

Synthesis of acyl-amide natural products using a versatile natural product biocatalyst

Supplementary Information

Piera M. Marchetti, Shona M. Richardson, Noor M. Kariem, and Dominic J.
Campopiano*

Materials and Methods:

TamA ANL Cloning and Expression Test

TamA was cloned from *Pseudoalteromonas tunicata* D2 genomic DNA. The ACP domain was deleted from the adenylation domain completely and the TamA ANL was cloned from the TamA plasmid into the pETHisTev vector (cite) using NcoI and XhoI restriction sites to yield recombinant TamA ANL with a Tobacco Etch Virus (TEV) protease cleavable N-terminal 6xHis-tag. Several constructs of differing lengths in pEHISTEV were produced (576, 581, 586 and 591 residues). The constructs were transformed into competent *E. coli* BL21 (DE3) cells. Test expressions were carried out on a small scale to determine optimal conditions of expression prior to large scale expression. A single colony of *E. coli* cells containing an expression plasmid was inoculated in 10 mL LB medium with Kanamycin and grown for ~18 h at 37 °C with orbital shaking. The optical density of the cells at 600 nm (OD₆₀₀) was measured and this culture added to 250 mL LB with appropriate antibiotic for a final OD₆₀₀ of 0.1. The new culture was grown to an OD₆₀₀ between 0.6-0.8 at 37 °C with shaking. Sterile filtered stock solution of IPTG (0.5 M) was added to 18x 5 mL aliquots of the culture for a final concentration of 0.1 mM, 0.5 mM and 1 mM IPTG for six cultures each. Half of each of the cultures were further shaken at 16 and 30 °C respectively for 3 h, 5 h and ~18 h time points before harvesting by centrifugation at 3500 xg for 10 min. The cell pellet was resuspended in 1 mL of phosphate buffered saline (PBS) and lysed by sonication for 1.5 min (10 sec on and 10 sec off). The pellet and supernatant were separated by centrifugation at 24000 xg with a Heraeus Multifuge X3R equipped with an F15-8x50 rotor. Each fraction was tested by denaturing 12% or 15% SDS-PAGE gels run at 200 V. Optimal conditions for expression were determined and applied to large scale expression.

TamA ANL Expression and Purification

A single colony of *E. coli* cells containing an expression plasmid were grown in 1 LB media with Kanamycin. Once the cultures reached OD₆₀₀ ~ 0.8 they were induced to express the gene using 0.1mM IPTG and left shaking at 16°C overnight. The cells were then harvested by centrifugation at 3500 xg for 10 min before being stored at -20°C. Purification was carried out on ice or at -4°C. Resuspension of the cells in the binding buffer (50 mM Tris pH 8, 250 mM NaCl, 10 mM imidazole, 10% glycerol) and lysed by sonication for 10 minutes (30 secs on and 30 secs off). This was then centrifuged at 14,000xg for 30 minutes to separate the pellets and supernatant and filtered through at 45µM filter. The cell-free extract was loaded onto a 1mL HisTrap Nickel affinity column (GE Healthcare) at 1mL/min. The column was washed with the binding buffer before eluting the protein by increasing the imidazole gradient (10 to 500mM). The fractions containing the protein were combined and dialysed with 1mg TEV against the dialysis buffer (50mM Tris pH8, 250mM NaCl and 10% glycerol) for 2h. The dialysed protein was then run through the HisTrap again to remove the cleaved HisTag, collecting the protein as flow through. The protein fractions was combined and concentrated down before loading onto a 120mL Superdex S200 gel filtration column with GFC buffer (50mM Tris pH8, 250mM NaCl and 10% glycerol). Those containing clean protein of interest were combined and concentrated for use or flash frozen at -80 °C. Purification was monitored by SDS-PAGE.

TamA Cloning, Expression and Purification

The *tamA* gene was cloned from *P. tunicata* D2 genomic DNA into the pEHISTEV vector to express the protein with a TEV-cleavable N-terminal 6xHis tag. This gene was expressed in 2L of BL21 (DE3) *E. coli* cells at 16 °C for ~18 h with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). TamA was then purified from the cell lysate using the same method described above. This includes nickel immobilized metal affinity chromatography (IMAC) followed by 6xHis tag cleavage by TEV protease. An additional step is the insertion of the 4'-phosphopantetheine (4'-PP) arm onto the ACP domain of TamA. This was done by adding the concentrated protein to 10 mL reaction containing 10 μM Sfp, 10 mM MgCl₂ and 500 μM CoASH in dialysis buffer. This reaction was left shaking for 18h at 4 °C. Protein purification was then continued as before by collecting the run through of the second HisTap, removing the sfp and TEV protease, before concentrating down and loading onto the HiLoad 16/60 Superdex S200 column.

Intact Protein LC ESI-MS

Bradford assay was used to determine protein concentration for liquid chromatography electrospray-MS (LC ESI-MS). Protein was diluted to a concentration of 10 μM and centrifuged for 10 min at 17000 xg. 5 μL of the supernatant was subjected to LC ESI-MS on a Synapt G2-Si Q-TOF (Waters) instrument with Phenomenex C4 3.6 μm LC column coupled to an ESI source. The LC gradient ran from 5% acetonitrile and water with 0.1% formic acid to 95% acetonitrile over 15 min. The MS source was set at 120 °C, backing pressure 2 mbar and sampling cone voltage of 54 V. Protein spectra are presented after subtraction using MassLynx V4.1 software. Peak *m/z* annotations were extracted from smoothed and centroided data which was also used in the component algorithm to extract the protein average masses.

TamA ANL Reactions

Reactions were carried out according to the methodology of the Flitsch group.²¹⁰ They contained 10 μg ANL, 1 mM fatty acid (from a 25 mM stock in DMSO), 5 mM ATP, 10 mM MgCl₂ and 100 mM amine (pH 9) in 100 mM sodium carbonate-sodium bicarbonate buffer (pH 9).

Amines and Amides LC ESI-MS

Amine or amide producing reactions were quenched with a 1:1 volume of acetonitrile with 0.01% TFA and centrifuged at 17000 xg. 5 μL of supernatant was subjected to LC ESI-MS on a Synapt G2-Si Q-TOF (Waters) instrument with Phenomenex Jupiter C18 5 μm 300 Å LC column coupled to an ESI source. The LC gradient ran from 5% acetonitrile and 95% water with 0.1% formic acid to 95% acetonitrile over 30 min. The MS source was set at 120 °C, backing pressure 2 mbar and sampling cone voltage of 54 V. Extracted ion chromatograms (EICs) and masses were determined on MassLynx V4.1 software.

