1,3-diketone analogs as selective lysyl hydroxylase 2 (LH2) antagonists

Juhoon Lee,[1] Hou-fu Guo,[*,2] Yazdan Maghsoud,[3] Erik Antonio Vázquez-Montelongo,[4] Zhifeng Jing,[5] Shike Wang,[6] Rachel M Sammons,[1] Eun Jeong Cho,[1] Pengyu Ren,[*,5] G. Andrés Cisneros[*,3,7] Jonathan M. Kurie,[*,6] and Kevin N. Dalby,[*,1]

[1]*Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy, University of Texas at Austin, Austin, Texas 78712, USA; Targeted Therapeutic Drug Discovery and Development Program, College of Pharmacy, University of Texas, Austin, TX 78712, USA*
*E-mail: dalby@austin.utexas.edu
[2]*Department of Molecular and Cellular Biochemistry, University of Kentucky College of Medicine, KY 40536, USA*
*E-mail: hguo3@g.uky.edu
[3]*Department of Chemistry and Biochemistry, The University of Texas at Dallas, Richardson, TX 75080, USA*
[4]*Department of Physical Medicine and Rehabilitation, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA*
[5]*Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX 78712, USA*
*E-mail: pren@utexas.edu
[6]*Departmnt of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA*
*E-mail: jkurie@mdanderson.org
[7]*Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA*
*E-mail: andres@utdallas.edu

Abstract

Lysyl hydroxylase 2 (LH2), an Fe(II) and α-ketoglutarate (αKG, also called 2-oxoglutarate, or 2OG)-dependent oxygenase, is an endoplasmic reticulum-resident enzyme that hydroxylates telopeptidyl lysine residues on fibrillar collagen molecules. It leads to the formation of hydroxylysine aldehyde-derived collagen cross-links (HLCCs), which are more stable than lysine aldehyde-derived collagen cross-links (LCCs) generated devoid of LH2. It has been reported that LH2 enhances lung cancer metastatic and invasive proclivity and modulates the types of collagen cross-links (HLCC-to-LCC) in the tumor stroma. Herein, we prepared a series of 1,3-diketone analogs **1**–**18** and identified **12** and **13** that inhibit the LH2-driven hydroxylation of a collagen peptide substrate with IC_{50} approximately 300 nM and 500 nM, respectively. **12** and **13** demonstrate a 9-fold selectivity for LH2 over LH1 and LH3. Quantum Mechanics/Molecular Mechanics (QM/MM) modeling indicates that in addition to the relatively stronger interactions between compounds **12** and **13** with the active site, the selectivity stems from non-covalent interactions like hydrogen bonding between the morpholine/piperazine rings with LH2-specific Arg661, where the corresponding residue in LH1 and LH3 is Pro. Migration assays in the 344SQ lung adenocarcinoma cell line reveal that **13** shows anti-migration activity.

Keywords: 1,3-diketones, LH2 inhibitors, Luciferase-based enzymatic activity assay, Anti-migration

1. Introduction

Fibrillar collagens are the most abundant extracellular matrix (ECM) proteins conferring mechanical strength to tissues and organs. Collagen biosynthesis involves several post-translational modifications. In particular, the modifications of lysine residues are partly accountable for regulating the fibrillar collagens' structural and mechanical properties. Lysines on helical and telopeptide domains of fibrillar collagens can be hydroxylated to form hydroxylysines which are further converted to the corresponding aldehydes by lysyl oxidase (LOXs)-mediated oxidative deamination.[1] These reactive aldehydes undergo multiple condensation reactions to form collagen crosslinks. Telopeptidyl lysines form lysine aldehyde-derived collagen crosslinks (LCCs).

In contrast, lysyl hydroxylase 2 (LH2)-mediated telopeptidyl hydroxylysines form structurally stubborn hydroxylysine aldehyde-derived collagen crosslinks (HLCCs) resistant to cleavage by collagenases.[2] It is well-known that reduced lysine hydroxylation is a cause of genetic connective tissue disorders such as Bruck Syndrome,[3-7] while abnormally elevated HLCC formation are associated with fibrosis and cancer metastasis.[8-12]

Three lysyl hydroxylase isoforms (LH1-3) play essential roles in collagen modifications in the endoplasmic reticulum (ER). Each isoform possesses two functionally distinct enzymatically active domains.[13] LH2 can hydroxylate Lys residues in the telopeptides, while all three LH isoforms hydroxylate lysine residues in the helical domain.[2] In fibrotic diseases of the lung[14-15] and liver,[16-18] HLCCs are the predominant types of cross-links attributable to undue hydroxylation of the telopeptidyl Lys, owing to the overexpression of LH2 in fibroblast.[9, 19] A similar switch is associated with tumor stroma where an increased level of LH2 leads to metastasis and is correlated with lower survival.[8, 10, 20-23] Heretofore, LH2 was thought to be localized in ER, where it hydroxylates procollagen Lys residues before the triple helix formation.[24] However, it was recently found to modify collagen in the extracellular space as it is secreted by carcinoma cells[25] and shown to be expressed in cancer-associated fibroblasts and contribute to a switch toward a high-HLCC, low-LCC state in the tumor stroma.[26]

LH3 was thought to be unique in its ability to modify helical hydroxylysine into 1,2-glucosylgalactosyl-5-hydroxylysine cooperatively with procollagen galactosyltransferase 1 and 2 (GLT25D1 and GLT25D2).[27-30] However, the longer LH2 isoform, LH2b, a driving factor in cancer, is overexpressed in lung adenocarcinomas and possesses robust collagen galactosyl hydroxy lysyl glucosyltransferase (GGT) activity. Notably, CRISPR/Cas-9-mediated inhibition of the GGT domain decreases the growth and metastasis of LH2b-expressing lung adenocarcinomas in a syngeneic immunocompetent orthotopic mouse model of lung cancer,[13] suggesting a role for the GGT activity in lung adenocarcinomas progression.[13]

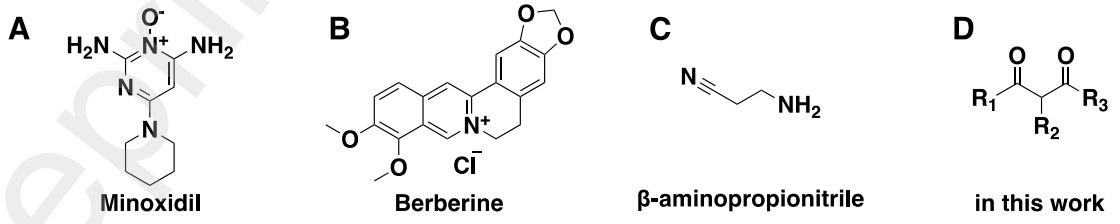


Figure 1. Compounds reported as LH2

Minoxidil (Figure 1A) has been considered an anti-fibrotic agent that could attenuate the expression of LHs, resulting in reduced formation of hydroxylysylpyridinoline cross-linking.[31-32] However, minoxidil suppresses the expression of proteins beyond the LH family. Moreover, Pfeffer *et al.* reported that minoxidil did not impact LH activity during postnatal mouse lung development.[33] A recent report demonstrated that berberine (Figure 1B) inhibits LH2 expression at both mRNA and protein levels, leading to decreased triple-negative breast cancer (TNBC) cell proliferation, motility, and glycolysis process. The regulation of LH2 by berberine results from the inhibition of the secretion of IL-6, thereby suppressing LH2 indirectly.[34] β-aminopropionitrile (Figure 1C) was reported to downregulate the expression of LH2 and inhibit the formation of a stable matrix, presumably via directly inhibiting LOX family members.[35]

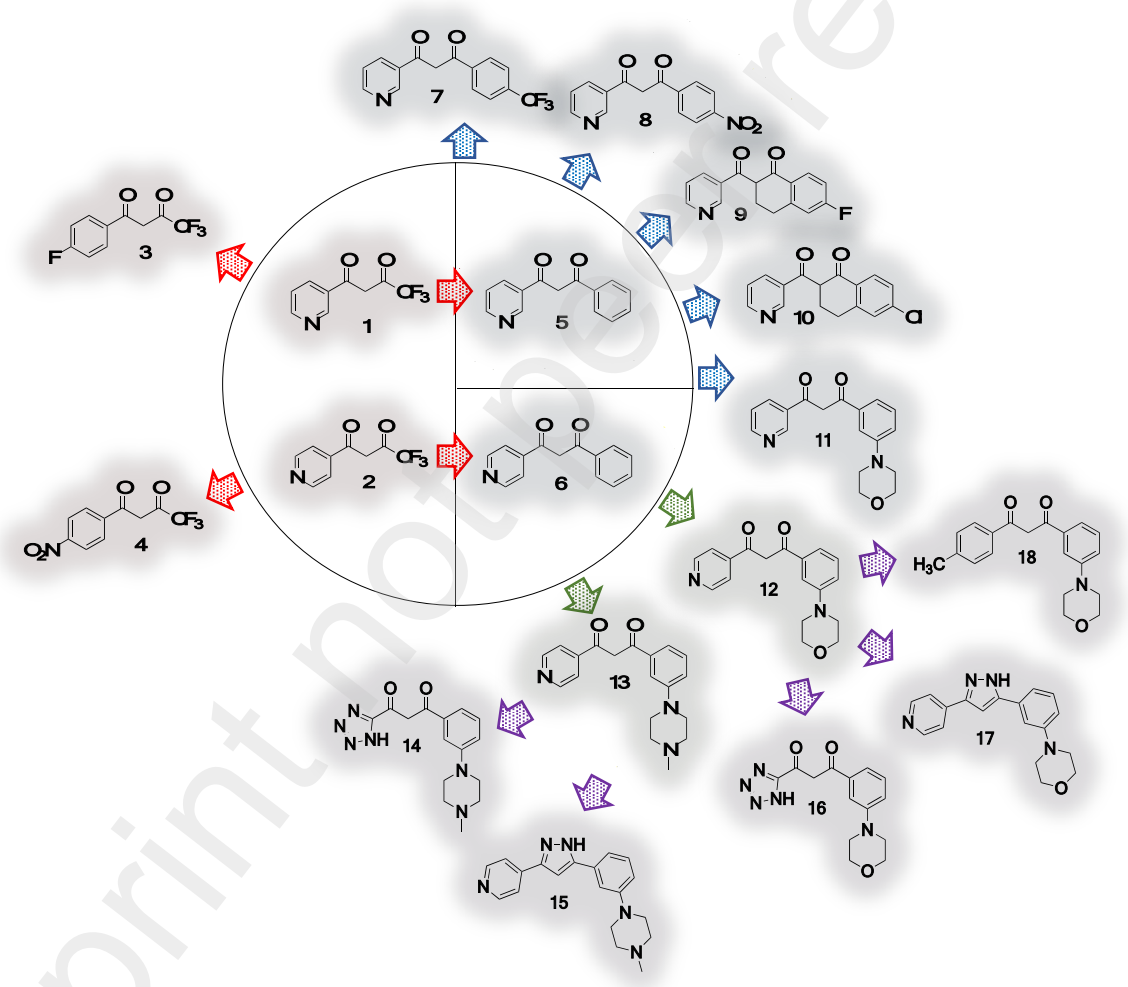


Figure 2. LH2 inhibitors design flow.

1,3-diketones (Figure 1D) can be found in numerous natural products, exhibiting many pharmacological activities, including antibacterial, antioxidant, antiviral, insecticidal, and antifungal activity.[36-41] The 1,3-diketone has been widely studied in various medicinal and coordination chemistry fields. Its intrinsic ability to participate in metal chelation and hydrogen bonding owing to tautomerism makes the 1,3-diketone a promising scaffold for intermediates in various organic reactions, metal-organic hybrid materials, and novel drug development.[42-47]

Here we report the development of a series of antagonists (compounds **1**–**18** in Figure 2) based on 1,3-diketone analogs against LH2. Compound **13** was considered in detail and was shown to inhibit the enzymatic activity of LH2 *in vitro* with some selectivity over other isoforms (LH1 and LH3) and hinder the migration of 344SQ lung adenocarcinoma cells (LH2b highly expressed) in a dose-dependent manner. Quantum Mechanics/Molecular Mechanics (QM/MM) calculation studies revealed that the calculated interaction energy ($IE_{QM/MM}$) results between **12**, **13,** and the LH2 is –259.9 and –216.3 kcal mol^{-1}, respectively, suggesting the presence of strong interactions between the ligands and the enzyme.

2. Results and discussions

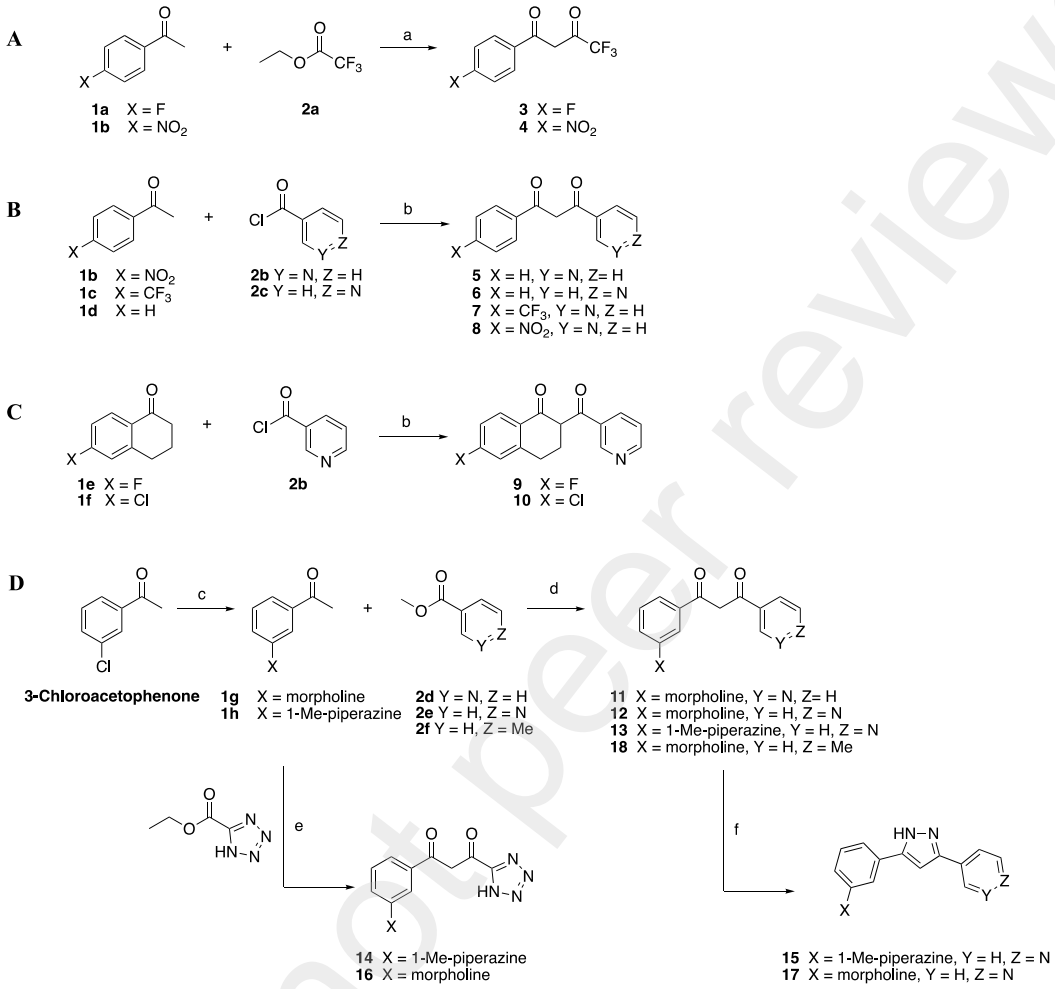
2.1. Synthesis

The 1,3-diketone analogs were synthesized from the acylation of enolizable carbonyl compounds following literature procedures.[47] **3** and **4** were prepared via modified Claisen intermolecular condensation. The para-substituted acetophenones were enolized by NaH and then acylated with ethyl trifluoroacetate to obtain the corresponding compounds (Scheme 1A). Compounds **5–10** were synthesized from acyl chlorides with enolated ketones by soft enolization[48] in the presence of magnesium bromide diethyletherate and Hünig's base, as shown in Schemes 1B and C. The reactions were performed at room temperature. pH adjustment (pH 4–5) using 1N HCl afforded the compounds. **11–13**, and **18** were synthesized as shown in Scheme 1D. Pd-catalyzed C–N bond formation reactions of 3-chloroacetopheone with morpholine or 1-methyl-piperazine in the presence of DavePhos (2-dicyclohexylphosphino-2'-(N,N-dimethylamino)biphenyl) and tripotassium phosphate gave the corresponding meta-substituted acetophenones **1g** and **1h**, respectively. Acylation with esters **2d–2f** in the presence of lithium bis(trimethylsilyl)amide afforded the desired compounds. The meta-substituted acetophenones also underwent an acylation reaction with 2-methoxypropyl-protected ethyl tetrazole-5-carboxylate to afford tetrazole-substituted variants **14** and **16**. Pyrazole-substituted derivatives (**15** and **17**) were prepared by a reaction of **12** or **13** with hydrazine monohydrate in ethanol.

2.2. Lead optimization based on enzymatic activity in vitro assay and QM/MM docking simulations

Previously, we reported the development of a luminescence-based *in vitro* assay strategy to monitor LH activity suitable for high-throughput screening (HTS).[49] Our efforts discovered **1** and **2** (IC_{50}s of 3.4 and 1.9 μM, respectively) as HTS hits that formed the basis for further developing LH2 antagonists with better potency and selectivity. Molecular docking studies of the HTS hit compounds to an LH2 homology model supported the notion that the dicarbonyl moiety is responsible for the chelation of the catalytic Fe(II) atom, with the pyridine nitrogen participating

in H-bonding interactions with LH2 active-site residues Y655 and R728 (human PLOD2 full length) (Figure 3A and B).



Scheme 1. Synthesis of inhibitors **3-18**. Reagents and conditions: (a) NaH, THF, 0 °C – RT,12 h; (b) $MgBr_2$•OEt, DIPEA, DCM, RT,12 h; (c) morphoine or 1-Me-piperazine, $Pd_2(dba)_3$, DavePhos, K_3PO_4, DME, Ar, 100 °C, 12h; (d) $LiN(SiMe_3)_2$, Ar, THF, –40 °C, 6h; (e) i) 2-methoxypropene, p-TSA, THF, RT, 1h. ii) $LiN(SiMe_3)_2$, Ar, THF, –40 °C, 6h, iii) HCl (f) NH_2NH_2•H_2O, ethanol, RT, 12h.

Based on the initial HTS and docking results, a computational study was performed in parallel on forty-four molecules, including most of the compounds of this paper.[59] Based on this previous study, we found that the enolate form of R_2 with specific substituents for R_1 and R_3 showed preferential interactions with LH2. Additionally, tunneling analysis to determine the existence of O_2 tunnels to transport the co-substrate to the active site revealed the possible existence of more than one tunnel. The more promising candidates from our previous study were observed to partially block these putative tunnels and have negatively charged enolate forms (IDs: 2- cpd1p and 4- cpd2p in Ref. 59), which have considerably strong interactions with the active site (–169.4 and –139.7 kcal mol^{-1}, respectively) as well as several NCIs (non-covalent interactions) with the residues of the active site including Y655 and R728.

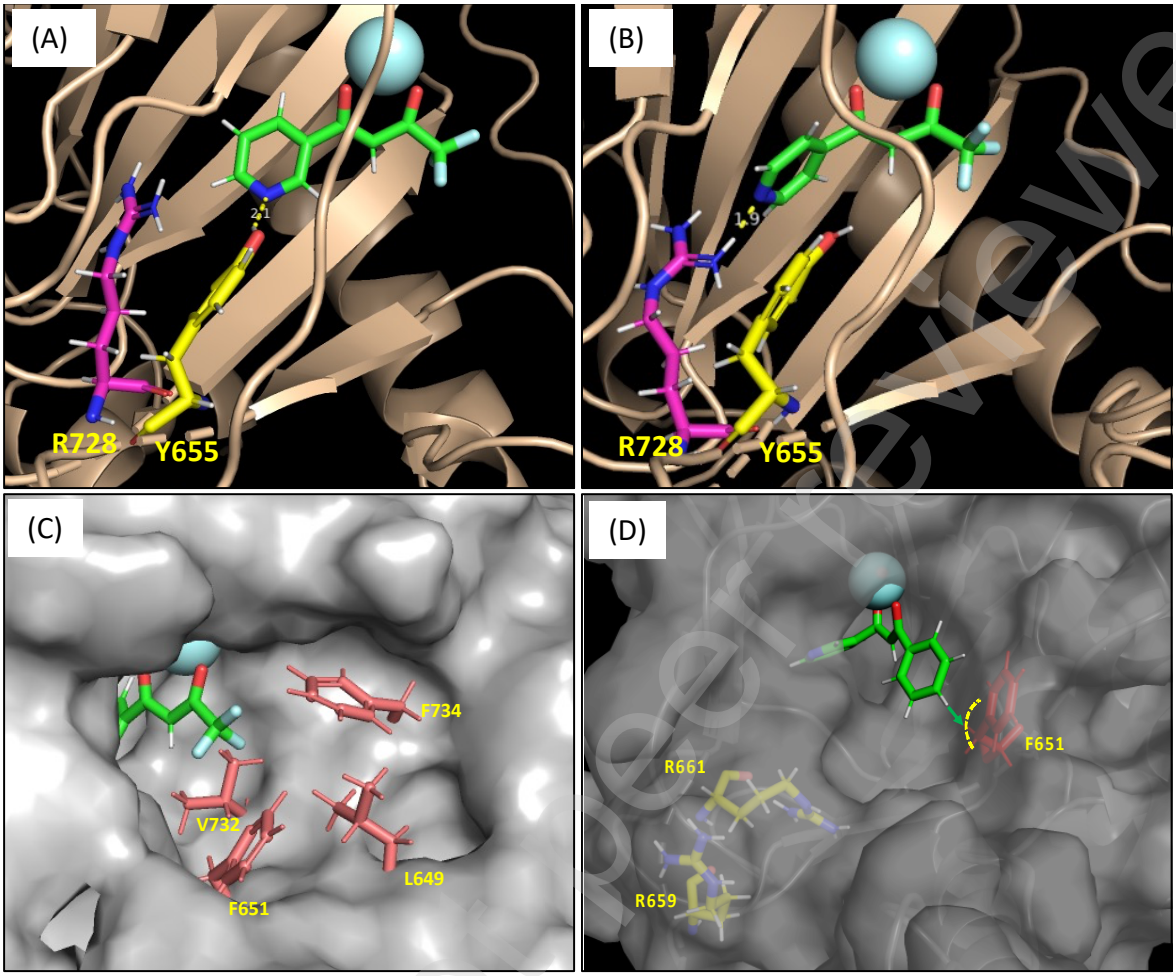


Figure 3. Binding interactions between HTS hit compounds **1** (**A** and **C**), **2** (**B**), and **5** (**C**) and LH2 active site predicted by molecular docking studies. **(A)** Calculated distance between the phenol hydrogen (Y655) and the pyridine nitrogen of **1** is 2.1 Å. **(B)** Calculated distance between the guanidine hydrogen (R728) and the pyridine nitrogen of **2** is 1.9 Å. **(C)** The entry of the active site of LH2 associating with **1** shows the hydrophobic environment in proximity to trifluoromethyl ($–CF_3$) group of **1**. **(D)** Docked model of **5** presenting the surface environment of benzene ring of the compound.

To test the role of pyridine in the interaction with the inner residues, compounds **3** and **4** were synthesized and assessed for inhibitory activities. Both compounds with para-substituted benzene rings on the R_1 position exhibited less potency than the original HTS compounds **1** and **2**, indicating that the pyridine has a vital role in associating with the LH2 active site. Calculated interactions for these two compounds with LH2 (IDs: 8- cpd6 and 9- cpd7 in Ref. 59) were also weak (–50.4 and –38.7 kcal mol^{-1}, respectively).

We then considered a modification of the R_3 position. We hypothesized that the trifluoromethyl ($–CF_3$) group plays a role not only in withdrawing electron density from the adjacent carbonyl group, which stabilizes the formation of an enolate that is considered to enhance the affinity towards Fe(II) metal, but also participates in hydrophobic interactions with nonpolar residues (L649, F651, V732, and F734) at the entry of the LH2 active site (Figure 3C). Accordingly,

we incorporated a benzene ring at the R_3 position in place of the CF_3, which is considered to possess similar characteristics as well as better accessibility for functionalization. Notably, the benzene-functionalized compounds **5** (IC_{50} = 5.0 μM) and **6** (IC_{50} = 1.0 μM) retained inhibitory activity compared to the HTS hits. In comparison with compounds **1** and **2**, compounds **5** (ID: 18-cpd1php in Ref. 59) and **6** (Table S2) show considerably stronger interactions with the enzyme (–185.2 and –200.1 kcal mol^{-1}, respectively).

Compounds **5** were further functionalized at the para position of the benzene ring (**7** and **8**) and the 6-position of the tetralin group that is tethered at the R_2 position (**9** and **10**). The presence of functional groups on the para position of benzene reduced the potency of the compounds (**7**; IC_{50} = N/A, **8**; IC_{50} = >10 μM, **9**; IC_{50} = >10 μM, **10**; IC_{50} = N/A), which suggests the environment of the para position of the benzene ring may have limited space for the functional groups to reside, as shown by the molecular docking of **5** to the LH2 homology model (Figure 3D). The calculated interaction energies for compounds **7**–**10** (IDs: 43- c24, 44- c25, 34- cpd26p, and 36- cpd27p, in Ref. 59) were –114.2, –190.5, –212.6, and –212.9 kcal mol^{-1}, respectively. This suggests that the presence of $–CF_3$ in the para position of the benzene ring in compound **7** (43- c24 in Ref. 59) led to a significant decrease in the interaction energies. At the same time, it was slightly increased for the other three compared with compound **5**. Interestingly, the number of residues having non-covalent interactions with these four compounds showed a slight decrease compared to compound **5**, especially in the case of compounds **7** and **9** (Tables S2 and S10 in Ref. 59). This suggested that our modifications at the para position of benzene ring do not produce favorable interaction effects between the inhibitor and the binding pocket's residues. Given the LH2–inhibitor interaction energies, non-covalent interactions, and distinct hindrance by F651 on the para-modification, we investigated whether a meta-substitution could enhance binding to LH2.

Besides the Fe(II) chelation and H-bonding interaction with the active site residues, the docked model of **5** revealed possible additional recognition sites, R661 and R659, proximal to the meta position of the benzene of **5** (Figure 3D). These two basic amino acid residues adjacent to the LH2 active site are specific to this member of the LH family (the corresponding residues in LH1 and LH3 are Glu and Pro, respectively). These residues are thought to play a role in the electrostatic interactions with acidic aspartate, and glutamate residues positioned adjacent to the Lys residues on fibrillar collagens.[13] We postulated that LH2's unique telopeptidyl lysyl hydroxylase activity and selectivity are determined partly by electrostatic interactions between the basic domain of LH2 and the acidic domain of collagen telopeptides. To achieve favorable interactions with the basic residues, morpholine (**11** and **12**) and 1-methylpiperazine (**13**) rings, which possess the intrinsic capability to participate in hydrogen bonding[50-51] were incorporated at the meta-position of the benzene of compounds **5** and **6**. Introducing the meta-substituted morpholine and methylpiperazine rings boosts the potencies (**11**; IC_{50} = 1.5 μM, **12**; IC_{50} = 0.3 μM, and **13**; IC_{50} = 0.5 μM) as compared with the parent inhibitors **5** and **6**. Notably, the weaker potency of **11** compared to **12** indicates that the pyridine nitrogen atom in 4 positions is optimal for interacting with inner residues in the active site.

The results of the interaction energies between compounds **12** and **13** with LH2 are –259.9 and –216.3 kcal mol^{-1}, respectively (see Table S2). The low values of $IE_{QM/MM}$ suggest a strong interaction between these two analogs and the enzyme. Table S2 also shows that the calculated interactions due to the QM region in compounds **12** and **13** are similar, –197.6 and –197.4 kcal mol^{-1}, respectively, but the interactions due to the MM region in compound **12** are about three times greater than compound **13** (–62.3 and –18.9 kcal mol^{-1}, respectively). In the case of **6**, the

interaction between the ligand and the active site is weaker than the other two, but the MM region stability is enhanced compared with the other two ligands.

