Electronic Supporting Information (ESI)

Protein self-assembly induces promiscuous nucleophilic biocatalysis in Morita-Baylis-Hillman (MBH) reaction

Pralhad N. Joshi, Landa Purushottam, Nirmal K. Das, Saptarshi Mukherjee, Vishal Rai*

Department of Chemistry, Indian Institute of Science Education and Research (IISER) Bhopal, Bhopal Bypass Road, Bhauri, Bhopal 462 066 India.

vrai@iiserb.ac.in

Contents

Supplementary methods

General information: The proteins, reagents and fluorophore (Thioflavin T, ThT) were purchased from Sigma-Aldrich. Aqueous buffers were prepared freshly using Millipore Grade I water (Resistivity > 5 MΩ cm, Conductivity < 0.2 μS/cm, TOC <30 ppb). The final pH was adjusted using pH meter Mettler Toledo (FE20). All the solvents used were reagent grade. The reaction mixture was vortexed in incubator-shaker (Thermo Scientific MaxQ 8000) 350 rpm.

Chromatography: Samples were purified by flash column chromatography (Combiflash Rf 200) with 230-400 mesh silica gel. Thin-layer chromatography (TLC) was performed on Merck TLC plates (Silica gel 60 F_{254}) and visualized using a UV lamp (254 nm) and stains such as 2,4-dinitrophenylhydrazine (2,4-DNP), iodine.

Nuclear magnetic resonance spectra: ^{1}H and ^{13}C spectra were recorded on Bruker Advance III 400 MHz and 500 MHz NMR spectrometer. ^{1}H NMR spectra were referenced to TMS (0 ppm), ^{13}C NMR spectra were referenced to $CDCl_3$ (77.16 ppm). Percentage conversions were elucidated from ^{1}H NMR in all the cases, unless specified otherwise. Peak multiplicities are designated by the following abbreviations: s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets; ddd, doublet of doublet of doublets. Spectra were recorded at 25 °C.

Mass spectrometry: Agilent Technologies 1200 series HPLC paired to a 6130 mass spectrometer with electron spray ionization (ESI) was used for LC-LRMS data. HRMS data was collected by Bruker Daltonics MicroTOF-Q-II. Matrix assisted laser desorption/ionisation (MALDI) time of flight mass spectrometry was performed with Bruker Daltonics UltrafleX treme, Software-Flex control version 3.4, Matrix-Sinapic acid and α-Cyano-4-hydroxycinnamic acid (HCCA).

Fluorescence spectroscopy: Steady-state absorption and emission measurements were carried out in a Carry-100 UV-Vis Spectrophotometer and HORIBA JOBIN YVON, FLUOROLOG 3-111, respectively. The fluorescence spectra were measured with a quartz cuvette of 1 mm path length. Lysozyme C and ThT were excited at 295 nm and 450 nm, respectively. The fluorescence emission was collected from 305 nm to 500 nm and 460 nm to 650 nm for Lysozyme C and ThT, respectively with an integration time of 0.1 s. The emission and the excitation slits were kept at 2 nm and 1 nm, respectively. In order to avoid inner filter effects, the fluorescence intensity was corrected according to the following equation [1]

$$I = I_{obs} \times e^{(A_{ex} + A_{em})/2} \quad (1)$$

where, I is the corrected fluorescence intensity and I_{obs} is the observed background-subtracted fluorescence intensity of the sample under investigation. A_{ex} and A_{em} are the measured absorbance at the excitation and emission wavelengths, respectively.

All solutions were prepared in 100 mM phosphate buffer, pH = 7.0. All experiments were carried out at 25 °C, unless stated otherwise. For spectroscopic measurements, the concentration of the protein was varied from 0.5 mM to 5 mM keeping the concentration of ThT fixed at ~6 μM.

Field Emission Scanning Electron Microscopy (FE-SEM): To carryout FE-SEM, the samples were drop casted on a carbon tape from the vial using a micropipette. The samples were dried inside a vacuum desiccator for 48 h and gold coating was done for 120 second. The images were recorded on Carl Zeiss (Ultraplus) FE-SEM at an accelerating voltage of 10 kV.

Dynamic Light Scattering (DLS): Solution of lysozyme C **5c** (2.5 mM) was prepared in phosphate buffer (0.1 M, pH 7.0) and filtered through a 0.45 μm syringe filter. Hydrodynamic diameter (D_H) was calculated for this solution in the 25 °C on a Delsa™ Nano (Beckman Coulter) instrument.

Molecular Docking: The three-dimensional structure of the lysozyme C **5c** (PDB code: 2LYZ[2]) was obtained from Protein Data Bank. A compatible PDB file corresponding to the structure of ThT was created and saved using ChemDraw software (Version 14.0.0.117) for reading in AutoDock4.2 software[3] program. A blind docking study was carried out by AutoDock4.2 software to obtained energy minimized docked conformation for the protein (Lysozyme C)-ligand (ThT) assemble, which uses the Lamarckian genetic algorithm (LGA).[4] As per the requirement for LGA, all the water molecules were removed and hydrogen atoms were added followed by computing Gasteiger charges. The parameters used for docking were as follows: grid size 126, 126, and 126 Å along the X, Y, and Z axes having 0.375 Å grid spacing and the centre of the grid box was taken corresponding to x,y,z = 1.428, 19.536, 8.602. The maximum number of energy evaluations, maximum number of generations, and GA crossover mode were fixed at 250000, 27000, and two points, respectively. The minimum binding energy conformer was chosen from a set of 10 different conformers for the docking simulation used for further analysis and its corresponding stereo representation was prepared in PyMOL.[5]

Methods

Procedure

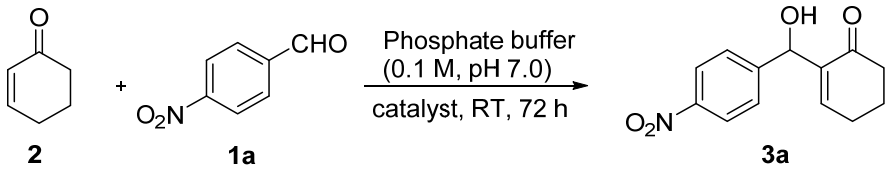
MBH Reaction: Protein was weighed (amount as per the tables) and dissolved in phosphate buffer (0.1 M, pH 7.0, concentration as per the tables) in a 2 ml Eppendorf tube. Subsequently, 2-cyclohexen-1-one (0.15 mmol) was withdrawn from a freshly prepared stock solution in DMSO and added. The mixture was stirred for five minutes followed by addition of aldehyde (0.05 mmol) withdrawn from a freshly prepared stock solution in DMSO. The concentration of 2-cyclohexen-1-one and aldehyde was maintained to control the final phosphate buffer:DMSO ratio as 9:1. The reaction was vortexed in incubator-shaker for 72 h. The reaction was quenched by dilution (active catalyst is low order protein assembly).The product was extracted by using ethylacetate:hexanes (70:30, 4×3 ml). The organic layer was dried over anhydrous sodium sulfate and concentrated on rotary evaporator followed by dried

under high vacuum. Purification was done by flash column chromatography by using 230-400 mesh silica gel.

Protein modification: Lysozyme C **5c** (14.3 mg, 0.001 mmol) was taken in phosphate buffer (7.9 ml, 0.1 M, pH 7.0) into aluminium foil covered conical flask. To this solution, 880 µl of DMSO containing 2,4′-dibromoacetophenone (0.227 mg, 0.001 mmol, from freshly prepared stock solution) was added and then the reaction mixture was vortexed at 25 ^{o}C in incubator-shaker for 8 h. The modification of lysozyme C **5c** was followed by MALDI-ToF-MS using sinapic acid solution as matrix. Unreacted 2,4′-dibromoacetophenone was extracted using ethyl acetate/hexane (6:4, 3×6 ml). Aqueous layer was further subjected to dialysis for desalting and concentrated by lyophilisation before use in the control experiment.

Additional results and discussion

Table S1. Control experiments with amino acids and peptides as catalyst

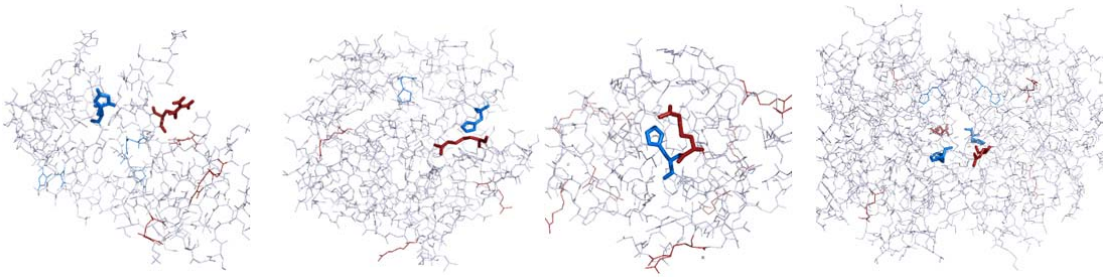


2 **1a** **3a**

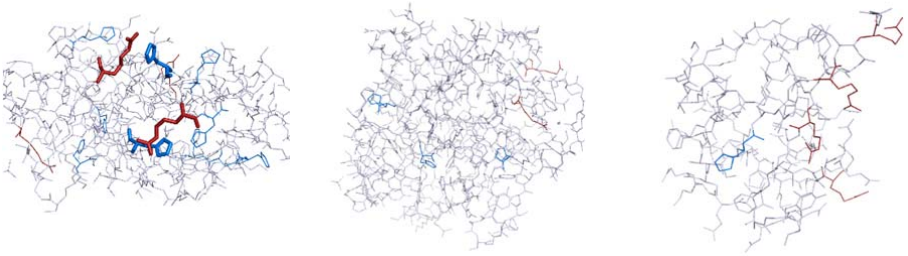
Entry	Catalyst (mol%)	% Conversion[a]
1	His (10)	14
2	Arg (10)	05
3	His (10)+Arg (10)	72
4	His (10)+Thr (10)	36
5	His (10)+Asn (10)	53
6	His (10)+Asp (10)	12
7	His(10)+Arg(10)+Thr(10)+Asn(10)+Asp(10)	18
8	His (10)+Ser (10)	30
9	His (10)+Tyr (10)	23
10	His (10)+Glu (10)	30
11	His (10)+Gln (10)	38
12	His (10)+Lys (10)	57
13	His (10)+Trp (10)	28
14	Fmoc-H(d)PAibG-NH_2 (10)	04
15	Fmoc-H(d)PAibR-NH_2 (10)	12
16	Fmoc-H(d)PAibFR-NH_2 (10)	36

[a] % Conversion determined by ^{1}H NMR.

Proteins contain nucleophilic amino acid residues such as methionine, cysteine, histidine, etc. have been utilized for nucleophilic catalysis.[6] We hypothesised that histidine can be a better catalyst for Morita-Baylis-Hillman reaction. According to the hypothesis, His along with other amino acid residues has been screened out of which histidine and arginine turned out to render desired reactivity. The control experiment supports that His and Arg pair are suitable catalophore where histidine can act as a nucleophilic catalyst and arginine would help in proton shuttling in the rate determining step.



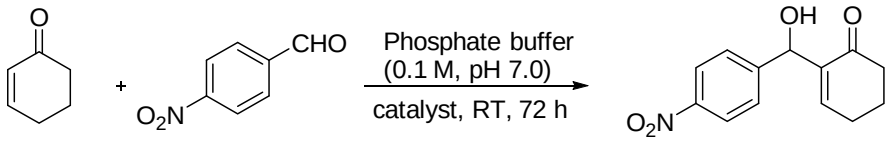
RNase A (5a) **α-Chymotrypsinogen (5b)** **Lysozyme C (5c)** **α-Chymotrypsin (5d)**



Myoglobin (5e) **Trypsin (5f)** **Ubiquitin (5g)**

Figure S1. His-Arg pairs in proteins **5a-g**

Table S2. Optimization of catalyst loading

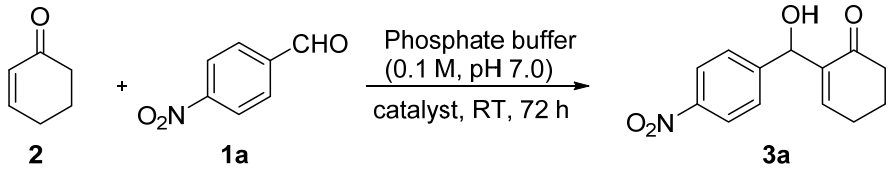


2 **1a** **3a**

Entry	mol%	% Conversion[a]	
		Lysozyme C	Myoglobin
1	0.5	11	46
2	1	27	78
3	2	33	82
4	5	23	80
5	10	72	79
6	15	75	78

[a] % Conversions determined by ^{1}H NMR; Concentration with respect to the protein is [b] 2.5 mM, [c] 1 mM.

Table S3. Optimization of the amount of catalyst



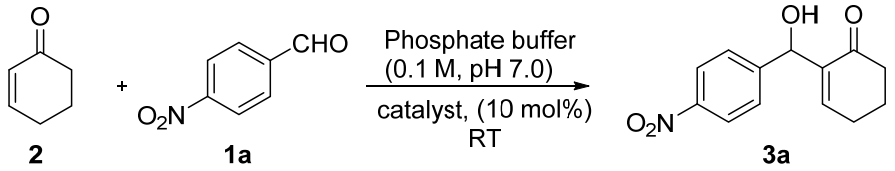
2 **1a** **3a**

Entry	mol%	% Conversion[a]	
		Lysozyme C	Myoglobin
1	0.5	00	00
2	1	05	07
3	2	08	10
4	5	20	78
5	10	72	65
6	15	33	59

[a] % Conversions by ^{1}H NMR. Constant substrate concentration.