High Performance Liquid Chromatography (HPLC) to determine conversion

Amide forming reactions were quenched by addition of a 1:1 volume of acetonitrile with 0.1% TFA. Standard samples were also made up in a 1:1 volume of acetonitrile to water. The samples were centrifuged at 17000 xg for 10 min before injection of 10 μL of the supernatant onto a Phenomenex Luna 5 μm C18 100 Å HPLC column. Samples were eluted with water and 0.1% TFA for 5 min followed by a 30 min gradient to 95% acetonitrile and 0.1% TFA. This concentration was maintained for 5 min and then returned to 100% water and 0.1% TFA for an additional 5 min. The eluent was monitored at 215 nm.

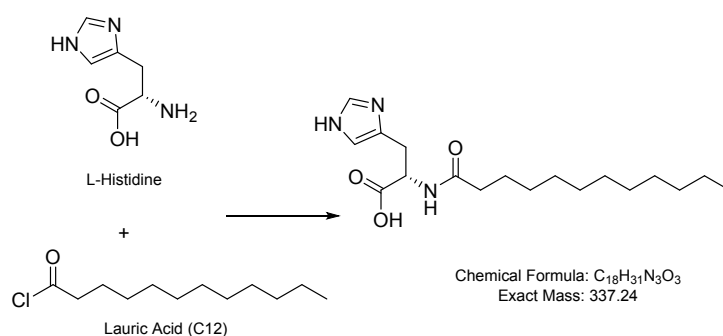
N-Dodecanoyl-L-Histidine (N-C12-L-His) Formation

NMR Spectra (^1H and ^{13}C) were recorded on an Avance III 500MHz spectrometer and DMSO- d_6 was used as the solvent. J values given in Hz. MS analysis recorded on a Synapt G2-Si Q-TOF (Waters) instrument and a Bruker Ultraflex MALDI TOF.

Chemical Synthesis

There are two published methods for the synthesis of the N-acyl His molecules- **(1)** Lei et al., Langmuir 2010 which couples L-His and the C12 acid chloride at 10g scale (L-His). The product was recrystallized and analysed with no further purification. **(2)** In contrast, Torring et al, ChemBioChem 2017, used Fmoc-His(Trt)-Wang resin and coupled this with the C6, C8 and C10 fatty acids but no scale and detailed reaction conditions were given. The resulting amide was isolated by HPLC-purification using a semi-prep C18 column but no solvent conditions and yield were reported.

We decided to follow the Lei et al method at $\sim 1/10^{\text{th}}$ scale.



L-Histidine (5.52mmol, 857mg) and NaOH (6.62 mmol, 265mg) were dissolved in a mixture of water (30 ml) and THF (10 ml) on ice. Dodecanoyl Chloride (5.03 mmol, 1.100 g) in THF (5 ml) was slowly added to the mixture. After 6 hours, a white precipitate had formed which was isolated and washed with water and ethanol. The reaction solution was then washed with petroleum ether (x 3), before the aqueous pH was altered to pH 5 upon which a white precipitate formed. The white precipitates were combined and recrystallized with 95% to give the final product N-Dodecanoyl-L-Histidine (90 mg g, 5.2%). δ H (500 MHz, DMSO- d_6) 8.00 (d, J 7.9, 1H), 7.57 (s, 1 H), 6.80 (s, 1 H), 4.40 (td, J 5.2 and 8.2, 1 H), 2.90 (m, 2H), 2.10 (t, J 7.4 2H), 1.46 (m, 2H), 1.25 (s, 16 H), 0.86 (t, J 6.9, 3 H). δ ^{13}C (500 MHz, DMSO- d_6) 174.96, 173.67, 135.09, 52.57, 35.62, 31.76, 29.51- 22.56, 14.42. m/z (ESI) for the $[\text{M}+\text{H}]^+$ is $[\text{C}_{18}\text{H}_{32}\text{N}_3\text{O}_3]$ predicted: 338.24, observed: 338.3. See Figures S9, S10 and S11.

Solubility issues with product isolation

In our hands we found the N-C12-L-His product to be insoluble in a range of solvents and only partially soluble (e.g. 100% DMSO). Therefore it was not possible to prepare a sufficiently concentrated stock to use preparative HPLC to further purify the compound. Nevertheless we characterised the Tama ANL catalysed product using NMR and MS (Figures S12).

TamA ANL Biotransformation

A scaled up reaction was attempted, which contained a total of 5 mg ANL in which 1mg was added every hour for 5 hours, 1 mM fatty acid (from a 25 mM stock in DMSO), 5 mM ATP, 10 mM MgCl₂ and 10 mM amine (pH 9) in 100 mM sodium carbonate-sodium bicarbonate buffer (pH 9). Reactions were shaken at 37°C for a total of 24 hours. After 24h, a white precipitate had formed which was isolated and washed with water and ethanol. The crude white precipitate of the final product N-Dodecanoyl-L-Histidine (0.029 g, 86%). δ H (500 MHz, DMSO-d⁶) 8.00 (d, *J* 4.8, 1H), 7.57 (s, 1 H), 6.81 (s, 1 H), 4.25 (m, 1 H), 2.93 (m, 2H), 2.06 (t, *J* 7.4 2H), 1.50 (m, 2H), 1.25 (s, 16 H), 0.86 (t, *J* 6.8, 3 H). *m/z* (MALDI) for [M+H]⁺ is [C₁₈H₃₂N₃O₃] predicted: 338.24, observed: [M+H]⁺ 338.239, 339.242 and 340.244