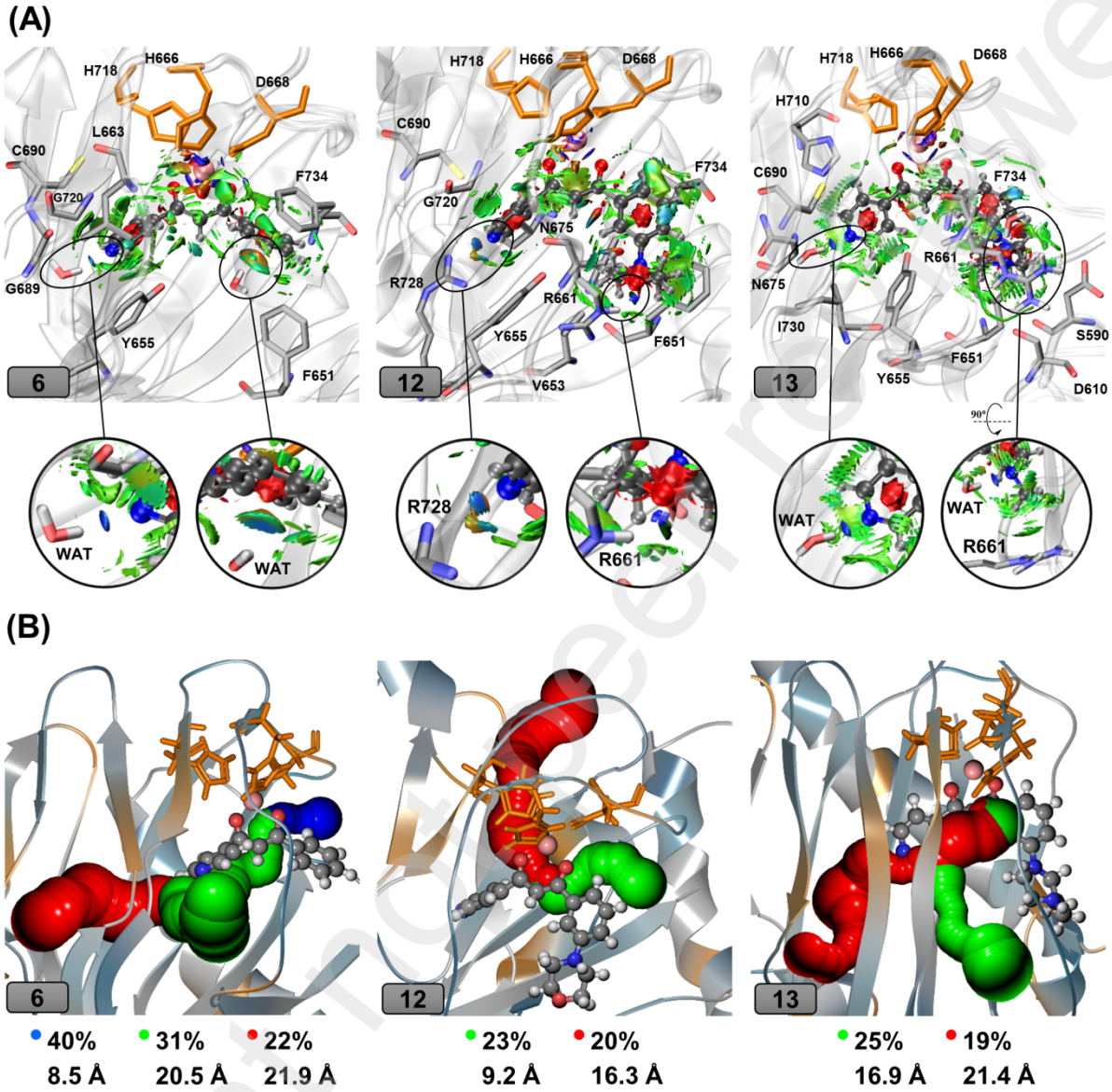


Figure 4. (A) The plot of the non-covalent interactions between compounds **6**, **12**, and **13** (labels are at the left bottom of each image) and the surrounding amino acid residues. The inhibitors are given in ball-and-sticks. The Fe(II) is shown in pink sphere, while the other residues of the active site (H666, D668, and H718) and water molecule (in **6** and **12**) are presented in orange sticks. Residues forming non-covalent interactions with the inhibitor are shown in sticks on a ribbon model background of the protein. The isovalue for the NCI visualization is 0.4 au with the color scale of –0.05 au < sign(λ_2)ρ < 0.05 au. Red surfaces are related to repulsive interactions, while green surfaces denote weak interactions like Van der Waals and blue ones show strong attractive interactions such as hydrogen bonds. **(B)** O_2-transporting tunnels with largest calculated percentages along the trajectory observed in compounds **6**, **12**, and **13**. Calculated tunnels are colored in blue, green, and red, respectively, based on the tunnel's length (Å) (blue: shortest, red: longest).

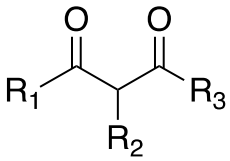
Furthermore, as shown in Figure 4A, compound **12** has attractive interactions with residues F651, V653, Y655, R661, N675, C690, G720, R728, and F734. Compound **13** shows interaction with S590 (weak), D610 (weak), F651, Y655, R661, N675, C690, H710, I730, and F734. The number of NCI surfaces for compound **6** with the rest of the protein is reduced and only shows interactions with F651, Y655, L663, G689, C690, G720, and F734. A closer look into the NCI plots shows compound **12** has two hydrogen bonds with R728 and R661 (Figure 4A-middle), and compound **13** has weak interactions with R661 (Figure 4A-right). In comparison, compound **6** only has two hydrogen bonds with two water molecules in the active site (Figure 4A-left). The interactions of **12** and **13**, particularly with R661 residue, suggest that their potencies might derive from the engagement of morpholine and methylpiperazine rings with LH2-specific R661. The NCI results and the calculated intermolecular interaction energies agree with the measured antagonist potencies of these three inhibitors.

Our previous study also suggested that three major tunnels exist in the apo-LH2, which can transport molecular oxygen to the active site with tunnel availability probabilities of 63% (blue tunnel in Ref. 59), 51% (green tunnel in Ref. 59), and 32% (red tunnel in Ref. 59).[59] Our results also showed that compounds **7** and **8** (IDs: 43- c24 and 44- c25 in Ref. 59) and some other studied molecules in that study not only considerably decrease the availability of these tunnels, but also elongate the tunnels' lengths (see Figure 5 in Ref. 59). As can be seen in Figure 4B, the tunneling results for compounds **6**, **12**, and **13** showed the same trend, in which the availability/length of the tunnels decrease/increase. Interestingly, in the case of compounds **12** and **13**, only two major putative tunnels are available for oxygen transportation.

A final set of modifications were performed, based on **12** and **13**, replacing the pyridine moiety with tetrazole (**16**; IC_{50} = >10 μM, **14**; IC_{50} = >10 μM) and the 1,3-diketone with pyrazole (**17**; IC_{50} = >10 μM, **15**; IC_{50} = >10 μM). The results indicated that pyridine and 1,3-diketone scaffolds are essential for the activity of the studied compounds. In addition, **18** with the toluene group on R_1 position shows no activity, which is in line with the importance of pyridine.

2.3. Selectivity profiles

The HTS hit compounds **(1** and **2)** and the four most potent compounds (**6**, **11**, **12**, and **13)** were evaluated using an enzymatic activity assay against all three LH isoforms for selectivity profiling (Table 1). Interestingly, **12** and **13** demonstrated up to 9-fold selectivity for LH2 over other LH isoforms, which indicates that the engagement of morpholine and methylpiperazine rings with LH2-specific residue R661 contributes to the selective binding to LH2. In the case of **6,** the selectivity shows 5-fold, while other compounds show less than 3-fold or no selectivity. Its moderate selectivity is attributed to the hydrophobic interactions between the benzene ring of **6** and hydrophobic residues (F651 and F734) at the entry of the active site to aid the stable binding to LH2, as shown in Figure 4A.



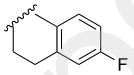
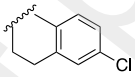
Compound	R_1	R_2	R_3	IC_{50} (μM)		
				LH2	LH1	LH3
1	3-Py*	H	CF_3	3.4 ± 1.1	4.3 ± 1.3	5.2 ± 4.0
2	4-Py	H	CF_3	1.9 ± 0.6	4.8 ± 0.9	4.8 ± 0.8
3	4-F-Ph*	H	CF_3	7.6 ± 2.9		
4	4-NO_2-Ph	H	CF_3	>10		
5	3-Py	H	Ph	5.0 ± 1.9		
6	4-Py	H	Ph	1.0 ± 0.9	6.4 ± 8.0	5.2 ± 6.0
7	3-Py	H	4-CF_3-Ph	N/A		
8	3-Py	H	4-NO_3-Ph	>10		
9	3-Py			>10		
10	3-Py			N/A		
11	3-Py	H	3-morpholine-Ph	1.5 ± 0.5	1.7 ± 0.8	4.8 ± 1.9
12	4-Py	H	3-morpholine-Ph	0.3 ± 0.1	2.3 ± 0.9	2.7 ± 2.9
13	4-Py	H	3-Me-piperazine-Ph	0.5 ± 0.1	1.6 ± 0.6	4.7 ± 8.8
14	Tetrazole	H	3-Me-piperazine-Ph	>10		
15	4-Py	pyrz	3-Me-piperazine-Ph	>10		
16	Tetrazole	H	3-morpholine-Ph	>10		
17	4-Py	pyrz	3-morpholine-Ph	>10		
18	4-Me-Ph	H	3-morpholine-Ph	N/A		

Table 1. Selectivity and potency of LH2 inhibitors measured by bioluminescence-based enzymatic activity *in vitro* assay. *Py = pyridyl, Ph = phenyl, pyrz = pyrazole.

To further assess the selectivity of compounds **6**, **12**, and **13**, we tested them against other Fe(II)/2OG-dependent enzymes. These enzymes included FTO (fat mass and obesity-associated protein, which preferentially demethylates N^6-methyladenosine in RNA[52]), JMJD2A (the histone demethylase Jumonji domain-containing protein 2A[53]), EGLN1 (Egl-9 family hypoxia-inducible factor 1, also called prolyl hydroxylase domain-containing protein 2 or PHD2[54]), and mimivirus L230, which hydroxylates lysine and glycosylates hydroxylysine residues on collagen[10]. The percent inhibition of each of these enzymes at a dose of 11 μM compound (**6**, **12**, or **13**) is shown in Figure 5. From these results, all three compounds show greater inhibition of LH2b over the other enzymes. Compound **6** shows lower selectivity for LH2b than the other compounds as it inhibits JMJD2A similarly to LH2b.

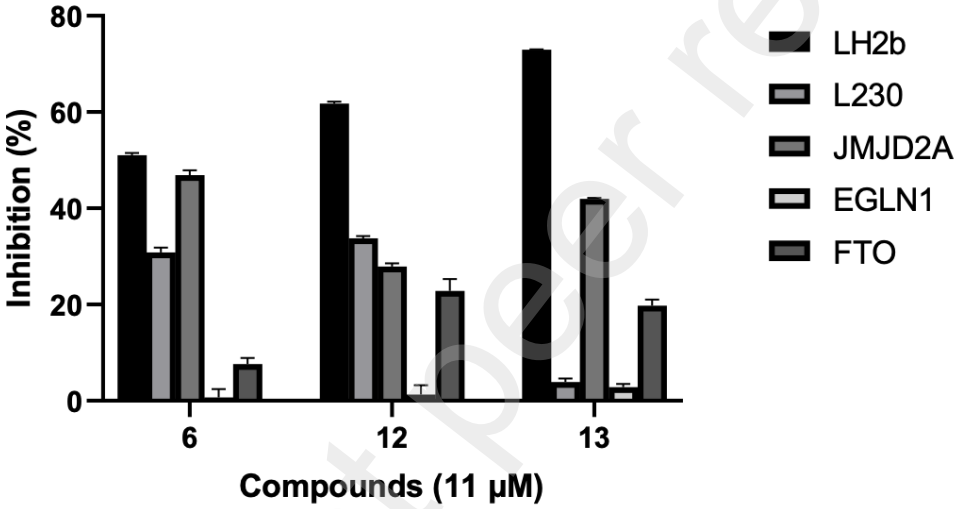


Figure 5. Selectivity profile of compounds against other Fe(II)/2OG-dependent enzymes. Compounds **6**, **12**, and **13** were tested at 11 μM for inhibition of LH2b and other Fe(II)/2OG-dependent enzymes L230, JMJD2A, EGLN1, and FTO. Inhibition (%) is the percent decrease in luminescence signal relative to DMSO controls in the luminescence-based enzymatic activity assay. All samples were performed in duplicate and error bars indicate standard error.

2.4. Migration assay

High expression of LH2 has been shown to drive cell migration.[56] To check whether the most potent compounds also have selective activity against LH2 in cells, we evaluated representative compound **13**'s inhibitory effect on the 344SQ WT and 344SQ-LH2KO cell lines. Compound **13** shows a dose-dependent inhibitory effect on the 344SQ WT cells (Figures 6C and D) at concentrations that do not influence cell proliferation of metastatic murine NSCLC (Non-small cell lung cancer) cell lines (344SQ and 531LN2, Figures 6A and B) with no inhibition on the migration of LH2KO cells (Figures 6C and D), strongly suggesting that **13** shows selectivity for LH2 *in vitro* and in cells.

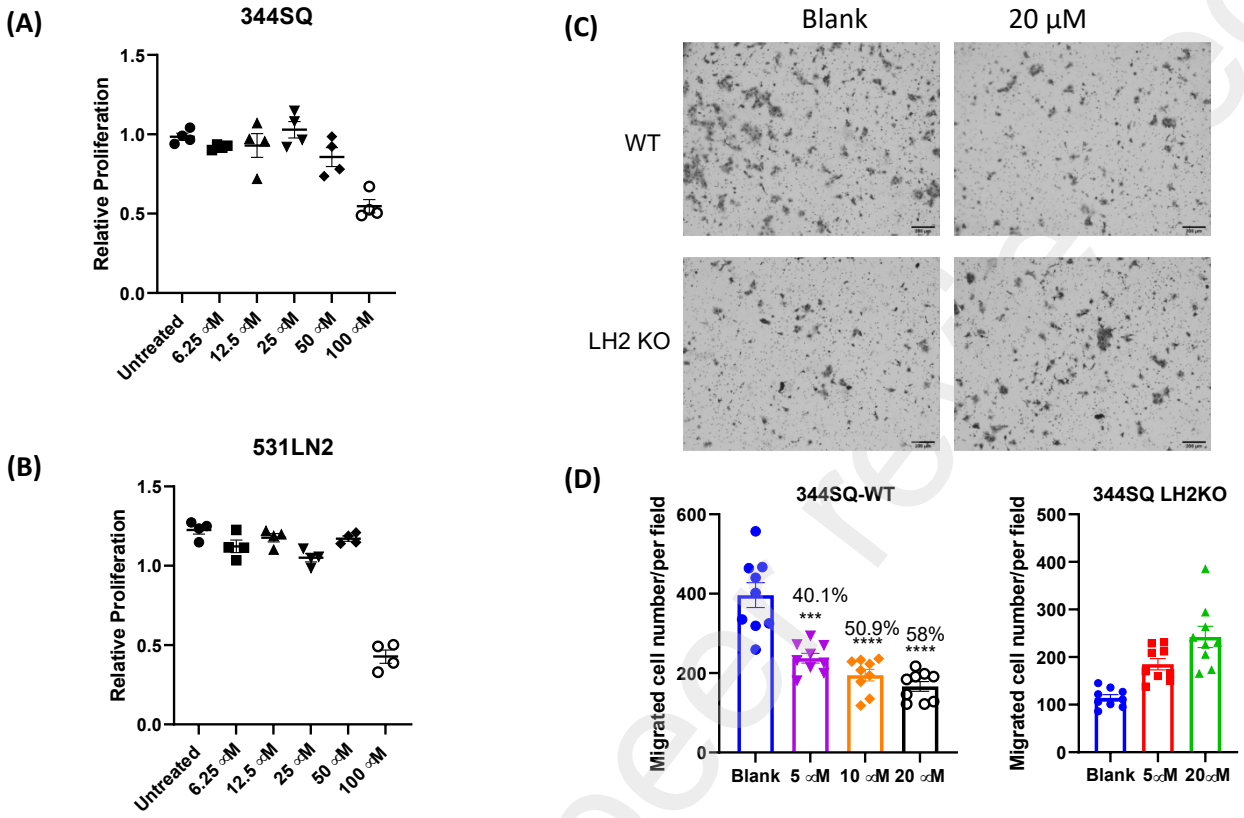


Figure 6. Compound **13** suppresses cell migration of 344SQ lung adenocarcinoma cells (LH2b highly expressed) but not proliferation. **(A and B)** The cell proliferation assay. Graphs showing the cell proliferation for **A)** 344SQ and **B)** 531LN2 with varying concentrations of **13**. **(C and D)** The cell migration assay. Migrated 344SQ cells (WT and LH2 KO) in Boyden chambers were **C)** imaged blank and 20 µM and **D)** counted after incubation with varying concentrations of **13**.

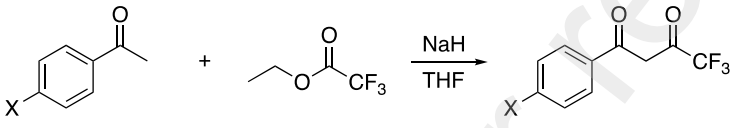
3. Conclusion

In summary, a series of eighteen 1,3-diketone analogs (compounds **1–18)** were designed, synthesized, and evaluated for LH2 inhibition by a luciferase-based *in vitro* enzymatic activity and cell migration assays. The pyridine group at R_1 was essential for H-bonding interactions with LH2 active-site residues. R_3 was optimized to incorporate a benzene ring to replace the trifluoromethyl group of HTS compounds. Meta substitution of the benzene ring benefits the inhibitory activity against LH2, while the modification of the para position reduces the potency due to the hindrance by F651. Compounds **12** and **13,** in which morpholine and 1-methyl piperazine were incorporated on the meta position of benzene, show enhanced potencies and selectivity for LH2 over the other two isoforms and Fe(II)/2OG-dependent enzymes. QM/MM and docking studies demonstrated the presence of strong interactions between the ligands and the enzyme, which supports the inhibitory activities and the selectivity of **12** and **13**. Migration assays in the 344SQ lung adenocarcinoma cell line revealed that **13** shows anti-migration activity in a dose-dependent manner. Since no effective LH2 inhibitors were reported, the results on 1,3-diketone analogs in this study should provide insights into developing antagonists to fight cancer metastasis.

4. Experimental section

4.1. Chemistry

All reagents and solvents were purchased from Aldrich, TCI, and Acros and used without further purification. **1** and **2** were purchased from Oakwood Chemical. NMR spectra were recorded on Agilent MR 400 MHz and Bruker AVIII 500 MHz instruments. The NMR spectra were referenced to solvent, and the spectroscopic solvents were purchased from Cambridge Isotope Laboratories. Chemical ionization (CI) and electrospray ionization (ESI) mass spectra were recorded on a VG ZAB-2E instrument and a VG AutoSpec apparatus. TLC analyses were carried out using Sorbent Technologies silica gel (200 mm) sheets. Column chromatography was performed on Sorbent silica gel 60 O (40–63 mm) or neutral alumina (50–200 mm, Brockmann grade II).



1a X = F
1b X = NO_2

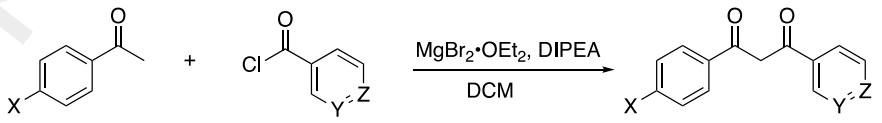
2a

3 X = F
4 X = NO_2

*General procedure for the syntheses of **3** and **4** (**4** as an example)*: A THF solution (5 mL) of sodium hydride (60 % dispersion in mineral oil) (88 mg, 2.2 mmol) was cooled to 0 ºC in an ice bath under Ar atmosphere. A THF solution (2 mL) of 4'-nitroacetophenone (330 mg, 2 mmol) was dropwise, and the mixture was stirred at 0 ºC for 30 min. Ethyl trifluoroacetate (0.262 mL, 2.2 mmol) was dropwise to the solution and stirred at RT for 12 h. The reaction was concentrated and diluted in ethyl acetate, washed with 1N HCl and brine, and dried over Na_2SO_4. The crude was purified by column chromatography over silica gel (ethyl acetate:hexane = 1:2, eluent) to obtain a yellow oily product recrystallized in ethyl acetate/dichloromethane/hexane mixture to afford the compound.

4,4,4-trifluoro-1-(4-fluorophenyl)butane-1,3-dione (***3***) : White solid; Yield: 85% ^{1}H NMR (400MHz, DMSO-d6): δ 8.23-8.19 (d, 2H, Ar*H*), 7.43-7.39 (d, 2H, Ar*H*), 7.00 (s, 1H, C=C*H*). ^{13}C NMR (100 MHz, DMSO-d_6): δ 185.1, 173.5, 173.1, 167.1, 164.6, 131.3, 129.4, 129.3, 118.7, 116.4, 116.2, 93.2. HRMS (ESI) *m/z* 235.0382 calcd for $C_{10}H_7F_4O_2$, found 235.0382.

*4,4,4-trifluoro-1-(4-nitrophenyl)butane-1,3-dione (**4**)* : Yellow solid; Yield: 75% ^{1}H NMR (500MHz, DMSO-d_6): δ 8.29-8.27 (d, 2H, Ar*H*), 8.13-8.11 (d, 2H, Ar*H*), 6.38 (s, 1H, C=C*H*). ^{13}C NMR (125 MHz, DMSO-d_6): δ 185.8, 172.3, 172.1, 171.9, 149.3, 144.0, 128.6, 123.8, 122.1, 120.0, 117.6, 115.3, 90.5. HRMS (M_{mi}) *m/z* 261.0249 (M+) calcd for $C_{10}H_6F_3NO_4$, found 261.0248



1b X = NO_2
1c X = CF_3
1f X = H

2d Y = N, Z = H
2e Y = H, Z = N

5 X = H, Y = N, Z= H
6 X = H, Y = H, Z = N
7 X = CF_3, Y = N, Z = H
8 X = NO_2, Y = N, Z = H

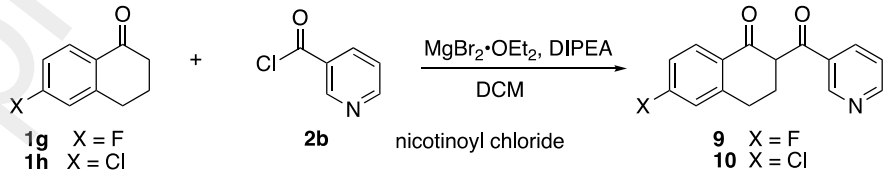
*General procedure for the syntheses of **5–8** (**5** as an example)*: To a dichloromethane solution (5 mL) of acetophenone (0.117 mL, 1 mmol), nicotinyl chloride hydrochloride (196 mg, 1.1 mmol) was added under Ar atmosphere. Magnesium bromide ethyl etherate (697.2 mg, 2.7 mmol) was added and stirred for 15 min. N,N-Diisopropylethylamine (0.82 mL, 4.7 mmol) was dropwise over 1 min. The reaction was stirred at RT for 12 h. The resulting crude was diluted with H_2O/methanol, and 1N HCl was added to the mixture until pH 4-5 was achieved. The aqueous crude was extracted by dichloromethane, and the organic phase was dried over Na_2SO_4. The crude was purified by column chromatography over silica gel (ethyl acetate:hexane = 1:3, eluent) to obtain a yellow solid.

1-phenyl-3-(pyridin-3-yl)propane-1,3-dione (**5**): Yellow solid, Yield: 29% ^{1}H NMR (500MHz, $CDCl_3$): δ 9.23 (s, 1H, Py*H*), 8.80-8.79 (d, 1H, Py*H*), 8.39-8.37(d, 1H, Py*H*), 8.02-8.00 (d, 2H, Ar*H*), 7.60-7.58 (m, 1H, Ar*H*), 7.53-7.52 (m, 3H, Ar*H,* Py*H*), 6.88 (s, 1H, C=C*H*). ^{13}C NMR (125MHz, $CDCl_3$): δ 186.7, 182.5, 151.3, 147.2, 135.8, 134.9, 133.1, 131.9, 128.9, 127.4, 124.1, 93.6. HRMS (ESI) *m/z* 226.0863 $(M+H)^+$ calcd for $C_{14}H_{11}NO_2$, found 226.0865.

1-phenyl-3-(pyridin-4-yl)propane-1,3-dione (**6**): Yellow solid, Yield: 22% ^{1}H NMR (400MHz, $CDCl_3$): δ 8.82-8.80 (d, 2H, Py*H*), 8.02-8.00 (d, 2H, Ar*H*), 7.83-7.81(d, 2H, Py*H*), 7.62-7.58 (m, 1H, Ar*H*), 7.54-7.50 (m, 2H, Ar*H*), 6.89 (s, 1H, C=C*H*). ^{13}C NMR (100MHz, $CDCl_3$): δ 188.5, 180.9, 150.1, 142.8, 135.2, 133.2, 128.8, 127.5, 120.6, 94.1. HRMS (ESI) *m/z* 226.0863 $(M+H)^+$ calcd for $C_{14}H_{11}NO_2$, found 226.0863.

1-(pyridin-3-yl)-3-(4-(trifluoromethyl)phenyl)propane-1,3-dione (**7**): Yellow solid, Yield: 20% ^{1}H NMR (400MHz, $CDCl_3$): δ 9.24 (s, 1H, Py*H*), 8.80 (d, 1H, Py*H*), 8.34-8.32 (d, 1H, Py*H*), 8.11-8.09 (d, 2H, Ar*H*), 7.78-7.76 (d, 2H, Ar*H*), 7.52-7.49 (m, 1H, Py*H*), 6.89 (s, 1H, C=C*H*). ^{13}C NMR (100MHz, $CDCl_3$): δ 184.3, 184.2, 152.5, 148.0, 138.1, 134.3, 134.0, 131.2, 127.6, 125.8, 124.9, 123.9, 122.2, 94.0. ^{19}F NMR (376 MHz, $CDCl_3$) δ -63.10. HRMS (ESI) *m/z* 294.0736 $(M+H)^+$ calcd for $C_{14}H_{10}N_2O_4$, found 294.0741.

1-(4-nitrophenyl)-3-(pyridin-3-yl)propane-1,3-dione (**8**): Yellow solid, Yield: 15% ^{1}H NMR (500MHz, DMSO-d_6): δ 9.40 (s, 1H, Py*H*), 8.85 (d, 1H, Py*H*), 8.57-8.55 (d, 1H, Py*H*), 8.44-8.38 (d, 4H, Ar*H*), 7.68-7.59 (d, 1H, Py*H*), 7.59 (s, 1H, C=C*H*). ^{13}C NMR (125MHz, DMSO-d_6): δ 186.1, 182.1, 153.7, 150.4, 149.1, 140.1, 136.0, 130.8, 129.3, 124.6, 124.4, 95.9. HRMS (ESI) *m/z* 271.0713 $(M+H)^+$ calcd for $C_{14}H_1N_2O_4$, found 271.0716.

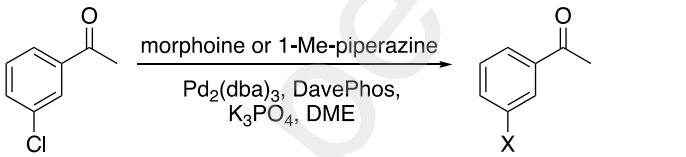


*General procedure for the syntheses of **9** and **10** (**9** as an example)*: To a dichloromethane solution (5 mL) of 6-fluoro-1-teralone (164.2 mg, 1 mmol), nicotinyl chloride hydrochloride (196

mg, 1.1 mmol) was added under Ar atmosphere. Magnesium bromide ethyl etherate (697.2 mg, 2.7 mmol) was added and stirred for 15 min. N,N-Diisopropylethylamine (0.82 mL, 4.7 mmol) was dropwise over 1 min. The reaction was stirred at RT for 12 h. The resulting crude was diluted with H_2O/methanol, and 1N HCl was added to the mixture until pH 4-5 was achieved. The aqueous crude was extracted by dichloromethane, and the organic phase was dried over Na_2SO_4. The crude was purified by column chromatography over silica gel (ethyl acetate:hexane = 1:3, eluent) to obtain a yellow solid.

6-fluoro-2-nicotinoyl-3,4-dihydronaphthalen-1(2H)-one (**9**): Yellow solid, Yield: 23% ^{1}H NMR (400MHz, $CDCl_3$): δ 8.85 (s, 1H, Py*H*), 8.74-8.73 (d, 1H, Py*H*), 8.08-8.05 (m, 2H, Py*H*), 7.56-7.53 (d, 1H, Ar*H*), 7.08-7.04 (t, 1H, Ar*H*), 6.95-6.92 (d, 1H, Ar*H*), 2.87-2.76 (m, 4H, CH_2). ^{19}F NMR (376 MHz, $CDCl_3$) δ -104.57. HRMS (ESI) *m/z* 270.0925 $(M+H)^+$ calcd for $C_{16}H_{12}FNO_2$, found 270.0932.

6-chloro-2-nicotinoyl-3,4-dihydronaphthalen-1(2H)-one (**10**): Yellow solid, Yield: 56% ^{1}H NMR (400MHz, $CDCl_3$): δ 8.85 (s, 1H, Py*H*), 8.74-8.73 (d, 1H, Py*H*), 8.03-8.01 (d, 1H, Py*H*), 7.98-7.96 (d, 1H, Py*H*) 7.53-7.50 (t, 1H, Ar*H*), 7.37-7.34 (d, 1H, Ar*H*), 7.25-7.24 (s, 1H, Ar*H*), 2.84-2.77 (m, 4H, CH_2). ^{13}C NMR (100MHz, $CDCl_3$): δ 184.0, 180.7, 150.0, 147.6, 143.1, 139.2, 136.8, 130.2, 128.2, 127.8, 127.5, 123.8, 106.5, 28.4, 24.4. HRMS (ESI) *m/z* 286.0629 $(M+H)^+$ calcd for $C_{16}H_{12}ClNO_2$, found 286.0634.



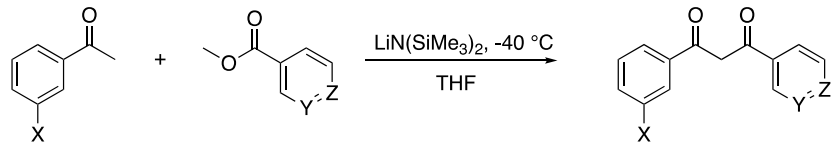
1i X = morpholine
1j X = 1-Me-piperazine

*General procedure for the syntheses of **1i** and **1j** (**1j** as an example)*: A two-neck round bottom flask was evacuated and backfilled with Ar. $Pd_2(dba)_3$ (82.4 mg, 0.09 mmol), DavePhos (70.84 mg, 0.18 mmol), and K_3PO_4 (891.5 mg, 4.2 mmol) were placed in the flask. The flask was evacuated, backfilled with Ar, and capped with a rubber septum. 3-chloroacetophenone (0.4 mL, 3,0 mmol) and 1-methylpiperazine (0.4 mL, 3.6 mmol) were added by using a syringe through the septum, and the mixture was stirred at 100 ºC under Ar for 12 h. The solution was cooled to RT, diluted with diethyl ether, and filtered through Celite. The crude was purified by silica gel column chromatography (dichloromethane:methanol = 40:1, eluent) to obtain a brown liquid.

