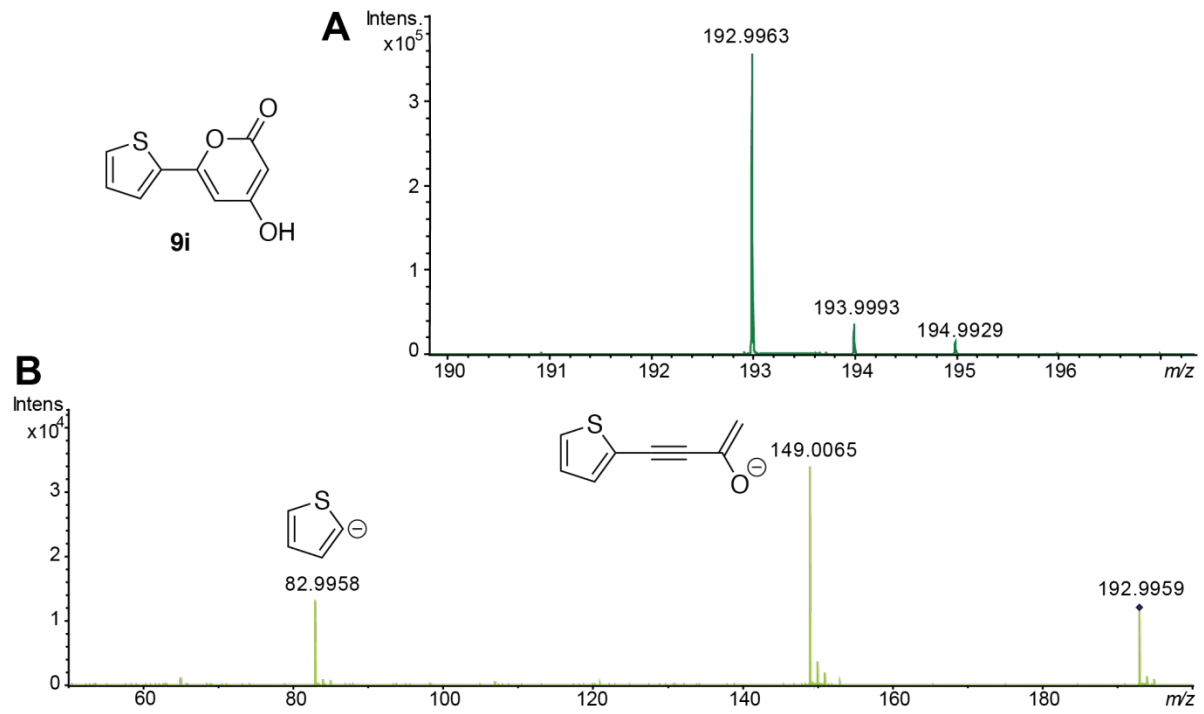


Fig. S51: MS^1 and MS^2 spectra of compound **9i**, product derived from substrate **8i**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9i** (*m/z* calculated for $C_9H_5O_3S$ 192.9965, found 192.9963). (B) MS^2 spectra of compound **9i** with rationalized structural annotations of fragment ions.