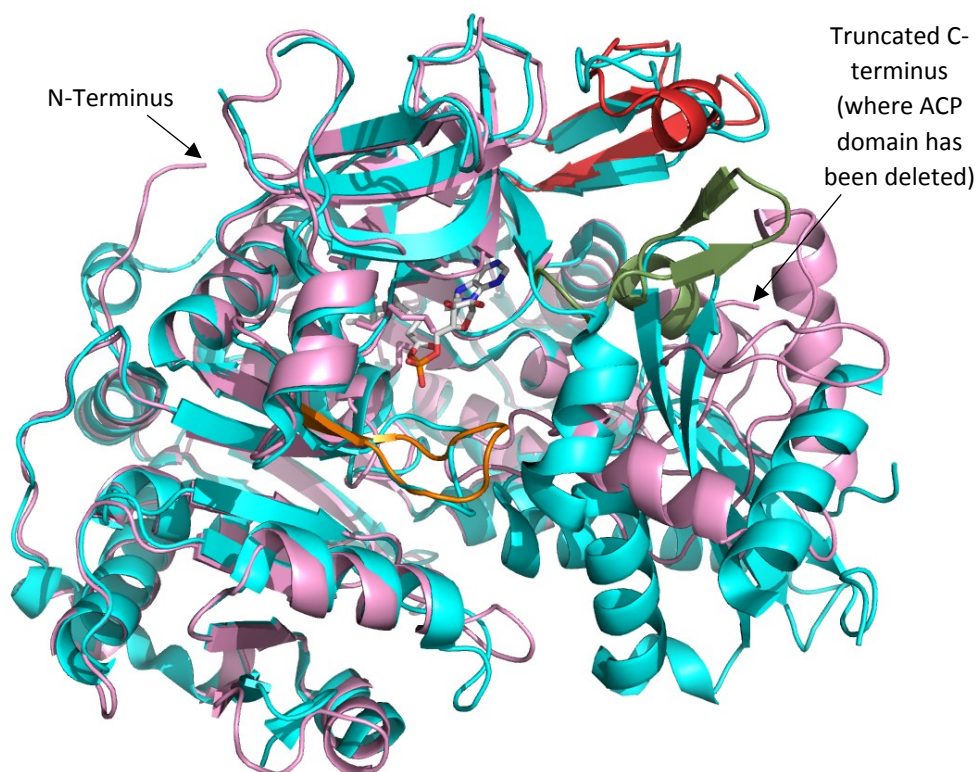
Supplementary Figures:

```
GAMECEASSLIDLLAFDARNKPSQE VFRFVSDNGESEASYDYQ TLSQEISRIAIGLQALIKT      60
SHNQDQALIVLPQGVQFVTAFYGCMAANVIAVPSFPPKSQLQIERLQFAITDLGNPIVITNR      122
DILPKLQEHIALDSVRWLLIEDLASVIAQPLSDFRTHEHSIALLOYSSGTTGKPKGVIITNQ      184
NIMENSELIRQSFHGKEDHTRMMLWLPPHDMGLVGGVMQGVYTYPTLLMPTDLFLRSQYR      146
WLKAVSDYRATTTGAPNFAYELAVKNIRESRLAELDLSSLENLFCGAEPINSHSINQFLDKF      308
APCGLKPEAFPLPCYGMAEATLMVSGKPHGQQYKQLCIDPELLKHGMVKPLNTPNAHSLWLVS      370
SGVVHSSLQARIVNPETGTEVAQQGVGEIWLQSSISPGYWQDAERTAINFGLPLAGYEETF      432
HRTGDLGFYHQDELFITGRLKEVVIIRGANFYPODLEYETT LAFPELNNCRSAAFSVPKEGK      494
EQLIMAIEVPRNVTEFNQYAKILNGRLVERFGIRADIILFLPRKTIKITSSGKLQRVAIKKA      556
YEEQQLPVYFQYQLOQE QIAPREVS LDISNQDSVAKWLVARVSELTGVAIAQISEHEPLTNV      618
GLDSVLAMEIILFRLEQQTG VYLLAPDVLVYSCNTPSLLAEQIIKVAGNVAEKELNLSC      674
```

Figure S1: Shown here is the TamA recombinant protein sequence from pEHISTEV expression construct. The truncated TamA ANL576 sequence is shown in blue, the ANL to ACP domain linker is underlined and in red is the ACP domain.

TamA ANL Homology Model

(a)



(b)

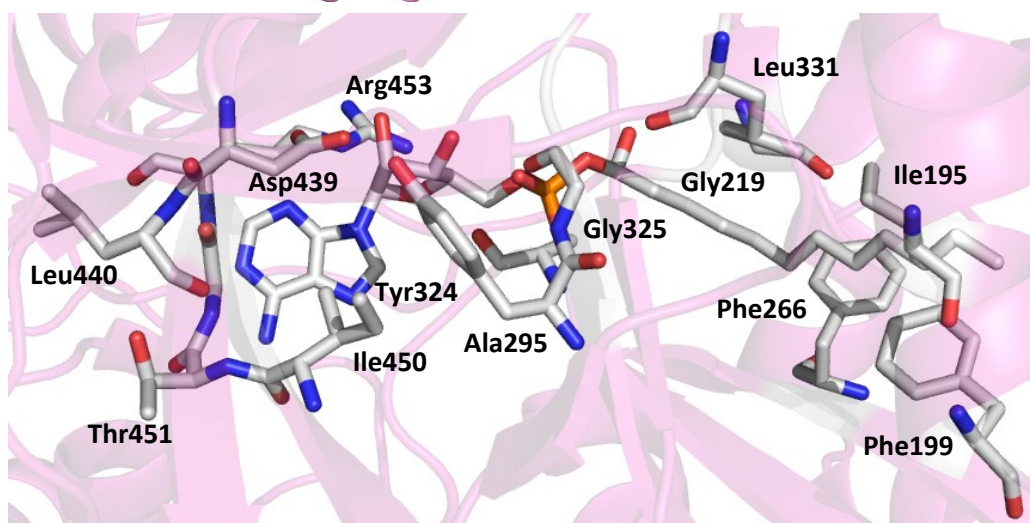


Figure S2: (a) Homology model of the TamA ANL domain structure constructed with Phyre2. We compare this with the fatty acid AMP ligase (FAAL) from *Legionella pneumophila* (blue, Uniprot code Q5ZTD3, PDB Code 3KXW) with the adenylate intermediate bound (multi-coloured). The TamA model is rendered with ANL domain (pink), the ATP-binding P-loop (orange), the hinge region (green) and the FAAL insertion loop (red). (b) The overlaid structures were then used to identify conserved active site residues between TamA ANL and FAAL and allowed a prediction of the TamA ANL active site residues less than 4Å away from the C12 acyl-adenylate intermediate. In particular residues Ile195, Phe 199 and Phe266 are predicted to be involved in binding the acyl chain. The figures were made using Pymol Molecular Graphics System, Version 1.8.

