

*Supporting Information for*

**Myoglobin-catalyzed intermolecular carbene N-H insertion with  
arylamine substrates**

Gopeekrishnan Sreenilayam and Rudi Fasan\*

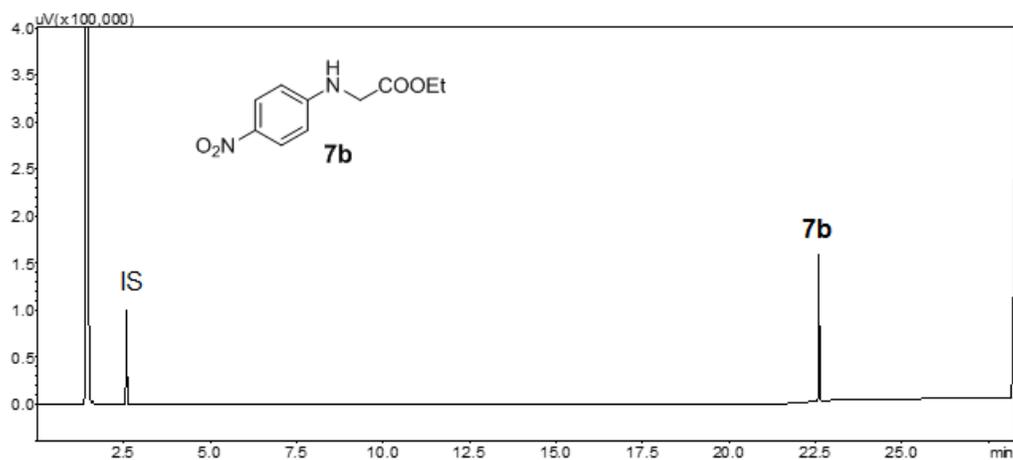
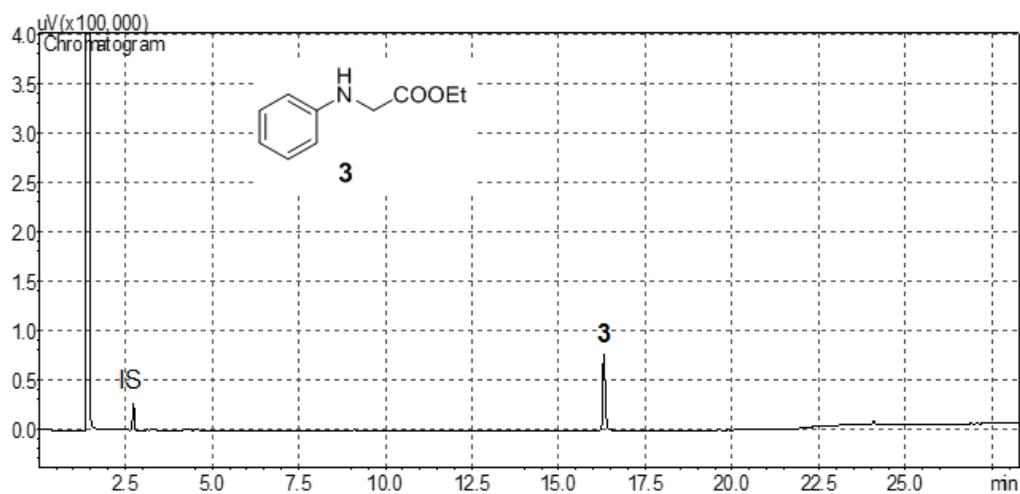
*Department of Chemistry, University of Rochester, 14627 Rochester, New York, USA*

