

In the format provided by the authors and unedited.

## Global profiling of lysine reactivity and ligandability in the human proteome

Stephan M. Hacker<sup>1\*</sup>, Keriann M. Backus<sup>1\*</sup>, Michael R. Lazear<sup>1</sup>, Stefano Forli<sup>2</sup>, Bruno E. Correia<sup>3</sup>, Benjamin F. Cravatt<sup>1</sup>

<sup>1</sup>Department of Chemical Physiology, The Scripps Research Institute. La Jolla, California 92307, USA.

<sup>2</sup>Department of Integrative Structural and Computational Biology, The Scripps Research Institute. La Jolla, California 92307, USA.

<sup>3</sup> Laboratory of Protein Design & Immunoengineering, Ecole Polytechnique Fédérale de Lausanne. 1015 Lausanne, Switzerland.

\* These authors contributed equally to this work.

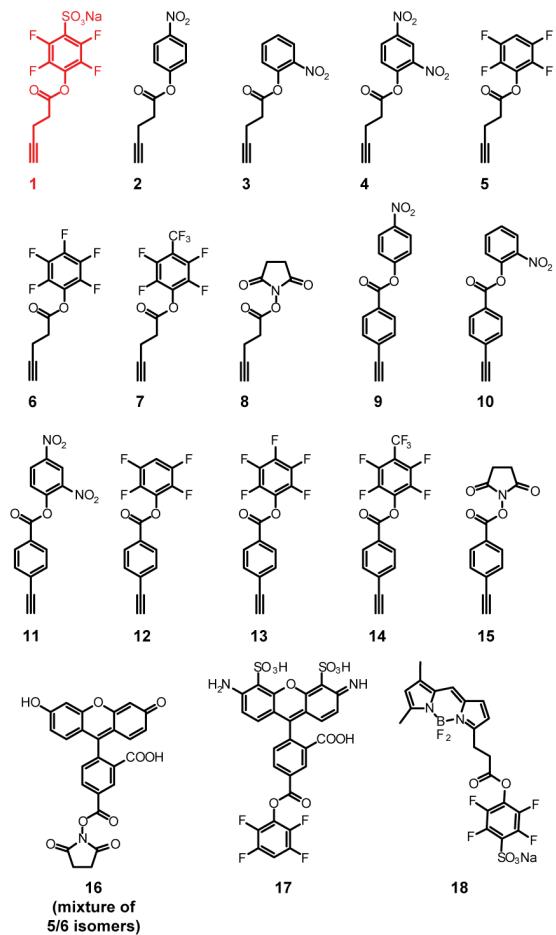
Correspondence and requests for materials should be addressed to S.M.H. (shacker@scripps.edu), K.M.B. (kbackus@scripps.edu) or B.F.C. (cravatt@scripps.edu).

### Content

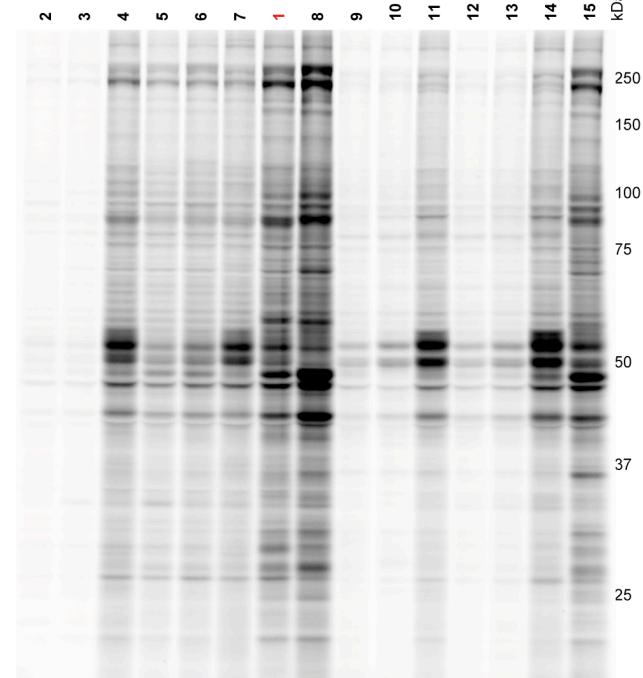
- (A) Supplementary Figures
- (B) Supplementary Table Legends
- (C) Biological Methods
- (D) Synthetic Methods
- (E) NMR spectra
- (F) References

**(A) Supplementary Figures**

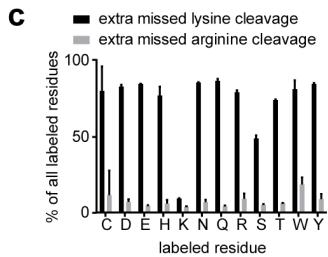
**a**



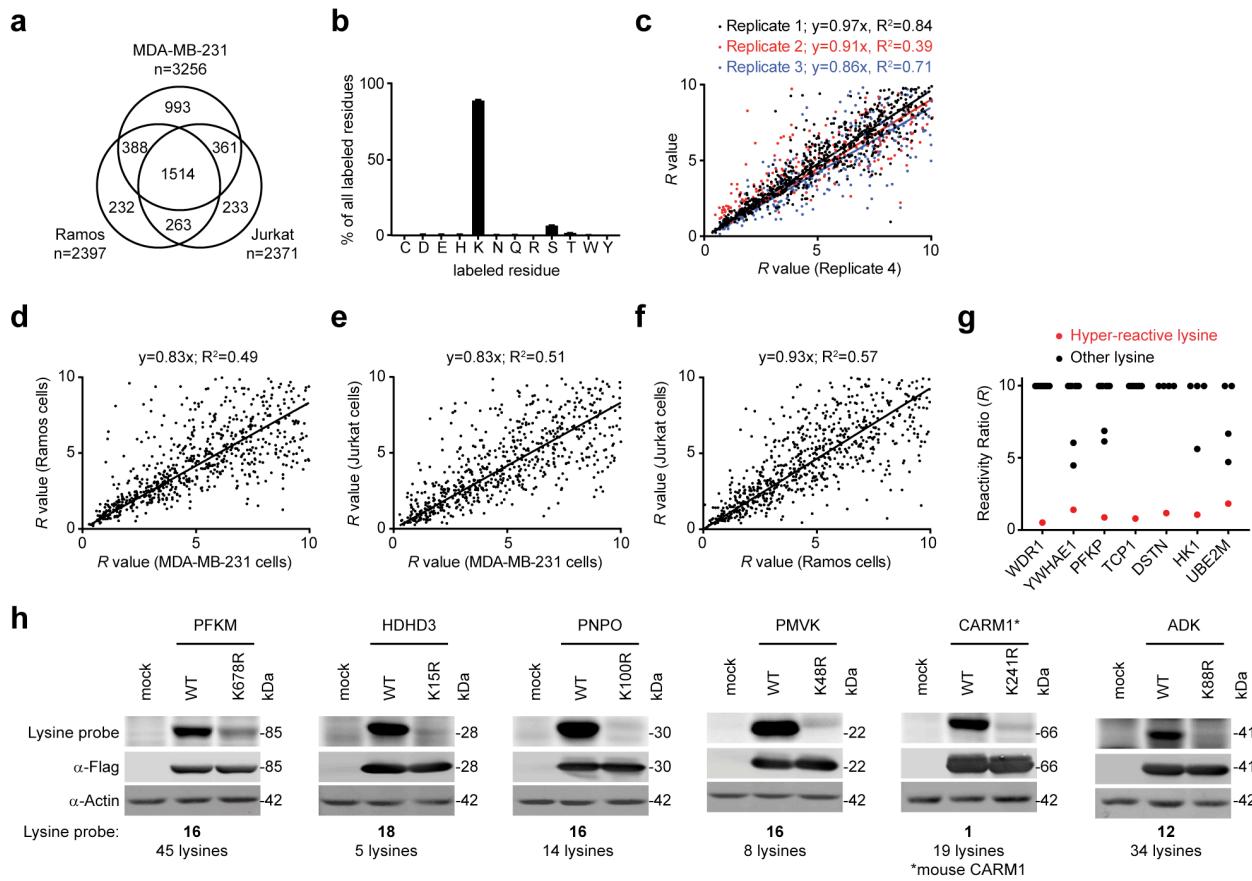
**b**



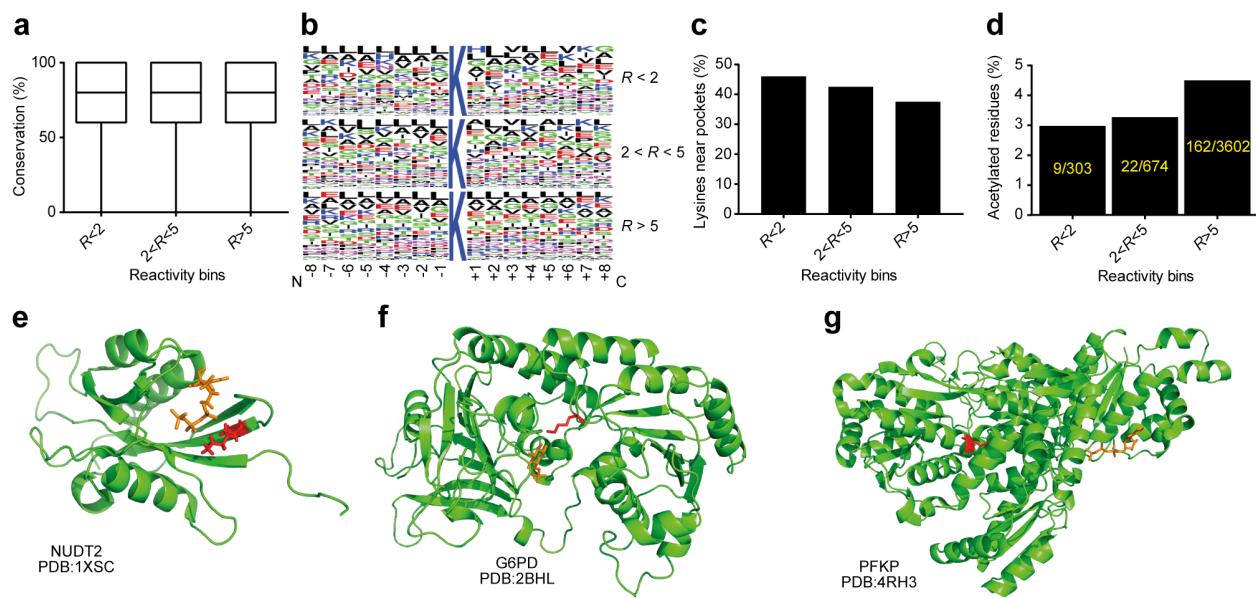
**c**



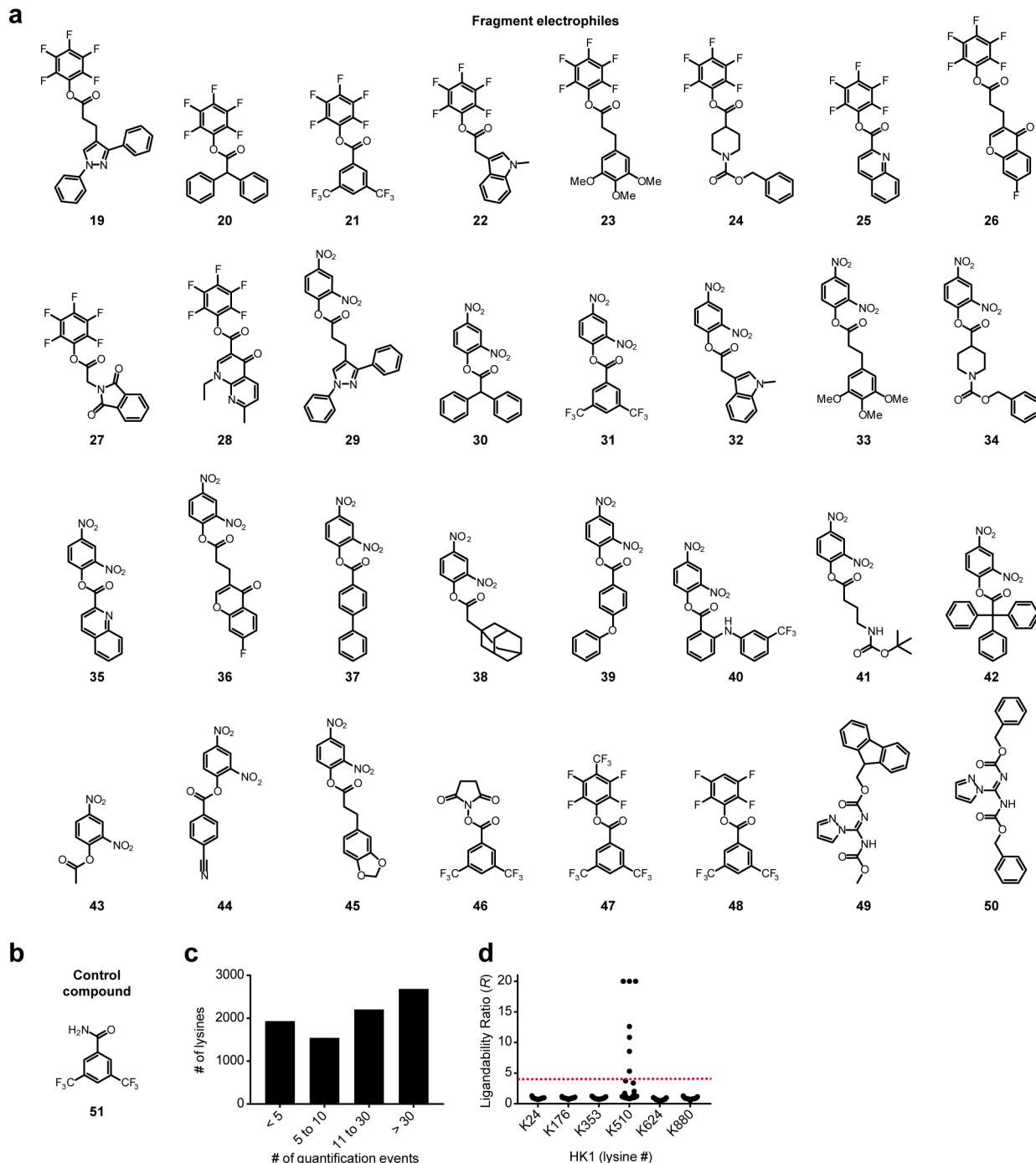
**Supplementary Figure 1.** Evaluating lysine-reactive probes for isoTOP-ABPP. **a**, Structures of various alkyne- (**1-15**) and fluorophore- (**16-18**) modified, amine-reactive probes. **b**, Qualitative assessment of respective proteomic reactivities of probes by SDS-PAGE and in-gel fluorescence scanning of MDA-MB-231 lysates treated with the indicated probes (100  $\mu$ M, 1 h), followed by conjugation to rhodamine-azide by copper-catalyzed azide-alkyne cycloaddition (CuACC). **c**, Most peptides detected as labeled by probe **1** on residues other than lysine contain missed tryptic cleavage events at unmodified lysine residues. Data represent mean values  $\pm$  standard deviation for three experiments.



**Supplementary Figure 2.** Proteome-wide quantification of lysine reactivity. **a**, Overlap of probe 1-labeled peptides detected in isoTOP-ABPP experiments performed with proteomes from the three indicated human cancer cell lines. **b**, Probe 1 also exhibits high selectivity for reacting with lysine in isoTOP-ABPP experiments comparing MDA-MB-231 cell lysates treated with 0.1 vs 1.0 mM of probe 1. Data was evaluated as in Fig. 1b, including requiring peptides to feature no missed cleavage sites at unmodified lysine residues. Data represent mean values  $\pm$  standard deviation for two experiments. **c-f**, Consistency of lysine reactivity ratios (*R* values) for isoTOP-ABPP experiments comparing 0.1 and 1.0 mM of probe 1 with (**c**) biological replicates of the same proteome (MDA-MB-231 lysates), or (**d-f**) proteomes from three different human cancer cell lines (MDA-MB-231, Ramos and Jurkat cells). **g**, *R* values for hyper-reactive (red) and medium/low-reactivity (black) lysines found within the same protein. Note that, for each protein, only a single hyper-reactive lysine is found (*R* values  $<$  2.0) among several quantified lysines. **h**, Hyper-reactive lysines can be site-selectively labeled by activated ester probes. HEK 293T cells expressing representative proteins with hyper-reactive lysines (or the corresponding lysine-to-arginine mutant) as FLAG-tagged fusion proteins were treated with the indicated lysine-reactive probe and analyzed by gel-based ABPP. \*mouse orthologue of CARM1.



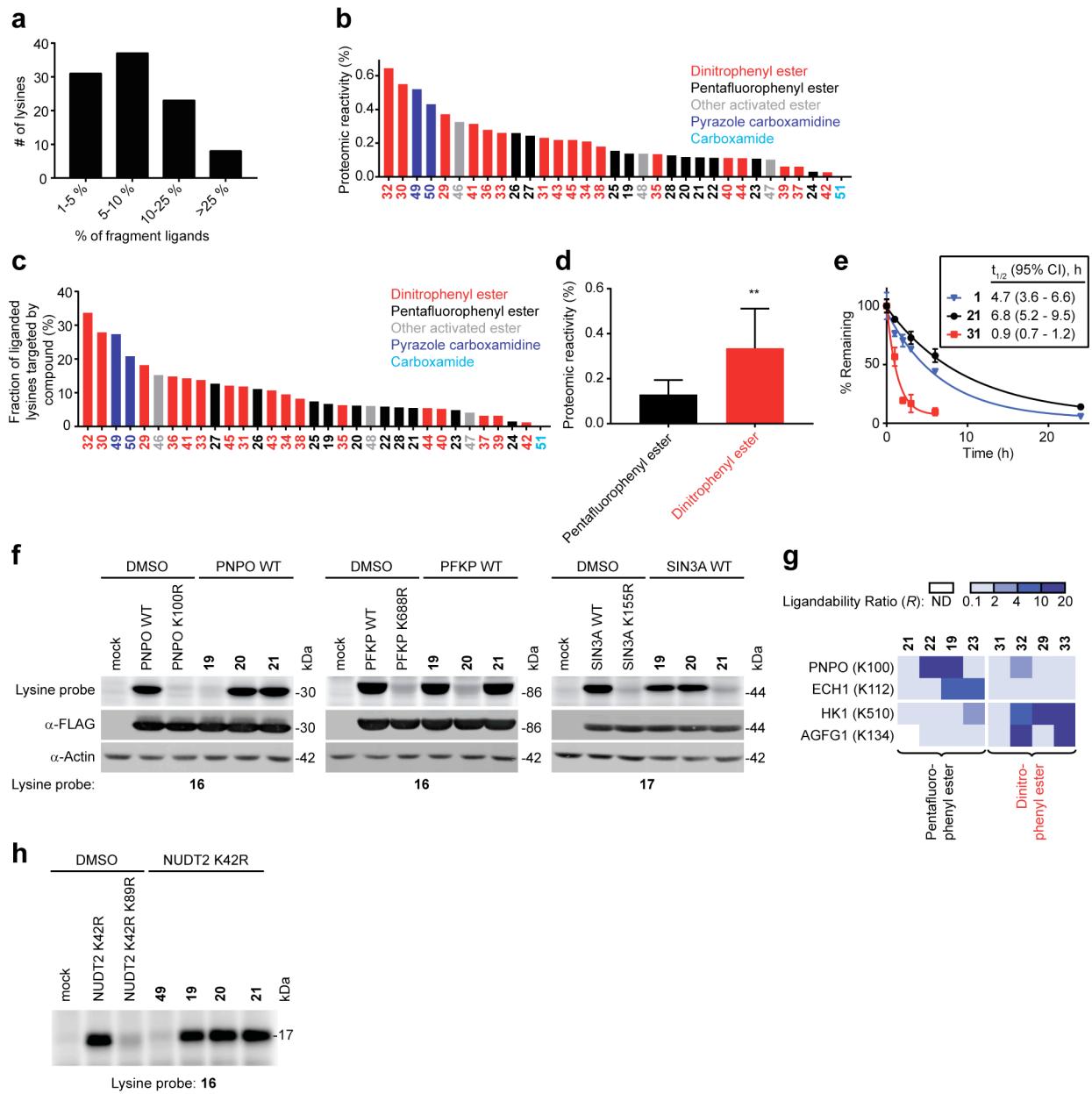
**Supplementary Figure 3.** Global and specific assessments of probe 1-reactive lysines. **a**, Box and whiskers plot showing the distribution of lysine conservation across *M. musculus*, *X. laevis*, *D. melanogaster*, *C. elegans* and *D. rerio* for probe 1-labeled lysines from different reactivity groups. **b**, Frequency plots showing no apparent conserved motifs for lysines from different reactivity groups. **c**, Hyper-reactive lysines are enriched near pockets. Pockets were detected using AutoSite and lysines were considered near a pocket, if they were less than 3.5 Å away from the pocket. **d**, Hyper-reactive lysines are less likely to be acetylated than lysines of lower reactivity. **e-g**, Structures of proteins with hyper-reactive lysines. Hyper-reactive lysines (K89 for NUDT2, K171 for G6PD and K688 for PFKP) are shown in red and molecules bound in the active site of the protein in orange (ATP for NUDT2, glucose-6-phosphate for G6PD and AMPPCP for PFKP).



**Supplementary Figure 4.** Proteome-wide screening of lysine-reactive fragment electrophiles.

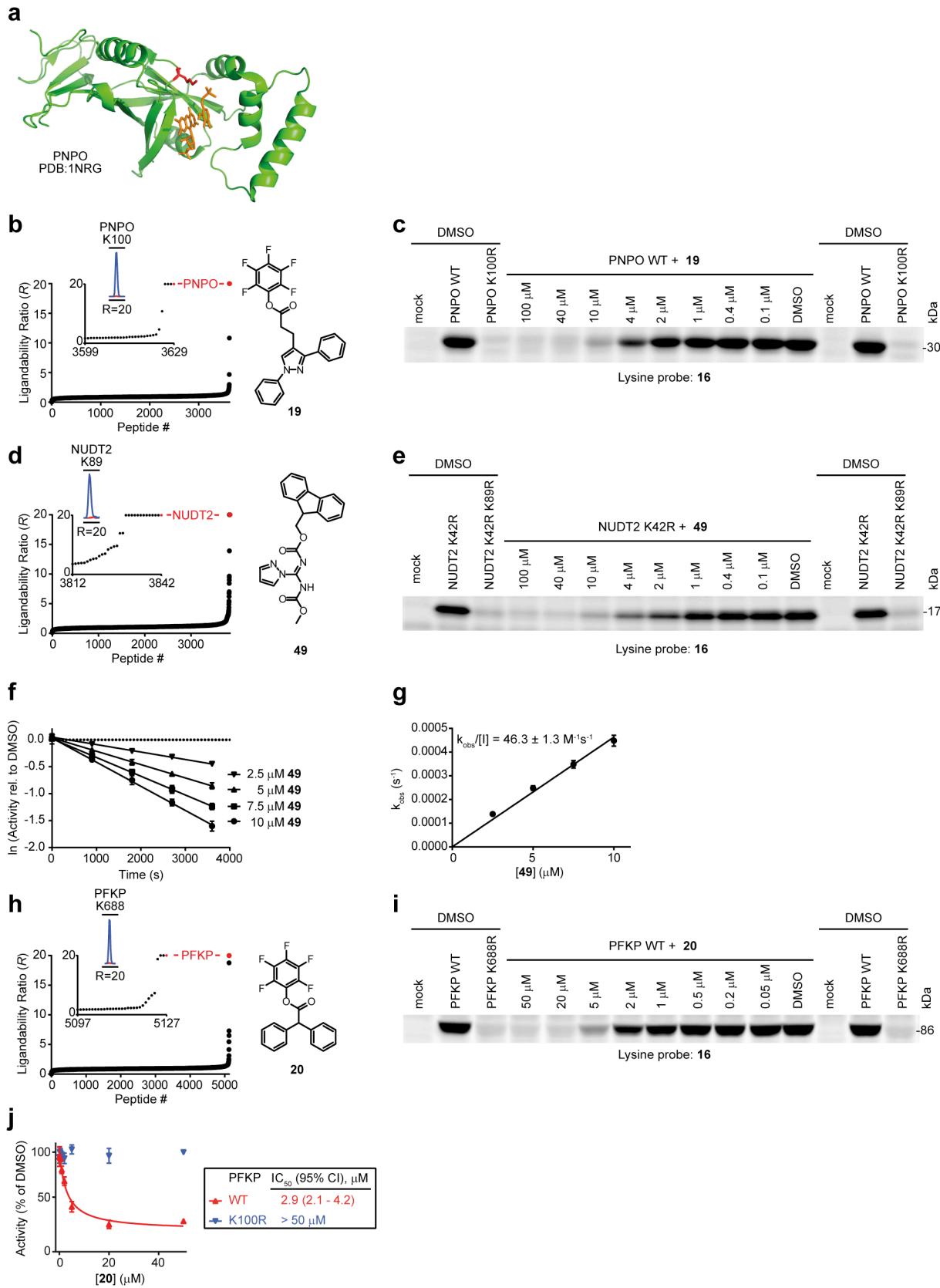
**a, b** Structures of compounds in the lysine-reactive fragment electrophile library, including non-electrophilic, amide-containing control compound **51** (**b**). **c**, Frequency of quantification of all lysines for the competitive isoTOP-ABPP experiments performed with fragment electrophiles. Note that lysines were required to have been quantified in at least two experiments for interpretation of ligandability. **d**, *R* values for six lysine residues in hexokinase-1 (HK1)

quantified by isoTOP-ABPP, identifying K510 as the only liganded lysine in HK1. Each point represents a distinct fragment-lysine interaction quantified by isoTOP-ABPP. The red dashed line marks an  $R$  value of 4 used to define a fragment liganding event.

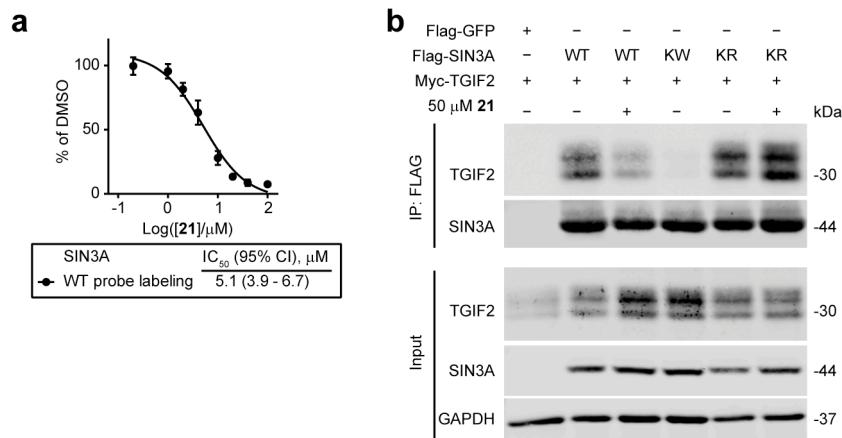


**Supplementary Figure 5.** Lysine-reactive fragment electrophiles exhibit distinct proteome-wide reactivity profiles. **a**, Most liganded lysines are targeted by a limited subset (< 10%) of the fragment electrophiles. Histogram depicting the number of liganded lysines targeted by different percentages of fragments. Percentage is the fraction of ligands among the fragments that this lysine was quantified for. For this analysis, liganded lysines were required to have been quantified in isoTOP-ABPP experiments for at least 10 different fragments. **b**, Rank order of reactivity values for fragment electrophiles calculated as the percentage of all quantified lysines with  $R \geq 4$  for each fragment. **c**, Rank order of reactivity values of fragment electrophiles calculated as the percentage of all liganded lysines with  $R \geq 4$  for

each fragment. **d**, Average proteomic reactivity values for eight pentafluorophenyl and eight dinitrophenyl esters that share common fragment-based binding elements. Data represent mean values  $\pm$  standard deviation for the eight fragment electrophiles in each group. Statistical significance was calculated by unpaired students t-test comparing dinitrophenyl to pentafluorophenyl esters; \*\*,  $p < 0.01$ . **e**, Study of the hydrolytic stability of STP alkyne (**1**) and two representative library members (**21**, **31**). Compounds were incubated in PBS at ambient temperature for indicated times and evaluated by HPLC with UV detection. Half-lives ( $t_{1/2}$ ) were determined by mono-exponential fit of the change in peak area for **1**, **21** and **31**, respectively. CI, confidence interval. **f**, Western blot analysis confirming equivalent protein expression for gel-based ABPP experiments depicted in **Fig. 4b**. SIN3A corresponds to a.a. 1-400 of SIN3A. **g**, Heat-map showing proteins that interact preferentially with dinitrophenyl and pentafluorophenyl esters, respectively. **h**, Probe **16**-labeling of K89 in NUDT2 is quantitatively blocked by guanidinylation fragment electrophile **49**, but not by the three tested activated ester fragment electrophiles. NUDT2 or the K89R mutant was expressed in HEK 293T cells, and cell lysates were treated with the indicated fragments (50  $\mu$ M, 1 h) followed by lysine-reactive probe **16** and gel-based ABPP. Note that all NUDT2 variants contain an additional K42R mutation to allow for gel-based detection of K89 labeling by probe **16**. For panel **e**, data represent mean values  $\pm$  standard deviation for three experiments.



**Supplementary Figure 6.** Site-specific fragment-lysine reactions and their functional effects on proteins. **a**, Structure of PNPO (PDB ID: 1NRG). Hyper-reactive lysine K100 is shown in red and FMN and pyridoxal-5'-phosphate bound in the active site are shown in orange. **b-e,h,i**, Competitive isoTOP-ABPP analysis (**b, d, h**) of MDA-MB-231 cell lysate treated with the indicated fragment electrophiles (50  $\mu$ M for 1 h (**b,h**) or 100  $\mu$ M for 4 h (**d**)) followed by probe **1**. Red dots ( $R$  value = 20) indicate full blockade of probe **1** labeling of the indicated lysines in PNPO (**b**), NUDT2 (**d**), and PFKP (**h**). Insets show lysines with the top 30  $R$  values and the MS1 chromatographic peaks for the fragment-competed lysines in PNPO, PFKP and NUDT2. **c,e,i**, Lysates from HEK 293T cells recombinantly expressing PNPO (**c**), NUDT2 (**e**), and PFKP (**i**) or the indicated lysine-to-arginine mutants were treated with the indicated concentrations of fragment electrophiles for 1 h followed by the indicated lysine-reactive probes, and analyzed by gel-based ABPP. Note that all NUDT2 variants contain an additional K42R mutation to allow gel-based detection of K89 labeling. **f**, Plot of the natural logarithm of the relative activity of NUDT2 after pre-incubation for indicated times with the indicated concentration of **49**.  $k_{obs}$  for the respective concentration of **49** is calculated from the slope of the linear fit. **g**, Plot of  $k_{obs}$  against the concentration of **49**.  $k_{obs}/[I]$  is calculated from the slope of the linear fit. **j**, Fragment **20** blocks the catalytic activity of PFKP in a concentration-dependent manner to produce a maximal inhibitory effect of ~ 80%. CI, confidence interval. For panels **f,g,j**, data represent mean values  $\pm$  standard deviation for three experiments.



**Supplementary Figure 7.** SIN3A is modulated site-specifically by lysine-reactive fragment electrophiles. **a**, IC<sub>50</sub> curve for blockade of probe **17**-labeling of SIN3A by fragment electrophile **21**. CI, confidence interval. **b**, Flag-SIN3A or the indicated Flag-SIN3A mutants (a.a. 1-400), or Flag-GFP, were co-expressed in HEK 293T with Myc-TGIF2. Cell lysates were treated with DMSO or **21** (50  $\mu M$ , 1 h), and proteins immunoprecipitated. The inputs (10% loading) and the immunoprecipitates were analyzed by Western blotting. For panel **a**, data represent mean values  $\pm$  standard deviation for three experiments.

## (B) Supplementary Table Legends

**Supplementary Table 1.** Complete proteomic data for lysine reactivity experiments. (**Tab 1**) Average Reactivity Ratios across all cell lines for all quantified peptides that meet the quality filters described in the Methods section. Shown are the average  $R_{10:1}$  and  $R_{1:1}$  ratios for all peptides as well as the designation of the functional class of the corresponding protein. (**Tab 2**) Median Reactivity Ratios in all cell lines for all quantified peptides that meet the quality filters described in the Methods section. Shown are the median  $R_{10:1}$  and  $R_{1:1}$  ratios for each peptide in the soluble proteome from MDA-MB-231, Jurkat and Ramos cells as well as in the membrane proteome from MDA-MB-231 cells. Furthermore, alternative assignments for lysines stemming from redundant peptides and the functional classes for the corresponding proteins are shown. (**Tab 3** and **Tab 4**) Median  $R_{10:1}$  (**Tab 3**) and  $R_{1:1}$  (**Tab 4**) ratios in the soluble proteome of MDA-MB-231 cells. Shown are the median  $R$  values, their standard deviation and the number of quantification events across all biological replicates for each peptide. (**Tab 5**) Individual Reactivity Ratios ( $R_{10:1}$ ) for all peptides quantified in each individual replicate in the soluble proteome from MDA-MB-231 cells. Shown are the individual  $R$  values in each experiment and the number of quantification events. (**Tab 6**) Gel-based ABPP experiments with representative proteins containing hyper-reactive lysines demonstrate that reactive lysines can be labeled selectively. Shown are the protein names and the residue number of their respective hyper-reactive lysine residues. Both wild-type and the lysine-to-arginine or lysine-to-alanine mutant of the indicated hyper-reactive lysine are exposed to a panel of lysine-reactive activated esters and the discrimination between wild-type and mutant protein labeling is evaluated by gel electrophoresis and in-gel fluorescence scanning. The number of lysines in the protein and the type of mutation (lysine-to-arginine: KR; lysine-to-alanine: KA) are also given.

**Supplementary Table 2.** Complete proteomic data for lysine ligandability experiments. (**Tab 1**) Average  $R$  values for all lysines that are quantified as liganded. Shown are the average  $R$  values for all compounds, alternative lysine assignments stemming from redundant peptides, the functional class of the corresponding protein and a designation, if the protein is found in Drugbank. (**Tab 2**) Information as in **Tab 1** for lysines that are quantified in competitive isoTOP-ABPP experiments, but not detected as liganded. (**Tab 3 – Tab 9**) Average  $R$  values for all peptides detected with the indicated competitor. Shown are the average  $R$  value, its standard deviation and the number of quantification events for each peptide with the respective

compound. (**Tab 10 – Tab 15**) Results of the site of labeling experiments performed with quantification using reductive dimethylation (ReDiMe) for the indicated proteins with the indicated compounds. Shown are the name of the protein, the number of the compound, the protease used, its amino acid specificity, the sequence coverage in the experiment, the sequence and calculated and detected mass for the modified peptide as well as the *R* values for all quantified unmodified peptides. (**Tab 16**) Median Enrichment Ratios for all quantified proteins in all co-immunoprecipitation experiments. Shown are the median enrichment ratios of all peptides from each protein identified and the number of spectral counts in the respective replicate. (**Tab 17**) List of all proteins that were preferentially enriched by wild-type SIN3A over GFP. Proteins were required to have a median Enrichment Ratio of at least 5 in the experiments comparing immunoprecipitation with wild-type SIN3A and GFP. Shown are the median Enrichment Ratios for each kind of experiment for each protein, as well as their standard deviations and the total number of quantified peptides.

## (C) Biological Methods

**Preparation of human cancer cell line proteomes.** All cell lines were obtained from ATCC, tested negative for mycoplasma contamination, and were used without further authentication, maintaining a low passage number (< 20 passages). Cell lines were grown at 37 °C with 5% CO<sub>2</sub>. MDA-MB-231 (ATCC: HTB-26), and HEK-293T (ATCC: CRL-3216) cells were grown in DMEM medium (Corning, 15-013-CV) supplemented with 10% fetal bovine serum (FBS, Omega Scientific, FB-11, Lot #441224), penicillin, streptomycin and glutamine. Jurkat A3 (ATCC: CRL-2570) and Ramos (ATCC: CRL-1596) cells were grown in RPMI-1640 medium (Corning, 15-040-CV) supplemented with 10% FBS, penicillin, streptomycin and glutamine. For *in vitro* labeling, cells were grown to 100% confluence for MDA-MB-231 cells or until cell density reached 1.5 million cells per ml for Ramos and Jurkat cells. Cells were washed with cold PBS, scraped with cold PBS and cell pellets were isolated by centrifugation (1,400g, 3 min, 4 °C), and stored at -80 °C until use. Cell pellets were resuspended in PBS, lysed by sonication and fractionated (100,000g, 45 min) to yield soluble and membrane fractions, which were then adjusted to a final protein concentration of 1.8 mg ml<sup>-1</sup> (soluble fraction) for compound screening by competitive isoTOP-ABPP and 1.5 mg ml<sup>-1</sup> (soluble fraction) or 3 mg ml<sup>-1</sup> (membrane fraction) for reactivity measurements by isoTOP-ABPP. For gel-based ABPP lysates were adjusted to 1.8 mg ml<sup>-1</sup> (soluble fraction) for MBA-MB-231 lysates and 1 mg ml<sup>-1</sup> (soluble fraction) for HEK 293T lysates expressing target proteins. The lysates were prepared fresh from frozen pellets directly before each experiment. Protein concentration was determined using the Bio-Rad DC™ protein assay kit.

### isoTOP-ABPP sample preparation.

**In vitro covalent fragment treatment for isoTOP-ABPP.** All compounds were made up as solutions in DMSO (100×) and were used at a final concentration of 50 µM for activated esters and 100 µM for guanidinyllating agents. For each profiling sample, 0.5 ml of lysate was treated with 5 µl of the 100× compound stock solution or 5 µl of DMSO. Samples were treated with activated esters for 1 h and with guanidinyllating agents for 4 h.

**STP-alkyne labeling and click chemistry.** For concentration-dependent reactivity measurements by isoTOP-ABPP, 0.5 ml proteome aliquots were treated at ambient temperature with 1 mM STP-alkyne **1** (5 µl of 100 mM stock in DMSO) and 0.1 mM STP alkyne **1** (5 µl of 10 mM stock in DMSO), respectively. For competitive isoTOP-ABPP, after *in vitro* fragment treatment (detailed above), the samples were labeled for 1 h at ambient temperature with 0.1

mM STP-alkyne **1** (5  $\mu$ l of 10 mM stock in DMSO). Samples were conjugated by copper-mediated azide-alkyne cycloaddition (CuAAC) to either the light (1 mM STP-alkyne or fragment treated) or heavy (0.1 mM STP-alkyne or DMSO treated) TEV tags (10  $\mu$ l of 5 mM stocks in DMSO, final concentration = 100  $\mu$ M) using tris(2-carboxyethyl)phosphine hydrochloride (TCEP; fresh 50 $\times$  stock in water, final concentration = 1 mM), TBTA ligand (17 $\times$  stock in DMSO:t-butanol 1:4, final concentration = 100  $\mu$ M) and CuSO<sub>4</sub> (50 $\times$  stock in water, final concentration = 1 mM). The samples were allowed to react for 1 h at room temperature, at which point the proteins from combined light and heavy samples were precipitated by chloroform-methanol extraction. The pellets were solubilized in PBS containing 1.2 % SDS (1 ml) with sonication and heating (5 min, 95 °C) and any insoluble material was removed by an additional centrifugation step at ambient temperature (5,000g, 10 min).

**Streptavidin enrichment.** For each sample, 100  $\mu$ l of streptavidin-agarose beads slurry (Pierce, 20349) was washed in 10 ml PBS (3 $\times$ ) and then resuspended in 6 ml PBS. The SDS-solubilized proteins were added to the suspension of streptavidin-agarose beads and the bead mixture was rotated for 3 h at ambient temperature. After incubation, the beads were pelleted by centrifugation (2,800g, 3 min) and were washed (1  $\times$  10 ml 0.2 % SDS in PBS, 2  $\times$  10 ml PBS and 2  $\times$  10 ml water).

**Trypsin and TEV digestion.** The beads were transferred to Eppendorf tubes with 1 ml PBS, centrifuged (20,000g, 1 min), and resuspended in PBS containing 6 M urea (500  $\mu$ l). To this was added 10 mM DTT (25  $\mu$ l of a 200 mM stock in water) and the beads were incubated at 65 °C for 15 min. 20 mM iodoacetamide (25  $\mu$ l of a 400 mM stock in water) was then added and allowed to react at 37 °C for 30 min with shaking. The bead mixture was diluted with 950  $\mu$ l PBS, pelleted by centrifugation (20,000g, 1 min), and resuspended in PBS containing 2M urea (200  $\mu$ l). To this was added 1 mM CaCl<sub>2</sub> (2  $\mu$ l of a 200 mM stock in water) and trypsin (2  $\mu$ g, Promega, sequencing grade in 4  $\mu$ l trypsin resuspension buffer) and the samples were allowed to digest overnight at 37 °C with shaking. The beads were separated from the digest with Micro Bio-Spin columns (Bio-Rad) by centrifugation (800g, 30 sec), washed (2  $\times$  1 ml PBS and 2  $\times$  1 ml water) and then transferred to fresh Eppendorf tubes with 1 ml water. The washed beads were washed once further in 140  $\mu$ l TEV buffer (50 mM Tris, pH 8, 0.5 mM EDTA, 1 mM DTT) and then resuspended in 140  $\mu$ l TEV buffer. 5  $\mu$ l TEV protease (80  $\mu$ M stock solution) was added and the reactions were rotated overnight at 30 °C. The TEV digest was separated from the beads with Micro Bio-Spin columns by centrifugation (8,000g, 3 min) and the beads were

washed once with water (100 µl). The samples were then acidified to a final concentration of 5% (v/v) formic acid and stored at -80 °C prior to analysis.

**Liquid-chromatography-mass-spectrometry (LC-MS) analysis of isoTOP-ABPP samples.**

TEV digests were pressure loaded onto a 250 µm (inner diameter) fused silica capillary columns packed with C18 resin (Aqua 5 µm, Phenomenex). The samples were analyzed by multidimensional liquid chromatography tandem mass spectrometry (MudPIT), using an LTQ-Velos Orbitrap mass spectrometer (Thermo Scientific) coupled to an Agilent 1200-series quaternary pump. The peptides were eluted onto a biphasic column with a 5 µm tip (100 µm fused silica, packed with C18 (10 cm) and bulk strong cation exchange resin (3 cm, SCX, Phenomenex)) in a 5-step MudPIT experiment, using 0%, 30%, 60%, 90%, and 100% salt bumps of 500 mM aqueous ammonium acetate and using a gradient of 5–100% buffer B in buffer A (buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; buffer B: 20% water, 80% acetonitrile, 0.1% formic acid) as has been described<sup>1</sup>. Data was collected in data-dependent acquisition mode with dynamic exclusion enabled (20 s, repeat count of 2). One full MS (MS1) scan (400-1800 m/z) was followed by 30 MS2 scans (ITMS) of the nth most abundant ions.

