Identification of a Lysine 4-Hydroxylase from the Glidobactin Biosynthesis and Evaluation of Its Biocatalytic Potential

*Alexander Amatuni, Hans Renata**

Contribution from the Department of Chemistry, The Scripps Research Institute Florida, 130 Scripps Way, Jupiter, FL 33458

SUPPLEMENTARY MATERIAL

Table of Contents

General materials and methods

Unless otherwise noted, all chemicals and reagents for chemical reactions were purchased at the highest commercial quality and used without further purification. Reactions were monitored by thin layer chromatography (TLC) and liquid chromatography/mass spectrometry (LC/MS). TLC was performed with 0.25 mm E. Merck silica plates (60F-254) using short-wave UV light as the visualizing agent, and ninhydrin, $KMnO_4$, or phosphomolybdic acid and heat as developing agents. LC/MS was performed with Agilent 1260 Infinity System equipped with Poroshell 120 EC-C18 column (3.0 x 50 mm, 2.7 micron). NMR spectra were recorded on a Bruker spetrometer and calibrated using residual undeuterated solvent. Optical rotations were measured on Autopol IV polarimeter (Rudolph Research Analytical). Enzymes (DpnI, Q5 polymerase) were purchased from New England Biolabs (NEB, Ipswich, MA). pET28a(+) GlbB expression vector was obtained via DNA synthesis from Genscript and was used directly to transform electrocompetent *E. coli* strain BL21(DE3). Sonication was performed using a Qsonica Q500 sonicator. Purified enzymes were accessed via immobilized metal ion affinity chromatography with HisTrap HP column.

Enzyme expression and purification

Two hundred milliliter TB_{amp} was inoculated with an overnight culture (1 mL, LB_{kan}) of recombinant *E. coli* BL21(DE3) cells harboring a pET28a(+) GlbB plasmid. The cultures were shaken at 250 rpm at 37 °C for roughly 2.5 h or until an optical density of 0.7–1.0 was reached. Cultures were cooled on ice (20 min) and then induced by adding IPTG to a final concentration of 0.025 mM. The cultures were allowed to continue for another 20 hours at 20 ºC and shaking at 250 rpm. Cell were harvested by centrifugation (4 °C, 15 min, 3,000xg), and the cell pellet was stored at –20 °C or below for at least 2 h.

Purification was performed with an AKTA pure FPLC system (GE Healthcare). For the purification of 6XHis tagged enzymes, the thawed cell pellet was resuspended in Ni-NTA buffer A (25 mM Tris.HCl, 200 mM NaCl, 25 mM imidazole, pH 8.0, 4 mL/g of cell wet weight) and lysed by sonication (3 min, 1 s On/4 s Off cycles, 50% amplitude). The lysate was centrifuged at 15,000xg for 30 min at 4 °C to remove cell debris. The collected supernatant was subjected to a Ni-NTA chromatography step using a Ni Sepharose column (HisTrap-HP, GE healthcare, Piscataway, NJ). The protein was eluted from the Ni Sepharose column using 25 mM Tris.HCl, 200 mM NaCl, 300 mM imidazole, pH 8.0. Ni-purified protein was buffer exchanged into 0.05 M phosphate buffer (pH = 8.0) using a 10 kDa MW cut-off centrifugal filter. Protein concentration was determined by A_{280} with calculated extinction coefficient (31190 $M^{-1}cm^{-1}$ for N-His-tagged GlbB). For storage, proteins were portioned into 100 µL aliquots, flash frozen on liquid N_2, and stored at –80 °C.

Protein and DNA Sequences

Protein sequence of GlbB (Uniprot accession code: A8KCI8)

MQLGVPFLKRESMSKMTGQEWAAAAPETEPGDVRAALQQRGWARFDATDMQVAVDEAADLQRLTE
YARSLPVDRFGTGGRHRSYAEGILTPRRKTIAWKAGARTPDGRVEIAYVQHSEFQPEHGGVVRNF
ARTREDILALPLVHRLIWYDLSLTPMFDAEGDLLCGFHMIRMQATPGAVARITPDCLHQDGQPFT
AVHLVERSHAEGGVNFIAPPRYTGRQFDEVPSHLLSAFVLGSPLQSYIIDDAAICHQVTAVSCSP
GASHGTRTVILIDFSPLNPASSAQPT

DNA sequence of GlbB (codon optimized using IDT Codon Optimization Tool)

ATGCAGCTGGGTGTGCCGTTTCTGAAGCGTGAGAGCATGAGCAAAATGACCGGCCAAGAATGGGCGGCGG
CGGCGCCGGAGACCGAACCGGGTGATGTTCGTGCGGCGCTGCAGCAACGTGGCTGGGCGCGTTTTGACGC
GACCGATATGCAGGTGGCGGTTGACGAGGCGGCGGATCTGCAACGTCTGACCGAATATGCGCGTAGCCTG
CCGGTGGACCGTTTTGGTACCGGTGGCCGTCACCGTAGCTATGCGGAGGGCATCCTGACCCCGCGTCGTA

AGACCATTGCGTGGAAGGCGGGTGCGCGTACCCCGGATGGTCGTGTGGAAATCGCGTACGTTCAGCACAG
CGAGTTCCAACCGGAACACGGTGGCGTGGTTCGTAACTTTGCGCGTACCCGTGAGGACATCCTGGCGCTG
CCGCTGGTTCACCGTCTGATTTGGTATGATCTGAGCCTGACCCCGATGTTCGACGCGGAAGGTGATCTGC
TGTGCGGCTTTCACATGATTCGTATGCAGGCGACCCCGGGTGCGGTGGCGCGTATTACCCCGGACTGCCT
GCACCAGGATGGCCAACCGTTTACCGCGGTGCACCTGGTTGAGCGTAGCCACGCGGAAGGTGGCGTTAAC
TTCATCGCGCCGCCGCGTTACACCGGTCGTCAATTTGATGAGGTGCCGAGCCACCTGCTGAGCGCGTTCG
TTCTGGGTAGCCCGCTGCAGAGCTATATCATTGACGATGCGGCGATTTGCCACCAAGTTACCGCGGTTAG
CTGCAGCCCGGGTGCGAGCCATGGTACCCGTACCGTGATCCTGATTGATTTCAGCCCGCTGAACCCGGCG
AGCAGCGCGCAGCCGACCTGA

For the construction of expression vector, the above sequence was inserted between the NdeI and XhoI restriction sites within the commercial pET28a(+) vector.

Construction of homology model for GlbB

A structure of PF10014 member from M. petroleiphilum has been reported (PDB ID: 3PL0, Uniprot ID: A2SJH7) by Xu *et al*. This structure served as a template for a homology model of GlbB, which was constructed using the Swiss-Model program.

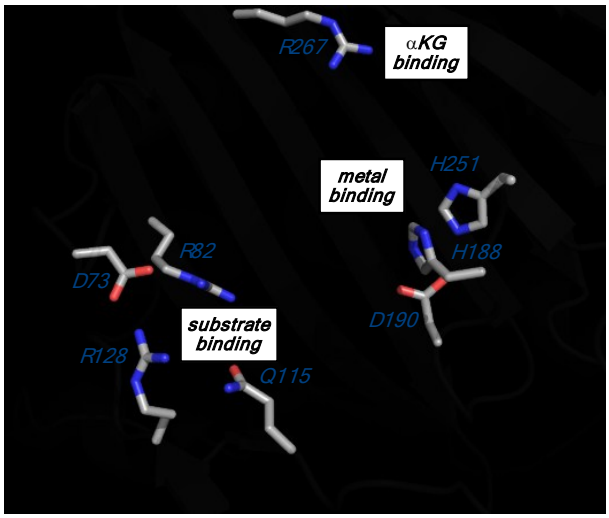


Figure S1. Homology model of GlbB, showing potential sites for substrate binding, metal binding, and αKG binding.