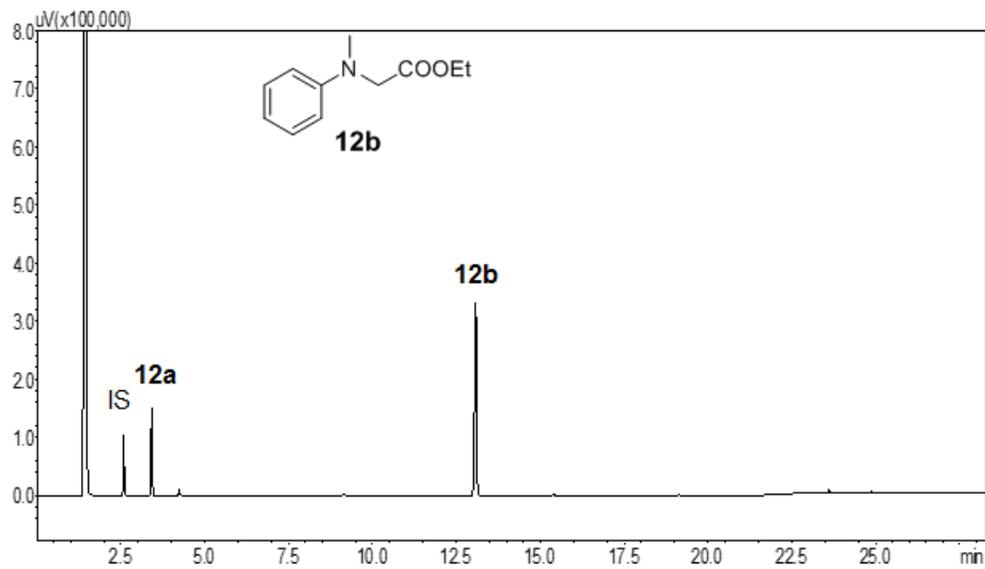
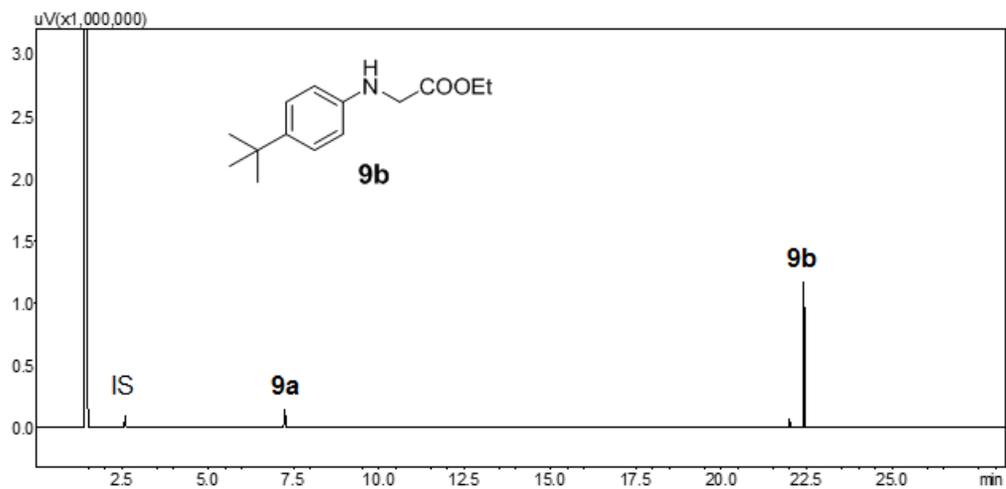
*Correspondence should be addressed to R.F. (fasan@chem.rochester.edu)*

**Table of contents:**

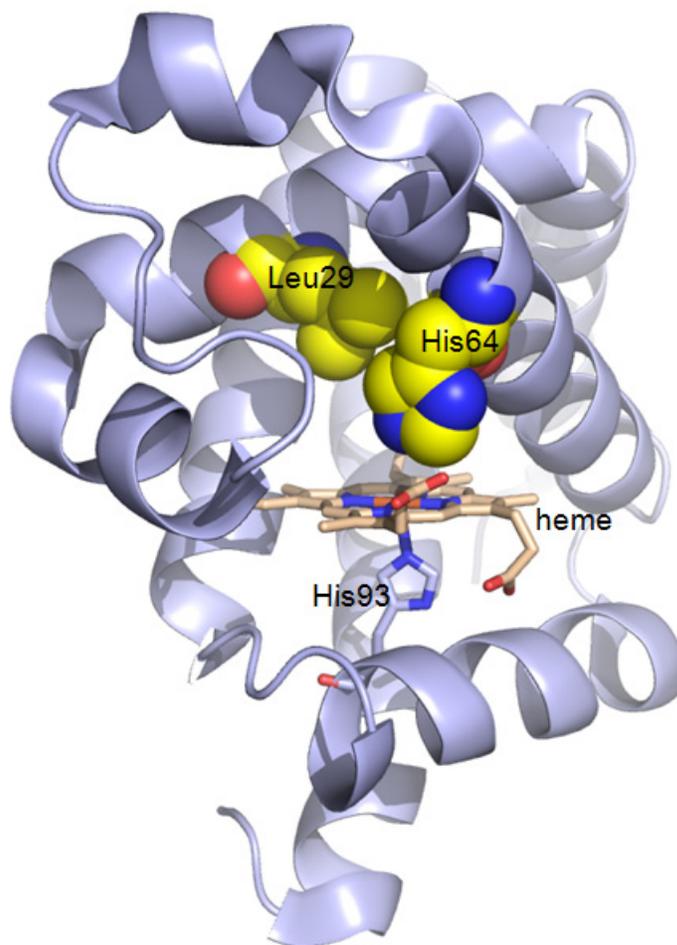
Supplementary Figures	Page S2-S4
Experimental Procedures	Pages S5-S8
Synthetic Procedures	Pages S9-S15
References	Page S16
NMR Spectra	Pages S17-S29