Table S2 and S3 suggest that the concentration of protein plays a critical role in catalysing the MBH reaction. In **Table S2**, the concentration is constant for proteins and variable for substrates, whereas in **Table S3** concentration was variable for protein and constant for substrates.

Table S4. Optimization of reaction time

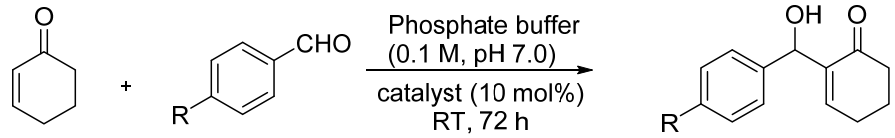


2 **1a** **3a**

Entry	Time (h)	% Conversion[a]	
		Lysozyme C[b]	Myoglobin[c]
1	12	08	10
2	24	18	20
3	48	52	60
4	72	72	79
5	96	74	78

[a] % Conversion determined by ^{1}H NMR. Concentration with respect to the protein is [b] 2.5 mM, [c] 1 mM.

Table S5. Unique catalytic activity of proteins

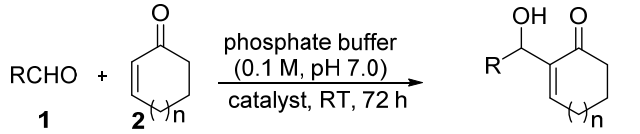


2

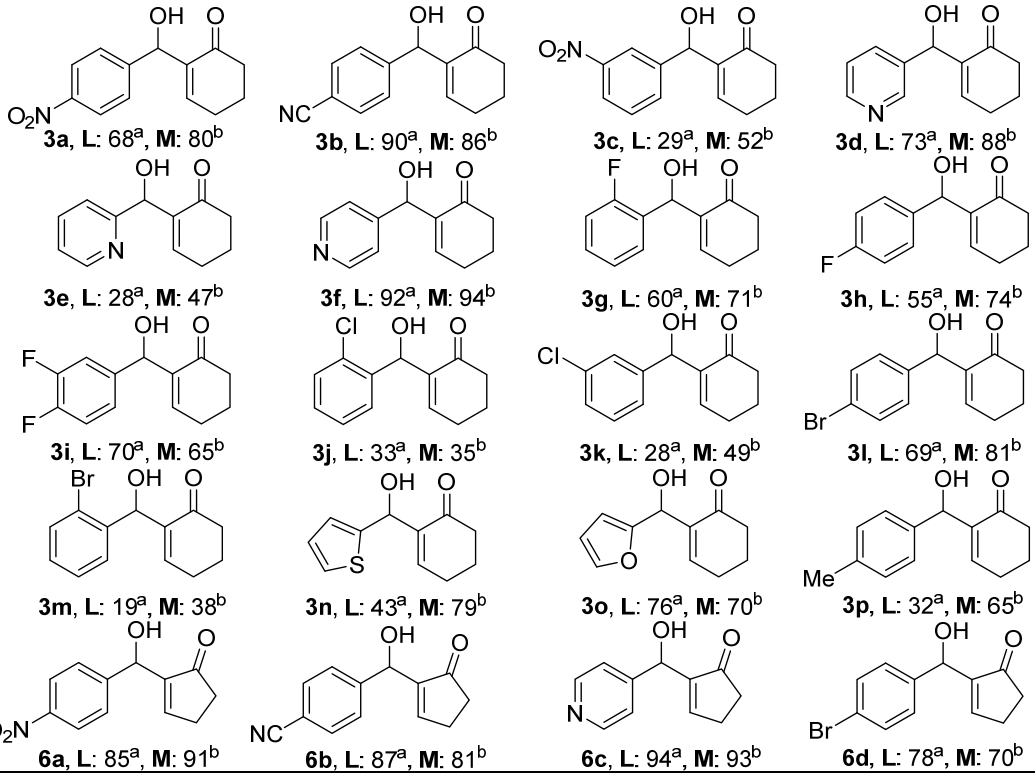
Entry	Catalyst	% Conversion[a]	
		3a	**3l**
1	Fmoc-His-NH_2	10	0
2	DBU	7	2
3	DABCO	2	1
4	DBN	3	1
5	Et_3N	2	0
6	PPh_3	0	0
7	DMAP	47	3
8	Imidazole	61	22
9	**Lysozyme C**	**72[b]**	**73**
10	**Myoglobin**	**79**	**79**

[a] % Conversions determined by ^{1}H NMR. [b] No enantioselectivity was observed.

Table S6. Isolated yields



1 **2**



% Isolated yield [a] Lysozyme C, **L** (2.5 mM, 10 mol%); [b] Myoglobin, **M** (1 mM, 2 mol%)

Fluorescence data

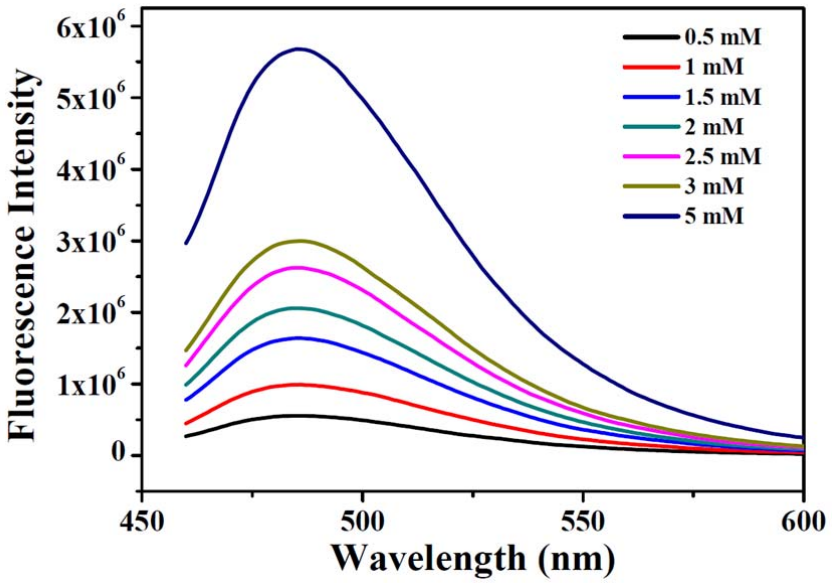


Figure S2: Representative ThT florescence spectra (at timescale of 1 day) with increasing concentration of lysozyme C **5c** marked in the figure.

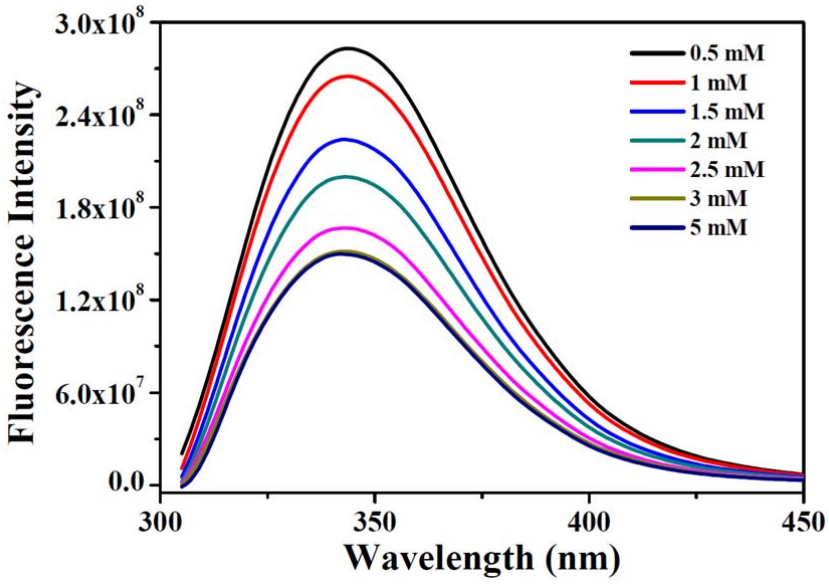


Figure S3: Representative Trp florescence spectra (at timescale of 1 day) with increasing concentration of lysozyme C **5c** marked in the figure.

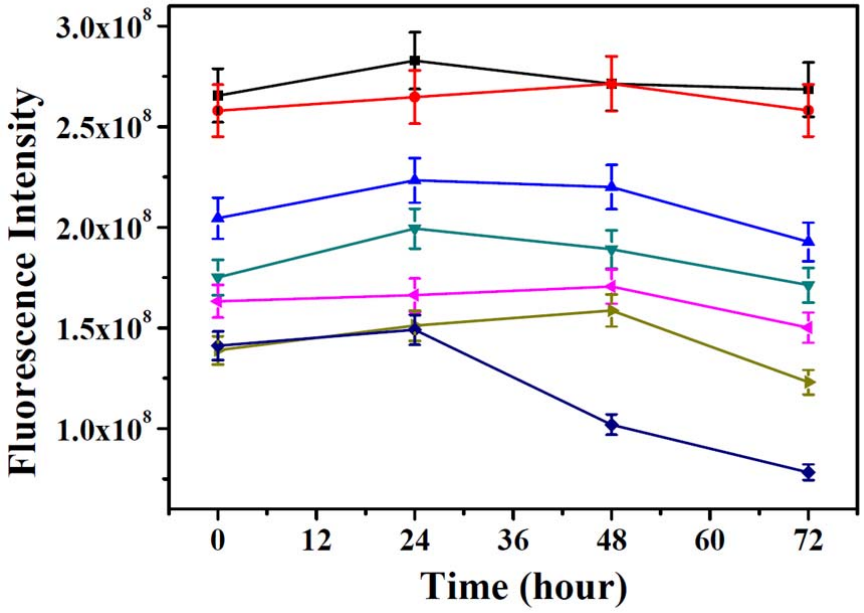


Figure S4: Trp fluorescence assay with increasing concentration of lysozyme C **5c** as a function of time. The colour black, red, blue, cyan, pink, yellow, and navy blue represents the concentration of lysozyme C 0.5, 1, 1.5, 2, 2.5, 3, 5 mM respectively.

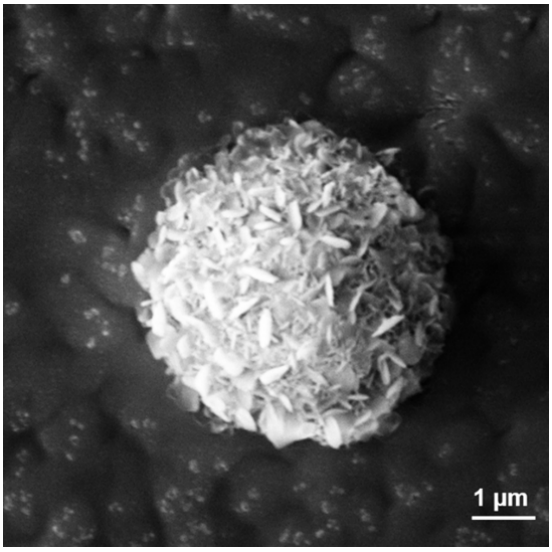


Figure S5: FE-SEM image at 2.5 mM concentration of lysozyme C **5c**.

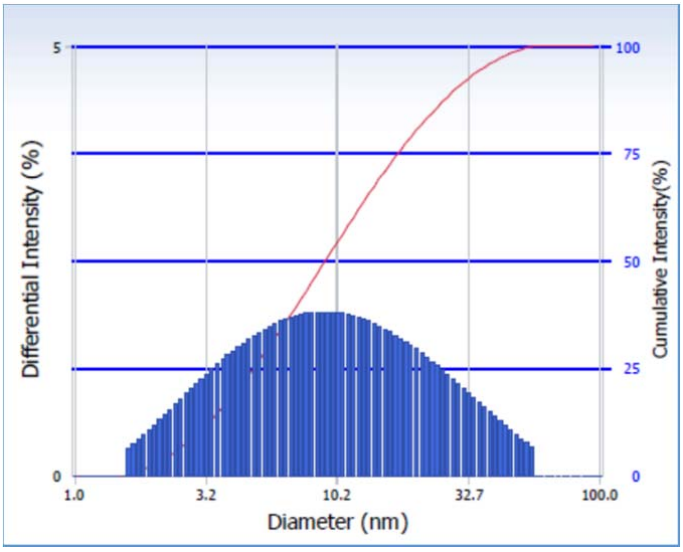
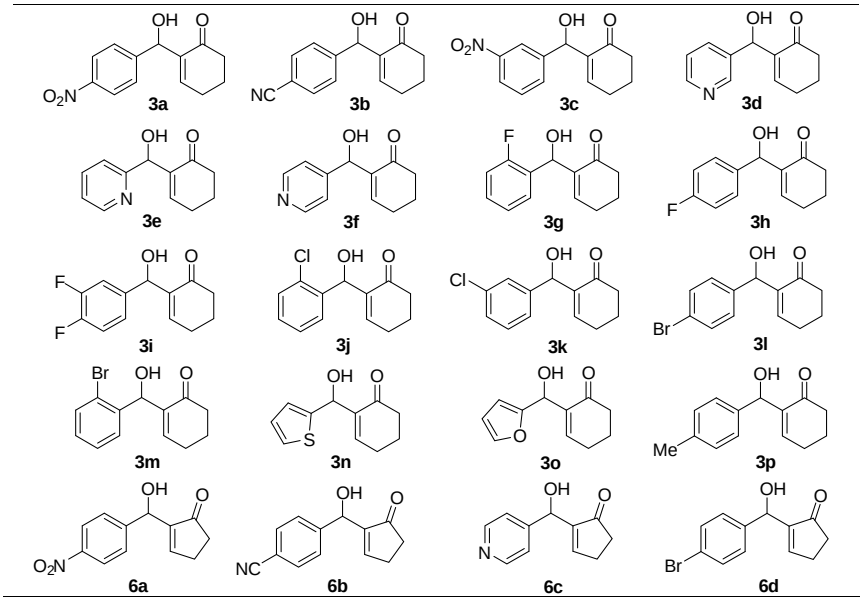


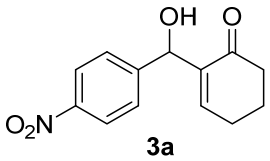
Figure S6: Dynamic Light Scattering (DLS) image at 2.5 mM concentration of lysozyme C **5c**.