1-(3-morpholinophenyl)ethan-1-one (**1i**): Brown solid, Yield: 63% ^{1}H NMR (400MHz, $CDCl_3$): δ 7.48-7.47 (s, 1H, Ar*H*), 7.42-7.40 (d, 1H, Ar*H*), 7.35-7.31 (t, 1H, Ar*H*), 7.10-7.07 (d, 1H, Ar*H*) 3.85-3.82 (m, 4H, CH_2), 3.19-3.16 (m, 4H, CH_2), 2.56 (s, 3H, C=O-CH_3). ^{13}C NMR (100MHz, $CDCl_3$): δ 198.3, 151.4, 138.0, 129.3, 120.3, 114.4, 66.7, 49.0, 26.7.

1-(3-(4-methylpiperazin-1-yl)phenyl)ethan-1-one (**1j**): Brown liquid, Yield : 74% ^{1}H NMR (400MHz, $CDCl_3$): δ 7.51-7.50 (s, 1H, Ar*H*), 7.43-7.41 (d, 1H, Ar*H*), 7.36-7.32 (t, 1H, Ar*H*), 7.13-7.11 (d, 1H, Ar*H*) 3.32-3.29 (m, 4H, CH_2), 2.66-2.64 (m, 4H, CH_2), 2.58 (s, 3H, C=O-CH_3),

2.40 (s, 3H, N-CH_3). ^{13}C NMR (100MHz, $CDCl_3$): δ 198.5, 151.2, 138.0, 129.3, 120.6, 120.1, 114.9, 54.8, 48.6, 45.9, 26.8.



1i X = morpholine
1j X = 1-Me-piperazine

2f Y = N, Z = H
2g Y = H, Z = N
2h Y = H, Z = Me

11 X = morpholine, Y = N, Z= H
12 X = morpholine, Y = H, Z = N
13 X = 1-Me-piperazine, Y = H, Z = N
18 X = morpholine, Y = H, Z = Me

*General procedure for the syntheses of **11–13**, and **18** (**12** as an example)*: 1-(3-morpholinophenyl)ethan-1-one (**1i**) (102.6 mg, 0.5 mmol) was placed in an 8 mL vial and evacuated and backfilled with Ar. THF (2 mL) was added to the vial and cooled down to -40 ºC. $LiN(SiMe_3)_2$ (1.0 M in THF) (1.1 mL, 1.1 mmol) was dropwise to the solution over 20 min. The mixture was stirred at -40 ºC for 10 min, and methyl isonicotinate (75.4 mg, 0.55 mmol) in THF (2 mL) was added dropwise and stirred at R. T. for 6 h. The resulting orange solution was dried *in vaccuo*, and the crude was purified by silica gel column chromatography (ethylacetate:hexanes = 1:5, eluent) to afford the compound.

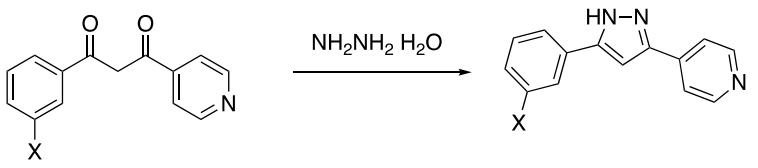
*1-(3-morpholinophenyl)-3-(pyridin-3-yl)propane-1,3-dione (**11**)*: Brown solid, Yield: 60%, ^{1}H NMR (500MHz, $CDCl_3$): δ 16.67 (s, 1H, O*H*), 9.24 (s, 1H, Py*H*), 8.78 (s, 1H, Py*H*), 8.38-8.36 (d, 1H, Py*H*), 7.58 (s, 1H, Ar*H*), 7.58-7.52 (t, 1H, Py*H*), 7.49-7.47 (d, 1H, Ar*H*), 7.42-7.38 (t, 1H, Ar*H*), 7.19-7.16 (d, 1H, Ar*H*), 6.86 (s, 1H, C=C*H*), 3.92-3.89 (m, 4H, CH_2), 3.27-3.25 (m, 4H, CH_2). ^{13}C NMR (125MHz, $CDCl_3$): δ 187.4, 182.1, 151.4, 151.2, 147.2, 136.1, 136.1, 132.0, 129.8, 124.4, 120.6, 119.5, 114.4, 94.0, 66.8, 49.5. HRMS (ESI) *m/z* 311.1390 $(M+H)^+$ calcd for $C_{18}H_{18}N_2O_3$, found 311.1389.

*1-(3-morpholinophenyl)-3-(pyridin-4-yl)propane-1,3-dione (**12**)*: Brown solid, Yield: 49%, ^{1}H NMR (400MHz, $CDCl_3$): δ 16.52 (s, 1H, O*H*), 8.80 (s, 2H, Py*H*), 7.82-7.81 (d, 2H, Py*H*), 7.55 (s, 1H, Ar*H*), 7.47-7.45 (d, 1H, Ar*H*), 7.41-7.38 (t, 1H, Ar*H*), 7.15-7.13 (d, 1H, Ar*H*), 6.86 (s, 1H, C=C*H*), 3.90-3.88 (m, 4H, CH_2), 3.26-3.23 (m, 4H, CH_2). ^{13}C NMR (100MHz, $CDCl_3$): δ 189.1, 180.3, 151.6, 150.0, 142.8, 136.2, 129.5, 120.6, 120.3, 119.0, 114.0, 94.4, 66.8, 49.0. HRMS (ESI) *m/z* 311.1390 $(M+H)^+$ calcd for $C_{18}H_{18}N_2O_3$, found 311.1393.

*1-(3-(4-methylpiperazin-1-yl)phenyl)-3-(pyridin-4-yl)propane-1,3-dione (**13**)*: Brown solid, Yield: 33%, ^{1}H NMR (400MHz, $CDCl_3$): δ 16.50 (s, 1H, O*H*), 8.81-8.79 (d, 2H, Py*H*), 7.79-7.77 (d, 2H, Py*H*), 7.56 (s, 1H, Ar*H*), 7.52-7.50 (d, 1H, Ar*H*), 7.43-7.39 (t, 1H, Ar*H*), 7.16-7.13 (d, 1H, Ar*H*), 6.85 (s, 1H, C=C*H*), 3.58-3.55 (m, 4H, CH_2), 3.05 (m, 4H, CH_2), 2.68 (s, 3H, CH_3). ^{13}C NMR (100MHz, $CDCl_3$): δ 189.1, 190.0, 151.5, 150.8, 142.4, 136.4, 129.6, 120.8, 120.5, 119.0, 114.7, 94.4, 54.9, 48.6, 45.9. HRMS (ESI) *m/z* 324.1707 $(M+H)^+$ calcd for $C_{19}H_{21}N_3O_2$, found 324.1708.

*1-(3-morpholinophenyl)-3-(p-tolyl)propane-1,3-dione (**18**)*: Yellow solid, Yield: 30%, ^{1}H NMR (400MHz, $CDCl_3$): δ 16.95 (s, 1H, O*H*), 7.90-7.88 (d, 2H, Ar*H*), 7.56 (s, 1H, Ar*H*), 7.47-7.45 (d, 1H, Ar*H*), 7.40-7.36 (t, 1H, Ar*H*), 7.30-7.28 (d, 2H, Ar*H*), 7.12-7.10 (d, 1H, Ar*H*), 6.81 (s, 1H, C=C*H*), 3.91-3.89 (m, 4H, CH_2), 3.26-3.23 (m, 4H, CH_2), 2.43 (s, 3H, CH_3). ^{13}C NMR

(100MHz, $CDCl_3$): δ 185.8, 185.6, 143.3, 136.7, 132.8, 129.4, 129.4, 127.2, 119.6, 118.8, 114.0, 93.0, 66.8, 49.2, 21.7. HRMS (ESI) *m/z* 324.1594 $(M+H)^+$ calcd for $C_{20}H_{21}NO_3$, found 324.1592.



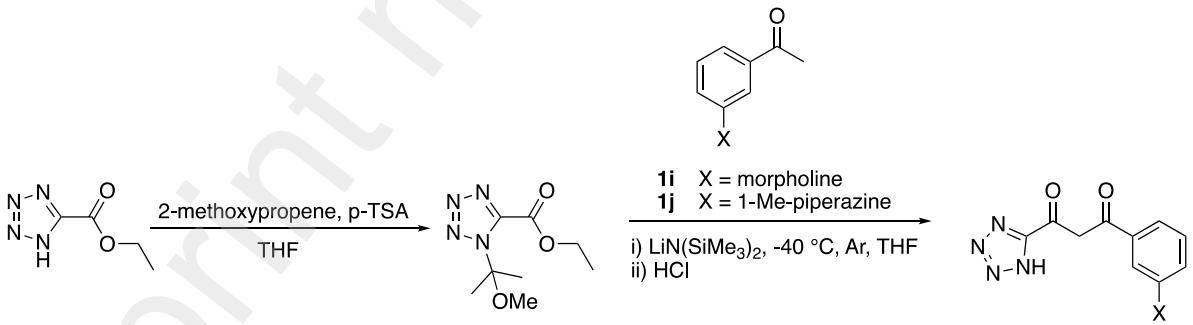
12 X = morpholine
13 X = 1-Me-piperazine

17 X = morpholine
15 X = 1-Me-piperazine

*General procedure for the syntheses of **15** and **17** (**17** as an example)*: 1-(3-morpholinophenyl)-3-(pyridin-4-yl)propane-1,3-dione (**12**) (31 mg, 0.1 mmol) and hydrazine monohydrate (0.1 mL) were refluxed in 10 mL ethanol for 12 h. The reaction was cooled to RT and concentrated *in vacuo*. The crude was recrystallized in dichloromethane/hexane mixture to obtain the product as a white solid.

*1-methyl-4-(3-(3-(pyridin-4-yl)-1H-pyrazol-5-yl)phenyl)piperazine (**15**)*: White solid, Yield: 99%, ^{1}H NMR (500MHz, DMSO-d_6): δ 13.57 (s, 1H, N*H*), 8.62-8.61 (d, 2H, Py*H*), 7.81-7.80 (d, 2H, Py*H*), 7.39 (s, 2H, Ar*H* and C=C*H*(Hpz)), 7.31-7.28 (t, 1H, Ar*H*), 7.23 (d, 1H, Ar*H*), 6.94-6.93 (d, 1H, Ar*H*), 3.23-3.21 (m, 4H, CH_2), 2.50-2.48 (m, 4H, CH_2), 2.24 (s, 3H, CH_3). ^{13}C NMR (125MHz, DMSO-d_6): δ 151.9, 150.6, 130.1, 119.9, 116.1, 115.6, 112.3, 101.2, 55.0, 48.3, 46.2. HRMS (ESI) *m/z* 320.1870 $(M+H)^+$ calcd for $C_{19}H_{21}N_5$, found 320.1879.

*4-(3-(3-(pyridin-4-yl)-1H-pyrazol-5-yl)phenyl)morpholine (**17**)*: White solid, Yield: 99%, ^{1}H NMR (500MHz, DMSO-d_6): δ 13.56 (s, 1H, N*H*), 8.62-8.61 (d, 2H, Py*H*), 7.81-7.80 (d, 2H, Py*H*), 7.39 (s, 2H, Ar*H* and C=C*H*(Hpz)), 7.34-7.31 (t, 1H, Ar*H*), 7.27-7.26 (d, 1H, Ar*H*), 6.96-6.94 (d, 1H, Ar*H*), 3.79-3.77 (m, 4H, CH_2), 3.21-3.19 (m, 4H, CH_2). ^{13}C NMR (125MHz, DMSO-d_6): δ 151.5, 150.2, 129.6, 119.4, 116.1, 114.9, 111.7, 101.0, 66.1, 48.3. HRMS (ESI) *m/z* 307.1553 $(M+H)^+$ calcd for $C_{18}H_{18}N_4O$, found 307.1561.



16 X = morpholine
14 X = 1-Me-piperazine

*General procedure for the syntheses of **14** and **16** (**16** as an example)*: To a THF solution (5 mL) of ethyl 1*H*-tetrazolium-5-carboxylate (213 mg, 1.5 mmol), p-toluenesulfonic acid monohydrate (3 mg, 0.015 mmol) and 2-methoxypropene (0.16 mL, 1.65 mmol) were added, and the mixture was stirred at RT for 1 hour. *(Mixture 1)*

1-(3-morpholinophenyl)ethan-1-one (**1i**) (308 mg, 1.5 mmol) was placed in a flask and evacuated and backfilled with Ar. THF (5 mL) was added to the vial and cooled down to -40 ºC. $LiN(SiMe_3)_2$ (1.0 M in THF) (3.3 mL, 3.3 mmol) was dropwise to the solution over 20 min. The mixture was stirred in -40 ºC for 10 min. *(Mixture 2)*
Mixture 2 was added dropwise to mixture 1 slowly and stirred at RT for 6h. The reaction was dried *in vacuo,* and diethyl ether was added (30 mL). The resulting precipitate was filtered and dissolved in 50 mL of water. 6 N HCl was dropwise with constant swirling until yellow precipitates started to form. The mixture flask was placed in an ice bath for 3 h. The resulting precipitates were filtered and dried to obtain the product as an orange solid.

*1-(3-(4-methylpiperazin-1-yl)phenyl)-3-(1H-tetrazol-5-yl)propane-1,3-dione (**14**)*: Yellow solid, Yield: 15%, ^{1}H NMR (500MHz, DMSO-d_6): δ 10.67 (s, 1H, N*H*), 7.64 (s, 1H, Ar*H*), 7.59-7.58 (d, 1H, Ar*H*), 7.50-7.47 (t, 1H, Ar*H*), 7.44 (s, 1H, C=C*H*), 7.35-7.34 (d, 1H, Ar*H*), 4.00-3.17 (m, 8H, CH_2), 2.83 (s, 3H, CH_3). ^{13}C NMR (125MHz, DMSO-d_6): δ 186.1, 174.0, 150.5, 135.0, 130.5, 121.6, 119.4, 114.6, 96.9, 52.5, 45.7, 42.4. HRMS (ESI) *m/z* 315.1564 $(M+H)^+$ calcd for $C_{15}H_{18}N_6O_2$, found 315.1570.

*1-(3-morpholinophenyl)-3-(1H-tetrazol-5-yl)propane-1,3-dione (**16**)*: Yellow solid, Yield: 13%, ^{1}H NMR (500MHz, DMSO-d_6): δ 7.53-7.52 (d, 1H, Ar*H*), 7.51 (s, 1H, Ar*H*), 7.45-7.42 (t, 1H, Ar*H*), 7.33 (s, 1H, C=C*H*), 7.29-7.27 (d, 1H, Ar*H*), 3.77-3.75 (m, 4H, CH_2), 3.21-3.20 (m, 4H, CH_2). ^{13}C NMR (125MHz, DMSO-d_6): δ 186.3, 173.8, 152.0, 134.6, 130.3, 121.0, 118.7, 113.4, 96.7, 66.5, 48.6. HRMS (ESI) *m/z* 324.1067 $(M+Na)^+$ calcd for $C_{14}H_{15}N_5O_3$, found 324.1069.

4.2. Protein purification and in vitro enzymatic activity assay

Human LH1-3 recombinant proteins were purified from CHO cell-derived conditioned medium samples as described previously.[11] In brief, LH1-3 expression vectors were transiently transfected in new Gibco™ ExpiCHO™ cells (Thermo Fisher Scientific, Waltham, MA) with polyethylenimine and expressed as a secreted protein with N-terminal His_8 and human growth hormone (hGH) tags. The LH1-3–containing conditioned medium samples were harvested by centrifugation at 7,000 rpm for 10 min, filtered through 0.22 μm EMD Millipore Stericup™ Sterile Vacuum Filter Units (EMD Millipore, Billerica, MA), concentrated and buffer-exchanged into Nickel-binding buffer (20 mM Tris, pH 8.0, 200 mM NaCl, 15 mM imidazole) using the Centramate™ & Centramate PE Lab Tangential Flow System (Pall Life Sciences, Ann Arbor, MI). The recombinant LH proteins were then purified with immobilized metal affinity chromatography and anion exchange chromatography.

LH enzymatic activity was measured using a luciferase-based assay as described.[57] In brief, the assay was performed in LH reaction buffer (50 mM HEPES pH 7.4, 150 mM NaCl) at 37 °C for 1 h with 1 μM LH enzymes, 10 μM $FeSO_4$, 100 μM 2-OG, 500 μM ascorbate, 1 mM DTT, 0.01% triton x-100, and 1 mM IKGIKGIKG collagen telopeptide mimics. Except for LH recombinant proteins, all reagents were prepared immediately before use. All reagents were dissolved in reaction buffer except for $FeSO_4$, which was prepared in 10 mM HCl, and the pH of the reaction mixture was checked with pH papers to ensure that HCl did not change the overall sample pH. All compounds were pre-incubated with LHs at RT for 30 min before LH enzymatic activity assay. LH activity was measured by detecting succinate production with Succinate-Glo™

kit reagents (Succinate-Glo™ JmjC Demethylase/Hydroxylase Assay, Promega, Madison, WI) according to manufacturers' instructions. Experiments were performed in triplicates, and IC_{50} values were determined by fitting the data using four-parameter logistic regression in SigmaPlot 14.5.

4.3. Selectivity test over other Fe(II)/2OG-dependent enzymes

Human EGLN1 (NP_071334.1) and human FTO (XP_011521615.1) were purchased from Active Motif (cat# 81065 and 31572, respectively). The substrate for EGLN1 (pepEGLN1) was a peptide of HIF-1α residues 556-574 and purchased from Anaspec (cat# AS-61528, DLDLEALAPYIPADDDFQL). The FTO substrate (oligoFTO, 5'-rCrUrU rGrUrC rA/iN6Me-rA/rC rArGrC rArGrA-3') was custom synthesized from Integrated DNA Technologies. JMJD2A (HDM) was a gift from Dr. Martinez at the University of Texas Southwestern Medical Center, and its substrate peptide H3K9me3 was purchased from Cayman (Cat# 10530, ARTKQTARK$(Me)_3$-STGGKA) respectively. L230 protein was synthesized as reported previously[58] and its peptide substrate (GTKGETGLKGII, abbreviated pepL230) was procured from LifeTein. All tested compounds were dissolved in DMSO to form 10 mM stock solutions.

With little modification, selectivity tests were conducted similarly to the activity assay described above. Assays were performed in 50 mM HEPES, pH 7.5, containing 10 μg/mL BSA, 0.01 % Tween-20, 1 mM DTT, 10 μM $FeCl_2$, and 100 μM ascorbate. Briefly, 2.5 μL of 5X compounds were incubated with 5 μL of 2.5X enzyme for 30 min at RT. Compounds **6**, **12**, and **13** were each tested at a final concentration of 11 μM. Enzymatic activity assay were initiated by adding 5 μL 2.5X substrate mixture (2OG with or without each appropriate enzyme-specific substrate). Final concentrations of the enzymes were 1 μM HDM, 0.2 μM EGLN1, 0.5 μM FTO, 1 μM LH2b and 1 μM L230, respectively. Enzyme-specific substrate concentrations were either 6 μM (H3K9me3, pepEGLN1, oligoFTO) or 1 mM (IKGIKGIKG, pepL230), and all reactions contained the additional substrate 2OG at a final concentration of 10 μM. Reactions proceeded for 1 h at RT, followed by processing with Succinate-Glo™ kit according to manufacturer instructions. Glo kit reagents were supplemented with 0.01% Tween-20. Luminescence measurements were collected on a Synergy Neo2 multi-mode plate reader (Agilent) with all samples in 384-well white polystyrene plates (Corning, cat# 3825).

4.4. Migration assay

The migration assay was done in 8.0 μm Boyden Chambers (Corning). Briefly, 20,000 cells were resuspended in 200 μL FBS-free RPMI 1640 (Corning) containing varying concentrations of compound **13,** and the cells were plated in the inner portion of the chambers. RPMI 1640, containing 10 % FBS, was added to the outer portion of the chambers. After incubating overnight, the cells in the chambers were harvested and washed with PBS. To fix the cells on the chamber membrane, the chambers were immersed in 90% ethanol for 20 min. Then the chambers were immersed in 0.1% crystal violet for 10 min. After staining, the chambers were dried, and taken representative images for cell counting. Triplicates were done for each condition, and the migration assay was repeated twice independently. The Student t-test was used to analyze the significance. *** indicates $P<0.001$.

4.5. Computational methods

The preparation of the systems, molecular docking, molecular dynamics simulations (MD), clustering, quantum mechanics/molecular mechanics calculations (QM/MM) and interaction energies ($IE_{QM/MM}$), and the non-covalent interactions (NCI) analyses have been explained in detail previously.[59] Briefly, the SWISS-MODEL server[60] was used to construct the homology model for LH2. After performing 1 ns of equilibration, 5 ns of MD production simulations were performed under NPT ensemble using the Tinker software[61] (Figure S39). The MD trajectories for each structure were then used for a six-dimensional cluster analysis with the *k*-means algorithm (see Figures S41–S46 and Table S1).[62] Each dimension in this analysis corresponds to a distance between the iron and the donor atoms of the coordinated residues ligand, a water molecule (in compounds **6** and **12**), two histidine residues, and an aspartate (see Figure S40). The QM/MM calculations were carried out using LICHEM,[63-64] combining Gaussian16[65] and TINKER.[66] The ωB97X-D/6–31G(d,p)[67-68] level of theory and AMOEBAbio18[69] force field were employed for the QM subsystem and the MM region of the calculations, respectively. Multiwfn[70] was employed to analyze the non-covalent interactions (NCI) between the ligand and the surrounding amino acid residues and solvent/molecular fragments using the promolecular density method.[71] Visual molecular dynamics (VMD) was used to create images and visualize the isosurface values.[72] Intramolecular tunnel analysis was performed and illustrated with Caver Analyst 2.0[73-74] to calculate the possibility of O_2-transporting tunnels in systems with compounds **6**, **12**, and **13**.

Declaration of competing interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by NIH CA070907, CPRIT RP210088, WELCH F-1390 (K.N.D); NIH R00CA225633, University of Kentucky Center for Cancer and Metabolism, funded through the NIH/NIGMS COBRE program under grant number P20 GM121327 (H.G.); NIH R01GM106137, R01GM114237 (P.R); NIH CA251067 (J.M.K); R01GM108583, Computing from XSEDE through allocation TG-CHE160044, UNT CASCaM (partially funded by the NSF Grant Nos. CHE1531468 and OAC-2117247) and UT-Dallas' Cyberinfrastructure and Research Services (G.A.C.).

References

1. Csiszar, K., Lysyl oxidases: a novel multifunctional amine oxidase family. *Prog Nucleic Acid Res Mol Biol* **2001,** *70*, 1-32.
2. Yamauchi, M.; Sricholpech, M., Lysine post-translational modifications of collagen. *Essays Biochem* **2012,** *52*, 113-33.
3. Scietti, L.; Campioni, M.; Forneris, F., SiMPLOD, a Structure-Integrated Database of Collagen Lysyl Hydroxylase (LH/PLOD) Enzyme Variants. *J Bone Miner Res* **2019,** *34* (7), 1376-1382.
4. Hyland, J.; Ala-Kokko, L.; Royce, P.; Steinmann, B.; Kivirikko, K. I.; Myllyla, R., A homozygous stop codon in the lysyl hydroxylase gene in two siblings with Ehlers-Danlos syndrome type VI. *Nat Genet* **1992,** *2* (3), 228-31.
5. Ha-Vinh, R.; Alanay, Y.; Bank, R. A.; Campos-Xavier, A. B.; Zankl, A.; Superti-Furga, A.; Bonafe, L., Phenotypic and molecular characterization of Bruck syndrome (osteogenesis imperfecta with contractures of the large joints) caused by a recessive mutation in PLOD2. *Am J Med Genet A* **2004,** *131* (2), 115-20.
6. Vahidnezhad, H.; Youssefian, L.; Saeidian, A. H.; Touati, A.; Pajouhanfar, S.; Baghdadi, T.; Shadmehri, A. A.; Giunta, C.; Kraenzlin, M.; Syx, D.; Malfait, F.; Has, C.; Lwin, S. M.; Karamzadeh, R.; Liu, L.; Guy, A.; Hamid, M.; Kariminejad, A.; Zeinali, S.; McGrath, J. A.; Uitto, J., Mutations in PLOD3, encoding lysyl hydroxylase 3, cause a complex connective tissue disorder including recessive dystrophic epidermolysis bullosa-like blistering phenotype with abnormal anchoring fibrils and type VII collagen deficiency. *Matrix Biol* **2019,** *81*, 91-106.
7. Salo, A. M.; Cox, H.; Farndon, P.; Moss, C.; Grindulis, H.; Risteli, M.; Robins, S. P.; Myllyla, R., A connective tissue disorder caused by mutations of the lysyl hydroxylase 3 gene. *Am J Hum Genet* **2008,** *83* (4), 495-503.
8. Chen, Y.; Terajima, M.; Yang, Y.; Sun, L.; Ahn, Y. H.; Pankova, D.; Puperi, D. S.; Watanabe, T.; Kim, M. P.; Blackmon, S. H.; Rodriguez, J.; Liu, H.; Behrens, C.; Wistuba, II; Minelli, R.; Scott, K. L.; Sanchez-Adams, J.; Guilak, F.; Pati, D.; Thilaganathan, N.; Burns, A. R.; Creighton, C. J.; Martinez, E. D.; Zal, T.; Grande-Allen, K. J.; Yamauchi, M.; Kurie, J. M., Lysyl hydroxylase 2 induces a collagen cross-link switch in tumor stroma. *J Clin Invest* **2015,** *125* (3), 1147-62.
9. van der Slot, A. J.; Zuurmond, A. M.; van den Bogaerdt, A. J.; Ulrich, M. M.; Middelkoop, E.; Boers, W.; Karel Ronday, H.; DeGroot, J.; Huizinga, T. W.; Bank, R. A., Increased formation of pyridinoline cross-links due to higher telopeptide lysyl hydroxylase levels is a general fibrotic phenomenon. *Matrix Biol* **2004,** *23* (4), 251-7.
10. Eisinger-Mathason, T. S.; Zhang, M.; Qiu, Q.; Skuli, N.; Nakazawa, M. S.; Karakasheva, T.; Mucaj, V.; Shay, J. E.; Stangenberg, L.; Sadri, N.; Pure, E.; Yoon, S. S.; Kirsch, D. G.; Simon, M. C., Hypoxia-dependent modification of collagen networks promotes sarcoma metastasis. *Cancer Discov* **2013,** *3* (10), 1190-205.
11. Yamauchi, M.; Barker, T. H.; Gibbons, D. L.; Kurie, J. M., The fibrotic tumor stroma. *J Clin Invest* **2018,** *128* (1), 16-25.
12. Piersma, B.; Bank, R. A., Collagen cross-linking mediated by lysyl hydroxylase 2: an enzymatic battlefield to combat fibrosis. *Essays Biochem* **2019,** *63* (3), 377-387.
13. Guo, H. F.; Bota-Rabassedas, N.; Terajima, M.; Leticia Rodriguez, B.; Gibbons, D. L.; Chen, Y.; Banerjee, P.; Tsai, C. L.; Tan, X.; Liu, X.; Yu, J.; Tokmina-Roszyk, M.; Stawikowska, R.; Fields, G. B.; Miller, M. D.; Wang, X.; Lee, J.; Dalby, K. N.; Creighton, C. J.; Phillips, G. N., Jr.; Tainer, J. A.; Yamauchi, M.; Kurie, J. M., A collagen glucosyltransferase drives lung adenocarcinoma progression in mice. *Commun Biol* **2021,** *4* (1), 482.
14. Last, J. A.; King, T. E., Jr.; Nerlich, A. G.; Reiser, K. M., Collagen cross-linking in adult patients with acute and chronic fibrotic lung disease. Molecular markers for fibrotic collagen. *Am Rev Respir Dis* **1990,** *141* (2), 307-13.
15. Gerriets, J. E.; Reiser, K. M.; Last, J. A., Lung collagen cross-links in rats with experimentally induced pulmonary fibrosis. *Biochim Biophys Acta* **1996,** *1316* (2), 121-31.