Sequence Alignment

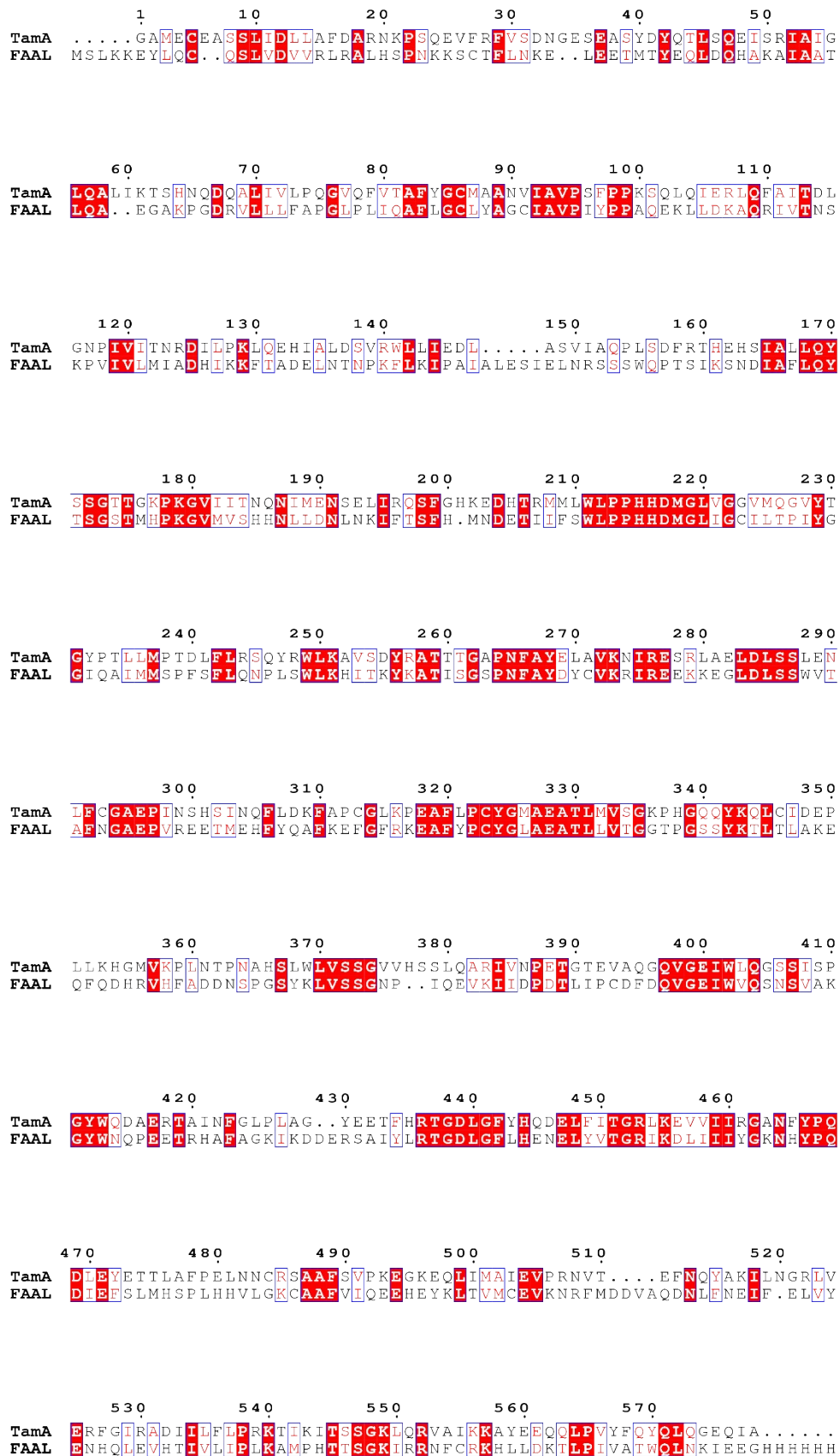


Figure S3: TamA ANL sequence alignment with closest homologue fatty acid AMP ligase .

TamA ANL Expression Test

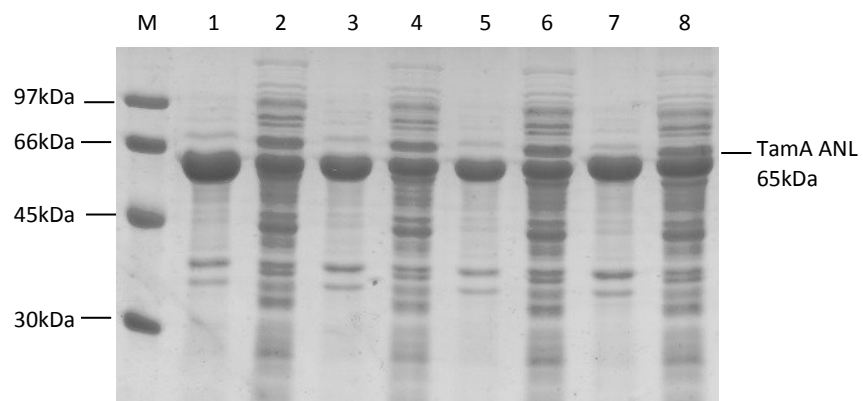


Figure S4: SDS-PAGE analysis of test expression of TamA ANL domain constructs (M) Marker (1) insoluble and (2) soluble fractions of expression of TamA ANL576, (3) insoluble and (4) soluble fractions of expression of TamA ANL581, (5) insoluble and (6) soluble fractions of expression of TamA ANL586, (7) insoluble and (8) soluble fractions of expression of TamA ANL591.

TamA ANL576 Purification

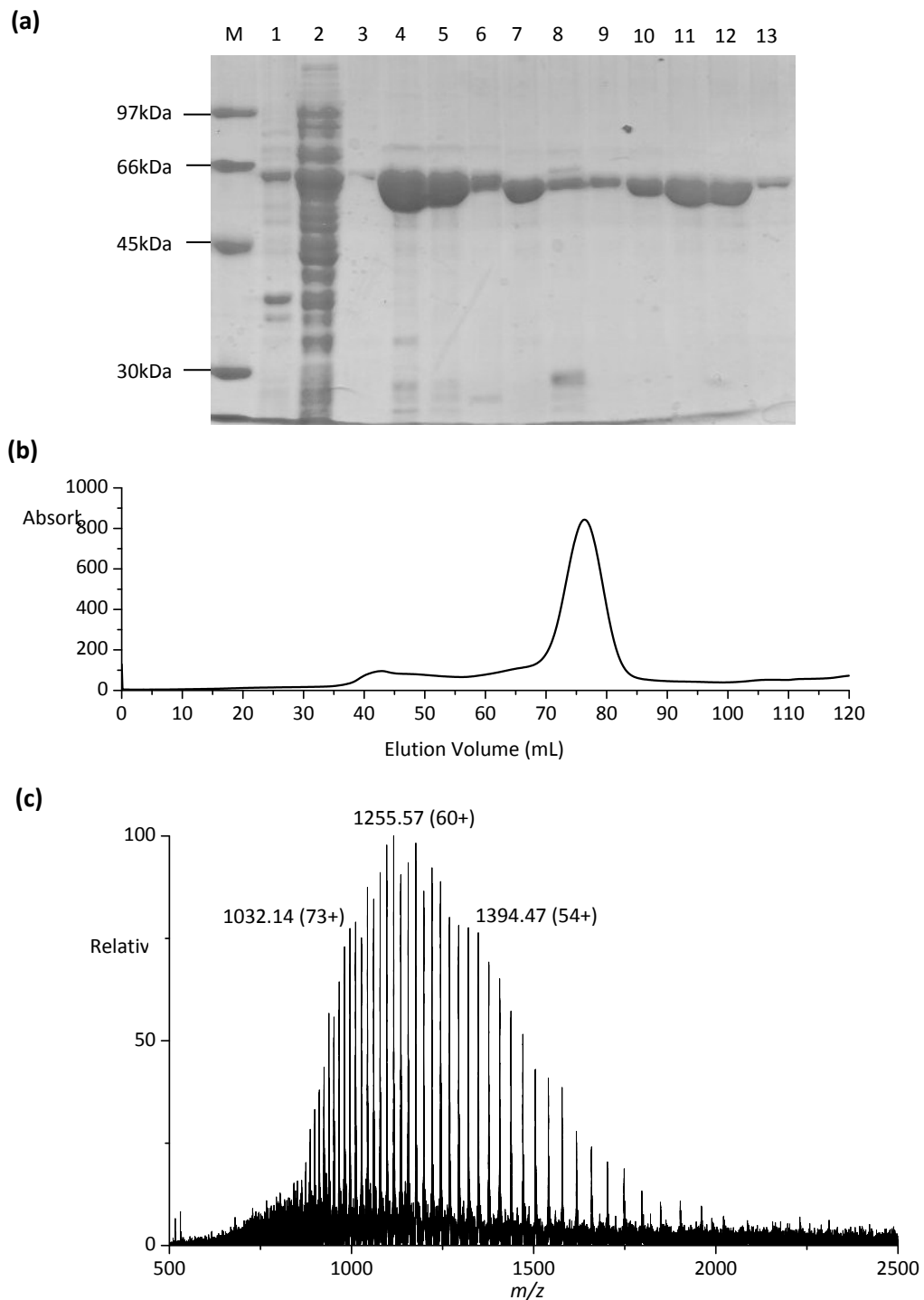
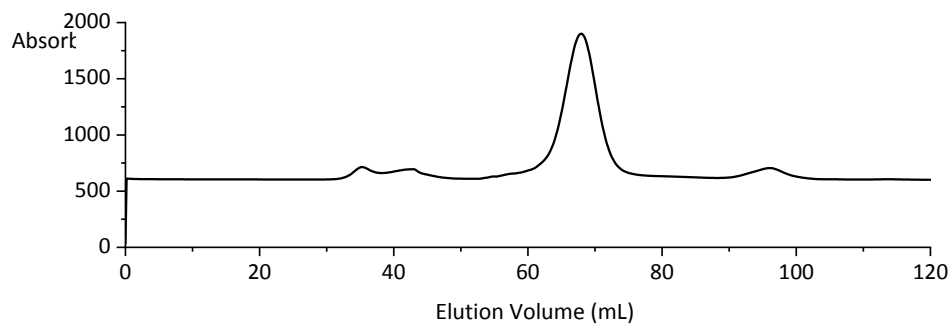