Legend



Compound characterization

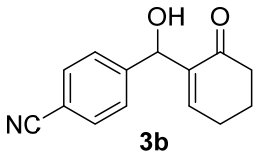
2-(hydroxy(4-nitrophenyl)methyl)cyclohex-2-enone



3a

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 8.16 (d, *J*=8.8 Hz, 2H), 7.52 (d, *J*=8.6 Hz, 2H), 6.83 (t, *J*=4.1 Hz, 1H), 5.60 (s, 1H), 3.62 (bs, 1H), 2.40-2.46 (m, 4H), 1.99 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 200.3, 149.4, 148.3, 147.4, 140.3, 127.3, 123.7, 72.3, 38.6, 25.9, 22.5 ppm. HRMS (ESI) $[MNa]^+$ calcd. for $C_{13}H_{13}NO_4Na$ 270.0742, found 270.0748.

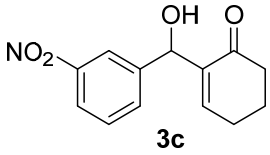
2-(hydroxy(4-isocyanophenyl)methyl)cyclohex-2-enone



3b

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.63 (d, *J*=8.4 Hz, 2H), 7.49 (d, *J*=8.0 Hz, 2H), 6.80 (t, *J*=3.9 Hz, 1H), 5.56 (s, 1H), 3.29 (bs, 1H), 2.41-2.47 (m, 4H), 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 200.3, 148.2, 147.4, 132.3, 127.2, 118.9, 111.4 72.3, 38.6, 25.9, 22.5 ppm. HRMS (ESI) $[MNa]^+$ calcd. for $C_{14}H_{13}NO_2Na$ 250.0844, found 250.0860.

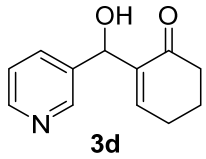
2-(hydroxy(3-nitrophenyl)methyl)cyclohex-2-enone



3c

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 8.21 (s, 1H), 8.12 (d, *J*=8.1 Hz, 1H), 7.73 (d, *J*=7.7 Hz, 1H), 7.51 (t, *J*=7.9 Hz, 1H), 6.85 (t, *J*=4.0 Hz, 1H), 5.59 (d, *J*=5.4 Hz, 1H), 3.58 (d, *J*=5.4 Hz, OH), 2.45 (m, 4H), 2.02 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.3, 148.3, 148.5, 144.4, 140.3, 132.7, 129.4, 122.6, 121.5, 72.1, 38.6, 25.9, 22.5 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{14}NO_4$ 248.09, found 248.49.

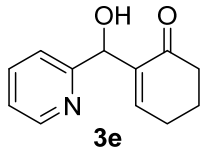
2-(hydroxy(pyridin-3-yl)methyl)cyclohex-2-enone



3d

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 8.52 (d, *J*=2.0 Hz, 1H), 8.45 (dd, *J*=4.8, 1.5 Hz, 1H), 7.73 (dd, *J*=6.3, 1.6 Hz, 1H), 7.25 (dd, *J*=7.9, 4.7 Hz, 1H), 6.86 (t, *J*=4.0 Hz, 1H), 5.57 (s, 1H), 2.39-2.45 (m, 4H), 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.1, 148.7, 148.3, 147.7, 140.6, 137.8, 134.4, 123.4, 70.5, 38.6, 25.9, 22.6 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{12}H_{14}NO_2$ 204.10, found 204.51.

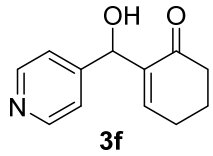
2-(hydroxy(pyridin-2-yl)methyl)cyclohex-2-enone



3e

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 8.51 (d, *J*=4.7 Hz, 1H), 7.66 (td, *J*=7.7, 1.7 Hz, 1H), 7.46 (d, *J*=7.9 Hz, 1H), 7.18 (m, 1H), 7.03 (t, *J*=4.1 Hz, 1H), 5.68 (s, 1H), 2.38-2.47 (m, 4H), 1.98 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 199.6, 160.5, 148.1, 147.6, 140.9, 139.9, 122.6, 121.5, 70.4, 38.7, 26.0, 22.7 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{12}H_{14}NO_2$ 204.10, found 204.34.

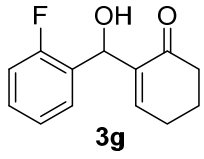
2-(hydroxy(pyridin-4-yl)methyl)cyclohex-2-enone



3f

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 8.48 (d, *J*=6.0 Hz, 2H), 7.28 (d, *J*=6.0 Hz, 1H), 6.86 (t, *J*=4.1 Hz, 1H), 5.51 (s, 1H), 2.42 (m, 4H), 1.98 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 199.9, 151.6, 149.6, 148.1, 140.3, 121.4, 71.1, 38.5, 25.9, 22.5 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{12}H_{14}NO_2$ 204. 10, found 204.23.

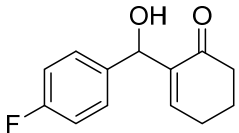
2-((2-fluorophenyl)(hydroxy)methyl)cyclohex-2-enone



3g

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 7.53 (td, *J*=7.6, 1.7 Hz, 1H), 7.27 (m, 1H), 7.17 (td, *J*=7.5, 1.1 Hz, 1H) 7.01 (m, 1H), 6.67 (t, *J*=4.2 Hz, 1H), 5.83 (s, 1H), 2.47 (m, 2H), 2.37 (m, 2H) 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.8, 159.8 (d, *J*=246.3 Hz), 147.6, 139.7, 129.2 (d, *J*=8.2 Hz), 128.6 (d, *J*=13.2 Hz), 128.3 (d, *J*=4.0 Hz), 124.3 (d, *J*=3.5 Hz), 115.2 (d, *J*=21.5 Hz), 67.0 (d, *J*=3.3 Hz), 38.6, 25.9, 22.5 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{14}FO_2$ 221.09, found 221.38.

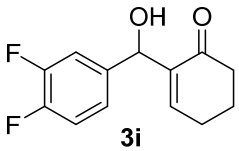
2-((4-fluorophenyl)(hydroxy)methyl)cyclohex-2-enone



3h

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 7.31 (dd, *J*=8.4, 5.5 Hz, 2H), 7.01 (t, *J*=8.7 Hz, 2H), 6.73 (t, *J*=4.0 Hz, 1H), 5.52 (s, 1H), 2.39-2.46 (m, 4H), 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.4, 162.2 (d, *J*=245.5 Hz), 147.3, 141.0, 137.5, 128.2, (d, *J*=8.1 Hz), 115.1 (d, *J*=21.4 Hz), 71.9, 38.5, 25.7, 22.5 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{13}FO_2$ 221.09, found 221.35.

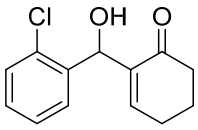
2-((3,4-difluorophenyl)(hydroxy)methyl)cyclohex-2-enone



3i

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.19 (ddd, *J*=11.2, 7.7, 2.0 Hz, 1H), 7.10 (m, 1H), 7.06 (m, 1H), 6.79 (t, *J*=4.2 Hz, 1H), 5.48 (s, 1H), 2.40-2.46 (m, 4H), 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 200.4, 151.0 (dd, *J*=81.4, 12.7 Hz), 149.0 (dd, *J*=80.9, 12.7 Hz), 147.7, 140.6, 139.1 (dd, *J*=4.9, 3.8 Hz), 122.5 (dd, *J*=6.3, 3.6 Hz), 117.0 (d, *J*=17.3 Hz), 115.7 (d, *J*=17.9 Hz), 71.6, 36.6, 25.8, 22.5 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{13}F_2O_2$ 239.08, found 239.38.

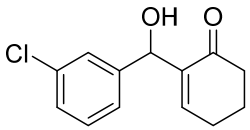
2-((2-chlorophenyl)(hydroxy)methyl)cyclohex-2-enone



3j

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 7.56 (d, *J*=6.5 Hz, 1H), 7.27 (m, 2H), 7.17 (m, 1H), 6.42 (t, *J*=4.0 Hz, 1H), 5.88 (d, *J*=2.9 Hz, 1H), 3.65 (d, *J*=4.4 Hz, OH), 2.43 (m, 2H), 2.29 (m, 2H), 1.93 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.8, 148.0, 139.5, 138.6, 132.5, 129.4, 128.8, 128.4, 127.1, 68.9, 38.6, 25.9, 22.6 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{14}ClO_2$ 237.06, found 237.66.

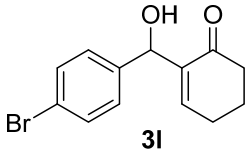
2-((3-chlorophenyl)(hydroxy)methyl)cyclohex-2-enone



3k

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 7.35 (s, 1H), 7.21-7.26 (m, 3H), 6.77 (t, *J*=4.1 Hz, 1H), 5.50 (s, 1H), 2.39-2.46 (m, 4H), 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.4, 147.8, 144.0, 140.7, 134.4, 129.7, 127.9, 126.7, 124.7, 72.1, 38.6, 25.9, 22.6 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{14}ClO_2$ 237.06, found 237.83.

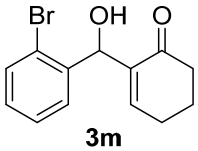
2-((4-bromophenyl)(hydroxy)methyl)cyclohex-2-enone



3l

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 7.46 (d, *J*=8.4 Hz, 2H), 7.24 (d, *J*=8.4 Hz, 2H), 6.74 (t, *J*=4.0 Hz, 1H), 5.50 (s, 1H), 3.46 (bs, 1H), 2.38-2.46 (m, 4H), 1.99 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.4, 147.6, 140.9, 140.8, 131.5, 128.3, 121.5, 72.2, 38.6, 25.9, 22.6 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{14}BrO_2$ 281.01, found 281.18.

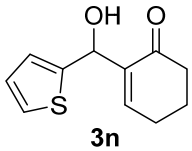
2-((2-bromophenyl)(hydroxy)methyl)cyclohex-2-enone



3m

^{1}H NMR ($CDCl_3$, 400 MHz) δ: 7.68 (dd, *J*=7.8, 1.3 Hz, 1H), 7.53 (dd, *J*=8.0, 0.8 Hz, 1H), 7.37 (t, *J*=7.2 Hz, 1H), 7.16 (td, *J*=7.8, 1.6 Hz, 1H), 6.45 (t, *J*=4.1 Hz, 1H), 5.91 (s, 1H), 3.66 (s, 1H), 2.51 (td, *J*=6.2, 1.6 Hz, 2H), 2.34 (m, 2H), 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 200.8, 148.2, 140.2, 139.6, 132.7, 129.1, 128.7, 127.7, 122.8, 70.9, 38.6, 25.9, 22.6 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{14}BrO_2$ 281.01, found 281.20.

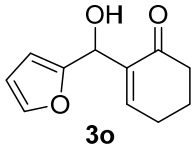
2-(hydroxy(thiophen-2-yl)methyl)cyclohex-2-enone



3n

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.23 (dd, *J*=5.0, 1.3Hz, 1H), 6.93-6.97 (m, 2H), 6.92 (t, *J*=4.2 Hz, 1H), 5.73 (s, 1H), 3.77 (s, 1H), 2.47 (m, 2H), 2.44 (m, 2H), 2.02 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 200.4, 147.5, 146.4, 140.4, 126.9, 125.0, 124.5, 69.9, 38.7, 25.9, 22.6 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{11}H_{13}SO_2$ 209.06, found 209.40.

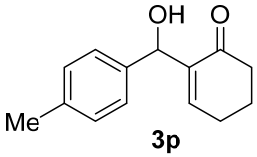
2-(furan-2-yl(hydroxy)methyl)cyclohex-2-enone



3o

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.35 (d, *J*=1.8 Hz, 1H), 6.90 (d, *J*=4.0 Hz, 1H), 6.32 (d, *J*=3.0, 1.8 Hz, 1H), 6.24 (d, *J*=3.0 Hz, 1H), 5.54 (s, 1H), 3.59 (bs, 1H), 2.41-2.48 (m, 4H), 2.00 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 200.1, 154.7, 148.0, 142.2, 138.5, 110.4, 107.0, 67.3, 38.5, 25.9, 22.6 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{11}H_{13}O_3$ 193.21, found 193.44.

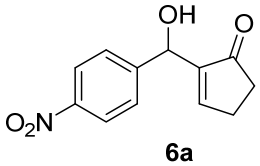
2-(hydroxy(p-tolyl)methyl)cyclohex-2-enone



3p

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.23 (d, *J*=8.0 Hz, 2H), 7.14 (d, *J*=8.0 Hz, 2H),6.75 (t, *J*=4.0 Hz, 1H), 5.52 (s, 1H), 3.41 (bs, OH), 2.44 (m, 2H), 2.38 (m, 2H), 2.34 (s, 3H) 1.99 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 200.6, 147.3, 141.3, 138.9, 137.2, 129.1, 126.5, 72.5, 38.7, 25.9, 22.7, 21.2 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{14}H_{17}O_2$ 217.12, found 217.34.

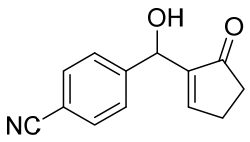
2-(hydroxy(4-nitrophenyl)methyl)cyclopent-2-enone



6a

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 8.21 (d, *J*=8.8 Hz, 2H), 7.58 (d, *J*=8.8 Hz, 2H), 7.30 (m, 1H), 5.67 (s, 1H), 2.63 (m, 2H), 2.48 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 209.4, 159.9, 148.6, 147.7, 146.8, 127.2, 123.9, 69.2, 35.3, 27.0 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{12}H_{12}NO_4$ 233.22, found 233.42.