16. Ricard-Blum, S.; Liance, M.; Houin, R.; Grimaud, J. A.; Vuitton, D. A., Covalent cross-linking of liver collagen by pyridinoline increases in the course of experimental alveolar echinococcosis. *Parasite* **1995,** *2* (2), 113-8.
17. Brenner, D. A.; Waterboer, T.; Choi, S. K.; Lindquist, J. N.; Stefanovic, B.; Burchardt, E.; Yamauchi, M.; Gillan, A.; Rippe, R. A., New aspects of hepatic fibrosis. *J Hepatol* **2000,** *32* (1 Suppl), 32-8.
18. Perepelyuk, M.; Terajima, M.; Wang, A. Y.; Georges, P. C.; Janmey, P. A.; Yamauchi, M.; Wells, R. G., Hepatic stellate cells and portal fibroblasts are the major cellular sources of collagens and lysyl oxidases in normal liver and early after injury. *Am J Physiol Gastrointest Liver Physiol* **2013,** *304* (6), G605-14.
19. van der Slot, A. J.; Zuurmond, A. M.; Bardoel, A. F.; Wijmenga, C.; Pruijs, H. E.; Sillence, D. O.; Brinckmann, J.; Abraham, D. J.; Black, C. M.; Verzijl, N.; DeGroot, J.; Hanemaaijer, R.; TeKoppele, J. M.; Huizinga, T. W.; Bank, R. A., Identification of PLOD2 as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. *J Biol Chem* **2003,** *278* (42), 40967-72.
20. Xu, Y.; Zhang, L.; Wei, Y.; Zhang, X.; Xu, R.; Han, M.; Huang, B.; Chen, A.; Li, W.; Zhang, Q.; Li, G.; Wang, J.; Zhao, P.; Li, X., Procollagen-lysine 2-oxoglutarate 5-dioxygenase 2 promotes hypoxia-induced glioma migration and invasion. *Oncotarget* **2017,** *8* (14), 23401-23413.
21. Gilkes, D. M.; Bajpai, S.; Wong, C. C.; Chaturvedi, P.; Hubbi, M. E.; Wirtz, D.; Semenza, G. L., Procollagen lysyl hydroxylase 2 is essential for hypoxia-induced breast cancer metastasis. *Mol Cancer Res* **2013,** *11* (5), 456-66.
22. Noda, T.; Yamamoto, H.; Takemasa, I.; Yamada, D.; Uemura, M.; Wada, H.; Kobayashi, S.; Marubashi, S.; Eguchi, H.; Tanemura, M.; Umeshita, K.; Doki, Y.; Mori, M.; Nagano, H., PLOD2 induced under hypoxia is a novel prognostic factor for hepatocellular carcinoma after curative resection. *Liver Int* **2012,** *32* (1), 110-8.
23. Sada, M.; Ohuchida, K.; Horioka, K.; Okumura, T.; Moriyama, T.; Miyasaka, Y.; Ohtsuka, T.; Mizumoto, K.; Oda, Y.; Nakamura, M., Hypoxic stellate cells of pancreatic cancer stroma regulate extracellular matrix fiber organization and cancer cell motility. *Cancer Lett* **2016,** *372* (2), 210-8.
24. Myllyharju, J.; Kivirikko, K. I., Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* **2004,** *20* (1), 33-43.
25. Chen, Y.; Guo, H.; Terajima, M.; Banerjee, P.; Liu, X.; Yu, J.; Momin, A. A.; Katayama, H.; Hanash, S. M.; Burns, A. R.; Fields, G. B.; Yamauchi, M.; Kurie, J. M., Lysyl Hydroxylase 2 Is Secreted by Tumor Cells and Can Modify Collagen in the Extracellular Space. *J Biol Chem* **2016,** *291* (50), 25799-25808.
26. Pankova, D.; Chen, Y.; Terajima, M.; Schliekelman, M. J.; Baird, B. N.; Fahrenholtz, M.; Sun, L.; Gill, B. J.; Vadakkan, T. J.; Kim, M. P.; Ahn, Y. H.; Roybal, J. D.; Liu, X.; Parra Cuentas, E. R.; Rodriguez, J.; Wistuba, II; Creighton, C. J.; Gibbons, D. L.; Hicks, J. M.; Dickinson, M. E.; West, J. L.; Grande-Allen, K. J.; Hanash, S. M.; Yamauchi, M.; Kurie, J. M., Cancer-Associated Fibroblasts Induce a Collagen Cross-link Switch in Tumor Stroma. *Mol Cancer Res* **2016,** *14* (3), 287-95.
27. Myllyla, R.; Wang, C.; Heikkinen, J.; Juffer, A.; Lampela, O.; Risteli, M.; Ruotsalainen, H.; Salo, A.; Sipila, L., Expanding the lysyl hydroxylase toolbox: new insights into the localization and activities of lysyl hydroxylase 3 (LH3). *J Cell Physiol* **2007,** *212* (2), 323-9.
28. Schegg, B.; Hulsmeier, A. J.; Rutschmann, C.; Maag, C.; Hennet, T., Core glycosylation of collagen is initiated by two beta(1-O)galactosyltransferases. *Mol Cell Biol* **2009,** *29* (4), 943-52.
29. Sricholpech, M.; Perdivara, I.; Nagaoka, H.; Yokoyama, M.; Tomer, K. B.; Yamauchi, M., Lysyl hydroxylase 3 glucosylates galactosylhydroxylysine residues in type I collagen in osteoblast culture. *J Biol Chem* **2011,** *286* (11), 8846-56.
30. Terajima, M.; Taga, Y.; Sricholpech, M.; Kayashima, Y.; Sumida, N.; Maeda, N.; Hattori, S.; Yamauchi, M., Role of Glycosyltransferase 25 Domain 1 in Type I Collagen Glycosylation and Molecular Phenotypes. *Biochemistry* **2019,** *58* (50), 5040-5051.
31. Zuurmond, A.-M; van der Slot-Verhoeven, A.J.; van Dura, E.A.; De Groot, J.; Bank, R.A., Minoxidil exerts different inhibitory effects on gene expression of lysyl hydroxylase 1, 2, and 3: implications for collagen cross-linking and treatment of fibrosis. Matrix Biol. **2005,** 24, 261–270.

32. Shao, S.; Zhang, X.; Duan, L.; Fang, H.; Rao, S.; Liu, W.; Guo, B.; Zhang, X., Lysyl Hydroxylase Inhibition by Minoxidil Blocks Collagen Deposition and Prevents Pulmonary Fibrosis via TGF-β_1/Smad3 Signaling Pathway. Med Sci Monit, **2018,** 24, 8592-8601.
33. Pfeffer, T.; Lignelli, E.; Inoue, H.; Mizikova, I.; Surate Solaligue, D.E.; Steenbock, H.; Myti, D.; Vadasz, I.; Herold, S.; Seeger, W.; Brinckmann, J.; Morty, R.E., Minoxidil Cannot Be Used To Target Lysyl Hydroxylases during Postnatal Mouse Lung Development: A Cautionary Note. *J Pharmacol Exp Ther.* **2020**, *375*, 478-487.
34. Du, Y.; Khan, M.; Fang, N.; Ma, F.; Du, H.; Tan, Z.; Wang, H.; Yin, S.; Wei, X., Berberine Attenuates Cell Motility via Inhibiting Inflammation-Mediated Lysyl Hydroxylase-2 and Glycolysis. *Front. Pharmacol.*, **2022**, *13*, 856777
35. Thaler, R.; Spitzer, S.; Rumpler, M.; Fratzl-Zelman, N.; Klaushofer, K.; Paschalis, E.P.; Varga, F., Differential effects of homocysteine and beta aminopropionitrile on preosteoblastic MC3T3-E1 cells, *Bone*, **2010**, *46*, 703–709.
36. Bennett, I.; Broom, N.J.P.; Cassels, R.; Elder, J.S.; Masson, N.D.; O'Hanlon P,J., Synthesis and antibacterial properties of b-diketone acrylate bioisosteres of pseudomonic acid A. *Bioorg Med Chem Lett* **1999**, *9*,1847–1852.
37. Sheikh, J.; Ingle, V.; Juneja, H., Synthesis of novel antibacterial agents 1-(20,40-dihydroxy-50-chlorophenyl)-3-arylpropane-1, 3-diones. *E J Chem* **2009**, *6*, 705–712
38. Sheikh, J.; Parvez, A.; Juneja, H.; Ingle, V.; Chohan, Z.; Youssoufi, M.; Hadda, T.B., Synthesis, biopharmaceutical characterization, antimicrobial and antioxidant activities of 1-(40-O-b-D-glucopyranosyloxy-20 -hydroxyphenyl)-3-aryl-propane-1,3-diones. *Eur J Med Chem* **2011**, *46*, 1390–1399
39. Diana, G.D.; Carabateas, P.M.; Johnson, R.E.; Williams, G.L.; Pancic, F.; Collins, J.C., Antiviral activity of some b-diketones. 4. Benzyldiketones. In vitro activity against both RNA and DNA viruses. *J Med Chem* **1978**, *21*, 889–894
40. Crouse, G.D.; McGowan, M.J.; Boisvenue, R.J., Polyflouro 1,3- diketones as systemic insecticides. *J Med Chem* **1989**, *32*, 2148–2151
41. Nishiyama, T.; Shiotsu, S.; Tsujita, H., Antioxidative activity and active site of 1,3-indanediones with the b-diketone moiety. *Polym Degrad Stab* **2002**, *76*, 435–439
42. Pivarcsik, T.; Toth, G.; Szemeredi, N.; Bogdanov, A.; Spengler G.; Kljun, J.; Kladnik, J.; Turel, I.; Enyedy, E. A., Comparison of Solution Chemical Properties and Biological Activity of Ruthenium Complexes of Selected β-Diketone, 8-Hydroxyquinoline and Pyrithione Ligands. *Pharmaceuticals* **2021**, *14*, 518.
43. Scapinello,L.;Vesco,G.; Nardo, L.; Maspero, A.; Vavassori, F.; Galli, S.; Penoni, A., Synthesis, Characterization and DNA-Binding Affinity of a New Zinc(II) Bis(5-methoxy-indol-3-yl) propane-1,3-dione Complex. *Pharmaceuticals* **2021**,*14*,760.
44. de Gonzalo,G.;Alcántara, A.R., Recent Developments in the Synthesis of β-Diketones. *Pharmaceuticals* **2021**, *14*, 1043.
45. Hansen,P.E., Structural Studies of β-Diketones and Their Implications on Biological Effects. *Pharmaceuticals* **2021**, *14*, 1189.
46. Vesco, G.; Brambati, M.; Scapinello, L.; Penoni, A.; Mella, M.; Masson, M.; Gaware, V.; Maspero, A.; Nardo, L., Asymmetric Phenyl Substitution: An Effective Strategy to Enhance the Photosensitizing Potential of Curcuminoids. *Pharmaceuticals* **2022**, *15*, 843.
47. Kel'in, A. V., Recent Advances in the Synthesis of 1,3-Diketones. *Curr Org Chem* **2003,** *7*, 1691-1711.
48. Lim, D.; Fang, F.; Zhou, G.; Coltart, D. M., Direct carbon-carbon bond formation via soft enolization: a facile and efficient synthesis of 1,3-diketones. *Org Lett* **2007,** *9* (21), 4139-42.
49. Devkota, A. K.; Veloria, J. R.; Guo, H. F.; Kurie, J. M.; Cho, E. J.; Dalby, K. N., Development of a High-Throughput Lysyl Hydroxylase (LH) Assay and Identification of Small-Molecule Inhibitors against LH2. *SLAS Discov* **2019,** *24* (4), 484-491.

50. Savastano, M.; García-Gallarín, C.; Giorgi, C.; Gratteri, P.; López de la Torre, M.D.; Bazzicalupi, C.; Bianchi, A.; Melguizo, M. Solid State and Solution Study on the Formation of Inorganic Anion Complexes with a Series of Tetrazine-Based Ligands. *Molecules* **2019**, *24*, 2247
51. Parkin, A.; Oswald, I. D. H.; Parsons, S. Structures of piperazine, piperidine and morpholine. *Acta Cryst.* **2004**. *B60*, 219-227
52. Lan, N.; Lu, Y.; Zhang, Y.; Pu, S.; Xi, H.; Nie, X.; Liu, J.; Yuan, W., FTO - A Common Genetic Basis for Obesity and Cancer. *Front Genet* **2020,** *11*, 559138.
53. Couture, J. F.; Collazo, E.; Ortiz-Tello, P. A.; Brunzelle, J. S.; Trievel, R. C., Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nat Struct Mol Biol* **2007,** *14* (8), 689-95.
54. Fong, G. H.; Takeda, K., Role and regulation of prolyl hydroxylase domain proteins. *Cell Death Differ* **2008,** *15* (4), 635-41.
55. Luther, K. B.; Hulsmeier, A. J.; Schegg, B.; Deuber, S. A.; Raoult, D.; Hennet, T., Mimivirus collagen is modified by bifunctional lysyl hydroxylase and glycosyltransferase enzyme. *J Biol Chem* **2011,** *286* (51), 43701-43709.
56. Liu, W.; Zhang, T.; Guo, L.; Wang, Y.; Yang, Y., Lysyl hydroxylases are transcription targets for GATA3 driving lung cancer cell metastasis. *Sci Rep* **2018,** *8*, 11905.
57. Guo, H. F.; Cho, E. J.; Devkota, A. K.; Chen, Y.; Russell, W.; Phillips, G. N., Jr.; Yamauchi, M.; Dalby, K. N.; Kurie, J. M., A scalable lysyl hydroxylase 2 expression system and luciferase-based enzymatic activity assay. *Arch Biochem Biophys* **2017,** *618*, 45-51.
58. Guo, H. F.; Tsai, C. L.; Terajima, M.; Tan, X.; Banerjee, P.; Miller, M. D.; Liu, X.; Yu, J.; Byemerwa, J.; Alvarado, S.; Kaoud, T. S.; Dalby, K. N.; Bota-Rabassedas, N.; Chen, Y.; Yamauchi, M.; Tainer, J. A.; Phillips, G. N., Jr.; Kurie, J. M., Pro-metastatic collagen lysyl hydroxylase dimer assemblies stabilized by Fe(2+)-binding. *Nat Commun* **2018,** *9* (1), 512.
59. Maghsoud, Y.; Vázquez-Montelongo, E. A.; Yang, X.; Liu, C.; Jing, Z.; Lee, J.; Harger, M.; Smith, A. K.; Espinoza, M.; Guo, H.-F., Computational Investigation of a Series of Small Molecules as Lead Compounds for Lysyl hydroxylase-2 (LH2) Inhibition. *ChemRxiv*. Cambridge: Cambridge Open Engage; **2022**; This content is a preprint and has not been peer-reviewed.
60. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L., SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research* **2018,** *46* (W1), W296-W303.
61. Harger, M.; Li, D.; Wang, Z.; Dalby, K.; Lagardère, L.; Piquemal, J.-P.; Ponder, J.; Ren, P., Tinker-OpenMM: Absolute and relative alchemical free energies using AMOEBA on GPUs. *Journal of Computational Chemistry* **2017,** *38* (23), 2047-2055.
62. Likas, A.; Vlassis, N.; Verbeek, J. J., The global k-means clustering algorithm. *Pattern recognition* **2003,** *36* (2), 451-461.
63. Kratz, E. G.; Walker, A. R.; Lagardère, L.; Lipparini, F.; Piquemal, J. P.; Andrés Cisneros, G., LICHEM: A QM/MM program for simulations with multipolar and polarizable force fields. *Journal of computational chemistry* **2016,** *37* (11), 1019-1029.
64. Gökcan, H.; Vázquez-Montelongo, E. A.; Cisneros, G. A., LICHEM 1.1: recent improvements and new capabilities. *Journal of chemical theory and computation* **2019,** *15* (5), 3056-3065.
65. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.;

Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.
66. Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J.-P.; Ren, P.; Ponder, J. W., Tinker 8: software tools for molecular design. *Journal of chemical theory and computation* **2018,** *14* (10), 5273-5289.
67. Chai, J.-D.; Head-Gordon, M., Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics* **2008,** *10* (44), 6615-6620.
68. Chai, J.-D.; Head-Gordon, M., Systematic optimization of long-range corrected hybrid density functionals. *The Journal of chemical physics* **2008,** *128* (8), 084106.
69. Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J.-P.; Ponder, J. W.; Ren, P., AMOEBA polarizable atomic multipole force field for nucleic acids. *Journal of chemical theory and computation* **2018,** *14* (4), 2084-2108.
70. Lu, T.; Chen, F., Multiwfn: a multifunctional wavefunction analyzer. *Journal of computational chemistry* **2012,** *33* (5), 580-592.
71. Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W., Revealing noncovalent interactions. *Journal of the American Chemical Society* **2010,** *132* (18), 6498-6506.
72. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *Journal of molecular graphics* **1996,** *14* (1), 33-38.
73. Jurcik, A.; Bednar, D.; Byska, J.; Marques, S. M.; Furmanova, K.; Daniel, L.; Kokkonen, P.; Brezovsky, J.; Strnad, O.; Stourac, J., CAVER Analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. *Bioinformatics* **2018,** *34* (20), 3586-3588.
74. Petřek, M.; Otyepka, M.; Banáš, P.; Košinová, P.; Koča, J.; Damborský, J., CAVER: a new tool to explore routes from protein clefts, pockets, and cavities. *BMC bioinformatics* **2006,** *7* (1), 1-9.

Supporting information for

1,3-diketone analogs as selective lysyl hydroxylase 2 (LH2) antagonists

Juhoon Lee,[1] Hou-fu Guo,[*,2] Yazdan Maghsoud,[3] Erik Antonio Vázquez-Montelongo, [4] Zhifeng Jing,[5] Shike Wang,[6] Rachel M Sammons,[1] Eun Jeong Cho,[1] Pengyu Ren,[*,5] G. Andrés Cisneros[*,3,7] Jonathan M. Kurie,[*,6] and Kevin N. Dalby,[*,1]

[1]*Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy, University of Texas at Austin, Austin, Texas 78712, USA; Targeted Therapeutic Drug Discovery and Development Program, College of Pharmacy, University of Texas, Austin, TX 78712, USA*
*E-mail: dalby@austin.utexas.edu
[2]*Department of Molecular and Cellular Biochemistry, University of Kentucky College of Medicine, KY 40536, USA*
*E-mail: hguo3@g.uky.edu
[3]*Department of Chemistry and Biochemistry, The University of Texas at Dallas, Richardson, TX 75080, USA*
[4]*Department of Physical Medicine and Rehabilitation, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA*
[5]*Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX 78712, USA*
*E-mail: pren@utexas.edu
[6]*Departmnt of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA*
*E-mail: jkurie@mdanderson.org
[7]*Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA*
*E-mail: andres@utdallas.edu

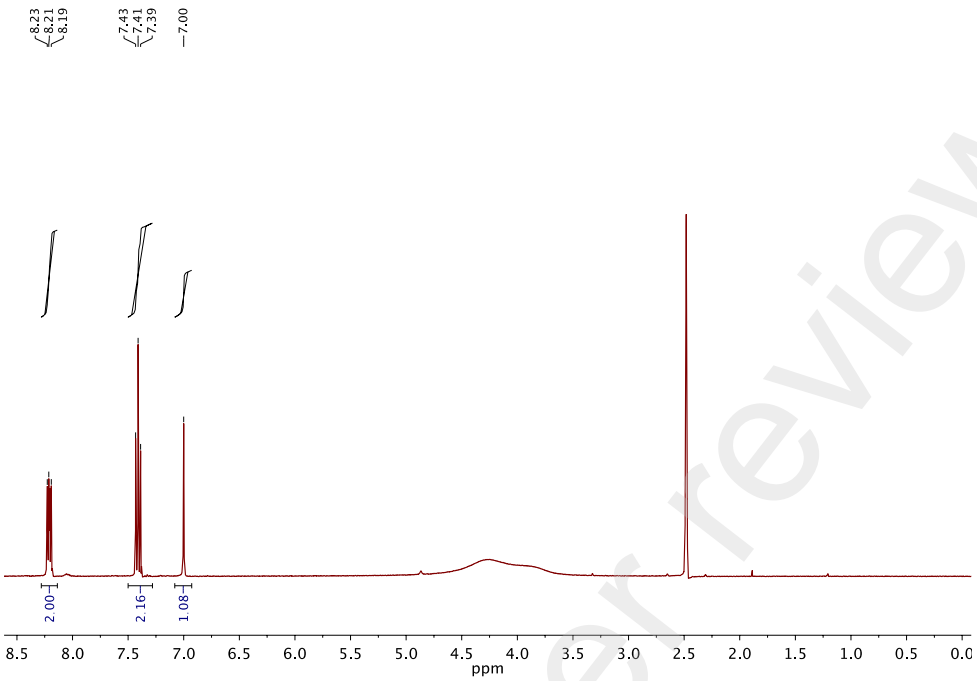


Figure S1. ^{1}H-NMR spectrum of **3** recorded in DMSO-d_6.

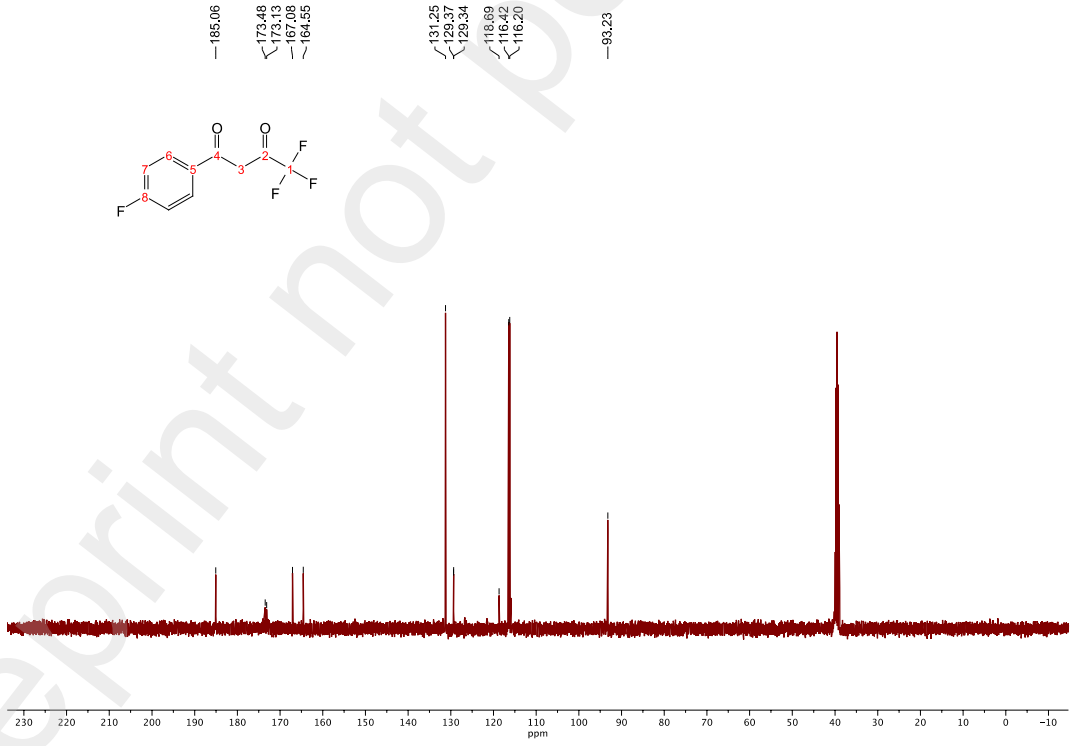


Figure S2. ^{13}C-NMR spectrum of **3** recorded in DMSO-d_6.

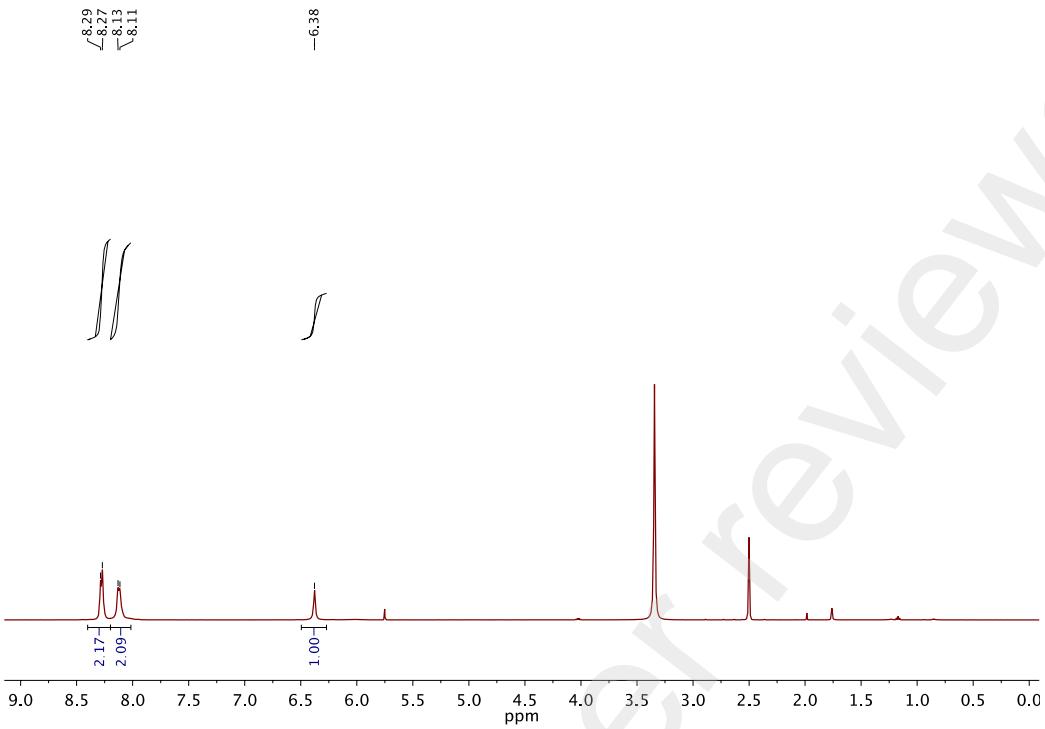


Figure S3. ^{1}H-NMR spectrum of **4** recorded in DMSO-d_6.

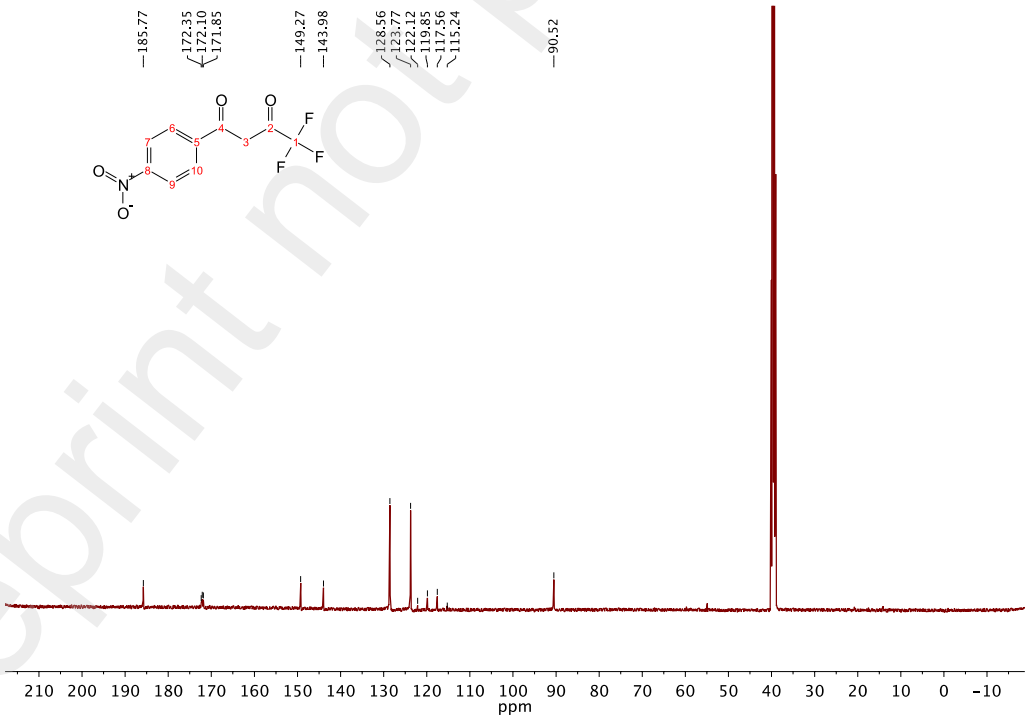


Figure S4. ^{13}C-NMR spectrum of **4** recorded in DMSO-d_6.

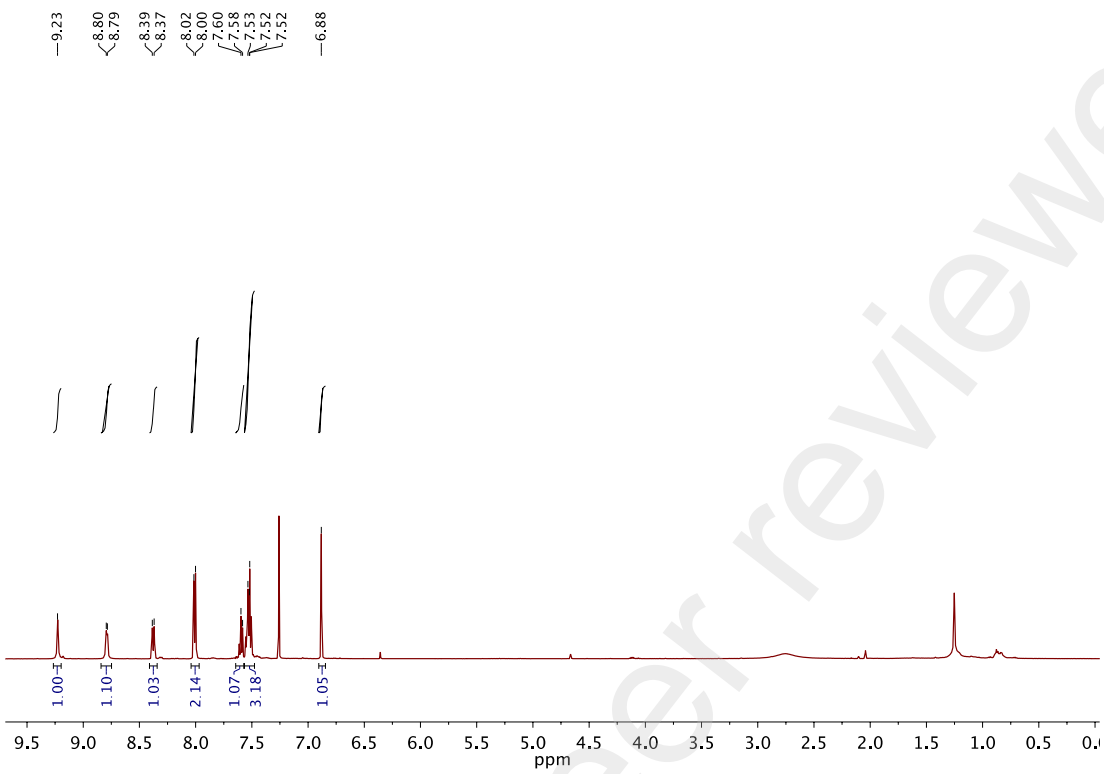


Figure S5. ^{1}H-NMR spectrum of **5** recorded in $CDCl_3$.