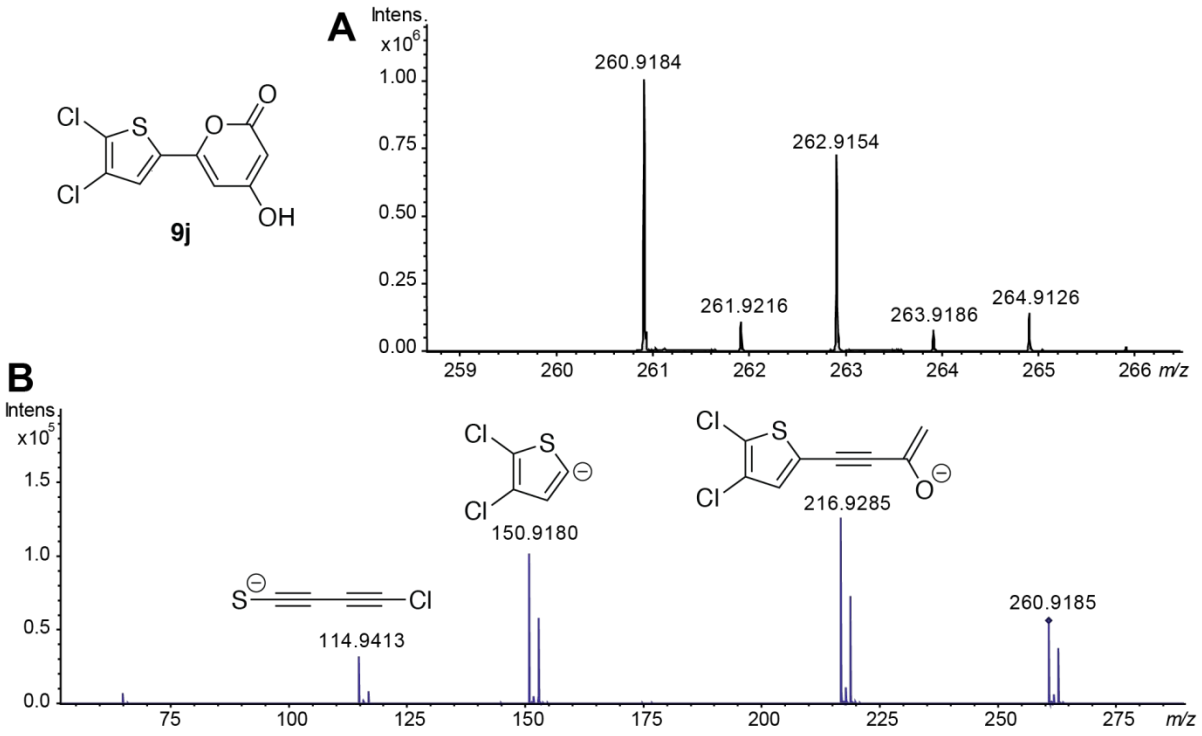


Fig. S52: MS^1 and MS^2 spectra of compound **9j**, product derived from substrate **8j**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9j** (*m/z* calculated for $C_9H_3Cl_2O_3S$ 260.9185, found 260.9184). (B) MS^2 spectra of compound **9j** with rationalized structural annotations of fragment ions.

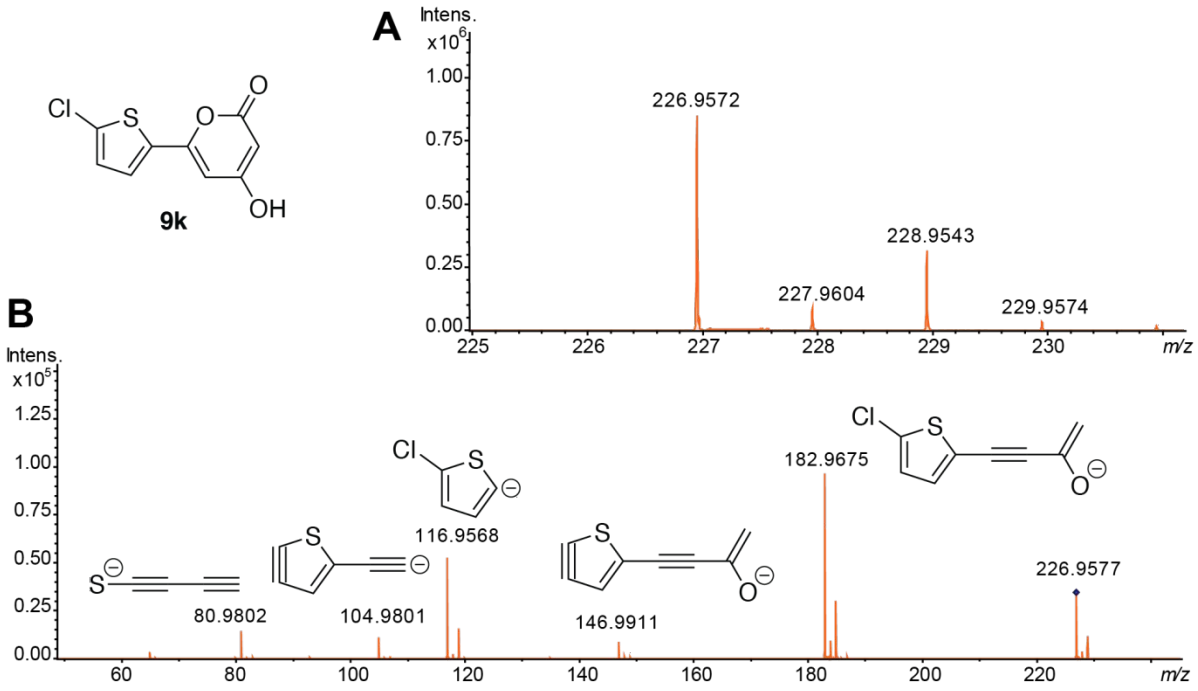


Fig. S53: MS^1 and MS^2 spectra of compound **9k**, product derived from substrate **8k**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9k** (*m/z* calculated for $C_9H_4ClO_3S$ 226.9575, found 226.9572). (B) MS^2 spectra of compound **9k** with rationalized structural annotations of fragment ions.