**Peptide and protein identification.** The MS2 spectra were extracted from the raw file using RAW Xtractor (version 1.9.9.2; available at <http://fields.scripps.edu/downloads.php>). MS2 spectra were searched using the ProLuCID algorithm (publicly available at <http://fields.scripps.edu/downloads.php>)<sup>2</sup> using a reverse concatenated, nonredundant variant of the Human UniProt database (release-2012\_11). Cysteine residues were searched with a static modification for carboxyamidomethylation (+57.02146). For all competitive and reactivity profiling experiments, lysine residues were searched with up to one differential modification for either the light or heavy TEV tags (+464.2491 or +470.26331, respectively). Peptides were required to have at least one tryptic terminus and to contain the TEV modification. ProLuCID data was filtered through DTASelect (version 2.0) to achieve a peptide false-positive rate below 1%.

**Differential labeling analysis of residues labeled by probe 1.** For analysis of the residues labeled by probe 1, peptide and protein identification was conducted as detailed above with differential modification for either the light or heavy TEV tags (+464.2491 or +470.26331, respectively) allowed on lysine, arginine, aspartate, glutamate, histidine, serine, threonine, tyrosine, asparagine, glutamine and tryptophan. Cysteine was searched with a differential modification for either the light or heavy TEV tags (+413.24185 and +407.22764, respectively).

**R value calculation and processing.** The ratios of light and heavy MS1 peaks for each unique peptide were quantified with in-house CIMAGE software<sup>3,4</sup>, using default parameters (3 MS1 acquisitions per peak and signal to noise threshold set to 2.5). For reactivity measurements by isoTOP-ABPP, the *R* value was calculated from the ratio of MS1 peak areas, comparing the 1 mM STP alkyne sample (light TEV tag) with the 0.1 mM STP alkyne sample (heavy TEV tag). For competitive isoTOP-ABPP, the *R* value was calculated from the ratio of MS1 peak areas, comparing the DMSO treated sample (heavy TEV tag) with the compound treated sample (light TEV tag). For peptides that showed a  $\geq 95\%$  reduction in MS1 peak area in both reactivity and compound treated samples a maximal ratio of 20 was assigned. Ratios for unique peptide entries are calculated for each experiment; overlapping peptides with the same modified lysine (for example, different charge states, MudPIT chromatographic steps or tryptic termini) are grouped together and the median ratio is reported as the final ratio (*R*). The peptide ratios reported by CIMAGE were further filtered to ensure the removal or correction of low-quality ratios in each individual data set. The quality filters applied were the following: removal of half tryptic peptides; for ratios with high standard deviations from the median (90% of the median or above) the lowest ratio was taken instead of the median; removal of peptides with *R* = 20 and only a single MS2 event triggered during the elution of the parent ion; manual annotation of all the peptides with ratios of 20, removing any peptides with low quality elution profiles that remained after the previous curation steps (only done for competitive isoTOP-ABPP).

**Cross-data processing for fragment screening.** For compound treated samples, biological replicates of the same condition were averaged, if the standard deviation was below 60% of the mean; otherwise, for lysines with at least one *R* value  $<4$  for a particular compound, the lowest value of the ratio set was taken. For lysines, where all *R* values for a particular compound were  $\geq 4$ , the average was reported. For peptides containing several possible modified lysines, the lysine with the highest number of quantification events was used for analysis and the remaining, redundant peptides were reported as alternative modification sites. Peptides included in the aggregate dataset (those used for further bioinformatics and statistical analyses) were required to have been quantified in 2 experiments for competitive isoTOP-ABPP. Lysines were categorized as liganded, if they had at least one ratio *R*  $\geq 4$  (hit fragments). For liganded lysines with *R* = 20 for all liganding events, lysines were required to have been quantified with *R* = 20 in two separate experiments and were further required to have been quantified with *R*  $<20$  in one additional experiment.

**Cross-data processing for reactivity profiling.** For reactivity profiling, the median of biological replicates of the same condition and cell-line was calculated. For peptides containing several possible modified lysines, the lysine with the highest number of quantification events was used for analysis and the remaining, redundant peptides were reported as alternative modification sites. Peptides were required to be detected in at least one 1 mM vs 0.1 mM and one 0.1 mM vs 0.1 mM data set with the latter *R* value being smaller than 2.5. All ratios derived from soluble reactivity experiments were averaged. If the lysine was not detected in any soluble fraction, the *R* value from the membrane fraction was taken. Additionally, all membrane-only lysines with reactivity values were further required to have been detected in at least one 0.1 mM vs 0.1 mM membrane profiling experiment. If the final reactivity value was >10, it was set to 10. Lysines were categorized based on the *R* values (hyper-reactive: *R*<2; moderately-reactive: *R*=2-5; low-reactive: *R*>5).

**Heatmap generation.** Heat maps were generated in R (v.3.1.3) using the heatmap.2 algorithm.

**DrugBank.** Proteins were queried against the DrugBank database (v. 5.0.3 released on 2016-10-24; group "All") and separated into DrugBank and non-DrugBank proteins.

**Protein class analysis.** To place each human protein into a distinct protein class, custom python scripts were written to parse the KEGG BRITE and Gene Ontology databases. Top level terms from KEGG were placed into a list for each protein. Enzymes were given preference for cases with multiple terms, and term-lists without enzymes were reduced by giving preference to the least frequently occurring term across the entire dataset. Gene Ontology terms and hierarchies were obtained from Superfamily, and the hierarchy tree was traversed to find more general terms for each protein. A library was constructed to place each Gene Ontology term into a category (Transporter, Channel and Receptors; Enzymes; Gene Expression and Nucleic Acid Binding; Scaffolding, Modulators and Adaptors). If a protein had Gene Ontology terms in different categories, the abovementioned order of categories was used to prioritize the protein class. If no Gene Ontology term was available that could be assigned to a category, the protein was sorted into the category "Uncategorized". For the final protein class, the KEGG BRITE term was used, if available. If no KEGG BRITE term was available, the Gene Ontology term was used.

**Functional annotation of lysines.** Lysines proximal to functional sites were defined as any lysine with a C $\alpha$  atom within 10 Å of an annotated ligand binding site in an X-ray or NMR structure. Custom Python scripts were developed to collect relevant NMR and X-ray structures,

including any co-crystallized small molecules, from the RCSB Protein Data Bank (PDB)<sup>5</sup>. The following small molecules were excluded from this analysis: MES, EDO, DTT, BME, ACR, ACY, ACE and MPD. Histograms of the frequency of functional sites for hyper-reactive, moderately-reactive and low reactive lysines were calculated.

**Analysis of lysine conservation.** Sequences of all human proteins were downloaded from UniProtKB. Orthologs of human proteins were obtained using the HUGO Gene Name Consortium's database, or the DRSC Integrative Ortholog Prediction Tool, provided by Harvard Medical School. Clustal Omega was used to generate multiple sequence alignments for each human protein and its orthologs, and in-house software was used to calculate the conservation of individual lysines. Only proteins with orthologues in all five organisms evaluated (*M. musculus*, *X. laevis*, *D. melanogaster*, *C. elegans* and *D. rerio*) were considered for the conservation analysis.

**Analysis of lysine ubiquitylation and acetylation.** Custom python scripts were used to compile ubiquitylation and acetylation sites and the frequency of modification at each lysine for human, mouse and rat proteomes available from the PhosphoSitePlus®, [www.phosphosite.org](http://www.phosphosite.org) (release-060616)<sup>6</sup>. To be considered acetylated or ubiquitylated, lysines were required to be modified with the respective PTM with a frequency of 10 or greater detection events. The percentage of total lysines modified within each reactivity range (hyper-reactive: R<2; moderately-reactive: R=2-5; low-reactive: R>5) was calculated.

**Pocket analysis.** Proteins, for which crystallographic structures were available and labeled lysines were detected, were selected for the structural analysis. UniProt accession codes were used to filter the PDB<sup>5</sup>, selecting structures determined by X-ray crystallography (resolution 3.5 Å or better). Results were then filtered to select entries with the largest sequence coverage. The following proteins have been analyzed (PDB-ID in parentheses): O00299 (3o3t), O14737 (2k6b), P00367 (1l1f), P04179 (1pl4), P04181 (1gbn), P04632 (4phj), P07195 (1i0z), P07355 (1w7b), P07954 (3e04), P08133 (1m9i), P08237 (4omt), P08758 (2xo2), P09429 (2yrq), P11413 (1qli), P11766 (2fzw), P12268 (1nf7), P12956 (3rzx), P13804 (2a1u), P15121 (4lbs), P15311 (4rm8), P18669 (1yjx), P19367 (1cza), P19784 (3e3b), P20839 (1jcn), P23284 (3ici), P23368 (1pj3), P23381 (1r6t), P23919 (1nmy), P24941 (4ek4), P26038 (1e5w), P30040 (2qc7), P36551 (2aex), P39748 (1ul1), P42330 (1zq5), P49458 (4uyk), P50583 (4ick), P51580 (2bzg), P52292 (4wv6), P55145 (2w51), P55263 (4o1l), P58546 (3aaa), P60520 (4co7), P61081 (1y8x), P61978 (1zzk), P62258 (3ual), P62826 (4hat), P62937 (4n1m), P68036 (4q5e), P78417 (3vln), Q01469

(5hz5), Q01813 (4xyj), Q13011 (2vre), Q13630 (4e5y), Q14914 (2y05), Q16851 (4r7p), Q5VW32 (3zxp), Q6YN16 (3kvo), Q8WUM4 (2r05), Q92600 (4cru), Q96HE7 (3ahq), Q9BSH5 (3k1z), Q9GZQ8 (5d94), Q9NTK5 (2ohf), Q9NVS9 (1nrg), Q9UBT2 (5fq2), Q9Y2Q3 (1yzx), Q9Y696 (2d2z). Structural issues (i.e., missing atoms, non-standard residues) were fixed, and wild-type amino acids restored; biological units were built using the ProDy Python module<sup>7</sup>, and structures curated removing chemical entities other than standard amino acids or catalytic metals. Hydrogens were added using Reduce<sup>8</sup> using default 'build' options. Alternate conformations were removed, then AutoDock PDBQT files were generated following the standard protocol<sup>9</sup>. Pocket analysis was performed with AutoSite<sup>10</sup> using neighbor\_cutoff=16 for pocket clustering tolerance. For each pocket, lysines within 3.5 Å from any pocket volume points were considered adjacent.

**Sequence motifs.** For all lysines quantified in the reactivity profiling experiments, the flanking sequence ( $\pm$  8 amino acids) was determined with a custom python script, parsing the UniProtKB entries for all proteins identified. The sequences were binned by lysine reactivity (hyper-reactive: R<2; moderately-reactive: R=2-5; low-reactive: R>5) and evaluated for sequence motifs using WebLogo <<http://weblogo.berkeley.edu/logo.cgi>>. WebLogo was created by: Gavin E. Crooks, Gary Hon, John-Marc Chandonia and Steven E. Brenner, Computational Genomics Research Group, Department of Plant and Microbial Biology, University of California, Berkeley.

**Lysine reactivity and ligandability comparison.** Lysines found in both the reactivity and ligandability data sets were sorted on the basis of their reactivity values (lower ratio indicates higher reactivity). The moving average of the percentage of total liganded lysines within each reactivity bin (step-size 200) was taken.

**Subcloning and mutagenesis.** Unless noted below, genes were amplified from cDNA prepared from low passage HEK 293T cells using the Ribozol RNA extraction reagent (Amresco) and the iScript Reverse Transcription Supermix kit (Bio-Rad). For the following proteins cDNA clones were used for amplification instead: PFKP (5180268, Dharmacon), HK1 (BC008730, transomic), SIN3A (BC137098, transomic), G6PD (BC000337, transomic) and TGIF1 (BC031268, transomic). Mouse CARM1 in pFLAG-CMV-6c was a kind gift from the Mowen lab (TSRI). NUDT2 was obtained as synthesized gene (IDT). DNA was amplified with custom forward and reverse primers (see below) using phusion polymerase (NEB, M0530S), following the manufacturers' instructions, digested with the indicated restriction enzyme and ligated into pFLAG-CMV-6c or pRK5 with the appropriate affinity tag. Lysine mutants were

generated using QuikChange site-directed mutagenesis using Phusion® High-Fidelity DNA Polymerase and primers containing the desired mutations and their respective complements. The cloning of TTR and its K35A mutant has been described earlier<sup>11</sup>. TTR was expressed in *E. coli* and purified as described. For gel-based experiments 1 µM TTR was added into 1 mg ml<sup>-1</sup> soluble MDA-MB-231 lysate.

### **Primer for amplification.**

protein	fwd or rev	sequence	vector	restriction sites	affinity tag
AGFG1	fwd	TGGTGCCTCGACGATGGCGGCCAGCGCGAACGCG	pRK5	Sall/NotI	N-terminal Flag
AGFG1	rev	TACTTAGCGGCCGCCATAAGAAAGGATTGGTTGATGAGCTTCC	pRK5	Sall/NotI	N-terminal Flag
ATOX1	fwd	TGGTGCCTCGACGATGCCGAAGCACGAGTTCTCTG	pRK5	Sall/NotI	N-terminal Flag
ATOX1	rev	TACTTAGCGGCCGCCACTCAAGGCCAAGGTAGGAAACAGT	pRK5	Sall/NotI	N-terminal Flag
FAH	fwd	TGGTGCCTCGACGATGTCCTCATCCCGGTGGC	pRK5	Sall/NotI	N-terminal Flag
FAH	rev	TACTTAGCGGCCGCCCTCATGATGGCAGGAGAGCAGG	pRK5	Sall/NotI	N-terminal Flag
G6PD	fwd	TGGTGCCTCGACGATGGCAGAGCAGGTGGCCCT	pRK5	Sall/NotI	N-terminal Flag
G6PD	rev	TACTTAGCGGCCGCCCTCAGAGCTTGTTGGGGTTCAACC	pRK5	Sall/NotI	N-terminal Flag
GSTO1	fwd	TGGTGCCTCGACGATGTCCGGGGAGTCAGCCAG	pRK5	Sall/NotI	N-terminal Flag
GSTO1	rev	TACTTAGCGGCCGCCCTCAGAGCCCATAAGTCACAGGCC	pRK5	Sall/NotI	N-terminal Flag
HDHD3	fwd	TGGTGCCTCGACGATGGCACACCCGGCTGCAGAT	pRK5	Sall/NotI	C-terminal Flag
HDHD3	rev	TACTTAGCGGCCGCCCTCAAAGCCCTGGAGTTGAGCC	pRK5	Sall/NotI	C-terminal Flag
HK1	fwd	AAGCTTGGCCCGCATGATGCCCGCGCAGCTCCT	pFLAG-CMV-6c	NotI/KpnI	N-terminal Flag
HK1	rev	ATCGATGGTACCTTAGCTGCTTGCCTCTGTGCGTA	pFLAG-CMV-6c	NotI/KpnI	N-terminal Flag

NUDT2	fwd	TGGTGCCTGACGATGGCGCTCCGTGCATGCGG	pRK5	Sall/NotI	N-terminal Flag
NUDT2	rev	TACTTAGCGGCCGCTTACGCTTCATTGAGCACAGAAATT	pRK5	Sall/NotI	N-terminal Flag
PFKM	fwd	AGCTTGCGGCCGATGACCCATGAAGAGCACCATG	pFLAG-CMV-6c	NotI/BamHI	N-terminal Flag
PFKM	rev	CTGCAGGGATCCTAGACGGCAGCTCCCCGG	pFLAG-CMV-6c	NotI/BamHI	N-terminal Flag
PFKP	fwd	AGCTCAAAGCTTGCATGGACCGGGACGACTCCCG	pFLAG-CMV-6c	HindIII/BamHI	N-terminal Flag
PFKP	rev	CTGCAGGGATCCTCAGACACTCCAGGGCTGCAC	pFLAG-CMV-6c	HindIII/BamHI	N-terminal Flag
PMVK	fwd	TGGTGCCTGACGATGGCCCCGCTGGAGGGCG	pRK5	Sall/NotI	N-terminal Flag
PMVK	rev	TACTTAGCGGCCGCTAAAGTCTGGAGCGGATAAATTCTATCAGG	pRK5	Sall/NotI	N-terminal Flag
PNPO	fwd	TGGTGCCTGACGATGACGTGCTGGCTGGGG	pRK5	Sall/NotI	N-terminal Flag
PNPO	rev	TACTTAGCGGCCGCTTAAGGTGCAAGTCTCATAGAGCC	pRK5	Sall/NotI	N-terminal Flag
SIN3A full	fwd	CTGCAGGTGACGATGAAGCGGCGTTGGATGA	pRK5	Sall/NotI	N-terminal Flag
SIN3A full	rev	GTTAGCGCGCCGCTTAAGGGGCTTGAATACTGTGCC	pRK5	Sall/NotI	N-terminal Flag
SIN3A (1-400)	fwd	TGGTGCCTGACGATGAAGCGGCGTTGGATGA	pRK5	Sall/NotI	N-terminal Flag
SIN3A (1-400)	rev	TACTTAGCGGCCGCTTACACAGAACCTCTCAGCAGTTG	pRK5	Sall/NotI	N-terminal Flag
TGIF1	fwd	TGGTGCCTGACGATGGTTCTAGCGCAGAGCCG	pRK5	Sall/NotI	N-terminal Myc
TGIF1	rev	TACTTAGCGGCCGCTTAAGCTGTAAGTTGCCTGAAGCTCC	pRK5	Sall/NotI	N-terminal Myc
TGIF2	fwd	TGGTGCCTGACGATGTCGGACAGTGATCTAGGTGAGG	pRK5	Sall/NotI	N-terminal Myc

TGIF2	rev	TACTTAGCGGCCGCTTACTGGGGATTTCAGAGACTAAAGG G	pRK5	Sall/NotI	N-terminal Myc
-------	-----	---	------	-----------	----------------

**Primer for site-directed mutagenesis.**

protein	mutation	fwd or rev	sequence
ADK	K88R	fwd	GGTGGCTCTACCCAGAATTCAATTGGGTGGCTCAGTGGATGATTG
ADK	K88R	rev	GAATCATCCACTGAGCCACCCGAATTGAATTCTGGTAGAGCCACC
AGFG1	K134R	fwd	CGCCAGAACAGCCCGCGTCGTGGCATCAGTTC
AGFG1	K134R	rev	GAACTGATGCCACGACGGGGCTTGTCTGGCG
ATOX1	K60R	fwd	GCAACCCTGAAGAAAACAGGACGCCTGTTCCCTACCTGGCCTTG
ATOX1	K60R	rev	CAAGGCCAAGGTAGGAAACAGTGCCTGTTCTCAGGGTTGC
CARM1	K241R	fwd	CGTGGTCATCCCTGGCCCGTAGAGGAGGTCTCATTG
CARM1	K241R	rev	CAATGAGACCTCCTCTACGCGGCCAGGGATGACCACG
FAH	K241R	fwd	GGAGTGCACGAGACATTAGCGCTGGAGTATGCCCTCTC
FAH	K241R	rev	GAGAGGGACATACTCCCAGCGCTGAATGTCTCGCACTCC
G6PD	K171R	fwd	GCATCATCGTGGAGCGCCCCCTCGGGAGGG
G6PD	K171R	rev	CCCTCCGAAGGGCGCTCCACGATGATGC
GSTO1	K57R	fwd	GTCATCAATACACCTGCGCAATAAGCCTGAGTGGTTC
GSTO1	K57R	rev	GAACCACTCAGGCTTATTGCGCAGGTTGATATTGATGAC
HDHD3	K15R	fwd	CTGACGTGGATGTGCGCGACACGCTGCTCAG
HDHD3	K15R	rev	CTGAGCAGCGTGTGCGCACATCCCACGTCAG
HK1	K510R	fwd	CACAACAATGCCGTGGTCGCATGCTGCCCTCCTTC
HK1	K510R	rev	GAAGGAGGGCAGCATGCGAACCAACGGCATTGTTGTG

NUDT2	K89R	fwd	GTGGCCCGCAACAAACCCGTACGGTGATCTATTGGCTG
NUDT2	K89R	rev	CAGCCAATAGATCACCGTACGGGTTGTCGGGCCAC
NUDT2	K42R	fwd	GGACGCCGCCGCCGGGGCATGTGGAAC
NUDT2	K42R	rev	GTTCCACATGCCCGCGCGCGTCC
PFKM	K678R	fwd	GATAGGAATTTGCCACTCGGATGGCGCCAAGGCTATGA
PFKM	K678R	rev	TCATAGCCTGGCGCCCATCCGAGTGGCAAATTCCATC
PFKP	K688R	fwd	CCATTGATAGAAACTTGAACCCGTATCTGCCAGAGCTATGGAGTGG
PFKP	K688R	rev	CCACTCCATAGCTCTGGCAGAGATACTGGTTCAAAGTTCTATCAAATGG
PMVK	K48R	fwd	GGCTCTCTGGTCCACTCCCGAGCAGTATGCTCAGGAGC
PMVK	K48R	rev	GCTCCTGAGCATACTGCTCGGGAGTGGACCAGAGAGCC
PNPO	K100R	fwd	CGCATGTTGCTGCTGCAGGGCTTCGGAAAGATG
PNPO	K100R	rev	CATCTTCCCGAAGCCCCGCAGCAGCAACATGCG
SIN3A	K155R	fwd	CCTTGACATCATGAAGGAATTCTCGTTCTAGAGCATCGACACCCC
SIN3A	K155R	rev	GGGGTGTGATGCTCTGAGAACGAAATTCTTCATGATGTCAAGG
TPMT	K32R	fwd	GGAAGAATGGCAAGACCGCTGGGTAAACGGCAAGACTG
TPMT	K32R	rev	CAGTCTGCCGTTACCCAGCGGTCTGCCATTCTCC

### Synthesized sequence for NUDT2.

AGCTCATATGGCGCTCCGTGCATGCAGGGCTTATCATTTCGTCGTTAATCCCGAAA  
 GTTGACAACAACCGCGATTGAATTCTCCTCTGCAGGCATCGGATGGCATTACCACTGGA  
 CGCCGCCGAAAGGGCATGTGGAACCAGCGAAGATGATCTGGAGACCGCACTGCGCGAA  
 ACCCAGGAAGAGGCAGGCATTGAGGCCGCCAGTTGACGATCATCGAAGGGTTAACGCG  
 GAGCTGAACTATGTGGCCCGCAACAAACCCAAAACGGTGATCTATTGGCTGGCCGAAGTC  
 AAAGATTATGACGTAGAGATTGCCTGAGCCACGAACATCAGGCGTATCGCTGGTTGGGG

CTGGAGGAAGCGTGCCAGTTGGCGCAATTAGGAAATGAAAGCTGCTCTTCAGGAGGGG  
CATCAATTCTGTGCTCAATTGAAGCGTAAGGATCCTTCAG

**Recombinant expression of proteins by transient transfection.** HEK 293T cells were grown to 50 % confluence in 10 ml DMEM supplemented with 10% fetal bovine serum (FBS), penicillin, streptomycin and glutamine in 10 cm tissue culture dishes. 3 µg of DNA was diluted in 500 µL DMEM and 30 µL of PEI (MW 40,000, 1 mg ml<sup>-1</sup>, Polysciences) were added. The mixture was incubated at room temperature for 30 min and added dropwise to the cells. Cells were grown for 48 h at 37 °C with 5% CO<sub>2</sub>. Cells were washed with cold PBS, scraped with cold PBS and cell pellets were isolated by centrifugation (1,400g, 3 min, 4 °C), and stored at -80 °C until use. Cell pellets were resuspended in PBS, lysed by sonication and fractionated (100,000g, 45 min) to yield soluble and membrane fractions. The soluble fraction was adjusted to a final protein concentration of 1 mg ml<sup>-1</sup> for gel-based ABPP experiments.

**Assessment of the reactivity of alkyne-containing ester probes.** 50 µL of soluble MDA-MB-231 proteome (1.8 mg ml<sup>-1</sup>) were treated with 100 µM of the indicated probe (**1-15**) for 1 h at room temperature. Copper-mediated azide-alkyne cycloaddition (CuAAC) was performed with 25 µM rhodamine-azide (50× stock in DMSO), tris(2-carboxyethyl)phosphine hydrochloride (TCEP; fresh 50× stock in water, final concentration = 1 mM), TBTA ligand (17× stock in DMSO:t-butanol 1:4, final concentration = 100 µM) and CuSO<sub>4</sub> (50× stock in water, final concentration = 1 mM). Samples were allowed to react for 1 h at ambient temperature. The reactions were quenched with 20 µl of 4× SDS-PAGE loading buffer and the quenched samples analyzed by SDS-PAGE (10%, 14% or 16% polyacrylamide; 20 µl of sample/lane) and visualized by in-gel fluorescence using a flatbed fluorescent scanner (BioRad ChemiDoc™ MP).

**Compound stability assays.** In triplicate, DMF solutions of the indicated compounds (**1** (10 mM), **21** (10 mM) and **31** (1 mM)) were diluted 1:10 in PBS buffer (pH 7.4) and incubated at ambient temperature. At the indicated times, 20 µL aliquots of the solvolysis reactions were removed and analyzed by analytical High Performance Liquid Chromatography (HPLC) on an Agilent 1200 Series HPLC system with an in line variable UV absorbance detector. Compound **1** was evaluated on a Phenomenex Luna C18 column (150 mm x 4.6 mm, 4 µm), with a flow rate of 1 ml min<sup>-1</sup>, and with a gradient beginning at 100% A (Water/Acetonitrile 95:5 (v/v) with 0.1% TFA (v/v)) for two minutes and then rising in a linear gradient to 10% B (Acetonitrile/Water 95:5 (v/v) with 0.1% TFA(v/v)) from two to five minutes and then continuing to 70% B at ten minutes. The solvent composition was then increased in a linear gradient to 100% B over three minutes,

held at 100% B for one minute and then returned to 100% A for a further two minutes. Compound **1** eluted at 10.5 minutes. Compounds **21** and **31** were evaluated on a Phenomenex Gemini C18 column (50 mm x 4.6 mm, 5  $\mu$ m), with a flow rate of 1 ml min<sup>-1</sup>, and with a gradient beginning at 100% A (Water/Acetonitrile 95:5 (v/v) with 0.1% TFA (v/v)) for two minutes and then rising in a linear gradient to 50% B (Acetonitrile /Water 95:5 (v/v) with 0.1% TFA(v/v)) from two to 5 minutes and then continuing to 100% B at thirteen minutes. The solvent composition was then held at 100% B for two minutes and then returned to 100% A for a further two minutes. Compound **21** eluted at 11.1 minutes and compound **31** eluted at 10.05 minutes. The decrease in peak area (254 nm) for the indicated compounds was monitored over time and was used to calculate the rates of solvolysis of each compound, fitted to the standard equation of non-linear one phase decay in GraphPad Prism 7.03.

**Direct labeling of recombinantly expressed proteins by gel-based ABPP.** 50  $\mu$ L of soluble HEK 293T proteome (1 mg ml<sup>-1</sup>) expressing the respective protein (WT or KR mutant) or transfected with an empty vector were treated with 10  $\mu$ M of the indicated probe for 1 h at room temperature. The samples were analyzed as described in the previous section. For quantification of relative labeling of the different protein variants, the intensity of labeling was determined by quantifying the integrated optical intensity of the bands using ImageLab 5.2.1 software (BioRad).

**Competitive gel-based ABPP and apparent IC<sub>50</sub> values.** 50  $\mu$ l of soluble proteome (1 mg ml<sup>-1</sup>) expressing the indicated protein were treated with fragment electrophiles (1  $\mu$ l of 50× stock solution in DMSO) at ambient temperature for 1 h. The indicated probe (fluorophore or alkyne-containing, 1  $\mu$ l of a 500  $\mu$ M solution, final concentration = 10  $\mu$ M) was then added and allowed to react for an additional 1 h. CuAAC and in-gel fluorescence analysis were performed as described above. For quantification of inhibition and apparent IC<sub>50</sub> determination, the percentage of labeling was determined by quantifying the integrated optical intensity of the bands using ImageLab 5.2.1 software (BioRad). Nonlinear regression analysis was used to determine the IC<sub>50</sub> values from a dose-response curve generated using GraphPad Prism 7.

**PFKP functional assay<sup>12</sup>.** For inhibitor experiments, 50  $\mu$ l of soluble proteome (initial total protein concentration: 1 mg ml<sup>-1</sup>) from HEK 293T cells expressing PFKP (WT or K688R mutant) or mock transfected cells (empty vector; negative control) were incubated with 1  $\mu$ l 50× of the compound in DMSO or DMSO for the positive or negative control for 1 h at room temperature. Lysates were diluted 40× with dilution buffer (PBS containing 0.2 mg ml<sup>-1</sup> BSA and 5 mM MgCl<sub>2</sub>)

and 40  $\mu$ l were added into a clear bottom 384 well plate. 10  $\mu$ l of a mixture of 3.5  $\mu$ l PBS, 2.5  $\mu$ l fructose-6-phosphate (100 mM), 1  $\mu$ l NADH (20 mM), 1  $\mu$ l ATP (50 mM), 1  $\mu$ l aldolase (50 U ml<sup>-1</sup>) and 1  $\mu$ l GDH/TPI (500 U ml<sup>-1</sup> TPI, 50 U ml<sup>-1</sup> GDH) were added to start the reaction. The absorbance of NADH was measured at 340 nm every minute for 30 min.

**PNPO functional assay.** 80  $\mu$ l of soluble proteome (total protein concentration: 1 mg ml<sup>-1</sup>) from HEK 293T cells expressing PNPO (WT or K100R mutant) or mock transfected cells (empty vector; negative control) were added into a clear bottom 384 well plate. For compound treatments, 1  $\mu$ l of the inhibitor (80 $\times$  solution in DMSO) or 1  $\mu$ l of DMSO (positive control) were added and the reactions were incubated for 1 h at room temperature. 10  $\mu$ l of 0.1 M Tris in PBS were added and the reaction was started by addition of 10  $\mu$ l 5 mM pyridoxine phosphate (PNP) in water (PNP was prepared as reported earlier<sup>13</sup>). The absorbance of the Schiff Base between pyridoxal phosphate and Tris was measured at 388 nm every minute for 30 min.

**G6PD functional assay.** Soluble proteome (initial total protein concentration: 1 mg ml<sup>-1</sup>) from HEK 293T cells expressing G6PD (WT or K171R mutant) or mock transfected cells (empty vector; negative control) were diluted 1000 $\times$  with dilution buffer. 88  $\mu$ l of this were added into a clear bottom 384 well plate. 12  $\mu$ l of a mixture of 8  $\mu$ l water, 2  $\mu$ l 60 mM glucose-6-phosphate and 2  $\mu$ l 20 mM NADP were added to start the reaction. The absorbance of NADPH was measured at 340 nm every minute for 30 min.

**NUDT2 functional assay.** NUDT2 activity was measured with a published assay using a fluorogenic substrate<sup>14</sup>. For inhibitor experiments, 50  $\mu$ l of soluble proteome (initial total protein concentration: 1 mg ml<sup>-1</sup>) from HEK 293T cells expressing NUDT2 (WT or K89R mutant) or mock transfected cells (empty vector; negative control) were incubated with 1  $\mu$ l 50 $\times$  of the compound in DMSO or DMSO for the positive or negative control (lysate transfected with empty vector) for 1 h at room temperature. Lysates were diluted 4000 $\times$  with dilution buffer and 64  $\mu$ l were added into a black 96 well, half volume plate. 16  $\mu$ l of fluorogenic substrate (5  $\mu$ M) were added to start the reaction. The fluorescence intensity with excitation at 530 nm and emission at 563 nm was measured every minute for 30 min.

**Calculation of relative activity or percent inhibition.** For PNPO, PFKP, NUDT2 and G6PD, the slope of the linear regression of the linear portion of the absorbance or fluorescence over time was used as measure their activity. Apparent activity was calculated relative to the WT. Percent inhibition was calculated relative to the positive and negative control and used to

calculate IC<sub>50</sub> values by nonlinear regression analysis from a dose-response curve generated using GraphPad Prism 7.

**Calculation of k<sub>obs</sub>/[I] for NUDT2.** 50 µl of soluble proteome from HEK 293T cells (initial total protein concentration: 1 mg ml<sup>-1</sup>) expressing NUDT2 WT were pre-incubated with indicated concentrations of **49** (1 µl of a 50× stock) for indicated times. At the specific time, the reaction mixture was diluted 4000× with dilution buffer and the activity of NUDT2 was determined as detailed above. Relative activity of NUDT2 at the indicated time-point of pre-incubation with the indicated concentration of **49** was calculated relative to a DMSO control as described above. The natural logarithm of the relative activity was calculated and plotted against the pre-incubation time. k<sub>obs</sub> was calculated for each concentration of **49** as the slope of this curve. k<sub>obs</sub> was plotted against the inhibitor concentration and k<sub>obs</sub>/[I] was calculated as the slope of this linear curve.

**Site of labeling of recombinantly expressed proteins by reductive dimethylation (ReDiMe).** 500 µl of soluble proteome from HEK 293T cells expressing the indicated proteins (1 mg ml<sup>-1</sup> total protein concentration; see recombinant expression of proteins by transient transfection for additional details) were treated with the indicated compound at 50 µM (5 µl of 5 mM stock in DMSO) or DMSO for 1 h at ambient temperature. For each sample, 20 µl anti-FLAG® M1 Agarose Affinity Gel (Sigma, A4596) slurry was washed once by centrifugation with 500 µl 0.1 M glycine pH 3.5 and three times with 500 µl PBS (8,000g, 3 min). The compound- and DMSO-treated reactions were separately enriched on anti-FLAG resin for 4 h at 4 °C while rotating. The beads were collected by centrifugation (8,000g, 3 min) and washed three times with PBS. The beads were resuspended in 80 µl 6 M Urea in TEAB (pH 8.0, 100 mM) and rotated at room temperature for 30 min to elute the captured proteins. After separation of the beads, 10 mM DTT (4 µl of 200 mM) were added and the reaction was incubated at 65 °C for 15 minutes following which 20 mM iodoacetamide (4 µl of 400 mM) was added and the reaction incubated for 30 minutes at 37 °C. The samples were then diluted with TEAB (232 µl) and to this was added the appropriate restriction enzyme (trypsin (10 µl, 5 µg total) for HDHD3, HK1, SIN3A and XRCC6 or rLysC (10 µl, 5 µg total, Promega, V1671) for PNPO and PFKP) and the samples were allowed to digest over night at 37 °C with shaking. Reductive dimethylation was performed as previously described<sup>15</sup>. Briefly, DMSO-treated samples were labeled with heavy-formaldehyde (<sup>13</sup>C,D<sub>2</sub>-) and compound-treated samples with light formaldehyde (<sup>12</sup>C,H<sub>2</sub>) (0.15% formaldehyde) and sodium cyanoborohydride (22.2 mM). After 1 h at ambient temperature with shaking, the reactions were quenched by addition of NH<sub>4</sub>OH (2.3%) for 10 min followed by

acidification with formic acid (5%). The samples were then combined and analyzed by LC/MS analysis. The MS2 spectra data were extracted from the raw file using RAW Xtractor (version 1.9.9.2; available at <http://fields.scripps.edu/downloads.php>). MS2 spectra data were searched using the ProLuCID algorithm (publicly available at <http://fields.scripps.edu/downloads.php>) using a reverse concatenated, nonredundant variant of the Human UniProt database (release-2012\_11). Cysteine residues were searched with a static modification for carboxyamidomethylation (+57.02146 C). Searches also included methionine oxidation as a differential modification (+15.9949 M). Peptides were searched with a static modification for dimethylation of lysine residues (+28.0313 K) and the N-terminus (+28.0313 N-term) and for ReDiMe labeled amino acids (+6.03181 K, +6.03181 N-term). Peptides were also searched with a differential modification on lysine to detect the directly labeled peptide-compound adducts (+246.07931 for **19**, +194.05791 for **33**, +166.04186 for **20**, +211.96968 for **21** and +143.03711 for **32**). Peptides were required to have at least one cognate proteolytic terminus and unlimited missed cleavage sites. ProLuCID data was filtered through DTASelect (version 2.0) to achieve a peptide false-positive rate below 1%. Ratios of heavy/light (DMSO/test compound) peaks were calculated using in-house CIMAGE software<sup>3,4</sup>. Unmodified peptides were included in the final analysis, if they stemmed from the expressed protein, contained cognate cleavage sites on both ends, contained no internal missed cleavage sites and had at least one lysine as the cleavage site.

**ABPP-SILAC IP experiment for SIN3A interacting proteins.** All SILAC experiments were performed using the isotopically labeled human HEK 293T cell line generated by 8 passages in either light (100 µg ml<sup>-1</sup> each of L-arginine and L-lysine) or heavy (100 µg ml<sup>-1</sup> each of [<sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>4</sub>]L-arginine and [<sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>2</sub>]L-lysine) SILAC DMEM media (Thermo Scientific) supplemented with 10% dialyzed fetal calf serum, penicillin, streptomycin and glutamine. 2×10<sup>-5</sup> SILAC HEK 293T cells were plated in 6 cm dishes in either heavy or light labeled SILAC media. Cells were transfected the next day with 1 µg of FLAG-GFP, or FLAG-SIN3A wild type, K155R, or K155W constructs as indicated. After 48 hours, cells were rinsed with ice-cold PBS and suspended in cold IP-lysis buffer (0.5% Chaps, 50 mM Hepes pH 7.4, 150 mM NaCl, and EDTA-free protease inhibitors and phosphatase inhibitors (Roche)) by gentle sonication. Samples were rotated for 30 minutes at 4 °C to complete lysis. For compound treatment experiments, 50 µM (final concentration) of **21** was added to samples prior to rotation. Samples were clarified by centrifugation for 1 minute at 16,000 rpm, and protein concentration was measured using the DC Protein Assay kit (Bio-Rad). Samples were normalized to 2 mg/mL by

addition of cold IP-lysis buffer. 25  $\mu$ L of anti-FLAG-M2 beads was added to the clarified supernatant and incubated for 3 h while rotating at 4°C. Beads were washed three times with cold PBS, and then eluted with 40  $\mu$ L of 8 M urea for 10 min at 65 °C. Samples were combined and then reduced by addition of 12.5 mM DTT at 65 °C for 15 minutes. Samples were alkylated with 25 mM iodoacetamide at 37 °C for 15 minutes, then diluted to 2 M urea with PBS. Sequence grade trypsin (Promega) was reconstituted in trypsin buffer with CaCl<sub>2</sub>, as detailed above, and 2  $\mu$ g of trypsin was added to each samples. Samples were shaken at 37 °C overnight after which digests were acidified with formic acid to a final concentration of 5% (v/v). Samples were stored at -80°C until analysis by LC-MS. LC-MS spectra were collected and analyzed as described above with the following modifications. Cysteine residues were searched with a static modification for carboxyamidomethylation (+57.02146 C). Searches also included methionine oxidation as a differential modification (+15.9949 M) and mass shifts of SILAC labeled amino acids (+10.0083 R, +8.0142 K) and no enzyme specificity. Peptides were required to have at least one tryptic terminus and unlimited missed cleavage sites. 2 peptide identifications were required for each protein. *R* values for co-immunoprecipitation are presented as the median ratio of heavy/light peptides for all biological replicates. A list of all proteins enriched preferentially by SIN3A was generated from a comparison of SIN3A wild type vs GFP immunoprecipitations, including all proteins with at least two distinct quantified peptide sequences and a median ratio greater than or equal to 5 ( $R \geq 5$ ). For the wild type vs mutant or compound treatment experiments, proteins were considered for analysis, if they had been preferentially enriched in the SIN3A vs GFP experiments ( $R \geq 5$ ). Furthermore, if there were at least two quantified unique peptides, the median ratio of each protein's unique peptides (not occurring in any other human protein) were reported.

**Co-IP experiment for the interaction between SIN3A and TGIF1 and TGIF2.** 6 cm dishes of HEK 293T cells were transfected at 40% confluence with 600 ng of FLAG-GFP, FLAG-SIN3A WT, K155W, or K155R construct, and 600 ng of MYC-TGIF1 or MYC-TGIF2 as indicated. After 48 hours, cells were lysed and enriched as described above. Following elution in 40  $\mu$ L urea, 15  $\mu$ L of loading buffer was added to samples. 15  $\mu$ L of both input (10 %) and outputs were loaded onto an SDS-PAGE gel.

**Western blotting.** Proteins were resolved by SDS-PAGE (3 h, 300 V) and transferred to nitrocellulose membranes (90 min, 60 V), blocked with 5% milk in TBS-T and probed with the indicated antibodies in 5% milk in TBS-T. The primary antibodies and the dilutions used are as follows: anti-Flag (Sigma Aldrich, F1804, 1:3,000), anti-Myc (Cell Signalling, 2272S, 1:5,000),

anti-actin (Cell Signaling, 3700, 1:3,000) and anti-GAPDH (Santa Cruz, 32233, 1:10,000). Blots were incubated with primary antibodies overnight at 4 °C with rocking and were then washed (3 × 5 min, TBS-T) and incubated with secondary antibodies (LICOR, IRDye 800CW or IRDye 680LT, 1:10,000) for 1 h at ambient temperature. Blots were further washed (3 × 5 min, TBST) and visualized on a LICOR Odyssey Scanner. Relative band intensities were quantified using ImageJ software.

**Statistical analysis.** The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. No statistical methods were used to predetermine sample size. Data are shown as mean ± standard deviation of at least two experiments. Statistical significance was calculated with unpaired students t-tests; \*, p < 0.05, \*\*, p < 0.01, \*\*\*, p < 0.001, \*\*\*\*, p < 0.0001.

#### (D) Synthetic methods

Chemicals and reagents were purchased from a variety of vendors, including Sigma Aldrich, Acros, Fisher, Fluka, Santa Cruz, CombiBlocks, BioBlocks, and Matrix Scientific, and were used without further purification, unless noted otherwise. Anhydrous solvents were obtained as commercially available pre-dried, oxygen-free formulations. Flash chromatography was carried out using 230–400 mesh silica gel. Preparative thin layer chromatography (PTLC) was carried out using glass backed PTLC plates 500-2000 µm thickness (Analtech). All reactions were monitored by thin layer chromatography carried out on 0.25 mm E. Merck silica gel plates (60F-254) and visualized with UV light, or by ninhydrin, ethanolic phosphomolybdic acid, iodine, p-anisaldehyde or potassium permanganate stain. NMR spectra were recorded on Varian INOVA-400, Bruker DRX-600 or Bruker DRX-500 spectrometers in the indicated solvent. Multiplicities are reported with the following abbreviations: s singlet; d doublet; t triplet; q quartet; p pentet; m multiplet; br broad. Chemical shifts are reported in ppm relative to the residual solvent peak and J values are reported in Hz. Mass spectrometry data were collected on an Agilent ESI-TOF instrument (HRMS-ESI) or an Agilent 6520 Accurate-Mass Q-TOF (HRMS).