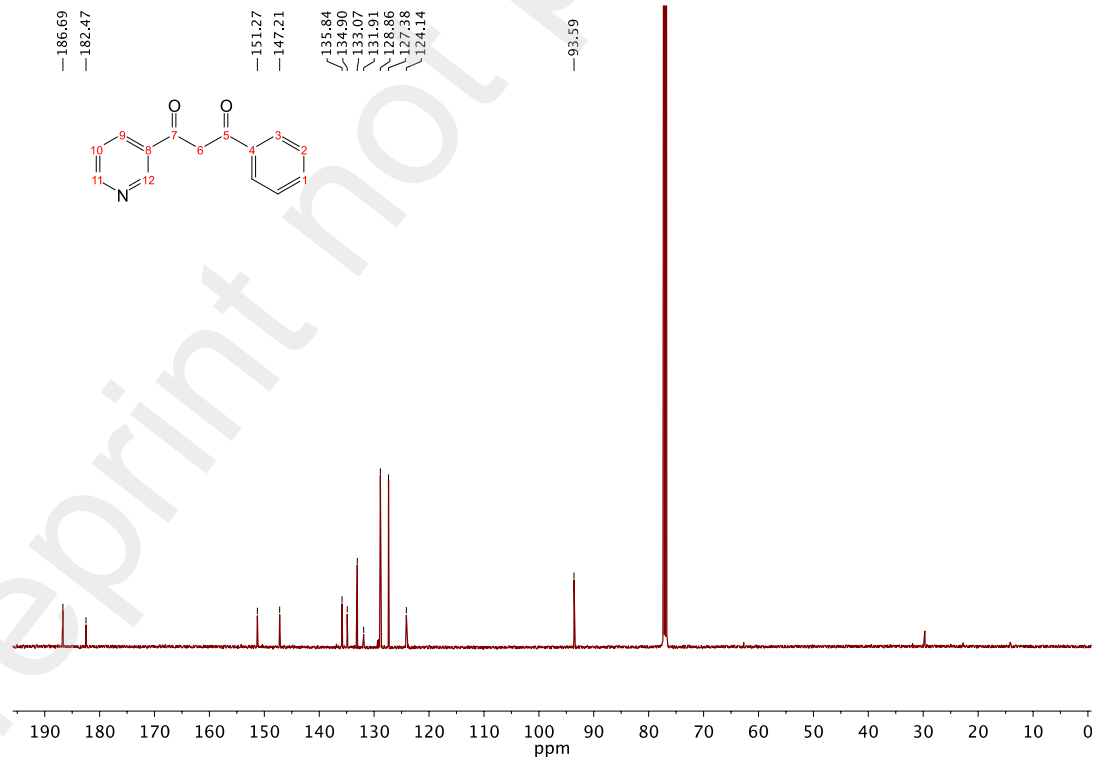


Figure S6. ^{13}C-NMR spectrum of **5** recorded in $CDCl_3$.

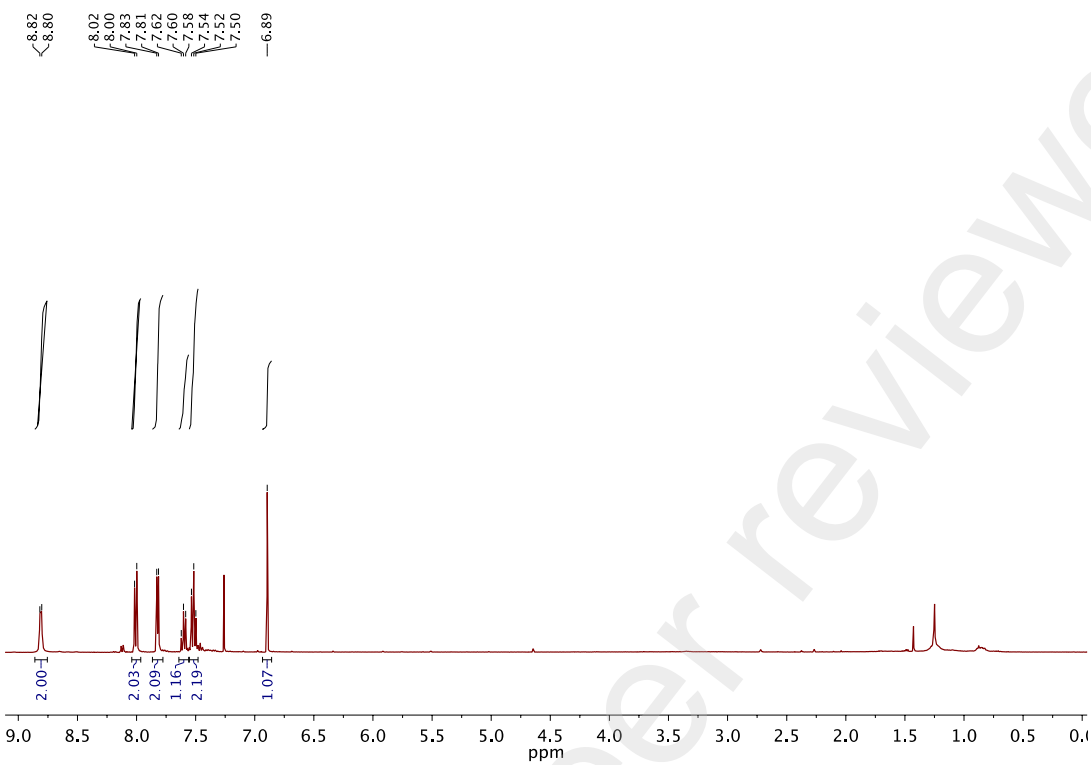


Figure S7. ^{1}H-NMR spectrum of **6** recorded in $CDCl_3$.

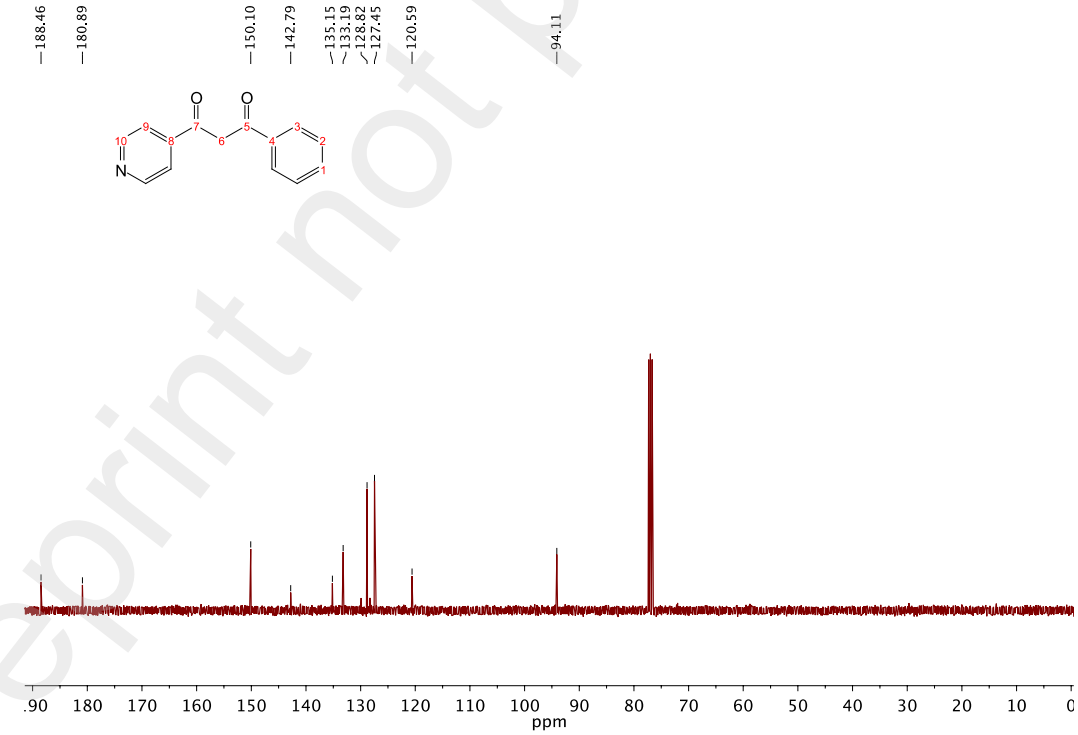


Figure S8. ^{13}C-NMR spectrum of **6** recorded in $CDCl_3$.

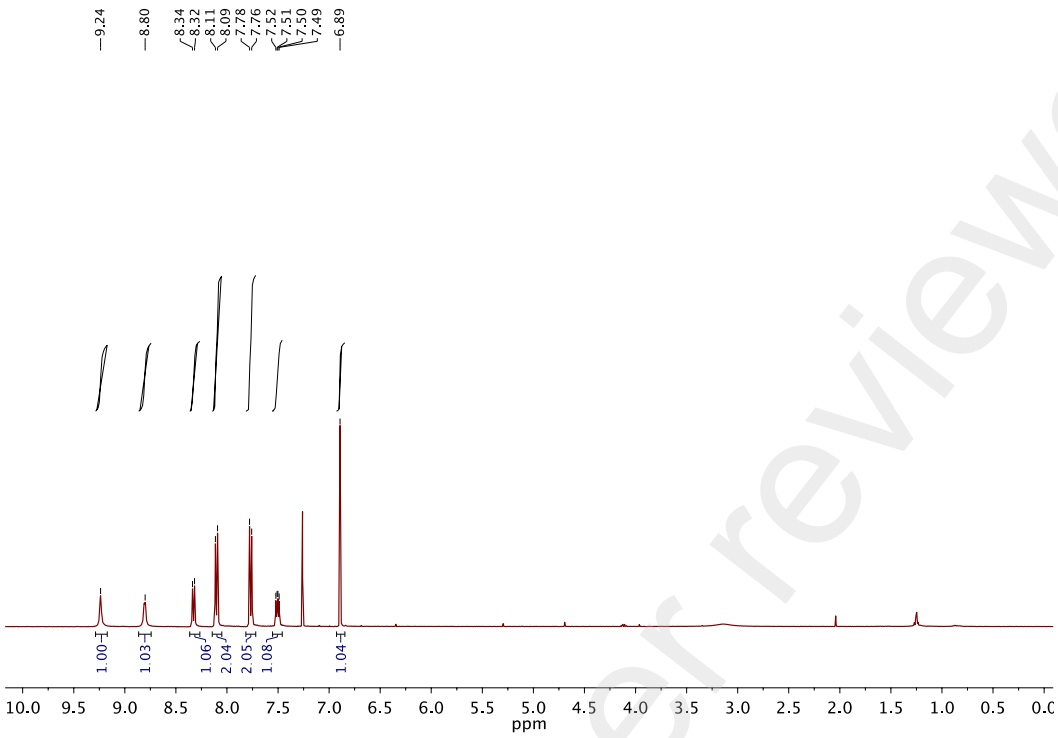


Figure S9. ^{1}H-NMR spectrum of **7** recorded in $CDCl_3$.

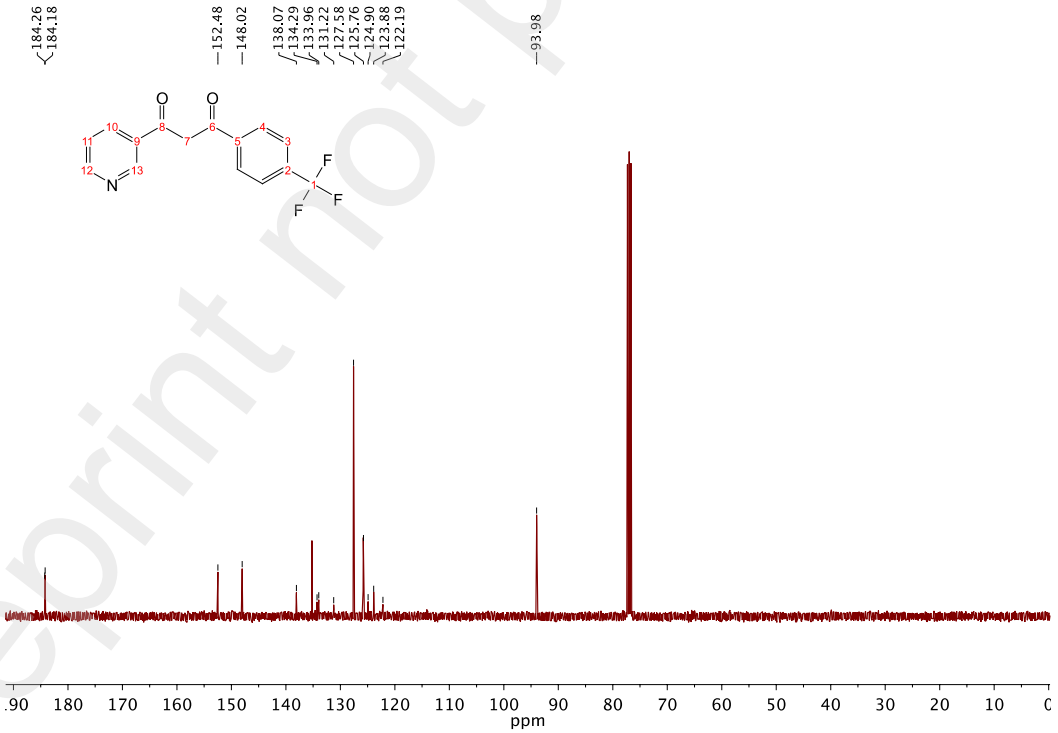


Figure S10. ^{13}C-NMR spectrum of **7** recorded in $CDCl_3$.

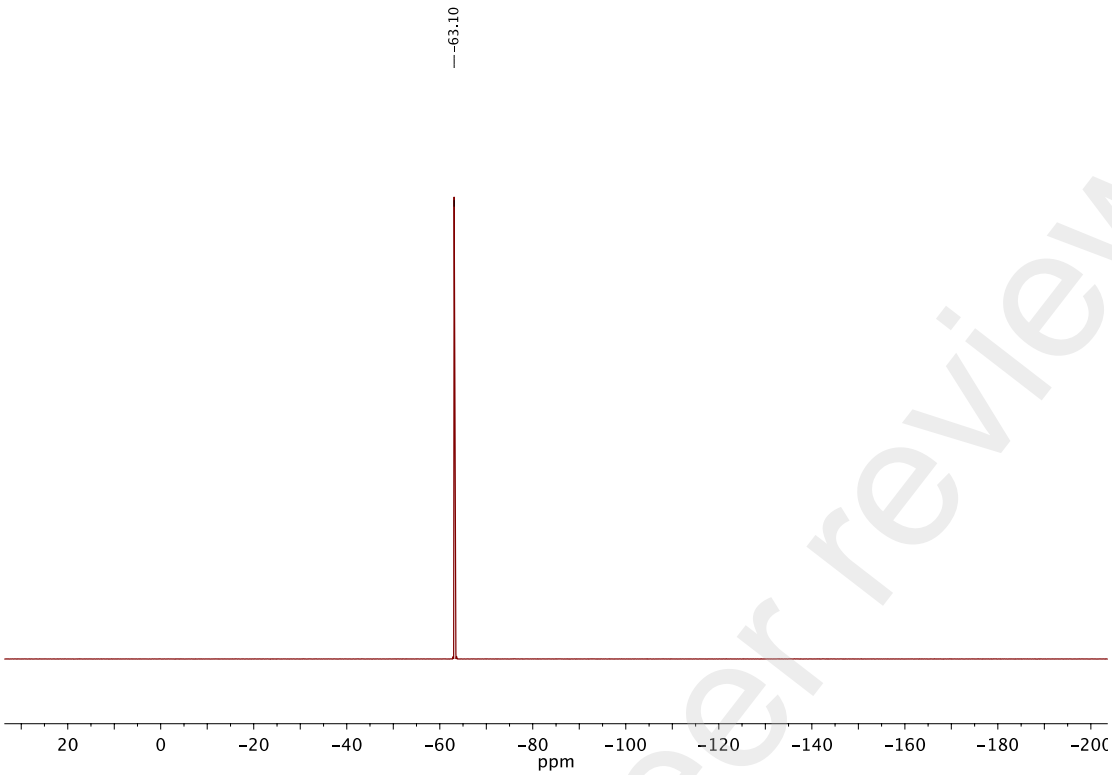


Figure S11. ^{19}F-NMR spectrum of **7** recorded in $CDCl_3$.

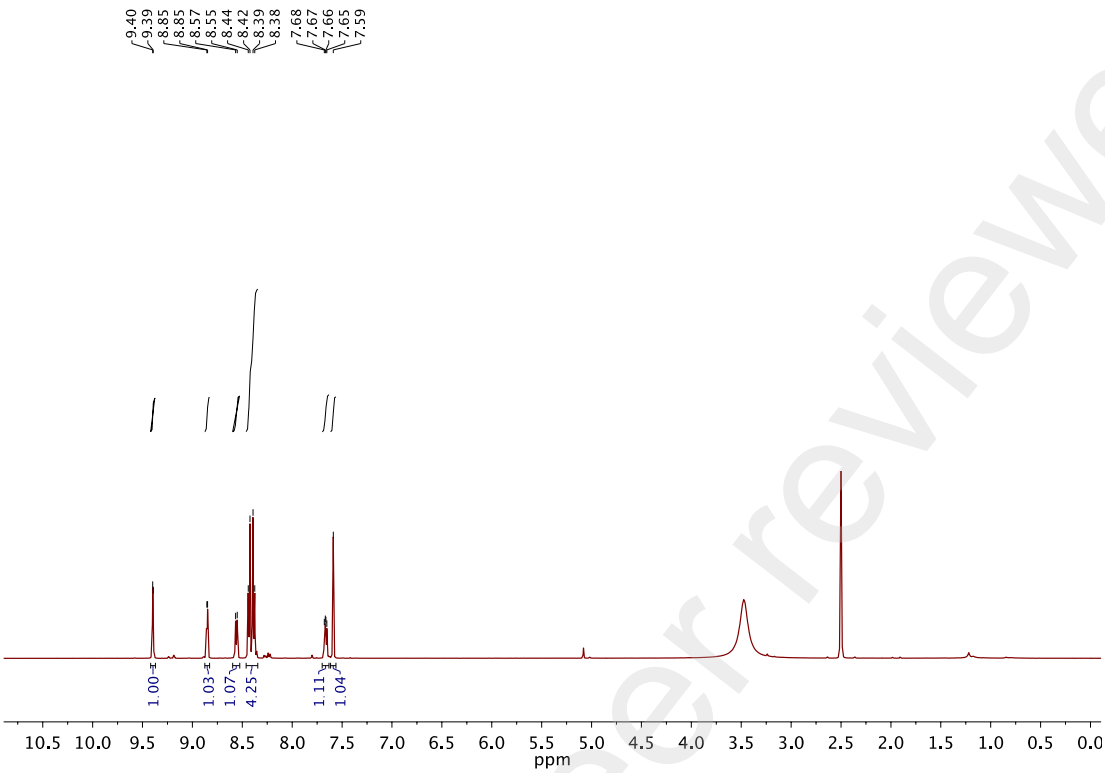


Figure S12. [1]H-NMR spectrum of **8** recorded in DMSO-d_6.

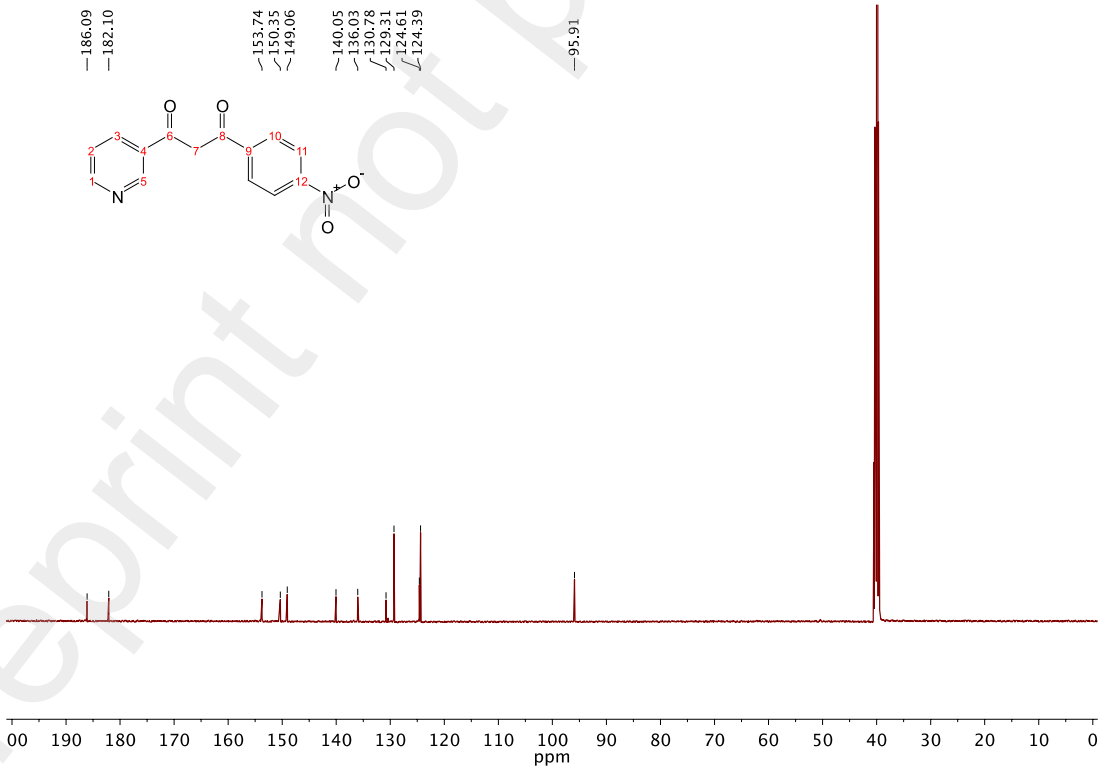


Figure S13. [13]C-NMR spectrum of **8** recorded in DMSO-d_6.

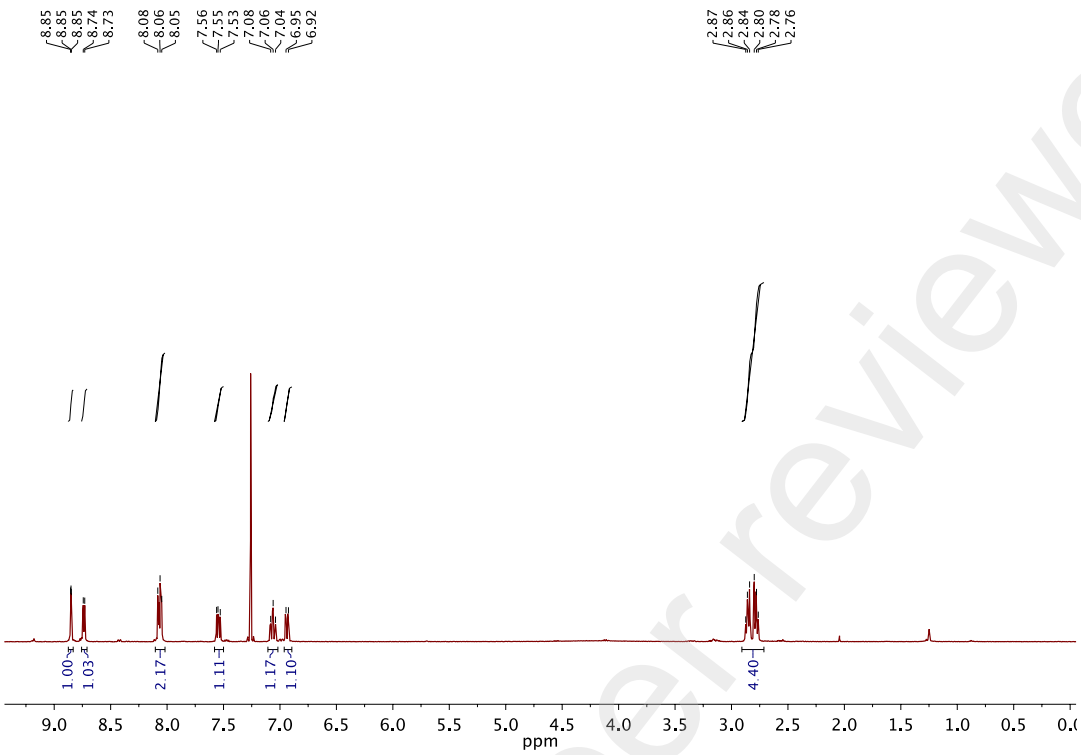


Figure S14. ^{1}H-NMR spectrum of **9** recorded in $CDCl_3$.

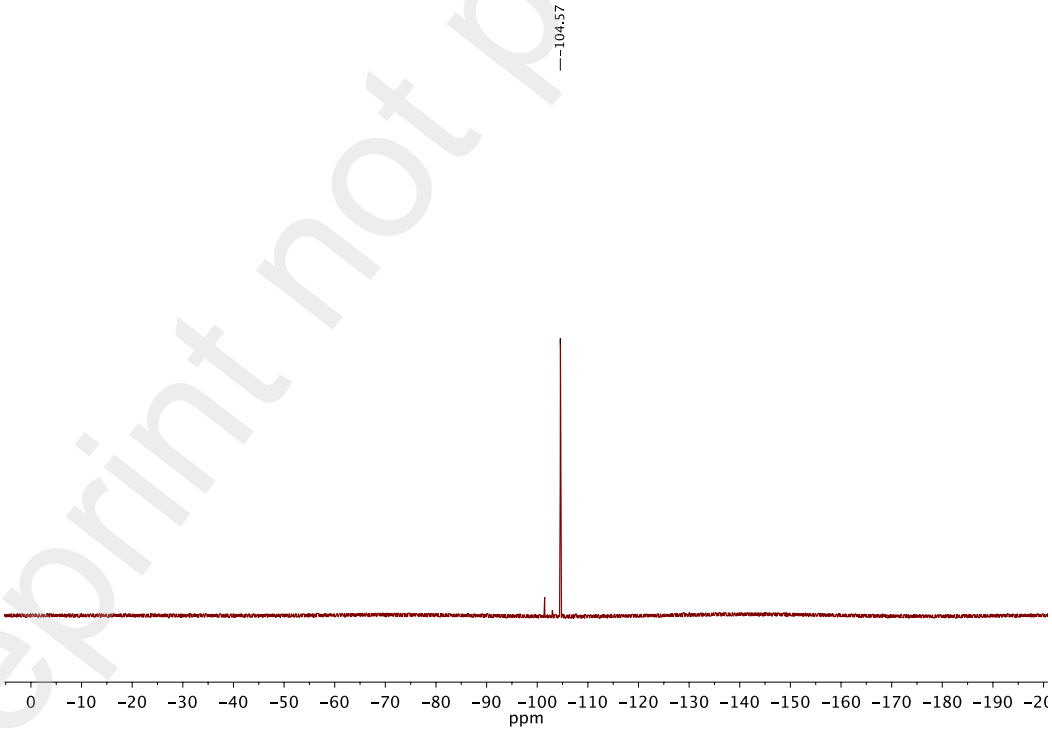


Figure S15. ^{19}F-NMR spectrum of **9** recorded in $CDCl_3$.

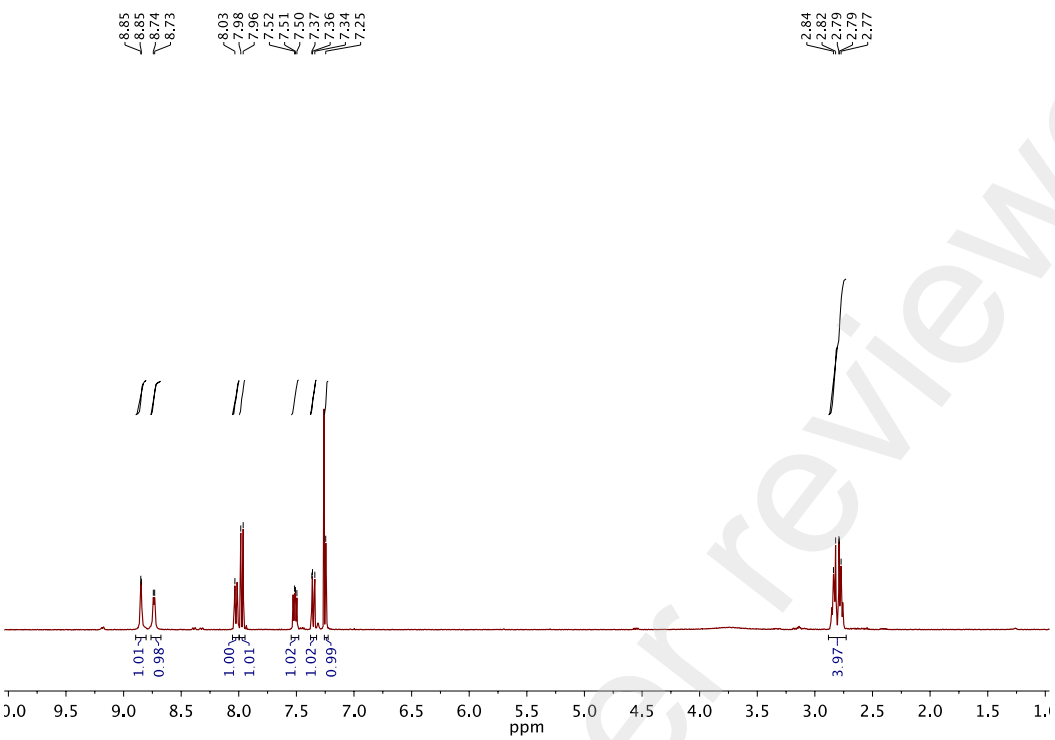


Figure S16. ^{1}H-NMR spectrum of **10** recorded in $CDCl_3$.