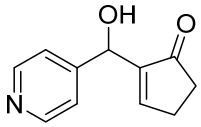
4-(hydroxy(5-oxocyclopent-1-en-1-yl)methyl)benzonitrile



6b

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.62 (d, *J*=8.4 Hz, 2H), 7.51 (d, *J*=8.4 Hz, 2H), 7.30 (m, 1H), 5.59 (s, 1H), 2.64 (m, 2H), 2.45 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 209.4, 159.9, 146.9, 146.8, 132.4, 127.1, 118.8, 111.6, 69.1, 35.2, 26.9 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{13}H_{12}NO_2$ 213.23, found 213.39.

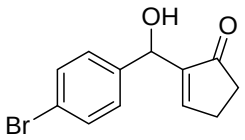
2-(hydroxy(pyridin-4-yl)methyl)cyclopent-2-enone



6c

^{1}H NMR ($CDCl_3$, 500 MHz) δ: 8.47 (d, *J*=6.0 Hz, 2H), 7.39 (m, 1H), 7.33 (d, *J*=6.0 Hz, 2H), 5.54 (s, 1H), 2.98 (s, OH), 2.59 (m, 2H), 2.42 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 209.0, 159.9, 149.7, 147.2, 121.3, 106.6, 67.8, 35.2, 26.9 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{11}H_{12}NO_2$ 189.21, found 189.46.

2-((4-bromophenyl)(hydroxy)methyl)cyclopent-2-enone

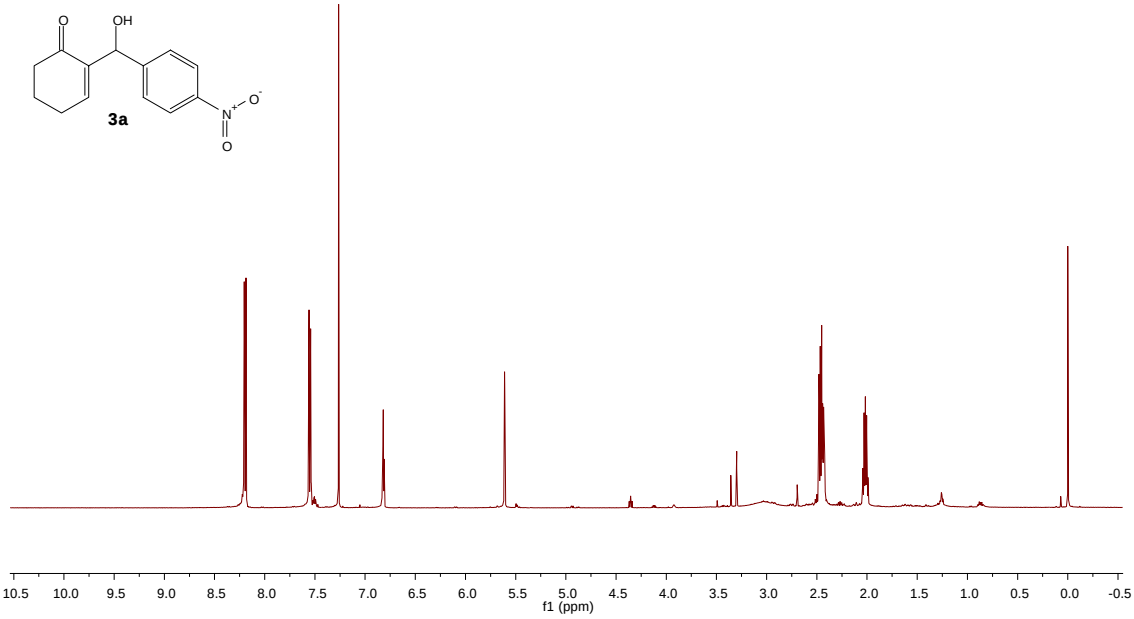


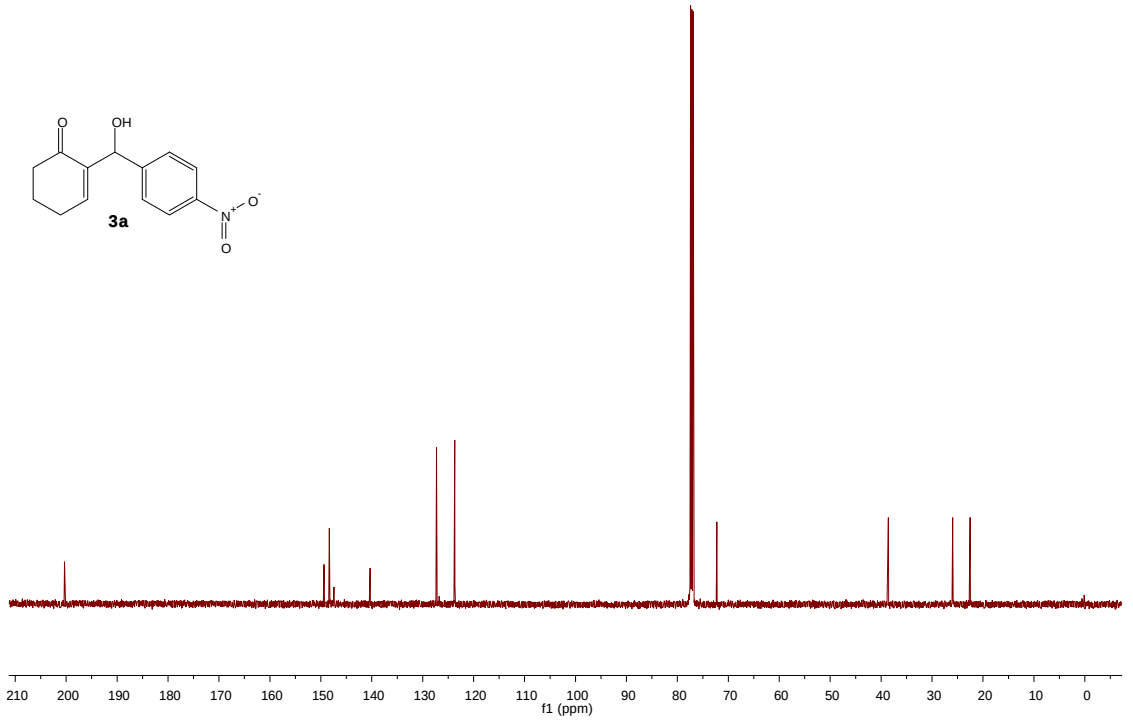
6d

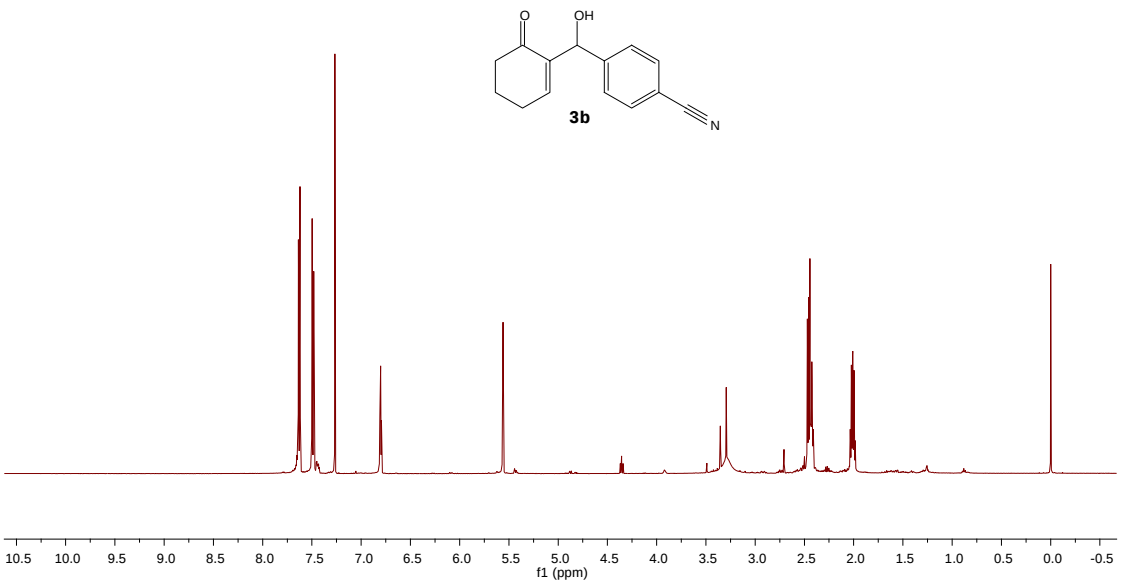
^{1}H NMR ($CDCl_3$, 500 MHz) δ: 7.47 (d, *J*=8.5 Hz, 2H), 7.25 (d, *J*=8.3 Hz, 2H), 7.24 (m, 1H), 5.60 (s, 1H), 2.59 (m, 2H), 2.45 (m, 2H) ppm. ^{13}C NMR ($CDCl_3$, 125 MHz) δ: 209.6, 159.6, 147.5, 140.5, 131.7, 128.2, 121.8, 69.3, 35.3, 26.8 ppm. MS (ESI) $[MH]^+$ calcd. for $C_{12}H_{12}BrO_2$ 267.12, found 267.35.

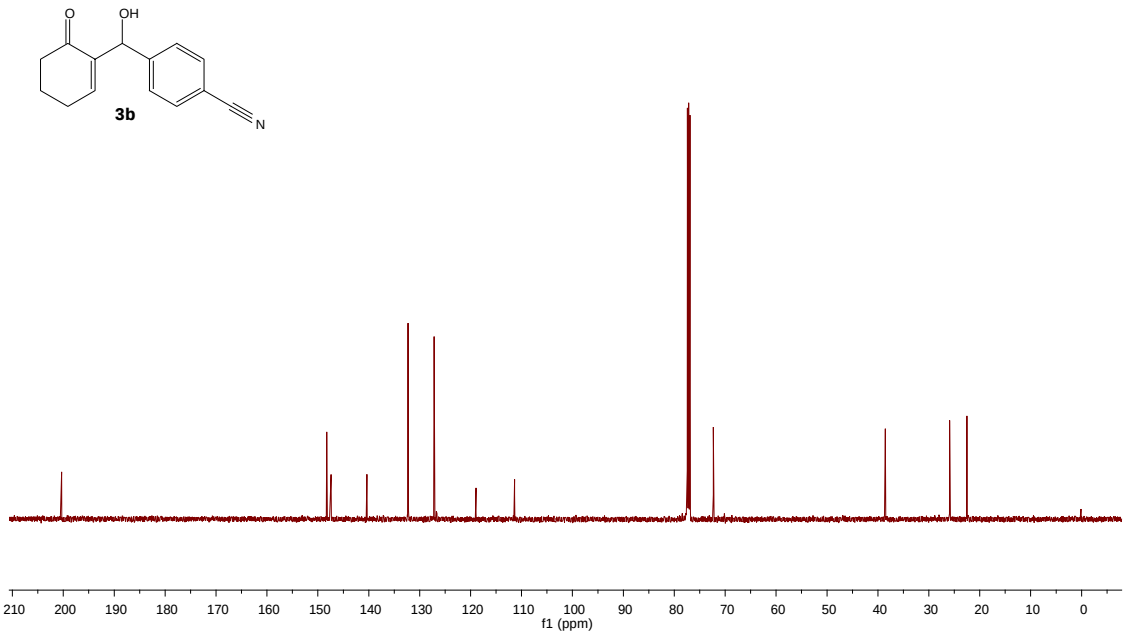
Spectral data

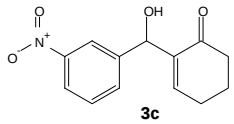
^{1}H and ^{13}C spectra



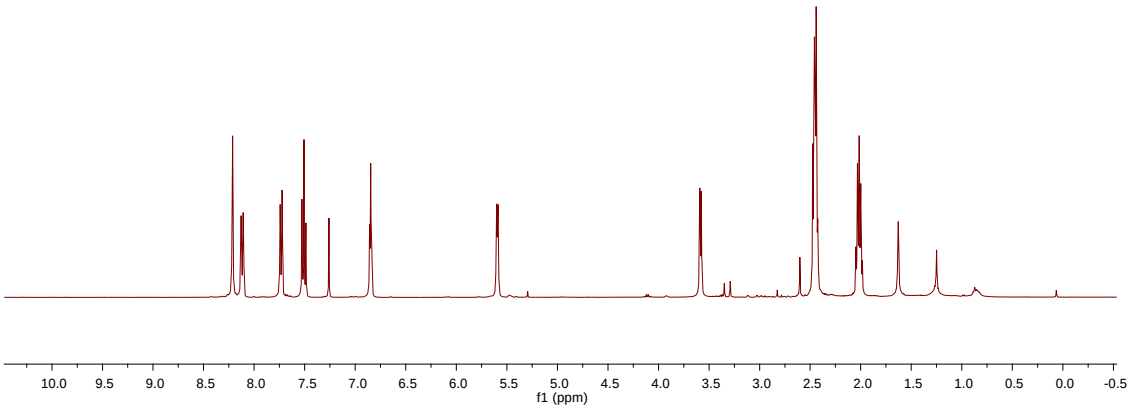


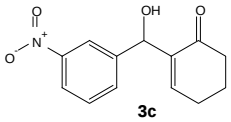

O
OH
N
3b
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)