The following molecules were purchased from commercial vendors: **1** (Lumiprobe, 40720), **16** (ThermoFisher Scientific, 46410), **17** (ThermoFisher Scientific, A37570), **18** (ThermoFisher Scientific, B10006), **50** (Sigma-Aldrich, 439428) and **51** (Sigma-Aldrich, 559997).

**General Procedure A.** 1.23 mmol of the carboxylic acid (1.5 eq.) and 0.82 mmol of the phenol (1.0 eq) or N-hydroxysuccinimide were dissolved in 5 ml DCM and 340 µl triethylamine (247 mg, 2.44 mmol, 3.0 eq.) were added. 418 mg 2-chloro-1-methylpyridinium iodide (1.64 mmol, 2.0 eq.) were added. The mixture was stirred over night at room temperature and directly loaded onto a preparative TLC. The TLC was run with the indicated solvent and the product was eluted from the silica. Evaporation of the solvent resulted in the desired ester.

**General Procedure B.** 0.82 mmol of the phenol or N-hydroxysuccinimide (1.0 eq.) were dissolved in 5 ml DCM and 340 µl triethylamine (247 mg, 2.44 mmol, 3.0 eq.) were added. To this 1.23 mmol of the carbonyl chloride were added and the mixture was stirred for 4 h at room temperature. The reaction was directly loaded onto a preparative TLC. The TLC was run with the indicated solvent and the product was eluted from the silica. Evaporation of the solvent resulted in the desired ester.

**4-Nitrophenyl 4-pentyoate (2).** This compound was synthesized according to General Procedure A starting from 4-pentyoic acid and 4-nitrophenol. The preparative TLC was run with n-hexane/DCM 1:3. 70 mg (39 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.28 (d,  $J$  = 8.7 Hz, 2H), 7.30 (d,  $J$  = 8.7 Hz, 2H), 2.86 (t,  $J$  = 7.3 Hz, 2H), 2.64 (t,  $J$  = 7.3 Hz, 2H), 2.07 – 2.04 (m, 1H); HRMS (m/z) calculated for  $\text{C}_{11}\text{H}_{10}\text{NO}_4$  [M+H]: 220.0604; found: 220.0602.

**2-Nitrophenyl 4-pentyoate (3).** This compound was synthesized according to General Procedure A starting from 4-pentyoic acid and 2-nitrophenol. The preparative TLC was run with n-hexane/DCM 1:3. 97 mg (54 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.12 (d,  $J$  = 8.3 Hz, 1H), 7.67 (t,  $J$  = 7.9 Hz, 1H), 7.42 (t,  $J$  = 8.0 Hz, 1H), 7.27 (d,  $J$  = 5.5 Hz, 1H), 2.92 (t,  $J$  = 7.3 Hz, 2H), 2.66 (d,  $J$  = 7.3 Hz, 2H), 2.08 – 2.03 (m, 1H); HRMS (m/z) calculated for  $\text{C}_{11}\text{H}_9\text{NNaO}_4$  [M+Na]: 242.0424; found: 242.0424.

**2,4-Dinitrophenyl 4-pentyoate (4).** This compound was synthesized according to General Procedure A starting from 4-pentyoic acid and 2,4-dinitrophenol. The preparative TLC was run with n-hexane/DCM 1:3. 192 mg (89 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.98 (d,  $J$  = 2.6 Hz, 1H), 8.53 (dd,  $J$  = 2.6 Hz,  $J$  = 8.9 Hz, 1H), 7.51 (d,  $J$  = 8.9 Hz, 1H), 2.96 (t,  $J$  = 7.3 Hz, 2H), 2.67 (dt,  $J$  = 2.6 Hz,  $J$  = 7.3 Hz, 2H), 2.07 (t,  $J$  = 2.6 Hz, 1H); HRMS (m/z) calculated for  $\text{C}_{11}\text{H}_9\text{N}_2\text{O}_6$  [M+H]: 265.0455; found: 265.0453.

**2,3,5,6-Tetrafluorophenyl 4-pentyoate (5).** This compound was synthesized according to General Procedure A starting from 4-pentyoic acid and 2,3,5,6-tetrafluorophenol. The preparative TLC was run with n-hexane/DCM 1:1. 185 mg (92 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.06 – 6.95 (m, 1H), 2.94 (t,  $J$  = 7.3 Hz, 2H), 2.66 (d,  $J$  = 7.3 Hz, 2H), 2.07 – 2.04 (m, 1H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -139.20 (dd,  $J$  = 12.3 Hz,  $J$  = 9.6 Hz, 2F), -153.07 (dd,  $J$  = 12.3 Hz,  $J$  = 9.6 Hz, 2F); HRMS (m/z) calculated for  $\text{C}_{11}\text{H}_7\text{F}_4\text{O}_2$  [M+H]: 247.0377; found: 247.0380.

**Pentafluorophenyl 4-pentyoate (6).** This compound was synthesized according to General Procedure A starting from 4-pentyoic acid and pentafluorophenol. The preparative TLC was run with n-hexane/DCM 1:1. 140 mg (65 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  2.93 (t,  $J$  = 7.3 Hz, 2H), 2.69 – 2.59 (m, 2H), 2.09 – 2.03 (m, 1H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -152.72 – -152.85 (m, 2F), -158.02 (t,  $J$  = 21.7 Hz, 1F), -162.39 – -162.60 (m, 2F); HRMS (m/z) calculated for  $\text{C}_{11}\text{H}_6\text{F}_5\text{O}_2$  [M+H]: 265.0283; found: 265.0280.

**4-Trifluoromethyl-2,3,5,6-tetrafluorophenyl 4-pentyoate (7).** This compound was synthesized according to General Procedure A starting from 4-pentyoic acid and 4-trifluoromethyl-2,3,5,6-tetrafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 168 mg (65 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  2.96 (t,  $J$  = 7.2 Hz, 2H), 2.66 (d,  $J$  = 7.2 Hz, 2H), 2.08 – 2.04 (m, 1H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -56.4 (t,  $J$  = 26.8 Hz, 3F), -140.43 – -140.76 (m, 2F), -150.35 – -150.50 (m, 2F); HRMS (m/z) calculated for  $\text{C}_{12}\text{H}_6\text{F}_7\text{O}_2$  [M+H]: 315.0251; found: 315.0252.

**4-Pentyoic acid NHS ester (8).** This compound was synthesized according to General Procedure A starting from 4-pentyoic acid and N-hydroxysuccinimide. The preparative TLC was run with DCM/ethyl acetate 4:1. 93 mg (58 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  2.88 (t,  $J$  = 2.88 Hz, 2H), 2.84 (s, 4H), 2.65 – 2.58 (m, 2H), 2.07 – 2.03 (m, 1H); HRMS (m/z) calculated for  $\text{C}_9\text{H}_{10}\text{NO}_4$  [M+H]: 196.0604; found: 196.0598.

**4-Nitrophenyl 4-ethynylbenzoate (9).** This compound was synthesized according to General Procedure A starting from 4-ethynylbenzoic acid and 4-nitrophenol. The preparative TLC was run with n-hexane/DCM 1:3. 74 mg (34 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.36 – 8.31 (m, 2H), 8.18 – 8.14 (m, 2H), 7.67 – 7.62 (m, 2H), 7.45 – 7.40 (m, 2H), 3.31 (s, 1H);  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ):  $\delta$  163.73, 155.68, 145.66, 132.58, 130.33, 128.59, 128.34, 125.47, 122.72, 82.61, 81.27; HRMS (m/z) calculated for  $\text{C}_{15}\text{H}_{10}\text{NO}_4$  [M+H]: 268.0604; found: 268.0605.

**2-Nitrophenyl 4-ethynylbenzoate (10).** This compound was synthesized according to General Procedure A starting from 4-ethynylbenzoic acid and 2-nitrophenol. The preparative TLC was run with n-hexane/DCM 1:3. 53 mg (24 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.20 – 8.09 (m, 3H), 7.71 (dt,  $J$  = 7.8, 1.2 Hz, 1H), 7.66 – 7.61 (m, 2H), 7.48 – 7.42 (m, 1H), 7.39 (dd,  $J$  = 8.2, 1.2 Hz, 1H), 3.30 (s, 1H);  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ):  $\delta$  163.87, 144.32, 141.99, 134.89, 132.49, 130.50, 128.50, 128.19, 126.92, 126.05, 125.45, 82.72, 81.06; HRMS (m/z) calculated for  $\text{C}_{15}\text{H}_{10}\text{NO}_4$  [M+H]: 268.0604; found: 268.0602.

**2,4-Dinitrophenyl 4-ethynylbenzoate (11).** This compound was synthesized according to General Procedure A starting from 4-ethynylbenzoic acid and 2,4-dinitrophenol. The preparative TLC was run with n-hexane/DCM 1:3. 151 mg (55 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.02 (s, 1H), 8.58 (d,  $J$  = 9.0 Hz, 1H), 8.15 (d,  $J$  = 8.1 Hz, 2H), 7.69 – 7.62 (m, 3H), 3.33 (s, 1H); HRMS (m/z) calculated for  $\text{C}_{15}\text{H}_9\text{N}_2\text{O}_6$  [M+H]: 313.0455; found: 313.0446.

**2,3,5,6-Tetrafluorophenyl 4-ethynylbenzoate (12).** This compound was synthesized according to General Procedure A starting from 4-ethynyl benzoic acid and 2,3,5,6-tetrafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 158 mg (66 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.19 – 8.15 (m, 2H), 7.67 – 7.62 (m, 2H), 7.06 (tt,  $J$  = 9.9 Hz,  $J$  = 7.1 Hz, 1H), 3.32 (s, 1H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -139.03 – -139.16 (m, 2F), -152.88 – -153.01 (m, 2F);  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ):  $\delta$  162.09, 146.24 (d,  $J$  = 248.7 Hz), 140.86 (d,  $J$  = 251.5 Hz), 132.61, 130.68, 129.93, 128.74, 127.19, 103.55 (t,  $J$  = 21.8 Hz), 82.54, 81.46; HRMS (m/z) calculated for  $\text{C}_{15}\text{H}_7\text{F}_4\text{O}_2$  [M+H]: 295.0377; found: 295.0374.

**Pentafluorophenyl 4-ethynylbenzoate (13).** This compound was synthesized according to General Procedure A starting from 4-ethynylbenzoic acid and pentafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 214 mg (84 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.16 (d,  $J$  = 8.2 Hz, 2H), 7.65 (d,  $J$  = 8.1 Hz, 2H), 3.33 (s, 1H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -152.61 – -152.73 (m, 2F), -157.90 (t,  $J$  = 21.8 Hz, 1F), -162.30 – -162.52 (m, 2F); HRMS (m/z) calculated for  $\text{C}_{15}\text{H}_6\text{F}_5\text{O}_2$  [M+H]: 313.0283; found: 313.0279.

**4-Trifluoromethyl-2,3,5,6-tetrafluorophenyl 4-ethynylbenzoate (14).** This compound was synthesized according to General Procedure A starting from 4-ethynylbenzoic acid and 4-trifluoromethyl-2,3,5,6-tetrafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 148 mg (50 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.16 (d,  $J$  = 8.5 Hz, 2H), 7.66 (d, 8.5 Hz, 2H), 3.34 (s, 1H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -56.32 (t,  $J$  = 22.0 Hz, 3F), -140.35 – -140.67 (m, 2F), -150.23 – -150.38 (m, 2F); HRMS (m/z) calculated for  $\text{C}_{16}\text{H}_6\text{F}_7\text{O}_2$  [M+H]: 363.0251; found: 363.0252.

**4-Ethynylbenzoic acid NHS ester (15).** This compound was synthesized according to General Procedure A starting from 4-ethynylbenzoic acid and N-hydroxysuccinimide. The preparative TLC was run with DCM/ethyl acetate 4:1. 94 mg (47 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.09 (d,  $J$  = 8.1 Hz, 2H), 7.61 (d,  $J$  = 8.1 Hz, 2H), 3.32 (s, 1H), 2.92 (s, 4H); HRMS (m/z) calculated for  $\text{C}_{13}\text{H}_{10}\text{NO}_4$  [M+H]: 244.0604; found: 244.0598.

**Pentafluorophenyl 3-(1,3-diphenyl-1H-pyrazol-4-yl)propanoate (19).** This compound was synthesized according to General Procedure A starting from 3-(1,3-diphenyl-1H-pyrazol-4-yl)propanoic acid and pentafluorophenol. The preparative TLC was run with n-hexane/DCM 1:1. 358 mg (95 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.88 (s, 1H), 7.77 – 7.71 (m, 4H), 7.51 – 7.43 (m, 4H), 7.43 – 7.37 (m, 1H), 7.32 – 7.27 (m, 1H), 3.20 (t,  $J$  = 7.4 Hz, 2H), 2.99 (t,  $J$  = 7.4 Hz, 2H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -152.86 – -153.01 (m, 2F), -158.08

(t,  $J = 21.7$  Hz, 1F), -162.31 – -162.54 (m, 2F);  $^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  168.90, 151.58, 141.23 (d,  $J = 249.2$  Hz), 140.09, 139.62 (d, 237.6 Hz), 138.00 (d,  $J = 250.8$  Hz), 133.47, 129.55, 128.81, 128.18, 127.99, 126.58, 126.46, 125.08, 118.96, 118.74, 34.03, 20.01; HRMS-ESI ( $m/z$ ) calculated for  $\text{C}_{24}\text{H}_{16}\text{F}_5\text{N}_2\text{O}_2$  [M+H]: 459.1126; found: 459.1126.

**Pentafluorophenyl 2,2-diphenylacetate (20).** This compound was synthesized according to General Procedure B starting from 2,2-diphenylacetyl chloride and pentafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 274 mg (88 %) of the product were obtained.  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.42 – 7.30 (m, 10H), 5.39 (s, 1H);  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -152.40 – -152.53 (m, 2F), -157.92 (t,  $J = 21.7$  Hz, 1F), -162.37 – -162.67 (m, 2F);  $^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  168.83, 141.30 (d, 250.5 Hz), 139.7 (d, 246.9 Hz), 137.96 (d, 262.6 Hz), 137.09, 129.05, 128.71, 128.04, 125.22, 56.49; HRMS ( $m/z$ ) calculated for  $\text{C}_{20}\text{H}_{12}\text{F}_5\text{O}_2$  [M+H]: 379.0752; found: 379.0737.

**Pentafluorophenyl 3,5-bis(trifluoromethyl)benzoate (21).** This compound was synthesized according to General Procedure B starting from 3,5-bis(trifluoromethyl)benzoyl chloride and pentafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 244 mg (70 %) of the product were obtained.  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.65 (s, 2H), 8.22 (s, 1H);  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -63.33 (s, 6F), -152.41 – -152.53 (m, 2F), -156.57 (t,  $J = 21.7$  Hz, 1F), -161.53 – -161.71 (m, 2F);  $^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  160.40, 141.33 (d, 252.8 Hz), 140.22 (d, 256.3 Hz), 137.70 (d,  $J = 252.8$  Hz), 133.13 (q,  $J = 34.8$  Hz), 130.84, 129.39, 128.22, 124.79, 122.74 (q,  $J = 273.0$  Hz); HRMS ( $m/z$ ) calculated for  $\text{C}_{15}\text{H}_4\text{F}_{11}\text{O}_2$  [M+H]: 425.0030; found: 425.0036.

**Pentafluorophenyl 2-(1-methyl-1H-indol-3-yl)acetate (22).** This compound was synthesized according to General Procedure A starting from 2-(1-methyl-1H-indol-3-yl)acetic acid and pentafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 279 mg (96 %) of the product were obtained.  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.62 (d,  $J = 7.9$  Hz, 1H), 7.34 (d,  $J = 8.2$  Hz, 1H), 7.31 – 7.24 (m, 1H), 7.17 (t,  $J = 7.4$  Hz, 1H), 7.12 (s, 1H), 4.12 (s, 2H), 3.80 (s, 3H);  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -152.68 – -152.80 (m, 2F), -158.39 (t,  $J = 21.7$  Hz, 1F), -162.58 – -162.81 (m, 2F);  $^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  168.04, 141.27 (d,  $J = 255.0$  Hz), 139.60 (d,  $J = 241.9$  Hz), 137.94 (d,  $J = 255.0$  Hz), 137.07, 128.13, 127.50, 125.39, 122.21, 119.65, 118.72, 109.58, 104.91, 32.88, 30.35; HRMS-ESI ( $m/z$ ) calculated for  $\text{C}_{17}\text{H}_{11}\text{F}_5\text{NO}_2$  [M+H]: 356.0704; found: 356.0710.

**Pentafluorophenyl 3-(3,4,5-trimethoxyphenyl)propanoate (23).** This compound was synthesized according to General Procedure A starting from 3-(3,4,5-

trimethoxyphenyl)propanoic acid and pentafluorophenol. The preparative TLC was run with DCM. 284 mg (85 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  6.46 (s, 2H), 3.86 (s, 6H), 3.83 (s, 3H), 3.08 – 2.95 (m, 4H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -152.87 – -153.09 (m, 2F), -158.12 (t,  $J$  = 21.7 Hz, 1F), -162.38 – -162.59 (m, 2F);  $^{13}\text{C-NMR}$  (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  168.86, 153.51, 141.24 (d,  $J$  = 246.7 Hz), 139.61 (d,  $J$  = 239.1 Hz), 137.99 (d,  $J$  = 248.4 Hz), 136.90, 135.20, 125.13, 105.33, 60.98, 56.21, 35.24, 31.17; HRMS-ESI (m/z) calculated for  $\text{C}_{18}\text{H}_{16}\text{F}_5\text{O}_5$  [M+H]: 407.0912; found: 407.0914.

**1-Benzyl 4-(pentafluorophenyl) piperidine-1,4-dicarboxylate (24).** This compound was synthesized according to General Procedure A starting from 1-((benzyloxy)carbonyl)piperidine-4-carboxylic acid and pentafluorophenol. The preparative TLC was run with DCM. 304 mg (86%) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.41 – 7.29 (m, 5H), 5.14 (s, 2H), 4.13 (s, 2H), 3.07 (t,  $J$  = 11.8 Hz, 2H), 2.89 (dd,  $J$  = 10.2, 3.8 Hz, 1H), 2.17 – 1.98 (m, 2H), 1.93 – 1.75 (m, 2H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -153.33 – -153.49 (m, 2F), -157.99 (t,  $J$  = 21.7 Hz, 1F), -162.28 – -162.50 (m, 2F); HRMS-ESI (m/z) calculated for  $\text{C}_{20}\text{H}_{17}\text{F}_5\text{NO}_4$  [M+H]: 430.1072; found: 430.1071.

**Pentafluorophenyl quinoline-2-carboxylate (25).** This compound was synthesized according to General Procedure A starting from quinoline-2-carboxylic acid and pentafluorophenol. The preparative TLC was run with n-hexane/DCM 1:1. 230 mg (83 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.42 (d,  $J$  = 8.5 Hz, 1H), 8.37 (d,  $J$  = 8.6 Hz, 1H), 8.31 (d,  $J$  = 8.6 Hz, 1H), 7.96 (d,  $J$  = 8.2 Hz, 1H), 7.87 (t,  $J$  = 7.8 Hz, 1H), 7.74 (t,  $J$  = 7.6 Hz, 1H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -151.99 – -152.13 (m, 2F), -157.62 (t,  $J$  = 21.7 Hz, 1F), -162.18 – -162.38 (m, 2F);  $^{13}\text{C-NMR}$  (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  161.73, 147.94, 145.09, 141.45 (d,  $J$  = 249.6), 139.78 (d,  $J$  = 251.1 Hz), 138.12 (d,  $J$  = 249.6 Hz), 137.88, 131.01 (two overlapping signals), 129.95, 129.73, 127.81, 125.66, 121.75; HRMS-ESI (m/z) calculated for  $\text{C}_{16}\text{H}_7\text{F}_5\text{NO}_2$  [M+H]: 340.0391; found: 340.0389.

**Pentafluorophenyl 3-(7-fluoro-4-oxo-4H-chromen-3-yl)propanoate (26).** This compound was synthesized according to General Procedure A starting from 3-(7-fluoro-4-oxo-4H-chromen-3-yl)propanoic acid and pentafluorophenol. The preparative TLC was run with DCM. 307 mg (93 %) of the product were obtained.  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.93 (s, 1H), 7.86 (dd,  $J$  = 8.2, 2.7 Hz, 1H), 7.48 (dd,  $J$  = 9.3, 4.2 Hz, 1H), 7.44 – 7.37 (m, 1H), 3.08 (t,  $J$  = 6.9 Hz, 2H), 2.90 (t,  $J$  = 6.9 Hz, 2H);  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -115.29 (s, 1F), -152.79 – -152.91 (m, 2F), -158.13 (t,  $J$  = 21.7 Hz, 1F), -162.38 – -162.58 (m, 2F);  $^{13}\text{C-NMR}$  (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  176.98,

169.09, 159.66 (d,  $J$  = 247.6 Hz), 153.68, 152.91, 141.17 (d,  $J$  = 245.2 Hz), 139.58 (d,  $J$  = 240.6 Hz), 137.97 (d,  $J$  = 245.2 Hz), 125.05 (two overlapping signals), 122.16 (d,  $J$  = 25.5 Hz), 121.28, 120.44, 110.65 (d,  $J$  = 23.7 Hz), 31.80, 21.91; HRMS-ESI ( $m/z$ ) calculated for  $C_{18}H_9F_6O_4$  [M+H]: 403.0400; found: 403.0400.

**Pentafluorophenyl 2-(1,3-dioxoisooindolin-2-yl)acetate (27).** This compound was synthesized according to General Procedure A starting from 2-(1,3-dioxoisooindolin-2-yl)acetic acid and pentafluorophenol. The preparative TLC was run with DCM. 257 mg (84 %) of the product were obtained.  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  7.96 – 7.90 (m, 2H), 7.82 – 7.75 (m, 2H), 4.81 (s, 2H);  $^{19}F$ -NMR (376 MHz,  $CDCl_3$ )  $\delta$  -152.01 – -152.17 (m, 2F), -157.15 (t,  $J$  = 21.6 Hz, 1F), -161.89 – -162.14 (m, 2F);  $^{13}C$ -NMR (151 MHz,  $CDCl_3$ ):  $\delta$  167.07, 163.85, 141.13 (d,  $J$  = 252.1 Hz), 140.00 (d,  $J$  = 255.1 Hz), 138.03 (d,  $J$  = 252.1 Hz), 134.68, 131.92, 124.62, 124.03, 38.29; HRMS-ESI ( $m/z$ ) calculated for  $C_{16}H_7F_5NO_4$  [M+H]: 372.0290; found: 372.0280.

**Pentafluorophenyl 1-ethyl-7-methyl-4-oxo-1,4-dihydro-1,8-naphthyridine-3-carboxylate (28).** This compound was synthesized according to General Procedure A starting from 1-ethyl-7-methyl-4-oxo-1,4-dihydro-1,8-naphthyridine-3-carboxylic acid and pentafluorophenol. The preparative TLC was run with ethyl acetate/DCM 1:4. 245 mg (75 %) of the product were obtained.  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  8.79 (s, 1H), 8.68 (d,  $J$  = 8.1 Hz, 1H), 7.31 (d,  $J$  = 8.1 Hz, 1H), 4.55 (q,  $J$  = 7.2 Hz, 2H), 2.70 (s, 3H), 1.55 (t,  $J$  = 7.2 Hz, 3H);  $^{19}F$ -NMR (376 MHz,  $CDCl_3$ )  $\delta$  -152.27 – -152.46 (m, 2F), -158.73 (t,  $J$  = 21.5 Hz, 1F), -162.91 – -163.10 (m, 2F);  $^{13}C$ -NMR (151 MHz,  $CDCl_3$ ):  $\delta$  174.36, 163.44, 160.90, 150.32, 148.68, 141.61 (d,  $J$  = 256.3 Hz), 139.59 (d,  $J$  = 230.6 Hz), 137.96 (d,  $J$  = 262.6 Hz), 137.08, 125.46, 122.04, 121.70, 108.51, 47.32, 25.23, 15.38; HRMS-ESI ( $m/z$ ) calculated for  $C_{18}H_{12}F_5N_2O_3$  [M+H]: 399.0763; found: 399.0764.

**2,4-Dinitrophenyl 3-(1,3-diphenyl-1H-pyrazol-4-yl)propanoate (29).** This compound was synthesized according to General Procedure A starting from 3-(1,3-diphenyl-1H-pyrazol-4-yl)propanoic acid and 2,4-dinitrophenol. The preparative TLC was run with ethyl acetate/n-hexane 2:3. A second preparative TLC was run with DCM/ethyl acetate 5:1. 142 mg (38 %) of the product were obtained.  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  8.95 (d,  $J$  = 2.7 Hz, 1H), 8.48 (dd,  $J$  = 8.9, 2.7 Hz, 1H), 7.90 (s, 1H), 7.79 – 7.72 (m, 4H), 7.51 – 7.43 (m, 4H), 7.42 – 7.35 (m, 2H), 7.31 – 7.26 (m, 1H), 3.20 (t,  $J$  = 7.4 Hz, 2H), 3.01 (t,  $J$  = 7.4 Hz, 2H);  $^{13}C$ -NMR (151 MHz,  $CDCl_3$ ):  $\delta$  169.73, 151.47, 148.50, 145.16, 141.69, 140.01, 133.45, 129.53, 129.16, 128.81,

128.16, 127.92, 126.75, 126.68, 126.43, 121.80, 118.88, 118.79, 34.64, 19.63; HRMS-ESI (m/z) calculated for C<sub>24</sub>H<sub>19</sub>N<sub>4</sub>O<sub>6</sub> [M+H]: 459.1299; found: 459.1299.

**2,4-Dinitrophenyl 2,2-diphenylacetate (30).** This compound was synthesized according to General Procedure B starting from 2,2-diphenylacetyl chloride and 2,4-dinitrophenol. The preparative TLC was run with n-hexane/DCM 2:3. The product was further purified by column chromatography using n-hexane/DCM 3:2. 114 mg (37 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.95 (d, J = 2.7 Hz, 1H), 8.48 (dd, J = 8.9, 2.7 Hz, 1H), 7.43 – 7.31 (m, 11H), 5.40 (s, 1H); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 169.56, 148.60, 145.28, 141.92, 136.86, 129.09, 129.04, 128.84, 128.10, 126.57, 121.80, 56.81; HRMS-ESI (m/z) calculated for C<sub>20</sub>H<sub>14</sub>N<sub>2</sub>NaO<sub>6</sub>[M+Na]: 401.0744; found: 401.0746.

**2,4-Dinitrophenyl 3,5-bis(trifluoromethyl)benzoate (31).** This compound was synthesized according to General Procedure B starting from 3,5-bis(trifluoromethyl)benzoyl chloride and 2,4-dinitrophenol. The preparative TLC was run with n-hexane/DCM 2:3. The product was further purified by column chromatography using n-hexane/DCM 3:2. 114 mg (33 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.09 (d, J = 2.6 Hz, 1H), 8.68 – 8.60 (m, 3H), 8.22 (s, 1H), 7.67 (d, J = 8.9 Hz, 1H); <sup>19</sup>F-NMR (376 MHz, CDCl<sub>3</sub>) δ -63.28 (s, 6F). <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 161.40, 148.20, 145.83, 141.58, 133.11 (q, J = 33.9 Hz), 130.81, 129.90, 129.61, 128.26, 126.79, 122.73 (q, J = 273.9 Hz), 122.29; HRMS (m/z) calculated for C<sub>15</sub>H<sub>6</sub>F<sub>6</sub>N<sub>2</sub>NaO<sub>6</sub> [M+Na]: 447.0022; found: 447.0029.

**2,4-Dinitrophenyl 2-(1-methyl-1H-indol-3-yl)acetate (32).** This compound was synthesized according to General Procedure A starting from 2-(1-methyl-1H-indol-3-yl)acetic acid and 2,4-dinitrophenol. The preparative TLC was run with DCM/n-hexane 2:1. 234 mg (54 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.94 (d, J = 2.7 Hz, 1H), 8.45 (dd, J = 8.9, 2.7 Hz, 1H), 7.65 (d, J = 7.9 Hz, 1H), 7.40 (d, J = 8.9 Hz, 1H), 7.34 (d, J = 8.2 Hz, 1H), 7.27 (t, J = 7.2 Hz, 1H), 7.17 (t, J = 7.4 Hz, 2H), 4.15 (s, 2H), 3.80 (s, 3H); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 168.90, 148.96, 145.10, 141.75, 137.07, 129.04, 128.44, 127.59, 126.79, 122.21, 121.78, 119.71, 118.76, 109.65, 104.68, 32.95, 31.07; HRMS-ESI (m/z) calculated for C<sub>17</sub>H<sub>14</sub>N<sub>3</sub>O<sub>6</sub> [M+H]: 356.0877; found: 356.0878.

**2,4-Dinitrophenyl 3-(3,4,5-trimethoxyphenyl)propanoate (33).** This compound was synthesized according to General Procedure A starting from 3-(3,4,5-trimethoxyphenyl)propanoic acid and 2,4-dinitrophenol. The preparative TLC was run with ethyl acetate/n-hexane 2:3. 143 mg (43 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>):

$\delta$  8.97 (d,  $J$  = 2.1 Hz, 1H), 8.52 (dd,  $J$  = 9.0, 2.1 Hz, 1H), 7.40 (d,  $J$  = 9.0 Hz, 1H), 6.47 (s, 2H), 3.87 (s, 6H), 3.84 (s, 3H), 3.08 – 2.98 (m, 4H);  $^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  169.74, 153.49, 148.62, 145.22, 141.78, 136.86, 135.28, 129.19, 126.71, 121.85, 105.41, 60.99, 56.26, 35.48, 30.80; HRMS-ESI (m/z) calculated for  $\text{C}_{18}\text{H}_{19}\text{N}_2\text{O}_9$  [M+H]: 407.1085; found: 407.1087.

**1-Benzyl 4-(2,4-dinitrophenyl) piperidine-1,4-dicarboxylate (34).** This compound was synthesized according to General Procedure A starting from 1-((benzyloxy)carbonyl)piperidine-4-carboxylic acid and 2,4-dinitrophenol. The preparative TLC was run with ethyl acetates/DCM 1:9. 215 mg (61 %) of the product were obtained.  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.97 (d,  $J$  = 2.6 Hz, 1H), 8.52 (dd,  $J$  = 8.9, 2.7 Hz, 1H), 7.46 (d,  $J$  = 8.9 Hz, 1H), 7.39 – 7.29 (m, 5H), 5.15 (s, 2H), 4.21 (s, 2H), 3.02 (t,  $J$  = 12.6 Hz, 2H), 2.87 (tt,  $J$  = 11.0, 3.9 Hz, 1H), 2.17 – 2.05 (m, 2H), 1.92 – 1.77 (m, 2H); HRMS-ESI (m/z) calculated for  $\text{C}_{20}\text{H}_{20}\text{N}_3\text{O}_8$  [M+H]: 430.1245; found: 430.1243.

**2,4-Dinitrophenyl quinoline-2-carboxylate (35).** This compound was synthesized according to General Procedure A starting from quinoline-2-carboxylic acid and 2,4-dinitrophenol. The preparative TLC was run with DCM. 25 mg (9 %) of the product were obtained.  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.08 (d,  $J$  = 2.6 Hz, 1H), 8.62 (dd,  $J$  = 9.0, 2.7 Hz, 1H), 8.43 (d,  $J$  = 8.5 Hz, 1H), 8.36 (d,  $J$  = 8.6 Hz, 1H), 8.32 (d,  $J$  = 8.5 Hz, 1H), 7.97 (d,  $J$  = 8.2 Hz, 1H), 7.87 (t,  $J$  = 7.7 Hz, 1H), 7.79 – 7.70 (m, 2H); HRMS-ESI (m/z) calculated for  $\text{C}_{16}\text{H}_{10}\text{N}_3\text{O}_6$  [M+H]: 340.0564; found: 340.0565.

**2,4-Dinitrophenyl 3-(7-fluoro-4-oxo-4H-chromen-3-yl)propanoate (36).** This compound was synthesized according to General Procedure A starting from 3-(7-fluoro-4-oxo-4H-chromen-3-yl)propanoic acid and 2,4-dinitrophenol. The preparative TLC was run with  $\text{CHCl}_3$ /acetone 95:5. 62 mg (19 %) of the product were obtained.  $^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  8.97 (d,  $J$  = 2.7 Hz, 1H), 8.54 (dd,  $J$  = 8.9, 2.7 Hz, 1H), 7.97 (s, 1H), 7.89 (dd,  $J$  = 8.2, 3.1 Hz, 1H), 7.54 – 7.47 (m, 2H), 7.46 – 7.40 (m, 1H), 3.12 (t,  $J$  = 6.9 Hz, 2H), 2.93 (t,  $J$  = 6.9 Hz, 2H);  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -115.29 (s, 1F);  $^{13}\text{C}$ -NMR (151 MHz,  $\text{CDCl}_3$ ):  $\delta$  176.94, 169.84, 159.46 (d,  $J$  = 247.1 Hz), 153.62, 152.73, 148.43, 145.03, 141.60, 129.02, 126.64, 124.85 (d,  $J$  = 7.3 Hz), 121.97 (d,  $J$  = 25.7 Hz), 121.62, 121.23, 120.27 (d,  $J$  = 8.1 Hz), 110.45 (d,  $J$  = 23.6 Hz), 32.27, 21.35; HRMS-ESI (m/z) calculated for  $\text{C}_{18}\text{H}_{12}\text{FN}_2\text{O}_8$  [M+H]: 403.0572; found: 403.0575.

**2,4-Dinitrophenyl [1,1'-biphenyl]-4-carboxylate (37).** This compound was synthesized according to General Procedure A starting from 1,1'-biphenyl-4-carboxylic acid and 2,4-dinitrophenol. The preparative TLC was run with n-hexane/DCM 2:3. The product was further

purified by column chromatography using n-hexane/DCM 3:2. 57 mg (19 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.02 (d, J = 2.7 Hz, 1H), 8.59 (dd, J = 8.9, 2.7 Hz, 1H), 8.26 (d, J = 8.3 Hz, 2H), 7.78 (d, J = 8.3 Hz, 2H), 7.70 – 7.64 (m, 3H), 7.51 (t, J = 7.5 Hz, 2H), 7.45 (t, J = 7.3 Hz, 1H); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 163.48, 149.11, 147.80, 145.17, 142.04, 139.62, 131.44, 129.23, 129.08, 128.81, 127.72, 127.52, 126.90, 126.12, 121.91; HRMS-ESI (m/z) calculated for C<sub>19</sub>H<sub>12</sub>N<sub>2</sub>NaO<sub>6</sub> [M+Na]: 387.0588; found: 387.0588.

**2,4-Dinitrophenyl 2-(adamantan-1-yl)acetate (38).** This compound was synthesized according to General Procedure A starting from 2-(adamantan-1-yl)acetic acid and 2,4-dinitrophenol. The preparative TLC was run with n-hexane/DCM 2:3. The product was further purified by column chromatography using n-hexane/DCM 3:2. 143 mg (48 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.93 (d, J = 2.6 Hz, 1H), 8.50 (dd, J = 9.0, 2.6 Hz, 1H), 7.47 (d, J = 8.9 Hz, 1H), 2.45 (s, 2H), 2.03 (s, 3H), 1.81 – 1.63 (m, 12H); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 168.31, 148.73, 145.03, 142.16, 128.92, 126.79, 121.74, 48.07, 42.33, 36.74, 33.41, 28.66; HRMS (m/z) calculated for C<sub>18</sub>H<sub>20</sub>N<sub>2</sub>NaO<sub>6</sub> [M+Na]: 383.1213; found: 383.1204.

**2,4-Dinitrophenyl 4-phenoxybenzoate (39).** This compound was synthesized according to General Procedure A starting from 4-phenoxybenzoic acid and 2,4-dinitrophenol. The preparative TLC was run with n-hexane/DCM 2:3. A second preparative TLC was run with n-hexane/ethyl acetate 6:1. 70 mg (22 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.00 (d, J = 2.7 Hz, 1H), 8.56 (dd, J = 9.0, 2.8 Hz, 1H), 8.18 – 8.12 (m, 2H), 7.65 (d, J = 8.9 Hz, 1H), 7.44 (t, J = 7.7 Hz, 2H), 7.28 – 7.22 (m, 1H), 7.12 (d, J = 8.4 Hz, 2H), 7.07 (d, J = 9.0 Hz, 2H); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 163.77, 162.95, 155.06, 149.12, 145.05, 142.02, 133.20, 130.34, 129.01, 126.87, 125.25, 121.83, 121.32, 120.63, 117.53; HRMS-ESI (m/z) calculated for C<sub>19</sub>H<sub>12</sub>N<sub>2</sub>NaO<sub>7</sub> [M+Na]: 403.0537; found: 403.0537.

**2,4-Dinitrophenyl 2-((3-(trifluoromethyl)phenyl)amino)benzoate (40).** This compound was synthesized according to General Procedure A starting from 2-((3-(trifluoromethyl)phenyl)amino)benzoic acid and 2,4-dinitrophenol. The preparative TLC was run with DCM/n-hexane 3:2. 254 mg (69 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.11 (s, 1H), 9.01 (d, J = 2.7 Hz, 1H), 8.57 (dd, J = 8.9, 2.7 Hz, 1H), 8.20 (dd, J = 8.1, 1.7 Hz, 1H), 7.64 (d, J = 8.9 Hz, 1H), 7.53 – 7.45 (m, 3H), 7.44 – 7.36 (m, 2H), 7.28 (d, J = 8.6 Hz, 1H), 6.91 (t, J = 7.4 Hz, 1H); <sup>19</sup>F-NMR (376 MHz, CDCl<sub>3</sub>) δ -63.09 (s, 3F); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 165.12, 148.80, 148.68, 145.19, 142.10, 140.65, 136.53, 132.68, 132.15 (q, J = 32.8 Hz), 130.25, 129.08, 127.01, 125.91, 123.93 (q, 272.9 Hz), 121.86, 120.94 (q, J = 3.9 Hz),

119.40 (q,  $J = 3.8$  Hz), 118.72, 114.35, 109.65; HRMS-ESI ( $m/z$ ) calculated for  $C_{20}H_{13}F_3N_3O_6$  [M+H]: 448.0751; found: 448.0753.

**2,4-Dinitrophenyl 4-((tert-butoxycarbonyl)amino)butanoate (41).** This compound was synthesized according to General Procedure A starting from 4-((tert-butoxycarbonyl)amino)butanoic acid and 2,4-dinitrophenol. The preparative TLC was run with ethyl acetate/DCM 1:9. 126 mg (42 %) of the product were obtained.  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  8.96 (d,  $J = 2.6$  Hz, 1H), 8.52 (dd,  $J = 8.9, 2.7$  Hz, 1H), 7.54 (d,  $J = 8.9$  Hz, 1H), 4.68 (s, 1H), 3.27 (q,  $J = 6.6$  Hz, 2H), 2.75 (t,  $J = 7.2$  Hz, 2H), 1.96 (p,  $J = 7.0$  Hz, 2H), 1.45 (s, 9H); HRMS-ESI ( $m/z$ ) calculated for  $C_{15}H_{20}N_3O_8$  [M+H]: 370.1245; found: 370.1244.

**2,4-Dinitrophenyl 2,2,2-triphenylacetate (42).** This compound was synthesized according to General Procedure A starting from 2,2,2-triphenylacetic acid and 2,4-dinitrophenol. The preparative TLC was run with  $CHCl_3$ /acetone 95:5. A second preparative TLC was run with the same solvent mixture. 116 mg (31 %) of the product were obtained.  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  8.89 (d,  $J = 2.7$  Hz, 1H), 8.40 (dd,  $J = 9.0, 2.7$  Hz, 1H), 7.42 – 7.29 (m, 15H), 7.02 (d,  $J = 9.0$  Hz, 1H);  $^{13}C$ -NMR (151 MHz,  $CDCl_3$ ):  $\delta$  170.46, 148.82, 145.14, 142.51, 141.51, 130.29, 128.92, 128.36, 127.70, 125.63, 121.62, 68.30; HRMS-ESI ( $m/z$ ) calculated for  $C_{26}H_{18}N_2NaO_6$  [M+Na]: 477.1057; found: 477.1060.

**2,4-Dinitrophenyl acetate (43).** This compound was synthesized according to General Procedure B starting from acetyl chloride and 2,4-dinitrophenol. The preparative TLC was run with DCM/n-hexane 2:1. 57 mg (31 %) of the product were obtained.  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  8.97 (d,  $J = 2.7$  Hz, 1H), 8.52 (dd,  $J = 8.9, 2.7$  Hz, 1H), 7.48 (d,  $J = 8.9$  Hz, 1H), 2.43 (s, 3H);  $^{13}C$ -NMR (151 MHz,  $CDCl_3$ ):  $\delta$  167.73, 148.73, 145.23, 141.79, 129.19, 126.82, 121.88, 20.88; HRMS ( $m/z$ ) calculated for  $C_8H_6N_2NaO_6$  [M+Na]: 249.0118; found: 249.0116.

**2,4-Dinitrophenyl 4-cyanobenzoate (44).** This compound was synthesized according to General Procedure B starting from 4-cyanobenzoyl chloride and 2,4-dinitrophenol. Instead of a preparative TLC, the reaction was purified using column chromatography with DCM/n-hexane 4:1. 104 mg (41 %) of the product were obtained.  $^1H$ -NMR (400 MHz,  $CDCl_3$ ):  $\delta$  9.05 (d,  $J = 2.8$  Hz, 1H), 8.61 (dd,  $J = 8.9, 2.7$  Hz, 1H), 8.31 (d,  $J = 8.3$  Hz, 2H), 7.87 (d,  $J = 8.3$  Hz, 2H), 7.67 (d,  $J = 8.9$  Hz, 1H);  $^{13}C$ -NMR (151 MHz,  $CDCl_3$ ):  $\delta$  162.24, 148.42, 145.60, 141.69, 132.84, 131.37, 131.27, 129.40, 126.76, 122.13, 118.35, 117.61; HRMS ( $m/z$ ) calculated for  $C_{14}H_8N_3O_6$  [M+H]: 314.0408; found: 314.0406.