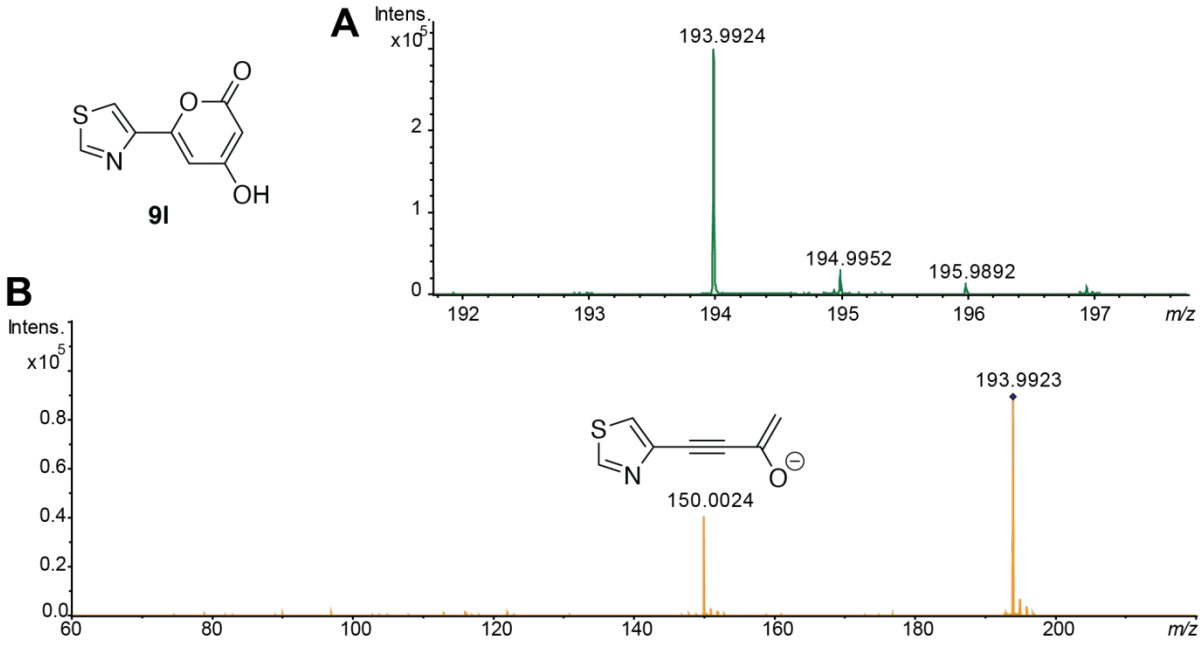


Fig. S54: MS^1 and MS^2 spectra of compound **9l**, product derived from substrate **8l**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9l** (*m/z* calculated for $C_8H_4NO_3S$ 193.9917, found 193.9924). (B) MS^2 spectra of compound **9l** with rationalized structural annotations of fragment ions.