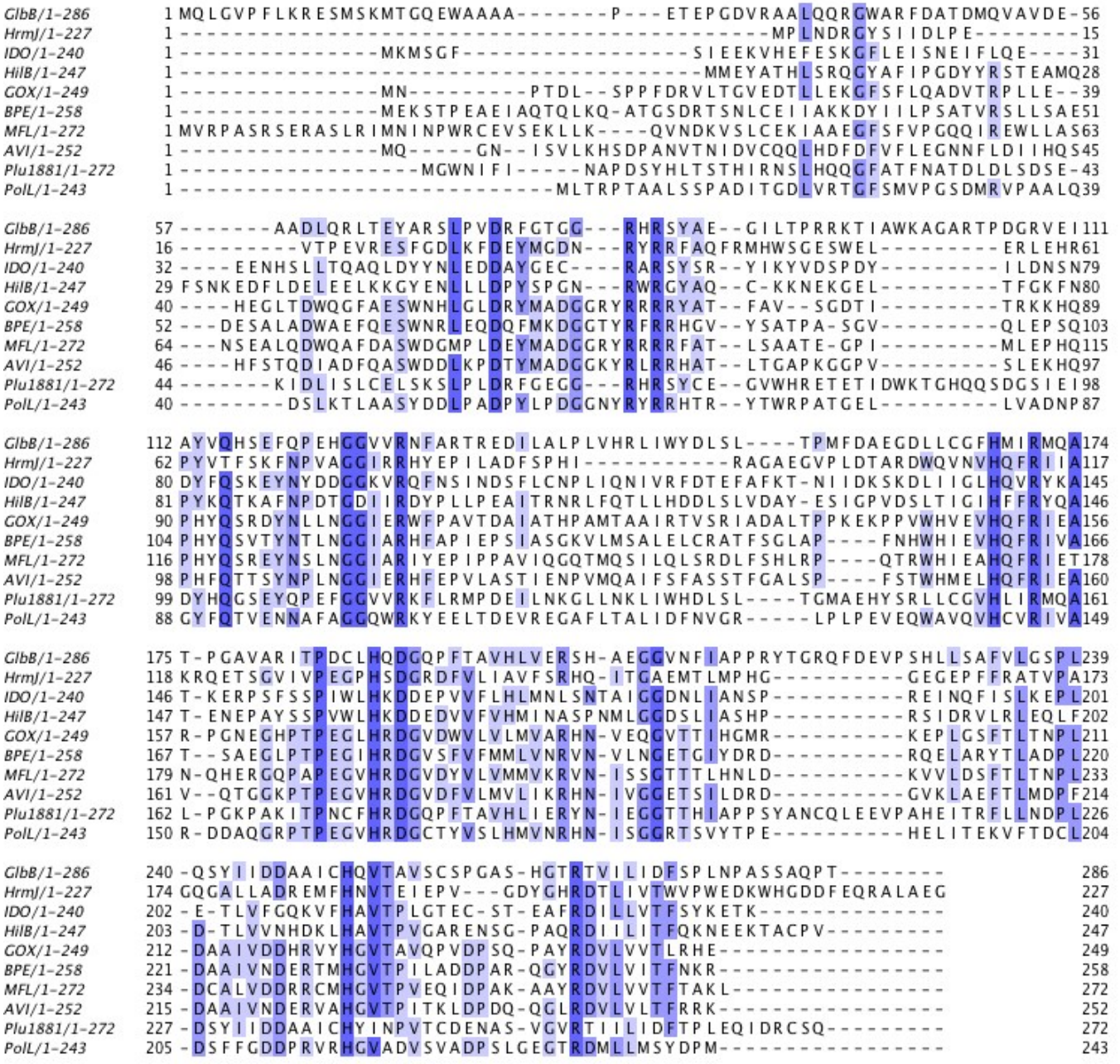


Figure S2. Multiple sequence alignment of select members of PF10014 using Clustal Omega algorithm. Uniprot accession codes: A8KCI8 (GlbB), F8S6W1 (HrmJ), E2GIN1 (IDO), A0A0H3L116 (HilB), Q5FQD2 (GOX), A9IE89 (BPE), Q1GXZ5 (MFL), Q0GPY6 (AVI), Q7N5R2 (Plu1881), C1IC27 (PolL).

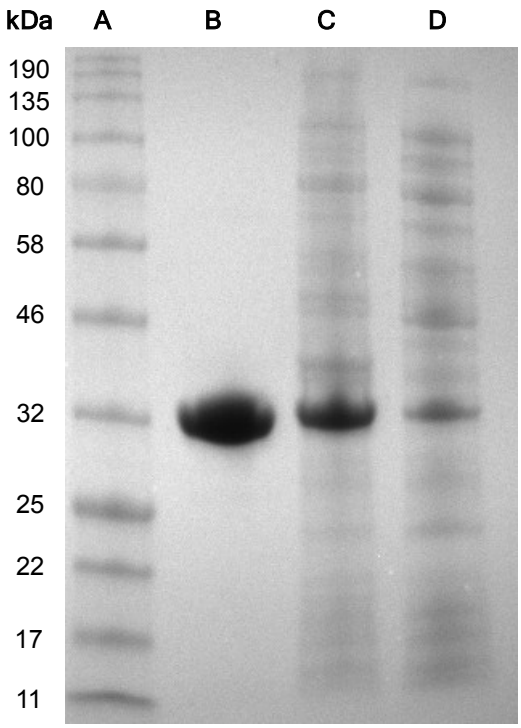


Figure S3. SDS PAGE gel analysis of purified GlbB (**B**), insoluble fraction (**C**), and soluble fraction (**D**). Molecular weight of N-His-tagged GlbB = 33.6 kDa.

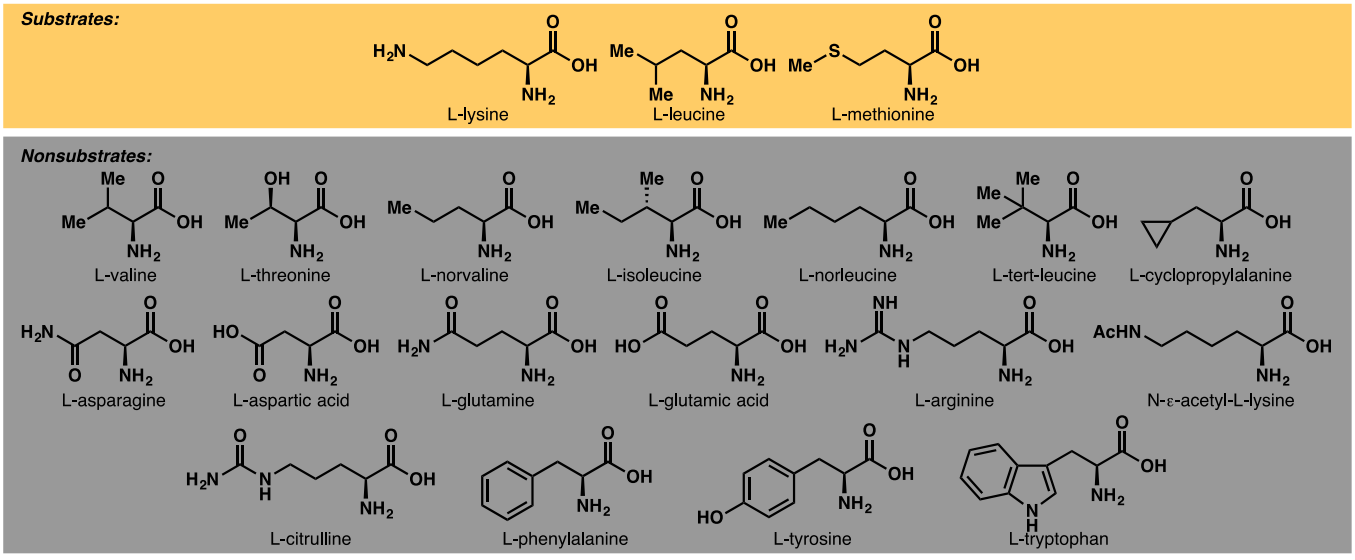


Figure S4. List of amino acid substrates that were tested for hydroxylation with GlbB. Amino acids accepted as substrates by GlbB are highlighted in orange, and amino acids not accepted as substrates are highlighted in grey.

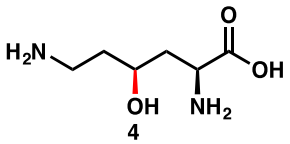
General procedure for amino acid purification with Dowex 50WX8 cation exchange resin

Dowex 50WX8 resin (20 mg/1 μmol of amino acid) was slurry-packed with 1 M NH_4OH in a flash chromatography column, and then washed with ~1 column volume (CV) of 1 M NH_4OH, and then ~ 2 CV of H_2O. Next, the resin was washed with 1 M HCl until pH = 1, and then with H_2O until pH = 6–7. The acidified crude reaction mixture was loaded directly onto the column, washing with H_2O until pH = 6–7, and with 1 M NH_4OH until product was no longer eluting from the column. Amino acid containing fractions were pooled, concentrated *in vacuo*, and then lyophilized.

Hydroxylation of various amino acids with purified GlbB

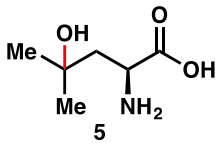
General procedure

A 20 ml scintillation vial was charged with the appropriate amino acid (60 μmol, 1.0 equiv, 20 mM final concentration), L-ascorbic acid (30 μmol, 0.5 equiv, 10 mM final concentration), and α-ketoglutaric acid (disodium salt dihydrate, 150 μmol, 2.5 equiv, 50 mM final concentration). 50 mM kPi buffer was added to the vial (pH 8.0, 3.0 mL), followed by 15 μL of $FeSO_4$ solution in H_2O (200 mM, 0.05 equiv, 1 mM final concentration). The reaction was started by the addition of GlbB stock solution (final concentration = 2 μM, 0.0001 equiv or 30 μM, 0.0015 equiv) and shaken for 12 h at 20 ºC, 250 rpm under air. After quenching with 300 μL of 1 M HCl, the crude reaction mixture was centrifuged (15,000 rpm, 5 min) and the supernatant was submitted directly to ion exchange purification on Dowex 50WX8 resin. Yields of hydroxylation products were determined using ^{1}H NMR analysis employing ethylene carbonate as an internal standard.



4-(OH)-lysine (4)

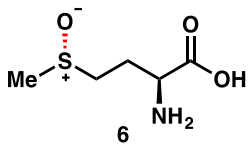
Biotransformation was performed using general procedure with 0.0001 equiv GlbB. Reaction afforded 59% yield (NMR) as shown in Figure S3. Full characterization of **4** is provided in the experimental procedure for preparative scale reaction.



5

4-(OH)-leucine (5)

Biotransformation was performed using general procedure with 0.0015 equiv GlbB. Reaction afforded 46% yield (NMR) as shown in Figure S4. NMR spectra of the desired product match those reported in literature.[S1]



6

L-methionine sulfoxide (6)

Biotransformation was performed using general procedure with 0.0015 equiv GlbB. Reaction afforded 49% yield (NMR) as shown in Figure S5. NMR spectra of the desired product match those of commercially available standard. Stereochemical assignment of the sulfoxide product was performed based on reported ^{1}H NMR of the diastereomers.[S2]

Determination of Michaelis-Menten parameters

For L-lysine: Kinetic parameters were determined by incubating 0.2 μM GlbB with 30 mM αKG, 10 mM L-ascorbic acid, 1 mM $FeSO_4$ and various concentrations of the amino acid substrates (0.10 mM to 0.50 mM) in 50 mM kPi buffer (pH 8.0, 2.5 mL total reaction volume). After 10 s, a 200 μL aliquot of the reaction was sampled and quenched with 100 μL of 12.5 mM Fmoc-OSu solution and 50 μL of sat. aq. $NaHCO_3$. After 15 min of incubation, 50 μL of 15 mM $TsNH_2$ was added as an internal standard for quantitation and the samples were centrifuged. The supernatants were concentrated using SpeedVac, resuspended in 100 μL of 1:1 MeCN:H_2O mixture, and analyzed by LCMS using the extracted ion chromatogram (EIC) mode, calibrating against standard curves made independently from Fmoc-derivatized product standards and $TsNH_2$. Apparent kinetic parameters were calculated from a nonlinear fit of the initial velocity data to the Michaelis-Menten equation with the KaleidaGraph data analysis software (Synergy). The error bars in the plots represent the standard deviations calculated from replicate measurements and the errors for the fitted parameters are standard errors for the fits.