O
OH
N
3b
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

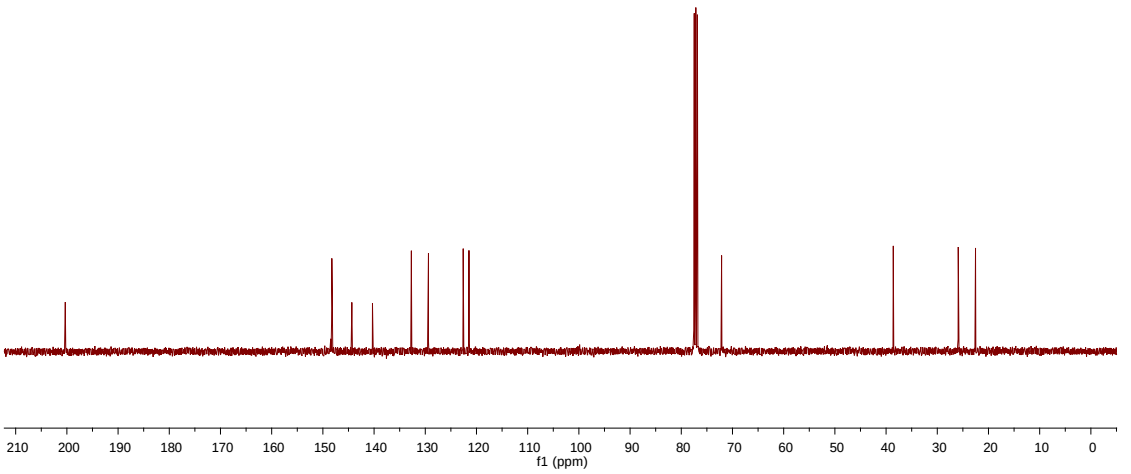

O
OH
O
N+
O-

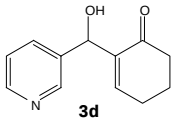
3c


10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

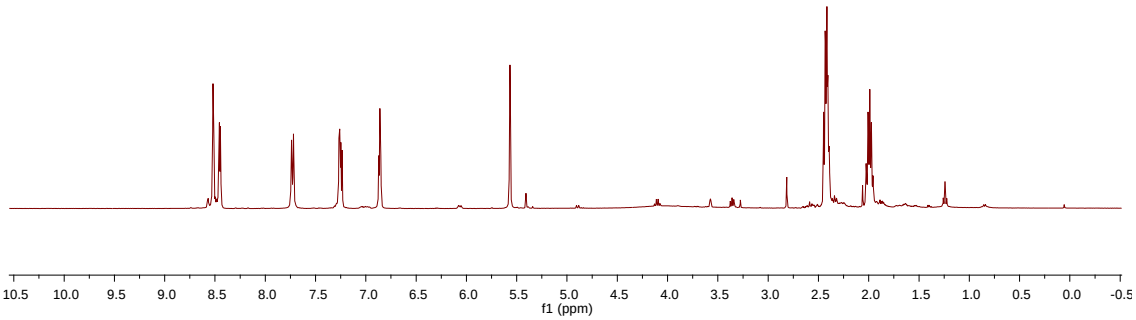

O
OH
O
N+
O-

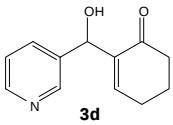
3c


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)


OH
O
N

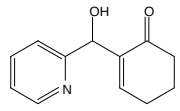
3d


10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

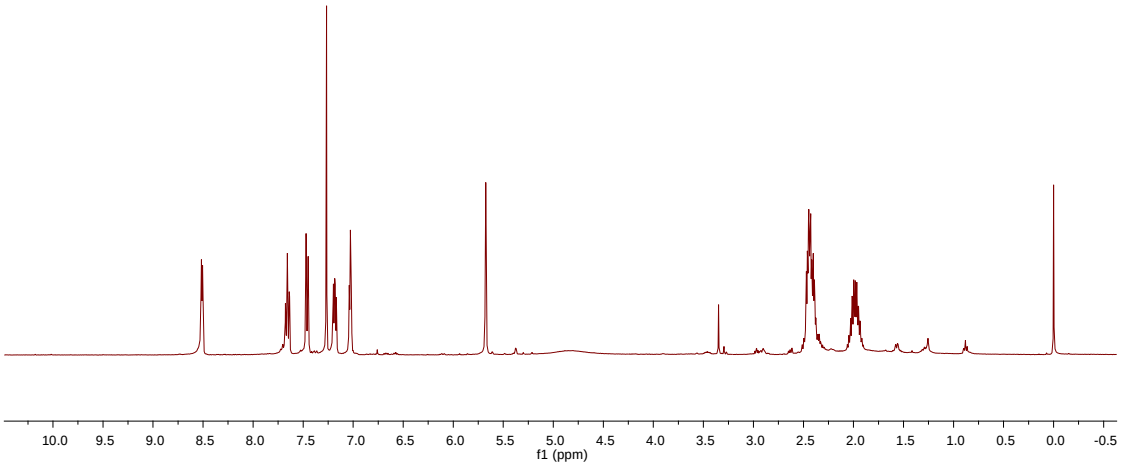

OH
O
N

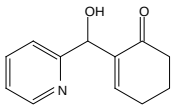
3d

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)


OH
O
N

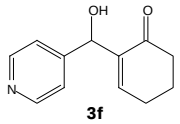
3e


10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

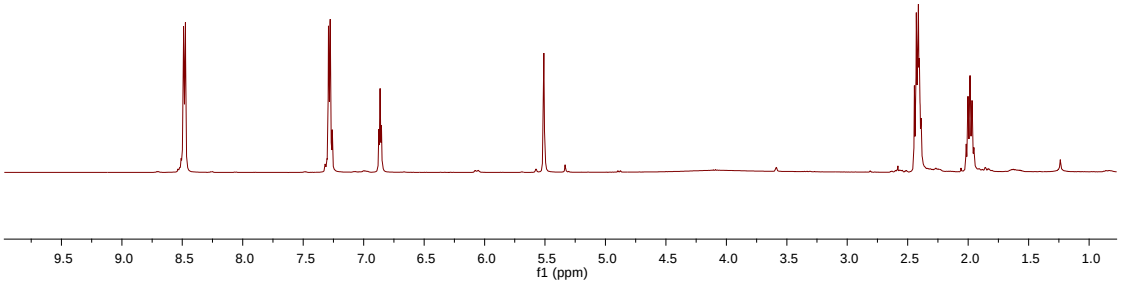

OH
O
N

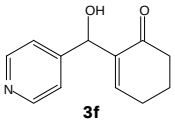
3e

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10
f1 (ppm)

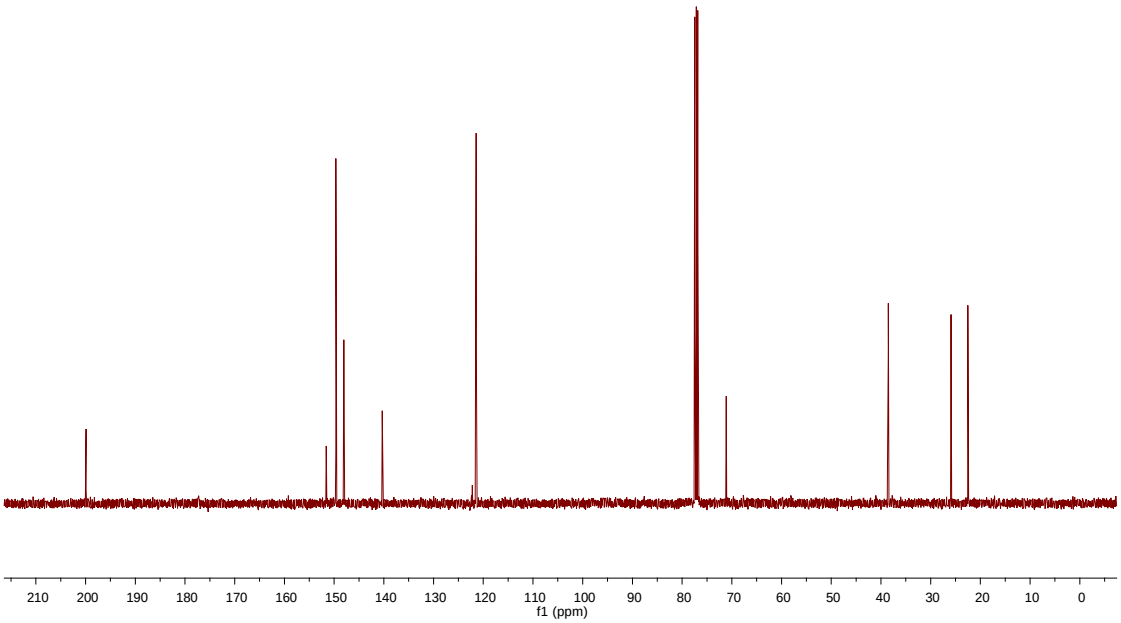

OH
O
N

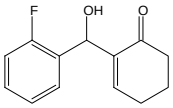
3f


9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0
f1 (ppm)

