Electronic Supplementary Information for

Azirine Ligation: Fast and Selective Protein Conjugation via Photoinduced Azirine-Alkene Cycloaddition

Reyna K. V. Lim and Qing Lin*

Department of Chemistry, State University of New York at Buffalo, Buffalo, New York 14260

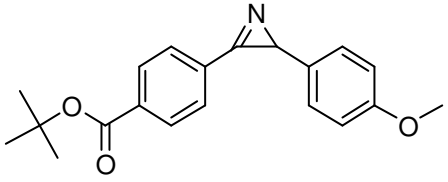
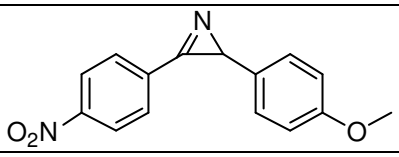
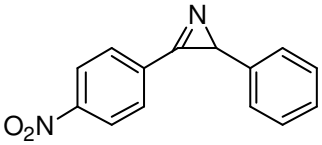
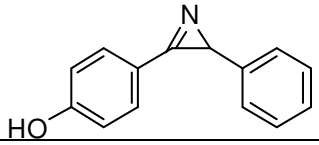
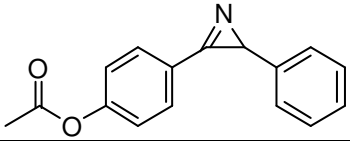
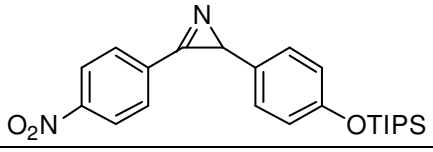
Materials and General Procedures

Unless otherwise stated, all chemicals and solvents were obtained from commercial sources and used without further purification. mPEG-SCM (MW ≈ 5000 Da) was obtained from Creative PEGworks. Chicken egg white lysozyme was obtained from Sigma. Analytical thin layer chromatography (TLC) was performed on Whatman AL SIL G/V254 flexible plates. Flash chromatography was performed with SiliCycle silica gel 60 Å (40-63 μm). All photoinduced cycloaddition reactions were carried out under ambient condition using oven-dried quartz test tubes with a magnetic stirrer. ^{1}H-NMR spectra were recorded with Inova-500 NMR spectrometer and chemical shifts were reported in ppm using either TMS or deuterated solvents as internal standards (TMS, 0.00; $CDCl_3$, 7.26; DMSO-d_6, 2.50). Multiplicity was reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad. ^{13}C-NMR spectra were recorded with Inova-300 spectrometer and chemical shifts were reported in ppm using $CDCl_3$ (δ 77.0) and DMSO-d_6 (δ 39.5) as internal standards. Kinetic study was performed using Keystone scientific BetaBasic-18 column (4.6 × 250 mm). Flow rate was set at 1.0 mL/min and UV detection wavelength was set at 254 nm. Electrospray LC-MS analysis was performed using a Finnigan LCQ Advantage IonTrap mass spectrometry coupled with a Surveyor HPLC system. Protein liquid chromatography was performed using a Phenomenex Jupiter 5 μm C4 300 Å reversed phase column (50 × 2.00 mm) with a flow rate of 250 μL/min. A linear gradient of 5-60% B over 30 min was applied for all runs in which solvent A was 0.1% aqueous formic acid and solvent B was 100% acetonitrile containing 0.1% formic acid. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of proteins was performed on XCell SureLock Mini-Cell apparatus with NuPAGE 12% Bis-Tris gels (Invitrogen). BenchMark Prestained Protein Ladder was loaded to one lane of the gel for estimation of apparent molecular weights.

Visualization of protein bands was achieved by staining the gel with the SimplyBlue SafeStain solution (Invitrogen).

Synthesis of Diarylazirines

Table S1. Various Azirines with their Isolated Yields.[a]

Azirine	Yield (%)
1 	21%
2 	12%
3 	3%[b]
4 	22%
5 	22%[c]
6 	13%

[a] Only the kinetic azirine products were isolated. [b] A large amount of vinyl azide was isolated. [c] Derived from acetylation of **4**.

General Procedure for the Heck Reaction: To a *para*-substituted styrene derivative in an oven-dried two-necked round bottom flask was added 1.2 equiv of aryl halide and 5 mol% $Pd(OAc)_2$. The mixture was dissolved in triethanolamine and stirred at 100 °C under argon overnight. After 5-min cooling, the reaction mixture was extracted with ethyl acetate. The combined organic layer was washed with brine, dried over sodium sulfate, concentrated under

reduced pressure, and purified *via* silica gel flash chromatography using hexane-ethyl acetate as the eluents to afford the stilbene product as a solid.

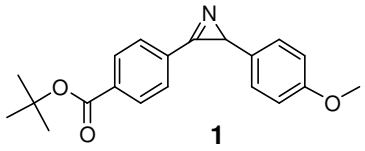
General Procedure for Iodoazidation: The stilbene analog was dissolved in acetonitrile (THF was used as co-solvent if the stilbene analog showed poor solubility in pure acetonitrile) along with 2.5 equiv of sodium azide in a two-necked round bottom flask placed in an ice bath. To the solution was added dropwise 1.2 equiv of iodine monochloride. The resulting mixture was stirred under argon overnight. The reaction was then quenched by addition of water and extracted 3× with diethyl ether (or ethyl acetate depending on solubility, 50 mL each time). The combined organic layer was washed with 150 mL of freshly prepared 5% sodium thiosulfate (pale yellow color turned colorless after washing), washed with brine, dried over anhydrous sodium sulfate, and concentrated under reduced pressure. The crude product was purified *via* silica gel flash chromatography using hexane-ethyl acetate as the eluents to afford the iodo azide as orange oil.

General Procedure for Vinyl Azide Formation: The iodoazide derivative was dissolved in diethyl ether followed by addition of 1.5 equiv potassium *tert*-butoxide. The reaction mixture was cooled in an ice bath under argon and left stirring overnight. The reaction was then quenched by addition of water and extracted with 3 × 50 mL diethyl ether (or ethyl acetate). The combined organic layer was washed with brine, dried over sodium sulfate, and concentrated under reduced pressure. The crude product was purified *via* silica gel flash chromatography using hexane-ethyl acetate as the eluents to afford the desired product as brown oil that decomposed slowly to give rise to azirine. During the column purification, kinetic and small amount of thermodynamic azirines were also isolated.

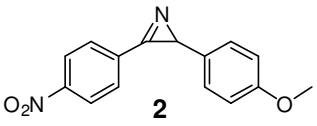
General Procedure for Azirine Formation: The vinyl azide was either refluxed in hexane or allowed to stand at room temperature for a few days before purification by silica gel column using hexane-ethyl acetate as the eluents. The desired azirines were obtained as solids.

General Procedure for Photoinduced Cycloaddition: 5 mg of azirine was dissolved in a 2 mL solvent (benzene, ethyl acetate, ethanol, ethanol/water (1:1) mixture) in a quartz test tube followed by addition of 5 equiv of suitable dipolarophile. The reaction mixture was irradiated at 302 nm using a handheld UV lamp (UVP, UVM-57, 302 nm, 115 V, 0.16 Amps) for 2 h. The

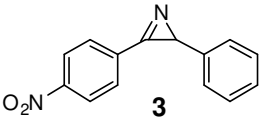
solvent and excess dipolarophile were evaporated under reduced pressure, and the residue was directly analyzed by ^{1}H-NMR.



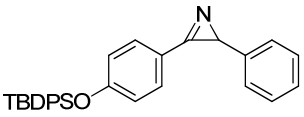
4-[3-(4-Methoxy-phenyl)-3*H*-azirin-2-yl]-benzoic acid *tert*-butyl ester: ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.89 (d, *J* = 8.0 Hz, 2H), 7.82 (d, *J* = 8.5 Hz, 2H), 7.17 (d, *J* = 8.5 Hz, 2H), 7.04 (d, *J* = 8.5 Hz, 2H), 3.89 (s, 3H), 3.29 (s, 1H), 1.57 (s, 3H); ^{13}C-NMR ($CDCl_3$, 75.4 MHz) δ 166.46, 163.72, 161.34, 146.68, 132.02, 129.50, 129.02, 125.79, 115.76, 114.87, 60.81, 55.56, 33.81, 14.31: HR-MS (EI) calcd for $C_{20}H_{21}NO_3$ 323.1516 [M^+], found 323.1523.



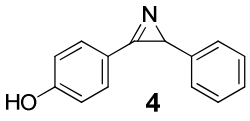
2-(4-Methoxy-phenyl)-3-(4-nitro-phenyl)-2*H*-azirine: ^{1}H-NMR (500 MHz, $CDCl_3$) δ 8.14 (d, *J* = 9.0 Hz, 2H), 7.83 (d, *J* = 9.0 Hz, 2H), 7.28 (d, *J* = 9.0 Hz, 2H), 7.07 (d, *J* = 8.5 Hz, 2H), 3.90 (s, 3H), 3.33 (s, 1H); ^{13}C-NMR (75.4 MHz, $CDCl_3$) δ 164.02, 160.68, 149.38, 146.88, 132.21, 126.45, 123.59, 115.03, 77.20, 55.63, 33.45; HR-MS (EI) calcd for $C_{15}H_{12}N_2O_3$ 268.0842 [M^+], found 268.0844.



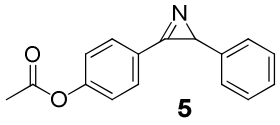
3-(4-Nitro-phenyl)-2-phenyl-2*H*-azirine: ^{1}H-NMR (500 MHz, $CDCl_3$) δ 8.16 (d, *J* = 9.0 Hz, 2H), 7.89 (d, *J* = 8.5 Hz, 2H), 7.66 (m, 1H), 7.59 (m, 2H), 7.30 (d, *J* = 9.0 Hz, 2H), 3.40 (s, 1H); ^{13}C-NMR (75.4 MHz, $CDCl_3$) δ 162.36, 148.85, 133.85, 130.15, 129.50, 126.53, 123.65, 122.96, 77.20, 33.82; HR-MS (ESI) calcd for $C_{14}H_{11}N_2O_2$ 239.0821 [M+ H^+], found 239.0825.



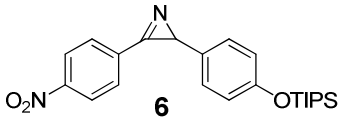
3-[4-(*tert*-Butyl-diphenyl-silanyloxy)-phenyl]-2-phenyl-2*H*-azirine: ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.71 (d, *J* = 7.0 Hz, 4H), 7.64 (d, *J* = 8.5 Hz, 2H), 7.42-7.44 (m, 2H), 7.36-7.39 (m, 4H), 7.24 (m, 2H), 7.19 (m, 1H), 7.10 (d, *J* = 7.5 Hz, 2H), 6.89 (d, *J* = 8.0 Hz, 2H), 3.18 (s, 1H), 1.11 (s, 9H); ^{13}C-NMR (75.4 MHz, $CDCl_3$) δ 162.06, 159.92, 141.17, 135.36, 131.92, 131.59, 130.21, 128.14, 127.95, 126.80, 125.95, 120.60, 116.73, 33.996, 26.35, 19.42.



4-(3-Phenyl-3*H*-azirin-2-yl)-phenol: ^{1}H-NMR (500 MHz, DMSO-d_6) δ 10.55 (s, 1H), 7.71 (d, J = 8.5 Hz, 2H), 7.27 (m, 2H), 7.20 (m, 1H), 7.07 (d, J = 7.0 Hz, 2H), 6.98 (d, J = 8.5 Hz, 2H), 3.23 (s, 1H); ^{13}C-NMR (DMSO-d_6, 125.7 MHz) δ 162.19, 160.84, 141.50, 131.86, 128.20, 126.65, 125.69, 116.52, 113.82, 32.62; LR-MS (ESI) calculated for $C_{14}H_{11}NO$ 209.08 [M^+], found 419.2 [2M + H^+].

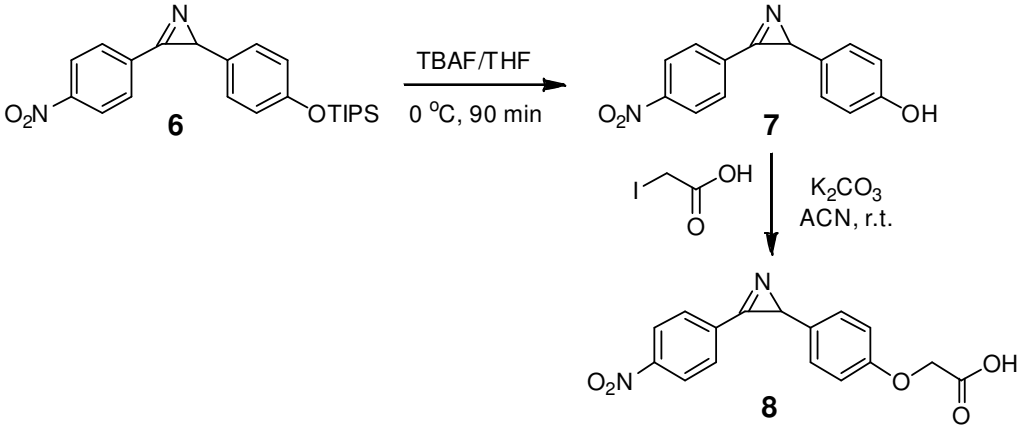


Acetic acid 4-(3-phenyl-2*H*-azirin-2-yl)-phenyl ester: ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.94 (d, J = 8.5 Hz, 2H), 7.23-7.31 (m, 5H), 7.14 (d, J = 7.0 Hz, 2H), 3.33 (s, 1H), 2.35 (s, 3H); ^{13}C NMR ($CDCl_3$, 75.4 MHz) δ 168.81, 162.74, 154.29, 140.58, 131.26, 128.29, 126.09, 122.72, 121.72, 77.21, 34.58, 21.12; LR-MS (ESI) calculated for $C_{16}H_{13}NO_2$ 251.09 [M^+], found 503.2 [2M + H^+].