**Supplementary Figure S1.** GC traces for representative Mb(H64V,V68A)-catalyzed N–H insertion reactions. The structure and peak of the desired N–H insertion product are shown. The peaks corresponding to the amine substrate and internal standard (IS) are also labeled.





**Supplementary Figure S1.** Crystal structure of wild-type sperm whale myoglobin (pdb 1A6K). Residues H64 and L29 are highlighted in yellow and displayed as sphere models. The heme cofactor (light brown) and proximal H93 ligand are displayed as stick models.



## Experimental Procedures

**Reagents and Analytical Methods.** All the chemicals and reagents were purchased from commercial suppliers (Sigma-Aldrich, ACS Scientific, Acros, Alfa Aesar) and used without any further purification, unless otherwise stated. All dry reactions were carried out under argon atmosphere in oven-dried glassware with magnetic stirring using standard gas-tight syringes, cannulae and septa.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were measured on Bruker DPX-400 (operating at 400 MHz for  $^1\text{H}$  and 100 MHz for  $^{13}\text{C}$ ) or Bruker DPX-500 (operating at 500 MHz for  $^1\text{H}$  and 125 MHz for  $^{13}\text{C}$ ). Tetramethylsilane (TMS) served as the internal standard (0 ppm) for  $^1\text{H}$  NMR and  $\text{CDCl}_3$  was used as the internal standard (77.0 ppm) for  $^{13}\text{C}$  NMR. Silica gel chromatography purifications were carried out using AMD Silica Gel 60 230-400 mesh. Gas chromatography (GC) analyses were carried out using a Shimadzu GC-2010 gas chromatograph equipped with a FID detector and a Shimadzu SHRXI-5MS column (15 m x 0.25 mm x 0.25  $\mu\text{m}$  film). Separation method: a) 1  $\mu\text{L}$  injection, injector temp.: 200  $^\circ\text{C}$ , detector temp: 300  $^\circ\text{C}$ . Gradient: column temperature set at 60  $^\circ\text{C}$  for 1 min, then to 200  $^\circ\text{C}$  at 10  $^\circ\text{C}/\text{min}$ , then to 290  $^\circ\text{C}$  at 30  $^\circ\text{C}/\text{min}$ . Total run time was 19.00 min. Separation method: b) using a Shimadzu GC-2010 gas chromatograph equipped with a FID detector, and a Cyclosil-B column (30 m x 0.25 mm x 0.25  $\mu\text{m}$  film). 1  $\mu\text{L}$  injection, injector temp.: 200  $^\circ\text{C}$ , detector temp: 300  $^\circ\text{C}$ . Gradient: column temperature set at 140  $^\circ\text{C}$  for 3 min, then to 160  $^\circ\text{C}$  at 1.8  $^\circ\text{C}/\text{min}$ , then to 165  $^\circ\text{C}$  at 1  $^\circ\text{C}/\text{min}$ , then to 245  $^\circ\text{C}$  at 25  $^\circ\text{C}/\text{min}$ . Total run time was 28 min.