**2,4-Dinitrophenyl 3-(benzo[d][1,3]dioxol-5-yl)propanoate (45).** This compound was synthesized according to General Procedure A starting from 3-(benzo[d][1,3]dioxol-5-yl)propanoic acid and 2,4-dinitrophenol. The preparative TLC was run with ethyl acetate/n-hexane 2:3. A second preparative TLC was run with DCM/ethyl acetate 5:1. 108 mg (37 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.96 (d, J = 2.7 Hz, 1H), 8.50 (dd, J = 8.9, 2.7 Hz, 1H), 7.40 (d, J = 8.9 Hz, 1H), 6.80 – 6.68 (m, 3H), 5.95 (s, 2H), 3.06 – 2.94 (m, 4H); HRMS-ESI (m/z) calculated for C<sub>16</sub>H<sub>12</sub>N<sub>2</sub>NaO<sub>8</sub> [M+Na]: 383.0486; found: 383.0488.

**3,5-Bis(trifluoromethyl)benzoic acid NHS ester (46).** This compound was synthesized according to General Procedure B starting from 3,5-bis(trifluoromethyl)benzoyl chloride and N-hydroxysuccinimide. The preparative TLC was run with DCM. 169 mg (58 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.58 (s, 2H), 8.19 (s, 1H), 2.95 (s, 4H); <sup>19</sup>F-NMR (376 MHz, CDCl<sub>3</sub>) δ -63.38 (s, 6F); HRMS-ESI (m/z) calculated for C<sub>13</sub>H<sub>8</sub>F<sub>6</sub>NO<sub>4</sub> [M+H]: 356.0352; found: 356.0352.

**2,3,5,6-Tetrafluoro-4-(trifluoromethyl)phenyl 3,5-bis(trifluoromethyl)benzoate (47).** This compound was synthesized according to General Procedure B starting from 3,5-bis(trifluoromethyl)benzoyl chloride and 2,3,5,6-tetrafluoro-4-(trifluoromethyl)phenol. The preparative TLC was run with n-hexane/DCM 2:1. 283 mg (73 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.65 (s, 2H), 8.23 (s, 1H); <sup>19</sup>F-NMR (376 MHz, CDCl<sub>3</sub>) δ -56.38 (t, J = 22.0 Hz, 3F), -63.36 (s, 6F), -139.52 – -139.92 (m, 2F), -149.93 – -150.20 (m, 2F); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 159.83, 144.89 (d, J = 265.2 Hz), 141.33 (d, J = 249.4 Hz), 133.28 (q, J = 34.9 Hz), 132.07, 130.94, 129.06, 128.48, 122.71 (q, J = 271.9 Hz), 120.77 (q, J = 276.2 Hz), 108.64; HRMS (m/z) calculated for C<sub>16</sub>H<sub>4</sub>F<sub>13</sub>O<sub>2</sub> [M+H]: 474.9998; found: 474.9968.

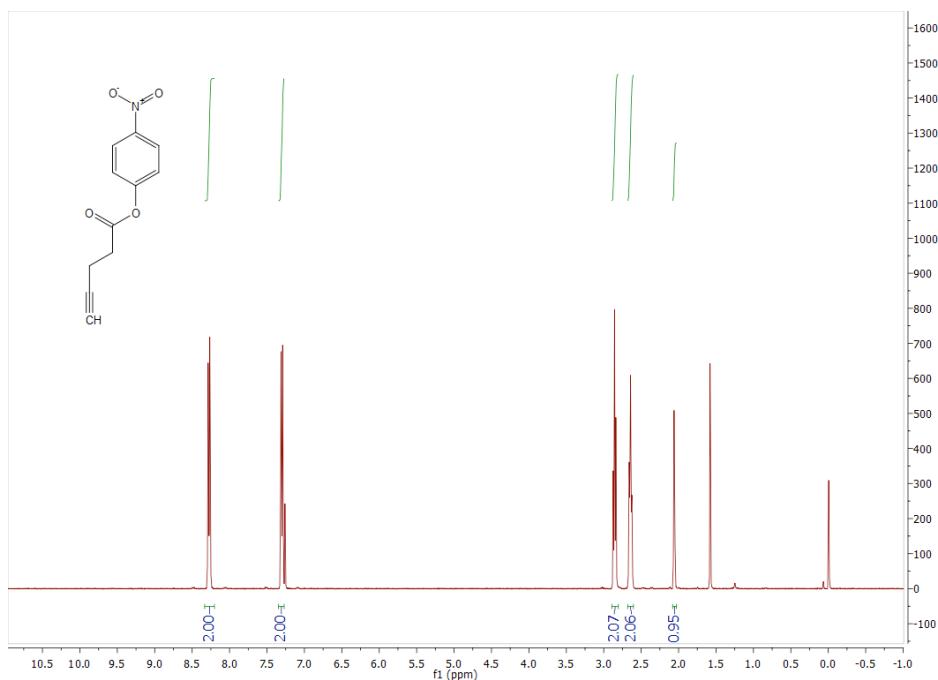
**2,3,5,6-Tetrafluorophenyl 3,5-bis(trifluoromethyl)benzoate (48).** This compound was synthesized according to General Procedure B starting from 3,5-bis(trifluoromethyl)benzoyl chloride and 2,3,5,6-tetrafluorophenol. The preparative TLC was run with n-hexane/DCM 2:1. 285 mg (86 %) of the product were obtained. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 8.66 (s, 2H), 8.21 (s, 1H), 7.11 (tt, J = 9.8, 6.9 Hz, 1H); <sup>19</sup>F-NMR (376 MHz, CDCl<sub>3</sub>) δ -63.31 (s, 6F), -138.31 – -138.44 (m, 2F), -152.69 – -152.82 (m, 2F); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 160.37, 146.33 (d, J = 250.0), 140.83 (d, J = 252.1 Hz), 133.10 (q, J = 34.4 Hz), 130.84, 129.65, 129.35, 128.11, 122.79 (q, J = 271.9 Hz), 104.23 (t, J = 23.1 Hz); HRMS (m/z) calculated for C<sub>15</sub>H<sub>5</sub>F<sub>10</sub>O<sub>2</sub> [M+H]: 407.0124; found: 407.0125.

**N-Methoxycarbonyl-pyrazole-1-carboxamidine (49a).** 2.94 g (20.1 mmol, 1 eq.) pyrazole-1-carboxamidine hydrochloride were dissolved in 20 ml DCM and 10.2 ml (7.55 g, 58 mmol, 2.9 eq.) DIPEA. 1.55 ml (1.9 g, 20.1 mmol, 1 eq.) methyl chloroformate were added and the solution was stirred at room temperature for 12h. The product was loaded onto silica gel without prior workup and purified by column chromatography using DCM as the eluent to give 2.47 g (73 %) of the product. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.04 (s, 1H), 8.44 (d, J = 2.8 Hz, 1H), 7.70 (d, J = 1.0 Hz, 1H), 7.65 (s, 1H), 6.43 (dd, J = 2.8, 1.0 Hz, 1H), 3.81 (s, 3H). <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 164.61, 155.45, 143.82, 128.88, 109.48, 53.02; HRMS (m/z) calculated for C<sub>6</sub>H<sub>9</sub>N<sub>4</sub>O<sub>2</sub> [M+H]: 169.0720; found: 169.0723.

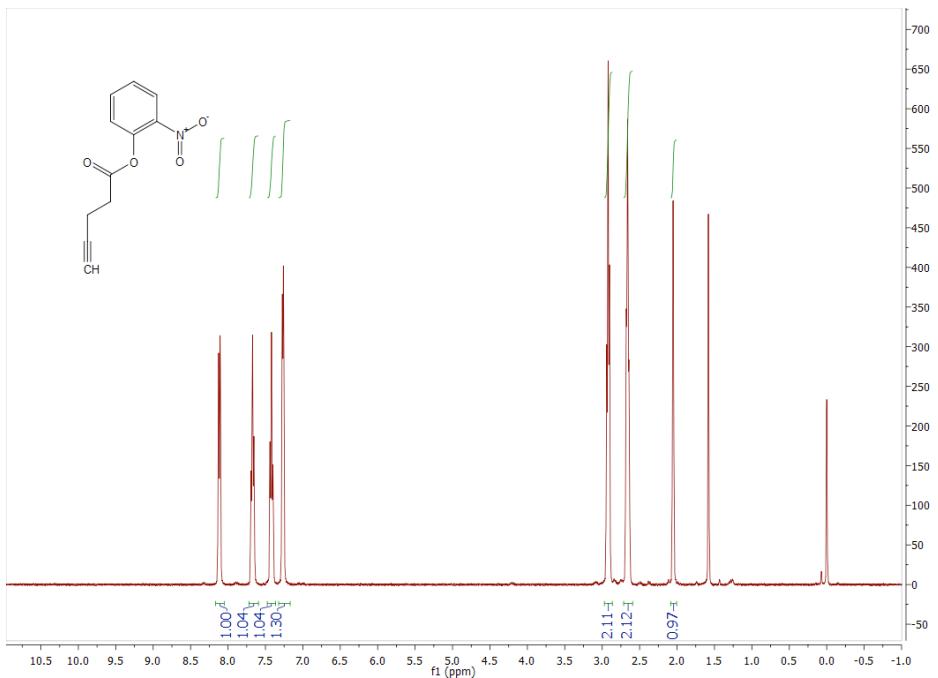
**N-Methoxycarbonyl-N'-9-fluorenylmethoxycarbonyl-pyrazole-1-carboxamidine (49).** 100 mg (0.6 mmol, 1 eq.) **49a** were dissolved in 4 ml anhydrous THF and cooled to 0 °C. To this, 35 mg sodium hydride (60 % in mineral oil, 0.88 mmol, 1.5 eq.) were added and the mixture was stirred at 0 °C for 1 h. 171 mg Fmoc-Cl (0.66 mmol, 1.1 eq.) were added and the reaction was warmed to room temperature over night and directly loaded onto a preparative TLC. The TLC was run with Et<sub>2</sub>O/hexanes 2:1. A second preparative TLC was run with ethyl acetate/n-hexane 1:1. 56 mg (24 %) of the product were obtained as a mixture of two tautomers (ratio of about 1.1:0.9). <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 9.47 – 9.27 (m, 1H both isomers), 8.38 (s, 1H major isomer), 8.32 (s, 1H minor isomer), 7.78 (d, J = 7.6 Hz, 2H both isomers), 7.73 – 7.67 (m, 2H both isomers), 7.65 – 7.56 (m, 1H both isomers), 7.48 – 7.37 (m, 2H both isomers), 7.37 – 7.28 (m, 2H both isomers), 6.51 (s, 1H both isomers), 4.56 – 4.46 (m, 2H both isomers), 4.45 – 4.36 (m, 1H major isomer), 4.34 – 4.25 (m, 1H minor isomer), 3.84 (s, 3H minor isomer), 3.74 (s, 3H major isomer); <sup>13</sup>C-NMR (151 MHz, CDCl<sub>3</sub>): δ 159.07, 158.54, 151.32, 150.88, 144.22, 143.21, 141.42, 138.53, 138.40, 129.10, 128.16, 127.78, 127.40, 127.19, 125.56, 125.15, 120.29, 120.04, 110.55, 69.01, 68.75, 53.86, 46.94, 46.71; HRMS (m/z) calculated for C<sub>21</sub>H<sub>19</sub>N<sub>4</sub>O<sub>4</sub> [M+H]: 391.1401; found: 391.1409.