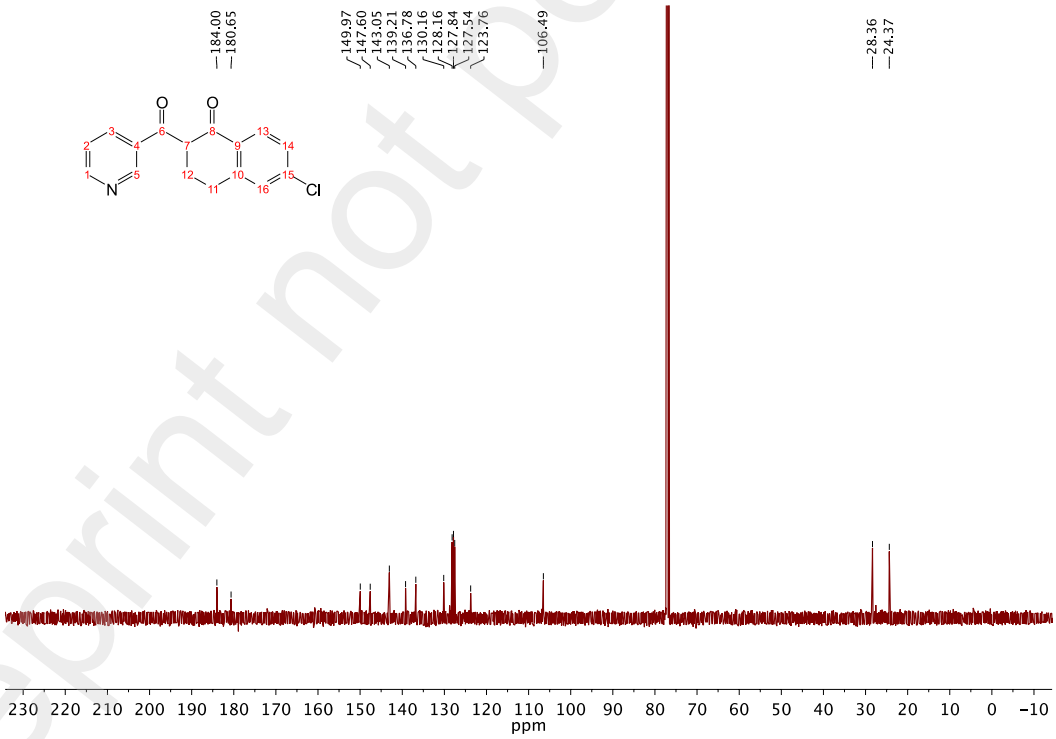


Figure S17. ^{13}C-NMR spectrum of **10** recorded in $CDCl_3$.

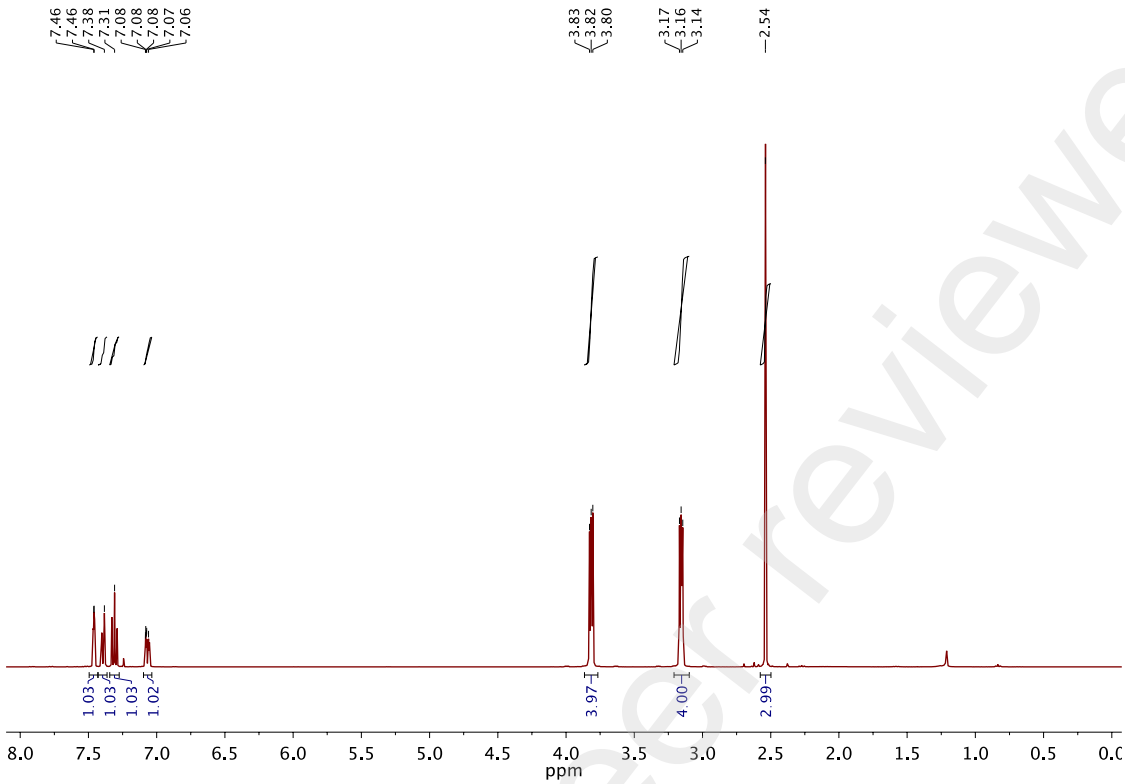


Figure S18. ^{1}H-NMR spectrum of **1i** recorded in $CDCl_3$.

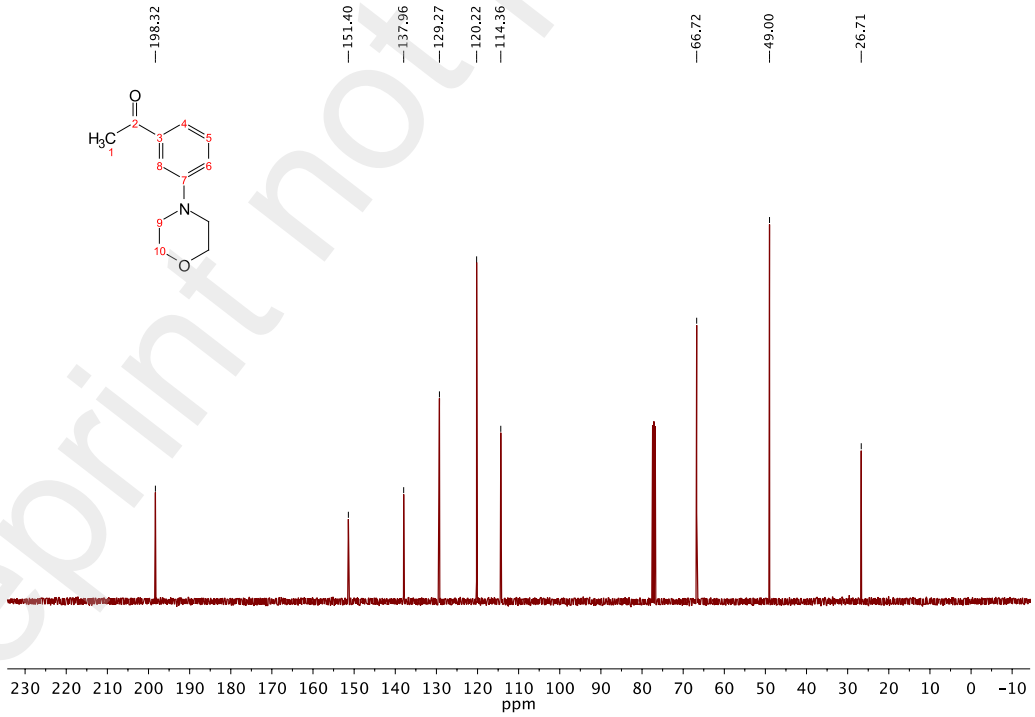


Figure S19. ^{13}C-NMR spectrum of **1i** recorded in $CDCl_3$.

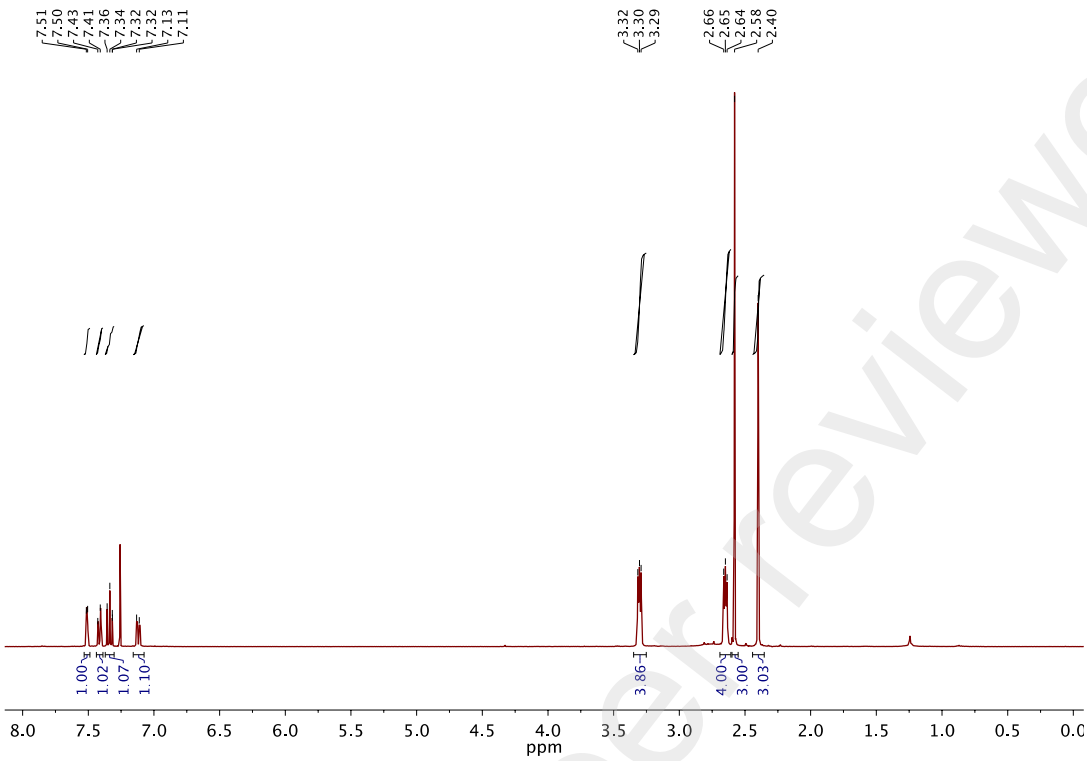


Figure S20. ^{1}H-NMR spectrum of **1j** recorded in $CDCl_3$.

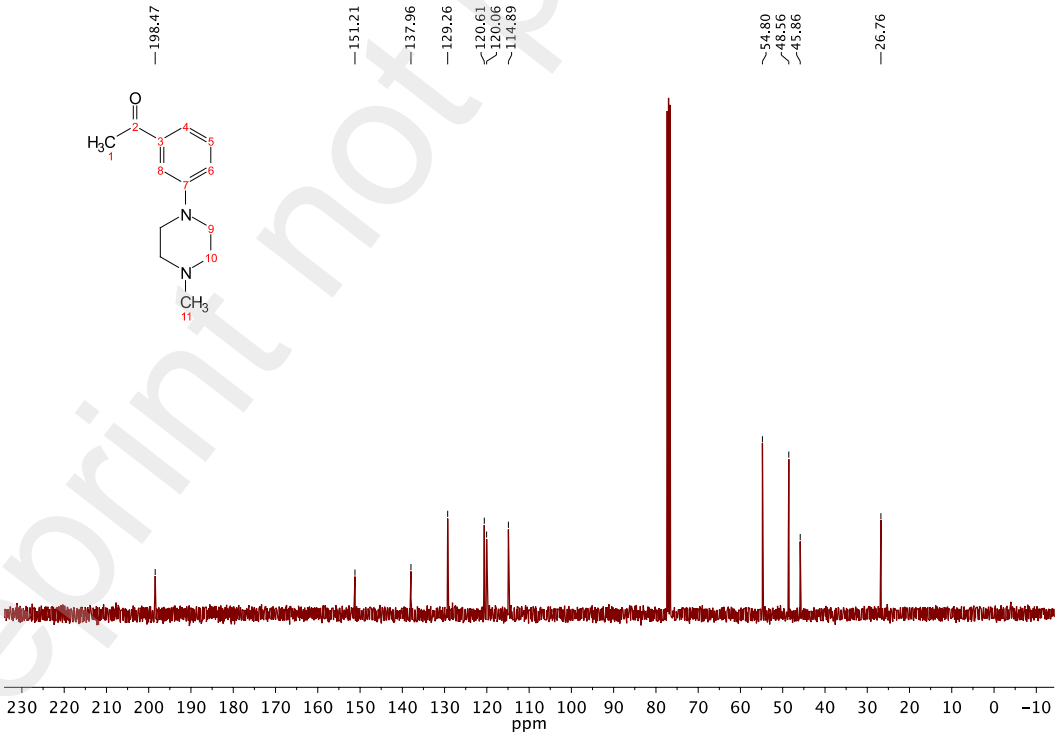


Figure S21. ^{13}C-NMR spectrum of **1j** recorded in $CDCl_3$.

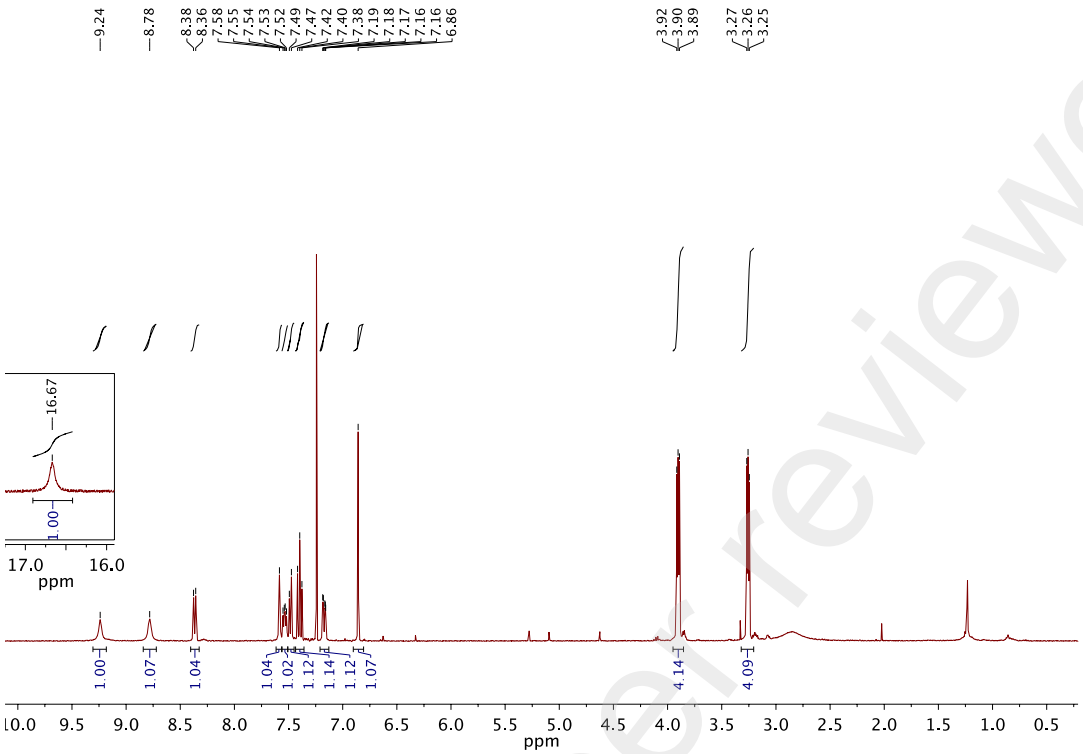


Figure S22. ^{1}H-NMR spectrum of **11** recorded in $CDCl_3$.

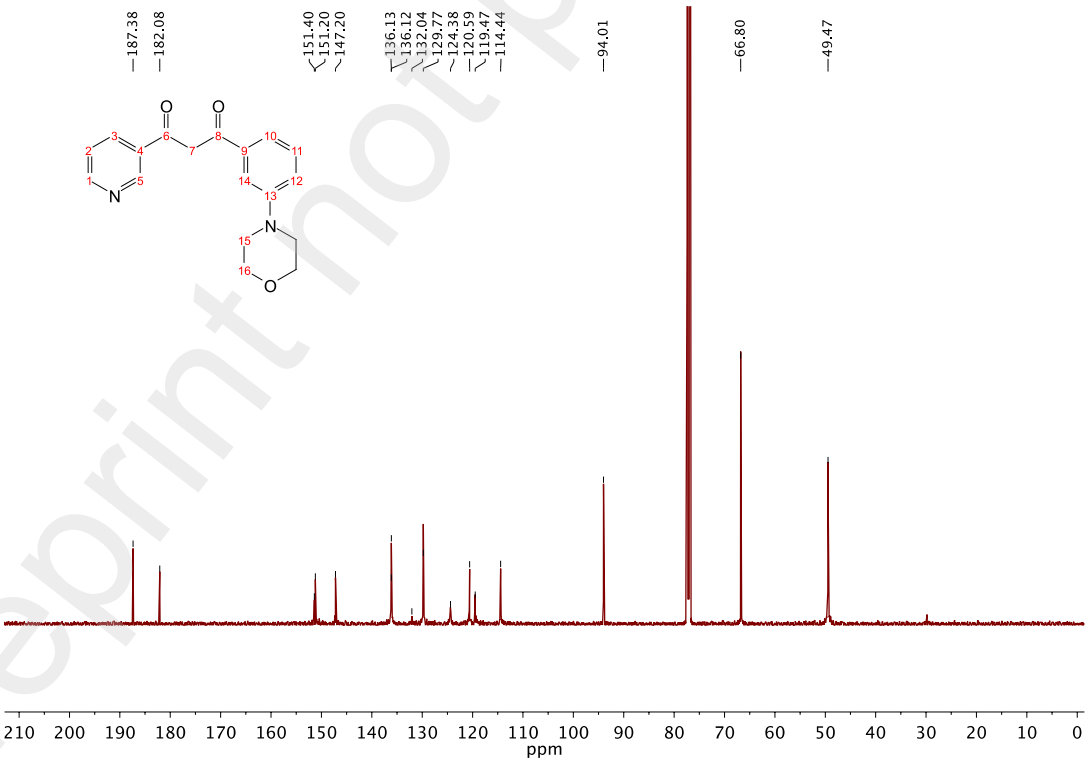


Figure S23. ^{13}C-NMR spectrum of **11** recorded in $CDCl_3$.

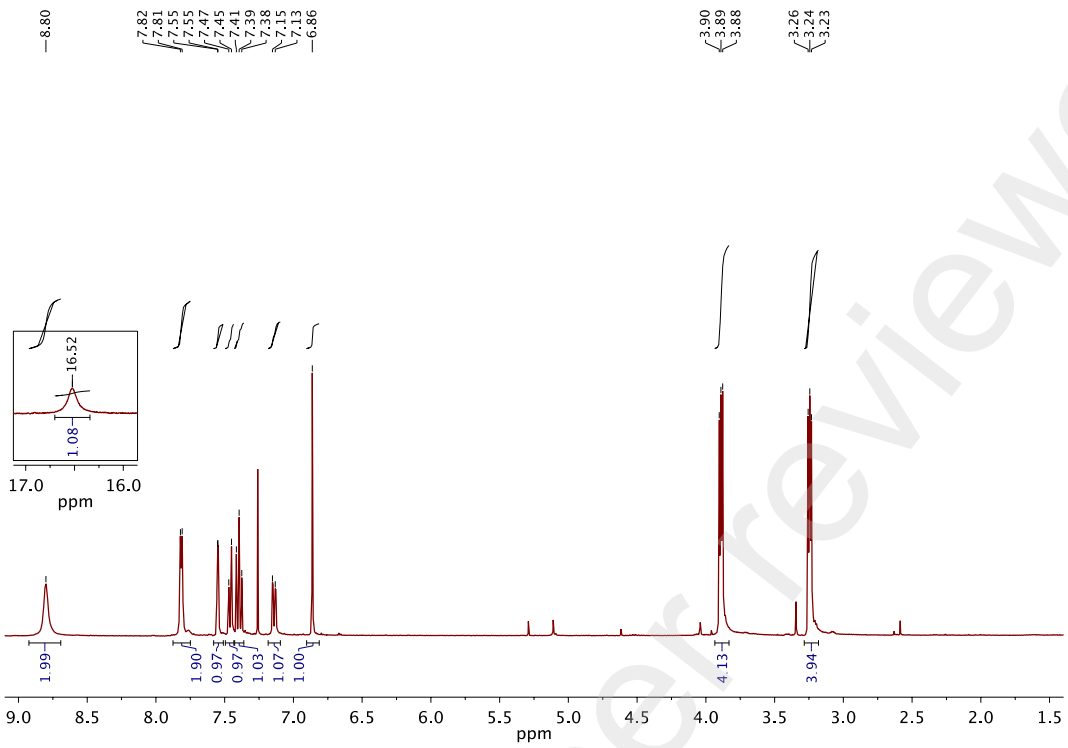


Figure S24. ^{1}H-NMR spectrum of **12** recorded in $CDCl_3$.

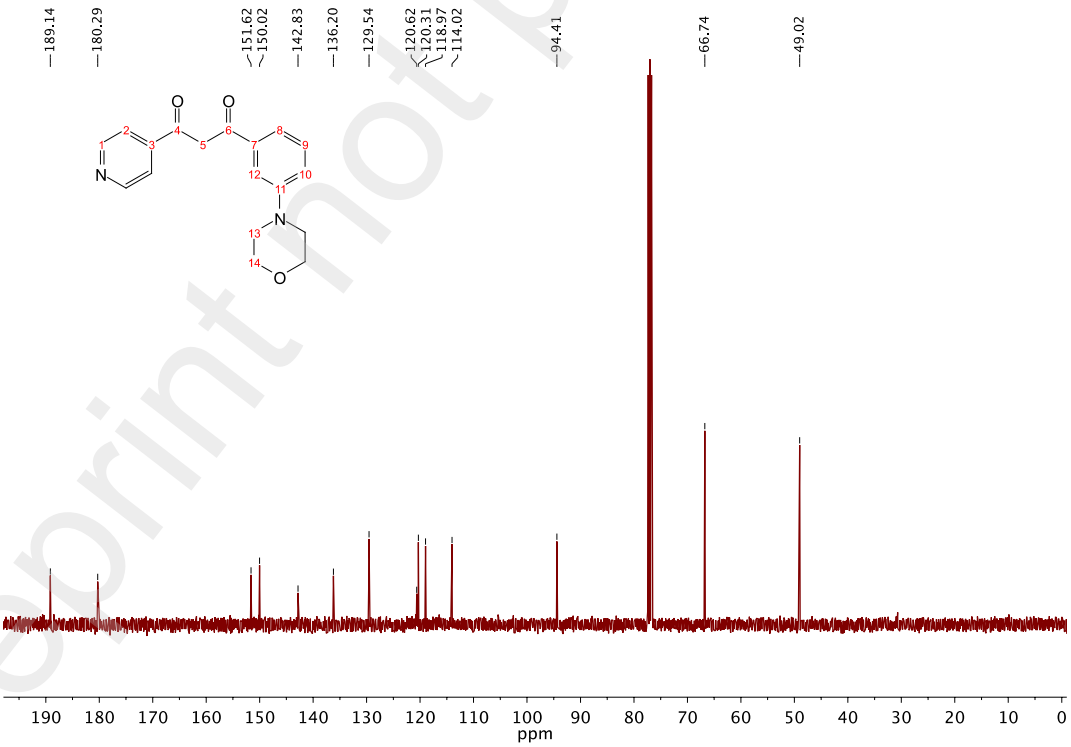


Figure S25. ^{13}C-NMR spectrum of **12** recorded in $CDCl_3$.

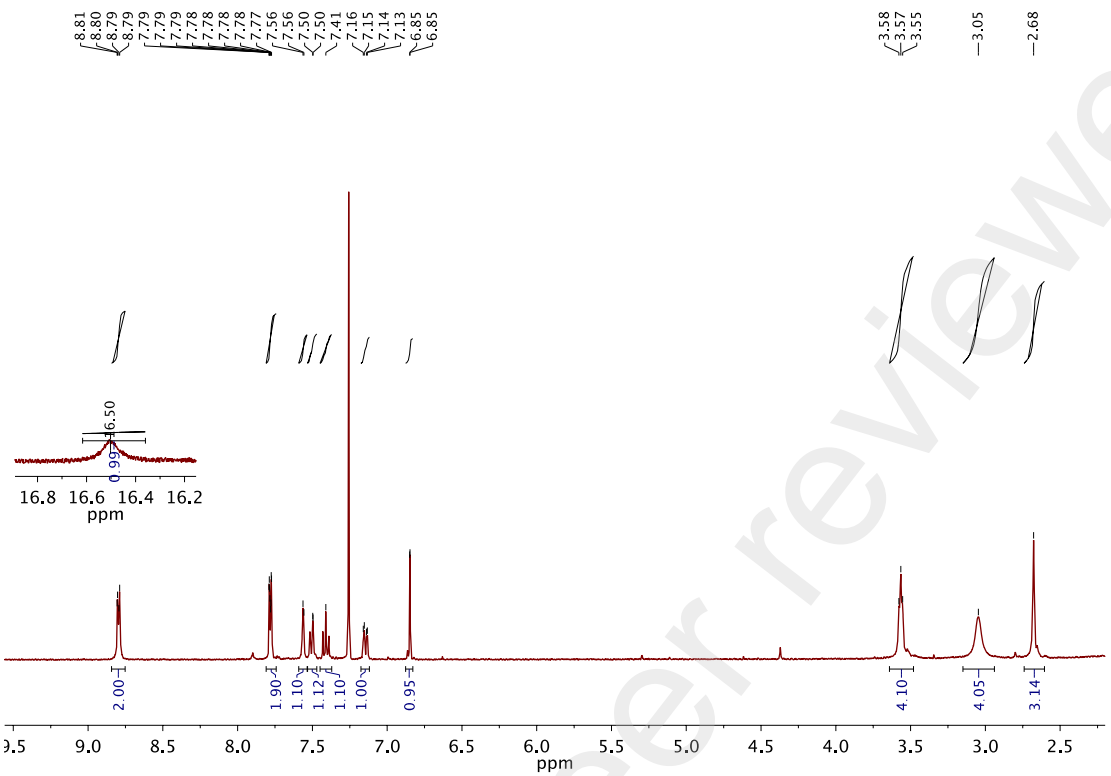


Figure S26. ^{1}H-NMR spectrum of **13** recorded in $CDCl_3$.

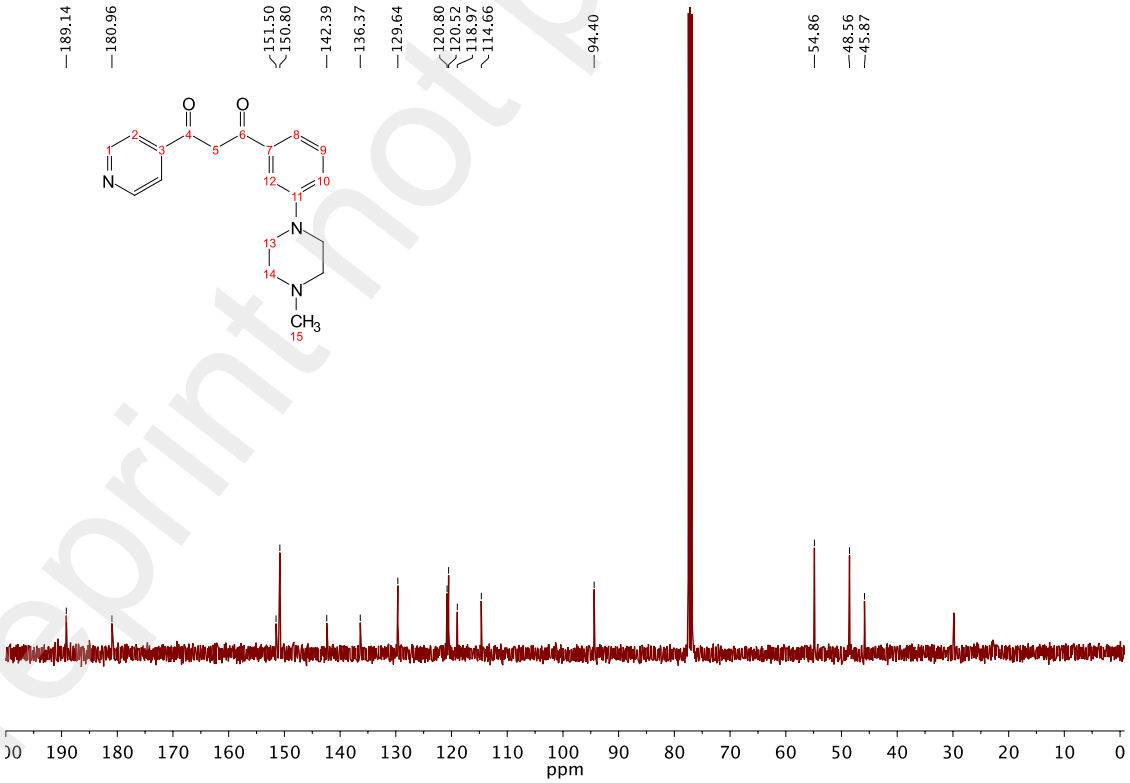


Figure S27. ^{13}C-NMR spectrum of **13** recorded in $CDCl_3$.