For L-leucine and L-methionine: Kinetic parameters were determined by incubating 10 μM GlbB with 30 mM αKG, 10 mM L-ascorbic acid, 1 mM $FeSO_4$ and various concentrations of the amino acid substrates (0.25 mM to 22 mM) in 50 mM kPi buffer (pH 8.0, 2.5 mL total reaction volume). At various time intervals, 200 μL aliquots of the reactions were sampled and quenched with 100 μL of 50 mM Fmoc-OSu solution and 50 μL of sat. aq. $NaHCO_3$. After 15 min of incubation, 50 μL of 30 mM $TsNH_2$ was added as an internal standard for quantitation. The samples were centrifuged at 15,000 rpm for 3 min and analyzed by LCMS using the extracted ion chromatogram (EIC) mode, calibrating against standard curves made independently from Fmoc-derivatized product standards and $TsNH_2$. Kinetic parameters were calculated as described above.

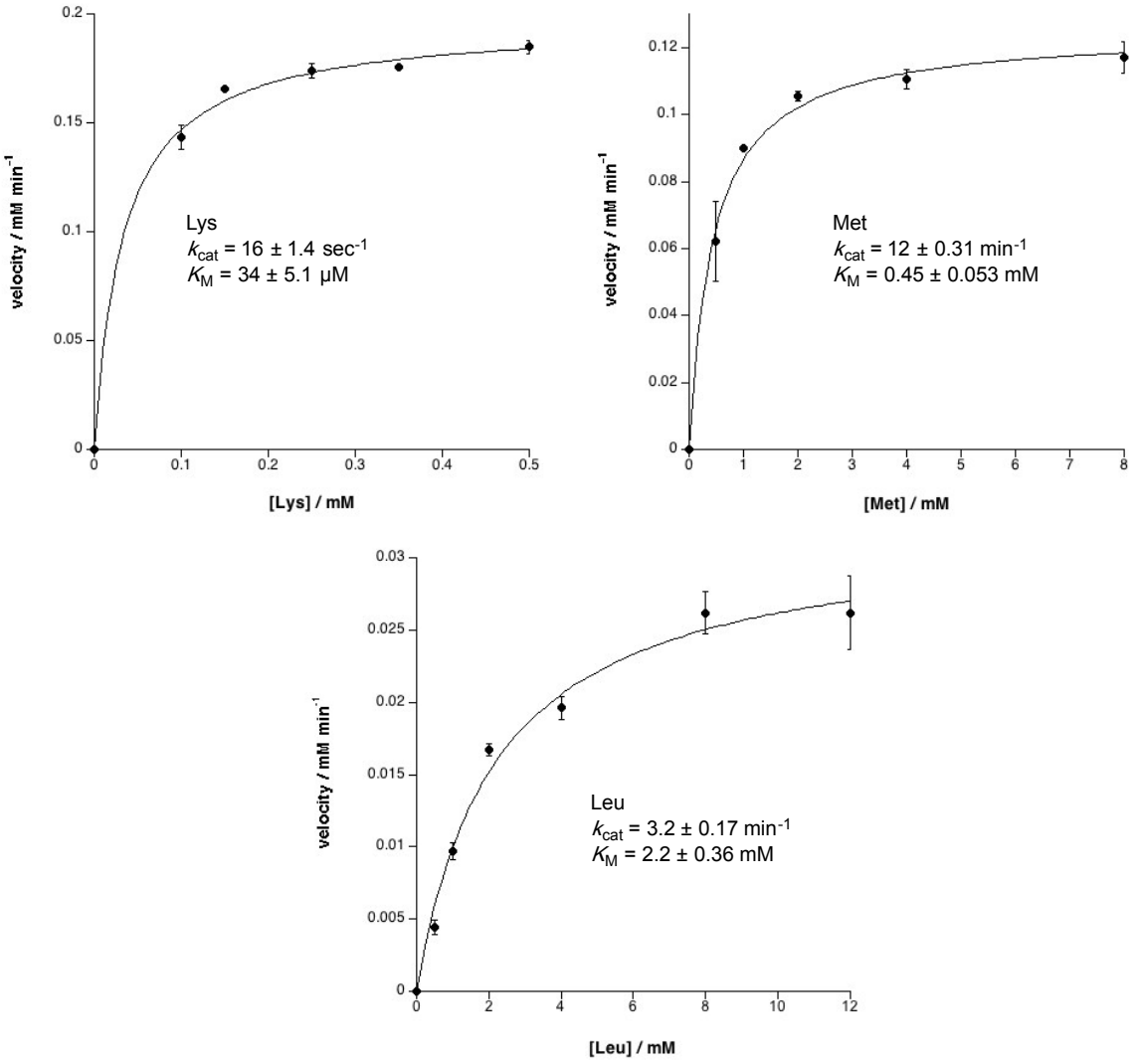
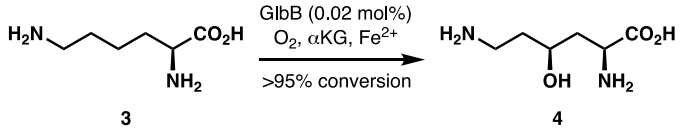


Figure S5. Michaelis-Menten plots of various amino acid hydroxylations with GlbB.

Experimental procedure for the synthesis of protected dipeptide 9



Preparative-scale lysine hydroxylation with purified GlbB

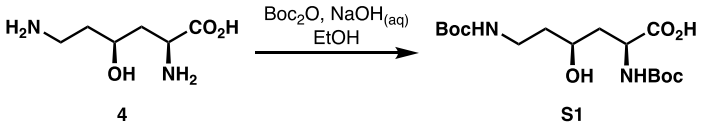
A 125 mL Erlenmeyer flask was charged with L-lysine (88 mg, 0.60 mmol, 1.0 equiv, 20 mM final concentration), α-ketoglutaric acid (disodium salt dihydrate, 339 mg, 1.50 mmol, 2.5 equiv, 50 mM final concentration) and L-ascorbic acid (53 mg, 0.30 mmol, 0.5 equiv, 10 mM final concentration). 50mM kPi buffer was added to the flask (pH 8, 30 mL), followed by 150 µL of $FeSO_4$ solution in H_2O (200 mM, 0.05 equiv, 1 mM final concentration). The reaction was started by addition of GlbB stock solution (200 µL of 600 µM stock, 0.0002 equiv, final concentration 4 µM) and set shaking for 12 h at 20 °C, 200 rpm under air. Upon completion, the reaction was acidified to pH 2 with 1 M HCl and centrifuged (4200 rpm, 15 min). After evaporation of the solvent, the crude product was carried forward to the next reaction without further purification. For characterization purposes, the crude product was purified by ion exchange chromatography with Dowex 50WX8 resin.

^{1}H NMR (600 MHz, D_2O w/ 1% TFA): δ 4.09–3.99 (m, 1H), 3.84 (dd, *J* = 7.9, 5.5 Hz, 1H), 3.21–3.08 (m, 2H), 2.18–2.09 (m, 1H), 1.97–1.91 (m, 1H), 1.91–1.80 (m, 2H).

^{13}C NMR (151 MHz, D_2O w/ 1% TFA): δ 174.4, 67.7, 53.5, 37.1, 36.6, 34.0.

$[\alpha]_D$ = 3.7° (c = 0.5, H_2O)

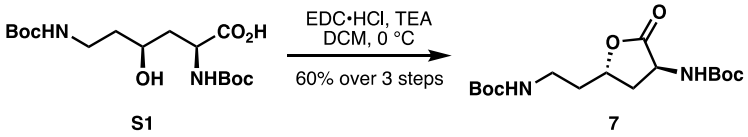
HRMS (ESI-TOF): calculated for $C_6H_{15}N_2O_3H^+$ ($[M+H]^+$) 163.1083, found 163.1078



Boc protection of 4-hydroxylysine

A solution of **4** (0.62 mmol, 1 equiv, 0.2 M final concentration) in 2.07 mL H_2O was basified to pH 11 with aqueous NaOH. After addition of a solution of Boc_2O (541 mg, 2.48 mmol, 4.0 equiv) in 1.03 mL EtOH, the reaction was stirred overnight at room temperature. EtOH was removed from the mixture *in vacuo* at 40 °C. The aqueous layer was acidified to pH 1 with HCl then extracted with ethyl acetate (3 x

10 mL). The combined organic layers were washed with brine, dried over Na_2SO_4 and concentrated to give a crude mixture of di-Boc 4-hydroxylysine **S1** and the corresponding γ-lactone **7**.



Lactonization of protected 4-hydroxylsyine

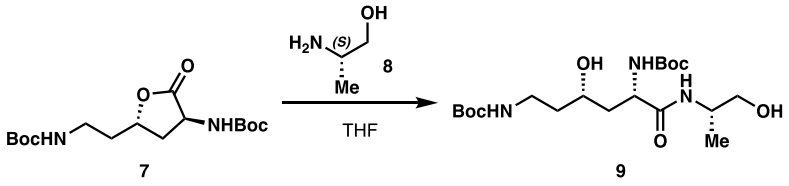
The crude material obtained from the above step was dissolved in dry DCM (3.10 mL, 0.2 M final concentration) and cooled to 0 °C. To this solution was added EDC•HCl (120 mg, 0.62 mmol, 1.0 equiv) as a solid, followed by triethylamine (86 µL, 0.62 mmol, 1.0 equiv). The reaction was stirred at 0 °C under argon for 2 hours or until completion as judged by TLC. The reaction was concentrated to dryness *in vacuo* at 40 °C. The residue was then dissolved in ethyl acetate (20 mL), washed with water, brine, and dried over Na_2SO_4. Purification by silica flash chromatography (60:40 hexanes:ethyl acetate) provided **7** as a white solid in a 9:1 mixture of rotamers (129 mg, 60% over 3 steps from L-lysine).

^{1}H NMR (400 MHz, DMSO-d_6): δ 7.45 (major rotamer, d, J = 8.3 Hz, 1H), 6.89 (major rotamer, t, J = 5.5 Hz, 1H), 4.56 (tdd, J = 8.1, 5.5, 3.7 Hz, 1H), 4.28 (major rotamer, q, J = 9.2 Hz, 1H), 3.08–2.89 (m, 2H), 2.31–2.09 (m, 2H), 1.82–1.60 (m, 2H), 1.38 (s, 9H), 1.37 (s, 9H).