Figure S5: TamA ANL576 purification (a) SDS-PAGE gel showing (M) marker, (1) insoluble and (2) soluble fractions of the cell pellet (3-6) 3mL fractions from nickel IMAC purification, (7) wash of TEV-cleaved TamA ANL, (8) elution of uncleaved TamA ANL and TEV protease (9-13) 3 mL fractions from the Superdex S200 GFC, (b) chromatogram of Superdex S200 GFC monitored at 280 nm showing peak elution at 76.3 mL consistent with the protein monomer (c) denaturing ESI-MS spectrum with a deconvoluted mass of 64653 ± 1 Da consistent with the predicted mass of 64651 Da.

TamA Purification

(a)



(b)

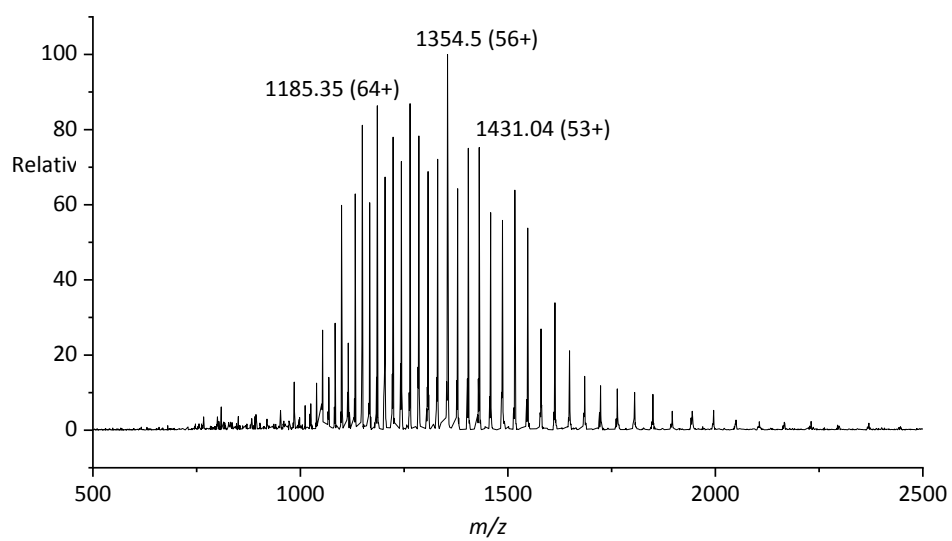


Figure S6: TamA purification (a) chromatogram of Superdex S200 GFC monitored at 280 nm showing peak elution at 70 mL consistent with the protein monomer (b) denaturing ESI-MS spectrum with a deconvoluted mass of 75789.95 ± 1.85 Da consistent with the predicted mass of the acylated enzymes 75270 Da + 4'-PP + C12.