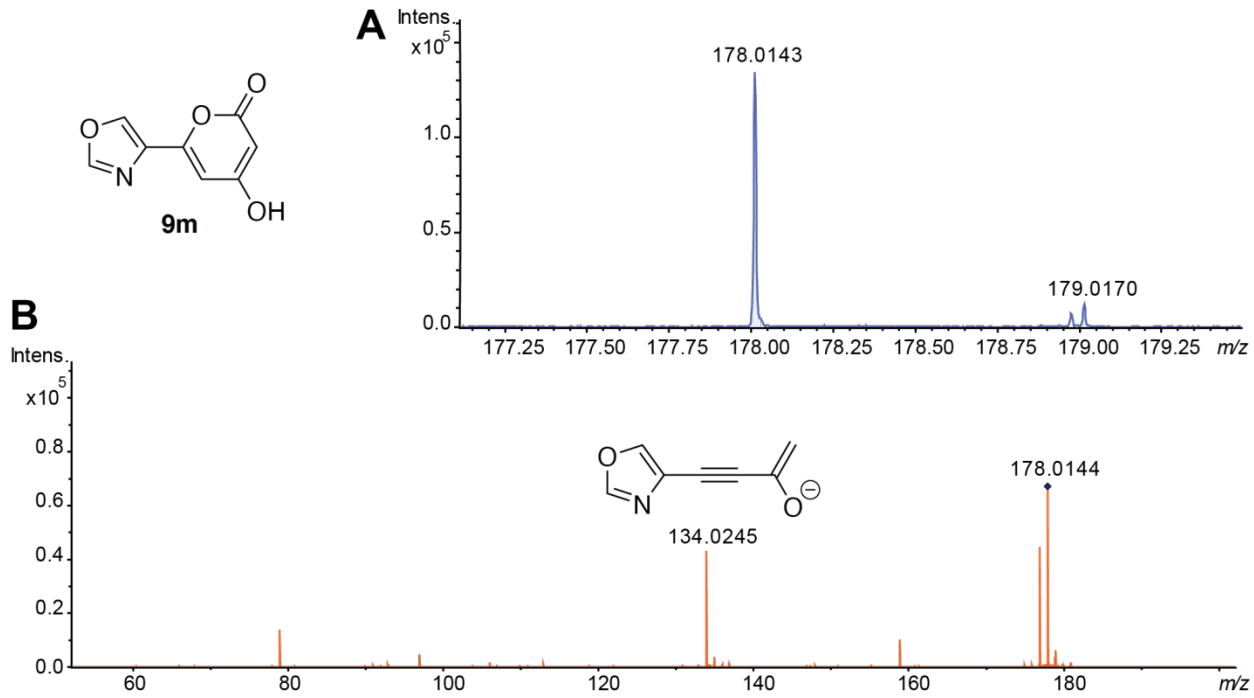


Fig. S55: MS^1 and MS^2 spectra of compound **9m**, product derived from substrate **8m**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9m** (*m/z* calculated for $C_8H_4NO_4$ 178.0146, found 178.0143). (B) MS^2 spectra of compound **9m** with rationalized structural annotations of fragment ions.

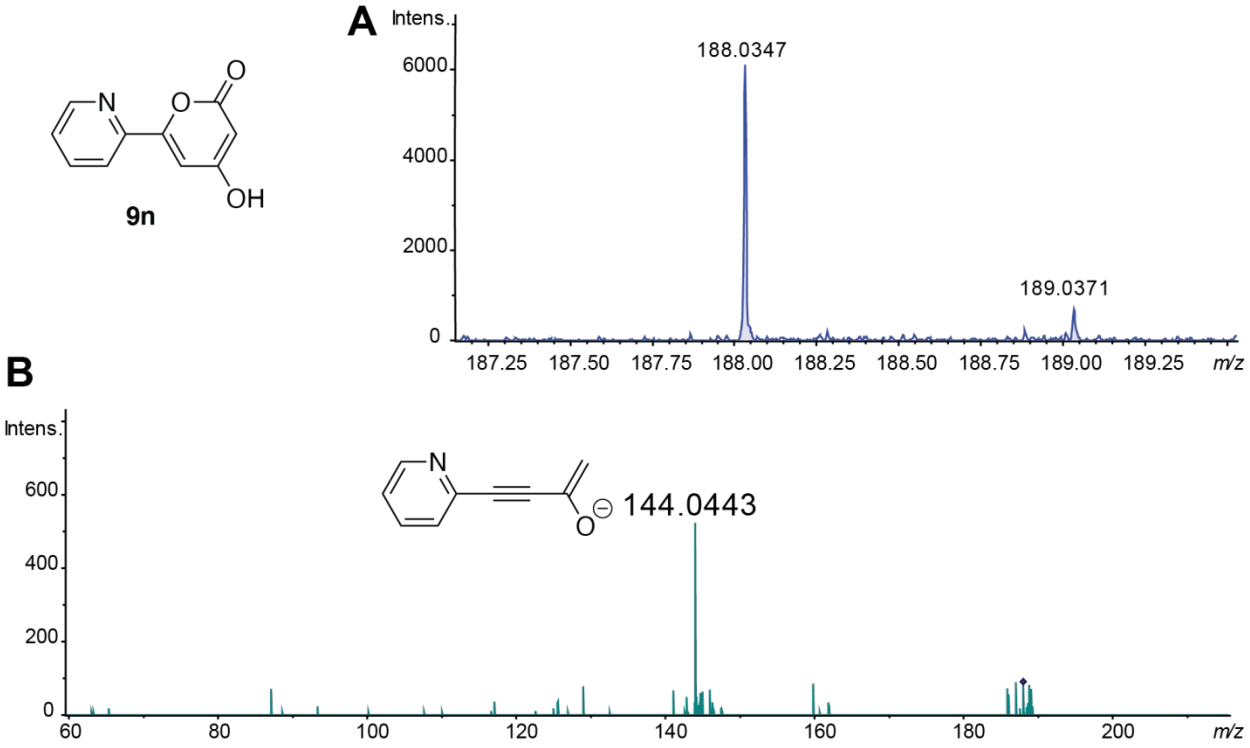


Fig. S56: MS^1 and MS^2 spectra of compound **9n**, product derived from substrate **8n**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9n** (*m/z* calculated for $C_{10}H_6NO_3$ 188.0353, found 188.0347). (B) MS^2 spectra of compound **9n** with rationalized structural annotations of fragment ions.