**(E) NMR spectra**

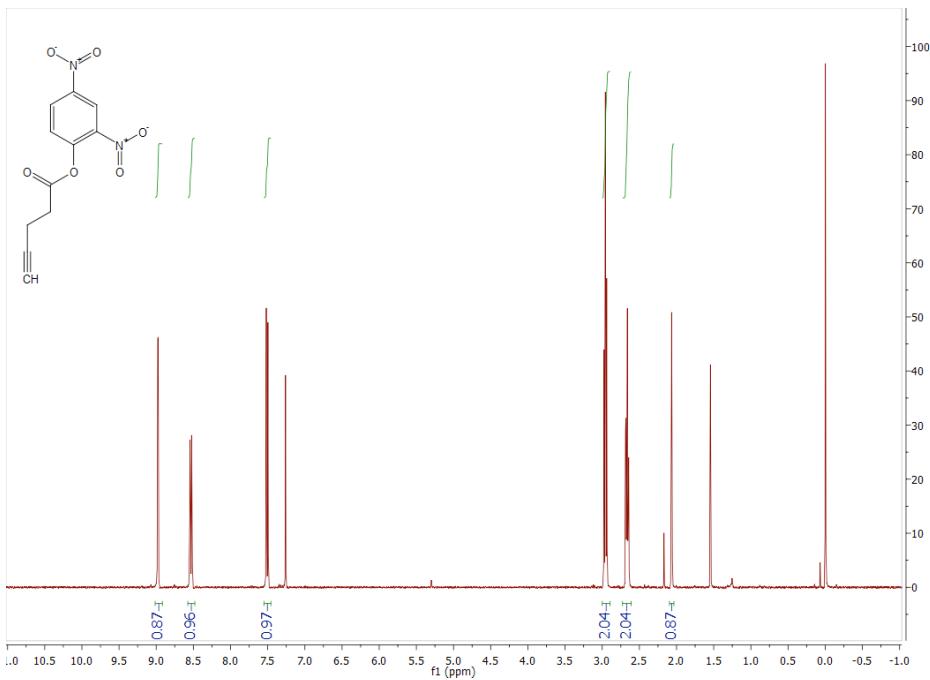
**$^1\text{H-NMR}$  for 2**



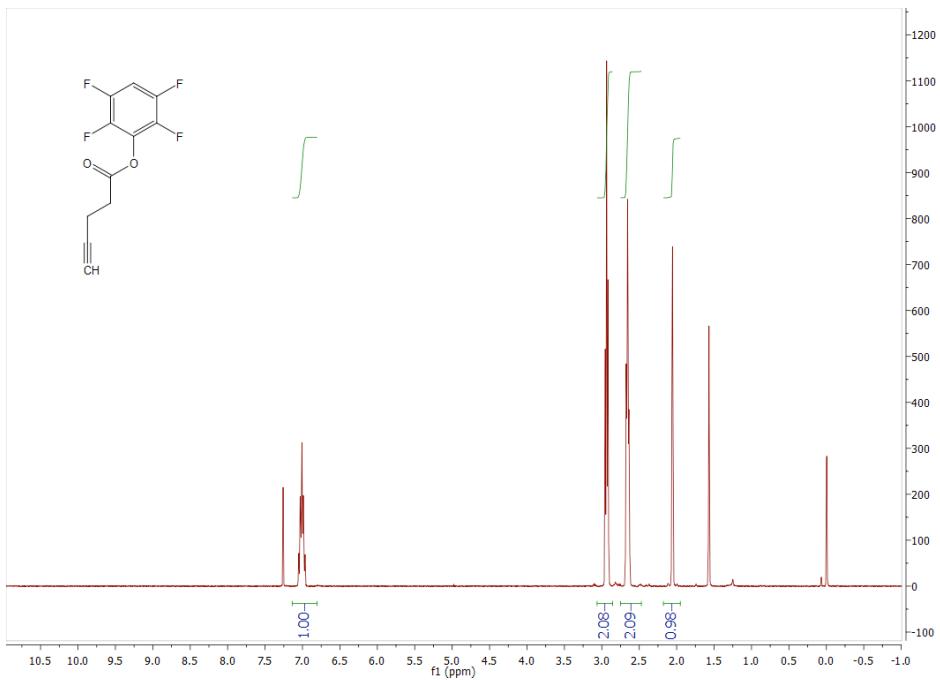
**<sup>1</sup>H-NMR for 3**



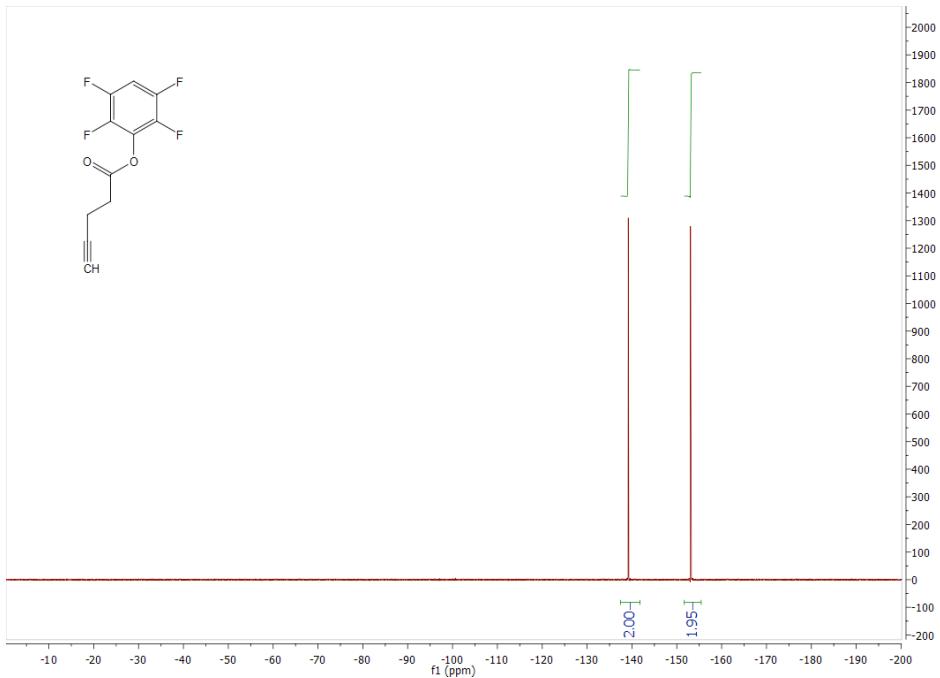
<sup>1</sup>H-NMR for 4



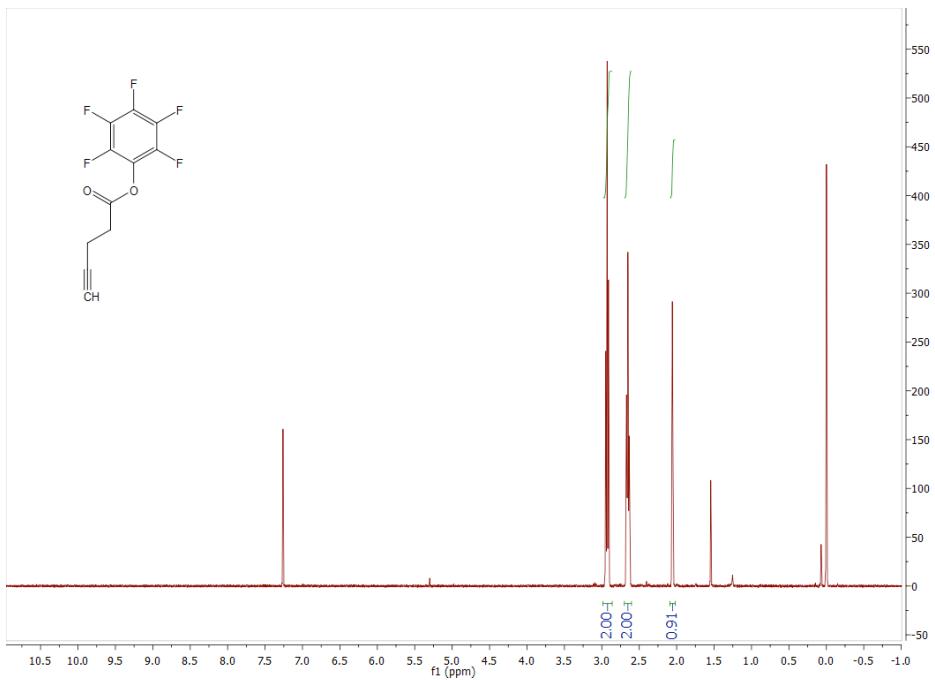
**<sup>1</sup>H-NMR for 5**



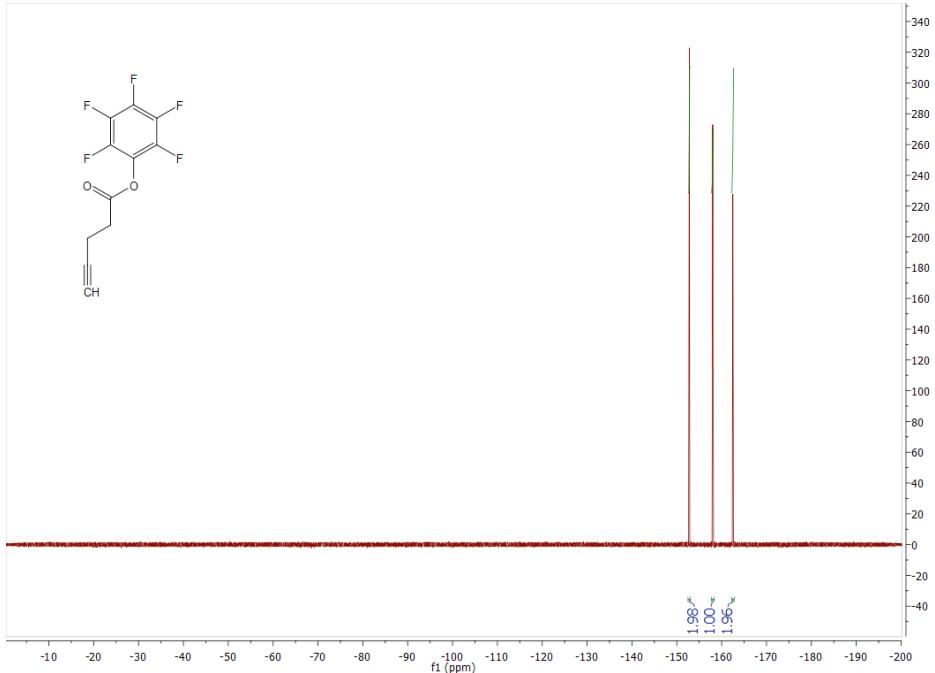
**<sup>19</sup>F-NMR for 5**



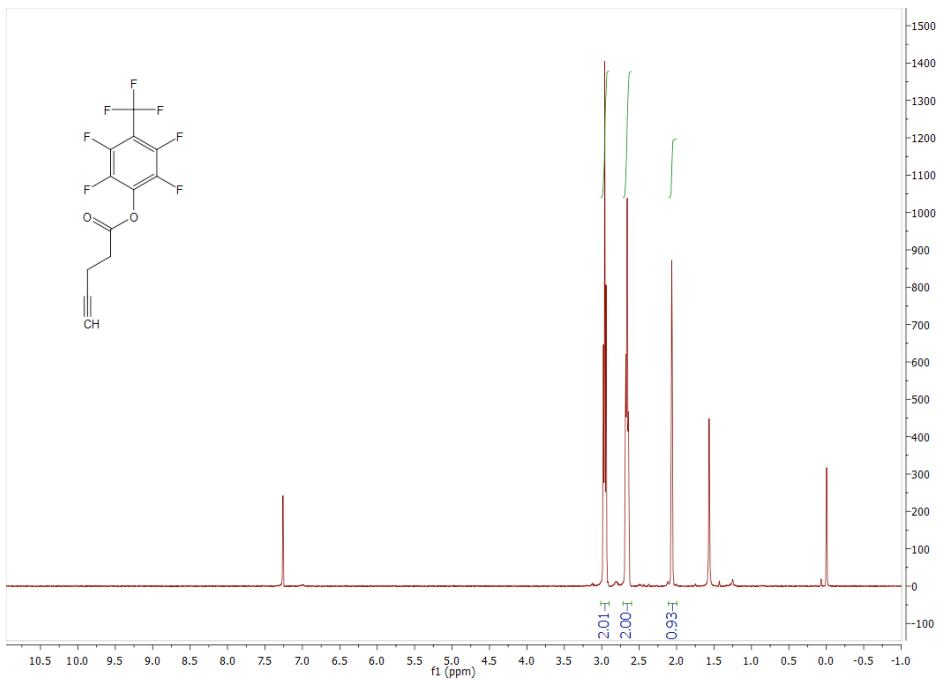
**<sup>1</sup>H-NMR for 6**



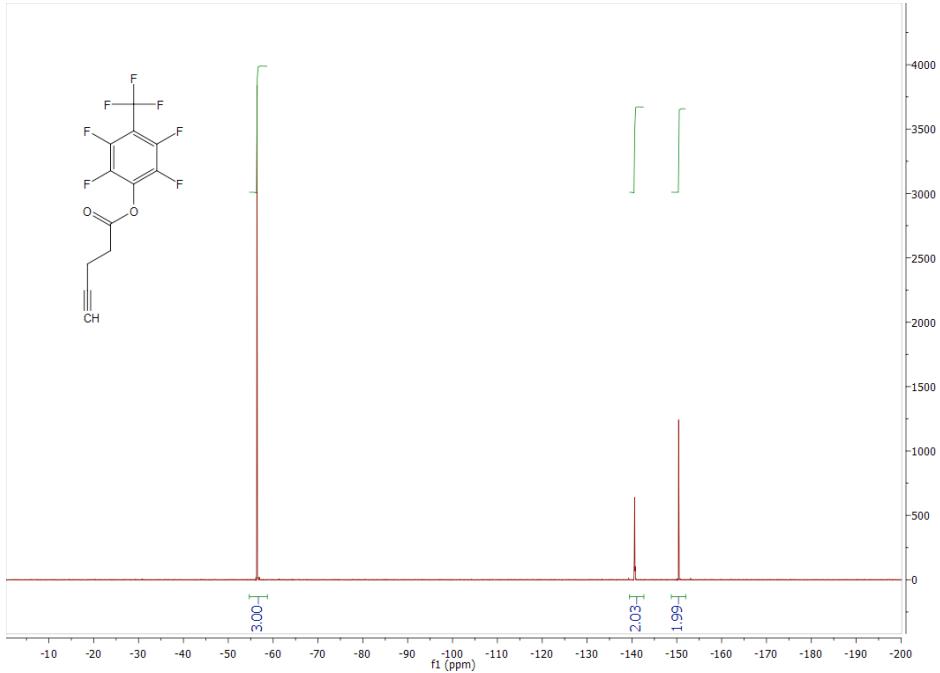
**<sup>19</sup>F-NMR for 6**



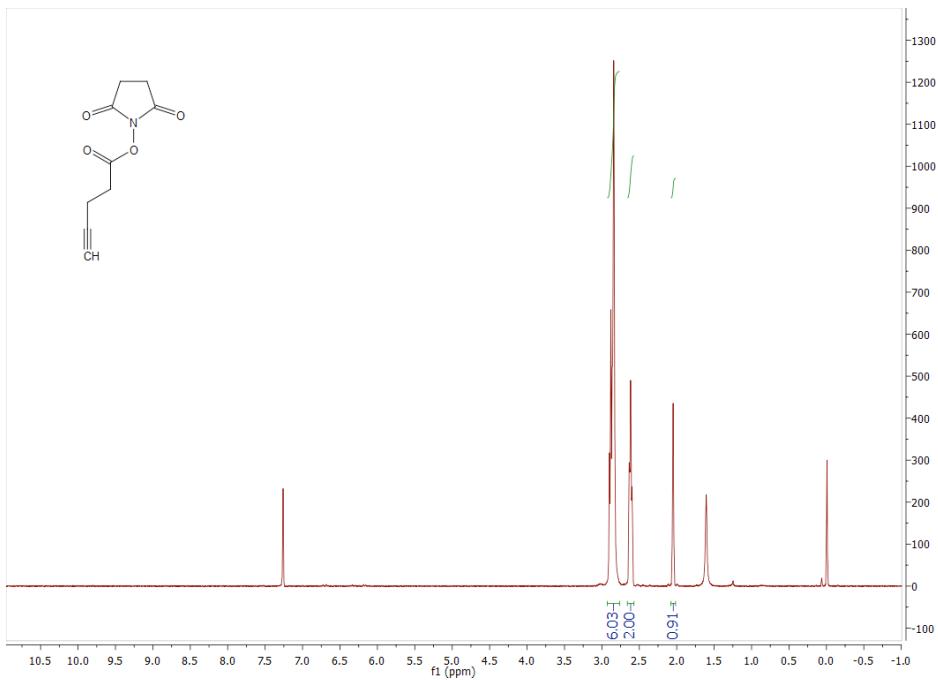
**<sup>1</sup>H-NMR for 7**



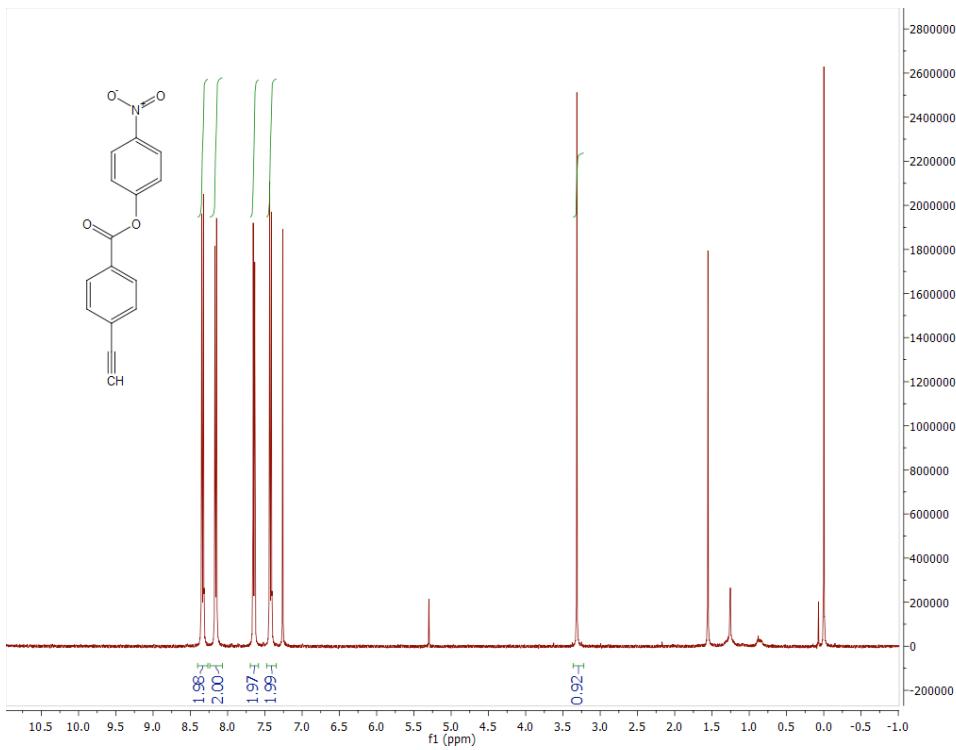
**<sup>19</sup>F-NMR for 7**



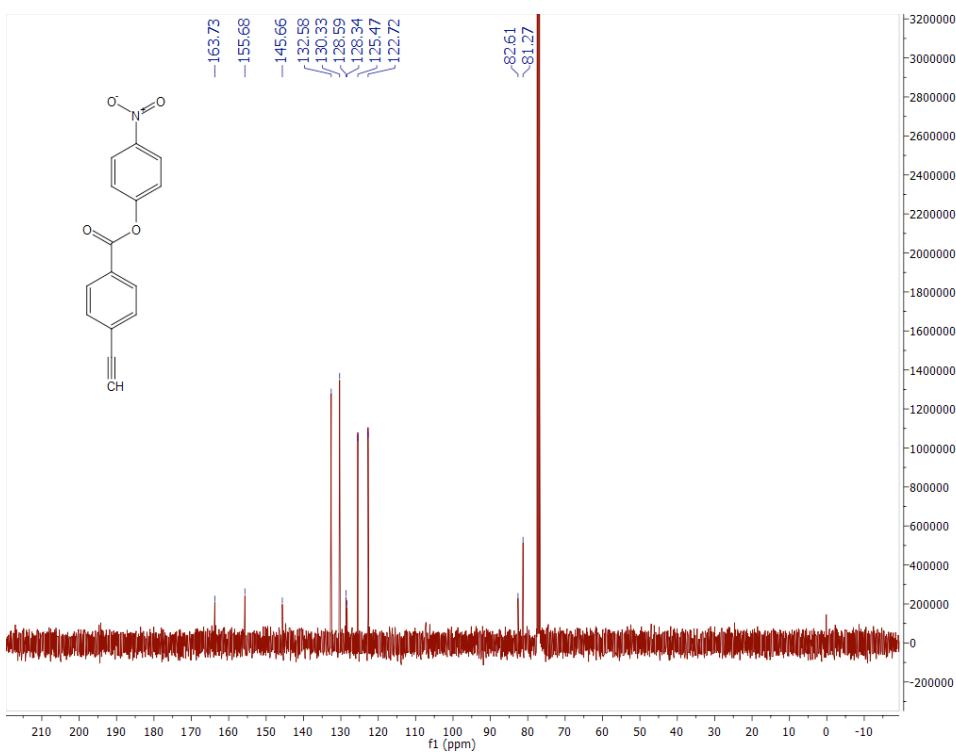
<sup>1</sup>H-NMR for 8



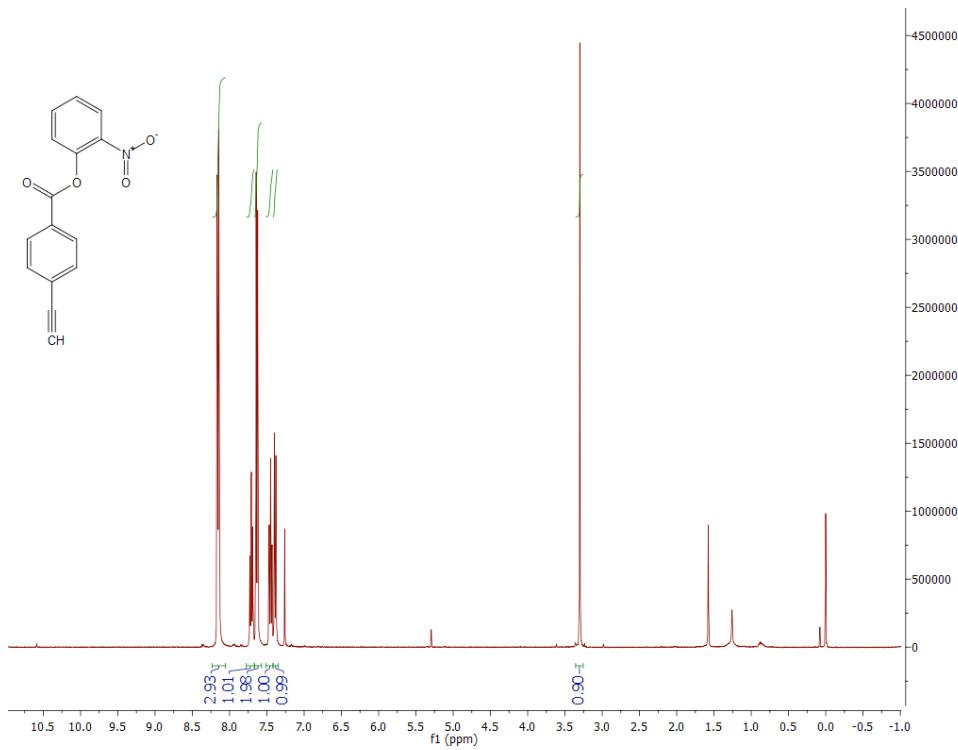
<sup>1</sup>H-NMR for 9



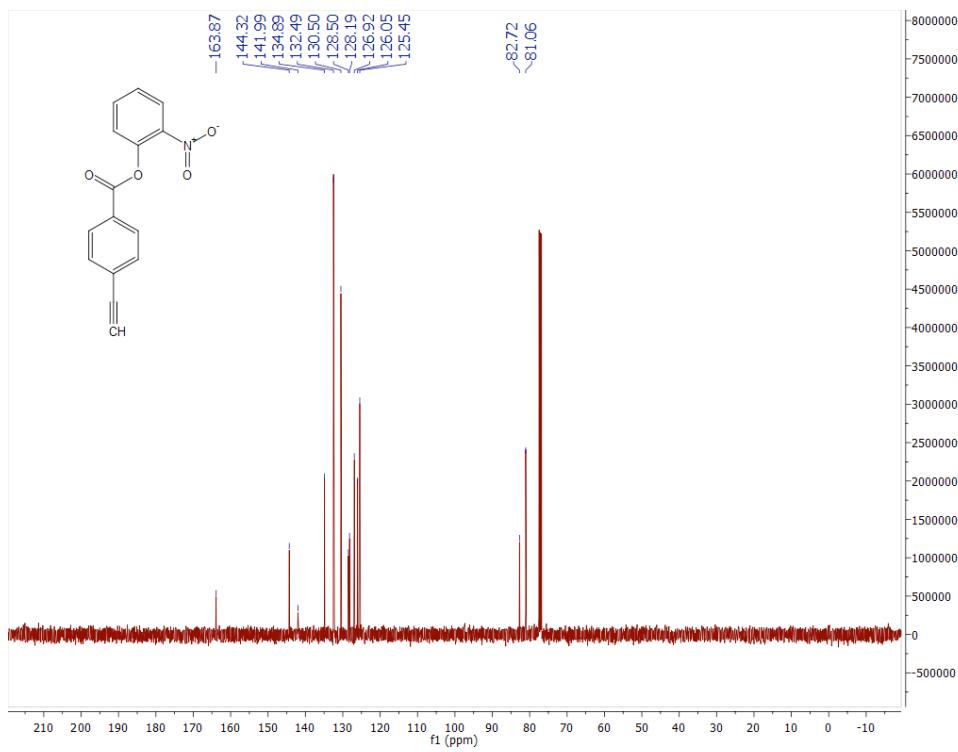
<sup>13</sup>C-NMR for 9



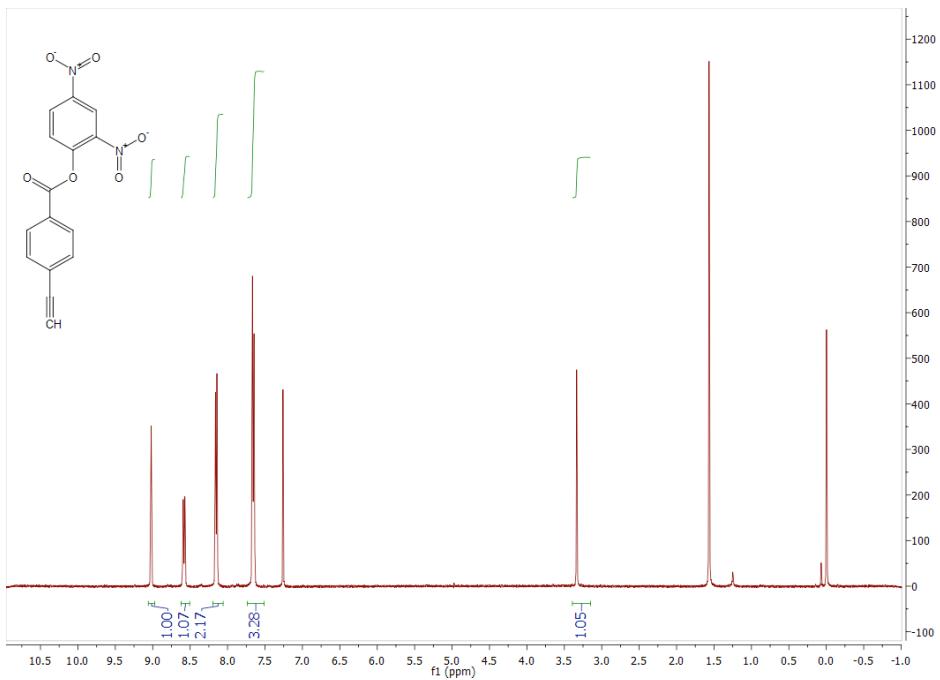
**<sup>1</sup>H-NMR for 10**



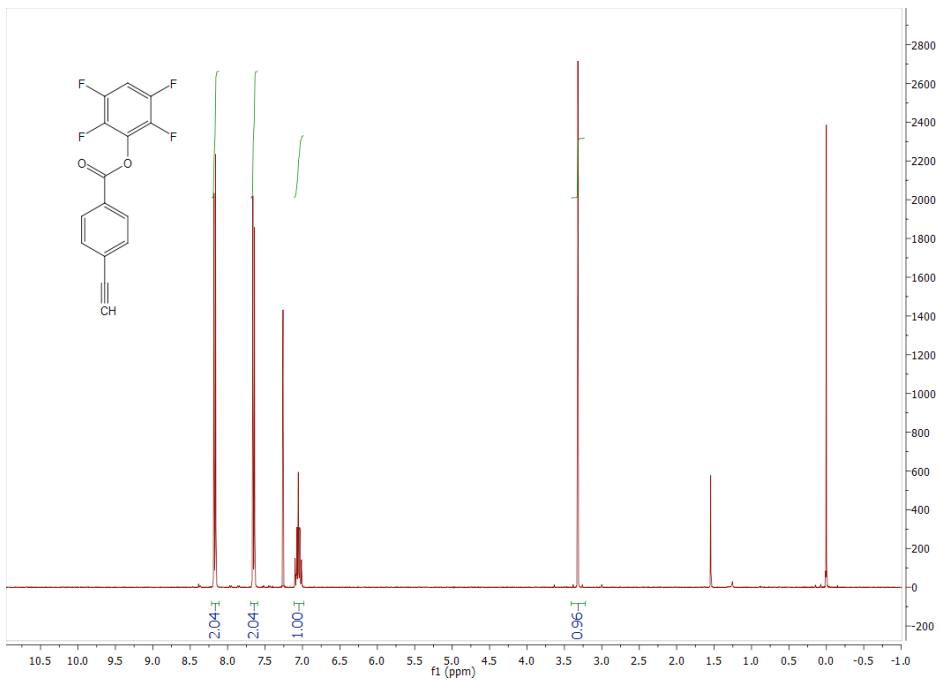
**<sup>13</sup>C-NMR for 10**



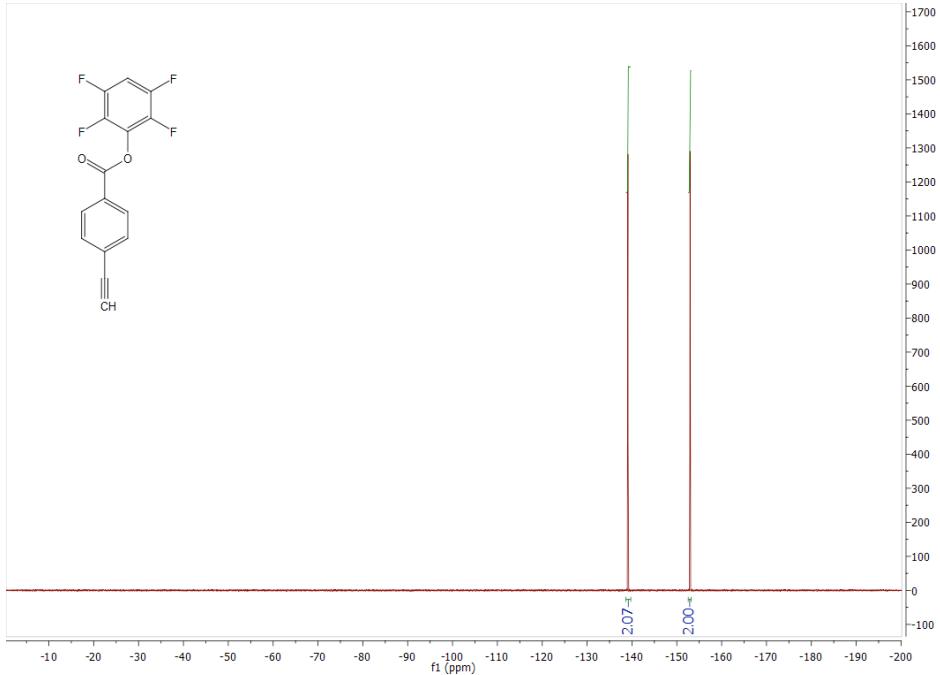
<sup>1</sup>H-NMR for 11



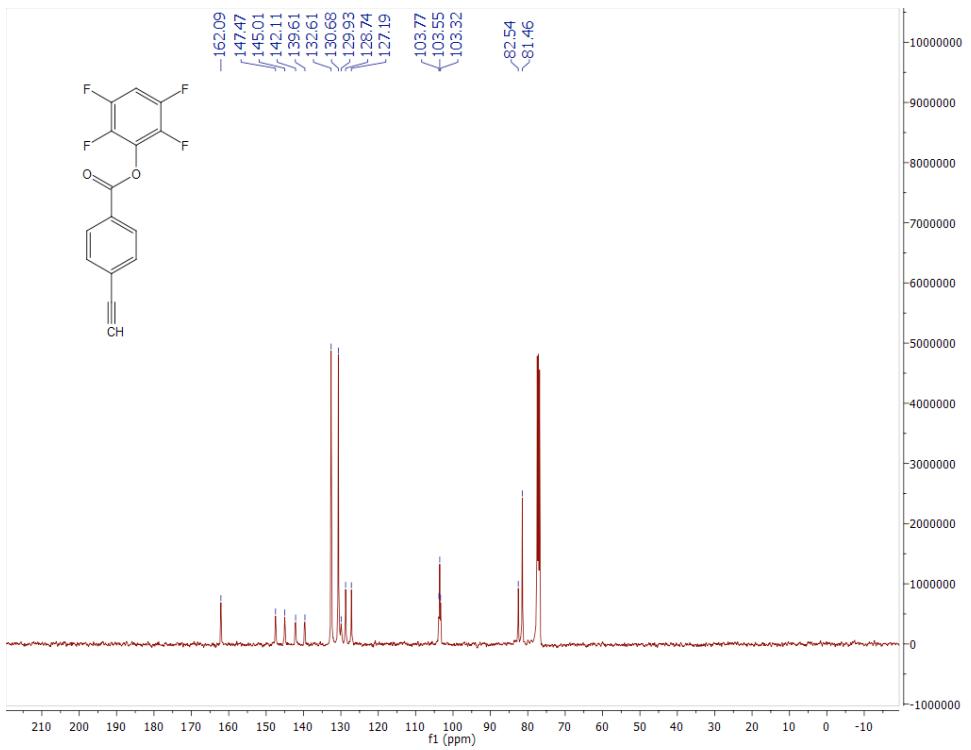
**<sup>1</sup>H-NMR for 12**



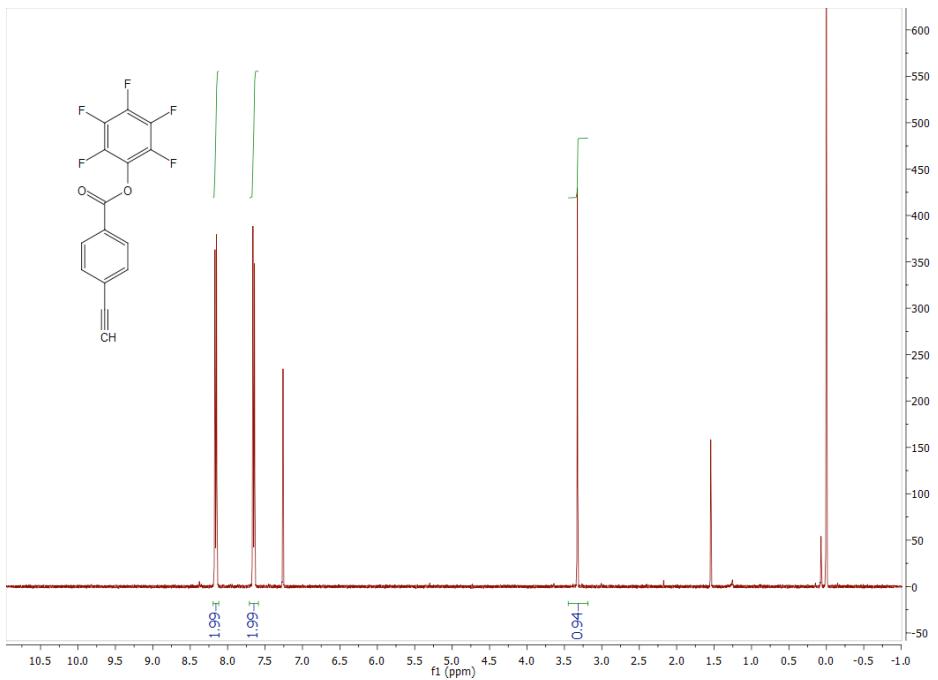
**<sup>19</sup>F-NMR for 12**



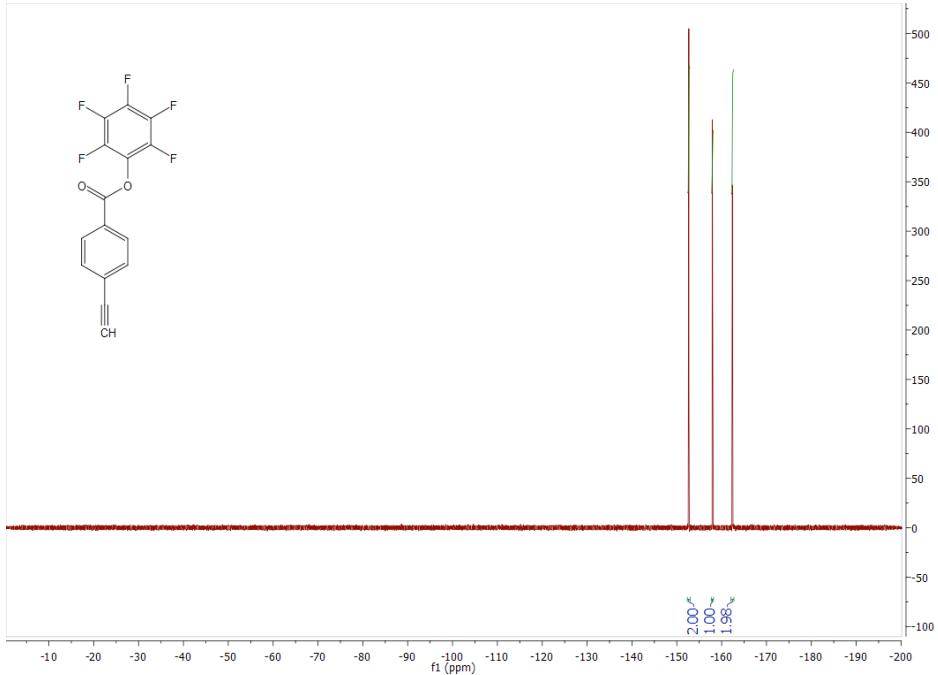
<sup>13</sup>C-NMR for 12



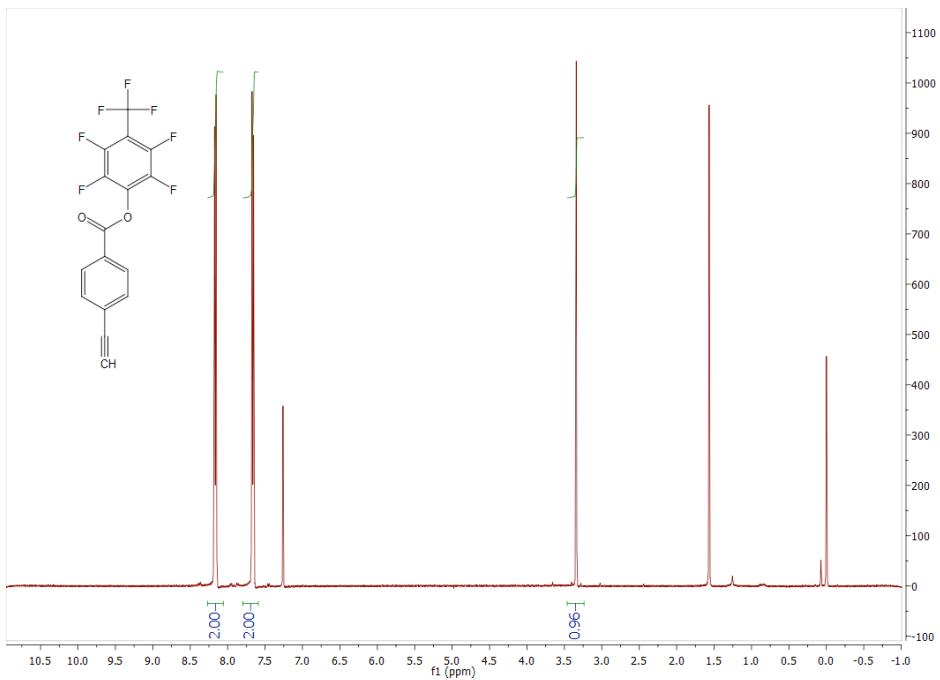
**<sup>1</sup>H-NMR for 13**



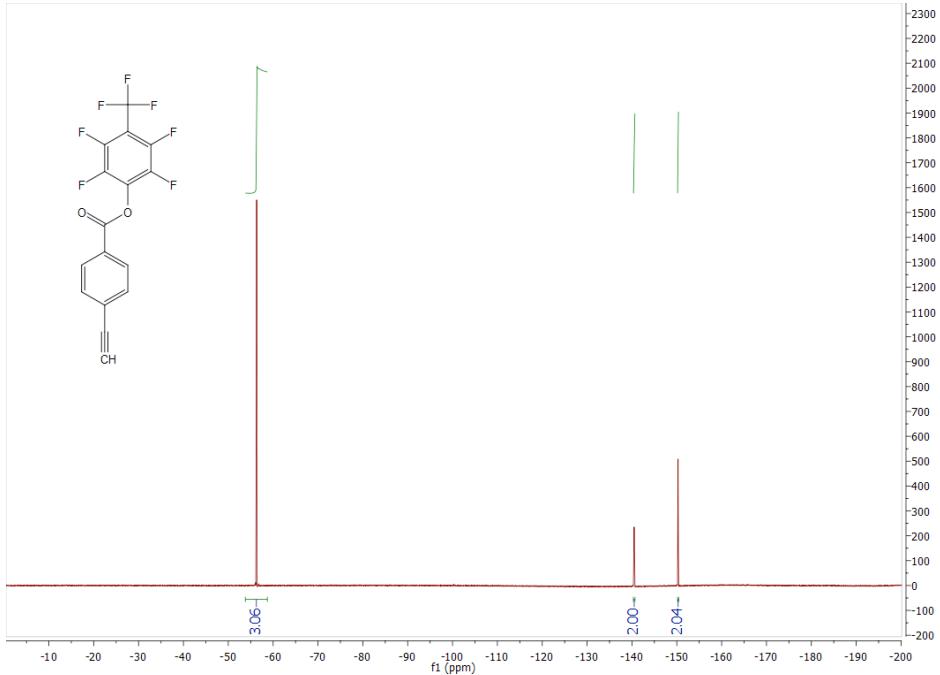
**<sup>19</sup>F-NMR for 13**



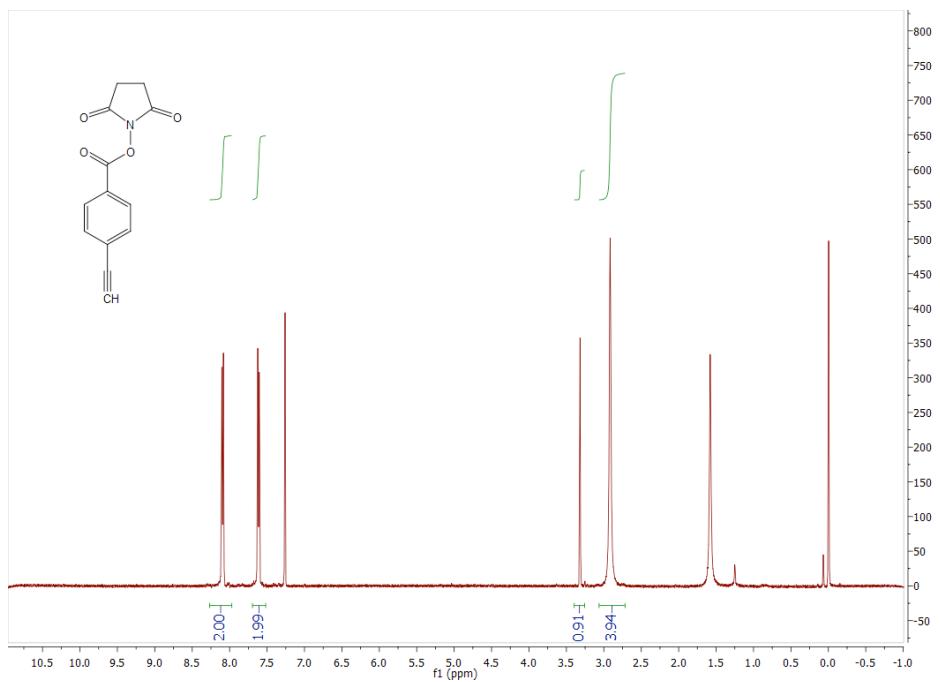
**<sup>1</sup>H-NMR for 14**



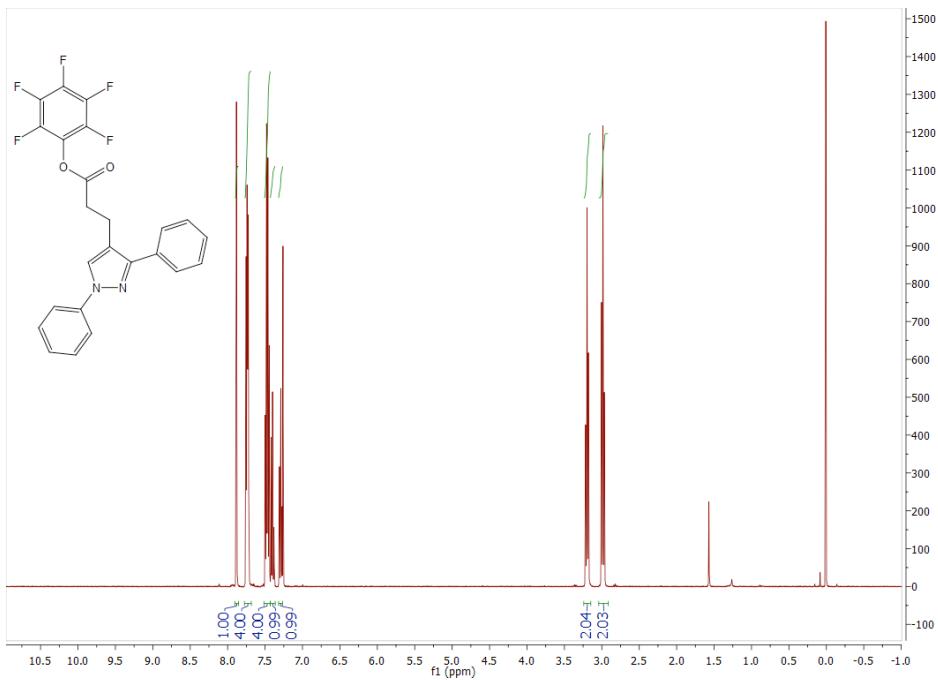
**<sup>19</sup>F-NMR for 14**



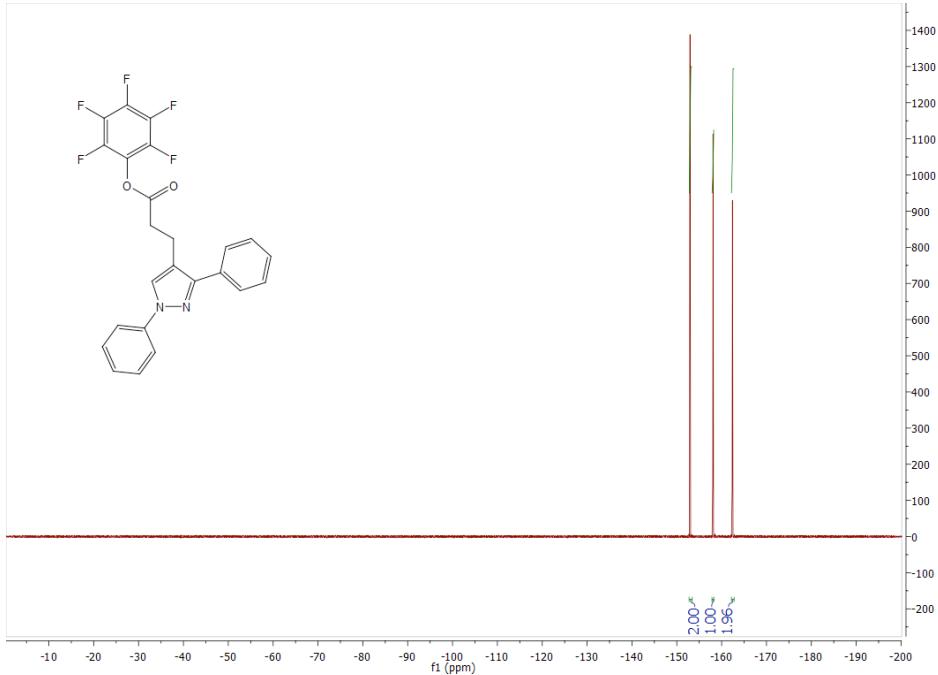
<sup>1</sup>H-NMR for 15



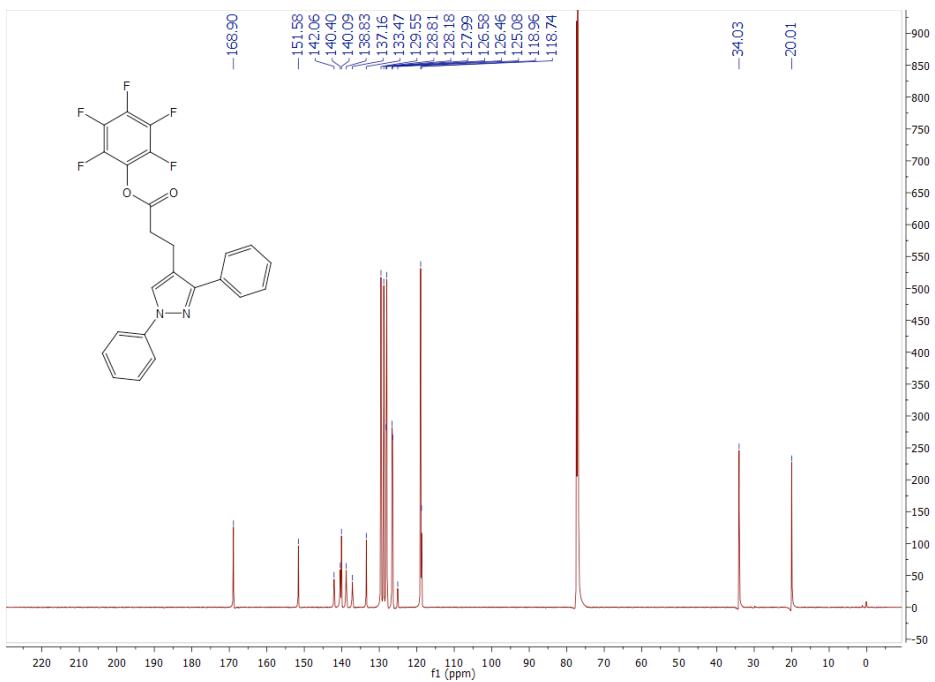
**<sup>1</sup>H-NMR for 19**



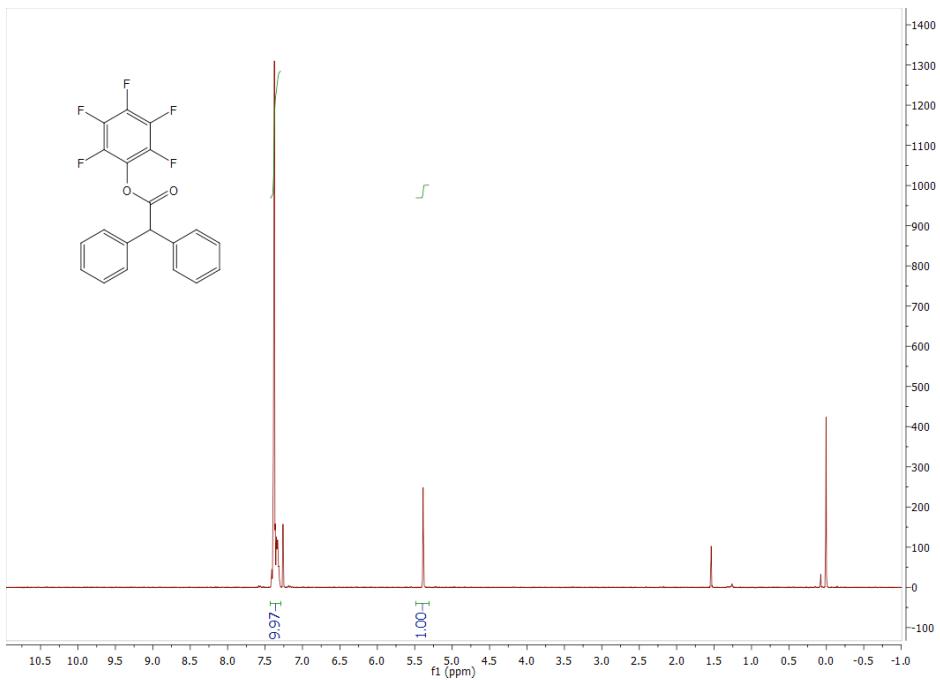
**<sup>19</sup>F-NMR for 19**



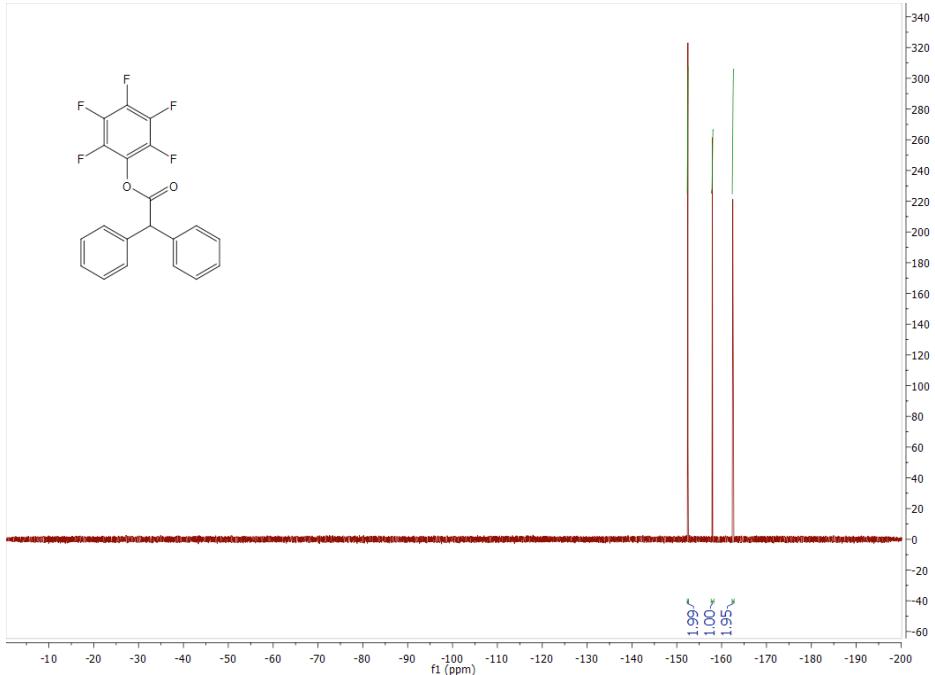
<sup>13</sup>C-NMR for 19



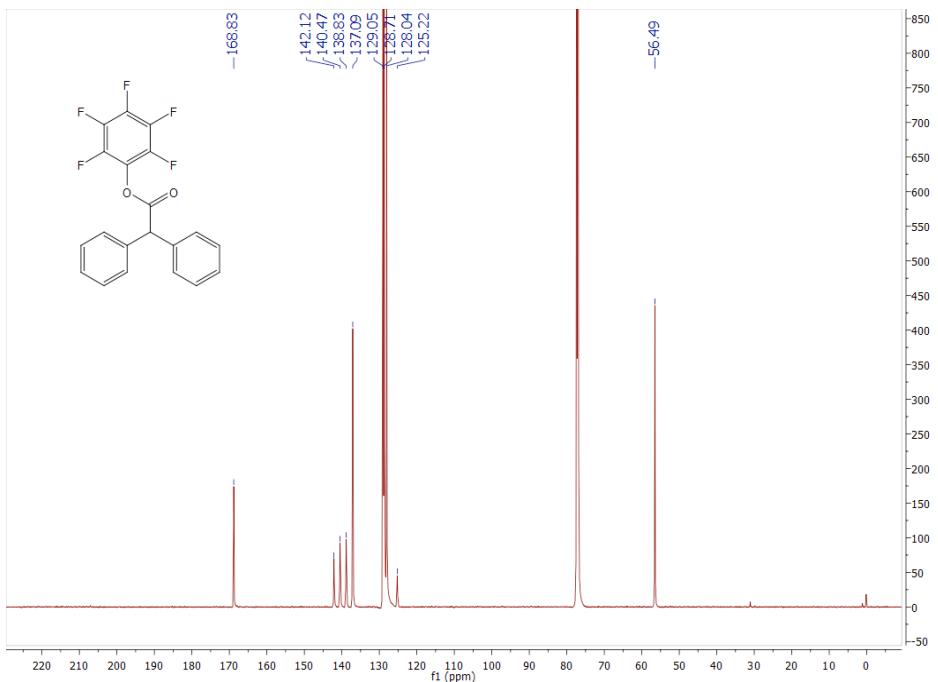
**<sup>1</sup>H-NMR for 20**



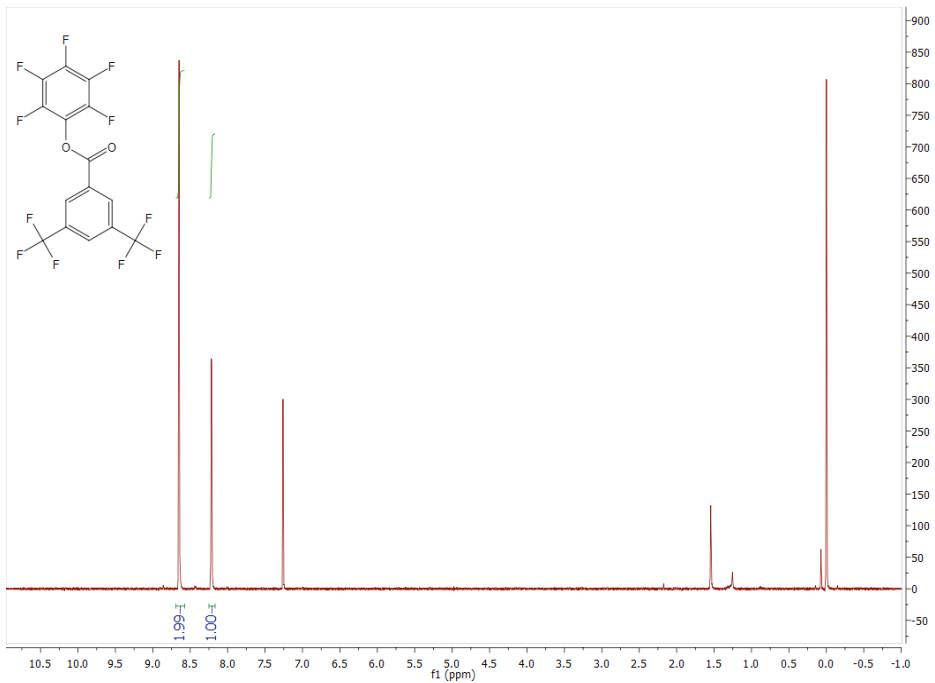
**<sup>19</sup>F-NMR for 20**



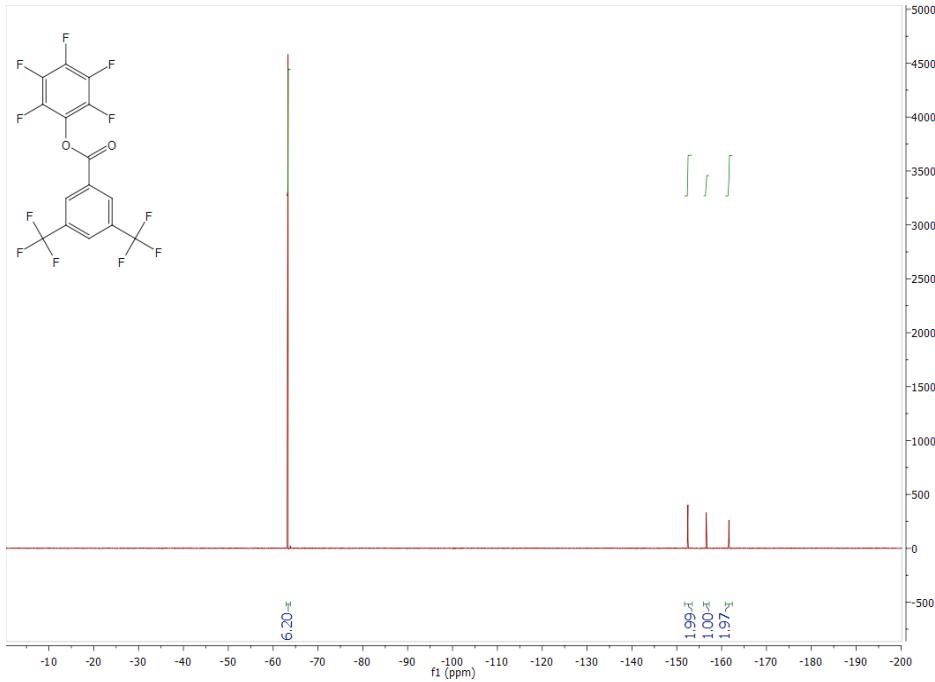
<sup>13</sup>C-NMR for 20



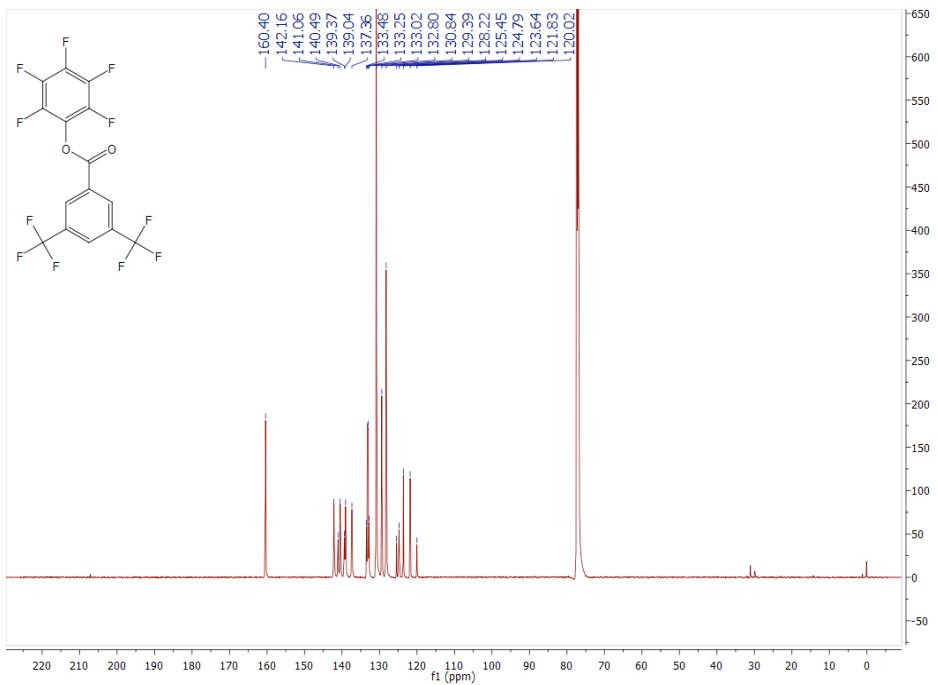
**<sup>1</sup>H-NMR for 21**



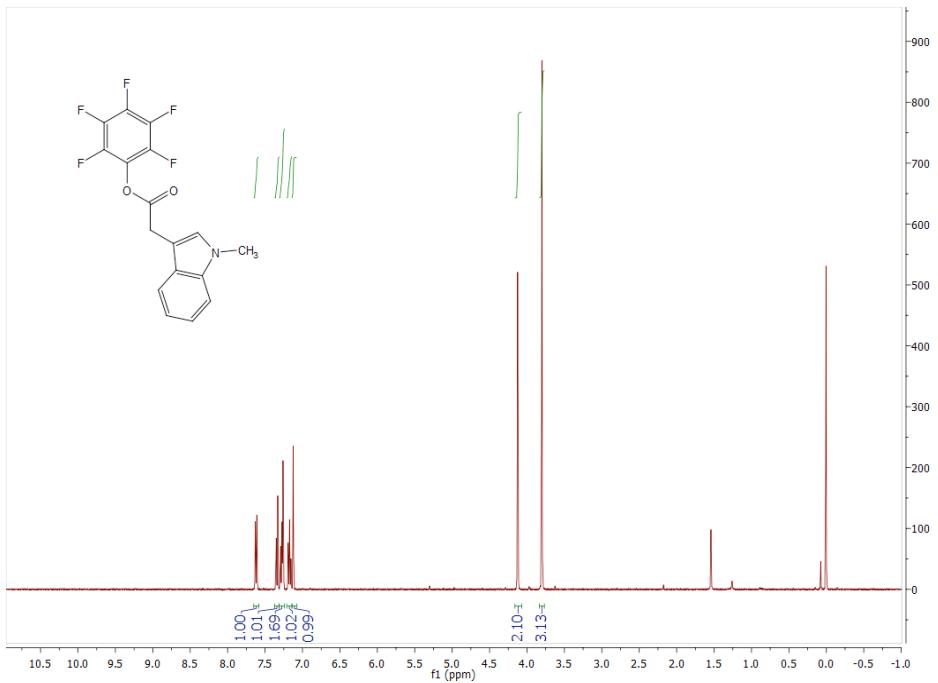
**<sup>19</sup>F-NMR for 21**



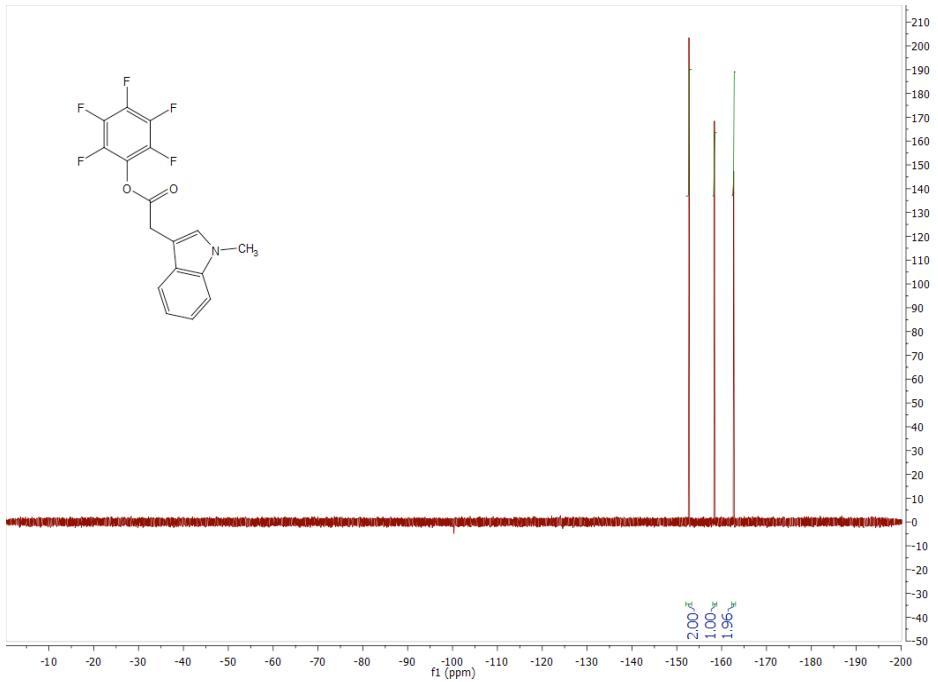