**Protein expression and purification.** Wild-type Mb and the engineered Mb variants were expressed in *E. coli* BL21(DE3) cells as described previously.<sup>[1]</sup> Briefly, cells were grown in TB medium (ampicillin, 100 mg  $\text{L}^{-1}$ ) at 37  $^\circ\text{C}$  (150 rpm) until  $\text{OD}_{600}$  reached 0.6. Cells were then induced with 0.25 mM  $\beta$ -D-1-thiogalactopyranoside (IPTG) and 0.3 mM  $\delta$ -aminolevulinic acid

(ALA). After induction, cultures were shaken at 150 rpm and 27 °C and harvested after 20 h by centrifugation at 4000 rpm at 4 °C. After cell lysis by sonication, the proteins were purified by Ni-affinity chromatography using the following buffers: loading buffer (50 mM Kpi, 800 mM NaCl, pH 7.0), wash buffer 1 (50 mM Kpi, 800 mM NaCl, pH 6.2), wash buffer 2 (50 mM Kpi, 800 mM NaCl, 250 mM glycine, pH 7.0) and elution buffer (50 mM Kpi, 800 mM NaCl, 300 mM L-histidine, pH 7.0). After buffer exchange (50 mM Kpi, pH 7.0), the proteins were stored at +4 °C. Myoglobin concentration was determined using an extinction coefficient  $\epsilon_{410} = 157 \text{ mM}^{-1} \text{ cm}^{-1}$ .<sup>[2]</sup>

**N-H insertion reactions.** Initial reactions (**Table 1**) were carried out at a 400  $\mu\text{L}$  scale using 20  $\mu\text{M}$  myoglobin, 10 mM aniline, 10 or 5 mM EDA, and 10 mM sodium dithionite. In a typical procedure, a solution containing sodium dithionite (100 mM stock solution) in potassium phosphate buffer (50 mM, pH 8.0) was degassed by bubbling argon into the mixture for 4 min in a sealed vial. A buffered solution containing myoglobin was carefully degassed in a similar manner in a separate vial. The two solutions were then mixed together via cannula. Reactions were initiated by addition of 10  $\mu\text{L}$  of aniline (from a 0.4 M stock solution in methanol), followed by the addition of 10  $\mu\text{L}$  or 5  $\mu\text{L}$  of EDA (from a 0.4 M stock solution in methanol) with a syringe, and the reaction mixture was stirred for 12 h at room temperature, under positive argon pressure. Reaction with hemin were carried out using an identical procedure with the exception that the purified Mb was replaced by 100  $\mu\text{L}$  of a hemin solution (80  $\mu\text{M}$  in DMSO:H<sub>2</sub>O, 1:1). For the optimization of the aniline:EDA ratio, reactions were performed according to the general procedure described above, using 20  $\mu\text{M}$  of protein Mb (H64V, V68A), 10 mM of aniline and variable amounts of EDA (2.5 mM EDA to 10mM EDA). Optimization of

the substrate loading was done in a similar manner, using 20  $\mu\text{M}$  Mb (H64V, V68A), variable quantities of aniline (from 10 to 160 mM final concentration), and variable quantities of EDA (from 10 to 160 mM final concentration), maintaining an aniline : EDA ratio of 1:1 at all times. Reactions for TTN determination were carried out according to the general procedure described above with 20  $\mu\text{M}$  of Mb (H64V, V68A), 10 mM aniline (10  $\mu\text{L}$  of 0.4 M stock solution in methanol), and 10 mM EDA (10  $\mu\text{L}$  of 0.4 M stock solution in methanol) were used. Optimization of the protein concentration was carried according to the general procedure along with varying the hemoprotein concentration from 20  $\mu\text{M}$  to 0.08  $\mu\text{M}$  of Mb (H64V, V68A) and 10 mM aniline (10  $\mu\text{L}$  of 0.4 M stock solution in methanol), and 10 mM EDA (10  $\mu\text{L}$  of 0.4 M stock solution in methanol).

**Preparative-scale reaction.** A solution containing sodium dithionate (100 mM stock solution, 1 mL, 10 mM) in potassium phosphate buffer (50 mM, pH 8.0, 7.6 mL) was degassed by bubbling argon into the mixture for 20 min in a sealed vial. A buffered solution containing 20  $\mu\text{M}$  of Mb(H64V, V68A) (1.34 mL of 150  $\mu\text{M}$  stock solution) was carefully degassed in a similar manner in a separate vial. The two solutions were then mixed together via cannula. Reactions were initiated by addition of 9.2  $\mu\text{L}$  of pure aniline, followed by the addition of 12  $\mu\text{L}$  of EDA with a syringe, and the reaction mixture was stirred for 12 h at room temperature, under positive argon pressure. The reaction mixture was extracted with dichloromethane (4 x 10 mL), organic layer evaporated under reduced pressure and the residue was purified by flash column chromatography (10% ethyl acetate in hexanes) to yield ethyl phenylglycinate **3** as colorless solid (14.5 mg, 80%). A small amount (0.9%) of the double insertion product was observed in the reaction mixture by GC analysis.