C9-C14 Amide Products ESI-MS

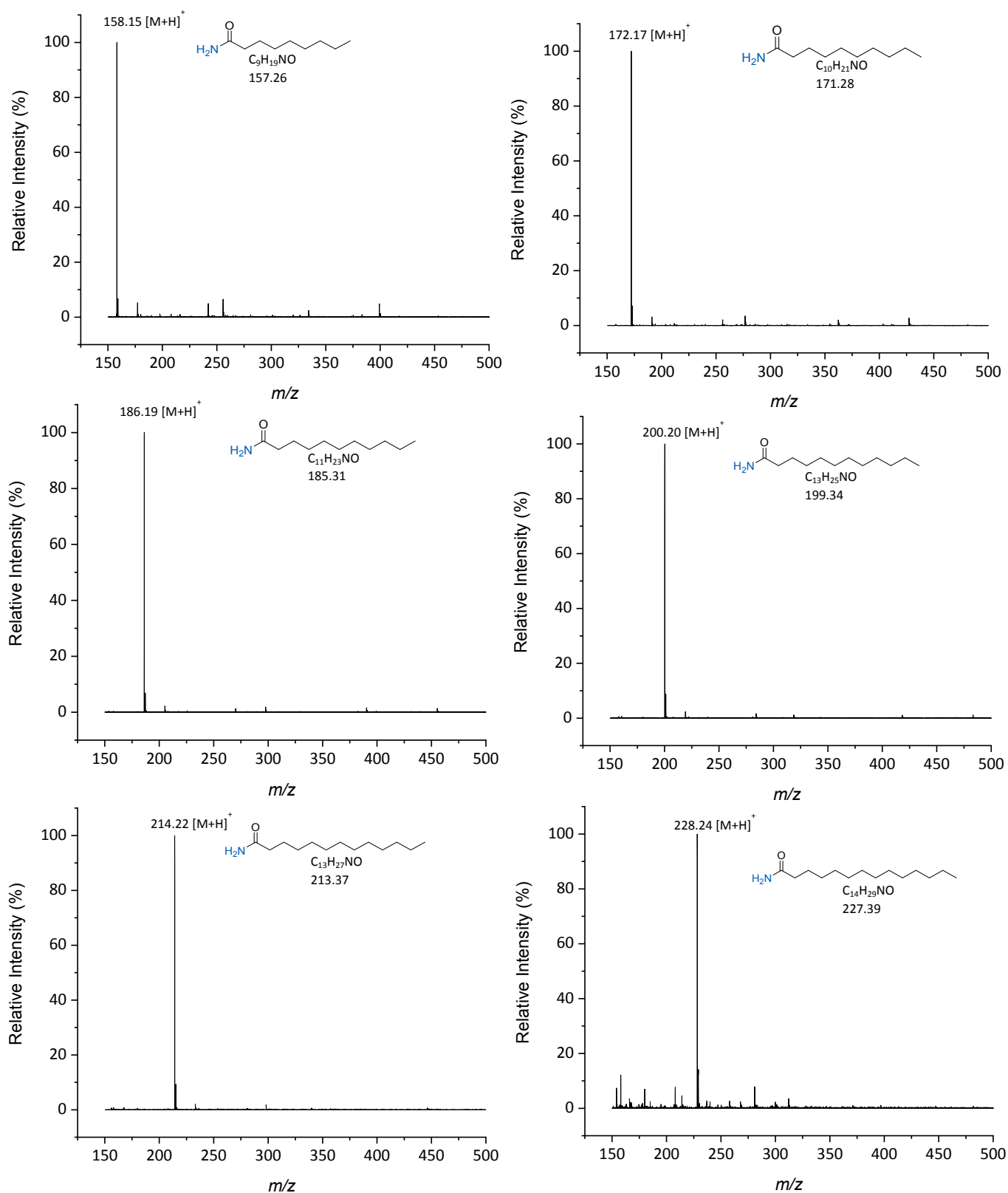
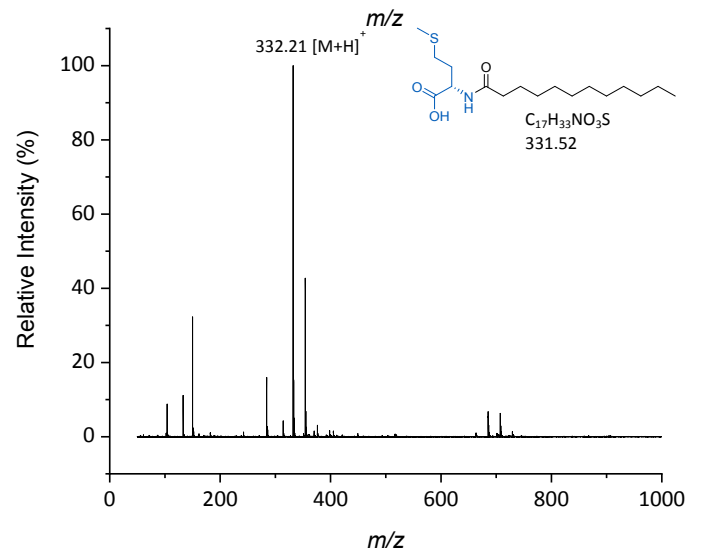
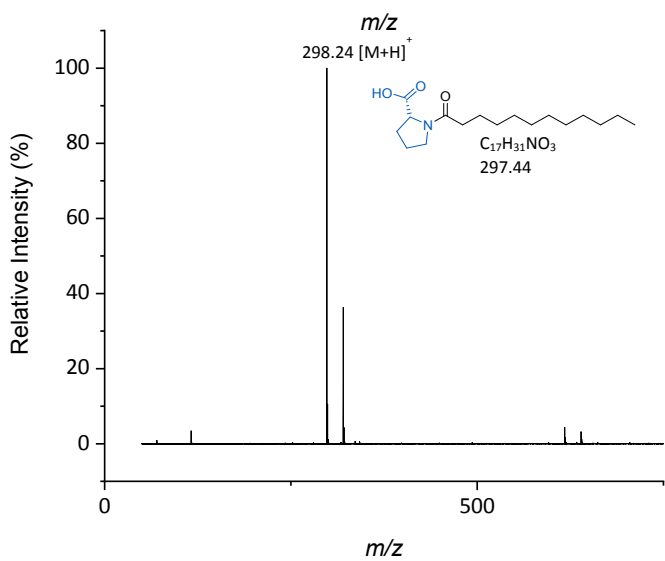
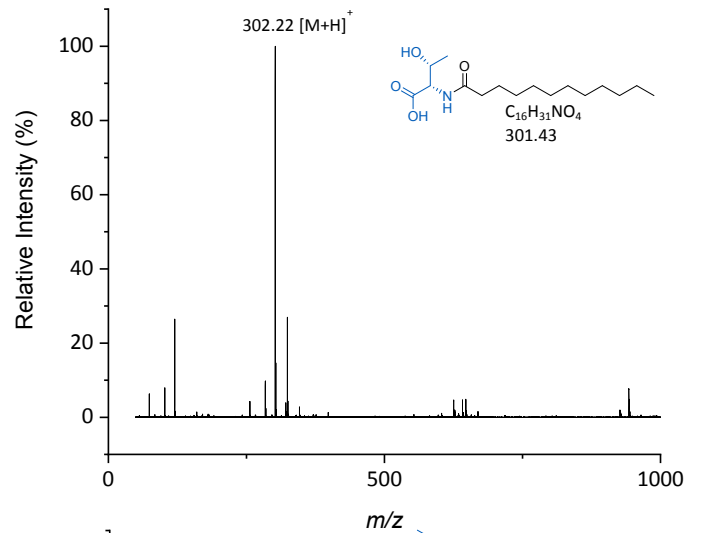
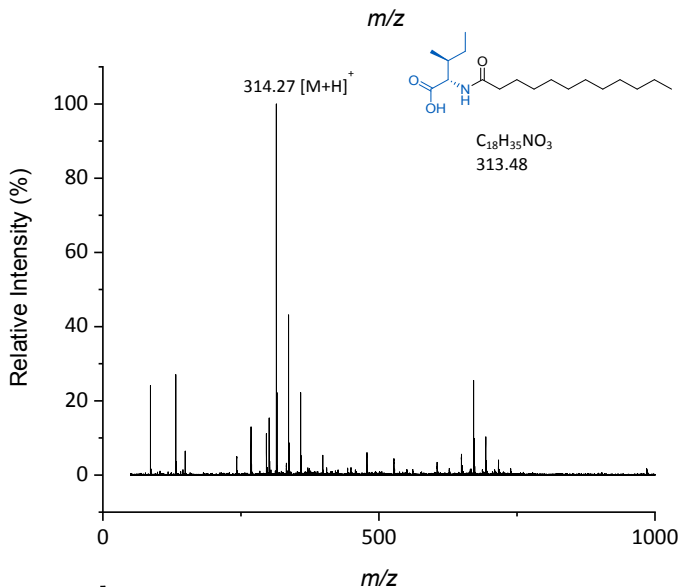
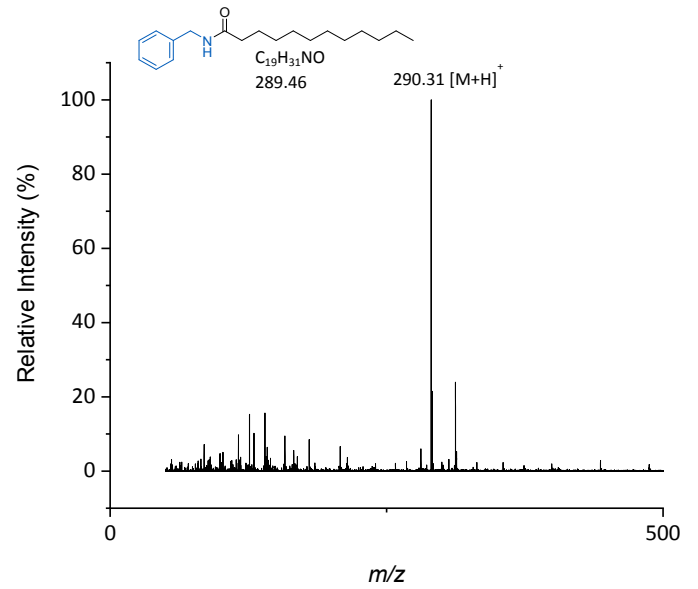
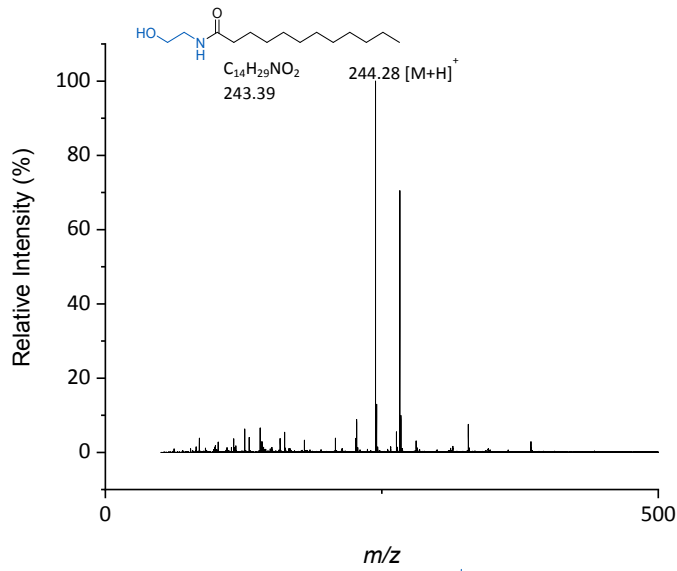