^{13}C NMR (101 MHz, DMSO-d_6): δ 175.6, 156.0, 155.5, 79.0, 78.1, 75.8, 48.9, 36.9, 35.6, 33.0, 28.7, 28.6.

$[\alpha]_D$ = 26.7° (c = 0.5, $CHCl_3$)

HRMS (ESI-TOF): calculated for $C_{16}H_{28}N_2O_6H^+$ ($[M+H]^+$) 345.2026, found 345.2026



*Synthesis of dipeptide **9***

Lactone **7** (200 mg, 0.58 mmol, 1 equiv) was dissolved in anhydrous THF in a flame-dried flask. After addition of L-alaninol **8** (218 mg, 2.90 mmol, 5.0 equiv), the reaction mixture was allowed to stir overnight under argon at room temperature. Upon completion by TLC, the reaction mixture was

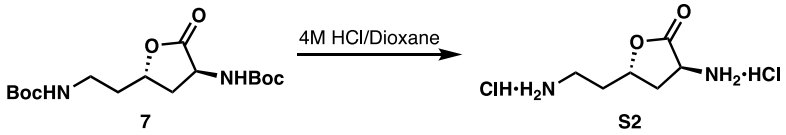
concentrated to dryness and filtered through a plug of silica in ethyl acetate to provide **9** as a white solid (223 mg, 92% yield)

^{1}H NMR (400 MHz, Methanol-d_4): δ 4.12 (t, *J* = 7.2 Hz, 1H), 4.00–3.87 (m, 1H), 3.74–3.61 (m, 1H), 3.48 (d, *J* = 5.4 Hz, 2H), 3.15 (t, *J* = 6.9 Hz, 2H), 1.93–1.76 (m, 1H), 1.74–1.62 (m, 2H), 1.62–1.49 (m, 2H), 1.44 (s, 9H), 1.43 (s, 9H), 1.15 (d, *J* = 6.8 Hz, 3H).

^{13}C NMR (101 MHz, Methanol-d_4): δ 174.4, 158.7, 157.7, 80.7, 80.0, 67.4, 66.0, 54.0, 48.5, 40.8, 38.4, 38.1, 28.8, 28.7, 17.0.

$[\alpha]_D$ = -14.1° (c = 0.3, $CHCl_3$)

HRMS (ESI-TOF): calculated for $C_{19}H_{37}N_3O_7H^+$ ($[M+H]^+$) 420.2710, found 420.2713



Deprotection of lactone 7

Lactone **7** (25 mg, 0.072 mmol) was stirred in an anhydrous 4M HCl/dioxane solution (0.80 mL) at room temperature. Upon completion by TLC, the reaction was concentrated to dryness to provide the fully deprotected lactone **S2** which was used for confirmation of stereochemistry by ^{1}H NOESY spectroscopy.

^{1}H NMR (400 MHz, DMSO-d_6): δ 9.03 (brs, 2H), 8.29 (brs, 2H), 4.87 (tdd, *J* = 8.5, 5.2, 2.6 Hz, 1H), 4.42 (t, *J* = 9.6 Hz, 1H), 2.85 (t, *J* = 7.5 Hz, 2H), 2.56–2.45 (m, 1H), 2.36 (ddd, *J* = 12.8, 9.5, 2.7 Hz, 1H), 2.11 – 1.91 (m, 2H).

References

S1. Smirnov, S. V.; Sokolov, P. M.; Kodera, T.; Sugiyama, M.; Hibi, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. *FEMS Microbiol. Lett*. **2012**, *331*, 97–104.

S2. Marzag, H.; Schuler, M.; Tatibouët, A.; Reboul, V. *Eur. J. Org. Chem.* **2017**, 896–900.

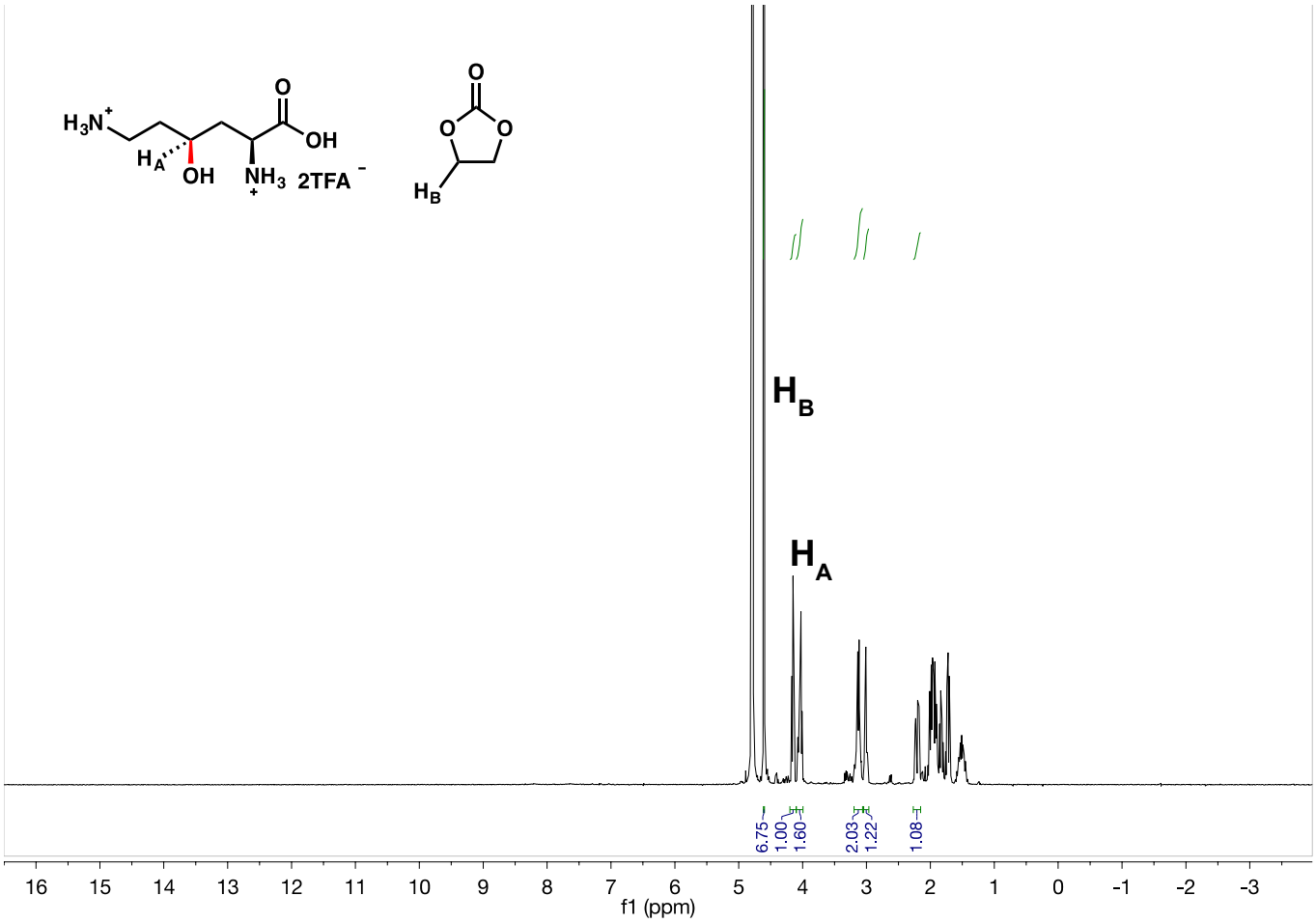


Figure S6. 20 mM L-lysine hydroxylation with 20 mM ethylene carbonate as internal standard (400 MHz, D_2O w/ 1% TFA).

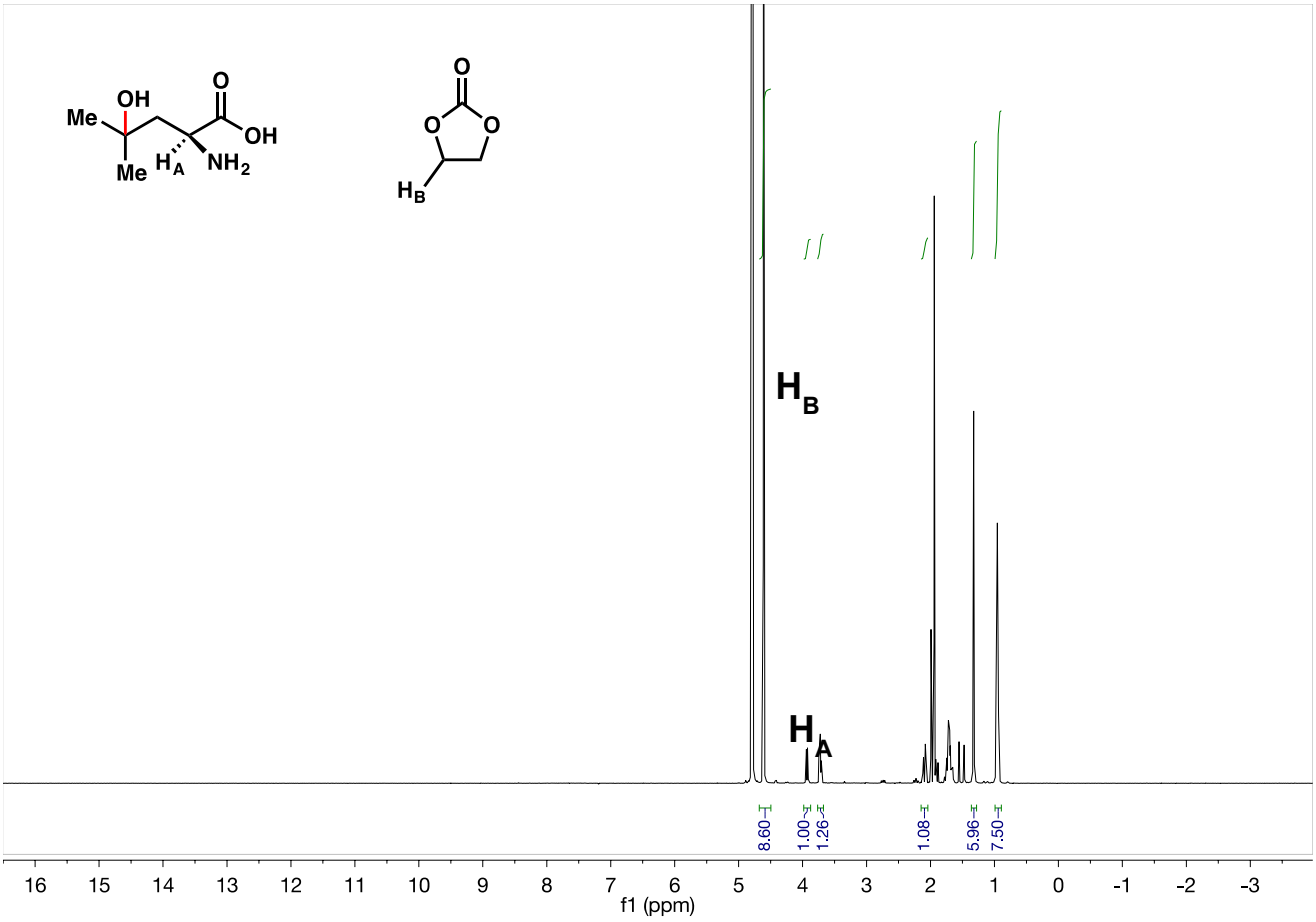


Figure S7. 20 mM L-leucine hydroxylation with 20 mM ethylene carbonate as internal standard (400 MHz, D_2O).