**<sup>13</sup>C-NMR for 21**



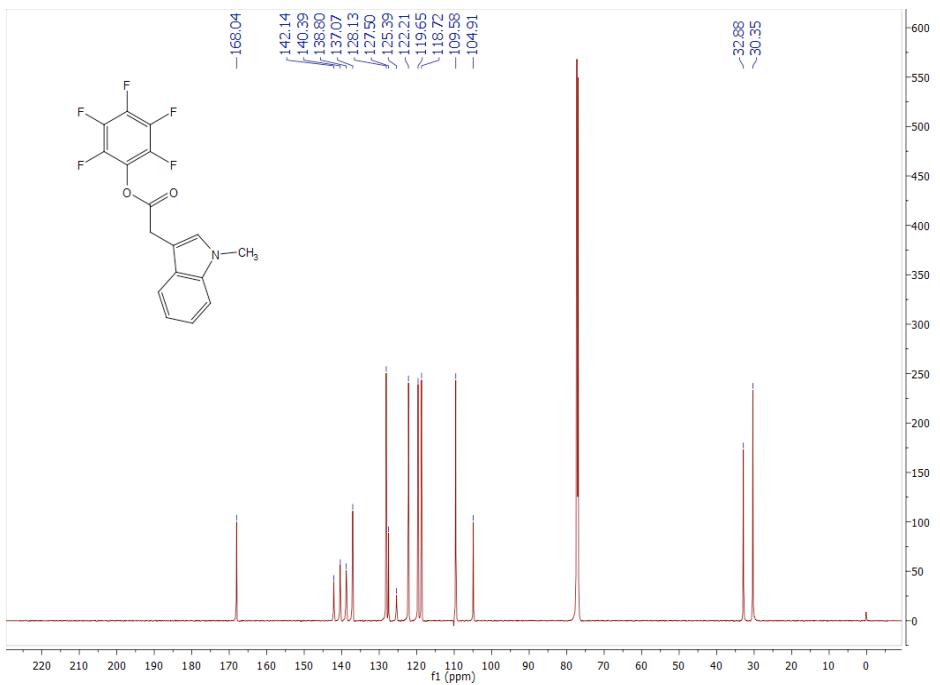
**<sup>1</sup>H-NMR for 22**



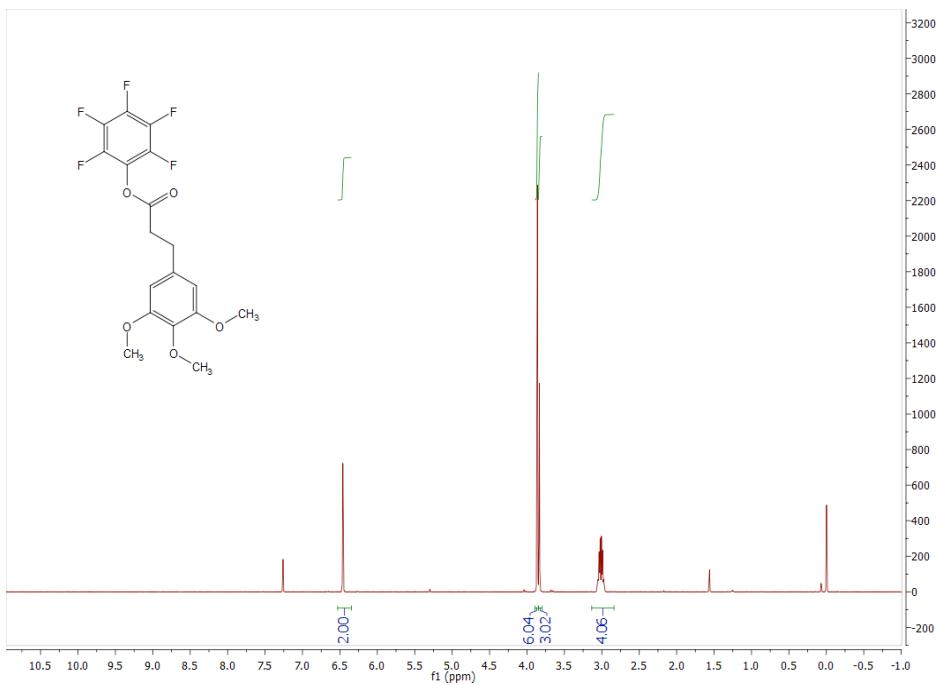
**<sup>19</sup>F-NMR for 22**



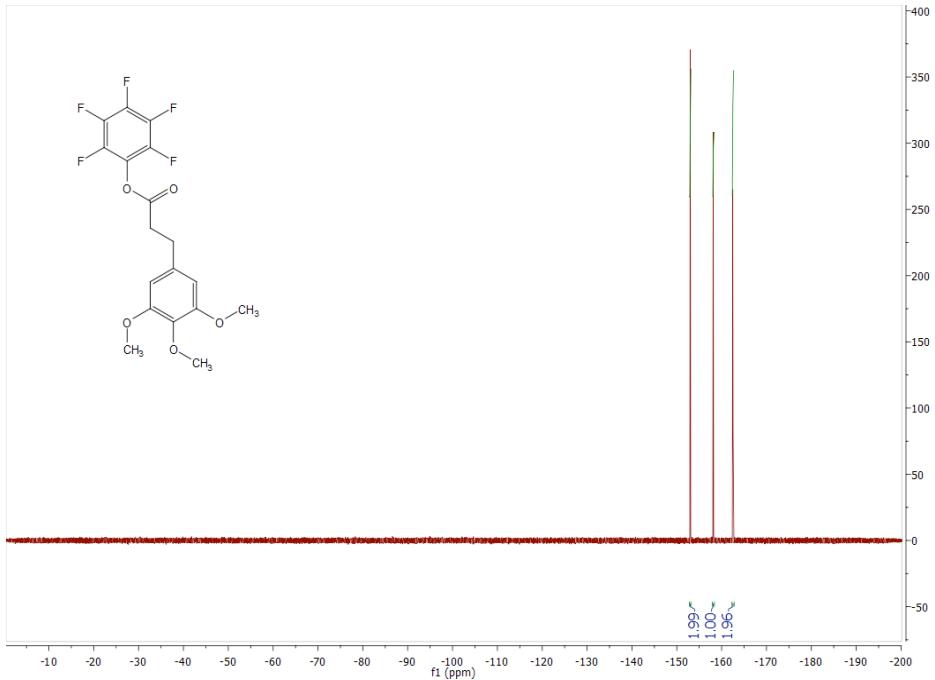
**<sup>13</sup>C-NMR for 22**



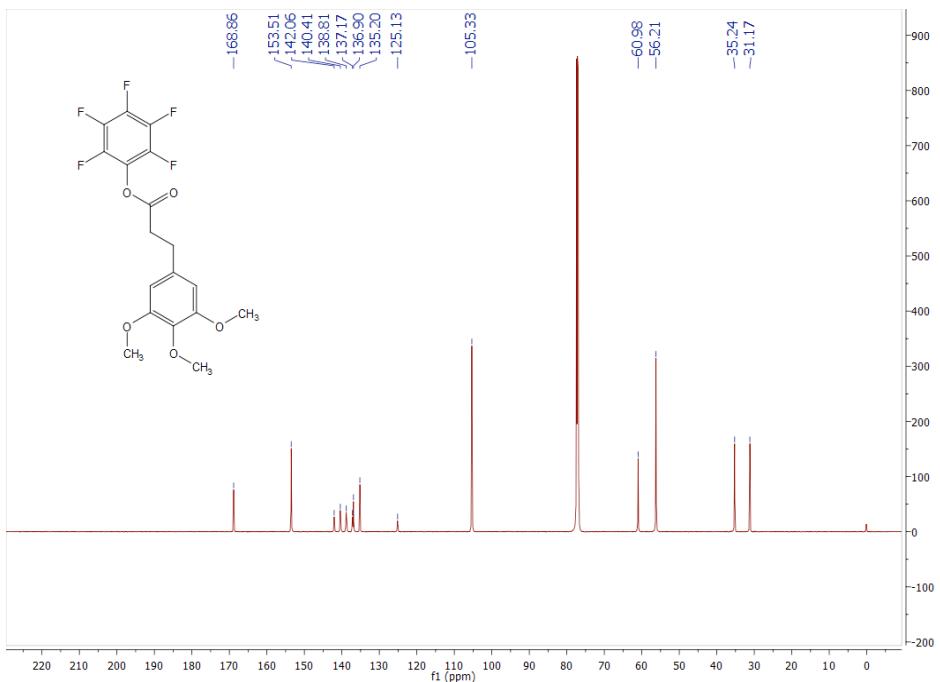
**<sup>1</sup>H-NMR for 23**



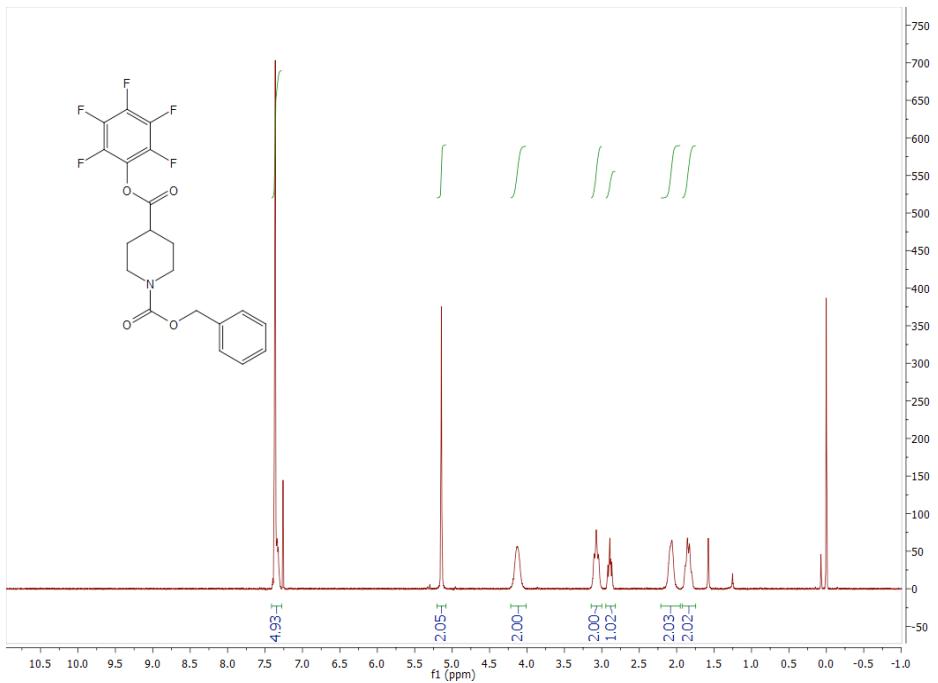
**<sup>19</sup>F-NMR for 23**



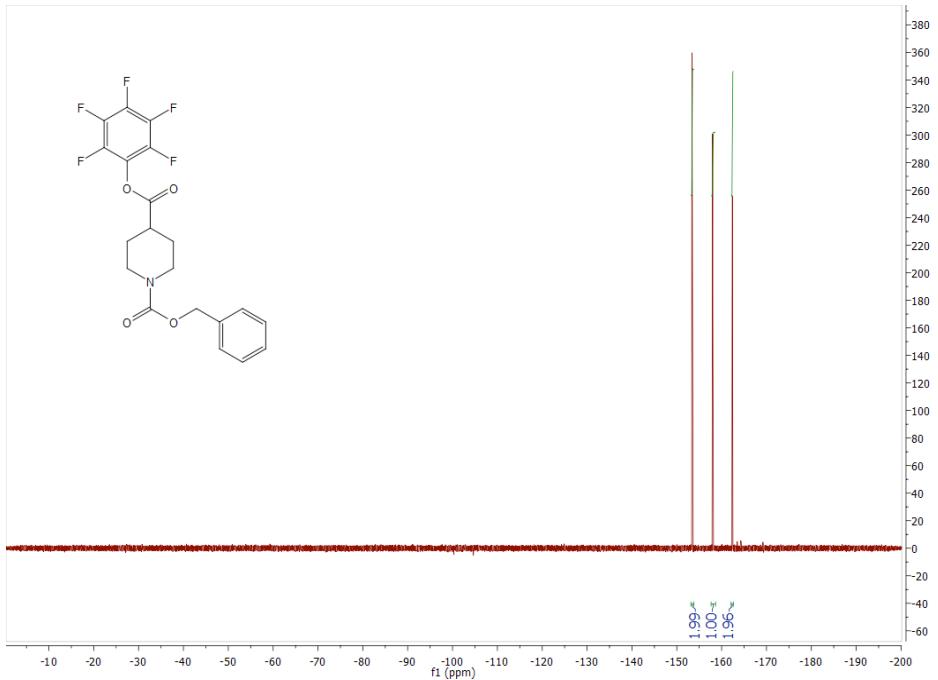
**<sup>13</sup>C-NMR for 23**



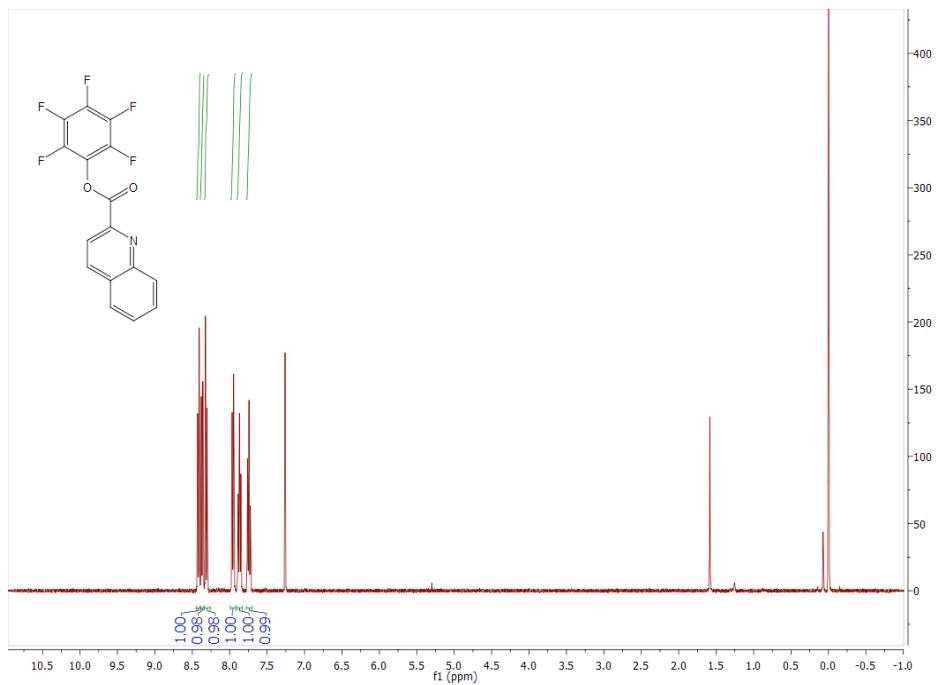
**<sup>1</sup>H-NMR for 24**



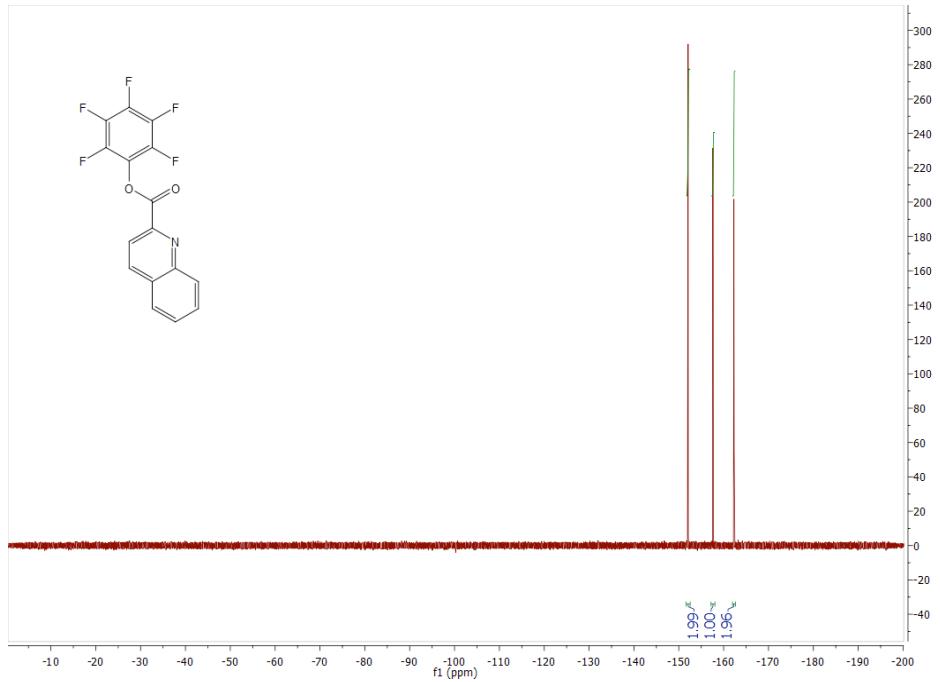
**<sup>19</sup>F-NMR for 24**



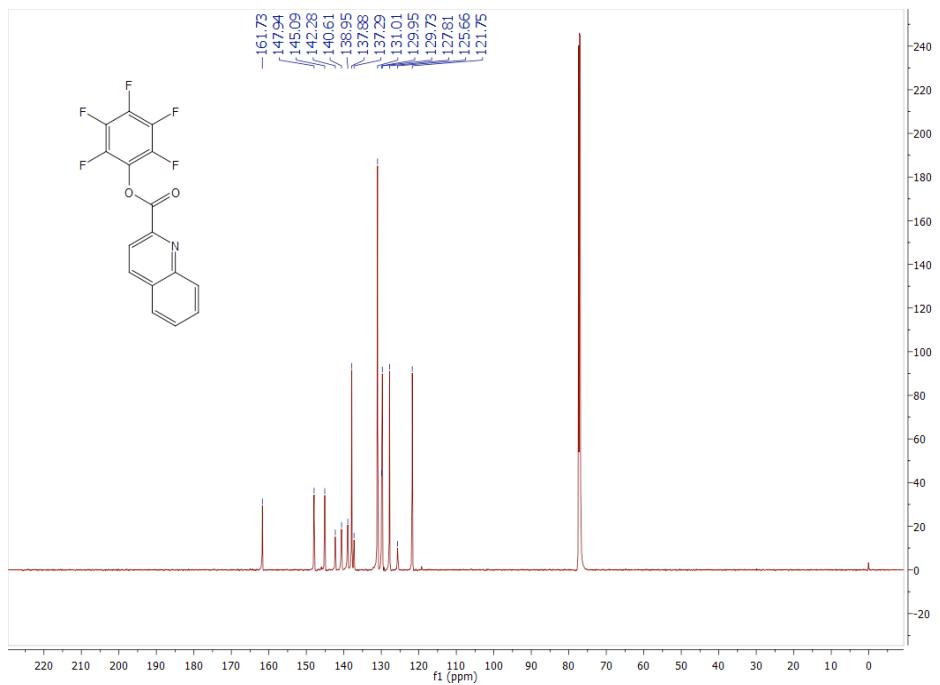
**<sup>1</sup>H-NMR for 25**



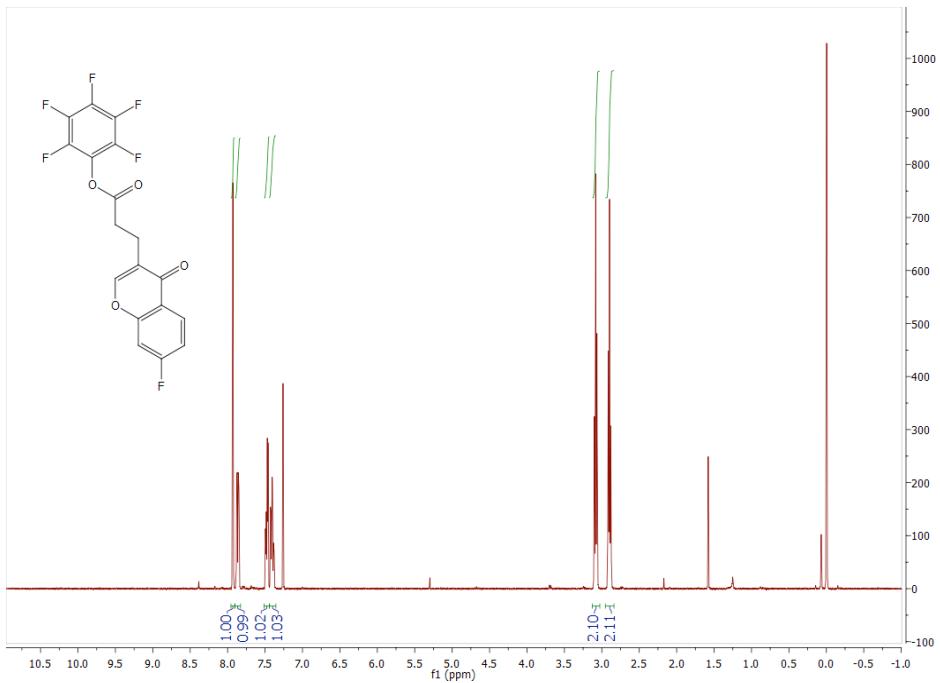
**<sup>19</sup>F-NMR for 25**



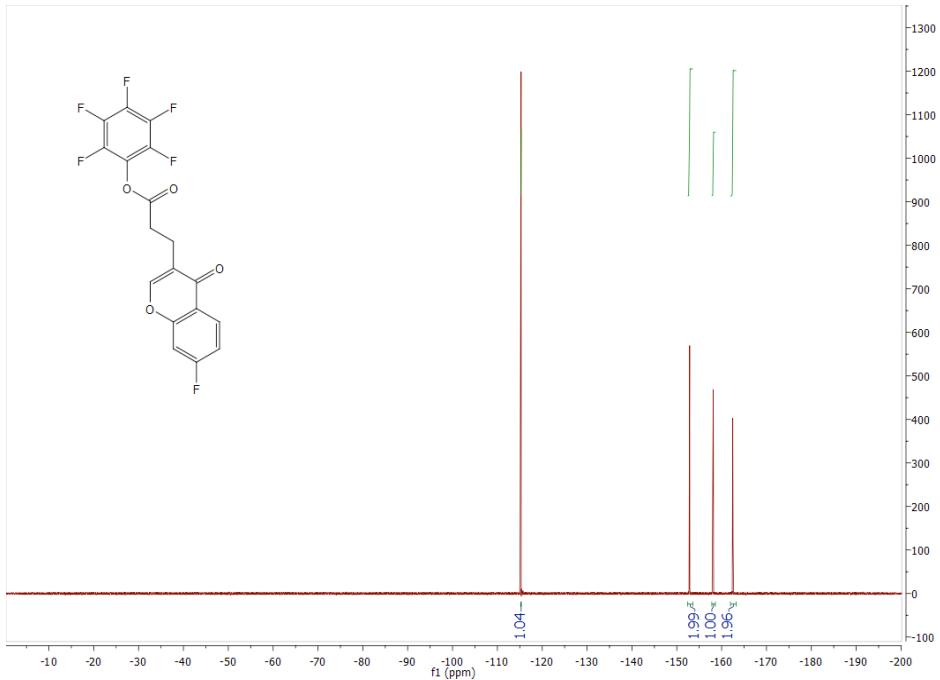
**<sup>13</sup>C-NMR for 25**



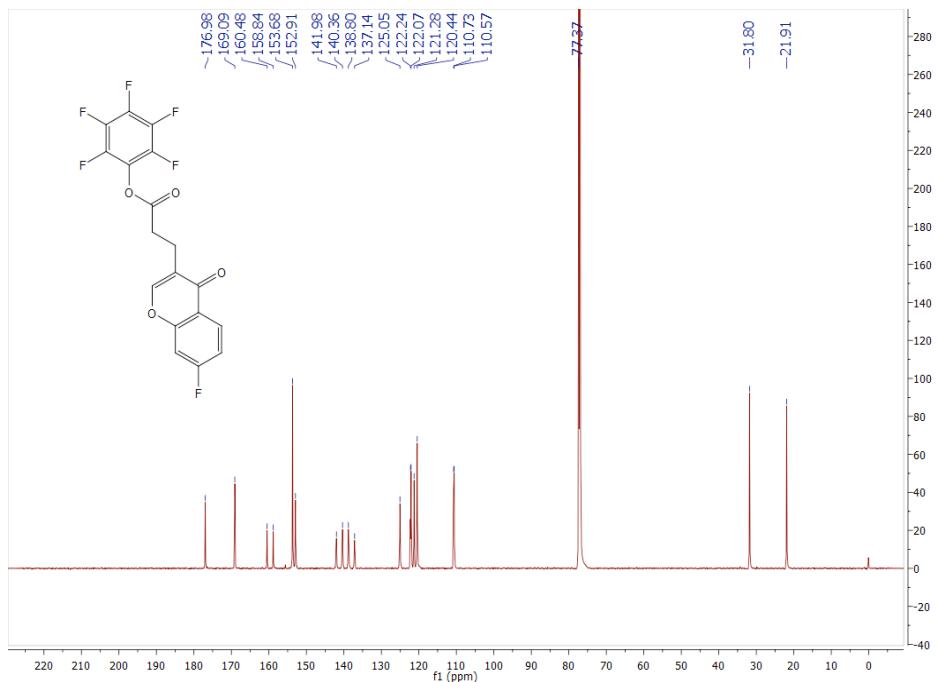
**<sup>1</sup>H-NMR for 26**



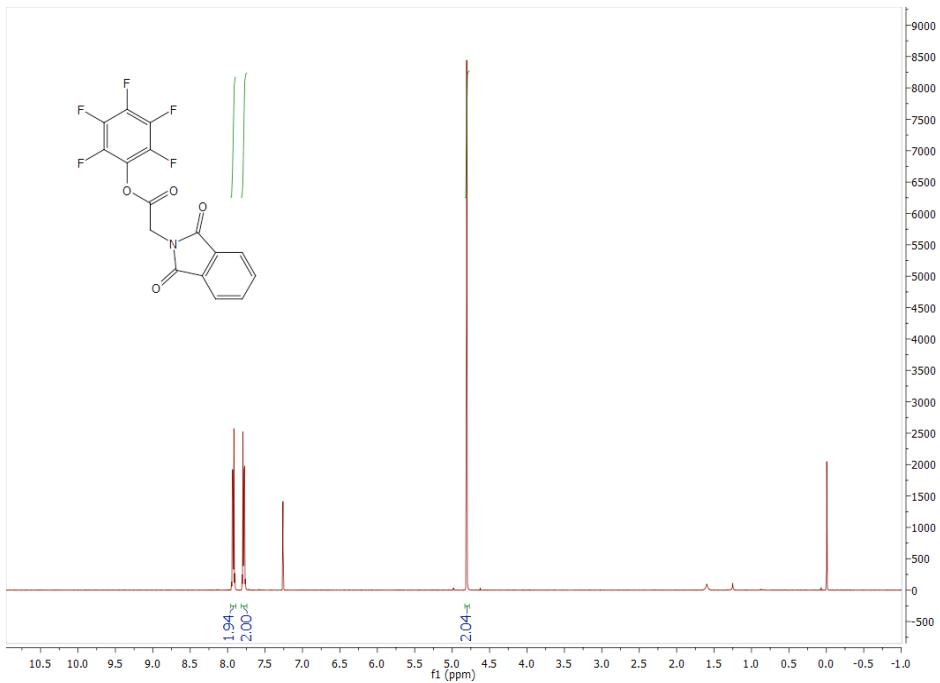
**<sup>19</sup>F-NMR for 26**



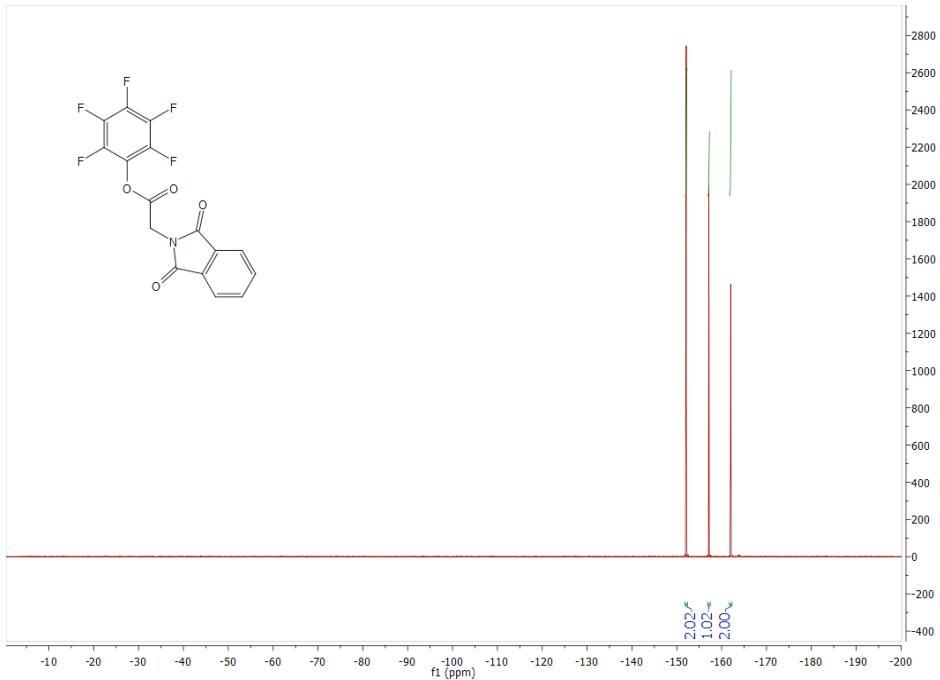
**<sup>13</sup>C-NMR for 26**



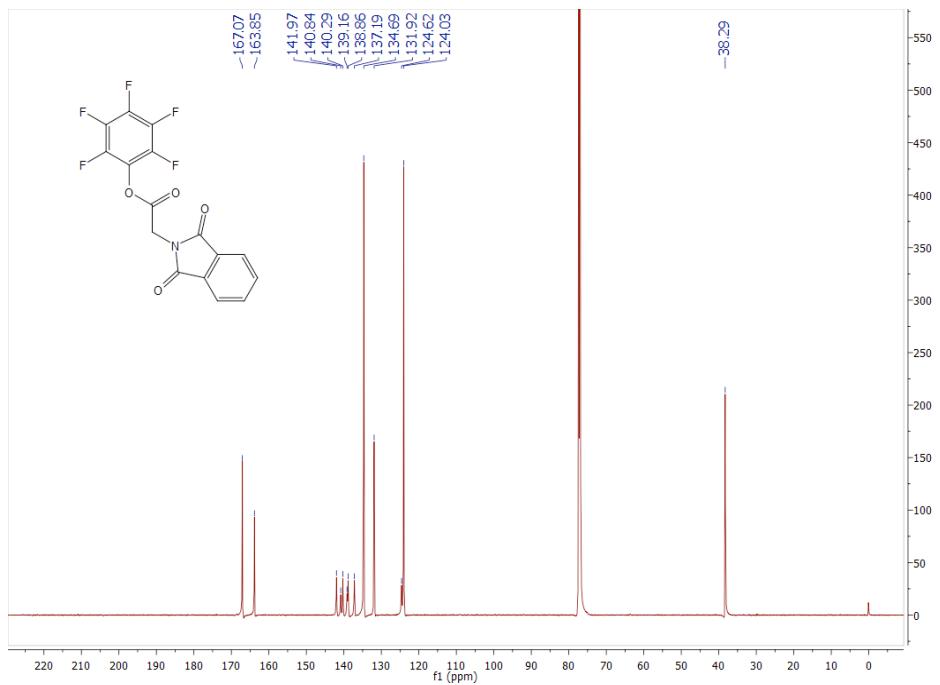
**<sup>1</sup>H-NMR for 27**



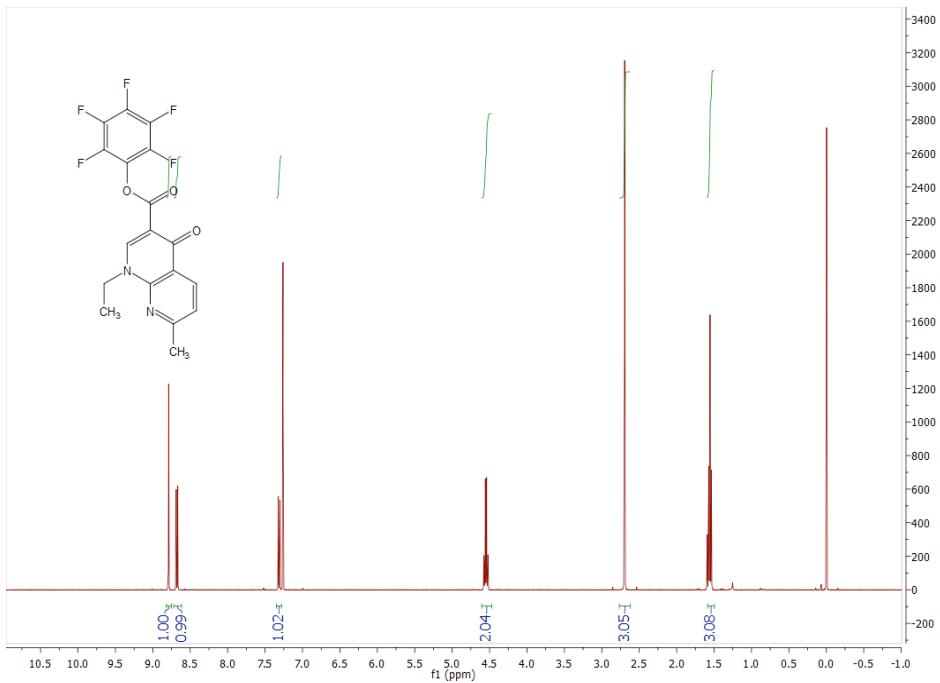
**<sup>19</sup>F-NMR for 27**



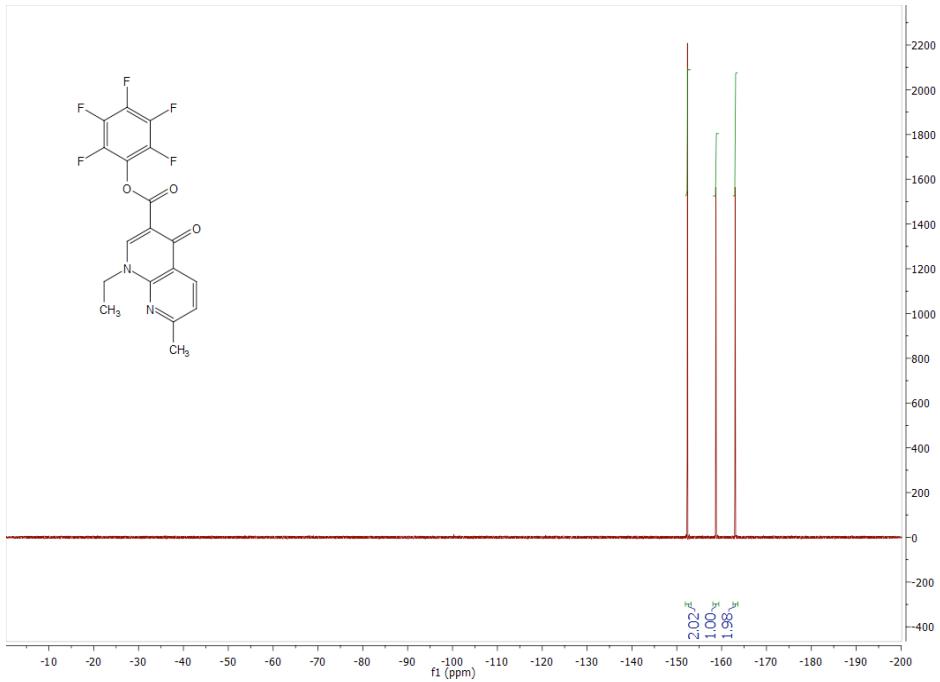
<sup>13</sup>C-NMR for 27



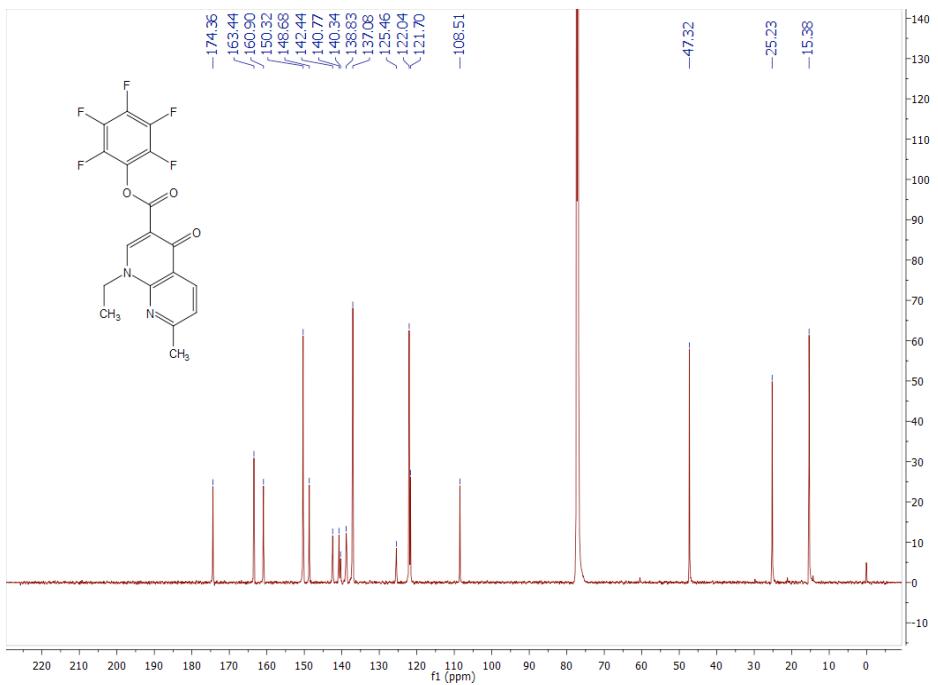
**<sup>1</sup>H-NMR for 28**



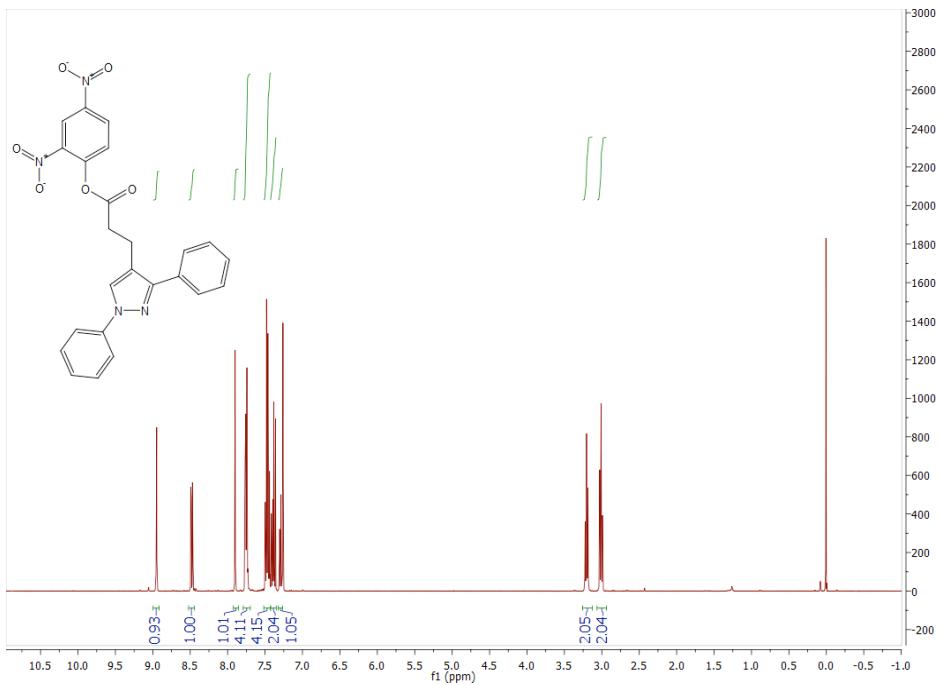
**<sup>19</sup>F-NMR for 28**



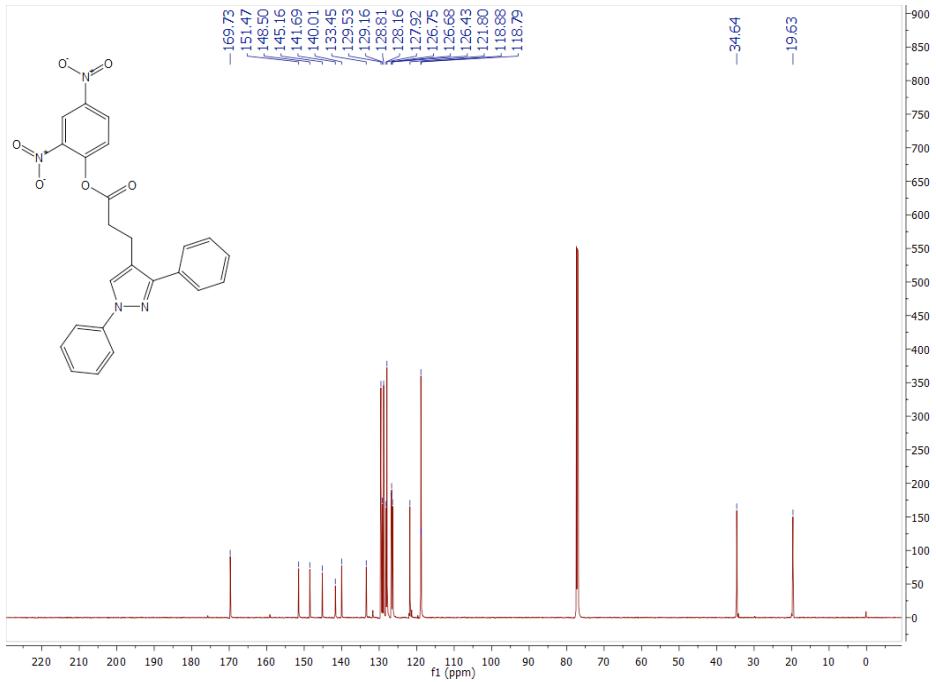
**<sup>13</sup>C-NMR for 28**



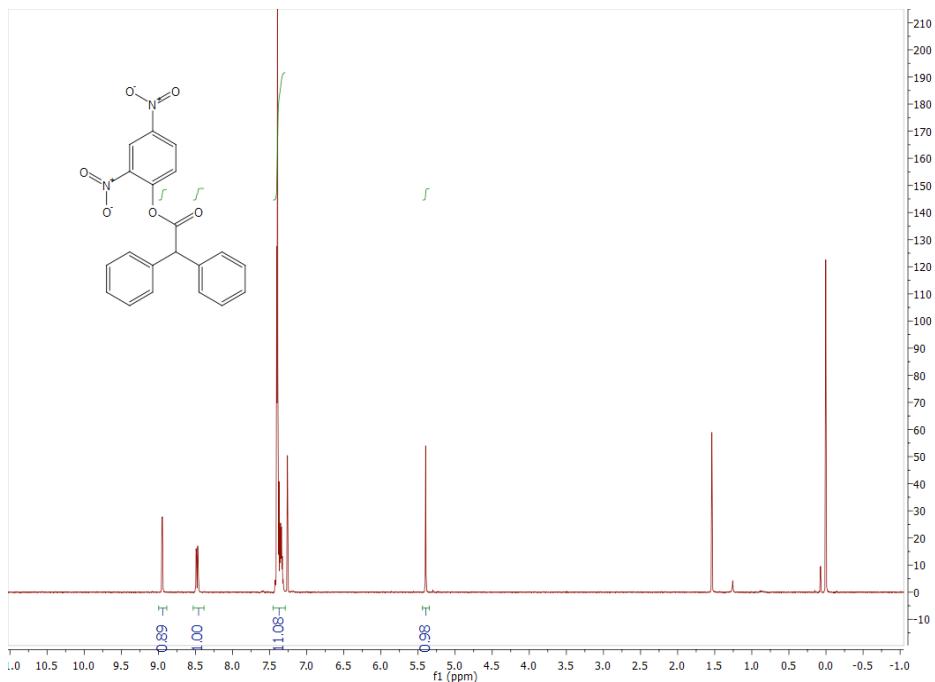
**<sup>1</sup>H-NMR for 29**



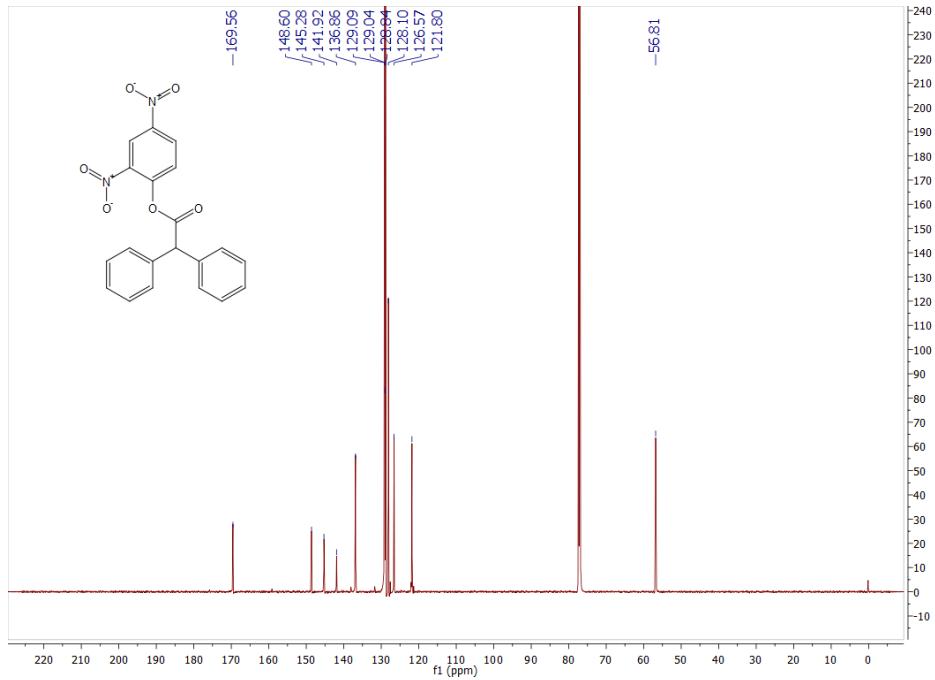
**<sup>13</sup>C-NMR for 29**



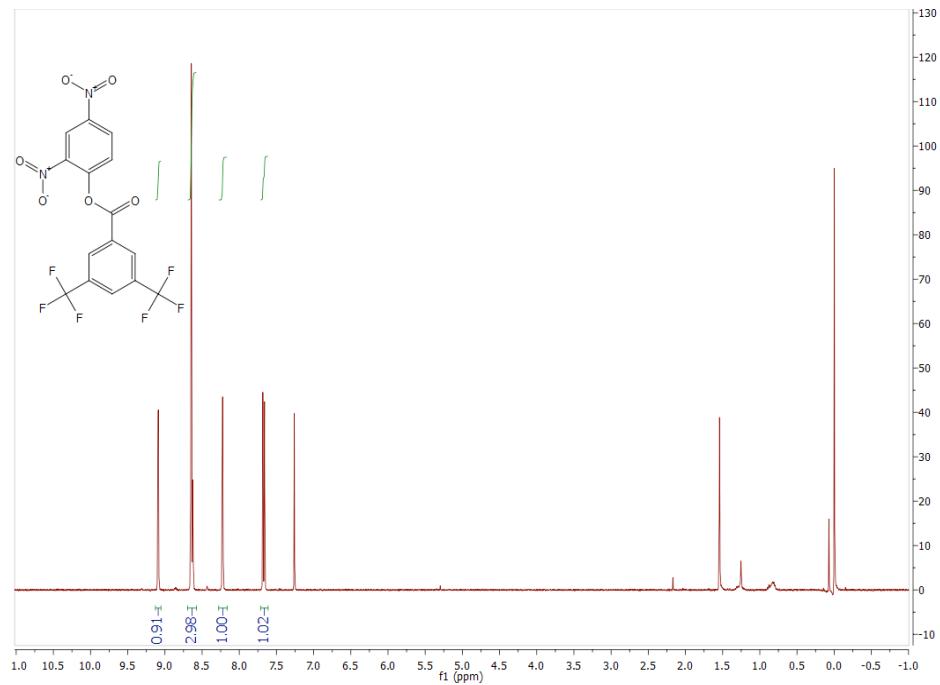
**<sup>1</sup>H-NMR for 30**



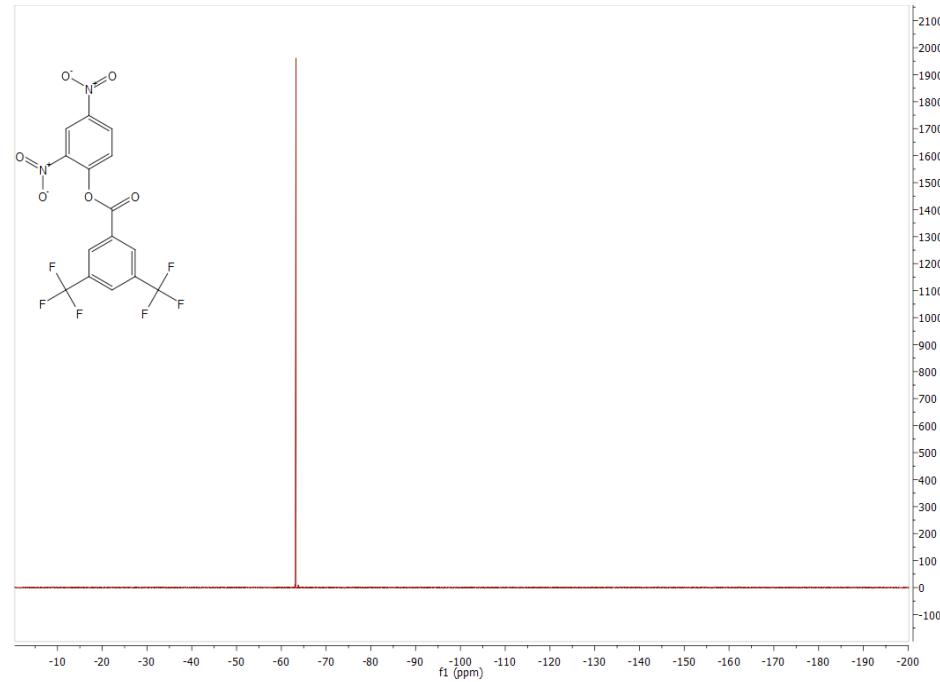
**<sup>13</sup>C-NMR for 30**



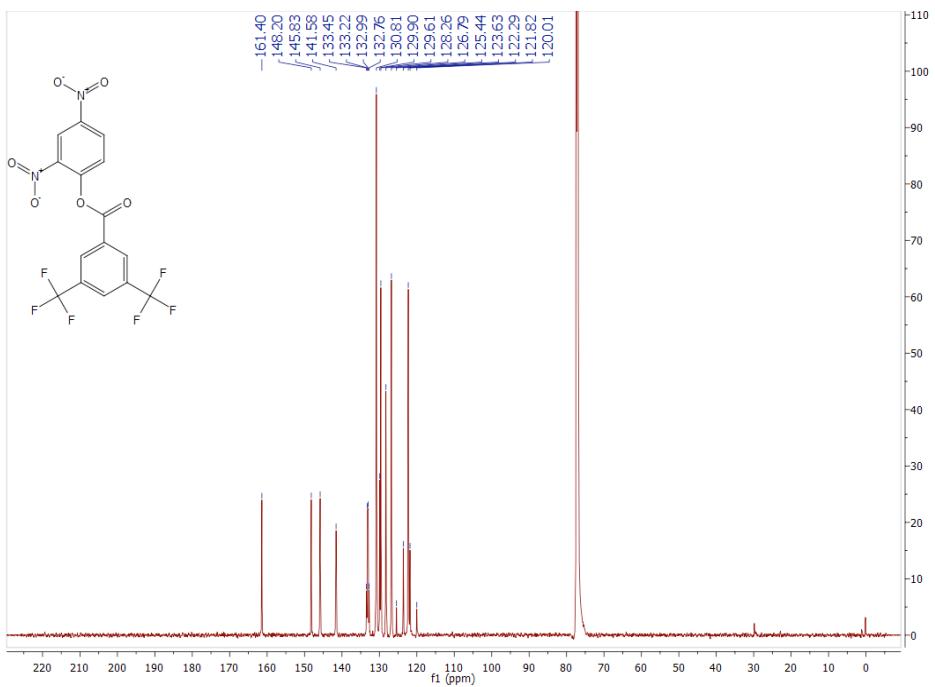
**<sup>1</sup>H-NMR for 31**



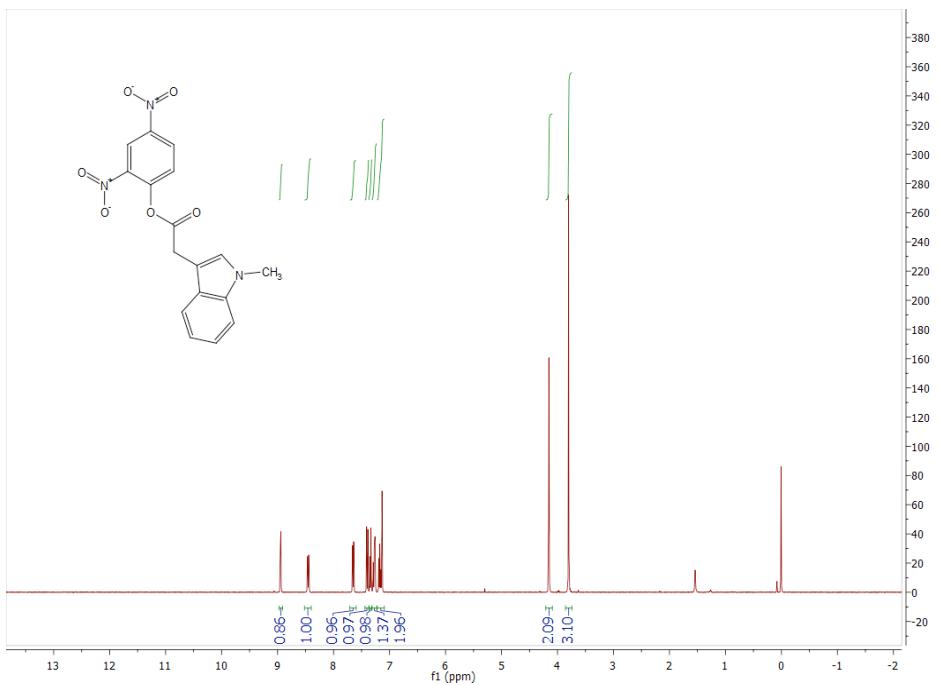
**<sup>19</sup>F-NMR for 31**



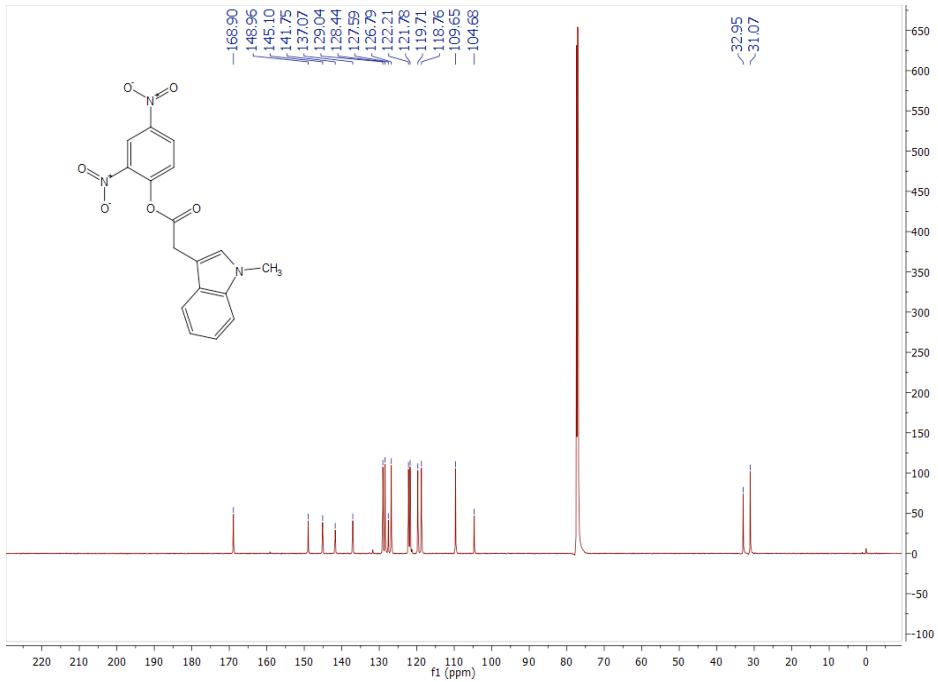
<sup>13</sup>C-NMR for 31



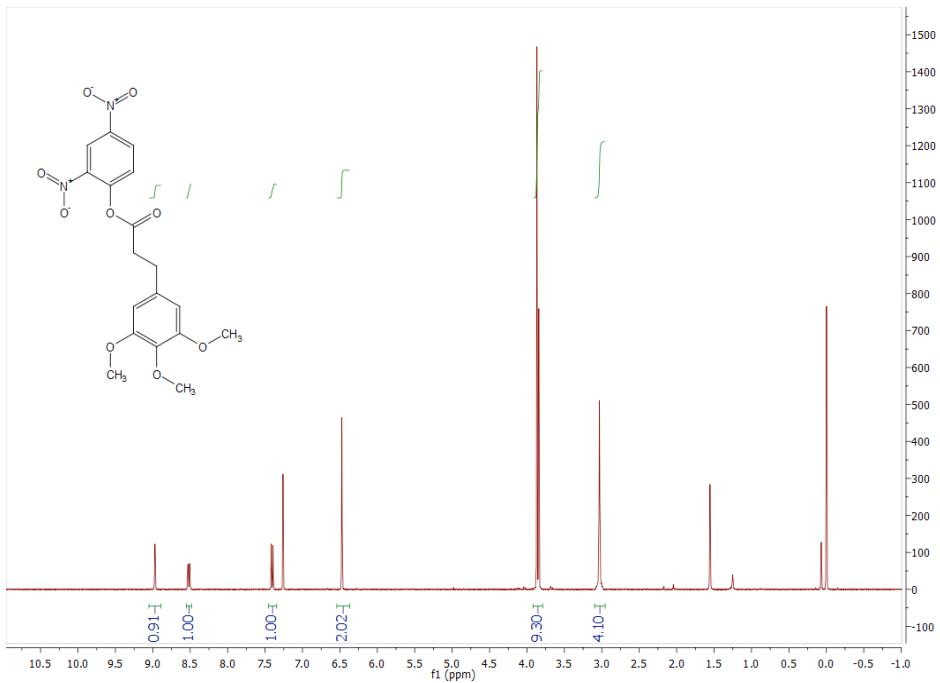
**<sup>1</sup>H-NMR for 32**



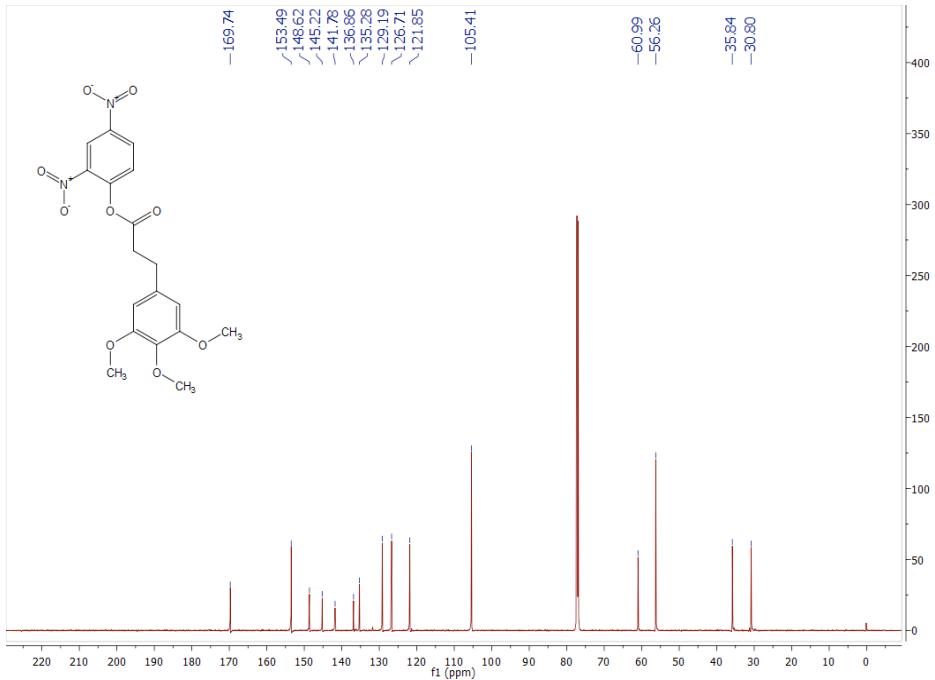
**<sup>13</sup>C-NMR for 32**



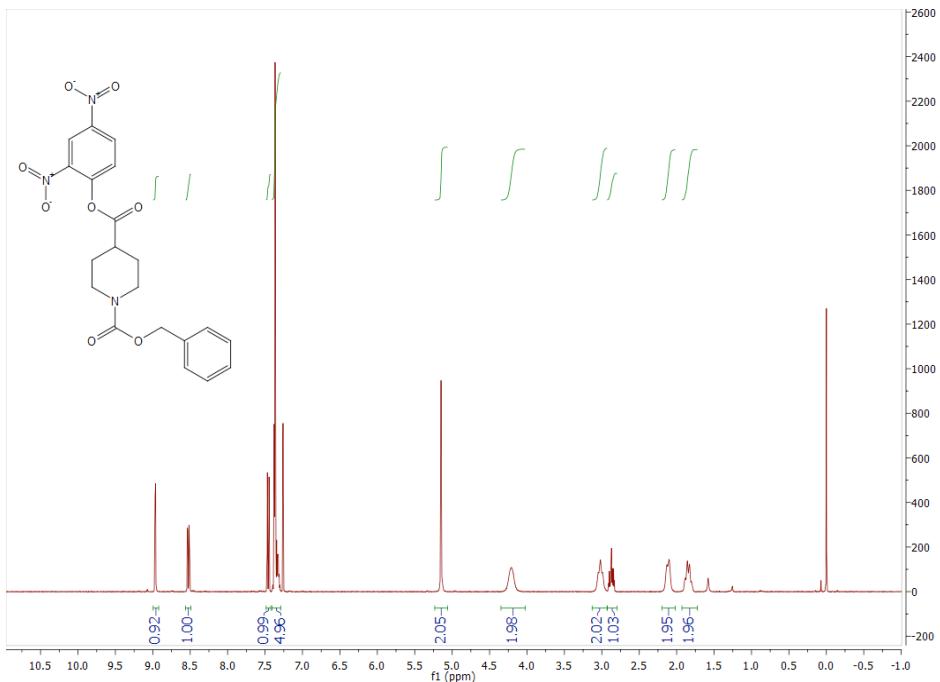