3-(4-Nitro-phenyl)-2-(4-triisopropylsilanyloxy-phenyl)-2*H*-azirine: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.15 (d, J = 8.5 Hz, 2H), 7.76 (d, J = 8.5 Hz, 2H), 7.29 (d, J = 9.0 Hz, 2H), 7.03 (d, J = 9.0 Hz, 2H), 3.31 (s, 1H), 1.29-1.32 (m, 3H), 1.11 (d, J = 7.0 Hz, 18H).

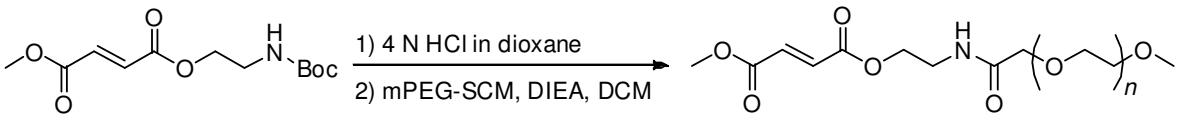
Synthesis of 4-[3-(4-Nitro-phenyl)-2*H*-azirin-2-yl]-phenoxy-acetic acid (8):



To a solution of azirine **6** (0.252 g, 0.61 mmol) in 7 mL THF was added 0.675 mL of tetrabutyl-ammonium fluoride in THF (1 M, 0.675 mmol). The reaction mixture was stirred over ice bath for 90 min. The solution was diluted with 20 mL DCM before quenching with 10 mL water. The solution was extracted with DCM (2×) and the organic layer was washed with brine, dried with Na_2SO_4, and concentrated under reduced pressure. The crude product was purified by silica gel

flash chromatography using ethyl acetate-hexane as eluents to give azirine **7** as a solid (0.146 g, 94% yield): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 8.15 (d, *J* = 9.0 Hz, 2H), 7.79 (d, *J* = 8.5 Hz, 2H), 7.27 (m, 3H), 7.01 (d, *J* = 8.5 Hz, 2H), 3.33 (s, 1H); ^{13}C-NMR ($CDCl_3$, 75.4 MHz) δ 160.69, 160.50, 149.25, 132.51, 126.45, 123.62, 116.63, 115.31, 33.49. To a solution of azirine **7** (13 mg, 0.05 mmol) in 2 mL acetonitrile was added potassium carbonate powder (208 mg, 1.5 mmol) with vigorous stirring, followed by slow addition of iodoacetic acid (52 mg, 0.28 mmol). The reaction mixture was stirred at room temperature for 18 h. To the mixture was then added a solution of 1 N NaOH until the mixture turned clear. The solution was titrated to pH 3-4 with 1 N HCl and then extracted with ethyl acetate (3×). The organic layer was dried with Na_2SO_4 and then evaporated under reduced pressure. The red-brown solid was recrystallized with ethyl acetate-hexane to afford a black solid (5.6 mg, 35 % yield): ^{1}H-NMR (500 MHz, CD_3OD) δ 8.17 (d, *J* = 9.0 Hz, 2H), 7.84 (d, *J* = 8.5 Hz, 2H), 7.35 (d, *J* = 8.5 Hz, 2H), 7.16 (d, *J* = 9.0 Hz, 2H), 4.59 (s, 2H), 3.42 (s, 1H); LR-MS (ESI negative ion mode) calculated for $C_{16}H_{11}N_2O_5$ 311.1 $[M-H]^-$, found 311.7.

Synthesis of mPEG-fumarate 9:



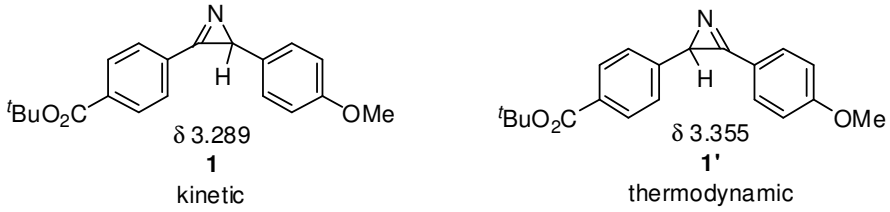
9

Boc-aminoethyl methyl fumarate[S1] (30 mg, 0.11 mmol) was treated with 1 mL anhydrous hydrochloride (4 N in dioxane) until the starting material disappeared as monitored by TLC. The solvent was removed under vacuum and triturated 3× with anhydrous DCM to remove the remaining HCl. The residue was redissolved in 2 mL DCM and diisopropylethylamine (70 µL, 0.4 mmol) was added followed by dropwise addition of the solution of mPEG-SCM (MW ≈ 5000 Da; 200 mg, 0.04 mmol) in 1 mL DCM. The resulting mixture was stirred overnight. To the cloudy solution was added pivaloyl chloride (50 µL, 0.4 mmol) and the solution became clear after 30 min stirring. Afterwards, 10 mL dry ethyl ether was added and the mixture was cooled over dry ice. The white precipitate that appeared at the bottom of the reaction vessel were collected by filtration, washed extensively with dry ethyl ether, and dried in vacuum to give the titled compound as a white powder (135 mg, 24% conversion): ^{1}H-NMR (500 MHz, $CDCl_3$) δ

6.88 (s, 2H), 4.31 (t, *J* = 5.5 Hz, 2H), 4.29 (s, 2H), 3.82 (s, 3H), 3.65 (s, 439H), 3.46 (m, 2H), 3.38 (s, 3H).

Azirine Structural Assignment

The azirine structures were assigned based on: 1) chemical shifts of the azirine ring proton in ^{1}H-NMR, and 2) heteronuclear multiple bond coupling (HMBC) 2D-NMR. For unsymmetrical azirines **1-3** and **5-8**, the isomer with a more downfield chemical shift for the azirine ring proton was assigned to the azirine structure with the C=N bond in conjugation with the more electron-rich phenyl ring. For example, azirine **1** (kinetic product, major product isolated) showed the ring proton chemical shift of 3.289 ppm *vs.* 3.355 ppm for the thermodynamic azirine **1'** (minor product isolated).



For azirine **4**, only one product isomer was isolated. To determine its structure, we carried out the HMBC experiment because HMBC detects the long-range couplings between a proton and a carbon that are 2 and 3 (sometimes 4) bonds away with great sensitivity. The full HMBC spectrum of **4** is shown in Figure S1. The spectrum was divided into four quadrants: yellow, blue, green, and red and the important cross peaks were encircled. There are two possible azirine structures, **4** and **4'**. Based on the vital cross-peaks between azirine ring hydrogen *d* and the neighboring phenyl ring carbons *7* and *8*, between phenyl ring hydrogen *a* and azirine *sp3* carbon (*6* for **4** or *5* for **4'**), and between phenyl hydrogens *e* and *f* and azirine *sp2* carbon (*5* for **4** or *6* for **4'**), we assigned the azirine structure to **4** as it explains better of the detected cross-peaks (Figure S1).

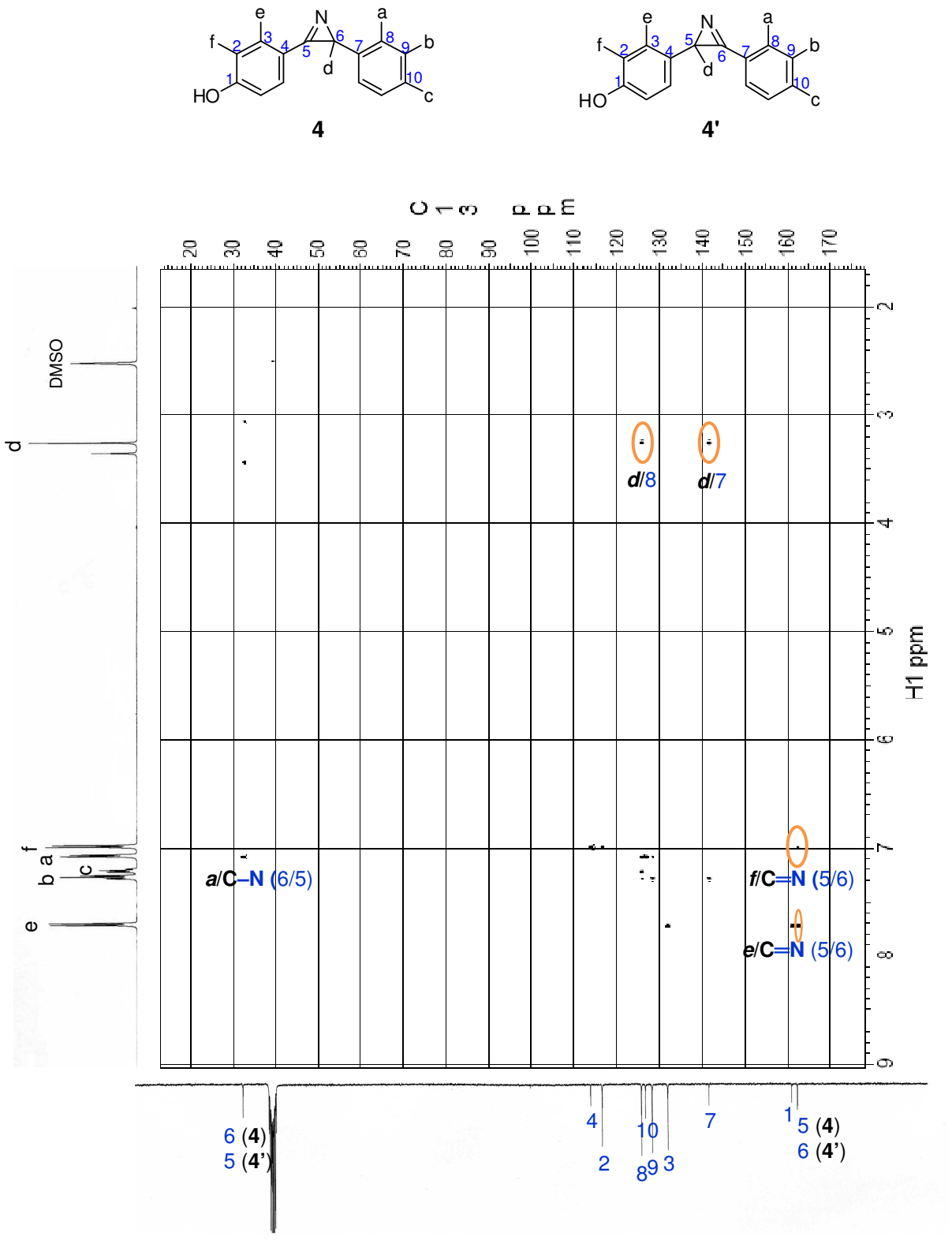


Figure S1. Heteronuclear multiple bond coupling (HMBC) spectrum of azirine **4**: The vital $^{1}H/^{13}C$ cross-peaks were marked on the spectrum.

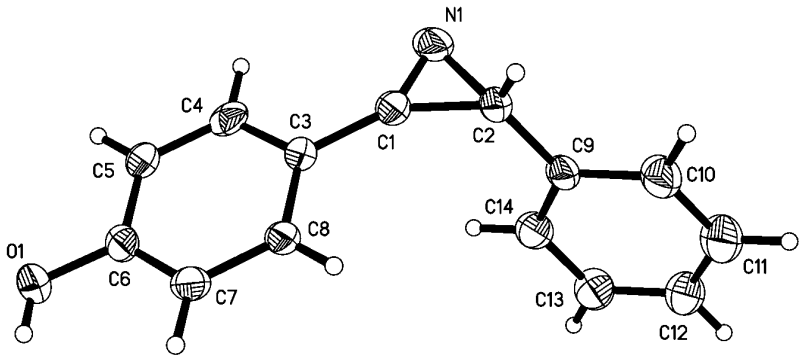


Figure S2. ORTEP diagrams of azirine **4** shown at 50% probability.

Table S2. Photoactivated 1,3-Dipolar Cycloaddition of **1** and Methyl Methacrylate in Various Solvents[a]

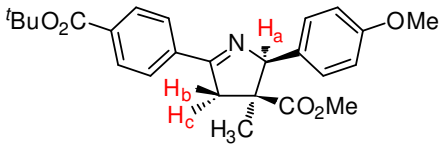
entry	*solvent*	*% conversion*[b]	*endo:exo*[b]
1	Benzene	100%	60:40
2	EtOAc	87%	55:45
3	EtOH	100%	50:50
4	EtOH/H_2O (1:1)	56%	50:50

[a] Reactions were conducted by irradiating 5 mg of **1** and 5 equiv of methyl methacrylate in 2 mL of solvent in quartz test tubes. [b] Based on ^{1}H-NMR of the crude product after removal of the solvent and excess reagent.