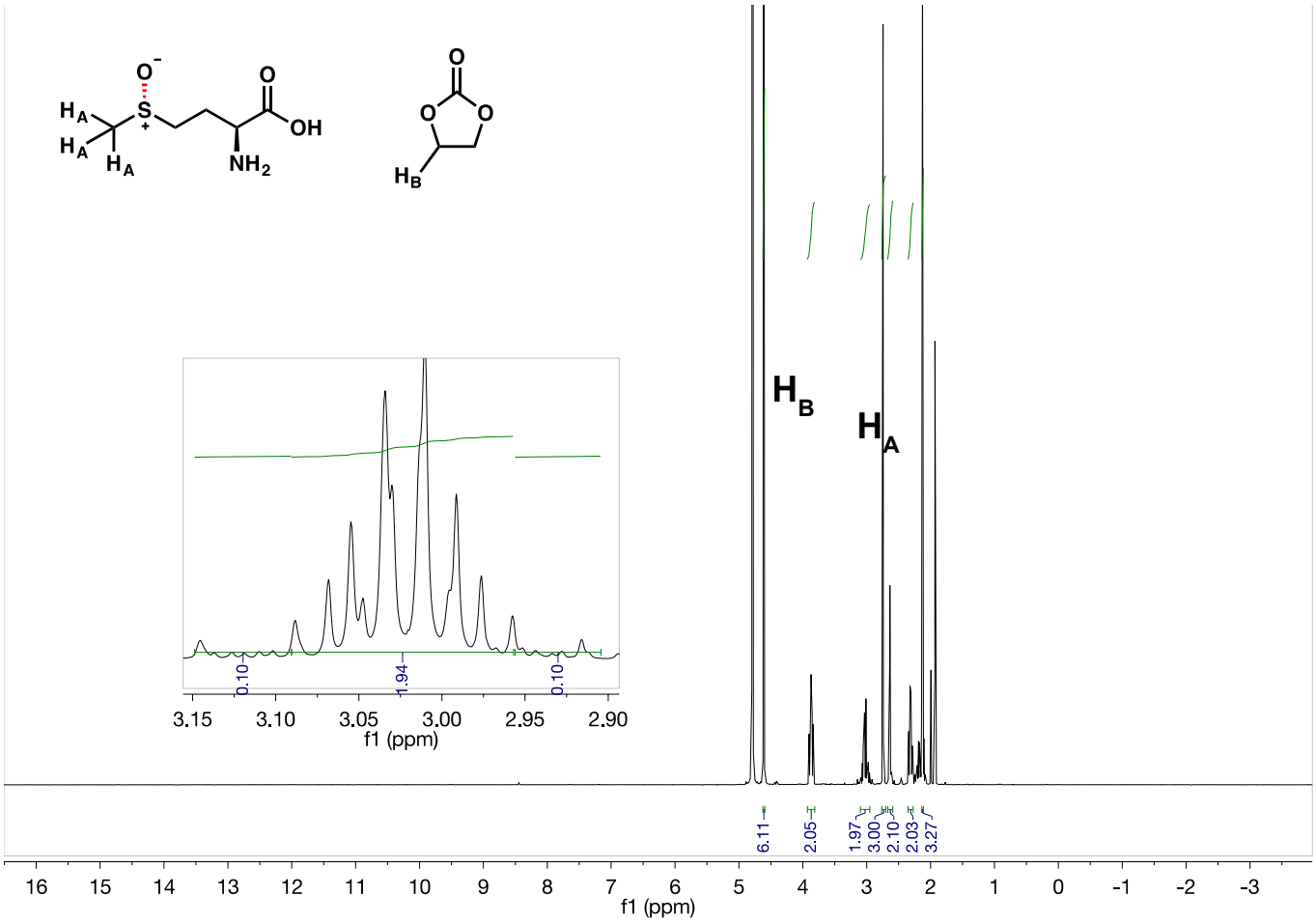


Figure S8. 20 mM L-methionine hydroxylation with 15 mM ethylene carbonate as internal standard (400 MHz, D_2O). Inset shows ^1H NMR analysis of the diastereoselectivity of the reaction.

Spectral Data

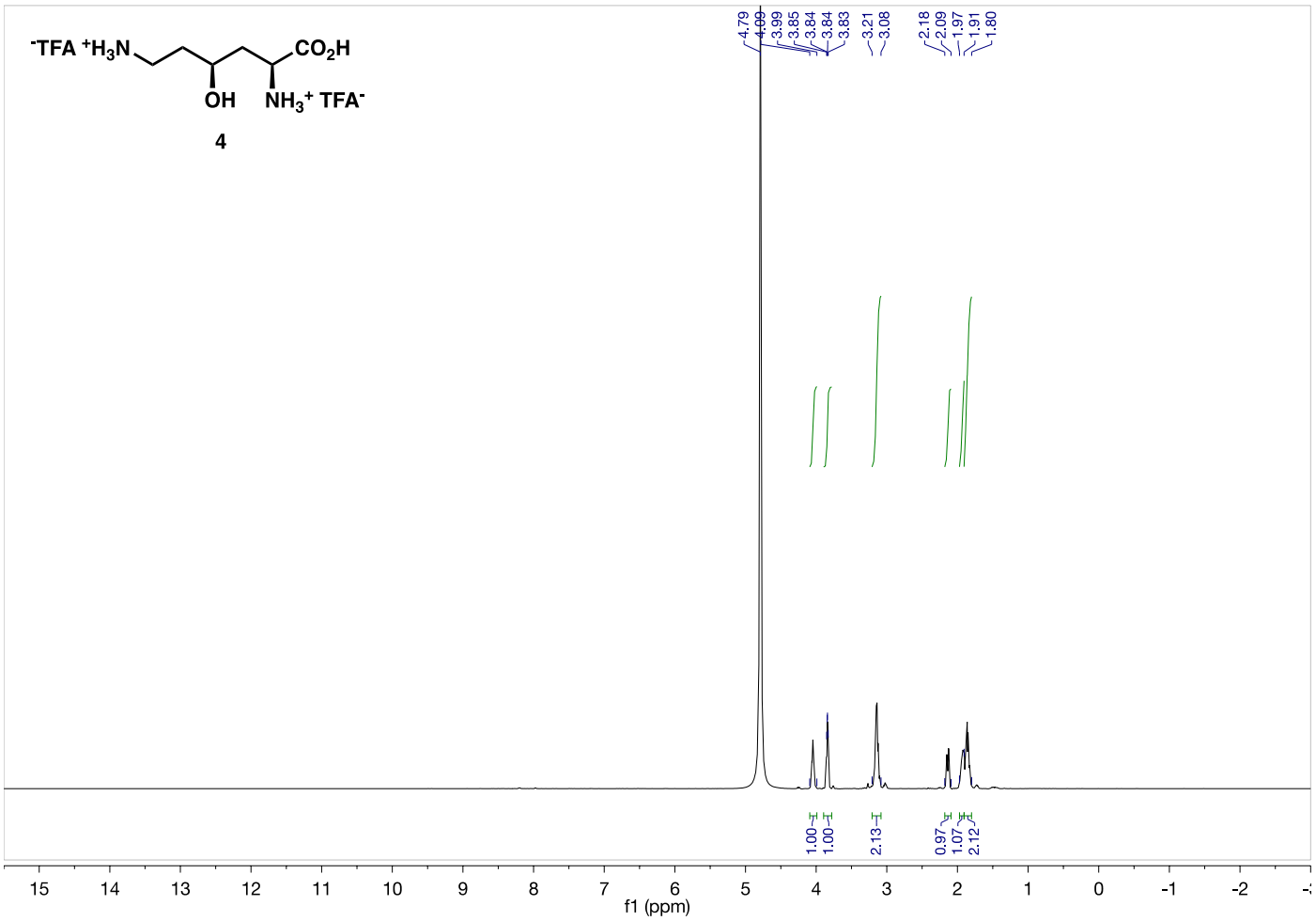


Figure S9. ^{1}H NMR of **4** (600 MHz, D_2O w/ 1% TFA)