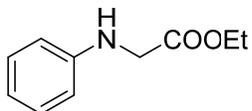
**Product analysis.** The reactions were analyzed by adding 20  $\mu\text{L}$  of internal standard (benzodioxole, 50 mM in methanol) to the reaction mixture, followed by extraction with 400  $\mu\text{L}$  of dichloromethane and separated organic layer was analyzed by GC-FID (see **Reagents and Analytical Methods** section for details on GC analyses). Calibration curves for quantification of the different N-H insertion products were constructed using authentic standards prepared synthetically using  $\text{Rh}_2(\text{OAc})_4$  as catalyst as described in **Synthetic Procedures**. All measurements were performed at least in duplicate. For each experiment, negative control samples containing either no hemoprotein or no reductant were included.

## Synthetic Procedures:

### General procedure for chemical synthesis of authentic N-H insertion product standards

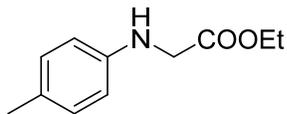
To a flame dried round bottom flask under argon, equipped with a stir bar was added amine (1 equiv.) and  $\text{Rh}_2(\text{OAc})_4$  (1 mol%) in toluene (2-3 mL). To this solution was added a solution of diazo compound (1 equiv.) in toluene (1-2 mL) over 30 minutes at 0 °C. The resulting mixture was heated at 80 °C for another 15-18 hrs. The solvent was removed under vacuum and the crude mixture was purified by flash column chromatography (hexanes/ethyl acetate) to provide N-H insertion products in good to excellent yield. The insertion products were characterized by GC-MS,  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR techniques.

### Ethyl phenylglycinate (3):



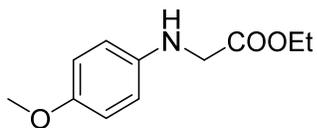
Following the standard procedure, pale brown solid, % yield (88), GC-MS  $m/z$  (% relative intensity): 179(32.1), 106(100), 77(21.2), 51(5.9);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.23 (t,  $J$  = 6.0 Hz, 2H), 6.79 (t,  $J$  = 6.5 Hz, 1H), 6.64 (d,  $J$  = 7.0 Hz, 2H), 4.34-4.24 (m, 3H), 3.91 (s, 2H), 1.33 (t,  $J$  = 6.5 Hz, 3H) ppm;  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  171.2, 147.1, 129.3, 118.1, 113.0, 61.3, 45.8, 14.2 ppm.

**Ethyl *p*-tolylglycinate (4b):**



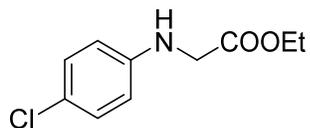
Following the standard procedure, light brown solid, % yield (86), GC-MS *m/z* (% relative intensity): 193(26.3), 120(100), 91(18.7), 65(5.7); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): δ 7.03 (d, *J* = 7.6 Hz, 2H), 6.56 (d, *J* = 8.0 Hz, 2H), 4.27 (q, *J* = 6.8 Hz, 2H), 3.88 (s, 2H), 2.26 (s, 3H), 1.32 (t, *J* = 7.2 Hz, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): δ 171.3, 144.8, 129.8, 127.4, 113.2, 61.2, 46.2, 20.4, 14.2 ppm.

**Ethyl (4-methoxyphenyl)glycinate (5b):**



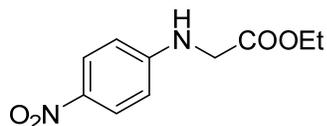
Following the standard procedure, colorless solid, % yield (80), GC-MS *m/z* (% relative intensity): 209(42.0), 136(100), 121(14.6), 108(13.0), 77(10.0); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): δ 6.79 (d, *J* = 8.4 Hz, 2H), 6.58 (d, *J* = 8.4 Hz, 2H), 4.24 (q, *J* = 7.2 Hz, 2H), 4.04 (br s, 1H), 3.84 (s, 2H), 3.73 (s, 3H), 1.29 (t, *J* = 7.2 Hz, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): δ 171.4, 152.6, 141.3, 114.8, 114.3, 61.2, 55.7, 46.8, 14.2 ppm.