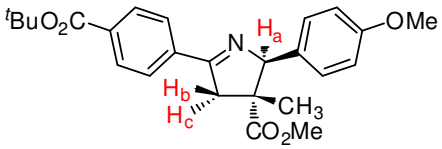
Pyrroline Cycloadduct Structural Assignment

^{1}H-NMR data support the formation of cycloadducts **1a** (endo) and **1a'** (exo) as Δ^1-pyrrolines. Moreover, formation of the other regioisomers, **1b** and **1b'**, was disregarded because ^{1}H-NMR spectra for the highlighted pyrroline ring protons for **1b** and **1b'** would be different from those for **1a** and **1a'**. It is expected that three protons in structures **1b** and **1b'** would couple to each

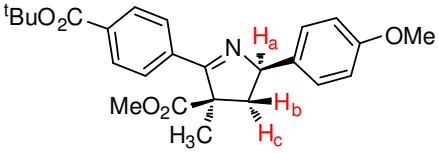
other to show the doublets of doublet split patterns; however, ^{1}H-NMR spectrum of the crude showed H_a as singlet and H_b and H_c as AB quartet, consistent with structures for **1a** and **1a'**.



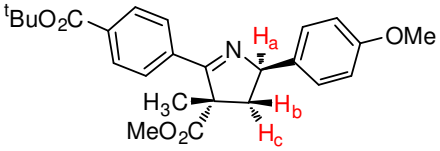
1a (endo)



1a' (exo)

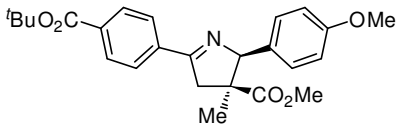


1b (endo)



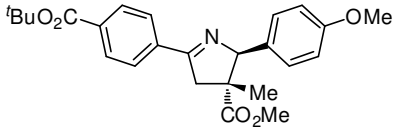
1b' (exo)

In accordance with Padwa and co-workers' previous reports,[S2] the assignments between **1a** and **1a'** were made based on the chemical shifts of the ring-methyl group. In the NMR spectra, there was a strong upfield shift for the pyrroline ring methyl group situating *syn* to the phenyl ring substituent. This strong upfield shift can be attributed to shielding effect by the neighboring phenyl ring π electrons as would be the case for **1a'**. Likewise, an upfield shift of the ester methoxy group would be expected for **1a** because methoxy group is *syn* to the phenyl ring. The assignments of the pyrroline cycloadducts with dimethyl fumarate were based on the same considerations.



1a (endo)

(2*S*,3*S*)-Methyl 5-(4-(*tert*-butoxycarbonyl)phenyl)-2-(4-methoxyphenyl)-3-methyl-3,4-dihydro-2*H*-pyrrole-3-carboxylate: ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.88-7.91 (m, 4H), 7.18 (d, J = 8.0 Hz, 2H), 6.97 (d, J = 8.5 Hz 2H), 5.26 (s, 1H), 3.88 (s, 3H); 3.86-3.90 (m, 1H), 3.17 (s, 3H), 2.90 (m, 1 H), 1.60 (s, 3H), 1.57 (s, 9H).



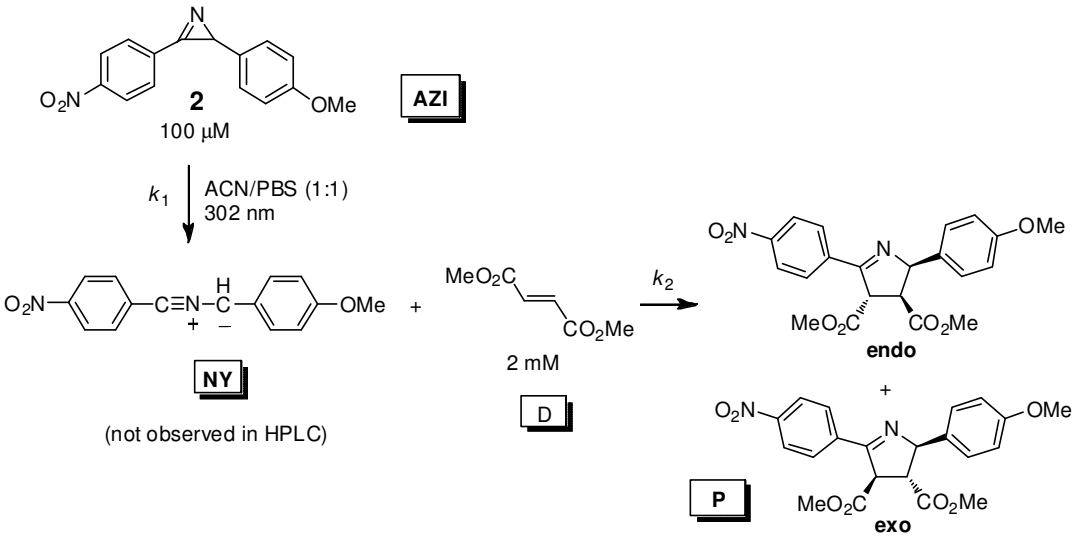
1a' (exo)

(2*S*,3*R*)-methyl 5-(4-(tert-butoxycarbonyl)phenyl)-2-(4-methoxyphenyl)-3-methyl-3,4-dihydro-2*H*-pyrrole-3-carboxylate: ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.96 (d, J = 8.5 Hz, 2H), 7.88 (d, J = 9.0 Hz, 2H), 7.34 (d, J = 9.0 Hz, 2H), 6.96 (d, J = 8.5 Hz,

2H), 5.70 (s, 1H), 3.87 (s, 3H), 3.82 (s, 3H), 3.70 (m, 1H), 3.02 (m, 1H), 1.60 (s, 9H), 0.80 (s, 3H).

Kinetic Study of the Azirine-Alkene Cycloaddition Reaction

A solution containing 100 μM azirine **2** and 2 mM dimethyl fumarate in 2 mL acetonitrile/PBS buffer (1:1) was prepared. An 150-μL solution in a quartz test tube was irradiated with a 302-nm UV lamp for 15 s (30 s for the next injection, and so on) and an aliquot of 100 μL was injected immediately into HPLC. The cycloadduct formation was monitored by UV absorbance at 254 nm. The amounts of the cycloadducts were calculated based on a calibration curve generated by injecting a reference compound (end/exo mixture) with the known concentrations run under the identical conditions.



The rate of azirine **2** (AZI) decomposition is described as d[AZI]/dt = -k_1[AZI] (eq. 1). The nitrile ylide intermediate (NY) is at a rate of k_1[AZI] but decays in the presence of dipolarophile (D) to give rise to the cycloadducts (P) at a rate of k_2[NY][D] (eq. 3). In the presence of excess D, the decay of nitrile ylide follows the first-order kinetics according to equation: r = k_2'[NY], where k_2' = k_2[D].

$$d[AZI]/dt = -k_1[AZI] \quad (1)$$

$$d[NY]/dt = k_1[AZI] - k_2[NY][D] \quad (2)$$

$$d[P]/dt = k_2[NY][D] \quad (3)$$

Unlike in flash photolysis[S3] where the nitrile ylide species was observed and constantly replenished, in the HPLC analysis nitrile ylide intermediate was not observed (see HPLC traces in Figure S2). This led us to assume that whenever the NY species is formed, it decays rapidly into P in the presence of excess dimethyl fumarate, and thus $k_2' >> k_1$. It follows that the rate of change of concentration of NY intermediate would be negligibly small, which leads to the steady-state approximation, $d[NY]/dt \approx 0$. Therefore, setting equation (2) to zero yields

$$k_1[AZI] = k_2[NY][D] \qquad (4)$$

By combining eq. 4 with equations 1 and 3, we derive

$$-d[AZI]/dt = d[P]/dt \qquad (5)$$

Eq. 5 suggests that the disappearance of azirine **2** corresponds to the quantitative cycloadducts formation. Indeed, this was confirmed by HPLC chromatogram showing the formation of the cycloadducts as the sole products along with the concurrent disappearance of the azirine starting materials. No side products, such as water-quenching and dimerization products, were observed. By plotting the starting materials and cycloadducts *vs.* time, the first-order rate constant for the azirine ring opening reaction (k_1) was determined to be 0.0379 s^{-1} and the pseudo-first-order rate constant (k_2') for the cycloadduct formation was 0.0373 s^{-1} (Figure S3).

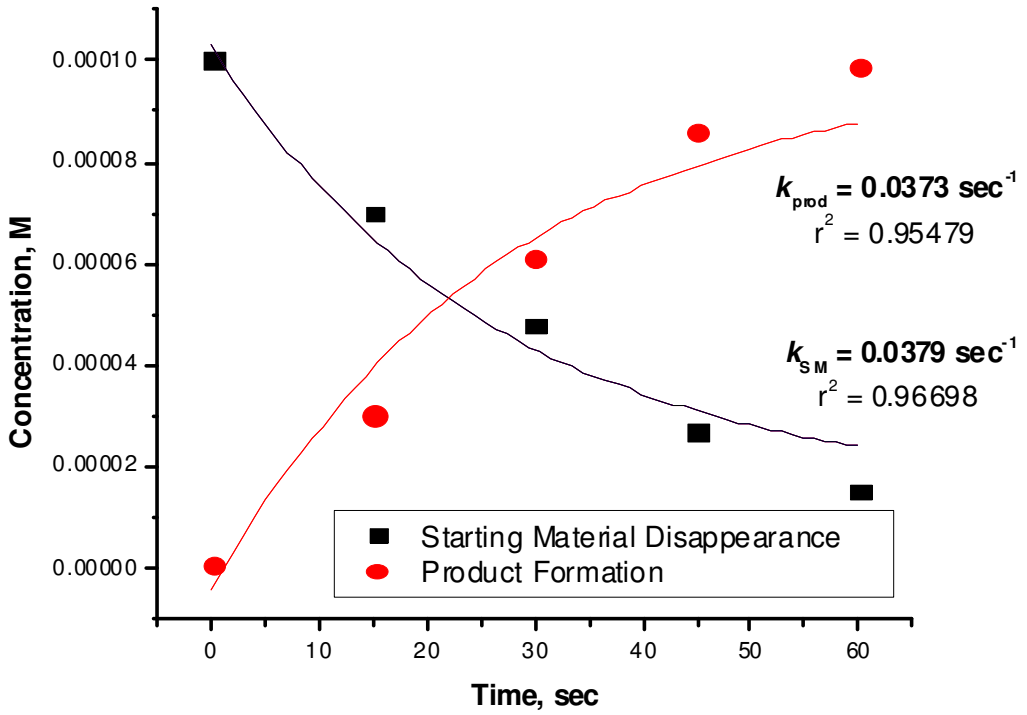


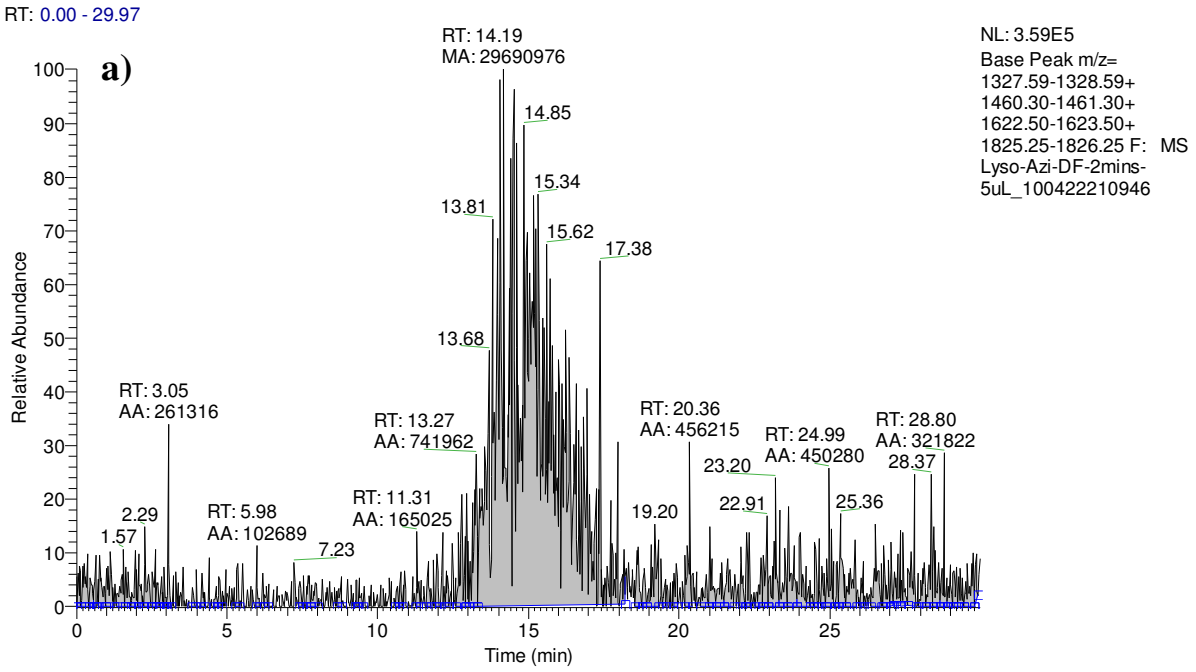
Figure S3. Plots of azirine disappearance (black square) and cycloadduct formation (red circle). The data were fitted to the exponential decay (or rise) and the first-order rate constants were derived and marked on the plots.