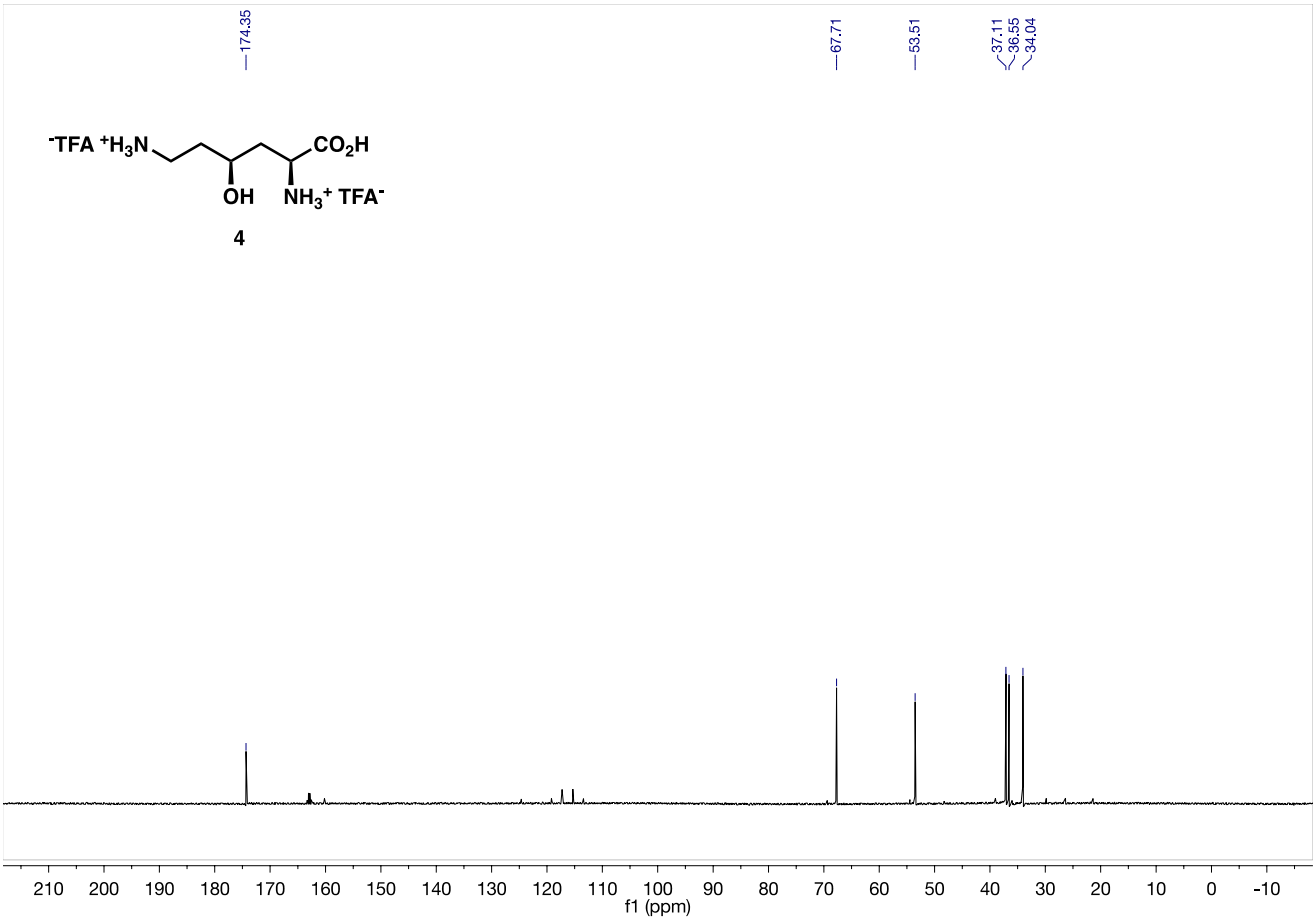


Figure S10. ^{13}C NMR of **4** (151 MHz, D_2O w/ 1% TFA)

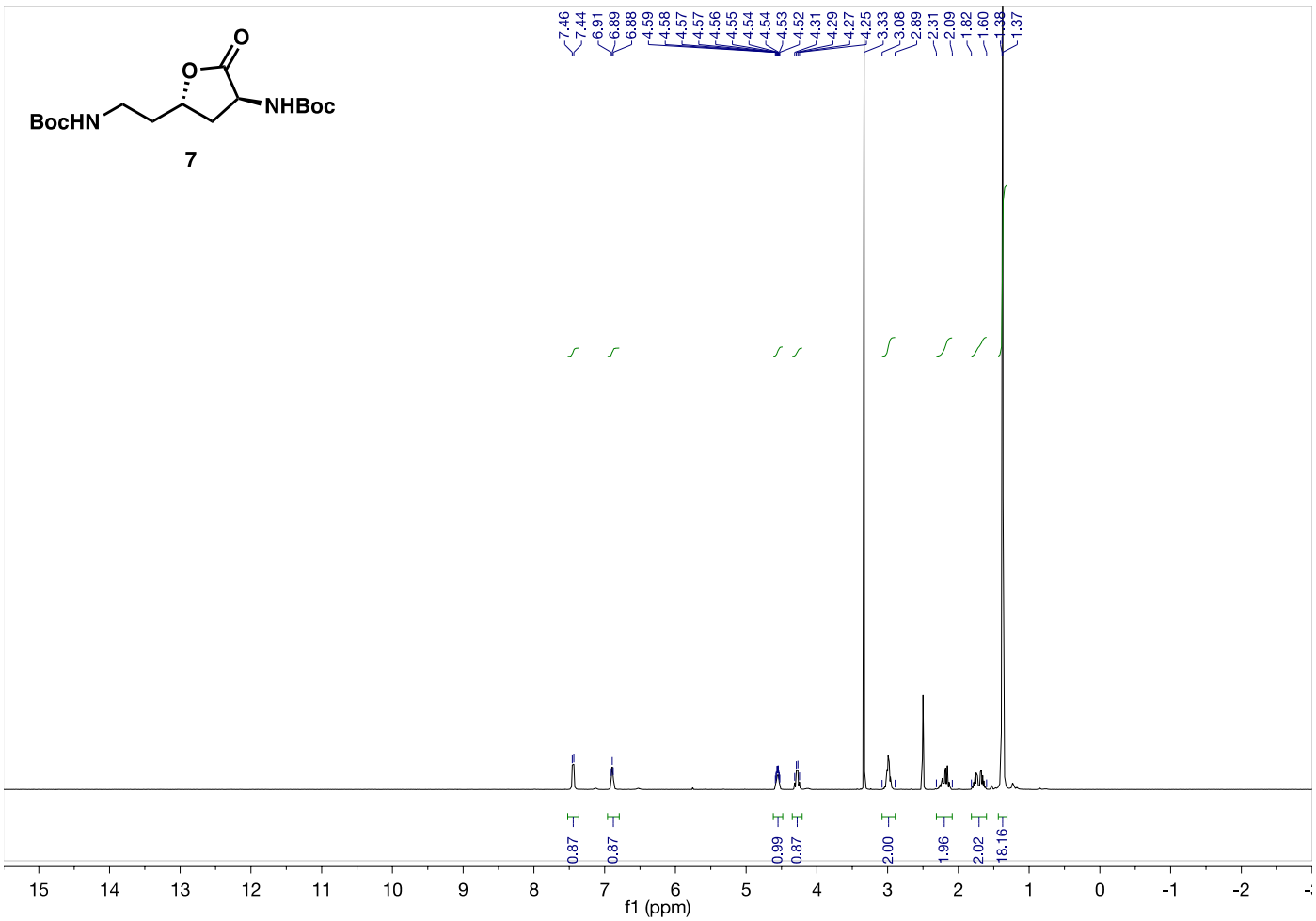


Figure S11. ^{1}H NMR of **7** (400 MHz, DMSO-d_6)

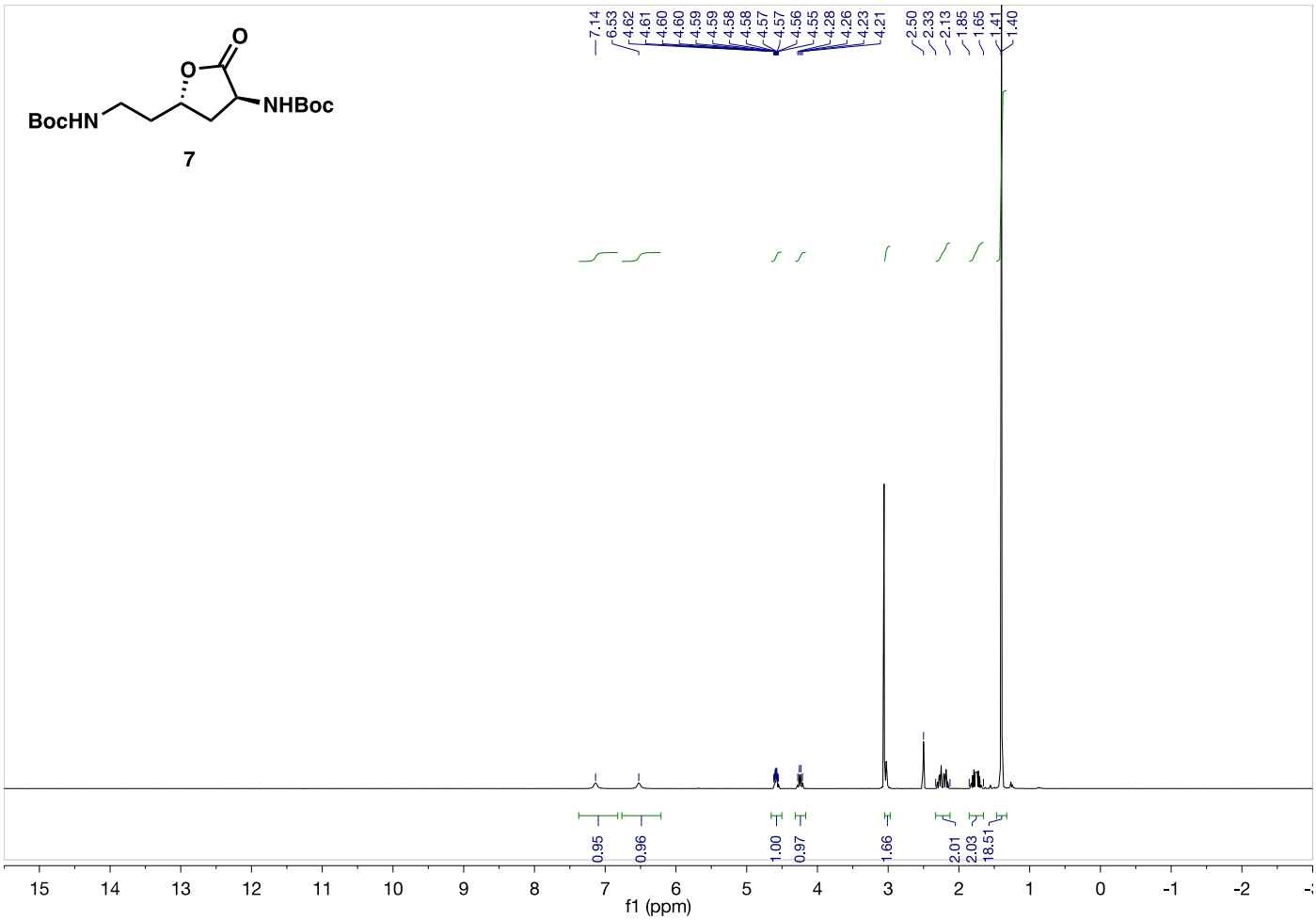


Figure S12. ^{1}H NMR of **7** (400 MHz, DMSO-d_6, 80 °C)