Figure S7: ESI-MS chromatograms of TamA ANL catalysed amide synthesis with fatty acid lengths C9-C14 and ammonia.

L-C12 Amide Products ESI-MS



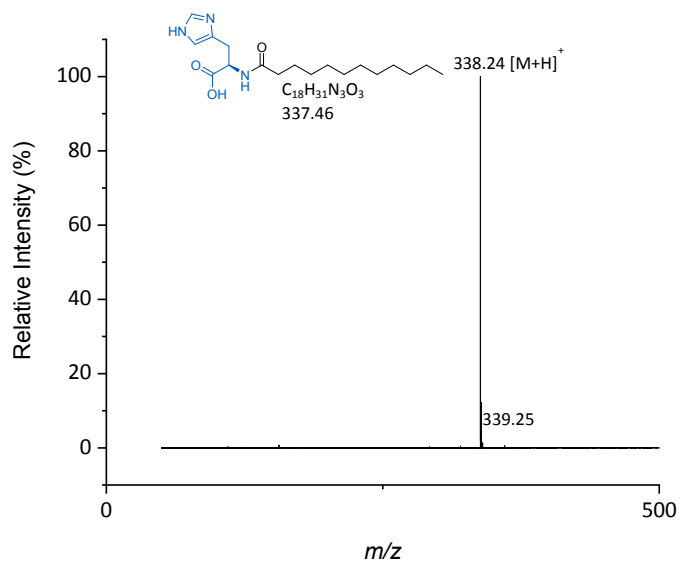
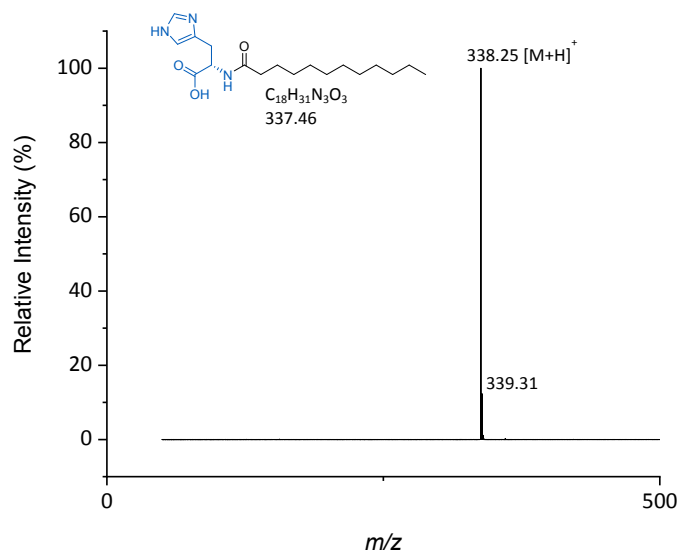
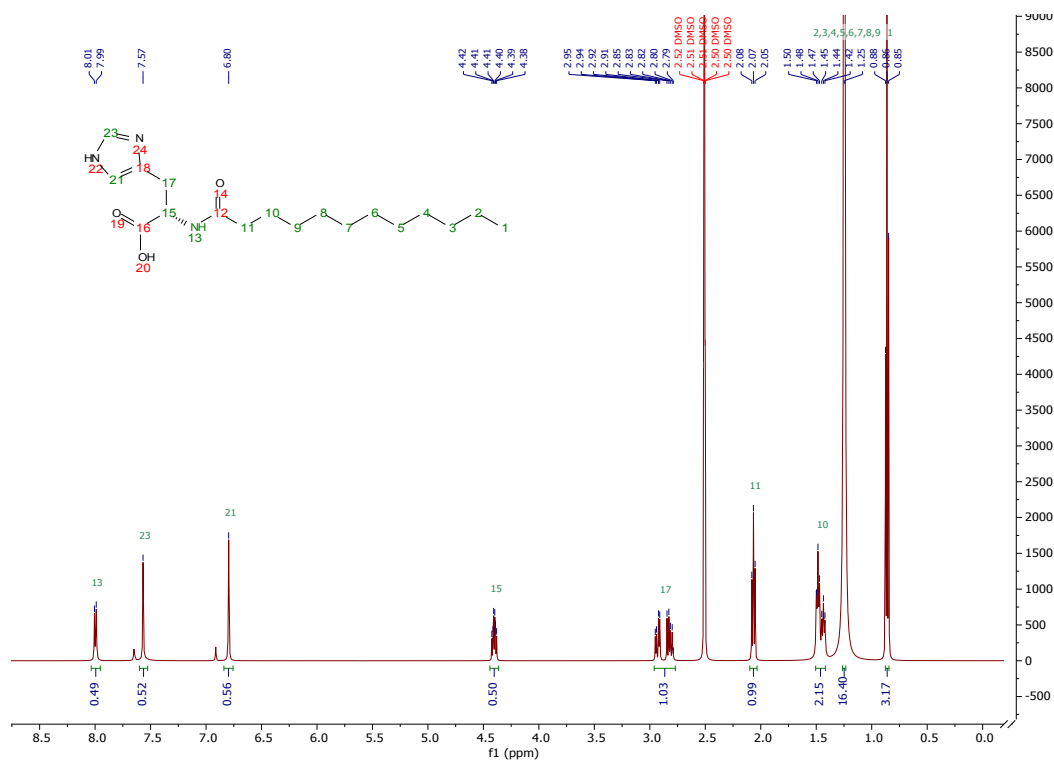