Protein Ligation Studies

Conjugating Azirine 8 to Chicken Lysozyme: To a 150-µL solution of azirine **8** (25 mM in DMF) was added *N*-hydroxysuccinimide (1.3 mg, 75 mM) and *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (1.1 mg, 38 mM). The reaction mixture was vortexed for 15 min to afford the activated azirine. Two 50-µL of this mixture were incubated with two separate 450-µL of egg white chicken lysozyme (125 µM in 10 mM NaH_2PO_4, 25 mM NaOAc, pH 8.5) with an additional 1 mg NHS to each to increase the solubility. Both reaction mixtures were tumbled for 3h at room temperature. Afterwards, excess small molecules were removed by using either ultracel-10K centrifugal filters (Millipore) or protein desalting spin columns (Pierce). The resulting solutions were used directly for the subsequent mass spectrometry and gel studies. The protein fractions were found to contain 44% of monoacylated lysozyme (Lyso-Azi), 22% diacylated lysozyme (Lyso-d-Azi), 7% triacylated lysozyme (Lyso-t-Azi), and 26% unreacted

lysozyme (Lyso) based on ion extractions in the LC-ESI/MS analysis. The identities of Lyso, Lyso-Azi, Lyso-d-Azi, and Lyso-t-Azi were confirmed by the deconvoluted intact masses: Lyso, calcd 14304 Da, found 14304.7 ± 0.3 Da; Lyso-Azi, calcd 14598 Da, found 14598.6 ± 0.5 Da; Lyso-d-Azi, calcd 14892 Da, found 14892 ± 1.0 Da; Lyso-t-Azi, calcd 15186 Da, found 15187 ± 1.0 Da.

Photoinduced Azirine Ligation with Dimethyl Fumarate: To an 100-µL azirine-containing lysozyme solution (125 µM Lyso-Azi) in PBS buffer, pH 7.4, in a quartz test tube was added 5.0 µL dimethyl fumarate (final concentration ~ 9 mM). Under vigorous stirring, the mixture was irradiated with a handheld 302-nm UV lamp for 2 min and then analyzed by LC-MS. Pyrroline cycloadduct formation was confirmed by the appearance of the charged ladders, which after deconvolution corresponds to the intact mass of Lyso- Pyr (14742.2 ± 0.6 Da, expected 14742 Da). The conversion from Lyso-Azi to Lyso-Pyr was estimated to be ~80% based on the extracted mass areas in the LC-MS according to the equation, $MA_{Lyso\text{-}Pyr}/(MA_{Lyso\text{-}Pyr} + MA_{Lyso\text{-}Azi})$ (Figure S4). No obvious side products, such as Michael addition products, were detected in the mass spectrometry.



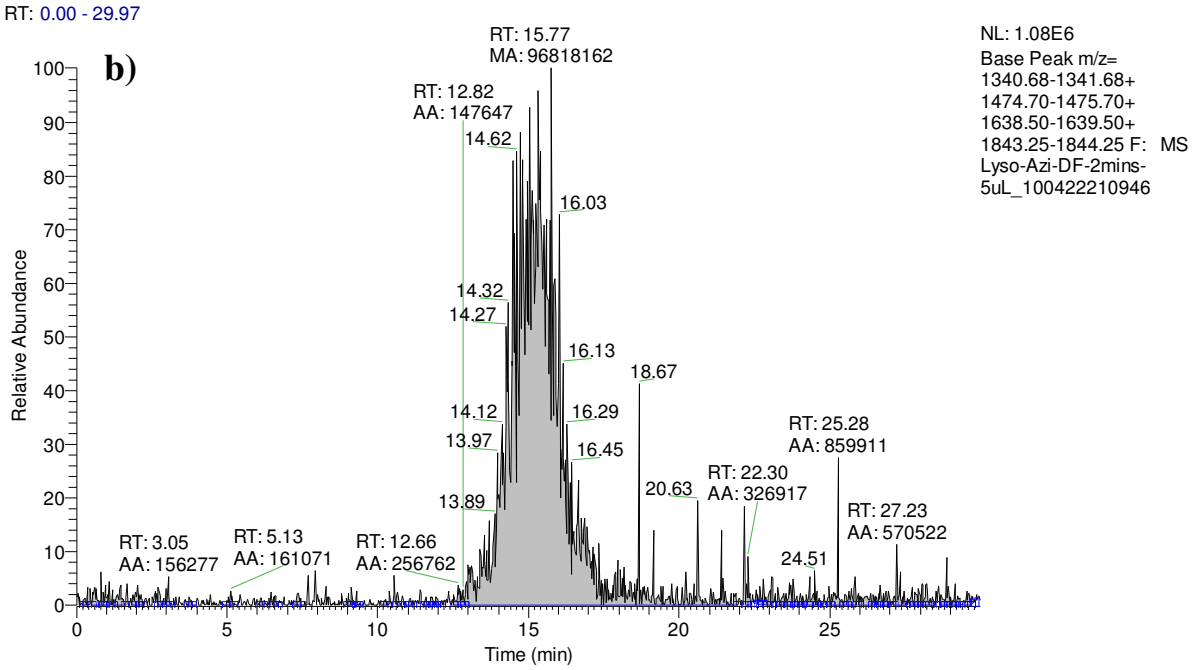


Figure S4. Ion extractions for Lyso-Azi (**a**) and Lyso-Pyr (**b**) with their corresponding mass areas marked on the top of the peaks.

Crystal Data Collection for Azirine 4: A crystal (0.18 x 0.16 x 0.06 mm^3) was placed onto the tip of a 0.1 mm diameter glass capillary tube or fiber and mounted on a Bruker SMART APEX II CCD Platform diffractometer for a data collection at 100.0(1) K. A preliminary set of cell constants and an orientation matrix were calculated from reflections harvested from three orthogonal wedges of reciprocal space. The full data collection was carried out using MoKα radiation (graphite monochromator) with a frame time of 90 sec and a detector distance of 4.02 cm. A randomly oriented region of reciprocal space was surveyed: six major sections of frames were collected with 0.50º steps in ω at six different ϕ settings and a detector position of -38º in 2θ. The intensity data were corrected for absorption. Final cell constants were calculated from the xyz centroids of 2465 strong reflections from the actual data collection after integration (see Table S2 for additional crystal and refinement information).

Crystal Structure Solution and Refinement: The structure was solved using SIR97 and refined using SHELXL-97. The space group $P2_1$ was determined based on systematic absences and intensity statistics. A direct-method solution was calculated which provided most non-hydrogen atoms from the E-map. Full-matrix least squares/difference Fourier cycles were performed which

located the remaining non-hydrogen atoms. All non-hydrogen atoms were refined with anisotropic displacement parameters. All hydrogen atoms were placed in ideal positions and refined as riding atoms with relative isotropic displacement parameters. Since the major enantiomer could not be determined by anomalous dispersion effects, it was assigned arbitrarily. All reflections, including Friedel opposites, were merged. The structure refined as a pseudo-merohedral twin, with a mass ratio of 53:47. The application of twin law, [1 0 0 / 0 1 0 / 0 0 1], a 180 degree rotation about direct lattice [1 0 0], reduced the $R1$ residual from 24.4 % to 5.9 %. The final full matrix least squares refinement converged to $R1 = 0.0591$ (F^2, $I > 2s(I)$) and $wR2 = 0.1686$ (F^2, all data). Unless noted otherwise all structural diagrams containing thermal displacement ellipsoids are drawn at the 50 % probability level.

Reference:

[S1] W. Song, Z. Yu, M. M. Madden, Q. Lin, *Mol. Biosyst.* **2010**, DOI: 10.1039/c003470c.

[S2] a) A. Padwa, J. Smolanoff, *J. Am. Chem. Soc.* **1971**, *93*, 548; b) A. Padwa, M. Dharan, J. Smolanoff, S. I. Wetmore, Jr., *J. Am. Chem. Soc.* **1973**, *95*, 1945.

[S3] N. J. Turro, D. A. Hrovat, I. R. Gould, A. Padwa, W. Dent, R. J. Rosenthal, *Angew. Chem. Int. Ed.* **1983**, *22*, 625.

Table S3. Crystal data and structure refinement for azirine **4**.

Identification code	azirine **4**	
Empirical formula	C14 H11 N O	
Formula weight	209.24	
Temperature	100.0(1) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	$P2_1$	
Unit cell dimensions	a = 5.4973(7) Å	$\alpha = 90°$
	b = 8.3407(11) Å	$\beta = 90.166(3)°$
	c = 11.8458(16) Å	$\gamma = 90°$
Volume	543.14(12) $Å^3$	
Z	2	
Density (calculated)	1.279 Mg/m^3	
Absorption coefficient	0.081 mm^{-1}	
$F(000)$	220	
Crystal color, morphology	colorless, rod	
Crystal size	0.18 x 0.16 x 0.06 mm^3	
Theta range for data collection	1.72 to 27.48°	
Index ranges	$-7 \leq h \leq 7$, $-10 \leq k \leq 10$, $-15 \leq l \leq 15$	
Reflections collected	9617	
Independent reflections	1325 [R(int) = 0.1271]	
Observed reflections	1011	
Completeness to theta = 27.48°	100.0%	
Absorption correction	Multi-scan	
Max. and min. transmission	0.9952 and 0.9856	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	1325 / 1 / 147	
Goodness-of-fit on F^2	1.010	
Final R indices [I>2sigma(I)]	$R1$ = 0.0591, $wR2$ = 0.1534	
R indices (all data)	$R1$ = 0.0776, $wR2$ = 0.1686	
Largest diff. peak and hole	0.286 and -0.319 $e.Å^{-3}$	

Table S4. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for azirine **4**. U_{eq} is defined as one third of the trace of the orthogonalized U_{ij} tensor.

	x	y	z	U_{eq}
O1	-2976(11)	9318(5)	5881(4)	37(1)
N1	-90(13)	1765(6)	6627(4)	40(2)
C1	591(14)	3177(7)	6736(4)	33(2)
C2	2398(14)	2098(7)	7249(5)	30(1)
C3	-259(12)	4780(8)	6510(5)	27(1)
C4	-2471(12)	5024(7)	5939(5)	30(1)
C5	-3351(13)	6532(7)	5724(4)	30(1)
C6	-1982(13)	7855(7)	6097(4)	29(1)
C7	219(13)	7646(7)	6651(5)	30(1)
C8	1099(13)	6116(6)	6858(5)	29(1)
C9	2515(14)	1844(6)	8476(4)	29(1)
C10	4472(14)	971(8)	8945(6)	40(2)
C11	4645(18)	734(9)	10090(6)	48(2)
C12	2899(13)	1368(8)	10784(5)	39(2)
C13	995(13)	2240(9)	10340(5)	37(2)
C14	797(13)	2463(7)	9185(5)	34(2)

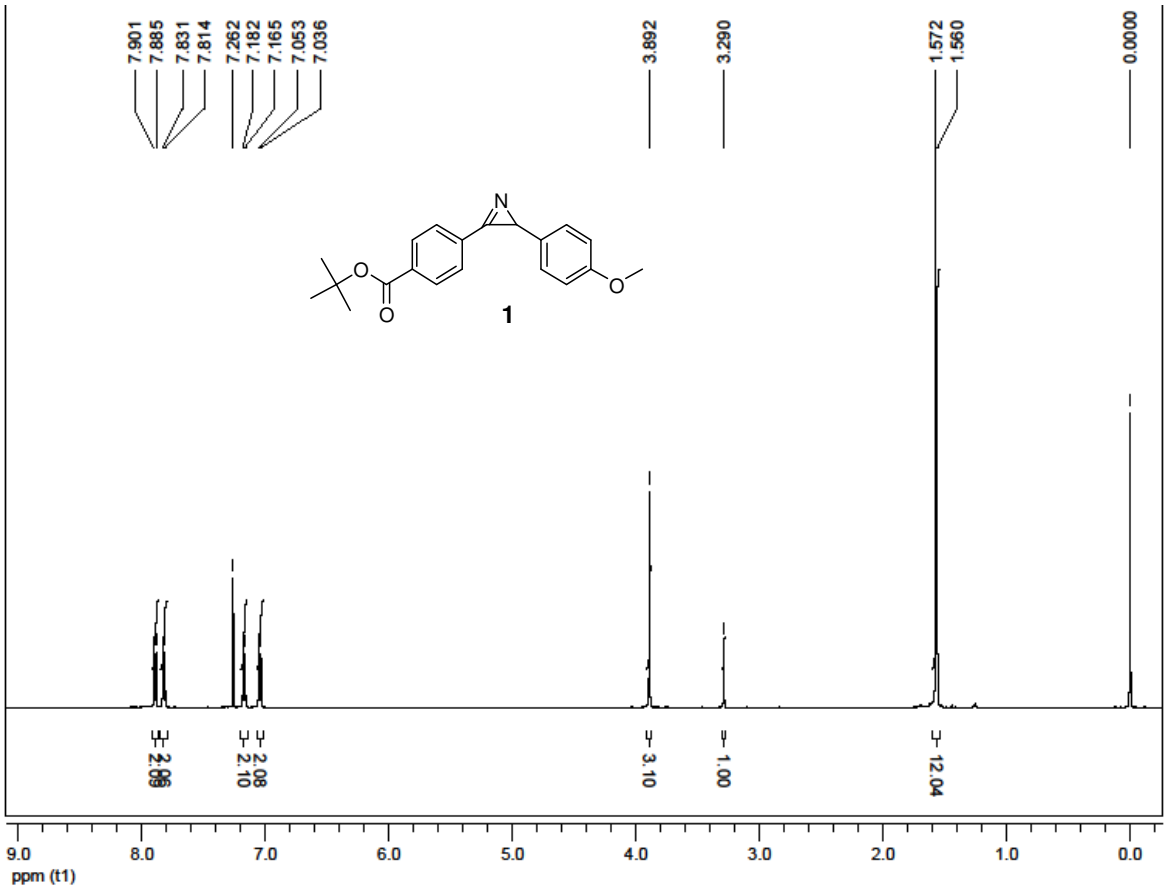
Table S5. Bond lengths [Å] and angles [°] for azirine **4**.