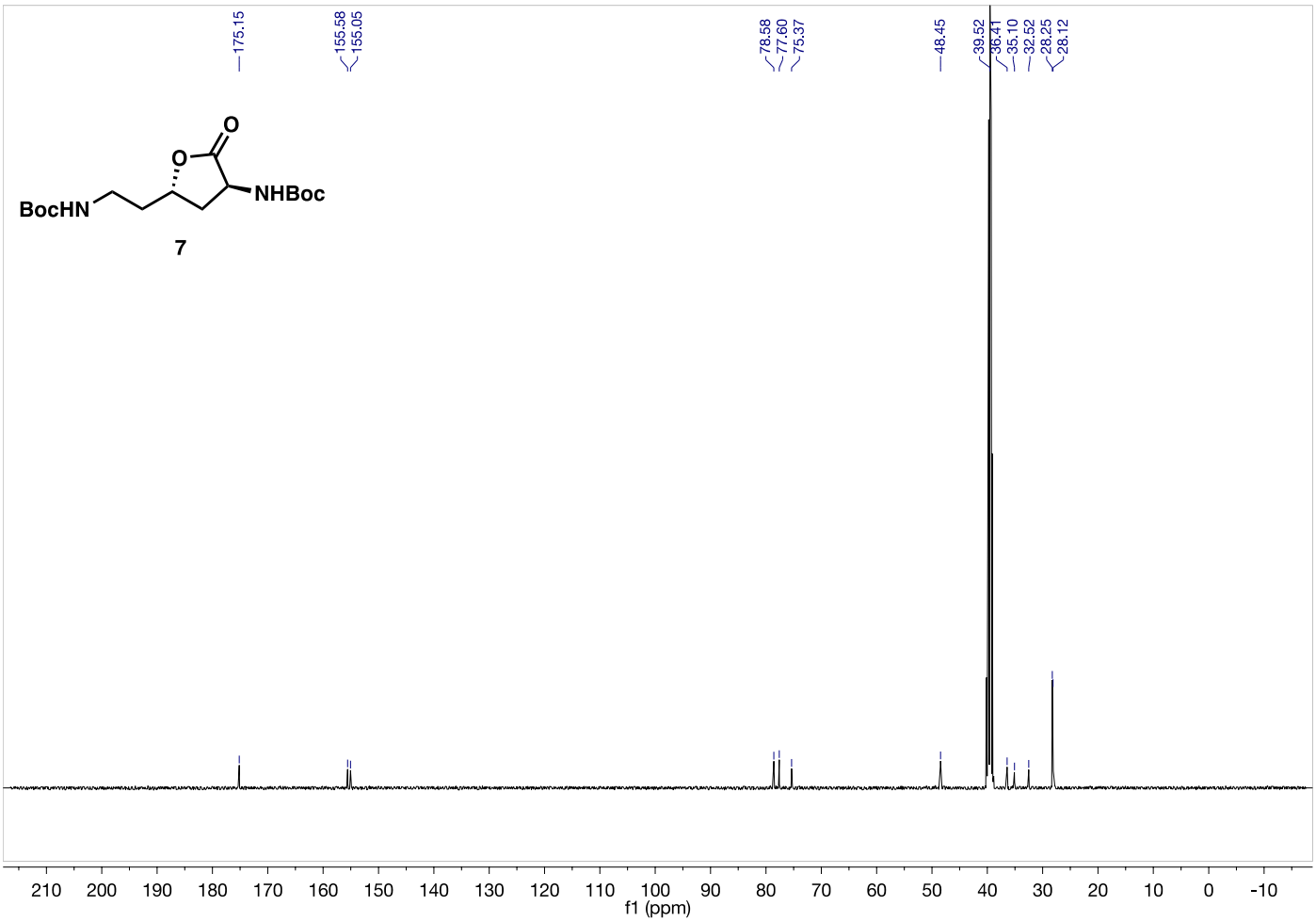


Figure S13. ^{13}C NMR of **7** (101 MHz, DMSO-d_6)

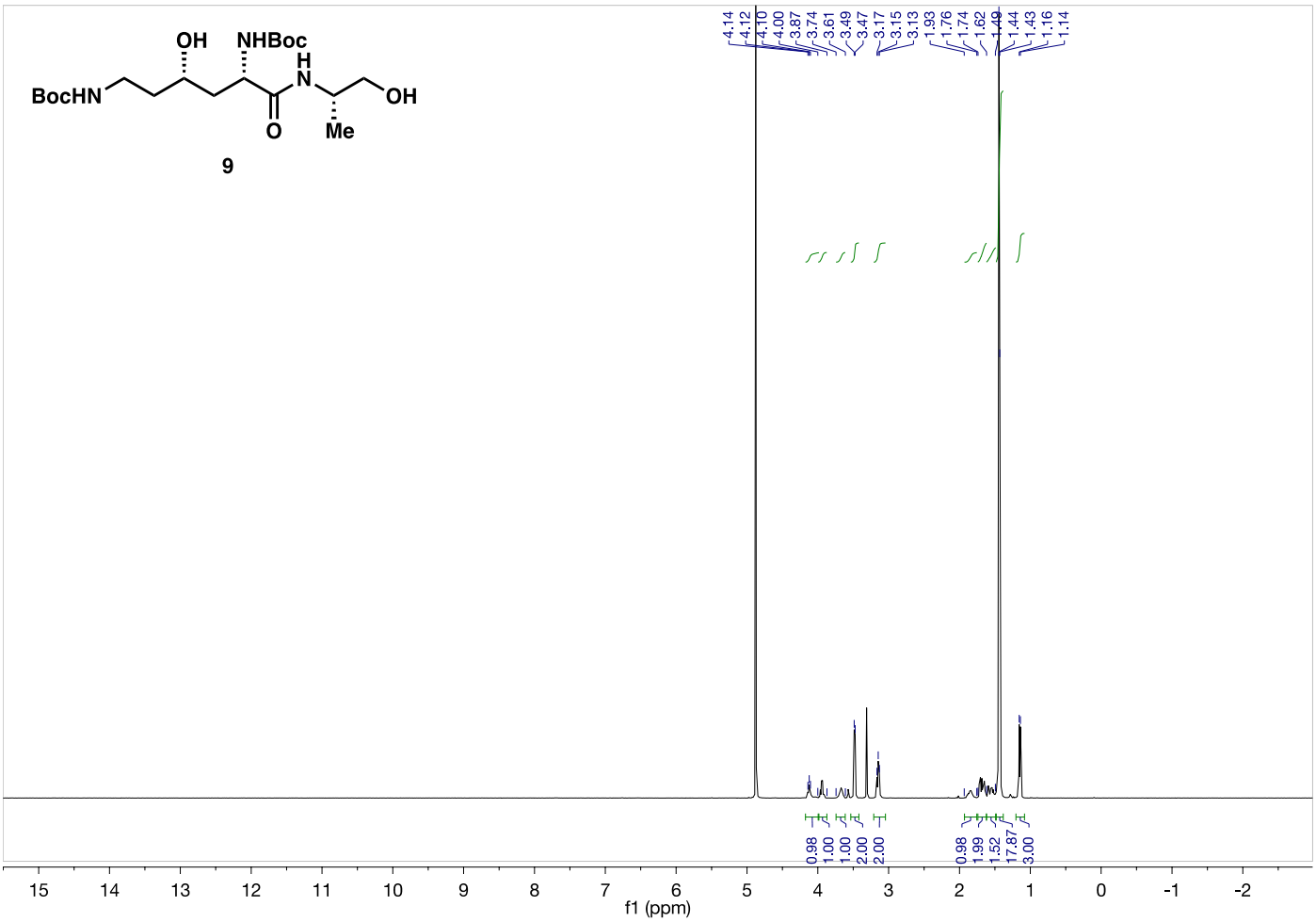


Figure S14. ^{1}H NMR of **9** (400 MHz, Methanol-d_4)