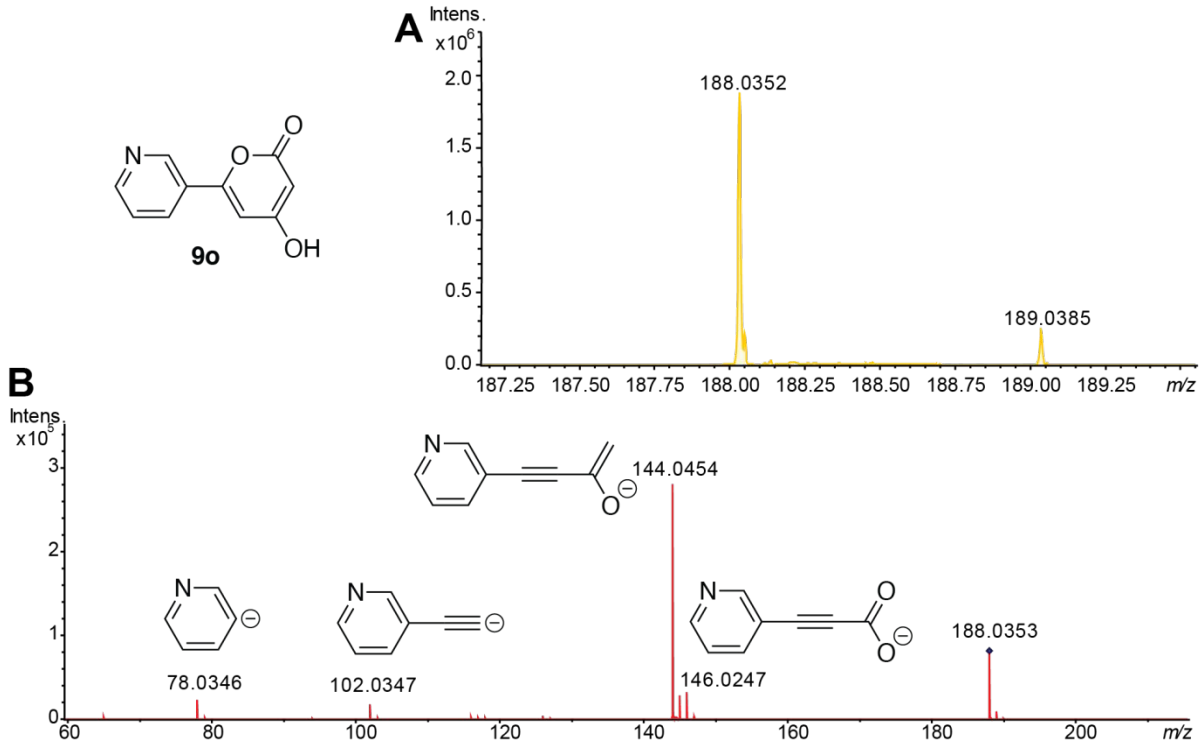


Fig. S57: MS^1 and MS^2 spectra of compound **9o**, product derived from substrate **8o**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9o** (*m/z* calculated for $C_{10}H_6NO_3$ 188.0353, found 188.0352). (B) MS^2 spectra of compound **9o** with rationalized structural annotations of fragment ions.