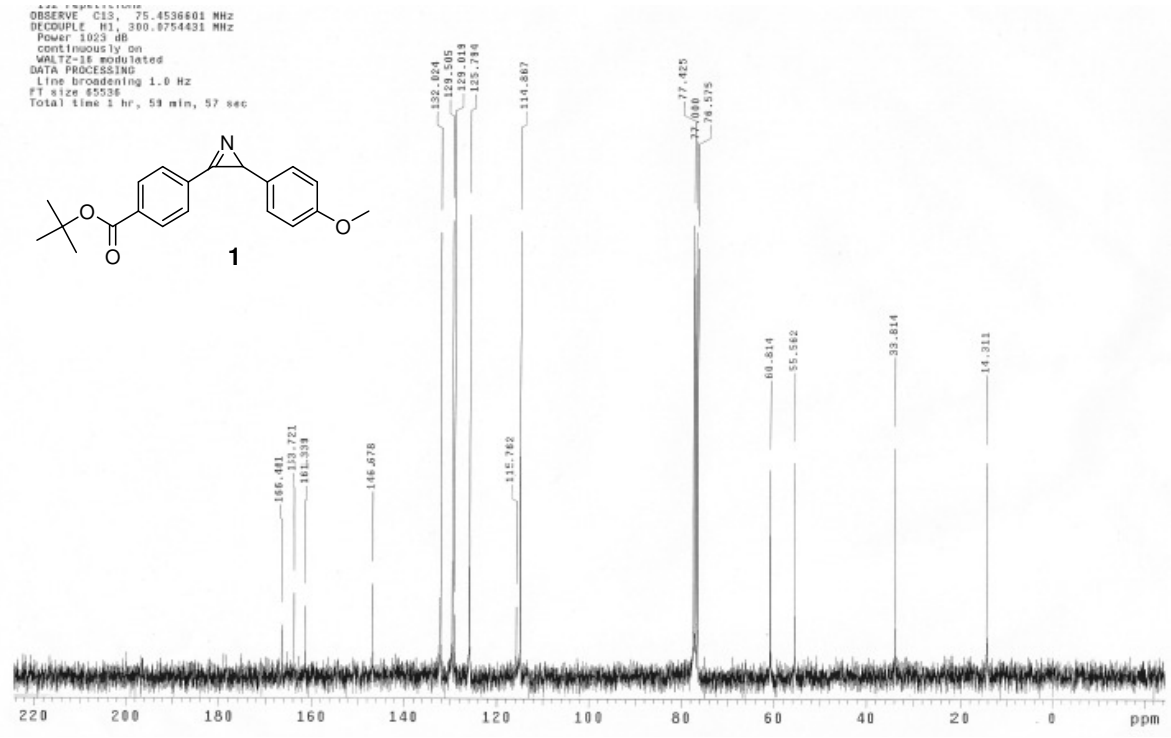
O(1)-C(6)	1.361(7)	C(1)-C(2)-C(9)	121.6(6)
O(1)-H(1)	0.8400	C(1)-C(2)-N(1)	48.0(4)
N(1)-C(1)	1.242(8)	C(9)-C(2)-N(1)	118.2(6)
N(1)-C(2)	1.576(10)	C(1)-C(2)-H(2)	117.0
C(1)-C(3)	1.442(9)	C(9)-C(2)-H(2)	117.0
C(1)-C(2)	1.470(9)	N(1)-C(2)-H(2)	117.0
C(2)-C(9)	1.470(7)	C(8)-C(3)-C(4)	119.1(6)
C(2)-H(2)	1.0000	C(8)-C(3)-C(1)	120.7(6)
C(3)-C(8)	1.402(9)	C(4)-C(3)-C(1)	120.3(6)
C(3)-C(4)	1.405(9)	C(5)-C(4)-C(3)	121.8(6)
C(4)-C(5)	1.370(9)	C(5)-C(4)-H(4)	119.1
C(4)-H(4)	0.9500	C(3)-C(4)-H(4)	119.1
C(5)-C(6)	1.407(9)	C(4)-C(5)-C(6)	118.3(6)
C(5)-H(5)	0.9500	C(4)-C(5)-H(5)	120.9
C(6)-C(7)	1.385(9)	C(6)-C(5)-H(5)	120.9
C(7)-C(8)	1.387(8)	O(1)-C(6)-C(7)	123.5(6)
C(7)-H(7)	0.9500	O(1)-C(6)-C(5)	115.5(6)
C(8)-H(8)	0.9500	C(7)-C(6)-C(5)	121.0(6)
C(9)-C(14)	1.367(8)	C(6)-C(7)-C(8)	120.2(6)
C(9)-C(10)	1.412(10)	C(6)-C(7)-H(7)	119.9
C(10)-C(11)	1.374(10)	C(8)-C(7)-H(7)	119.9
C(10)-H(10)	0.9500	C(7)-C(8)-C(3)	119.6(6)
C(11)-C(12)	1.371(10)	C(7)-C(8)-H(8)	120.2
C(11)-H(11)	0.9500	C(3)-C(8)-H(8)	120.2
C(12)-C(13)	1.378(10)	C(14)-C(9)-C(10)	118.7(5)
C(12)-H(12)	0.9500	C(14)-C(9)-C(2)	121.7(6)
C(13)-C(14)	1.386(9)	C(10)-C(9)-C(2)	119.7(6)
C(13)-H(13)	0.9500	C(11)-C(10)-C(9)	120.9(7)
C(14)-H(14)	0.9500	C(11)-C(10)-H(10)	119.5
C(6)-O(1)-H(1)	109.5	C(9)-C(10)-H(10)	119.5
C(1)-N(1)-C(2)	61.5(5)	C(12)-C(11)-C(10)	119.3(7)
N(1)-C(1)-C(3)	139.7(7)	C(12)-C(11)-H(11)	120.3
N(1)-C(1)-C(2)	70.5(5)	C(10)-C(11)-H(11)	120.3
C(3)-C(1)-C(2)	149.6(6)	C(11)-C(12)-C(13)	120.5(6)

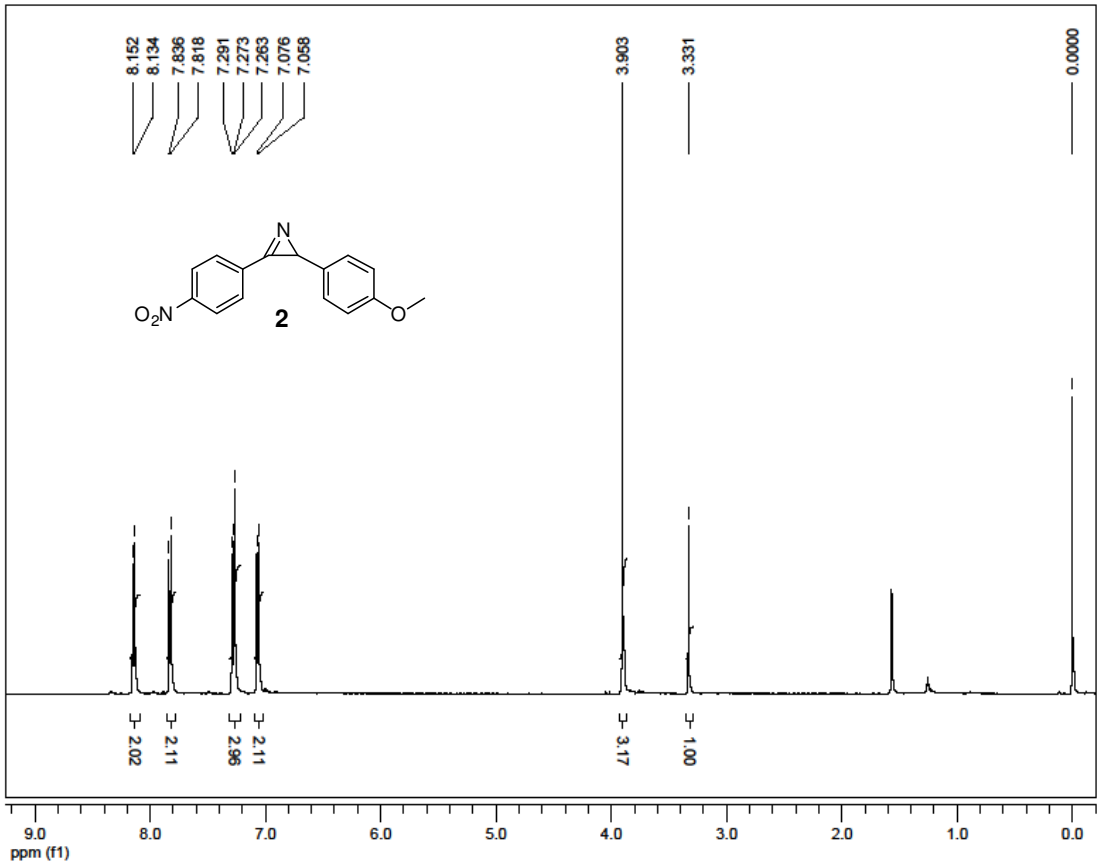
C(11)-C(12)-H(12)	119.8	C(14)-C(13)-H(13)	119.8
C(13)-C(12)-H(12)	119.8	C(9)-C(14)-C(13)	120.3(6)
C(12)-C(13)-C(14)	120.4(7)	C(9)-C(14)-H(14)	119.9
C(12)-C(13)-H(13)	119.8	C(13)-C(14)-H(14)	119.9

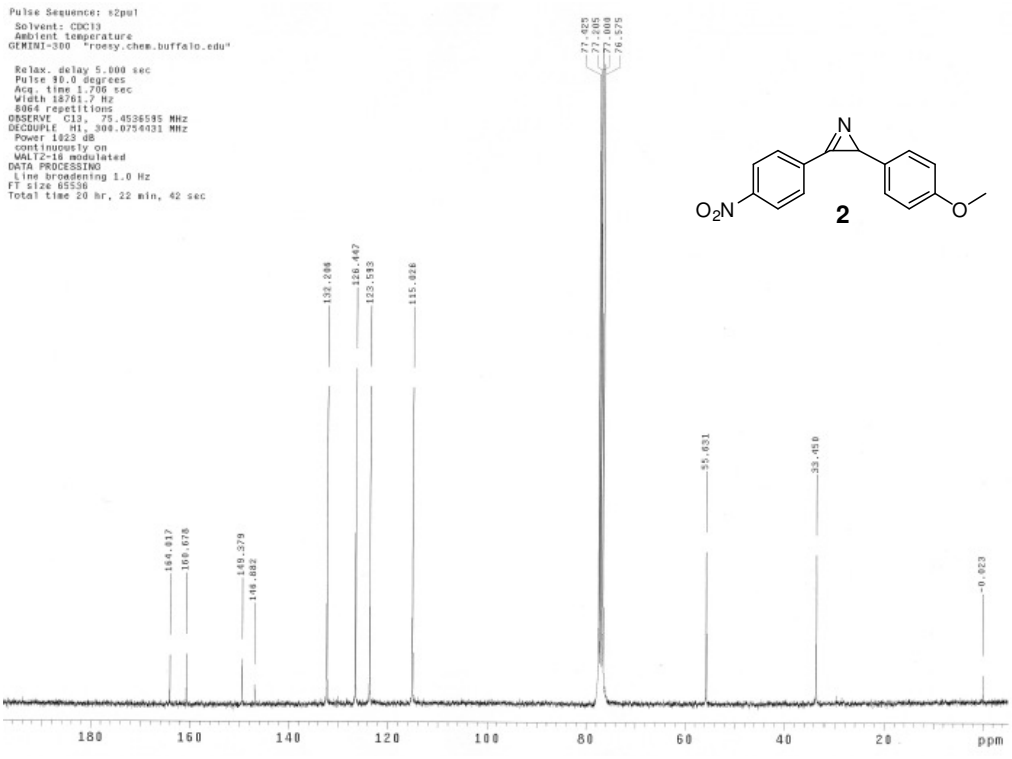
Table S6. Torsional angles [°] for azirine **4**.