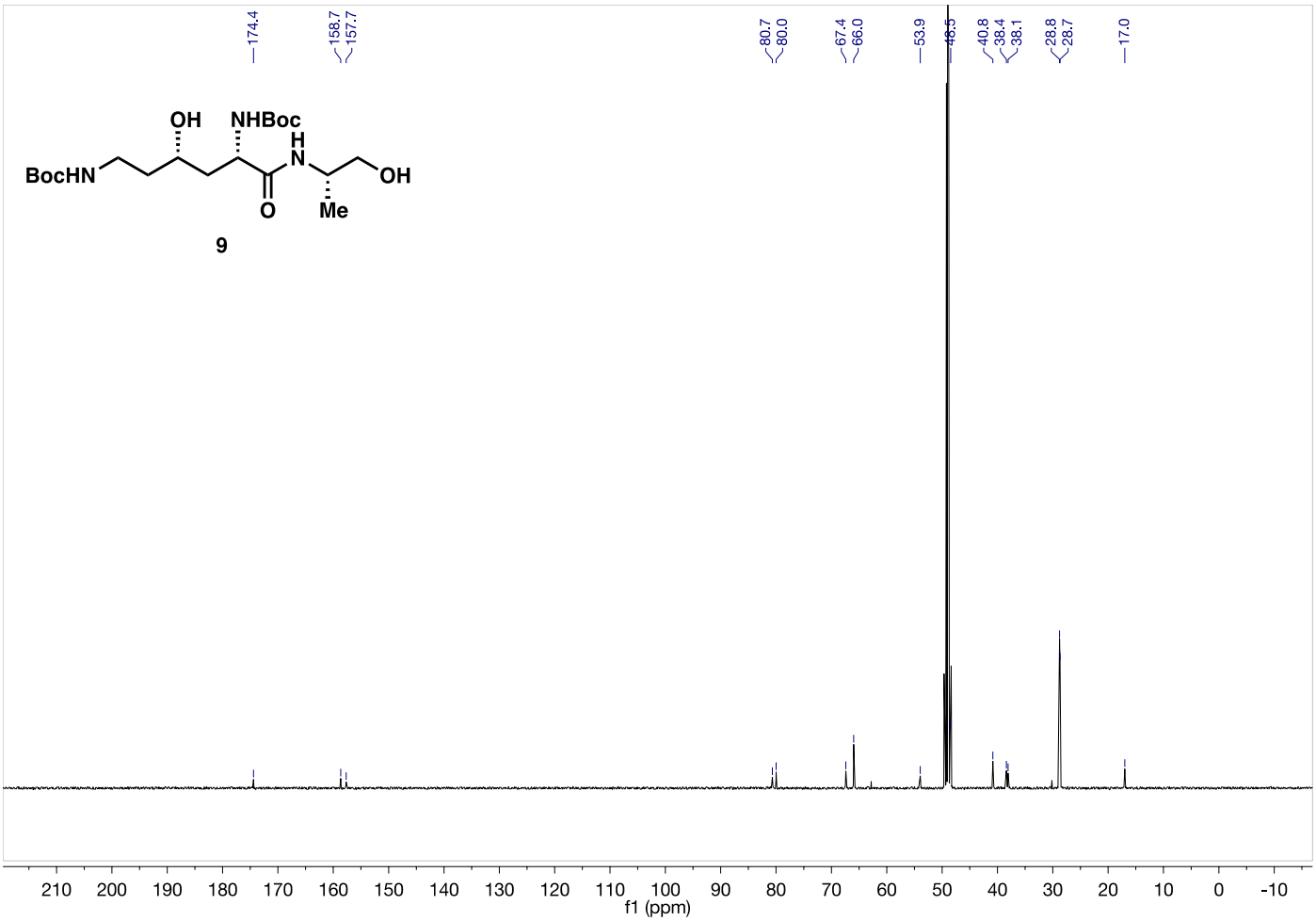


Figure S15. ^{13}C NMR of **9** (101 MHz, Methanol-d_4)

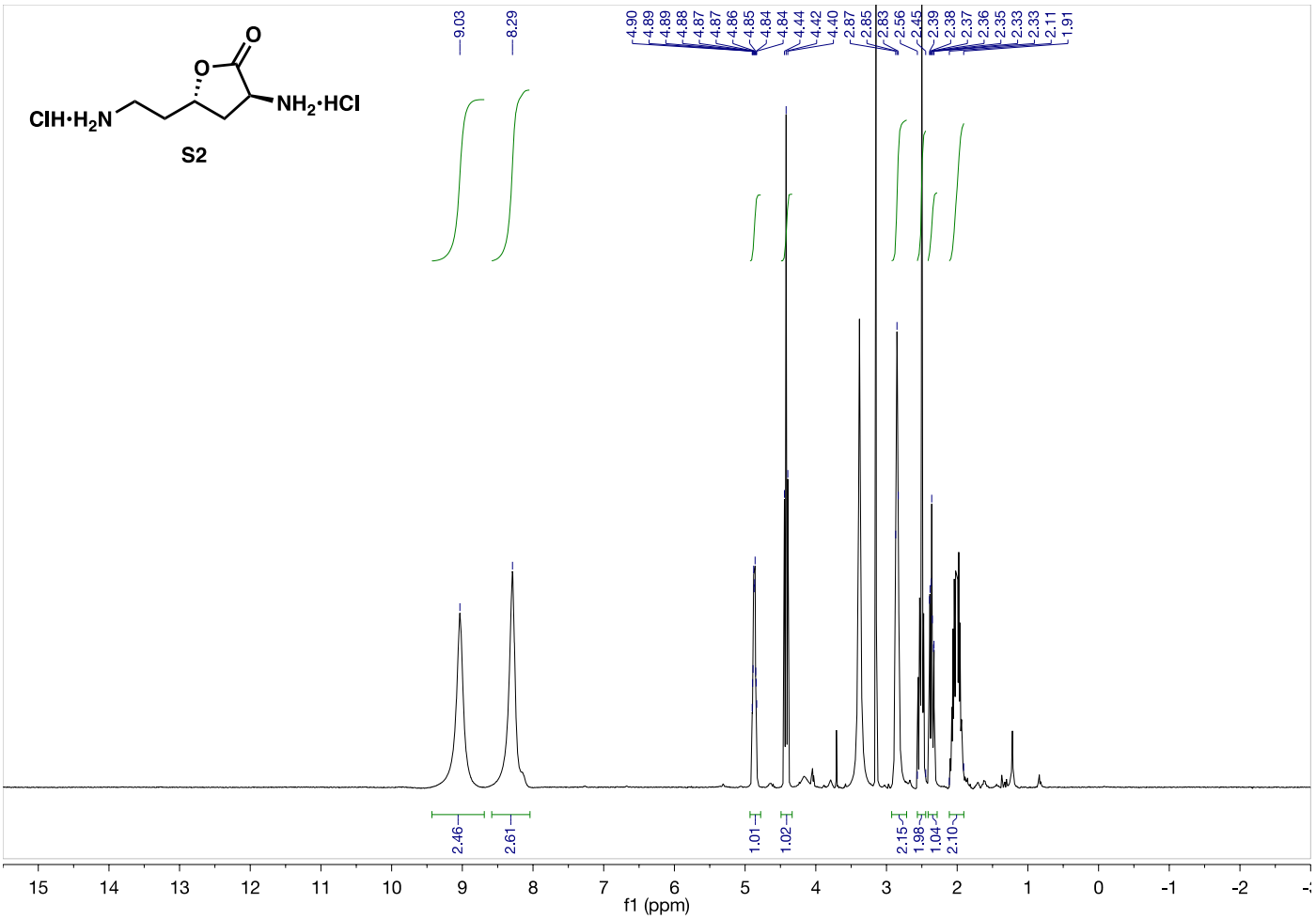


Figure S16. ^{1}H NMR of **S2** (400 MHz, DMSO-d_6)

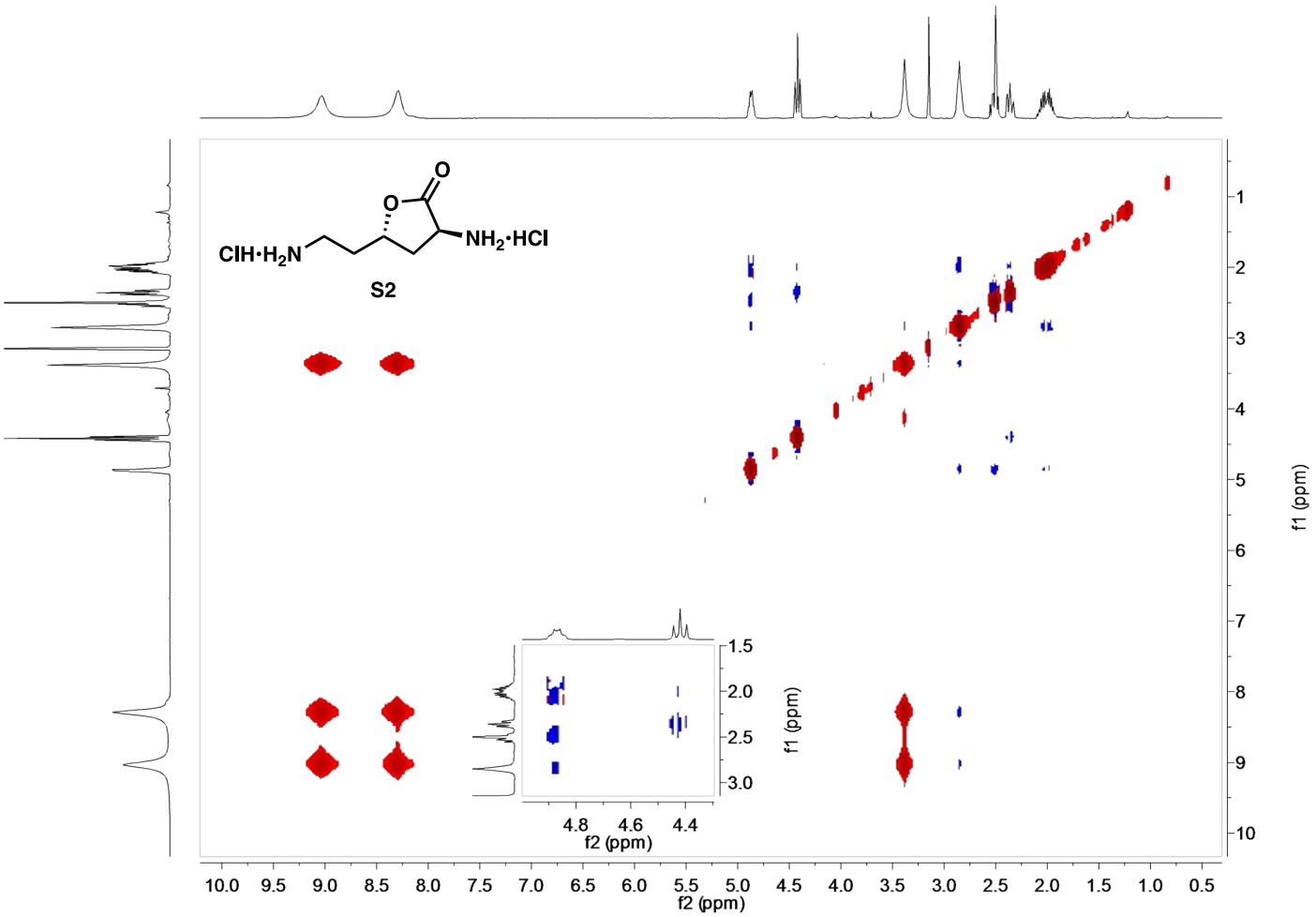


Figure S17. 2D NOESY ^{1}H NMR of **S2** (400 MHz, DMSO-d_6)