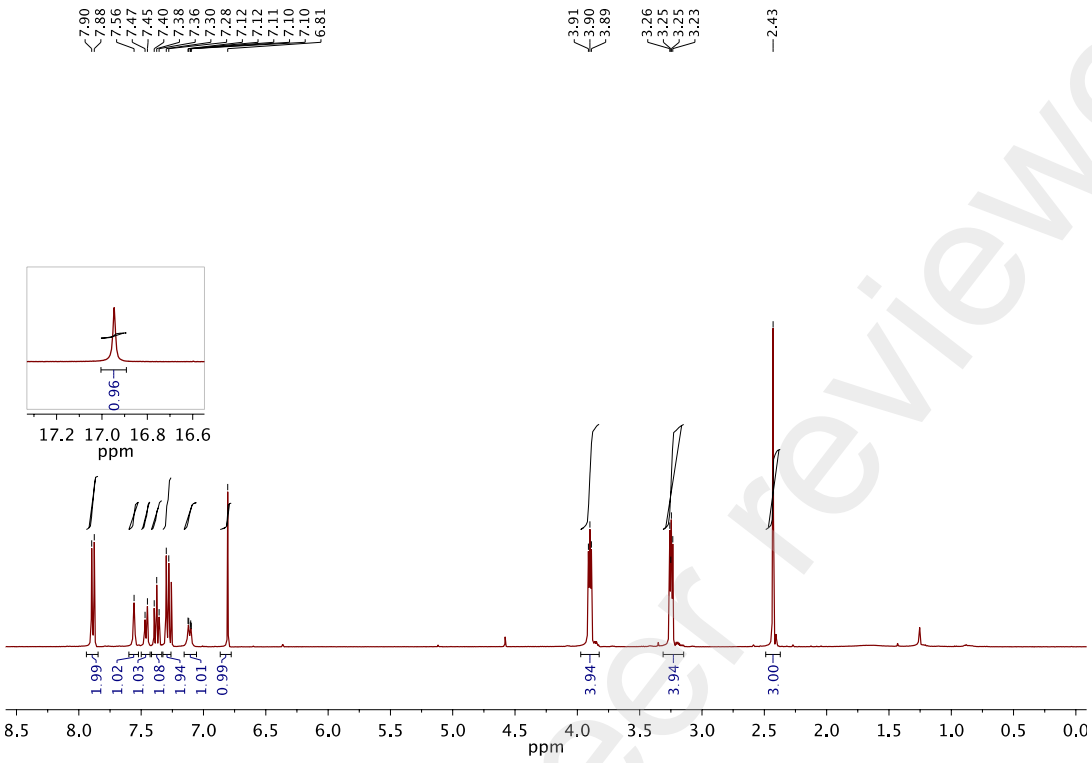


Figure S28. ^{1}H-NMR spectrum of **18** recorded in $CDCl_3$.

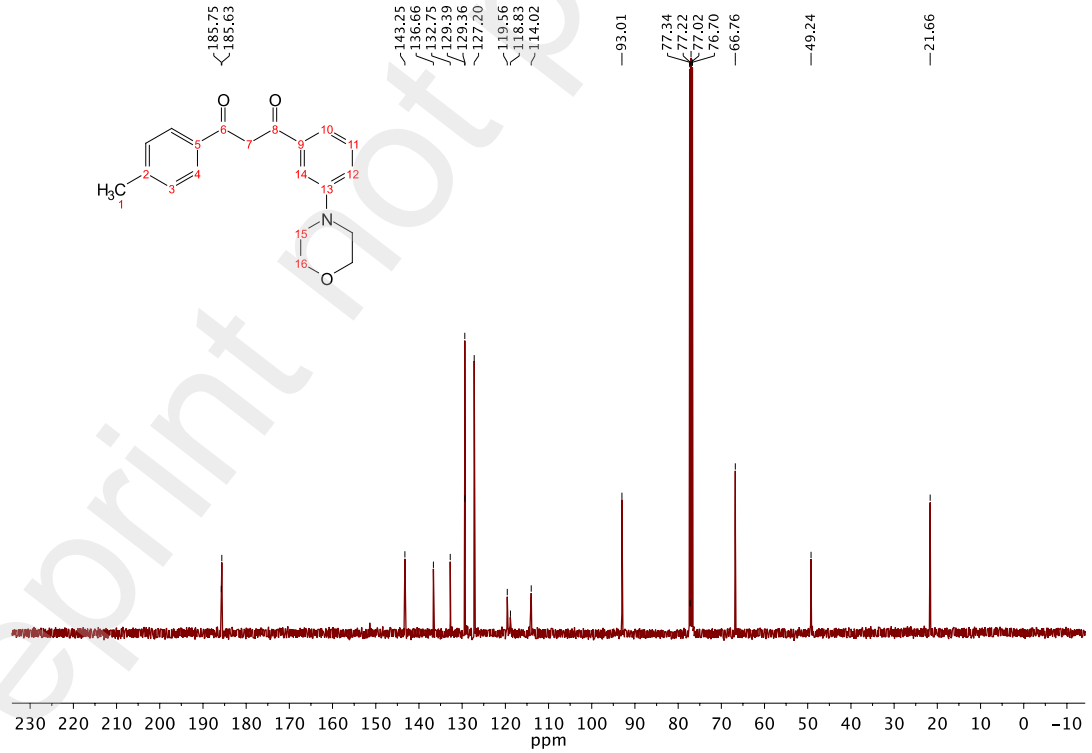


Figure S29. ^{13}C-NMR spectrum of **18** recorded in $CDCl_3$.

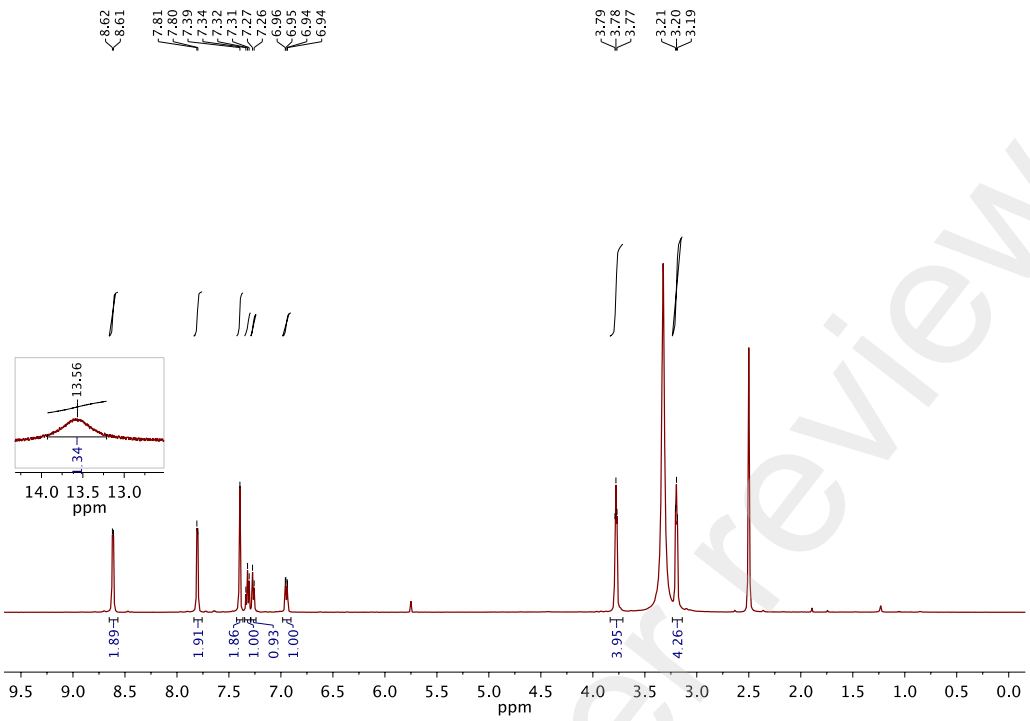


Figure S30. [1]H-NMR spectrum of **17** recorded in DMSO-d_6.

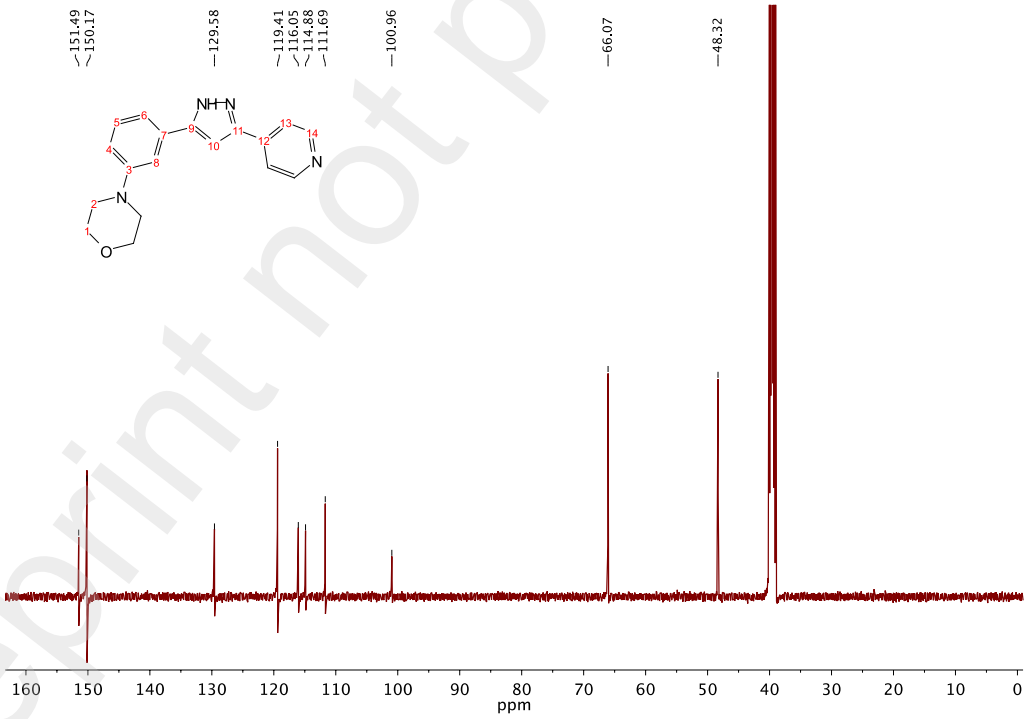


Figure S31. [13]C-NMR spectrum of **17** recorded in DMSO-d_6.

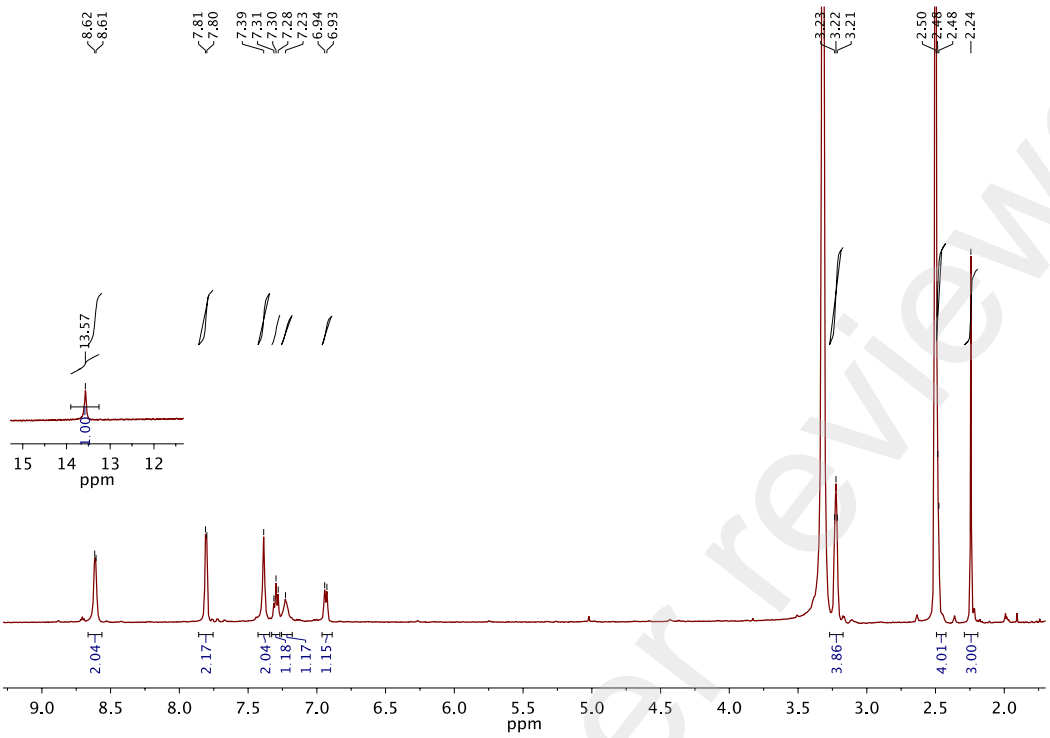


Figure S32. [1]H-NMR spectrum of **15** recorded in DMSO-d_6.

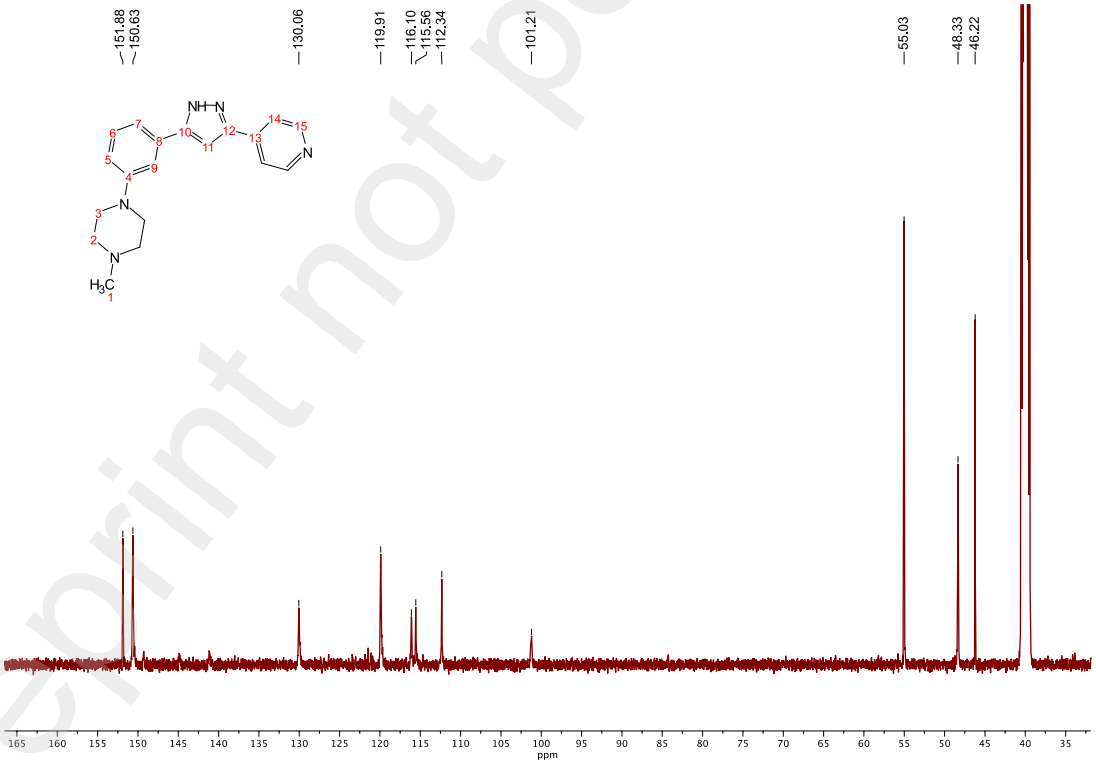


Figure S33. [13]C-NMR spectrum of **15** recorded in DMSO-d_6.

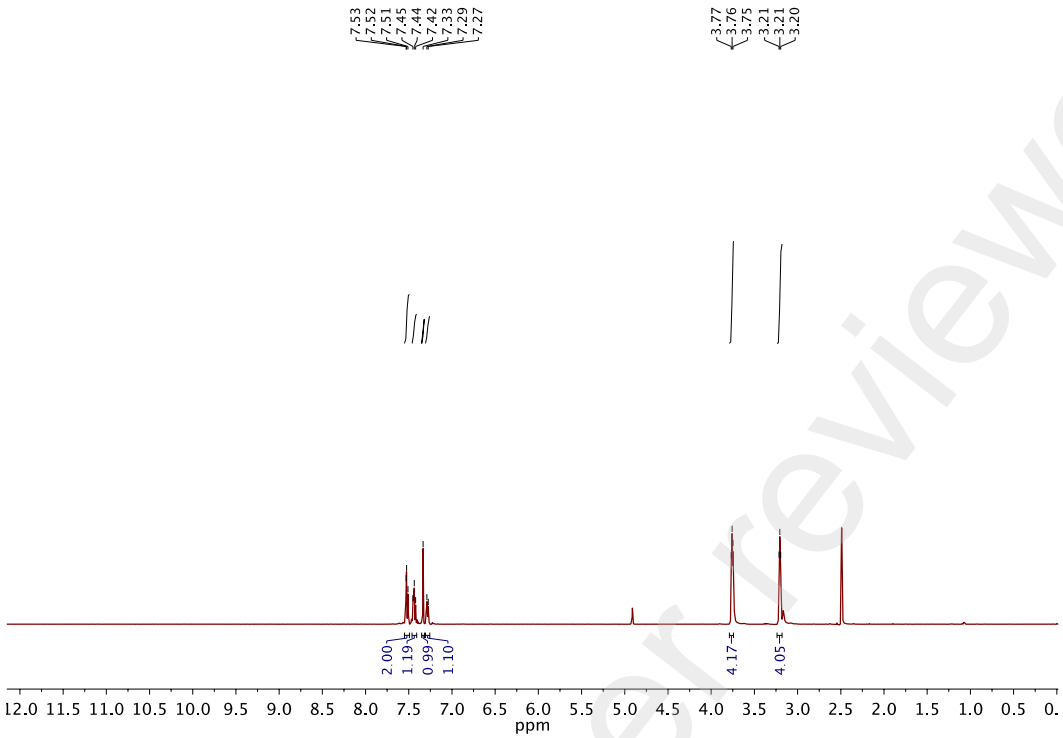


Figure S34. [1]H-NMR spectrum of **16** recorded in DMSO-d_6.

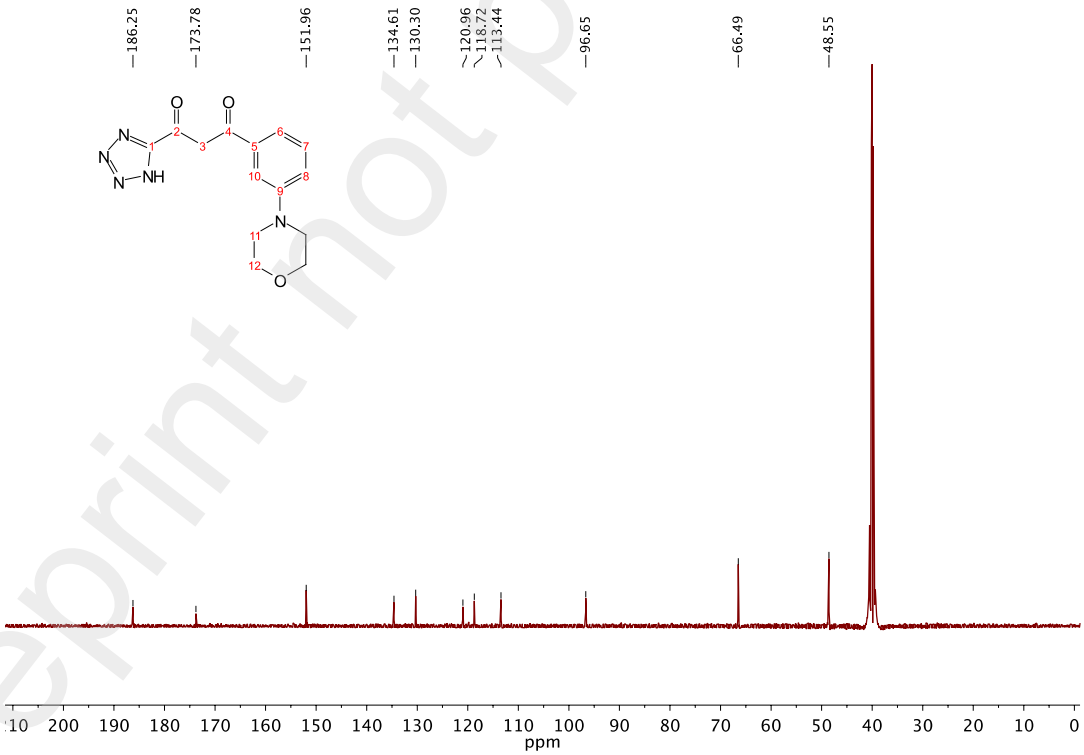


Figure S35. [13]C-NMR spectrum of **16** recorded in DMSO-d_6.

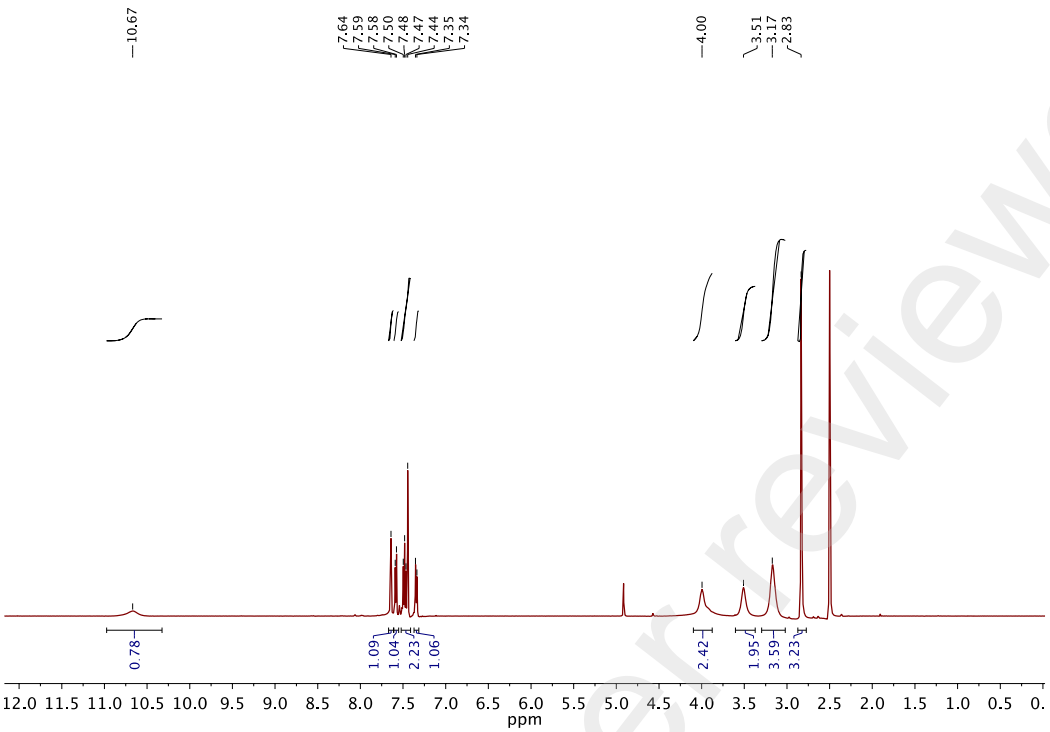


Figure S36. [1]H-NMR spectrum of **14** recorded in DMSO-d_6.

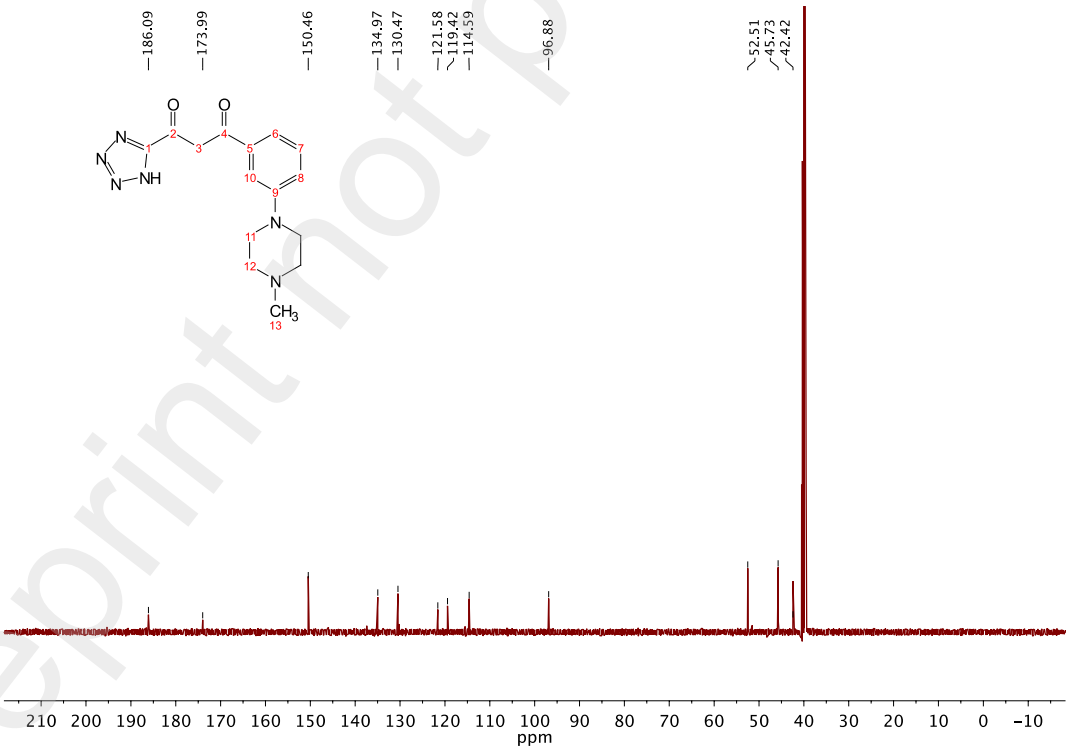


Figure S37. [13]C-NMR spectrum of **14** recorded in DMSO-d_6.

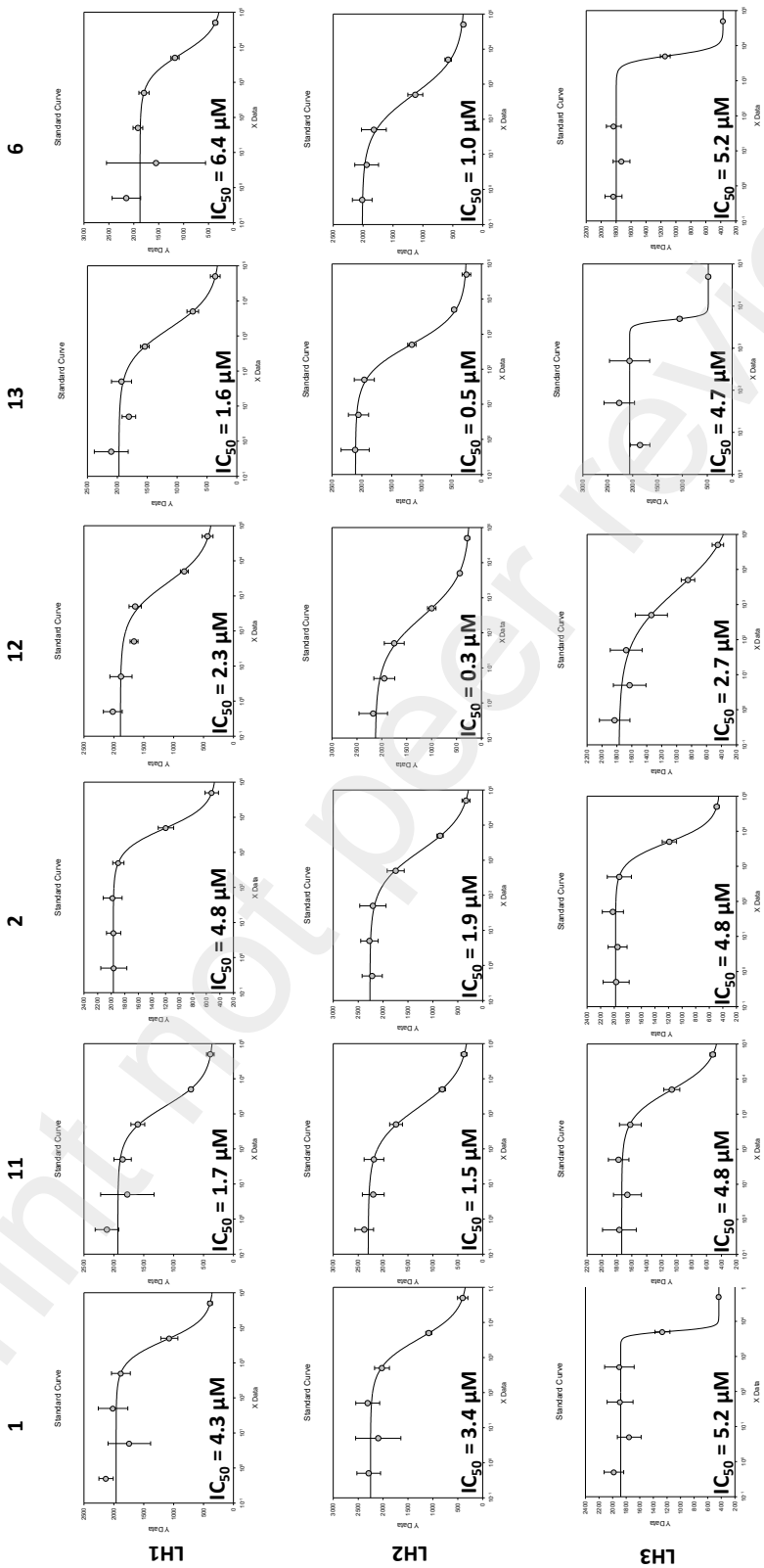
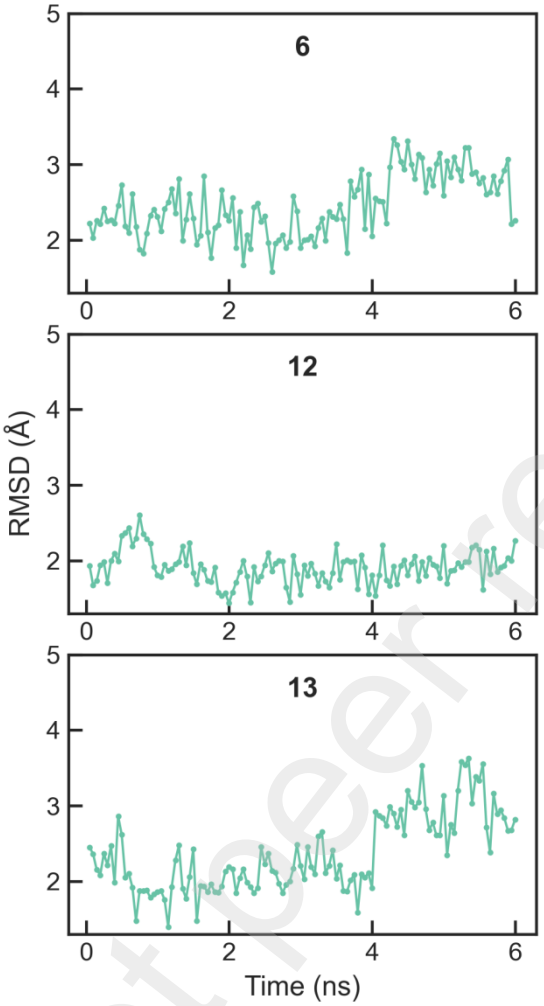


Figure S38. Enzymatic *in vitro* activity assay against all three LH isoforms (LH1-3). Compounds **1**, **2**, **6**, **11**, **12**, and **13** were evaluated.