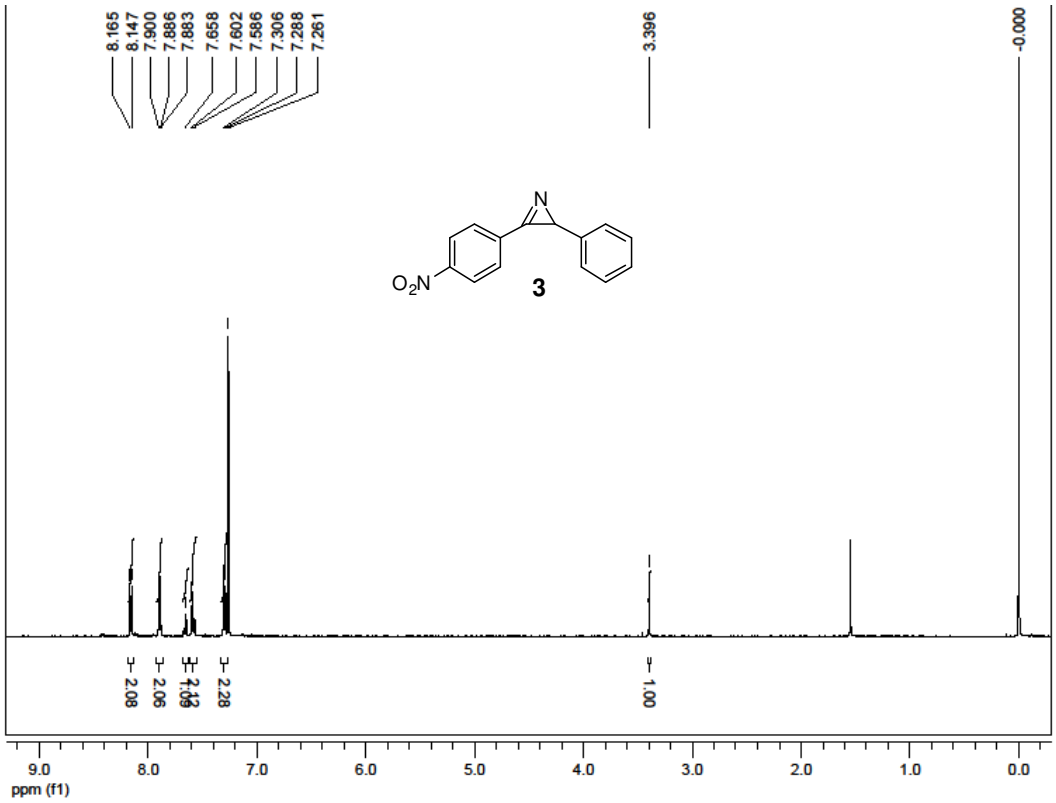
C2-N1-C1-C3	-175.8(10)	C10-C11-C12-C13	-0.3(12)
N1-C1-C2-C9	-101.0(7)	C11-C12-C13-C14	1.2(11)
C3-C1-C2-C9	73.6(13)	C10-C9-C14-C13	0.4(9)
C3-C1-C2-N1	174.7(13)	C2-C9-C14-C13	-178.2(7)
C1-N1-C2-C9	108.5(6)	C12-C13-C14-C9	-1.2(10)
N1-C1-C3-C8	175.4(8)		
C2-C1-C3-C8	3.2(14)		
N1-C1-C3-C4	-4.4(12)		
C2-C1-C3-C4	-176.7(9)		
C8-C3-C4-C5	-0.8(9)		
C1-C3-C4-C5	179.0(6)		
C3-C4-C5-C6	-0.3(8)		
C4-C5-C6-O1	-178.6(5)		
C4-C5-C6-C7	1.0(8)		
O1-C6-C7-C8	178.9(6)		
C5-C6-C7-C8	-0.7(8)		
C6-C7-C8-C3	-0.5(8)		
C4-C3-C8-C7	1.2(8)		
C1-C3-C8-C7	-178.7(6)		
C1-C2-C9-C14	7.0(9)		
N1-C2-C9-C14	-48.8(7)		
C1-C2-C9-C10	-171.6(6)		
N1-C2-C9-C10	132.6(6)		
C14-C9-C10-C11	0.5(10)		
C2-C9-C10-C11	179.1(7)		
C9-C10-C11-C12	-0.5(12)		

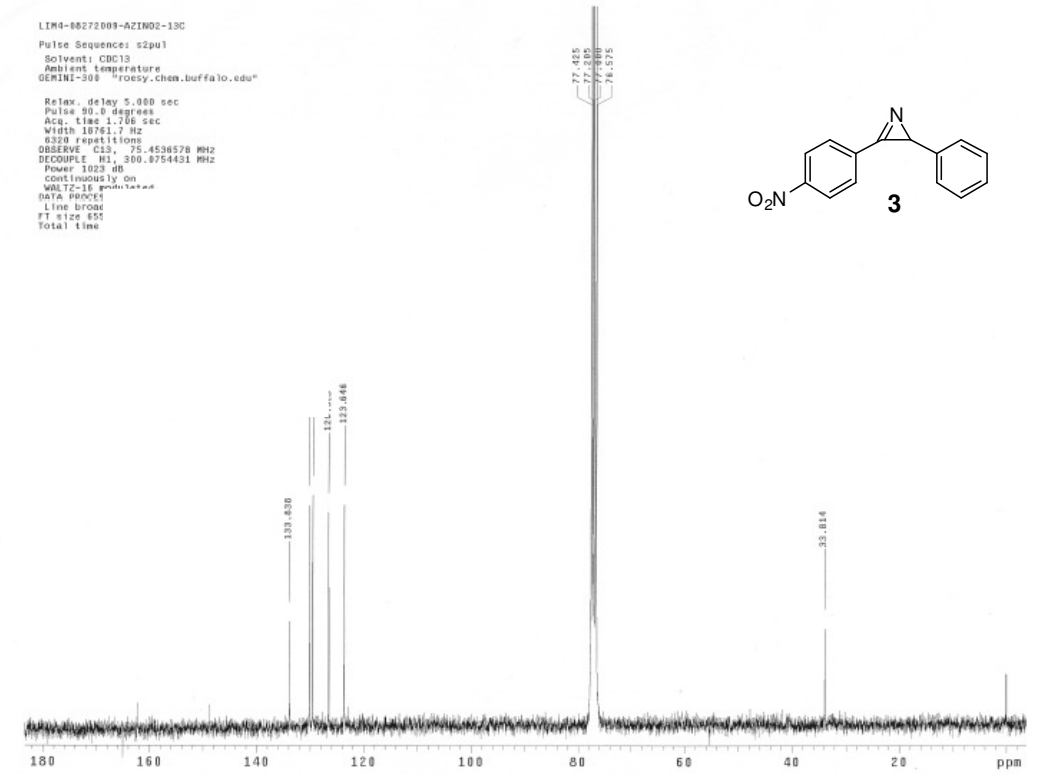

7.901
7.885
7.831
7.814
7.262
7.182
7.165
7.053
7.036
3.892
3.290
1.572
1.560
0.0000
2.08
2.10
2.08
3.10
1.00
12.04
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (t1)
1

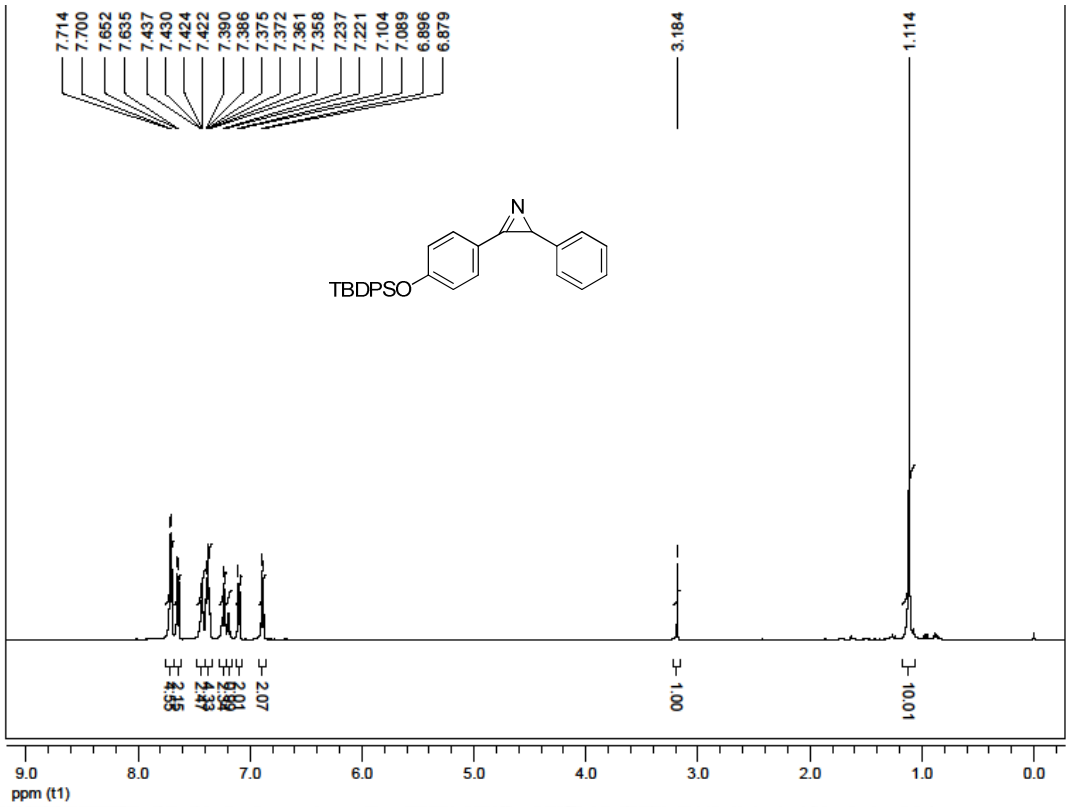

OBSERVE C13, 75.4536601 MHz
DECOUPLE H1, 300.0754431 MHz
Power 1023 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 1 hr, 59 min, 57 sec
166.481
163.721
161.339
146.878
132.824
129.505
129.014
125.794
115.762
114.867
77.425
77.080
76.575
60.814
55.562
33.814
14.311
220
200
180
160
140
120
100
80
60
40
20
0
ppm
1

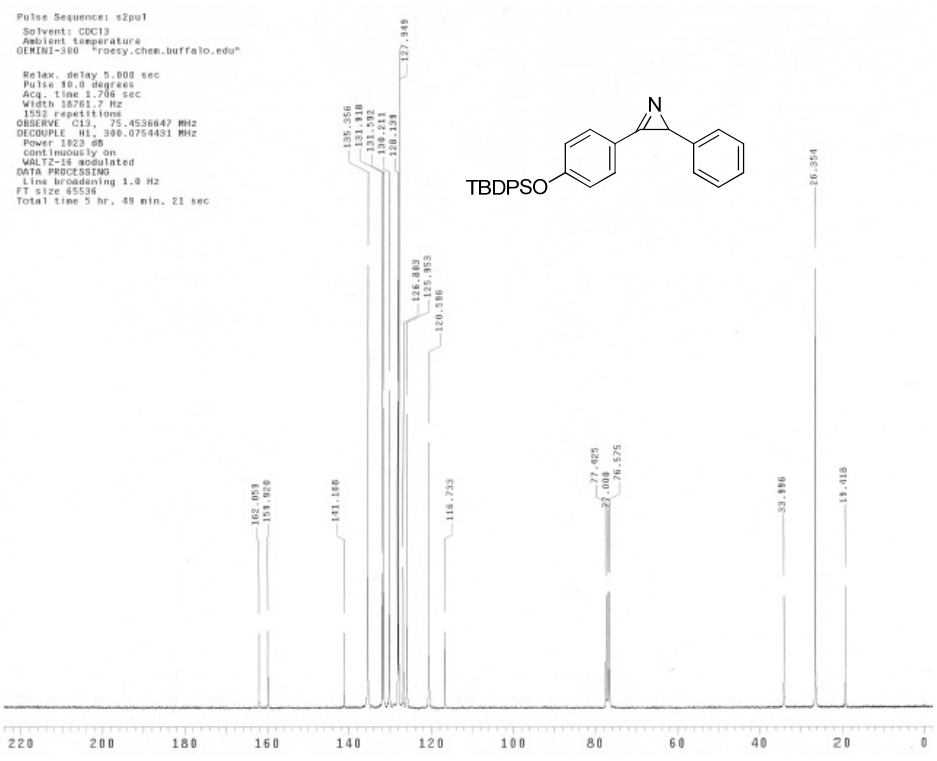

8.152
8.134
7.836
7.818
7.291
7.273
7.263
7.076
7.058
3.903
3.331
0.0000
N
O2N
O
2
2.02
2.11
2.96
2.11
3.17
1.00
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (f1)