**Ethyl (4-chlorophenyl)glycinate (6b):**



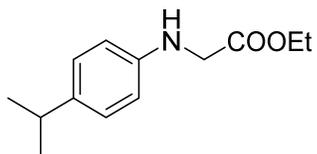
Following the standard procedure, colorless solid, % yield (82), GC-MS  $m/z$  (% relative intensity): 213(22.7), 142(42.5), 140(100), 105(14.1), 77(13.3);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.13 (d,  $J = 9.0$  Hz, 2H), 6.52 (d,  $J = 8.5$  Hz, 2H), 4.26 (q,  $J = 6.8$  Hz, 2H), 3.85 (s, 2H), 1.30 (t,  $J = 7.0$  Hz, 3H) ppm;  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  170.8, 145.6, 129.1, 122.8, 114.0, 61.4, 45.8, 14.2 ppm.

**Ethyl (4-nitrophenyl)glycinate (7b):**



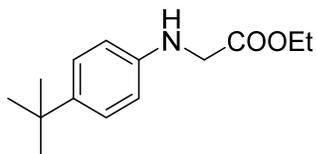
Following the standard procedure, yellow solid, % yield (75), GC-MS  $m/z$  (% relative intensity): 224(12.6), 151(100), 105(47.4), 76(4.2);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  8.12 (d,  $J = 9.0$  Hz, 2H), 6.56 (d,  $J = 8.5$  Hz, 2H), 5.08 (br s, 1H), 4.31 (q,  $J = 7.5$  Hz, 2H), 3.98 (s, 2H), 1.33 (t,  $J = 7.1$  Hz, 3H) ppm;  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  169.7, 151.9, 126.3, 111.5, 61.9, 44.9, 14.1 ppm.

**Ethyl (4-isopropylphenyl)glycinate (8b):**



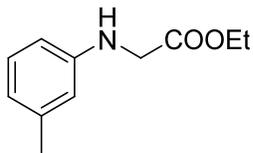
Following the standard procedure, light brown oil, % yield (82), GC-MS  $m/z$  (% relative intensity): 221(33.7), 206(32.2), 178(12.5), 148(100), 132(30.6);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz):  $\delta$  7.08 (d,  $J = 7.2$  Hz, 2H), 6.58 (d,  $J = 7.2$  Hz, 2H), 4.27-4.22 (m, 3H), 3.89 (s, 2H), 2.85-2.79 (m, 1H), 1.34-1.28 (m, 3H), 1.24-1.21 (m, 6H) ppm;  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz):  $\delta$  171.3, 145.1, 138.7, 127.2, 113.1, 61.2, 46.1, 33.2, 24.2, 14.2 ppm.

**Ethyl (4-(*tert*-butyl)phenyl)glycinate (9b):**



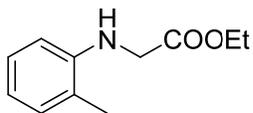
Following the standard procedure, brown oil, % yield (80), GC-MS  $m/z$  (% relative intensity): 235(34.6), 220(87.9), 192(26.3), 162(100), 146(36.1);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz):  $\delta$  7.23 (d,  $J = 8.4$  Hz, 2H), 6.58 (d,  $J = 8.0$  Hz, 2H), 4.26 (q,  $J = 7.2$  Hz, 2H), 3.88 (s, 2H), 1.31-1.27 (m, 12 H) ppm;  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz):  $\delta$  171.3, 144.6, 140.9, 126.1, 112.8, 61.3, 46.1, 33.9, 31.5, 31.3, 14.2 ppm.

**Ethyl *m*-tolylglycinate (10b):**



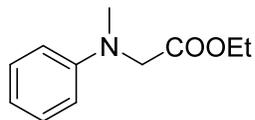
Following the standard procedure, colorless solid, % yield (83), GC-MS *m/z* (% relative intensity): 193(41.4), 120(100), 91(35.1), 65(11.2); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): δ 7.12 (t, *J* = 7.6 Hz, 1H), 6.61 (d, *J* = 7.2 Hz, 1H), 6.45-6.43 (m, 2H), 4.28-4.23 (m, 3H), 3.90 (s, 2H), 2.30 (s, 3H), 1.33 (t, *J* = 7.2 Hz, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): δ 171.2, 147.1, 139.1, 129.2, 119.1, 113.8, 110.1, 61.3, 45.9, 21.6, 14.2 ppm.