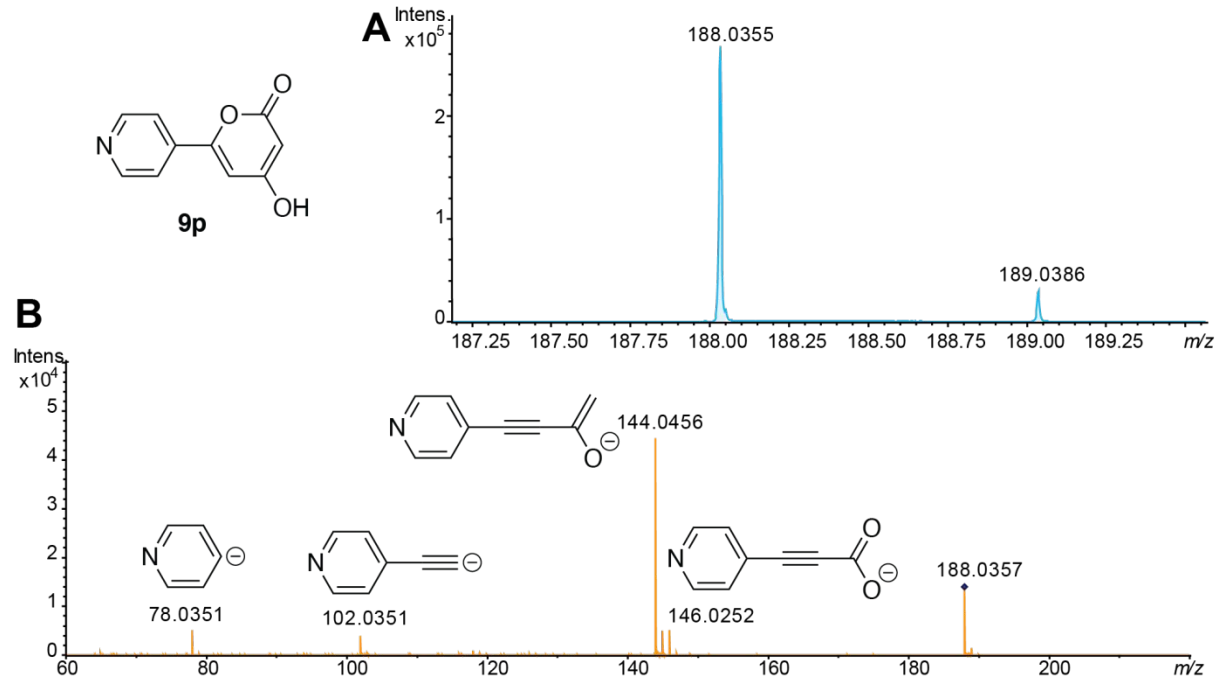


Fig. S58: MS^1 and MS^2 spectra of compound **9p**, product derived from substrate **8p**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9p** (*m/z* calculated for $C_{10}H_6NO_3$ 188.0353, found 188.0355). (B) MS^2 spectra of compound **9p** with rationalized structural annotations of fragment ions.

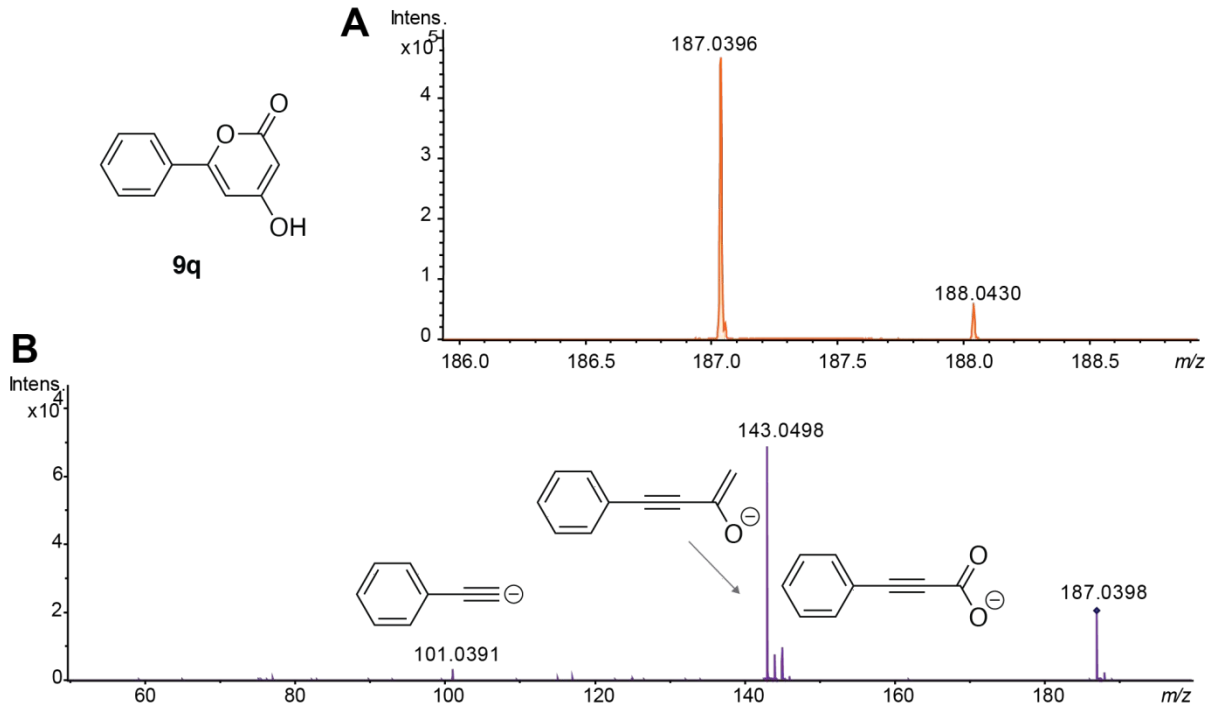


Fig. S59: MS^1 and MS^2 spectra of compound **9q**, product derived from substrate **8q**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9q** (*m/z* calculated for $C_{11}H_7O_3$ 187.0401, found 187.0396). (B) MS^2 spectra of compound **9q** with rationalized structural annotations of fragment ions.