Figures S39. Ligand root mean squared deviation (RMSD) for compounds **6**, **12,** and **13**. The structure was first aligned to the alpha carbons of the initial structure, and then the RMSD was calculated on the ligand atoms without fitting.

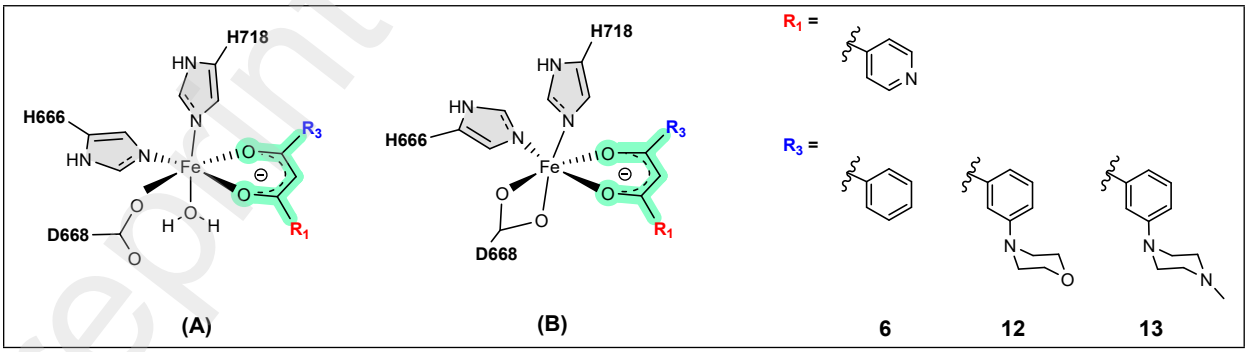


Figure S40. General stereochemistry of the active site in **6** and **12** (A), and **13** (B). The green highlighted residue in (A) and (B) is the enolate part of the ligands.

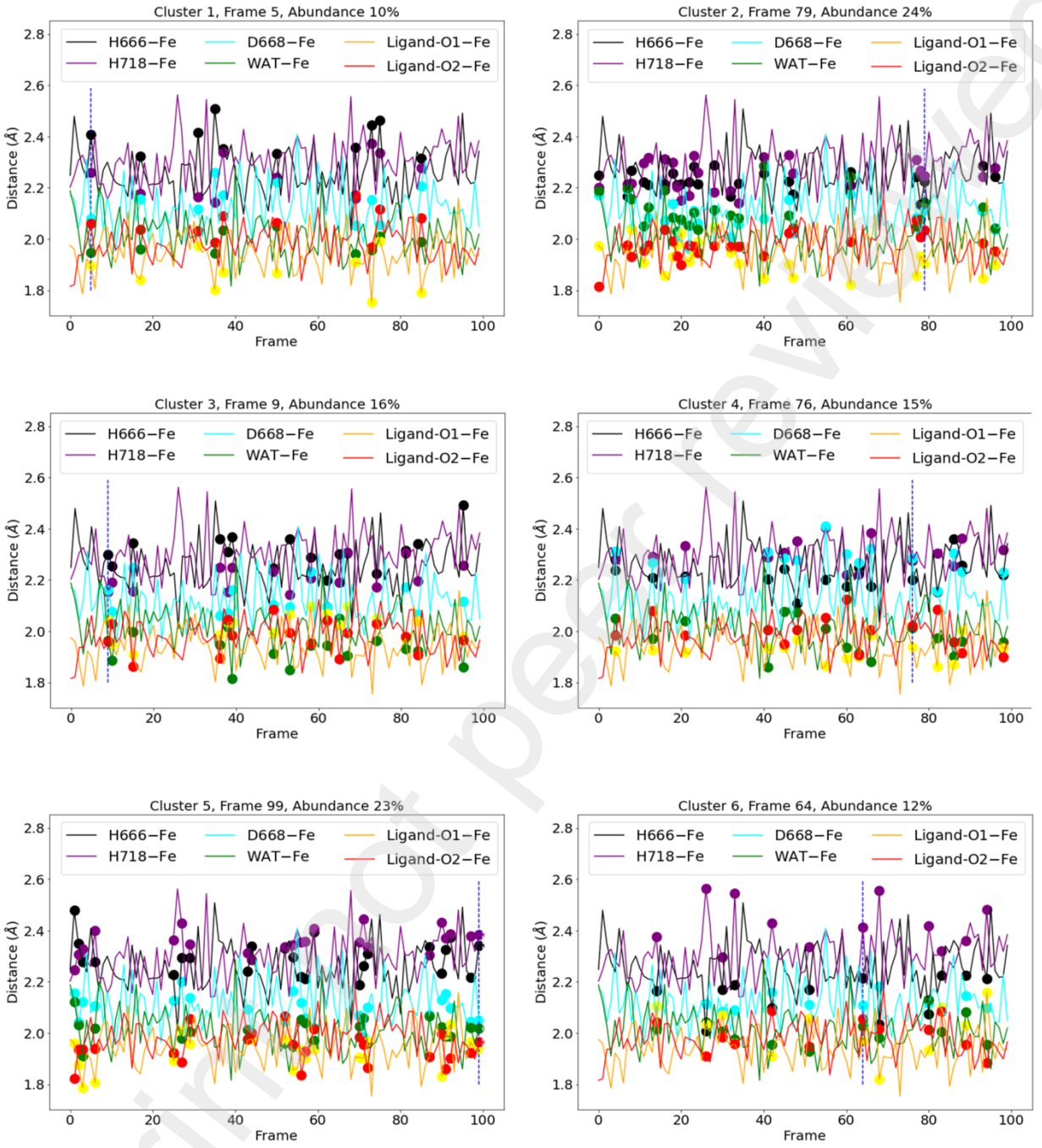


Figure S41. Clusters from the *k*-means analysis for compound **6**. The number of each cluster, related representative/frame, and the abundance percentage are given on top of each graph. The vertical dashed line in each graph shows the closest representative to the centroids of that cluster in all six dimensions. The bold dots show the other frames belong to the cluster.

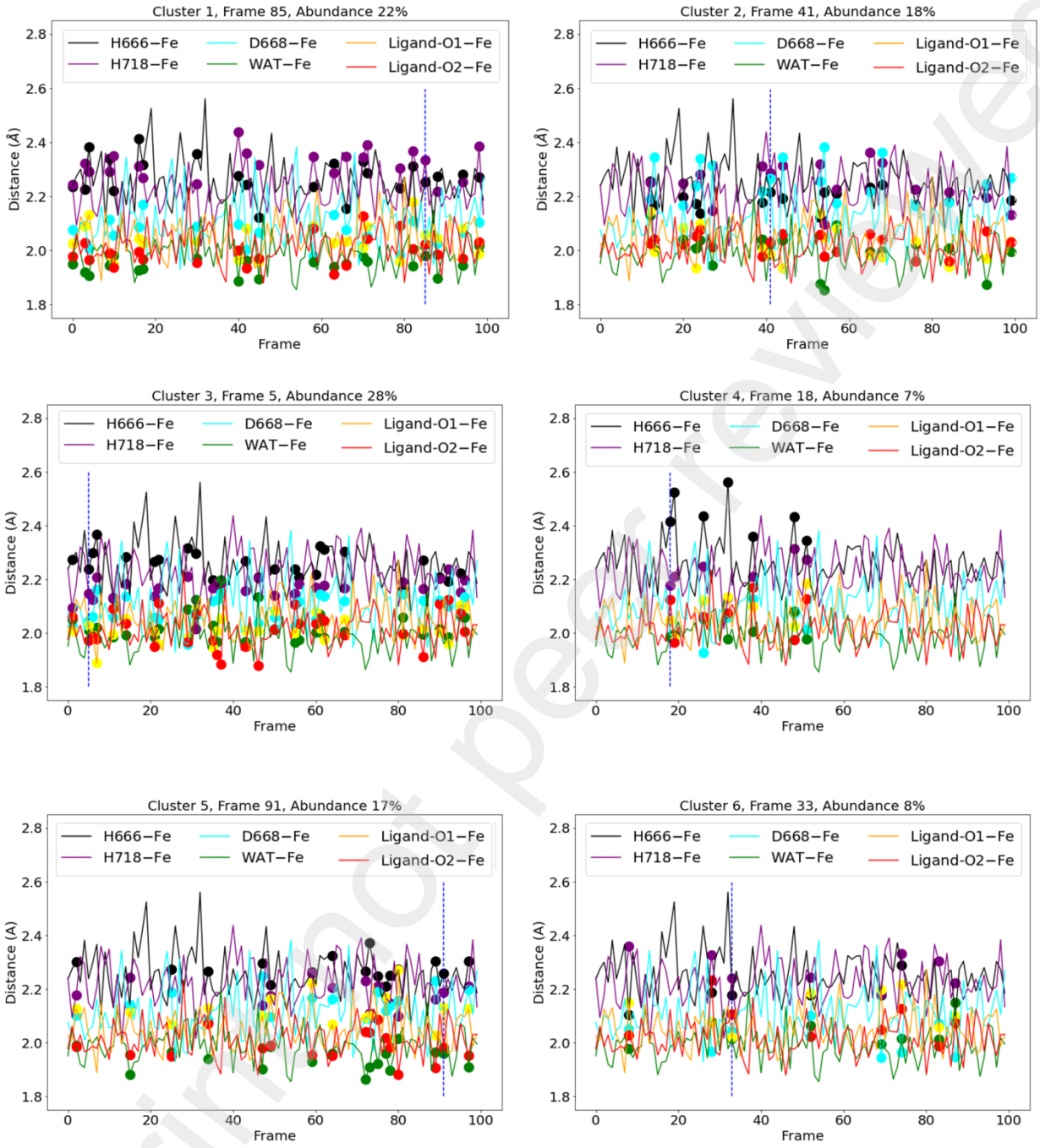


Figure S42. Clusters from the *k*-means analysis for compound **12**. The number of each cluster, related representative/frame, and the abundance percentage are given on top of each graph. The vertical dashed line in each graph shows the closest representative to the centroids of that cluster in all six dimensions. The bold dots show the other frames belong to the cluster.

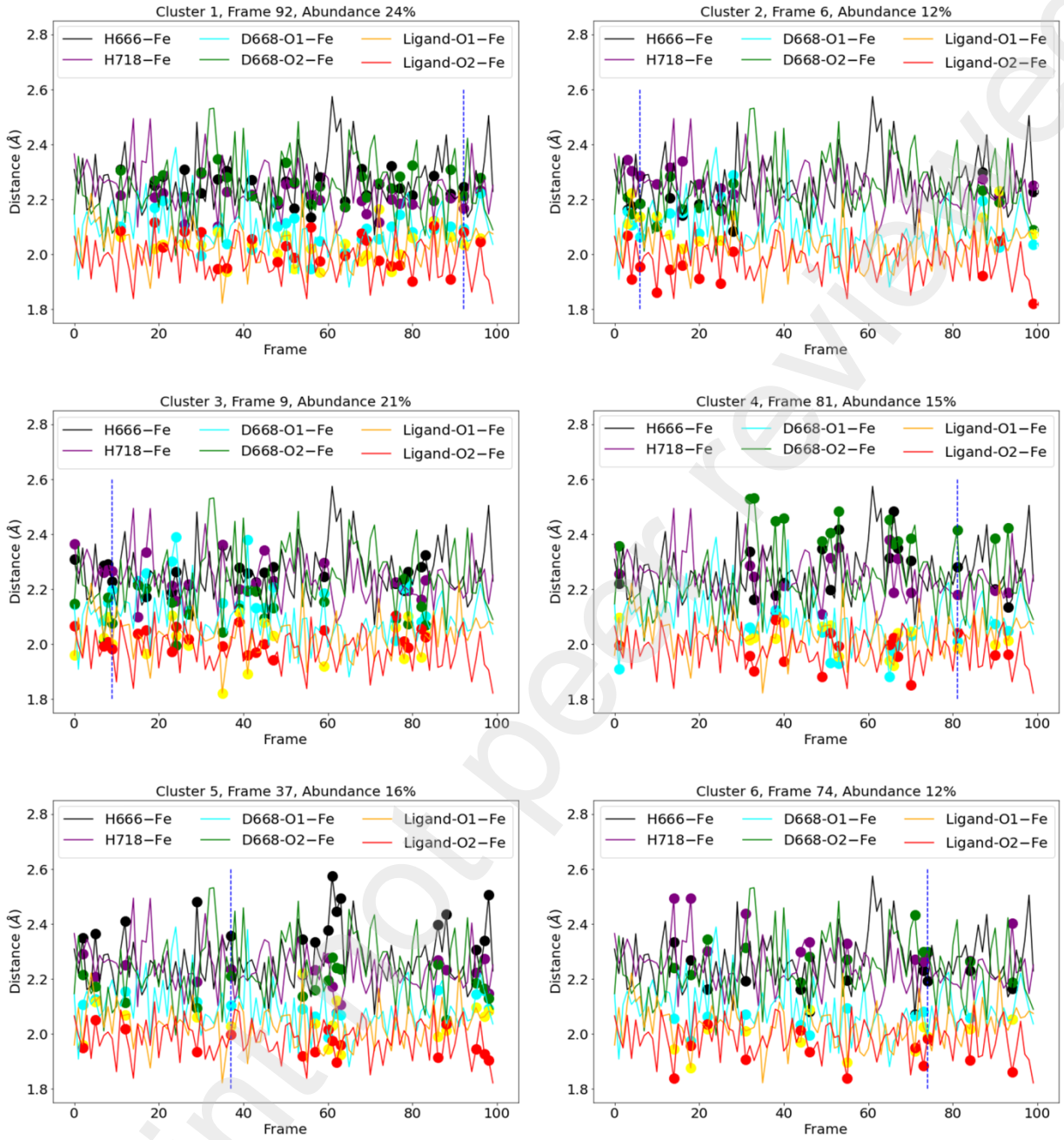
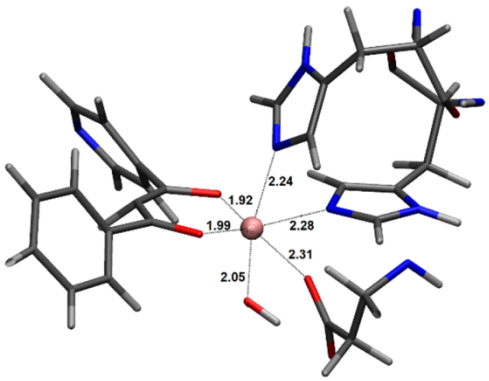
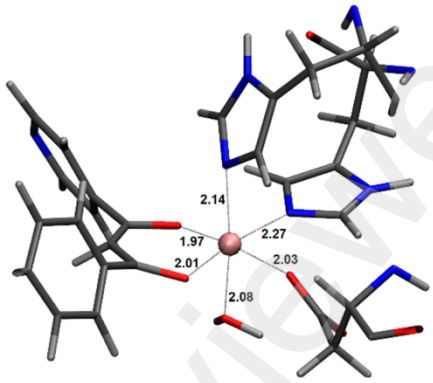


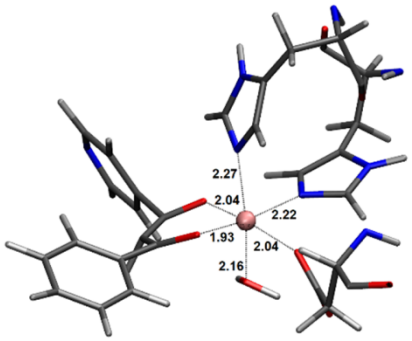
Figure S43. Clusters from the *k*-means analysis for compound **13**. The number of each cluster, related representative/frame, and the abundance percentage are given on top of each graph. The vertical dashed line in each graph shows the closest representative to the centroids of that cluster in all six dimensions. The bold dots show the other frames belong to the cluster.



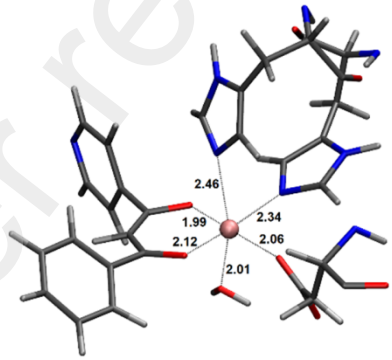
Cluster 1, Frame 5, 10%



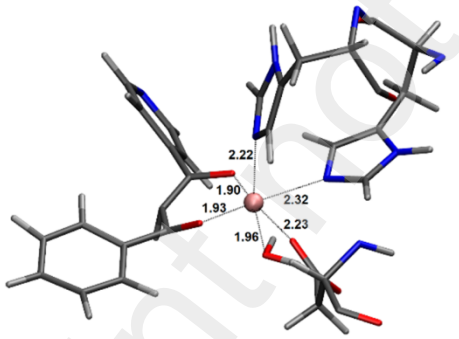
Cluster 2, Frame 79, 24%



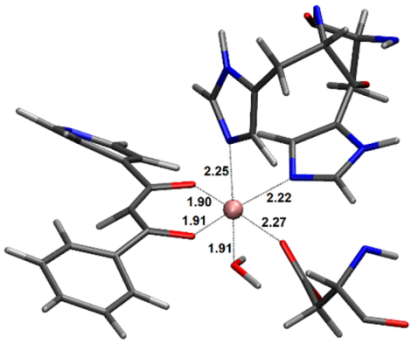
Cluster 3, Frame 9, 16%



Cluster4, Frame 76, 15%



Cluster 5, Frame 99, 23%



Cluster 6, Frame 64, 12%

Figure S44. Close-up to show the interactions between Fe(II) and residues of the active site in each representative structure of compound **6**. The distances between the iron and H666, D668, H718, water molecule, and ligand are given in each figure. The MM region is not shown for more clarity.

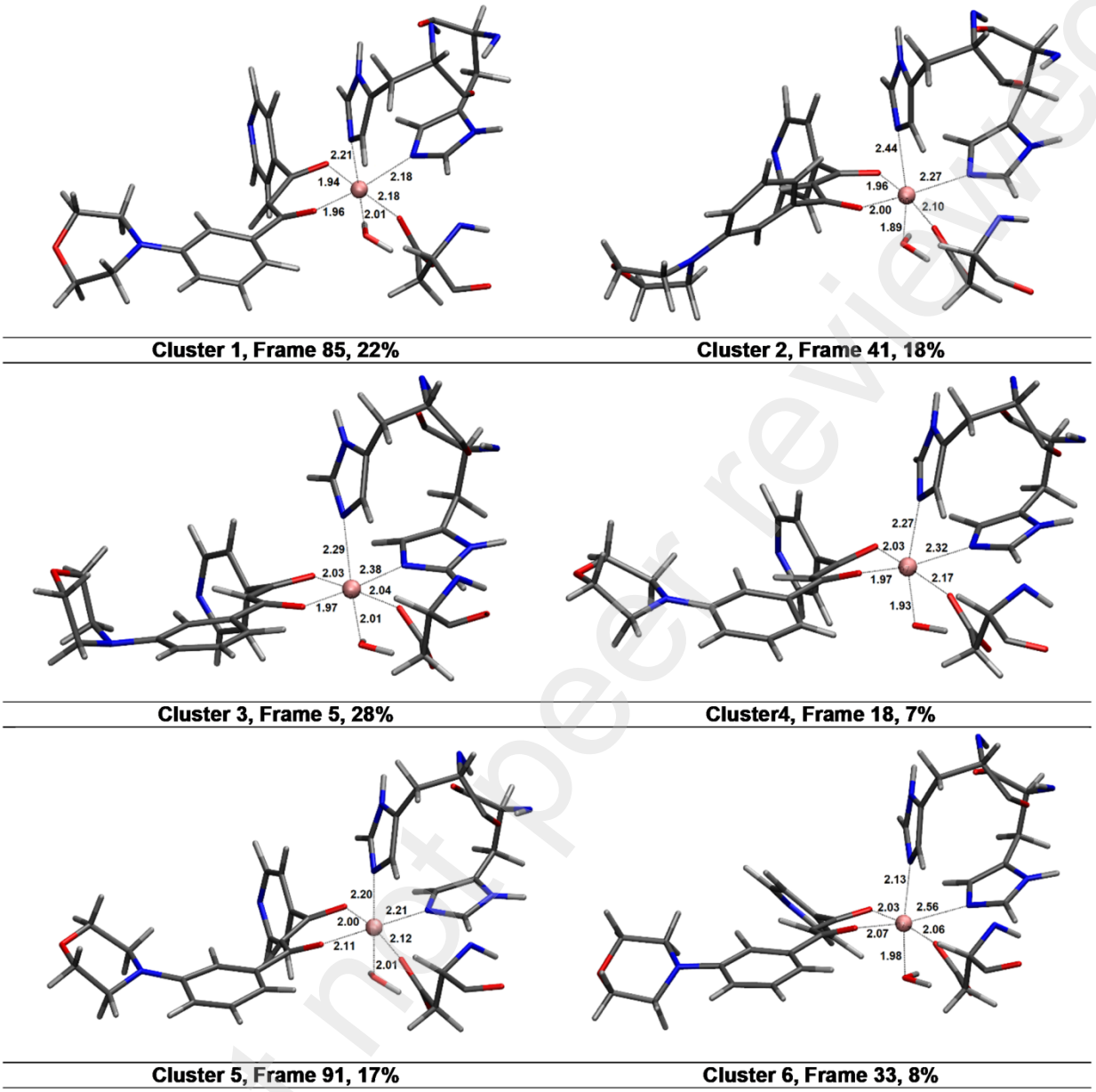


Figure S45. Close-up to show the interactions between Fe(II) and residues of the active site in each representative structure of compound **12**. The distances between the iron and H666, D668, H718, water molecule, and ligand are given in each figure. The MM region is not shown for more clarity.

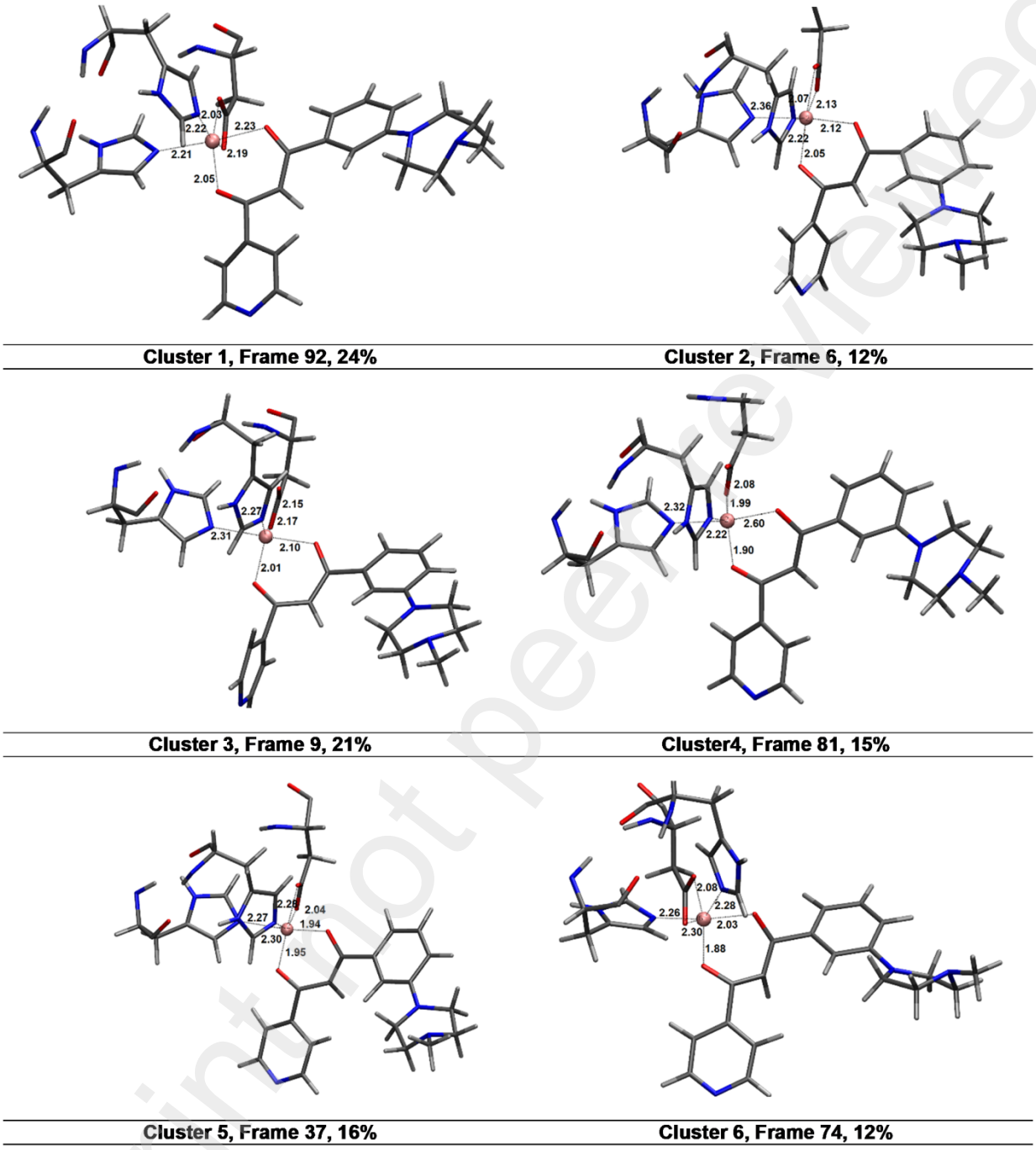


Figure S46. Close-up to show the interactions between Fe(II) and residues of the active site in each representative structure of compound **13**. The distances between the iron and H666, D668, H718, and ligand are given in each figure. The MM region is not shown for more clarity.

Table S1. The results of the *k*-means clustering analysis for each system include the number of clusters, position of the centroid in the trajectory, percentage, and average distance. The last column lists the calculated relative QM/MM optimization energies (eV) for each system's representative frames at ωB97X-D/6–31G(d,p) level of theory with AMOEBABIO18 Polarizable Force Field.

Comp.	No. cluster	Pos. (ns)	Pos. (frame)	Pct. (%)	Avg. dist. (Å)	Rel. opt. Energy (eV)
6	1	0.25	5	10	0.087	84
	2	3.95	79	24	0.082	90
	3	0.45	9	16	0.085	43
	4	3.80	76	15	0.058	39
	5	4.95	99	23	0.075	0
	6	3.20	64	12	0.102	87
12	1	4.25	85	22	0.071	140
	2	2.05	41	18	0.072	10
	3	0.25	5	28	0.062	0
	4	0.90	18	7	0.093	249
	5	4.55	91	17	0.034	125
	6	1.62	33	8	0.103	239
13	1	4.60	92*	24	0.082	140
	2	0.30	6	12	0.074	179
	3	0.45	9	21	0.073	236
	4	4.05	81	15	0.094	122
	5	1.85	37	16	0.077	171
	6	3.70	74*	12	0.099	0

*In the case of compound **13**, frame 74, with a cluster percentage of 12% is the most stable structure *but* with the lowest relative energy, while frame 92 has the highest cluster percentage of 24% but with higher energy of 140 eV. In this case, IE_{QMMM} (see table SX2) was calculated for both frames to consider both the representative with the lowest optimization energy and the representative with the highest population.

Table S2. Calculated values of the QM and MM components of the interaction energies (IE) for all three systems at ωB97X-D/6–31G(d,p) level of theory with AMOEBABIO18 Polarizable Force Field. All the values are in Hartree, otherwise stated.

QM/MM components	Compound			
	6	**12**	**13 (frame 74)***	**13 (frame 92)***
$QM_{Active\ site+Ligand}$	–38840.6206	–38303.8954	–38410.8483	–36907.7499
$MM_{LH2+Ligand}$	–202.3891	–200.8681	–195.7376	–198.0642
$QM_{Active\ site}$	–37809.3013	–37253.1502	–37360.1305	–36163.0193
QM_{Ligand}	–1031.0044	–1050.4306	–1050.4305	–744.4751
MM_{LH2}	–202.2898	–200.8380	–195.7060	–198.0008
IE_{QM}	–0.3149	–0.3146	–0.2873	–0.2555
IE_{MM}	–0.0993	–0.0301	−0.0317	–0.0634
IE_{QMMM}	–0.4142	–0.3447	−0.3190	–0.3189
IE_{QM} *(kcal mol^{-1})***	–160.3	–197.6	–197.4	–180.3
IE_{MM} *(kcal mol^{-1})***	–39.8	–62.3	–18.9	–19.9
$IE_{QM/MM}$ *(kcal mol^{-1})****	–200.1	–259.9	–216.3	–200.2

* Calculated values for two representatives of compound **13** (frames 74 and 92) indicate that frame 74, which corresponds to the lowest optimization energy (see Table SX1), also has the lowest $IE_{QM/MM}$. The ~16 kcal mol^{-1} calculated energy difference between frames 74 and 92 is due to a stronger interaction between **13** and the active site in frame 74, as opposed to a stabilization due to the MM subsystem. Based on these results, frame 74 was used for further study and calculations in the main text.

** Interactions due to the QM region (IE_{QM}) and the MM region (IE_{MM}).

*** $IE_{QM/MM} = [QM_{active\ site\ +\ ligand} - (QM_{active\ site} + QM_{ligand})] + [MM_{LH2\ +\ ligand} - MM_{LH2}]$