**Ethyl *o*-tolylglycinate (11b):**



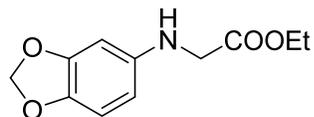
Following the standard procedure, light brown oil, % yield (85), GC-MS *m/z* (% relative intensity): 193(31.2), 120(100), 91(25.1), 65(7.0); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 7.16 (t, *J* = 7.5 Hz, 1H), 7.11 (d, *J* = 7.5 Hz, 1H), 6.75 (t, *J* = 7.5 Hz, 1H), 6.52 (d, *J* = 8.0 Hz, 1H), 4.30 (q, *J* = 7.0 Hz, 2H), 4.07 (br s, 1H), 3.96 (s, 2H), 2.24 (s, 3H), 1.35 (t, *J* = 7.0 Hz, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz): δ 171.3, 145.1, 130.2, 127.1, 122.5, 117.8, 109.9, 61.3, 45.9, 17.3, 14.2 ppm.

**Ethyl *N*-methyl-*N*-phenylglycinate (12b):**



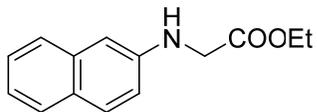
Following the standard procedure, brown oil, % yield (79), GC-MS *m/z* (% relative intensity): 193(27.1), 120(100), 91(18.5), 65(5.75); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 7.26 (t, *J* = 7.0 Hz, 2H), 6.78-6.70 (m, 3H), 4.21 (q, *J* = 7.0 Hz, 2H), 4.07 (s, 2H), 3.08 (s, 3H), 1.27 (t, *J* = 7.0 Hz, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz): δ 171.1, 148.9, 129.2, 117.3, 112.3, 60.8, 54.5, 39.5, 14.2 ppm.

**Ethyl benzo[*d*][1,3]dioxol-5-ylglycinate (13b):**



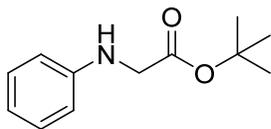
Following the standard procedure, light brown solid, % yield (75), GC-MS *m/z* (% relative intensity): 223(36.7), 150(100), 120(9.4), 92(12.7), 65(16.9); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): δ 6.66 (d, *J* = 8.0 Hz, 1H), 6.24 (s, 1H), 6.02 (d, *J* = 8.0 Hz, 1H), 5.85 (s, 2H), 4.21-4.19 (m, 2H), 4.07 (br s, 1H), 3.82 (s, 2H), 1.30-1.26 (m, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): δ 171.2, 148.4, 142.7, 140.2, 108.6, 104.5, 100.7, 96.3, 61.3, 46.7, 14.2 ppm.

**Ethyl naphthalen-2-ylglycinate (14b):**



Following the standard procedure, purple solid, % yield (78), GC-MS m/z (% relative intensity): 229(19.7), 156(100), 127(19.7); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): δ 7.71-7.63(m, 3H), 7.41 (t, *J* = 7.6 Hz, 1H), 7.25 (t, *J* = 6.4 Hz, 1H), 6.96 (d, *J* = 8.4 Hz, 1H), 6.75 (s, 1H), 4.50 (br s, 1H), 4.31 (q, *J* = 7.2 Hz, 2H), 4.01 (s, 2H), 1.34 (t, *J* = 7.2 Hz, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): δ 171.0, 144.7, 135.0, 129.1, 127.8, 127.7, 126.4, 126.0, 122.3, 117.9, 104.7, 61.4, 45.8, 14.2 ppm.

***tert*-butyl phenylglycinate (15):**



Following the standard procedure, yellow oil, % yield (89), GC-MS m/z (% relative intensity): 207(13.5), 151(37.3), 106(100), 77(18.7), 57(27.9), <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 7.21-7.18 (m, 2H), 6.76 (t, *J* = 7.5 Hz, 1H), 6.62 (d, *J* = 8.0 Hz, 2H), 3.80 (s, 2 H), 1.50 (s, 9H) ppm, <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz): δ 170.3, 147.2, 129.2, 118.0, 113.0, 81.9, 45.5, 28.1 ppm.

## References

- [1] M. Bordeaux, R. Singh, R. Fasan, *Bioorg. Med. Chem.* **2014**, *22*, 5697-5704.
- [2] C. Redaelli, E. Monzani, L. Santagostini, L. Casella, A. M. Sanangelantoni, R. Pierattelli, L. Banci, *Chembiochem* **2002**, *3*, 226-233.

# $^1\text{H}$ and $^{13}\text{C}$ NMR Spectra