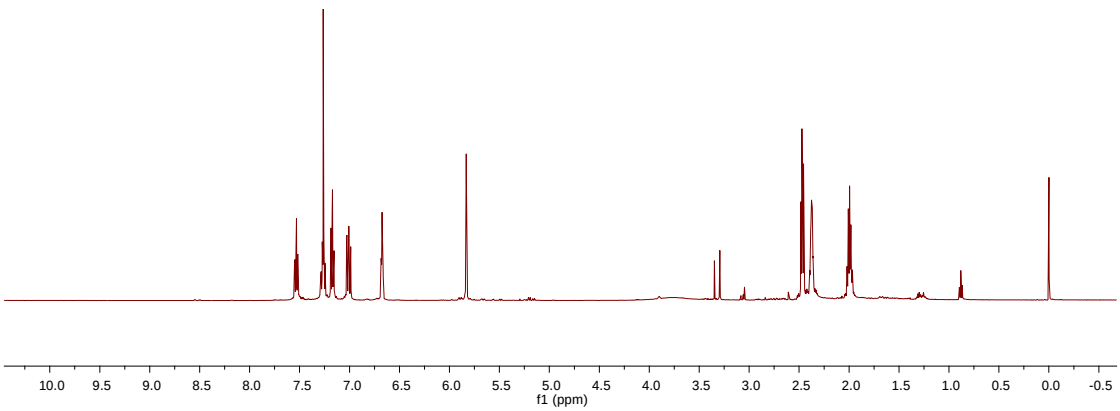

OH
O
N

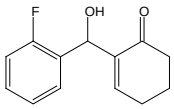
3f


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

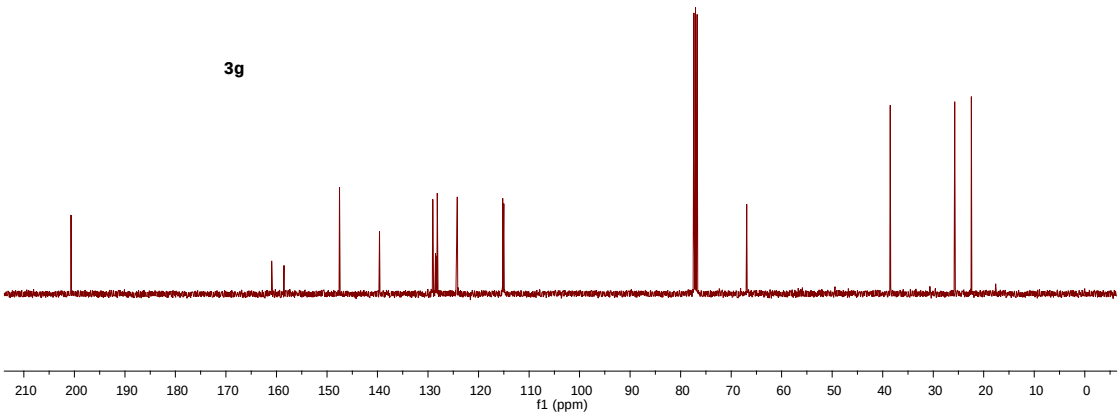

F
OH
O

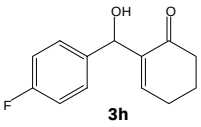
3g


10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

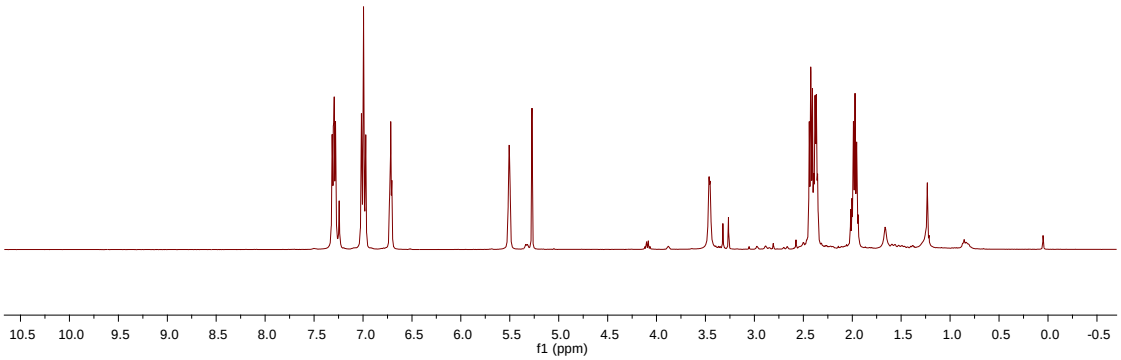

F
OH
O

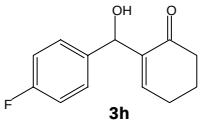
3g


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

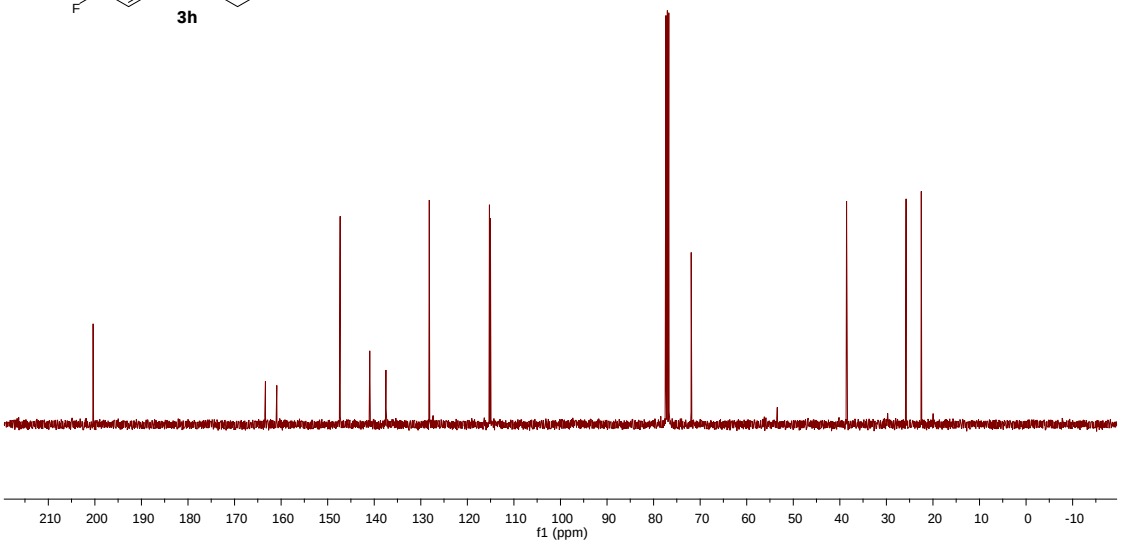

OH
O
F

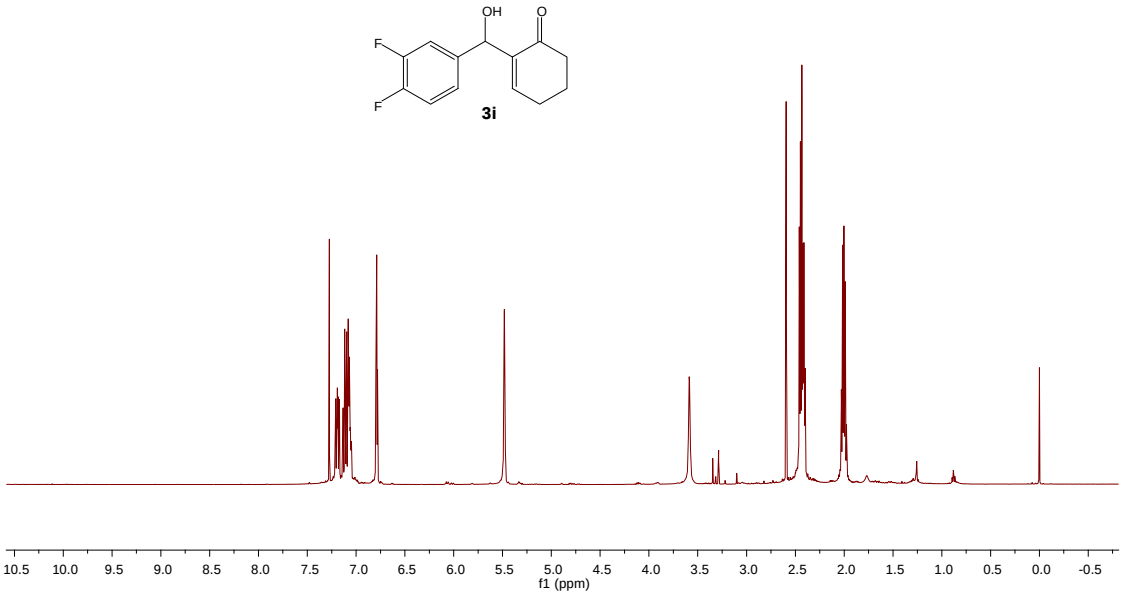
3h

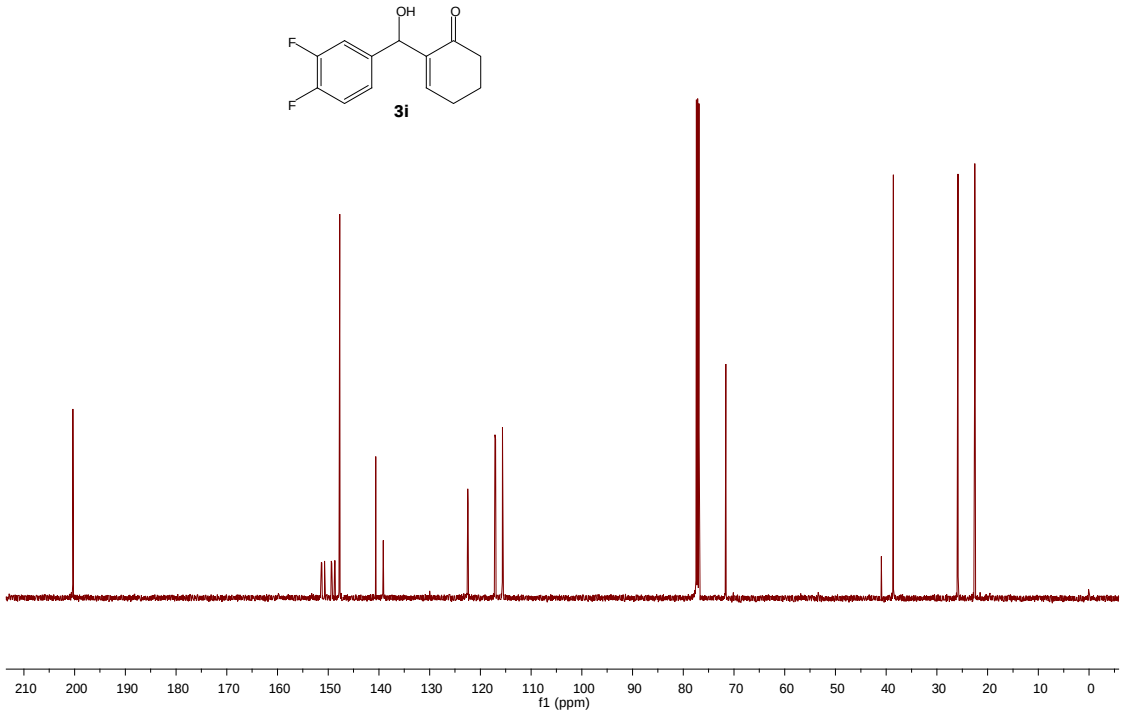

10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

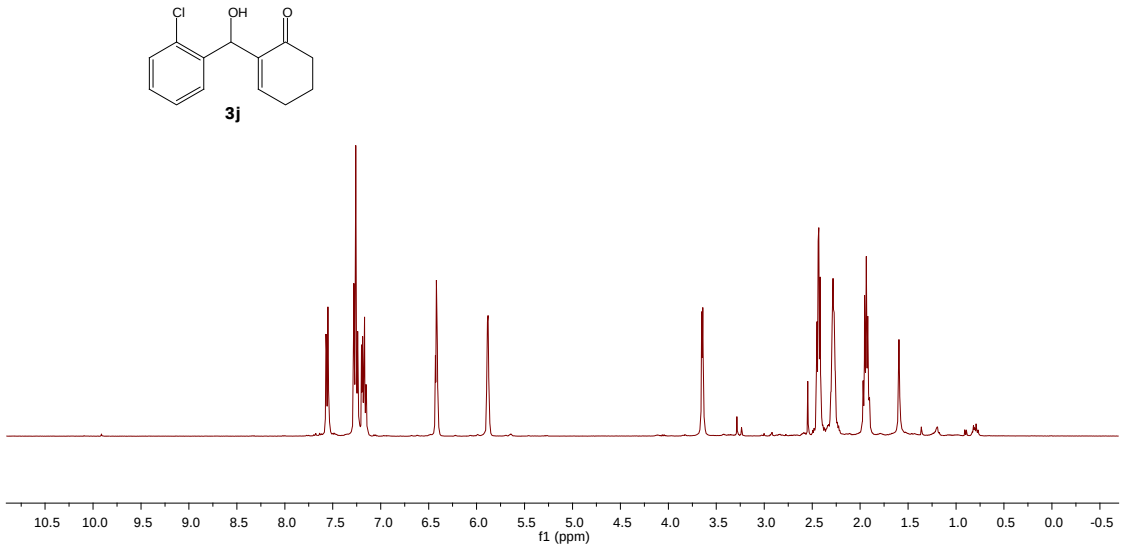

OH
O
F

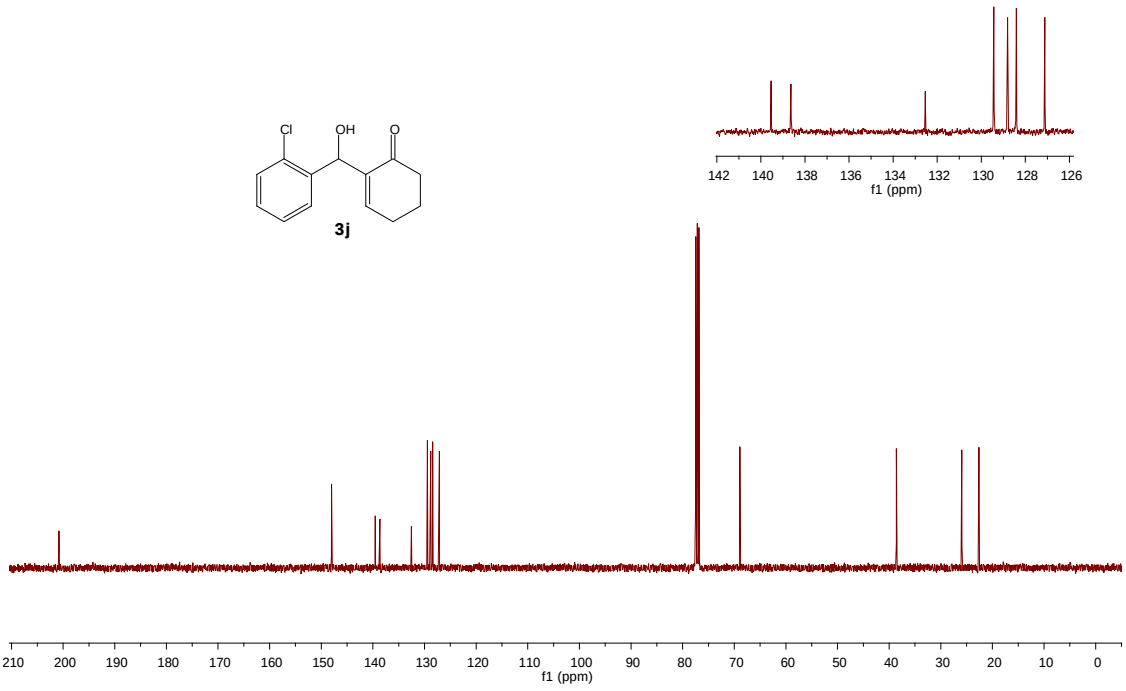
3h

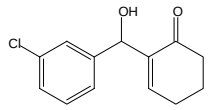

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)


OH
O
F
F
3i
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

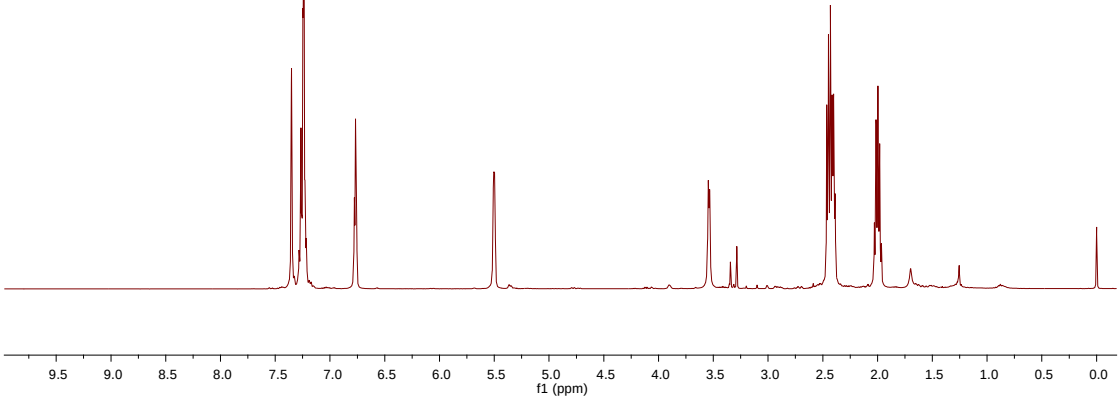

OH
O
F
F
3i
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

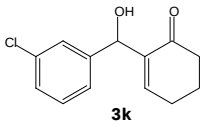

Cl
OH
O
3j
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)


Cl
OH
O
3j
142 140 138 136 134 132 130 128 126
f1 (ppm)
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