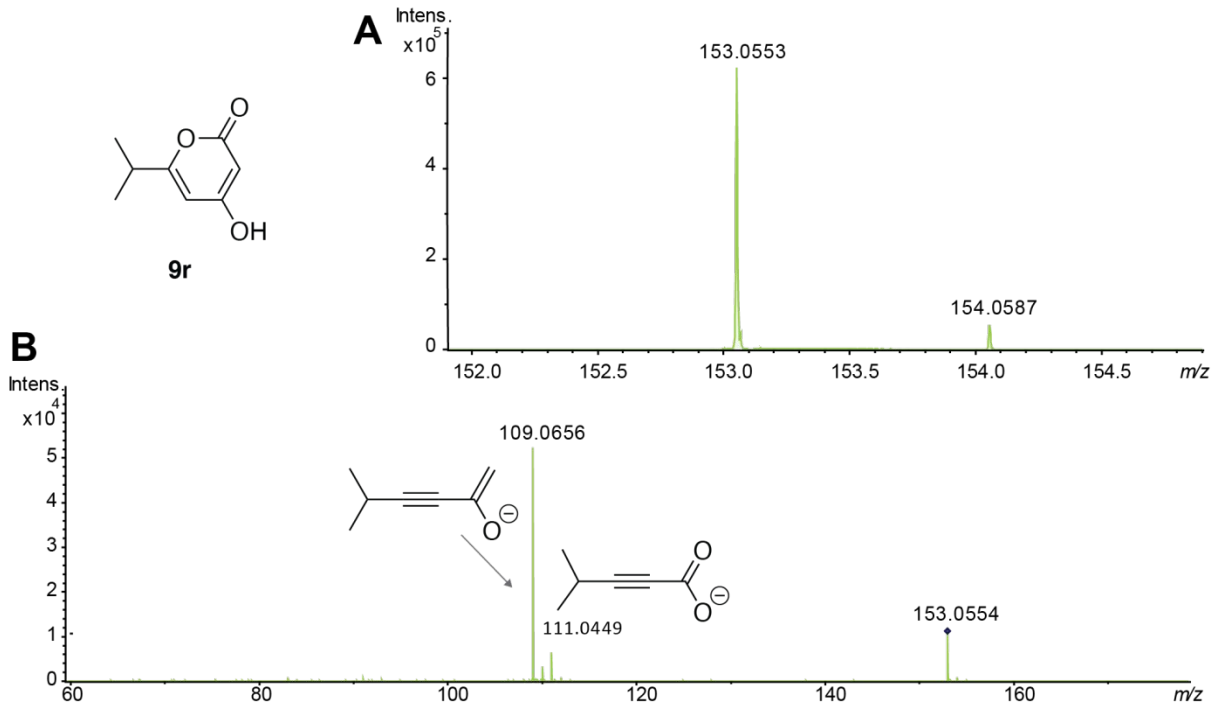


Fig. S60: MS^1 and MS^2 spectra of compound **9r**, product derived from substrate **8r**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9r** (*m/z* calculated for $C_8H_9O_3$ 153.0557, found 153.0553). (B) MS^2 spectra of compound **9r** with rationalized structural annotations of fragment ions.