**<sup>1</sup>H-NMR for 33**



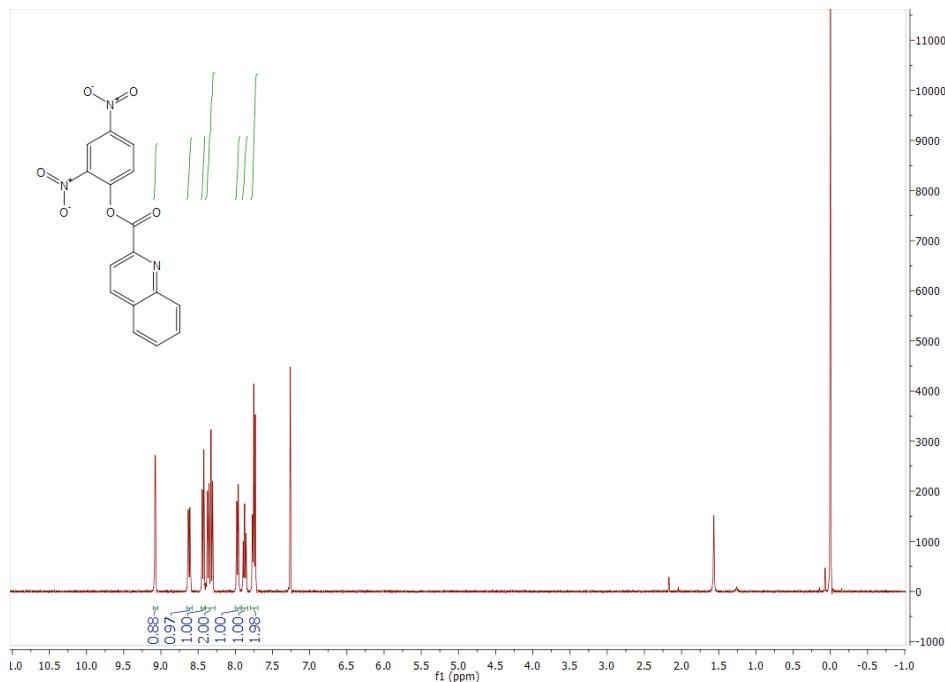
**<sup>13</sup>C-NMR for 33**



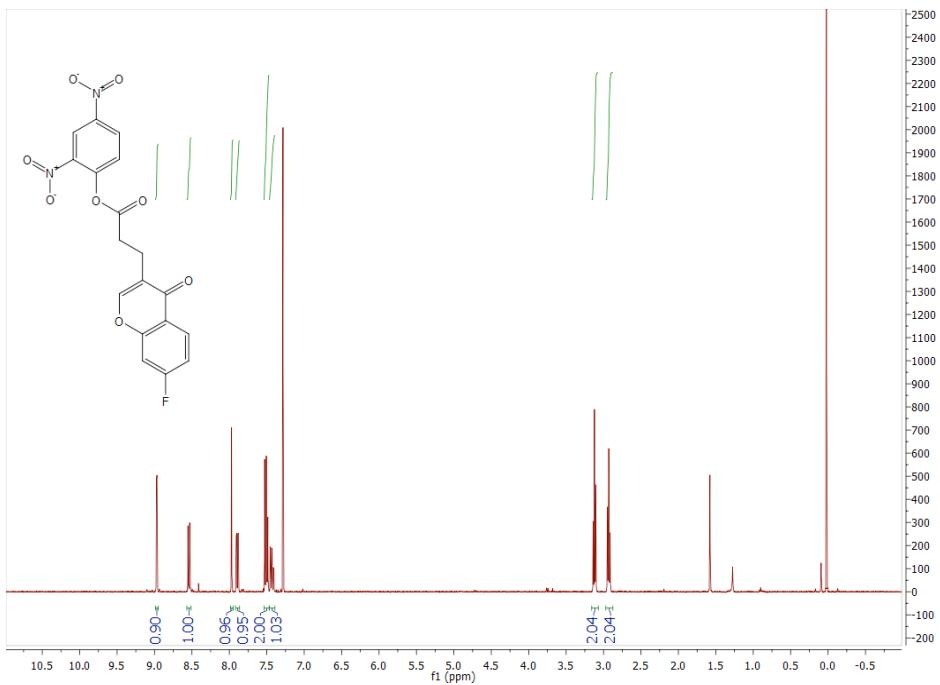
<sup>1</sup>H-NMR for 34



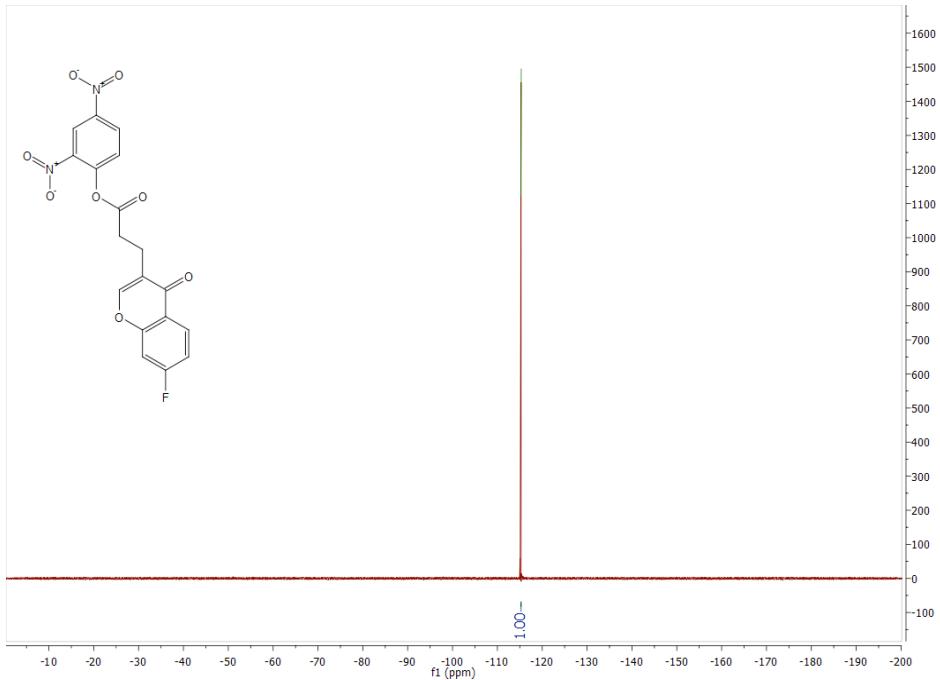
<sup>1</sup>H-NMR for 35



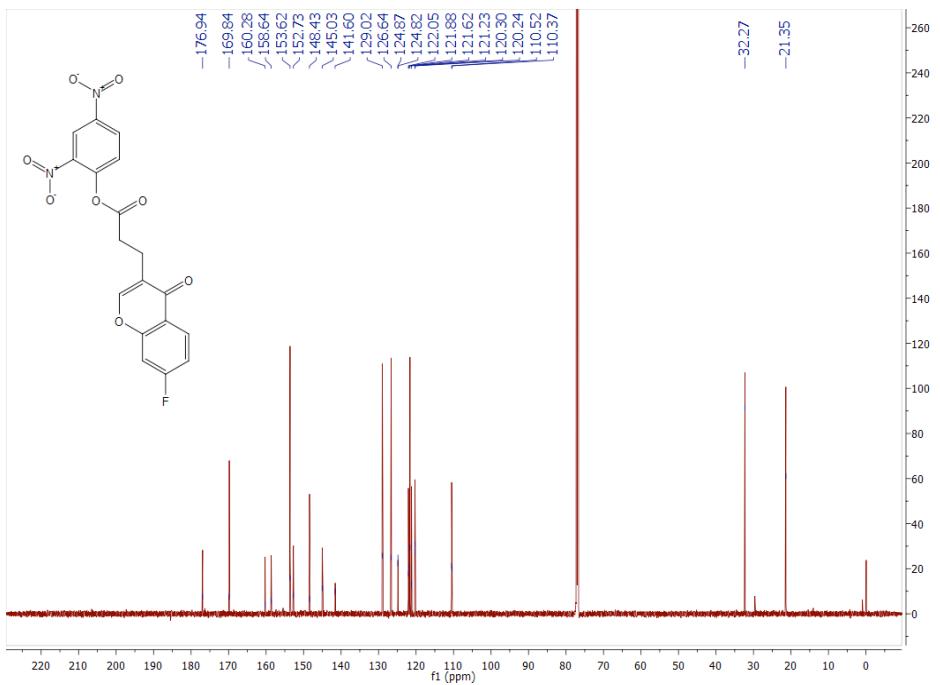
**<sup>1</sup>H-NMR for 36**



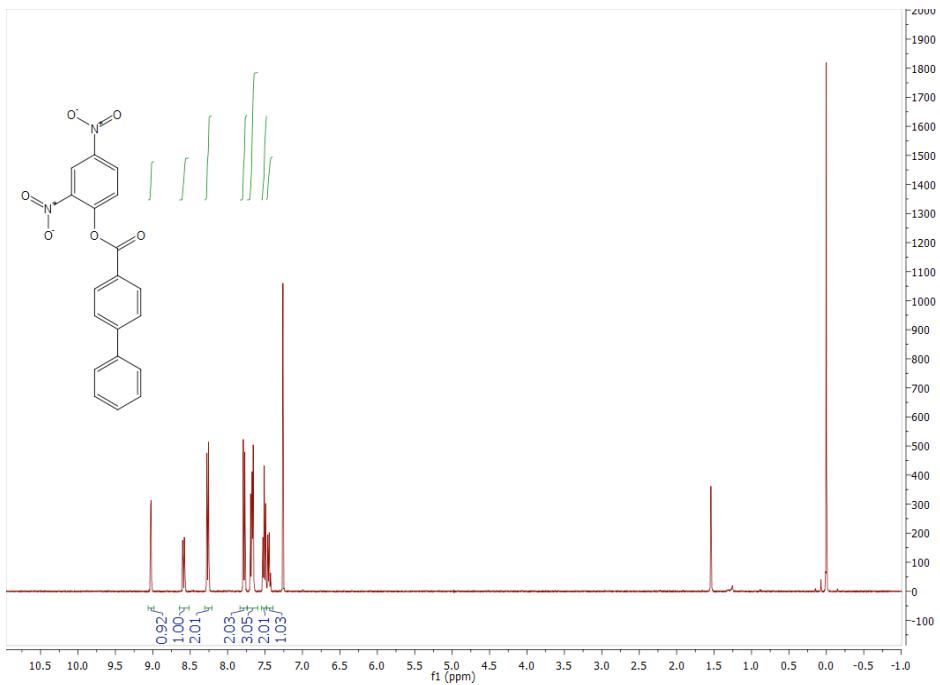
**<sup>19</sup>F-NMR for 36**



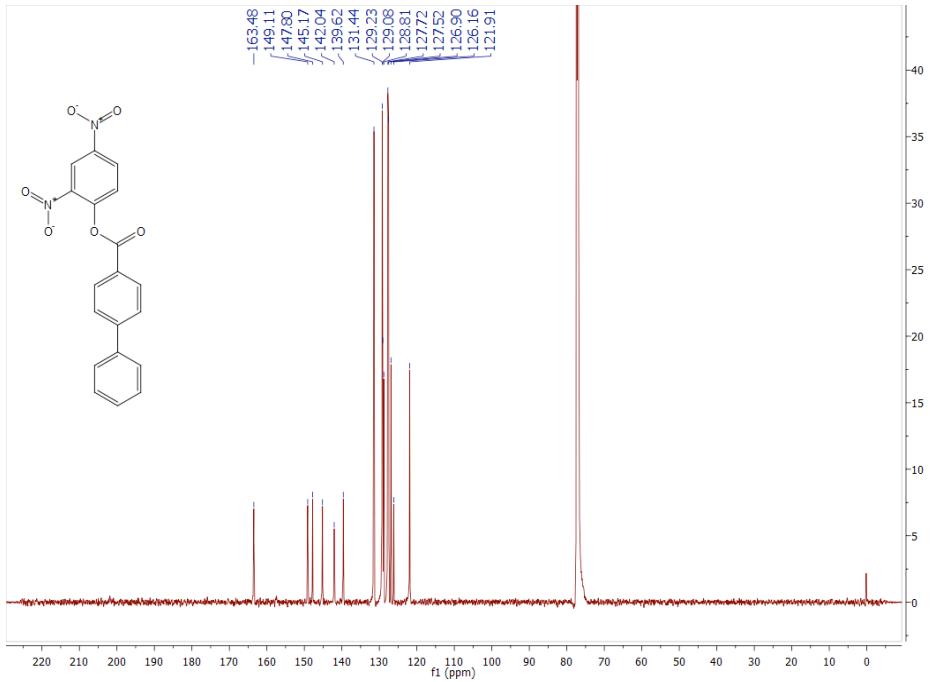
**<sup>13</sup>C-NMR for 36**



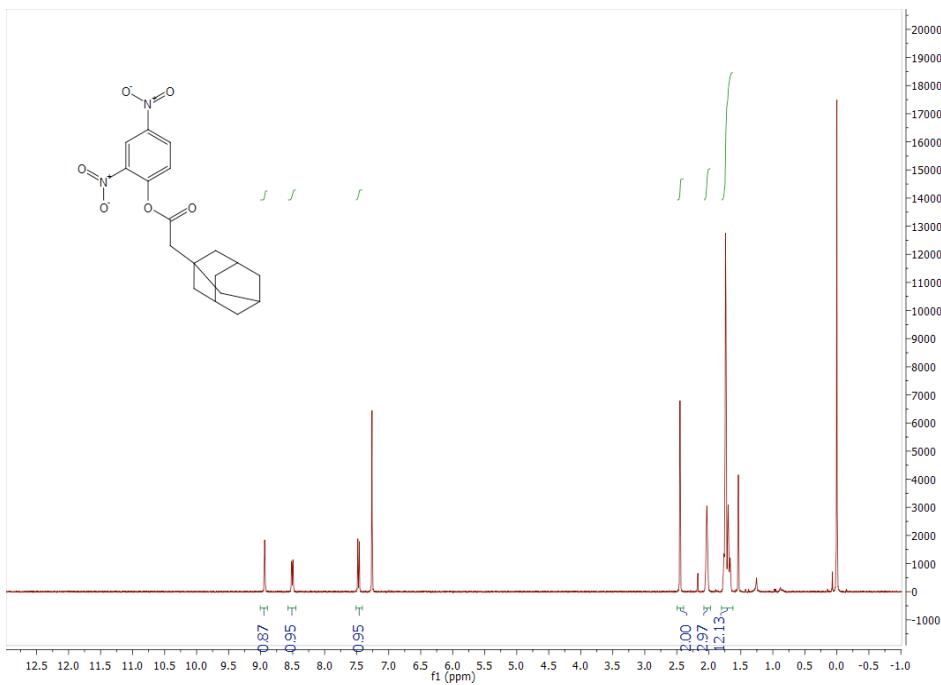
**<sup>1</sup>H-NMR for 37**



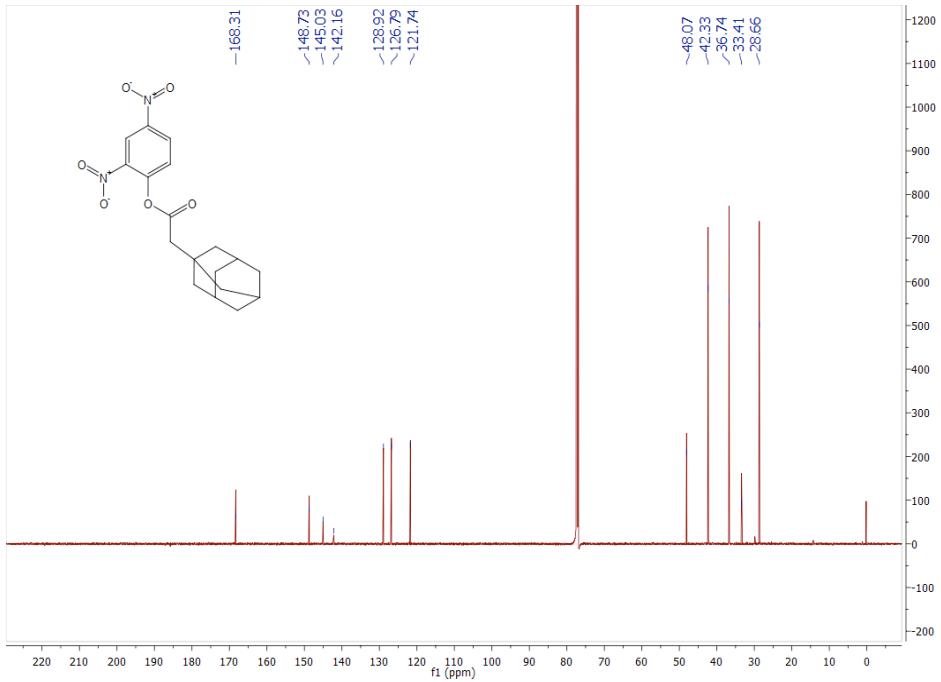
**<sup>13</sup>C-NMR for 37**



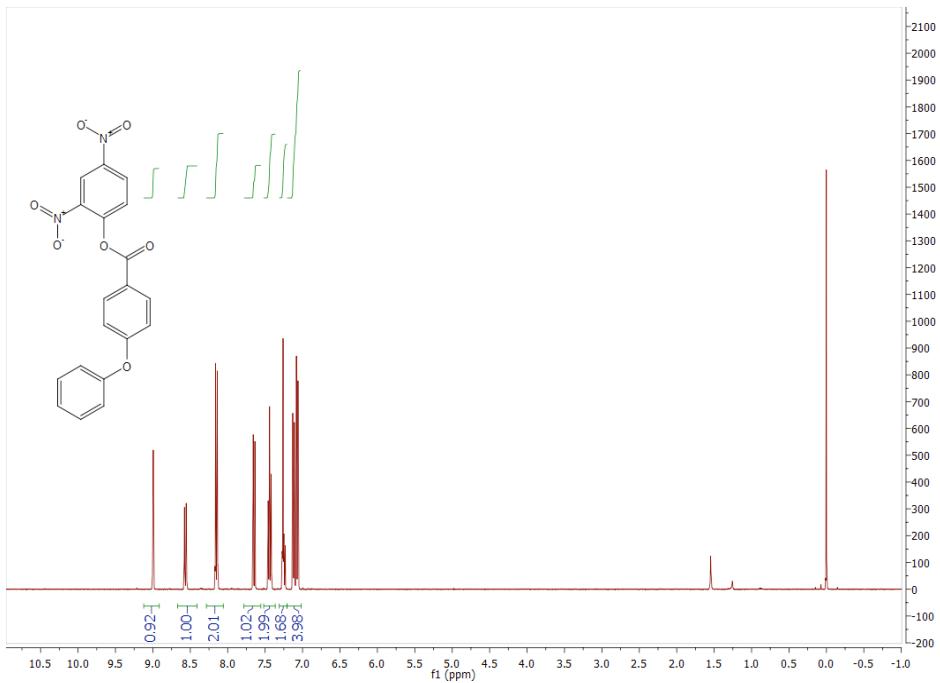
<sup>1</sup>H-NMR for 38



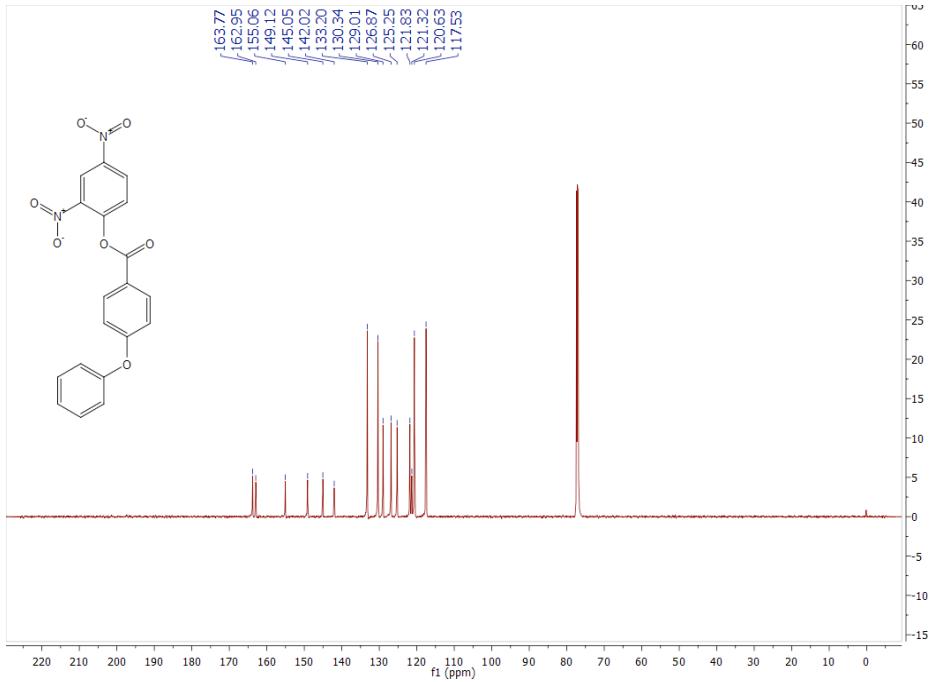
<sup>13</sup>C-NMR for 38



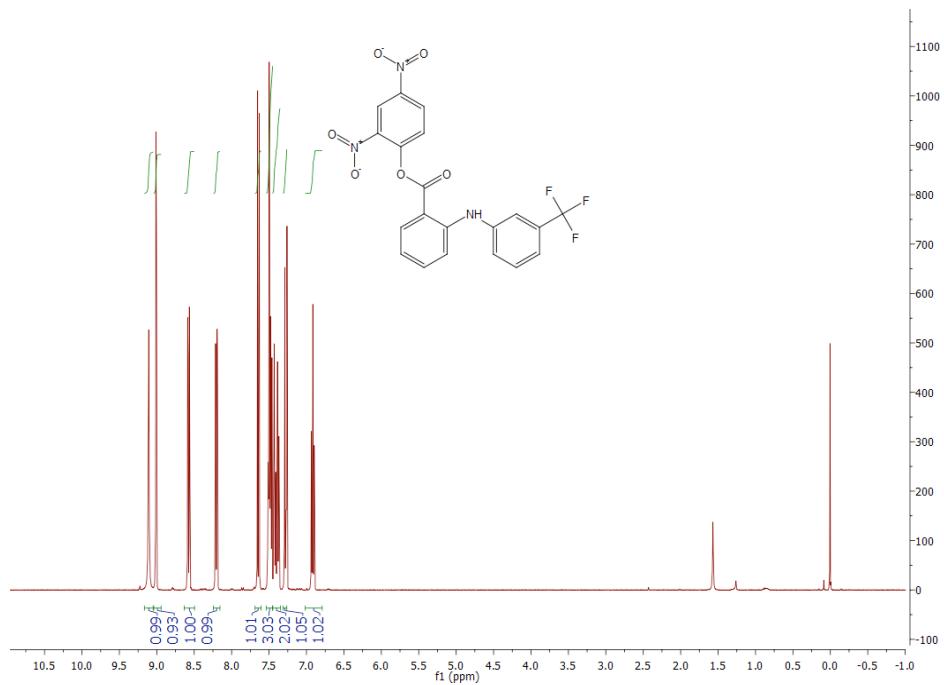
<sup>1</sup>H-NMR for 39



<sup>13</sup>C-NMR for 39



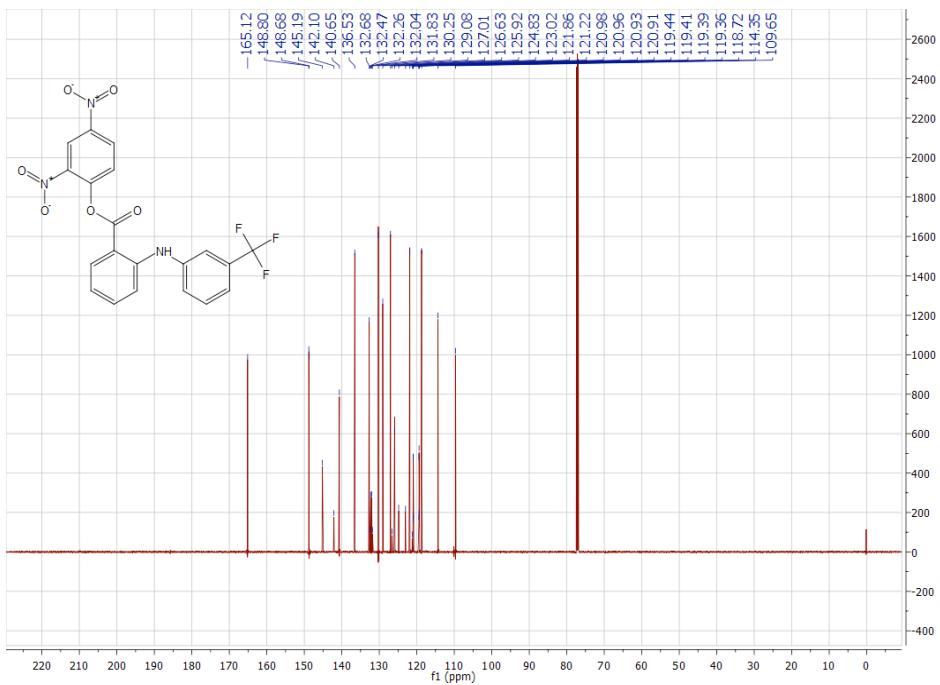
**<sup>1</sup>H-NMR for 40**



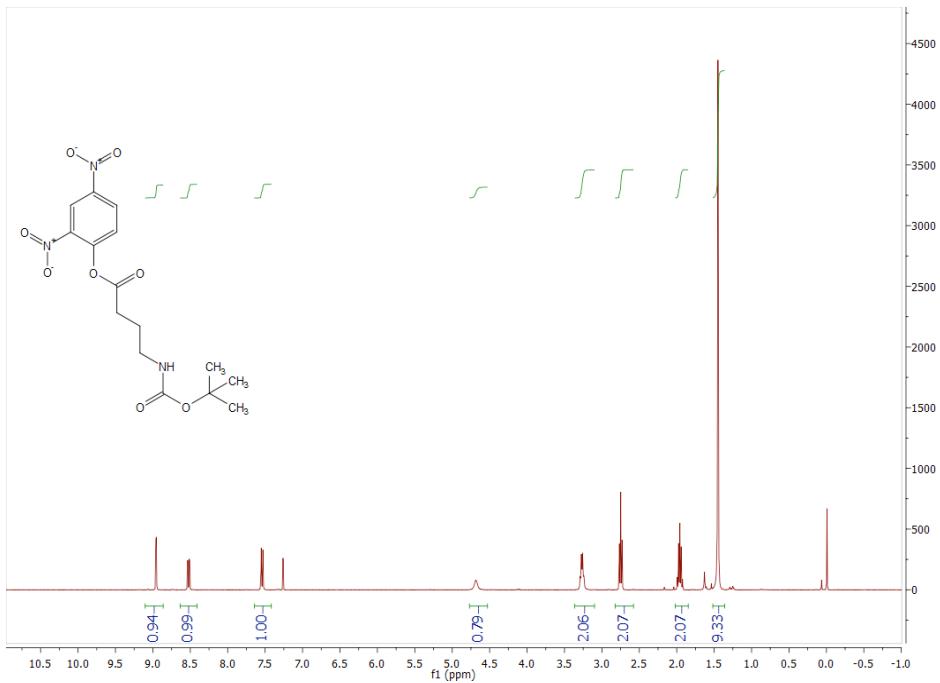
**<sup>19</sup>F-NMR for 40**



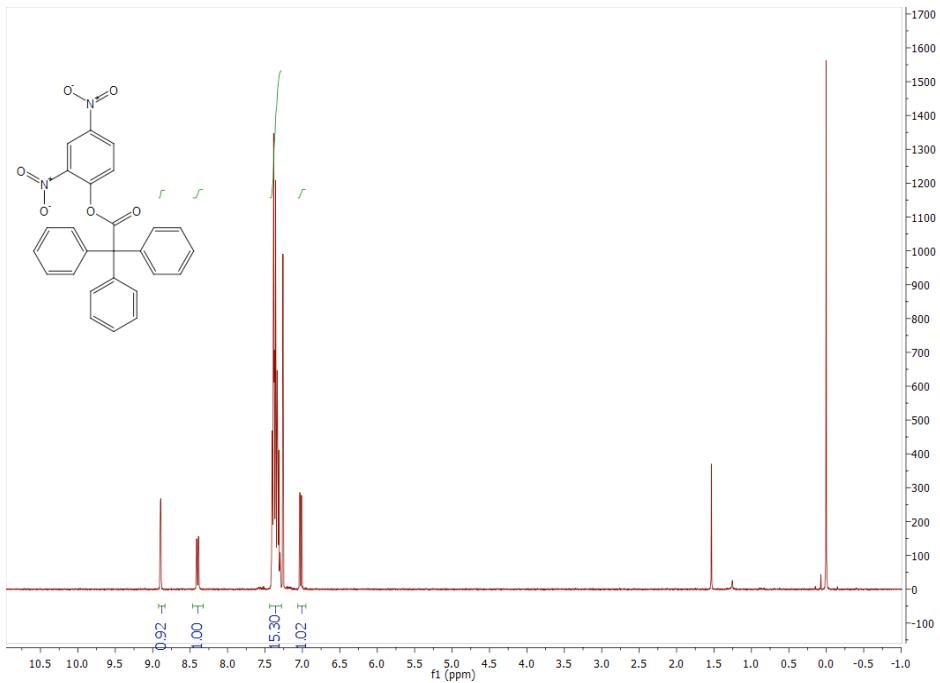
<sup>13</sup>C-NMR for 40



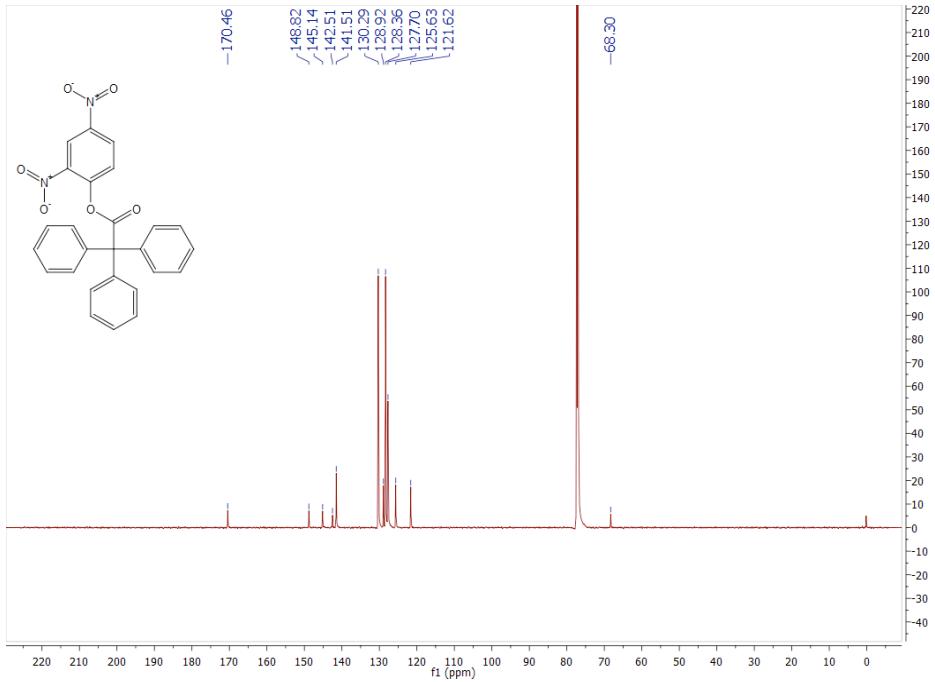
<sup>1</sup>H-NMR for 41



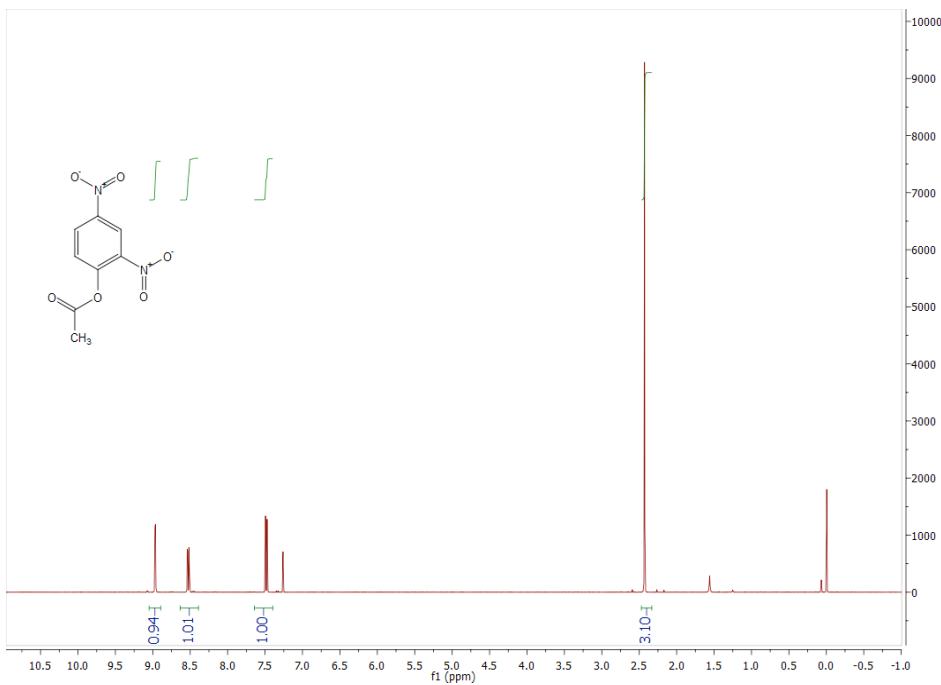
**<sup>1</sup>H-NMR for 42**



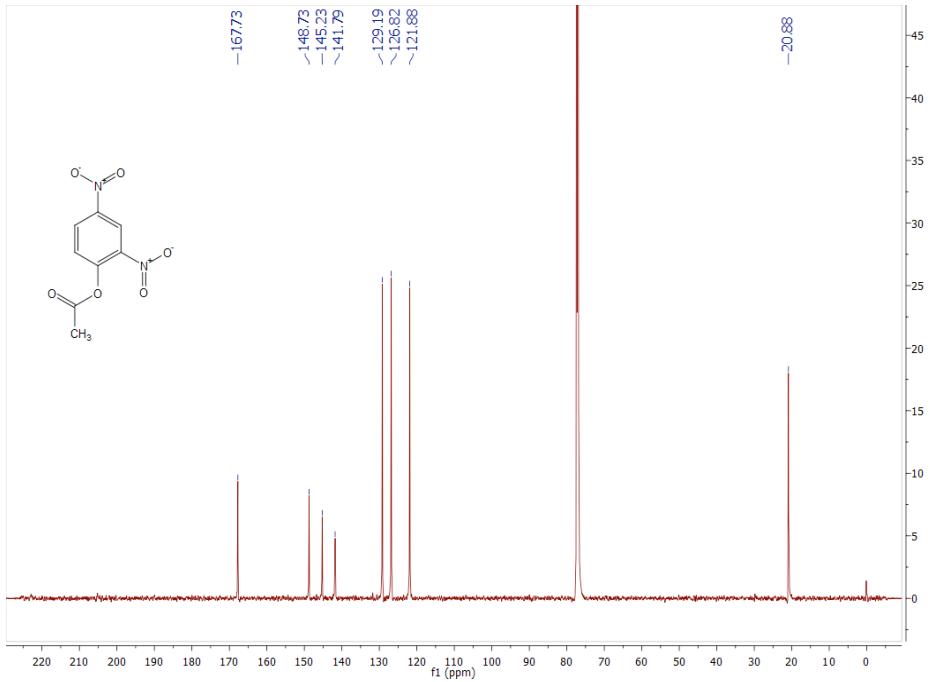
**<sup>13</sup>C-NMR for 42**



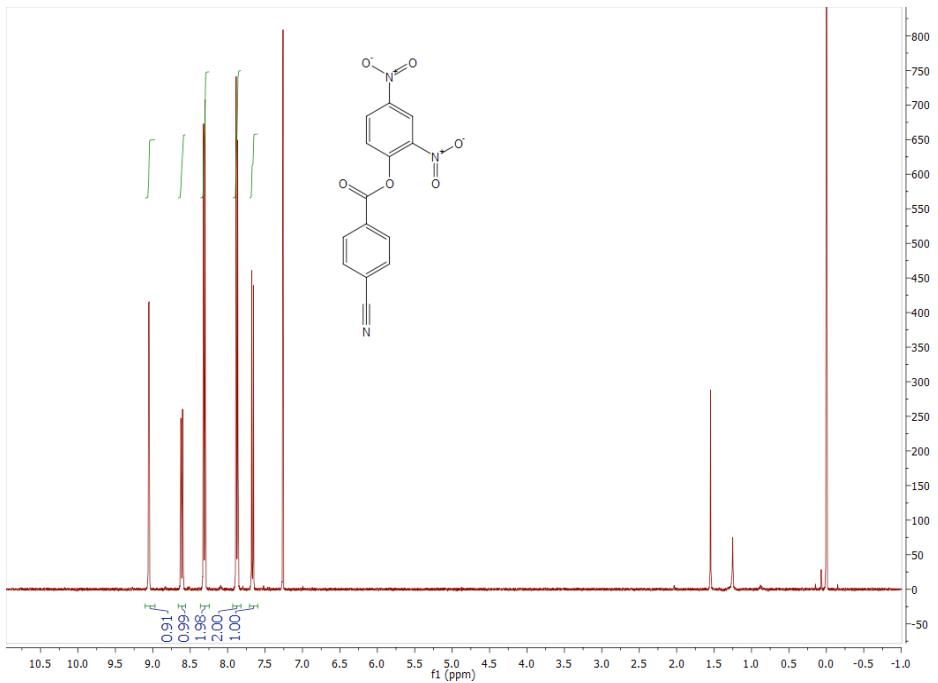
<sup>1</sup>H-NMR for 43



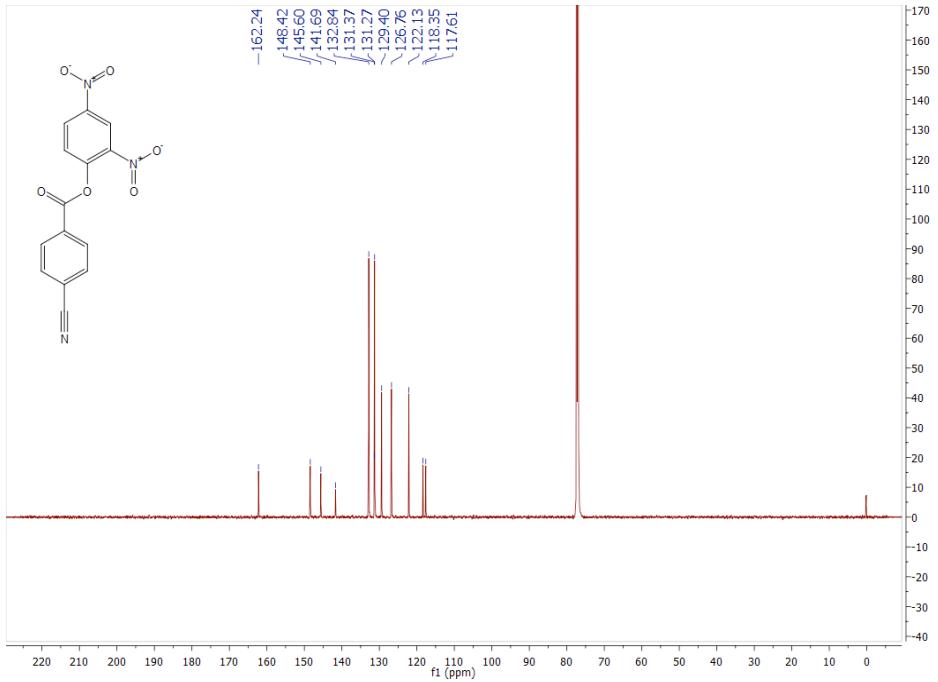
<sup>13</sup>C-NMR for 43



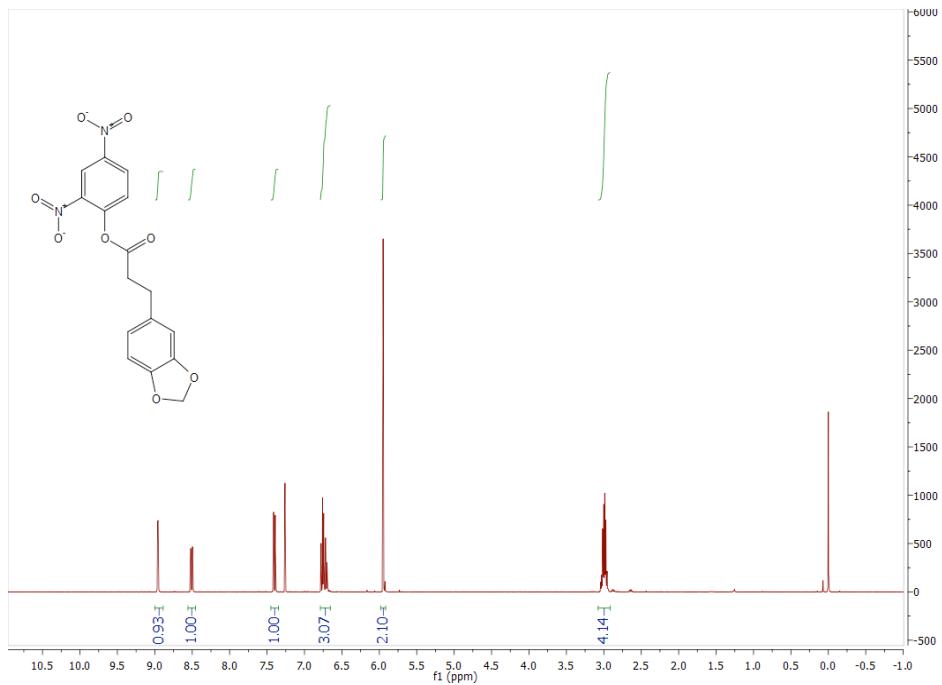
<sup>1</sup>H-NMR for 44



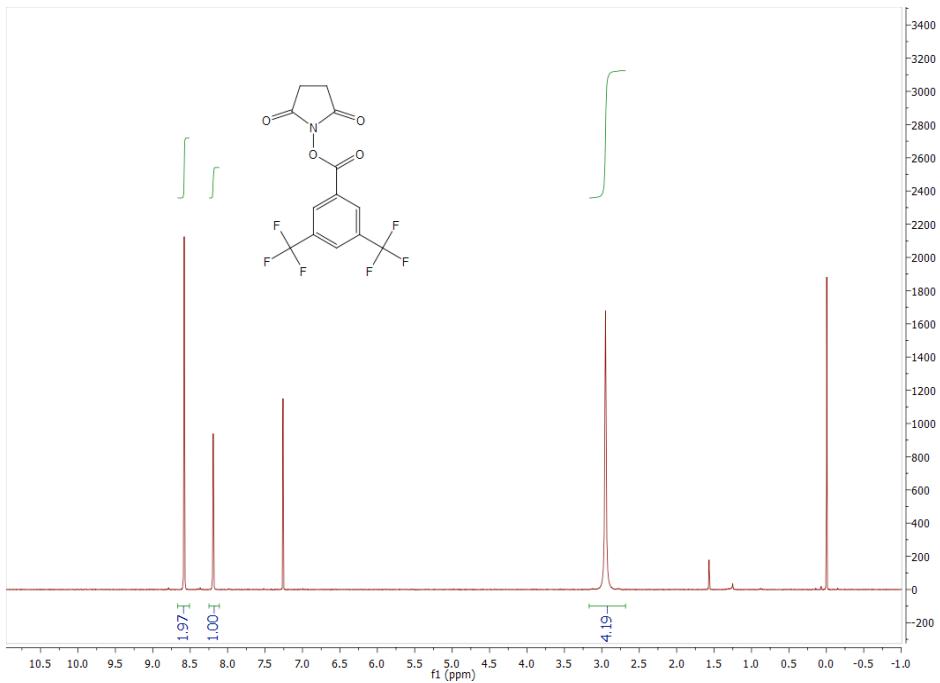
<sup>13</sup>C-NMR for 44



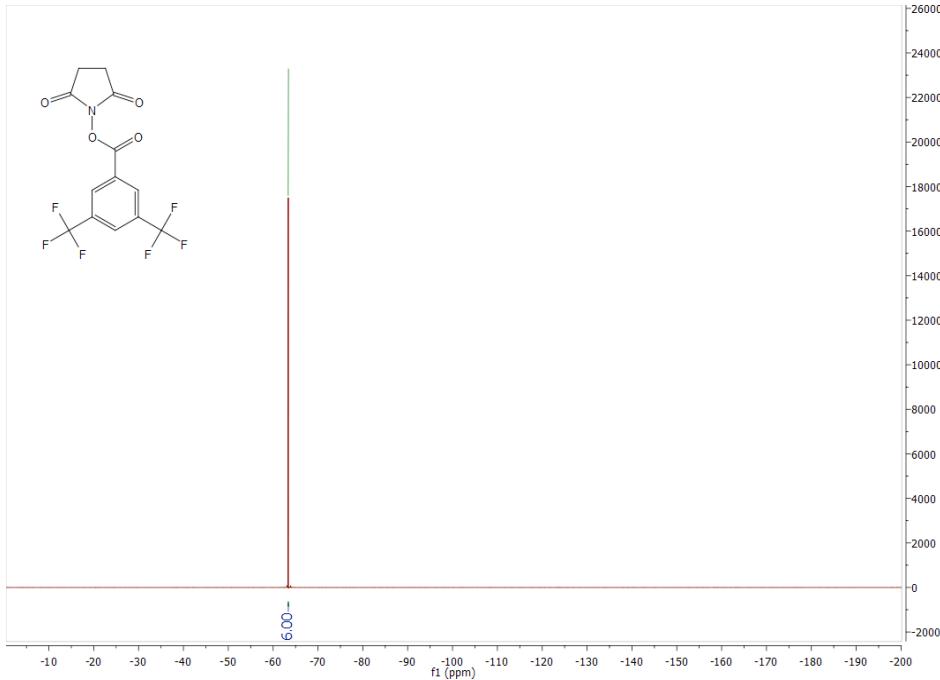
<sup>1</sup>H-NMR for 45



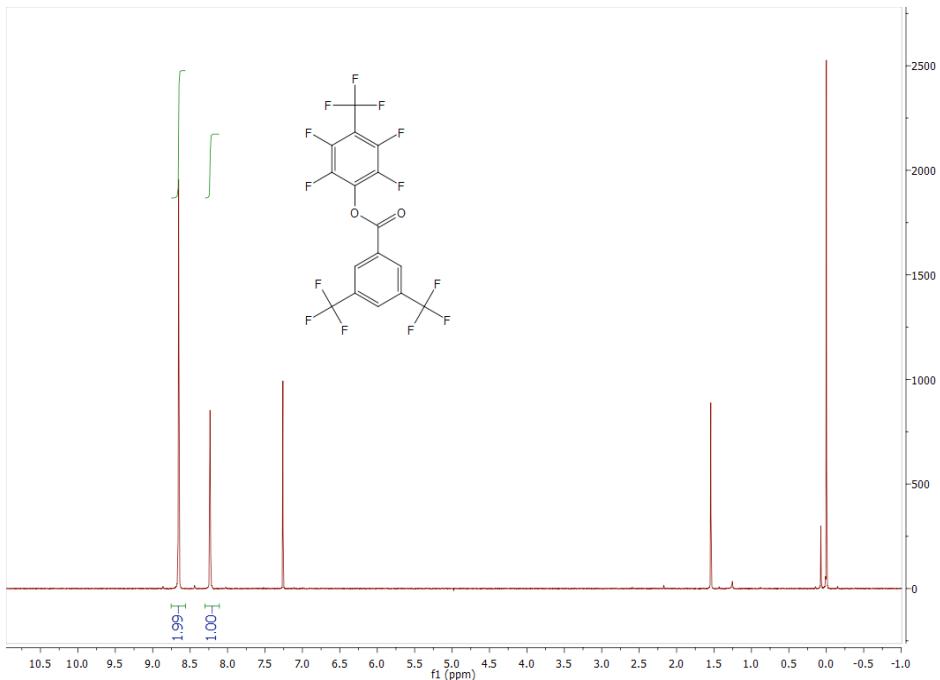
**<sup>1</sup>H-NMR for 46**



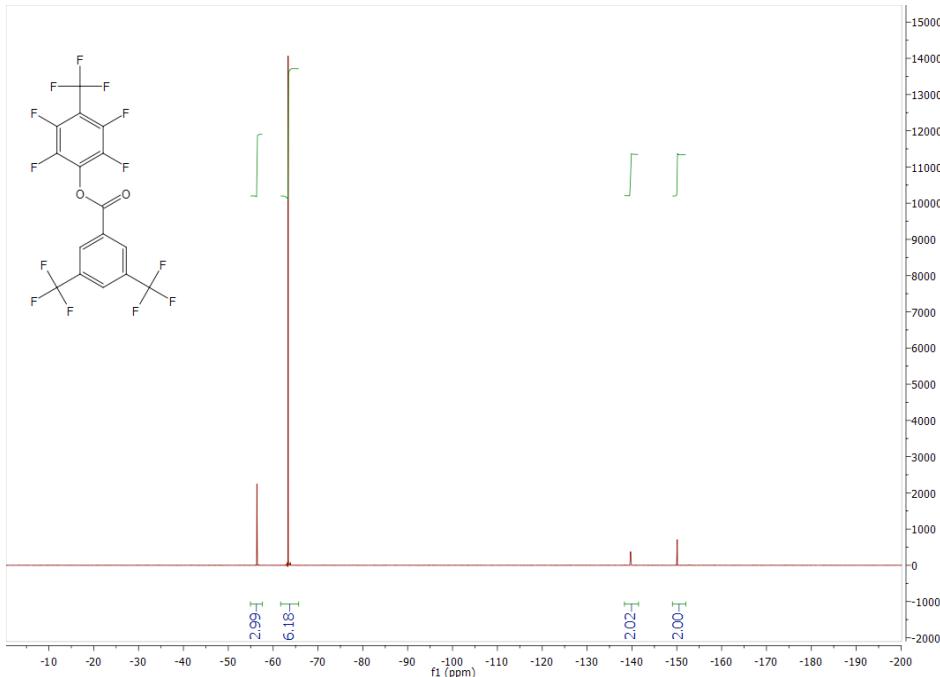
**<sup>19</sup>F-NMR for 46**



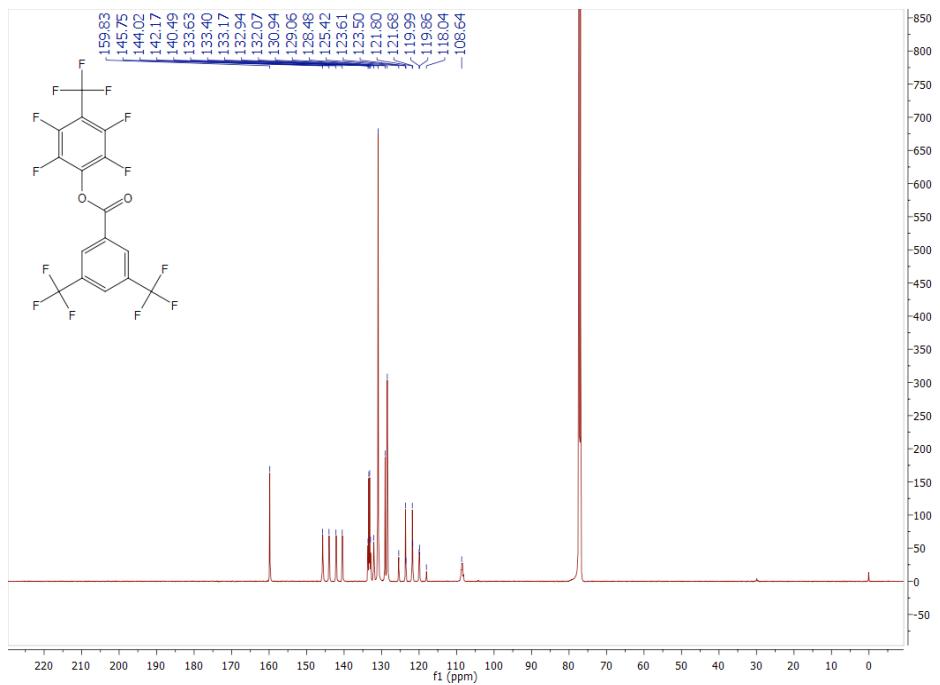
**$^1\text{H-NMR}$  for 47**



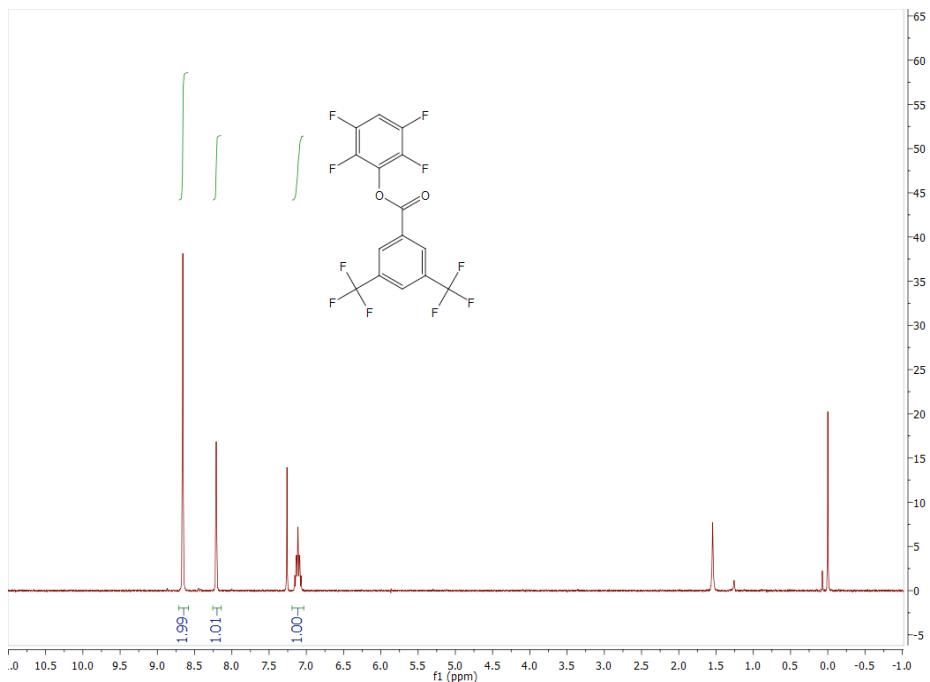
**$^{19}\text{F-NMR}$  for 47**



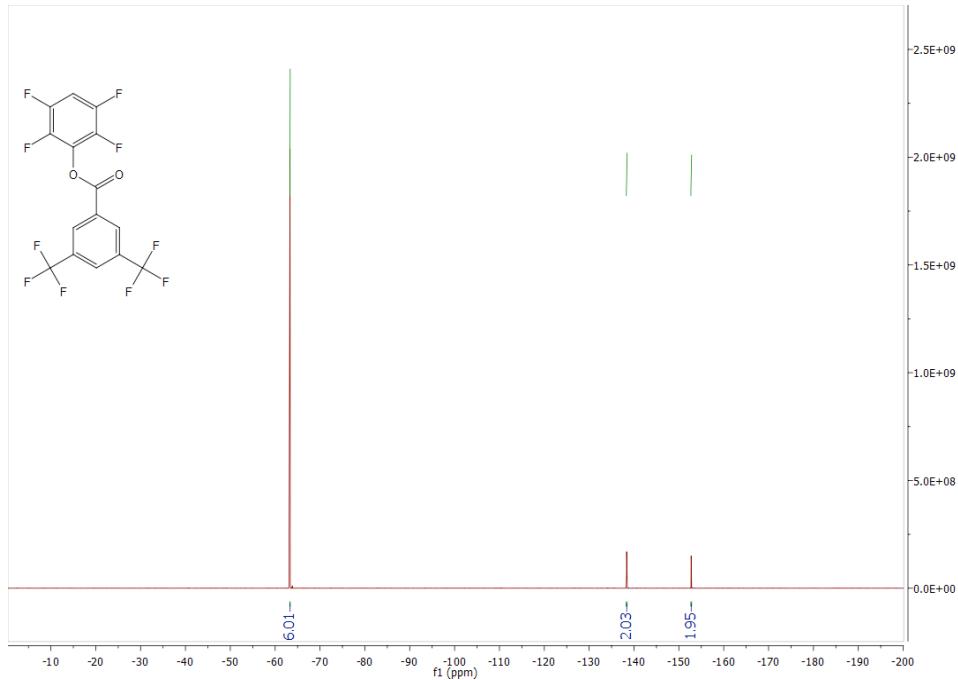
**<sup>13</sup>C-NMR for 47**



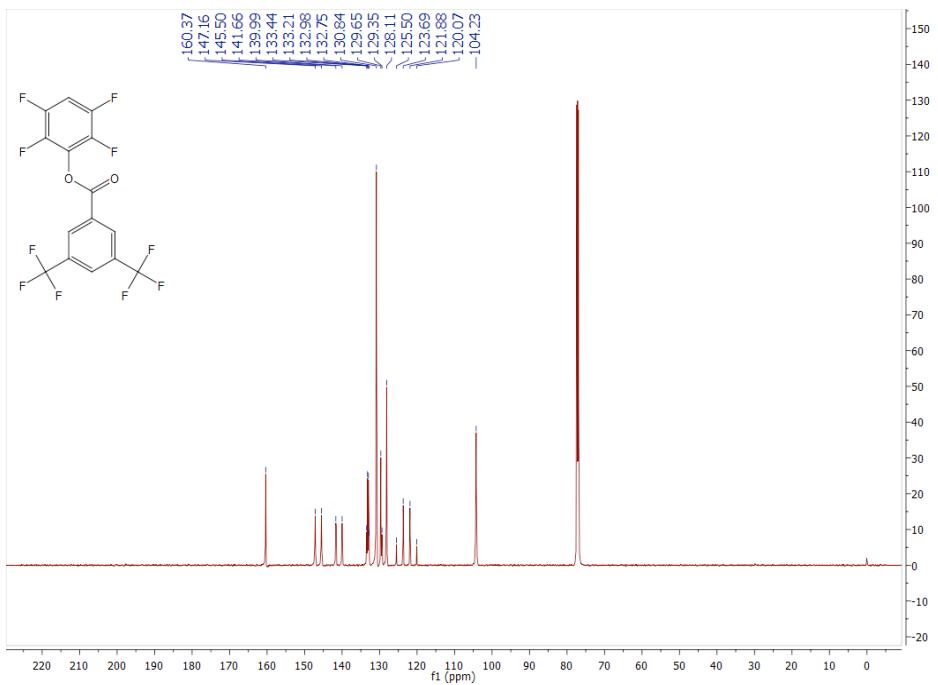
**<sup>1</sup>H-NMR for 48**



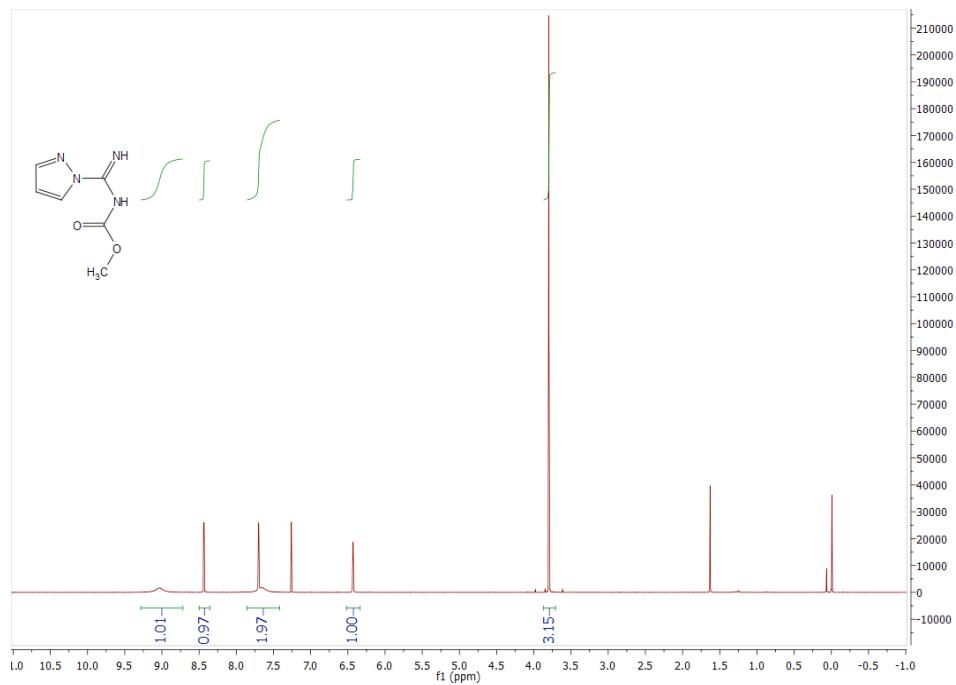
**<sup>19</sup>F-NMR for 48**



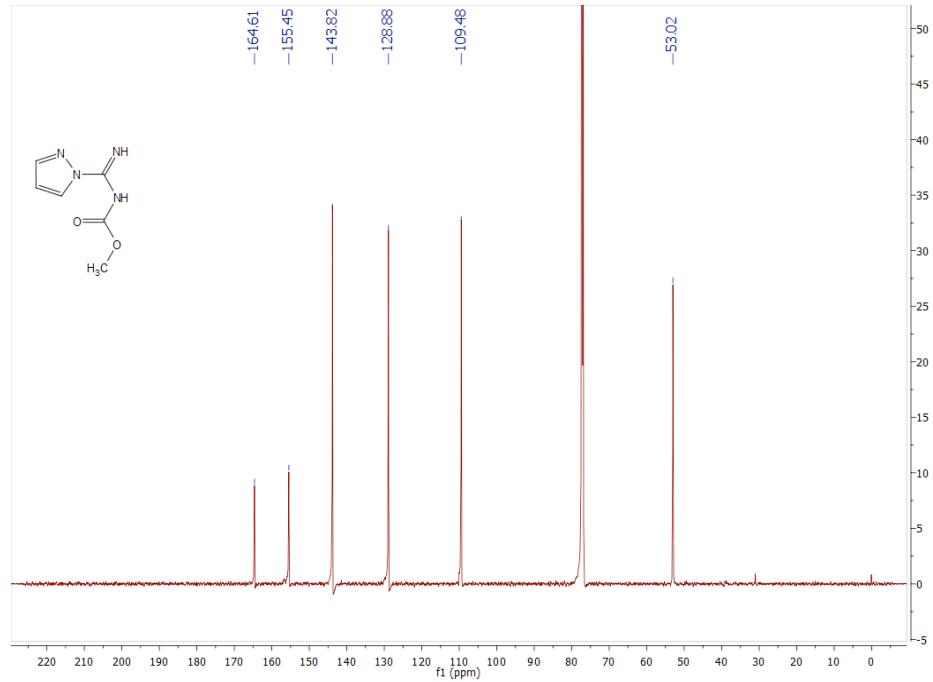
<sup>13</sup>C-NMR for 48



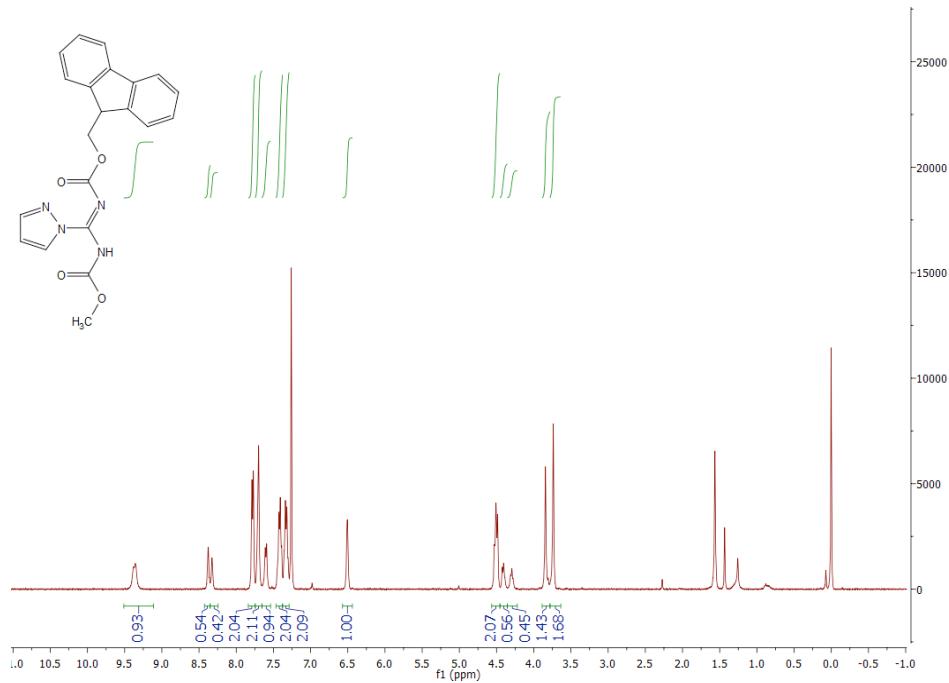
<sup>1</sup>H-NMR for 49a