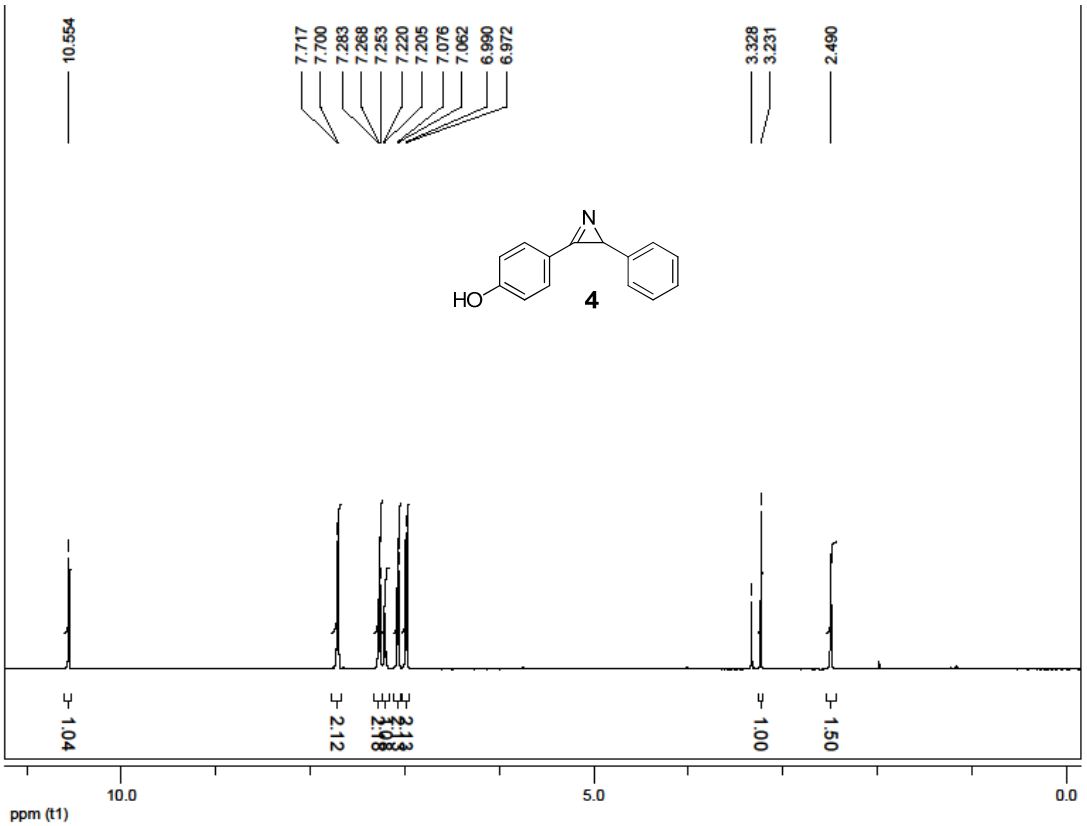

Pulse Sequence: s2pul
Solvent: CDCl3
Ambient temperature
GEMINI-300 "roesy.chem.buffalo.edu"
Relax. delay 5.000 sec
Pulse 90.0 degrees
Acq. time 1.706 sec
Width 18761.7 Hz
8064 repetitions
OBSERVE C13, 75.4536595 MHz
DECOUPLE H1, 300.0754031 MHz
Power 1023 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 20 hr, 22 min, 42 sec
77.425
77.005
76.585
76.575
132.284
126.447
123.593
115.026
55.631
33.450
164.017
160.678
149.379
148.882
-0.023
N
O2N
O
2
180
160
140
120
100
80
60
40
20
ppm

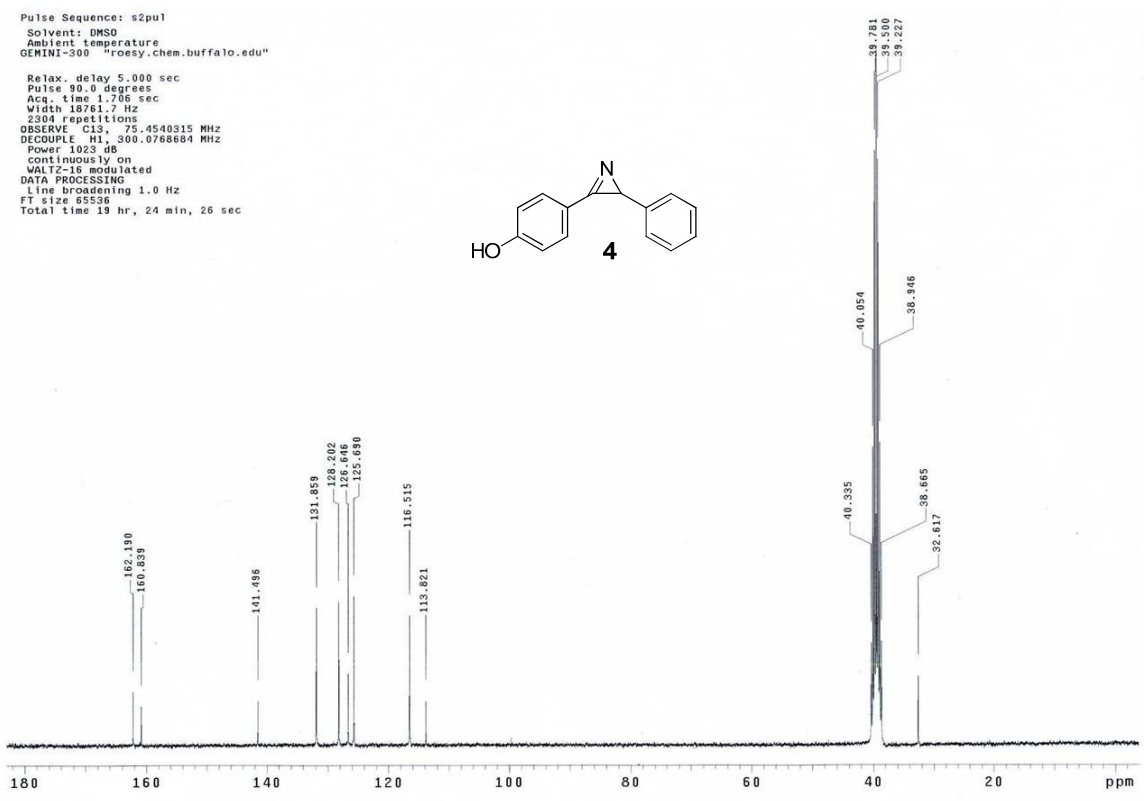

8.165
8.147
7.900
7.886
7.883
7.658
7.602
7.586
7.306
7.288
7.261
3.396
-0.000
N
O2N
3
2.08
2.06
2.28
1.00
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (f1)

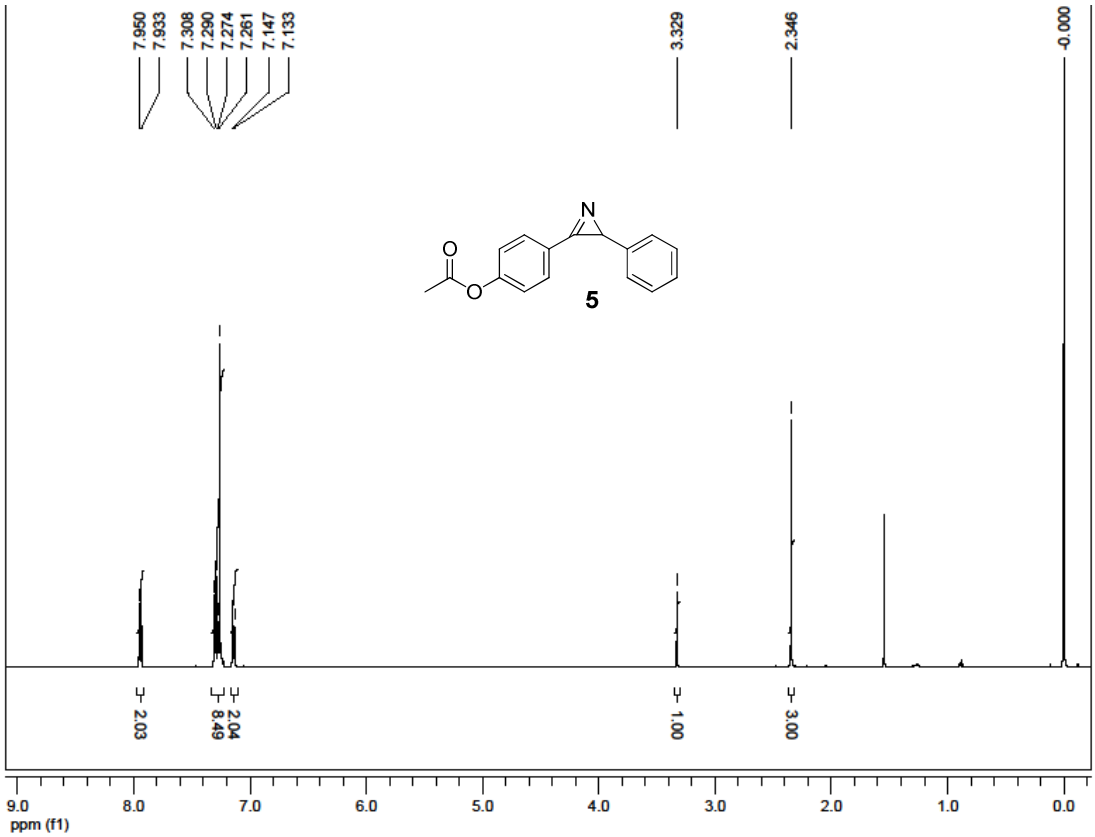

LIM4-80272009-AZINO2-13C
Pulse Sequence: s2pul
Solvent: CDCl3
Ambient temperature
GEMINI-300 "roesy.chem.buffalo.edu"
Relax. delay 5.000 sec
Pulse 90.0 degrees
Acq. time 1.706 sec
Width 10761.7 Hz
6320 repetitions
OBSERVE C13, 75.4538578 MHz
DECOUPLE H1, 300.0754431 MHz
Power 1023 dB
continuously on
WALTZ-16 modulated
DATA PROCES
Line broac
FT size 65
Total time
77.405
77.085
76.765
76.575
133.436
123.646
33.814
N
O2N
3
180
160
140
120
100
80
60
40
20
ppm

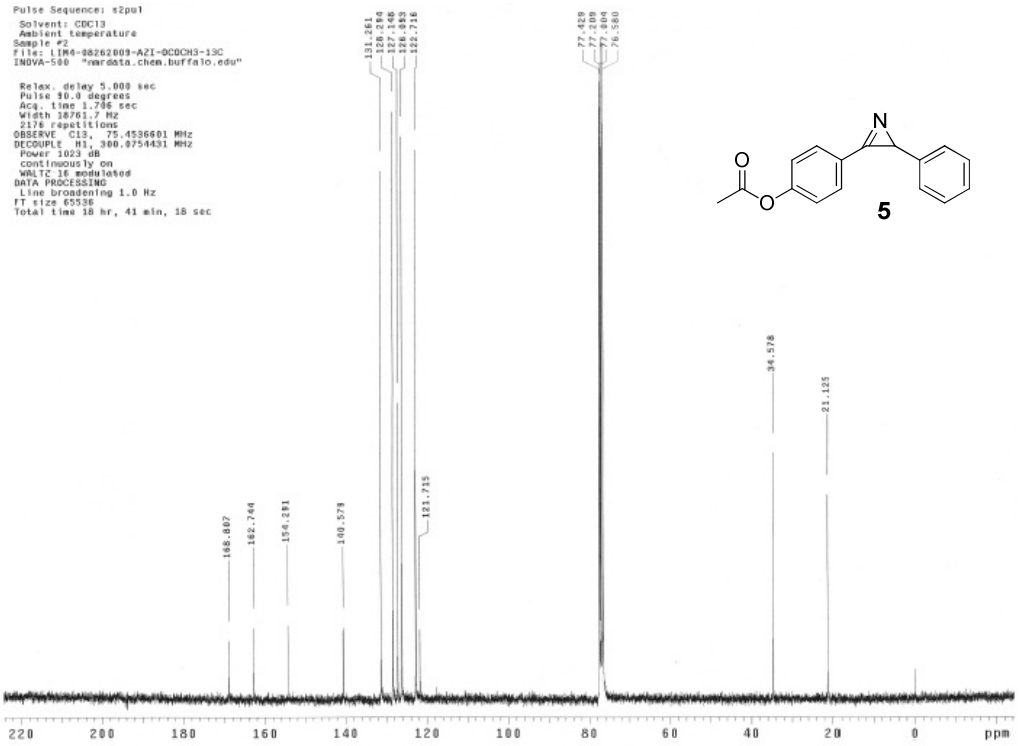

7.714
7.700
7.652
7.635
7.437
7.430
7.424
7.422
7.390
7.386
7.375
7.372
7.361
7.358
7.237
7.221
7.104
7.089
6.896
6.879
3.184
1.114
N
TBDPSO
2.07
1.00
10.01
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (t1)

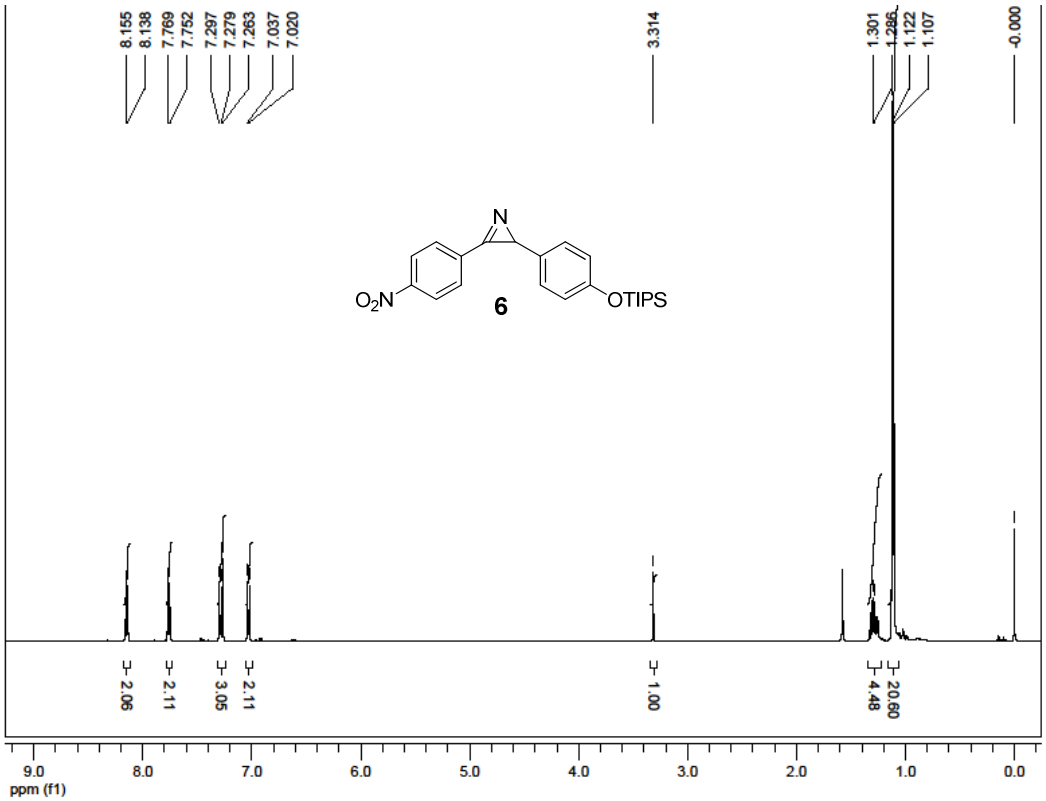

Pulse Sequence: s2pul
Solvent: CDCl3
Ambient temperature
GEMINI-300 "roesy.chem.buffalo.edu"
Relax. delay 5.000 sec
Pulse 30.0 degrees
Acq. time 1.708 sec
Width 18761.7 Hz
1552 repetitions
OBSERVE C13, 75.4536647 MHz
DECOUPLE H1, 300.0754431 MHz
Power 1023 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 5 hr, 49 min, 21 sec
N
TBDPSO
220
200
180
160
140
120
100
80
60
40
20
0