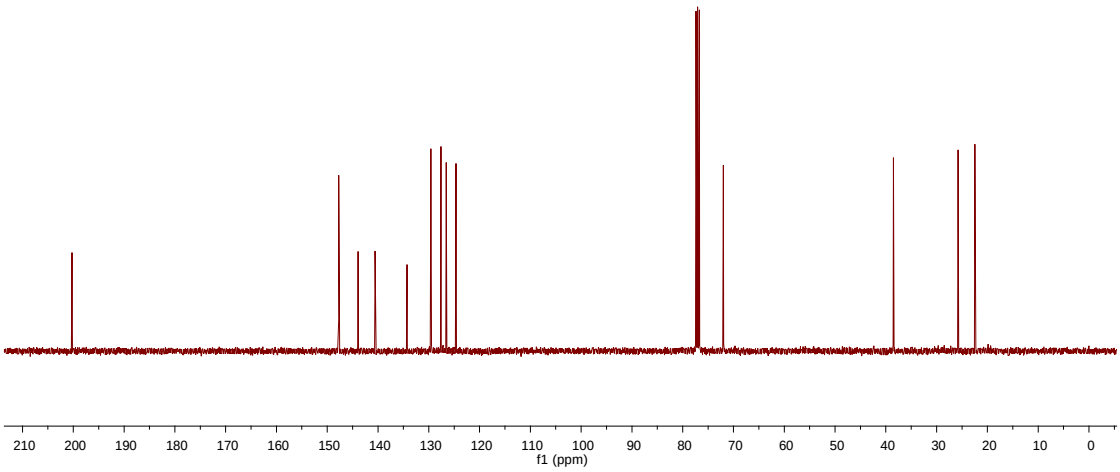

OH
O
Cl

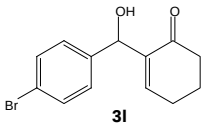
3k


9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

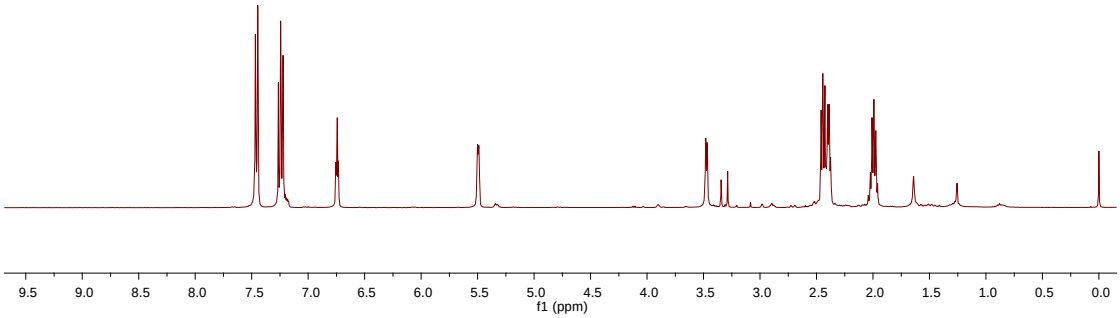

OH
O
Cl

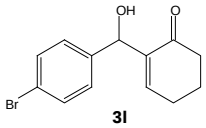
3k


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

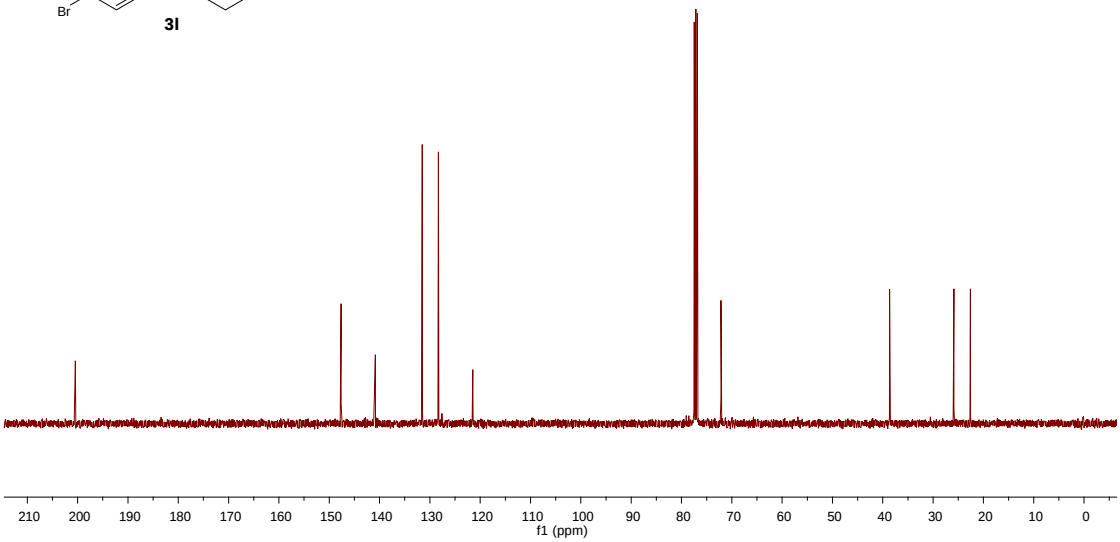

OH
O
Br

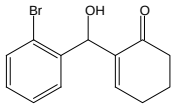
3l


9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

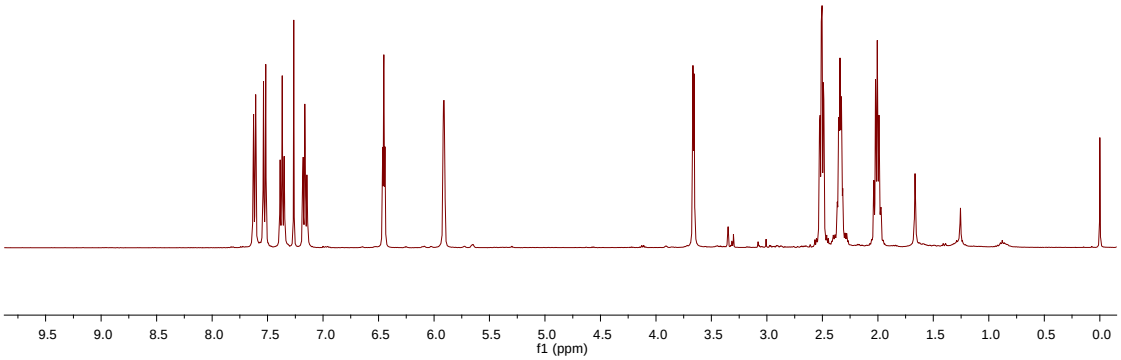

OH
O
Br

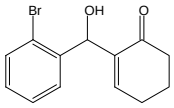
3l


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

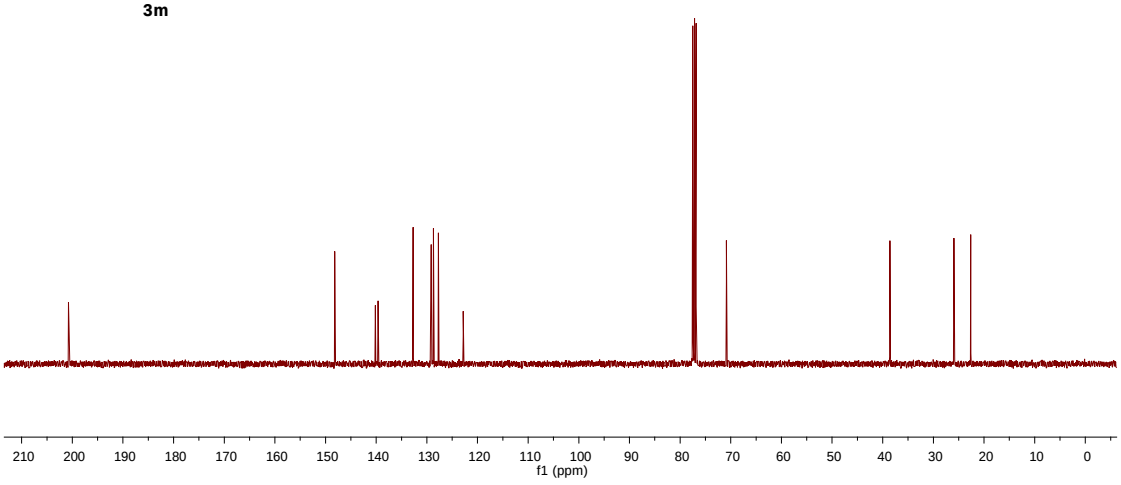

Br
OH
O

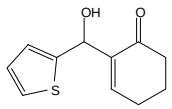
3m


9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

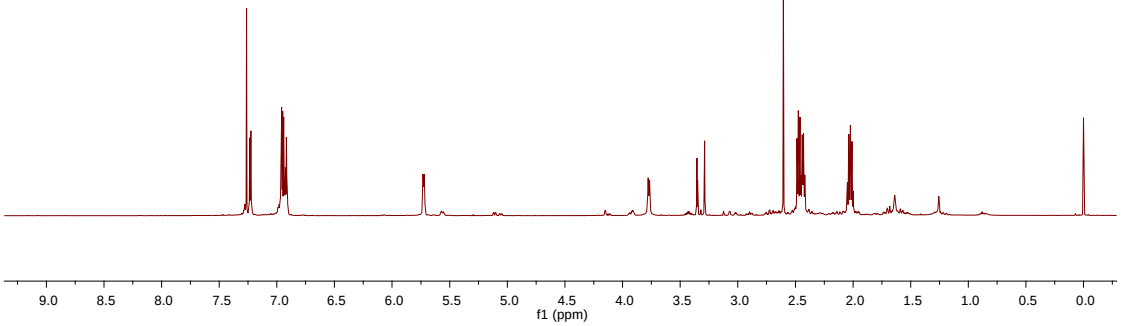

Br
OH
O

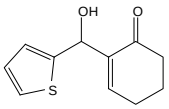
3m


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)