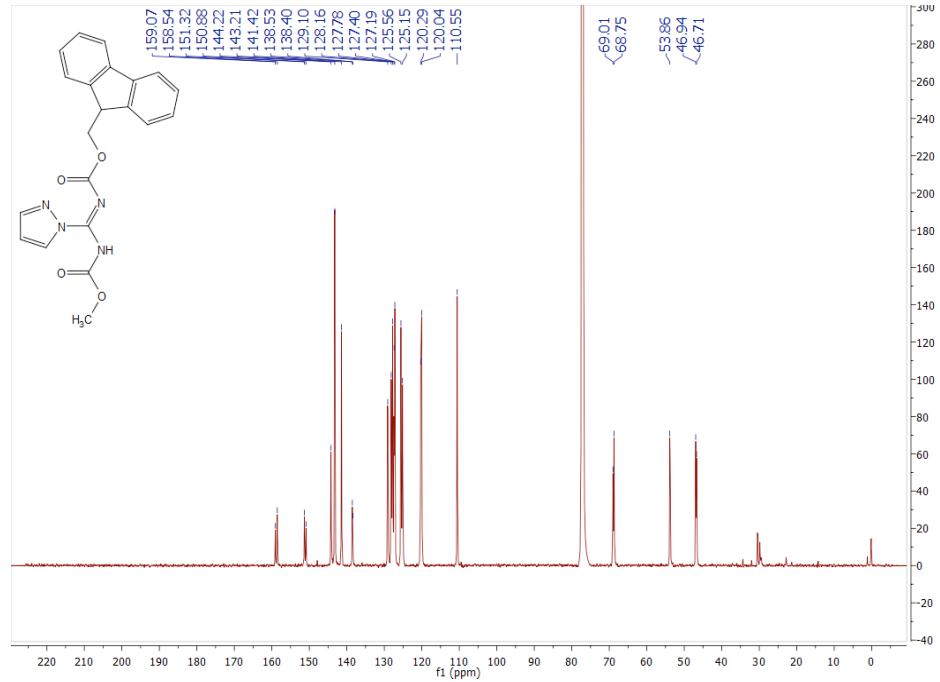
<sup>13</sup>C-NMR for 49a



**<sup>1</sup>H-NMR for 49**



**<sup>13</sup>C-NMR for 49**



## (F) References

- 1 Weerapana, E., Speers, A. E. & Cravatt, B. F. Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)--a general method for mapping sites of probe modification in proteomes. *Nat. Protoc.* **2**, 1414-1425 (2007).
- 2 Xu, T. *et al.* ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J. Proteomics* **129**, 16-24 (2015).
- 3 Weerapana, E. *et al.* Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790-795 (2010).
- 4 Wang, C., Weerapana, E., Blewett, M. M. & Cravatt, B. F. A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. *Nat. Meth.* **11**, 79-85 (2014).
- 5 Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-242 (2000).
- 6 Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512-520 (2015).
- 7 Bakan, A., Meireles, L. M. & Bahar, I. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575-1577 (2011).
- 8 Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **285**, 1735-1747 (1999).
- 9 Forli, S. *et al.* Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **11**, 905-919 (2016).
- 10 Ravindranath, P. A. & Sanner, M. F. AutoSite: an automated approach for pseudo-ligands prediction-from ligand-binding sites identification to predicting key ligand atoms. *Bioinformatics* **32**, 3142-3149 (2016).
- 11 Choi, S., Connelly, S., Reixach, N., Wilson, I. A. & Kelly, J. W. Chemoselective small molecules that covalently modify one lysine in a non-enzyme protein in plasma. *Nat. Chem. Biol.* **6**, 133-139 (2010).
- 12 Yi, W. *et al.* Phosphofructokinase 1 glycosylation regulates cell growth and metabolism. *Science* **337**, 975-980 (2012).
- 13 Argoudelis, C. J. Preparation of crystalline pyridoxine 5'-phosphate and some of its properties. *J. Agr. Food Chem.* **34**, 995-998 (1986).
- 14 Hacker, S. M., Buntz, A., Zumbusch, A. & Marx, A. Direct Monitoring of Nucleotide Turnover in Human Cell Extracts and Cells by Fluorogenic ATP Analogs. *ACS Chem. Biol.* **10**, 2544-2552 (2015).
- 15 Inloes, J. M. *et al.* The hereditary spastic paraparesia-related enzyme DDHD2 is a principal brain triglyceride lipase. *Proc. Natl. Acad. Sci. U S A* **111**, 14924-14929 (2014).