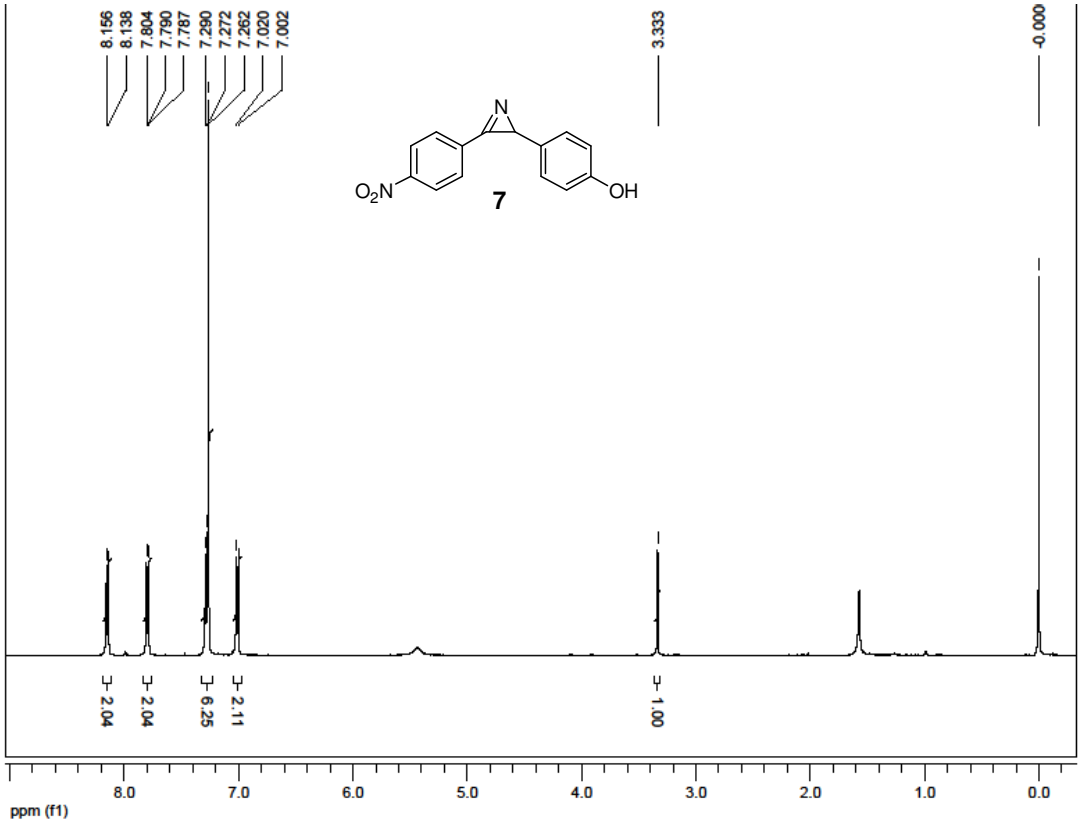

10.554
7.717
7.700
7.283
7.268
7.253
7.220
7.205
7.076
7.062
6.990
6.972
3.328
3.231
2.490
N
HO
4
1.04
2.12
1.00
1.50
10.0
5.0
0.0
ppm (t1)


Pulse Sequence: s2pul
Solvent: DMSO
Ambient temperature
GEMINI-300 "roesy.chem.buffalo.edu"
Relax. delay 5.000 sec
Pulse 90.0 degrees
Acq. time 1.706 sec
Width 18761.7 Hz
2304 repetitions
OBSERVE C13, 75.4540315 MHz
DECOUPLE H1, 300.0768684 MHz
Power 1023 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 19 hr, 24 min, 26 sec
N
HO
4
162.190
160.839
141.486
131.859
128.202
126.646
125.630
116.515
113.821
40.054
40.335
39.781
39.500
39.227
38.946
38.665
32.617
180
160
140
120
100
80
60
40
20
ppm

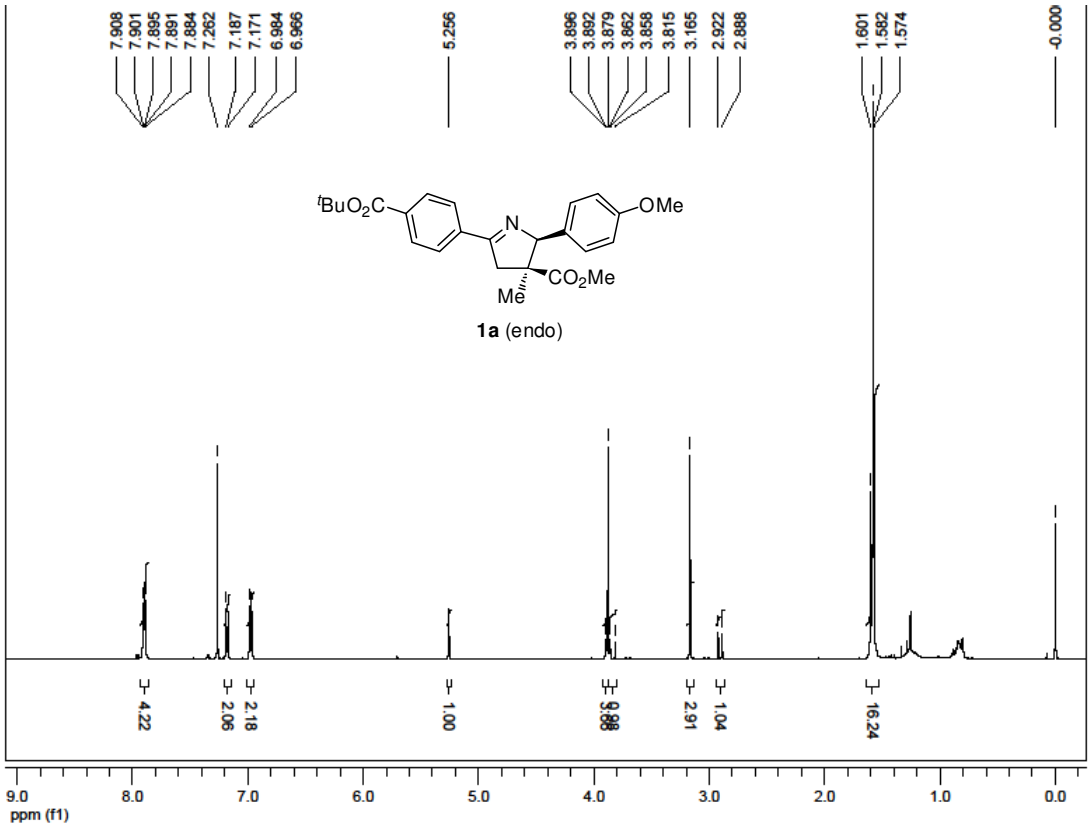

7.950
7.933
7.308
7.290
7.274
7.261
7.147
7.133
3.329
2.346
-0.000
N
O
O
5
2.03
8.49
2.04
1.00
3.00
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (f1)


Pulse Sequence: s2pul
Solvent: CDCl3
Ambient temperature
Sample #2
File: LIM4-08262009-AZI-OCOCH3-13C
INOVA-500 "nmrdata.chem.buffalo.edu"
Relax. delay 5.000 sec
Pulse 90.0 degrees
Acq. time 1.706 sec
Width 18761.7 Hz
2176 repetitions
OBSERVE C13, 75.4536601 MHz
DECOUPLE H1, 300.0754431 MHz
Power 1023 dB
continuously on
WALTZ 16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 18 hr, 41 min, 18 sec
131.261
128.744
128.580
127.148
126.853
122.716
77.429
77.209
77.004
76.580
34.578
21.125
168.867
162.744
154.281
140.578
121.735
N
O
O
5
220
200
180
160
140
120
100
80
60
40
20
0
ppm

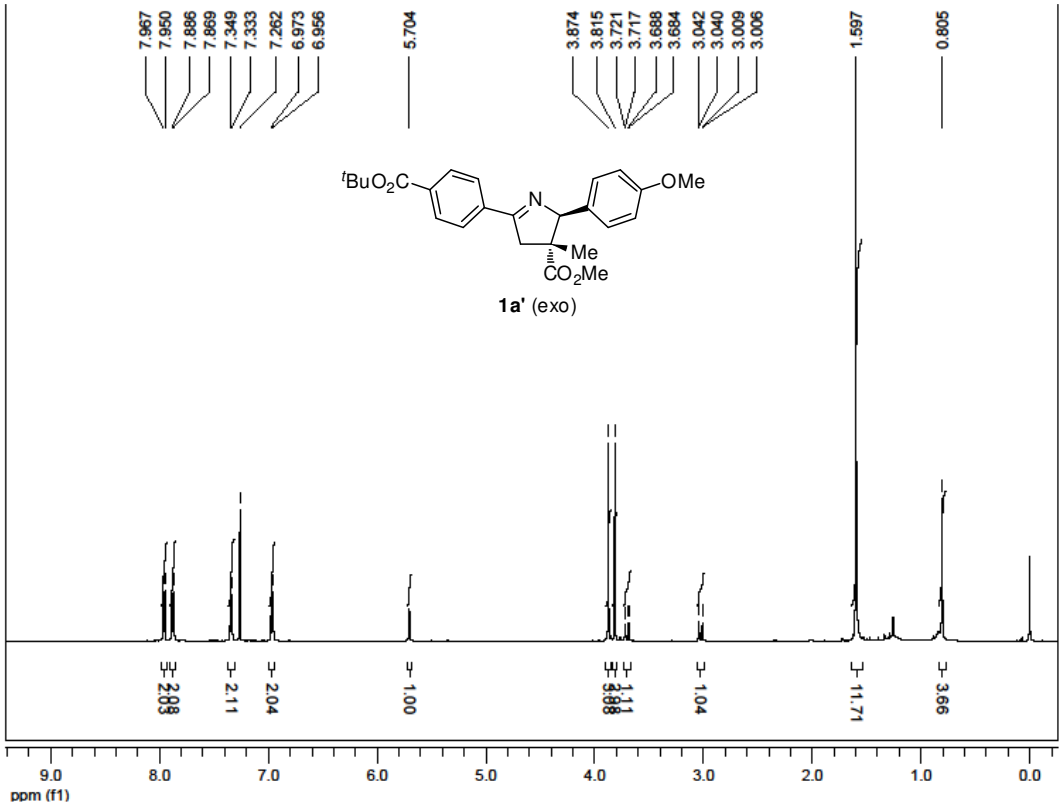

8.155
8.138
7.769
7.752
7.297
7.279
7.263
7.037
7.020
3.314
1.301
1.286
1.122
1.107
-0.000
N
O2N
OTIPS
6
2.06
2.11
3.05
2.11
1.00
4.48
20.60
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (f1)


8.156
8.138
7.804
7.790
7.787
7.290
7.272
7.262
7.020
7.002
3.333
0.000
N
O2N
OH
7
2.04
2.04
6.25
2.11
1.00
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (f1)


Pulse Sequence: s2pul
Solvent: CDCl3
Ambient temperature
GEMINI-300 "roesy.chem.buffalo.edu"
Relax. delay 5.000 sec
Pulse 90.0 degrees
Acq. time 1.706 sec
Width 18761.7 Hz
7520 repetitions
OBSERVE C13, 75.4536583 MHz
DECOUPLE H1, 300.0754431 MHz
Power 1023 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 19 hr, 26 min, 49 sec
77.425
77.000
76.575
160.686
160.504
149.250
132.510
124.454
123.616
116.627
115.307
33.488
-0.023
160
140
120
100
80
60
40
20
ppm
N
O2N
OH
7
8.177
8.159
7.846
7.829
7.360
7.343
7.164
7.146
4.886
4.594
3.417
3.316
3.313
3.310
3.307
3.304
O
O
2.07
2.07
2.08
2.06
2.12
1.00
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (f1)



1a (endo)



1a' (exo)


Solvent: CDCl3
Ambient temperature
File: yzp-240418-a275-mpeg-h-cdcl3
INOVA-500 "tocsy.chem.buffalo.edu"
Relax. delay 2.000 sec
Pulse 35.7 degrees
Acq. time 1.891 sec
Width 10000.0 Hz
64 repetitions
OBSERVE H1, 499.8884010 MHz
DATA PROCESSING
Line broadening 0.7 Hz
FT size 65536
Total time 4 min, 9 sec
9
ppm