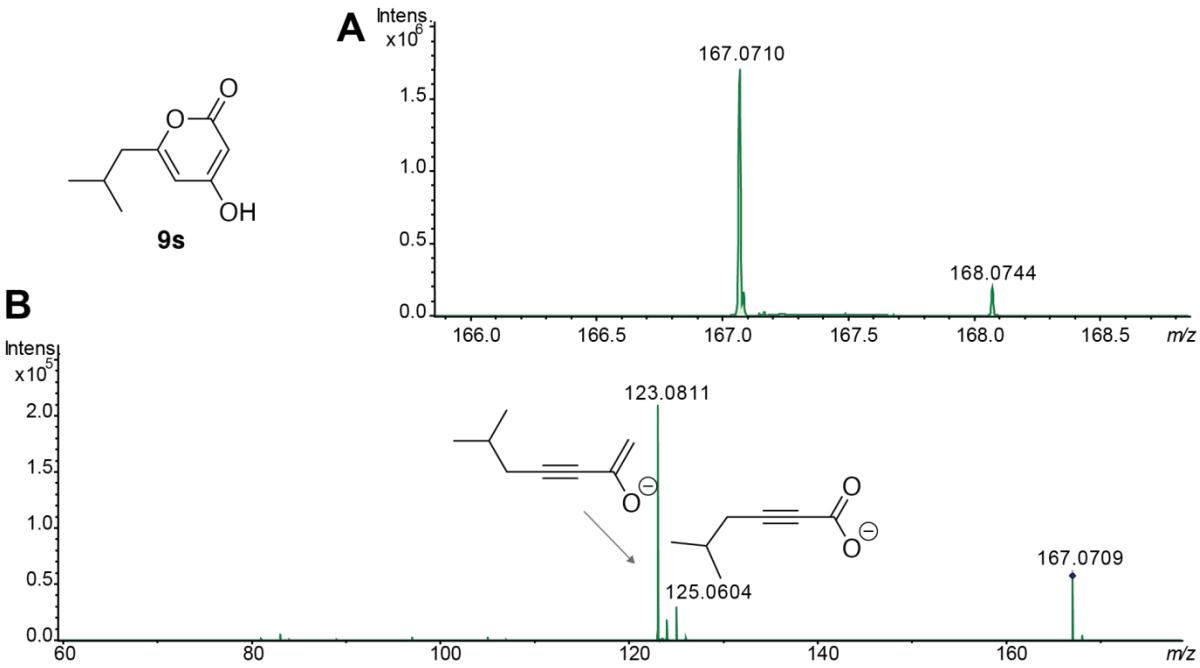


Fig. S61: MS^1 and MS^2 spectra of compound **9s**, product derived from substrate **8s**. (A) HRMS (ESI) identified molecule ions corresponding to $[M-H]^-$ for compound **9s** (*m/z* calculated for $C_9H_{11}O_3$ 167.0714, found 167.0710). (B) MS^2 spectra of compound **9s** with rationalized structural annotations of fragment ions.

SUPPLEMENTARY REFERENCES

1. Yi, D.; Acharya, A.; Gumbart, J. C.; Gutekunst, W. R.; Agarwal, V., Gatekeeping ketosynthases dictate initiation of assembly line biosynthesis of pyrrolic polyketides. *J. Am. Chem. Soc.* **2021,** *143* (20), 7617-7622.

2. Yi, D.; Niroula, D.; Gutekunst, W. R.; Loper, J. E.; Yan, Q.; Agarwal, V., A nonfunctional halogenase masquerades as an aromatizing dehydratase in biosynthesis of pyrrolic polyketides by type I polyketide synthases. *ACS Chem. Biol.* **2022,** *17* (6), 1351-1356.

3. Agarwal, V.; Diethelm, S.; Ray, L.; Garg, N.; Awakawa, T.; Dorrestein, P. C.; Moore, B. S., Chemoenzymatic synthesis of acyl coenzyme A substrates enables *in situ* labeling of small molecules and proteins. *Org. Lett.* **2015,** *17* (18), 4452-4455.

4. Wang, M. Z.; Xu, H.; Liu, T. W.; Feng, Q.; Yu, S. J.; Wang, S. H.; Li, Z. M., Design, synthesis and antifungal activities of novel pyrrole alkaloid analogs. *Eur. J. Med. Chem.* **2011,** *46* (5), 1463-72.

5. Meiser, P.; Weissman, K. J.; Bode, H. B.; Krug, D.; Dickschat, J. S.; Sandmann, A.; Müller, R., DKxanthene biosynthesis—understanding the basis for diversity-oriented synthesis in myxobacterial secondary metabolism. *Chem. Biol.* **2008,** *15* (8), 771-781.

6. Thiede, S.; Wosniok, P. R.; Herkommer, D.; Debnar, T.; Tian, M.; Wang, T.; Schrempp, M.; Menche, D., Total synthesis of leupyrrins A1 and B1, highly potent antifungal agents from the myxobacterium *Sorangium cellulosum*. *Chem. Eur. J* **2017,** *23* (14), 3300-3320.

7. Kusebauch, B.; Brendel, N.; Kirchner, H.; Dahse, H.-M.; Hertweck, C., Assessing oxazole bioisosteres as mutasynthons on the rhizoxin assembly line. *Chembiochem* **2011,** *12* (15), 2284-2288.