OH
O
S

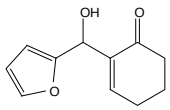
3n


9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

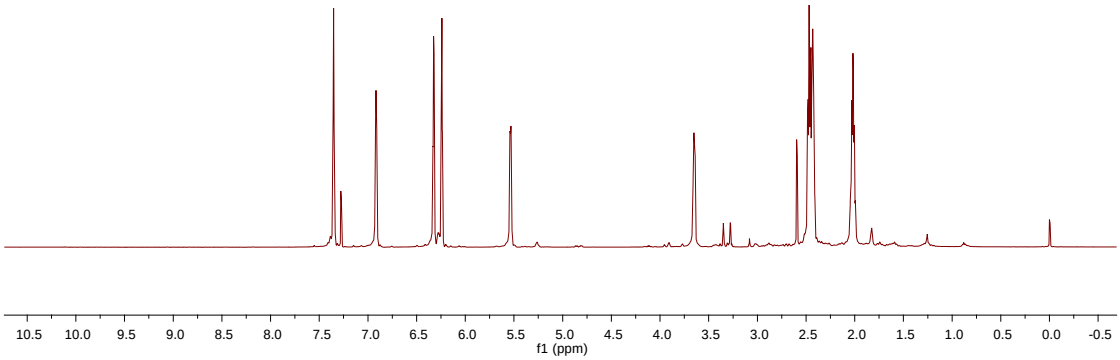

OH
O
S

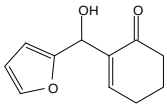
3n

210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

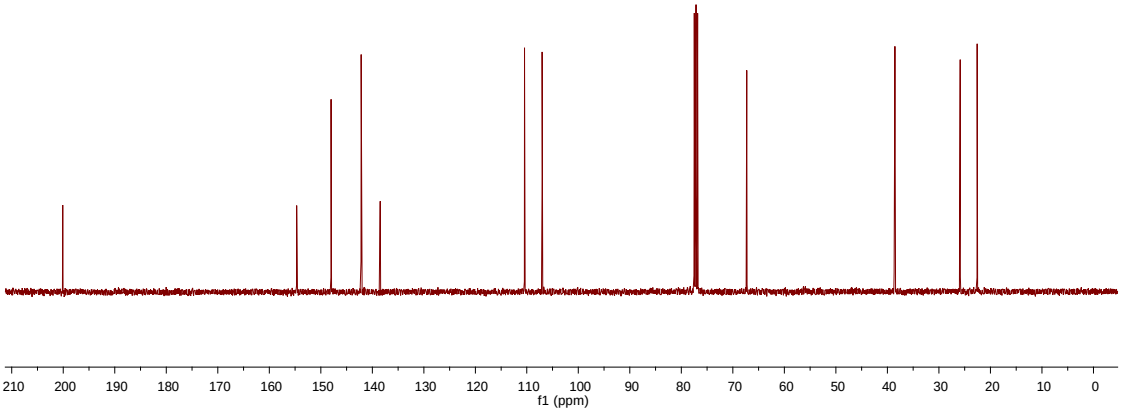

OH
O
O

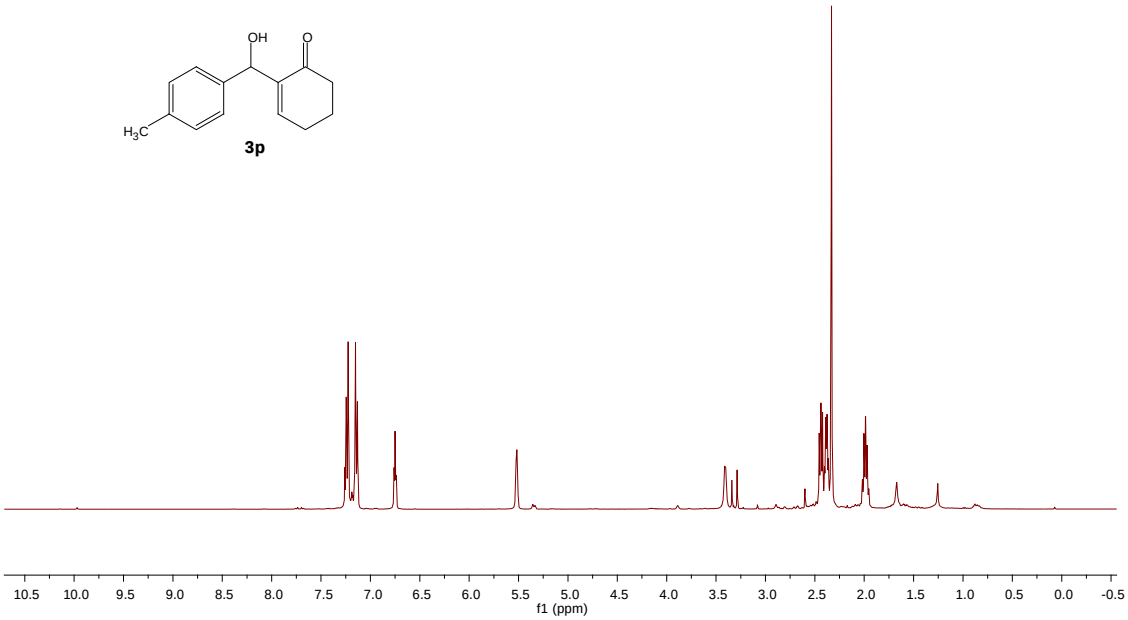
3o

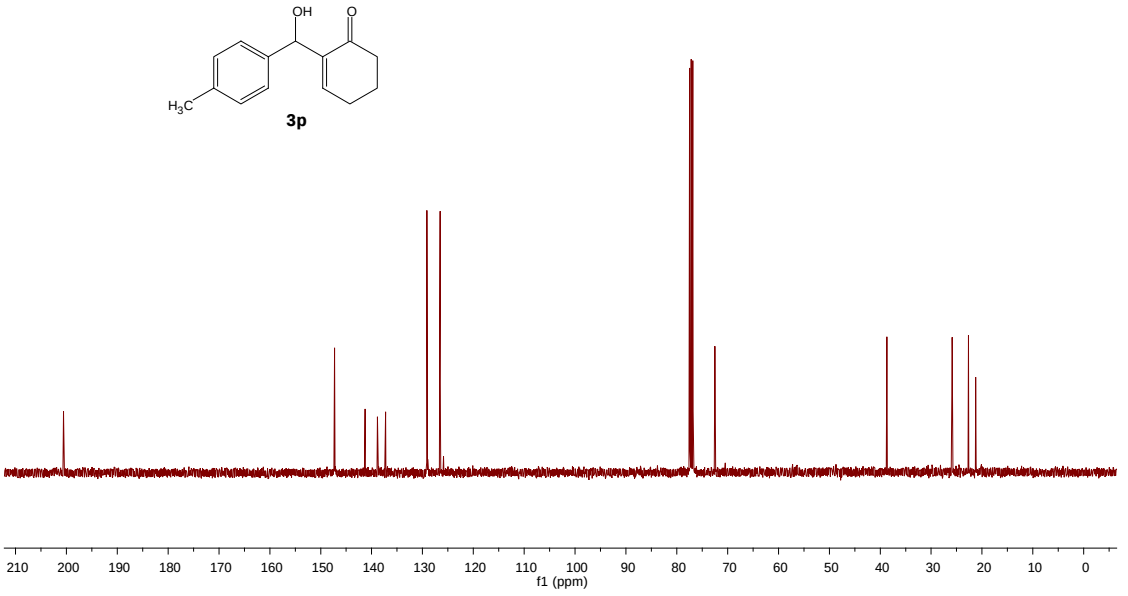

10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

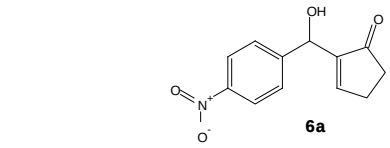

OH
O
O

3o

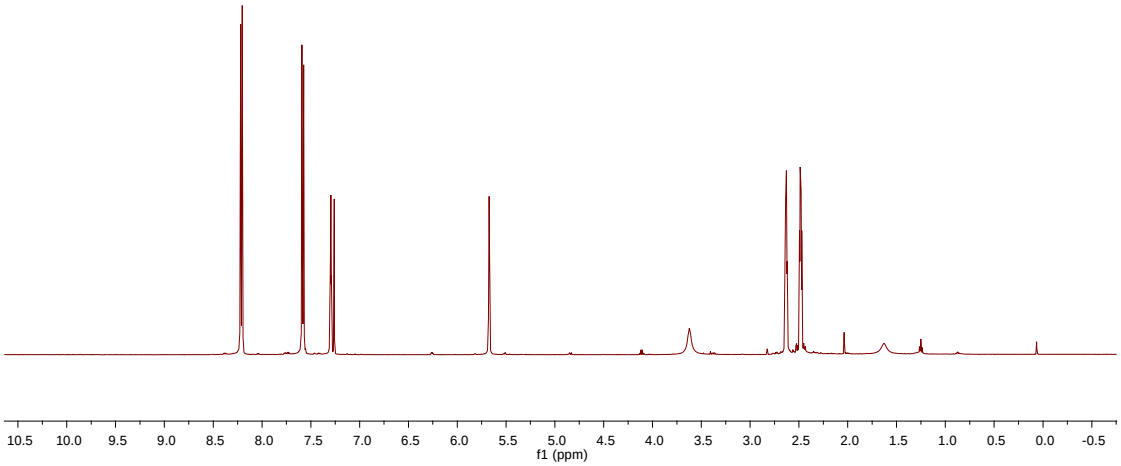

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

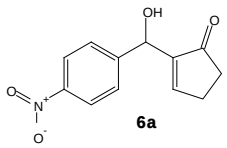

OH
O
H3C
3p
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)


OH
O
H3C
3p
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)


OH
O
O
N+
O-

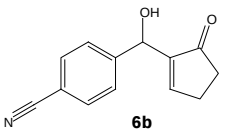
6a


10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

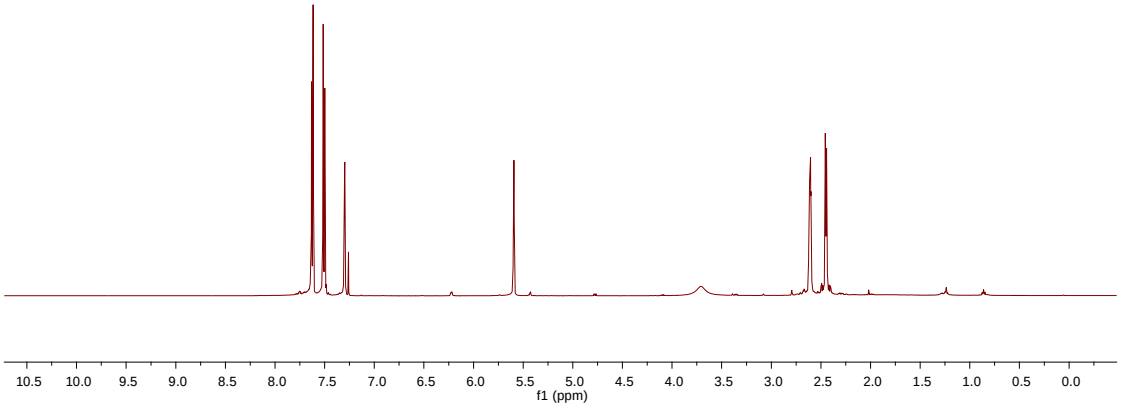

OH
O
O
N+
O-

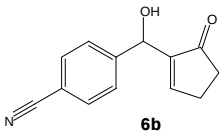
6a

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)


OH
O
N

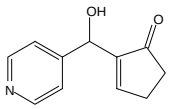
6b


10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

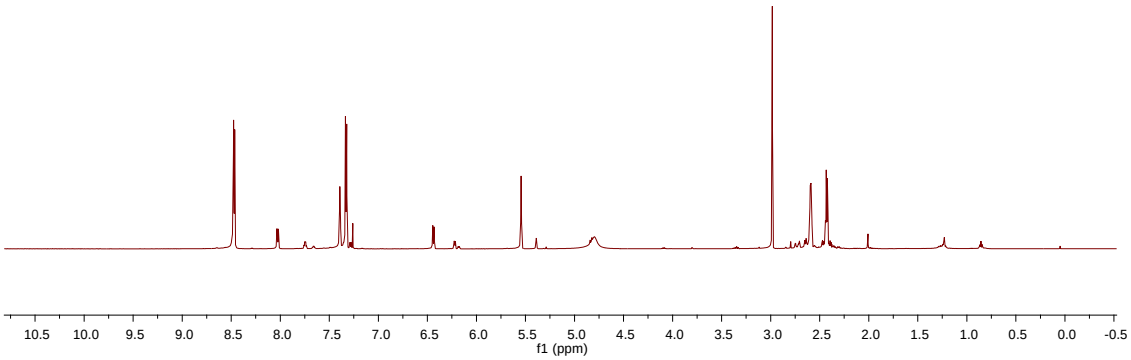

OH
O
N

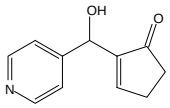
6b

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

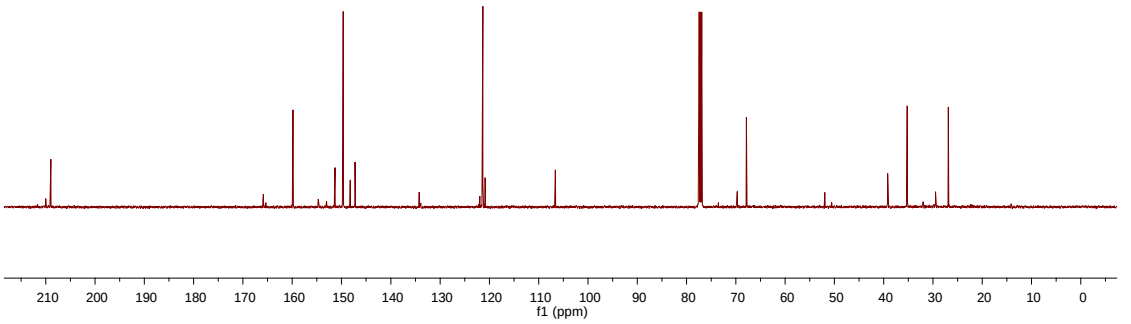

OH
O
N

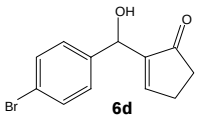
6c


10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

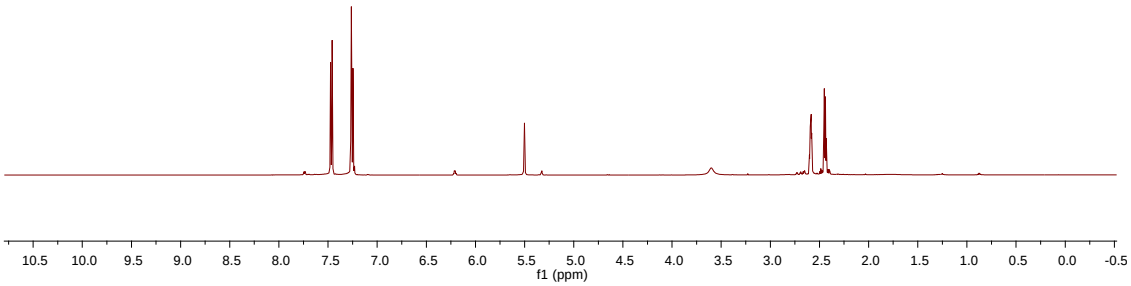

OH
O
N

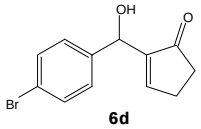
6c


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

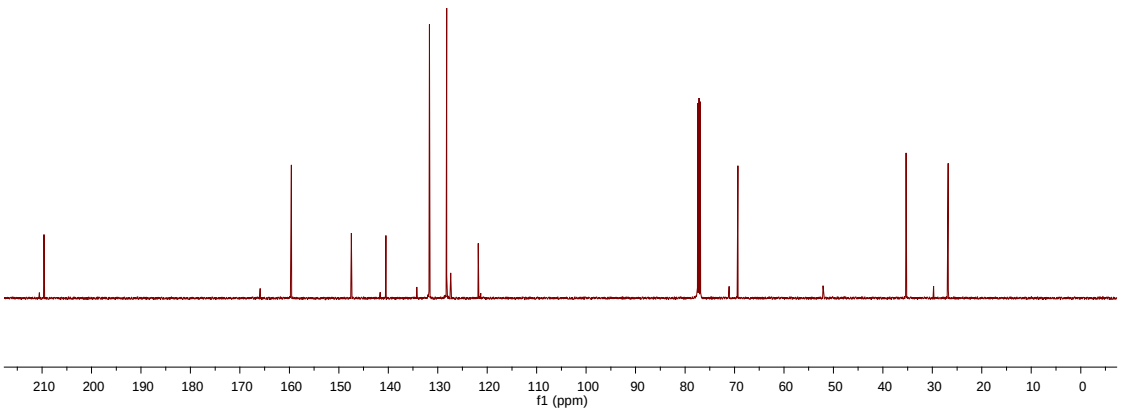

OH
O
Br

6d


10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)


OH
O
Br

6d


210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

References

[1] J. R. Lakowicz, in *Principles of Fluorescence Spectroscopy*, 2nd ed., Plenum, New York, 1999.
[2] R. Diamond, *J. Mol. Biol*. 1974, **82**, 371-374.
[3] G. M. Morris, D. S. Goodesell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Bellew, A. J. Olson, *J. Comput. Chem.* 1999, **19**, 1639-1662.
[4] N. K. Das, N. Ghosh, A. P. Kale, R. Mondal, U. Anand, S. Ghosh, V. K. Tiwari, *J. Phys. Chem. B* 2014, **118**, 7267-7276.
[5] W. L. De Lano in *The PyMOL Molecular Graphics System*; De Lano Scientific: San Carlos, CA, 2002.
[6] S. Bjelic, L. G. Nivon, N. Celebi-Olcum, G. Kiss, C. F. Rosewall, H. M. Lovick, E. L. Ingalls, J. L. Gallaher, J. Seetharaman, S. Lew, G. T. Montelione, J. F. Hunt, F. E. Michael, K. N. Houk, D. Baker, *ACS Chem. Biol.* 2013, **8**, 749-757.