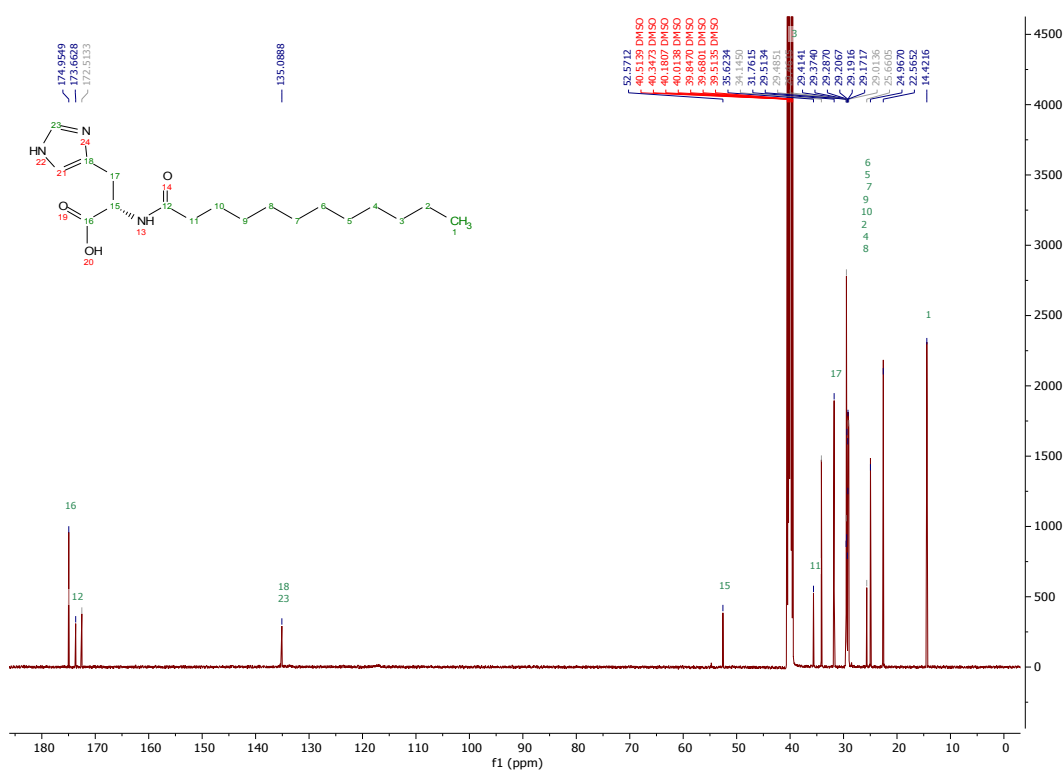
Figure S8: ESI-MS chromatograms of TamA ANL catalysed amide formation with fatty acid length C12 with a series of amines.

Chemically synthesised N-C12-L-histidine Characterisation

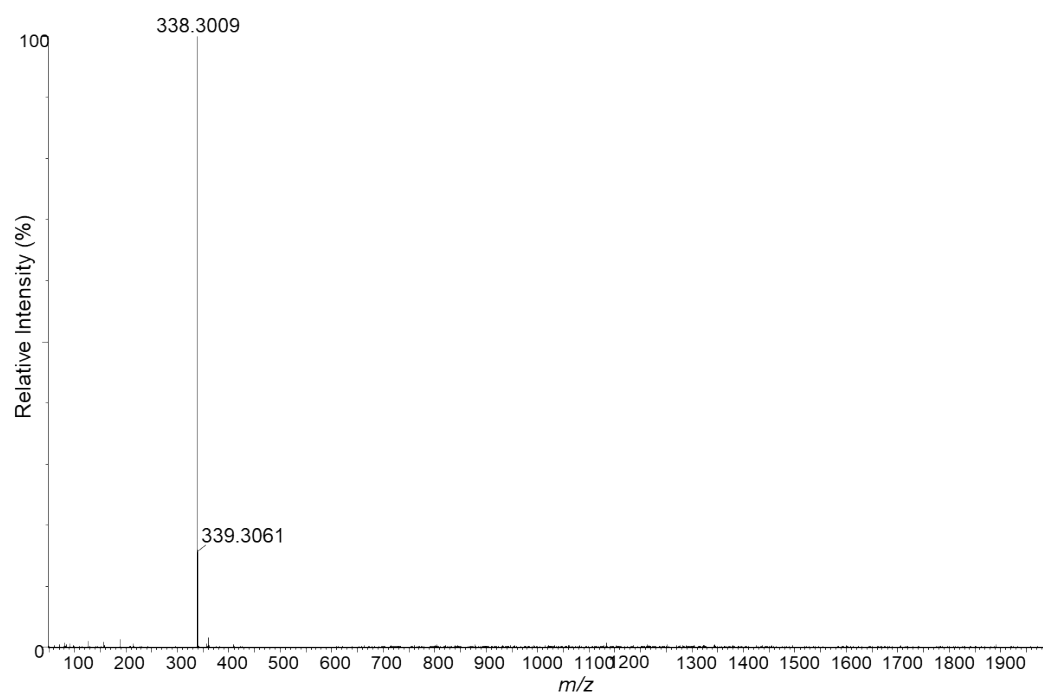
(a)



(b)



(c)

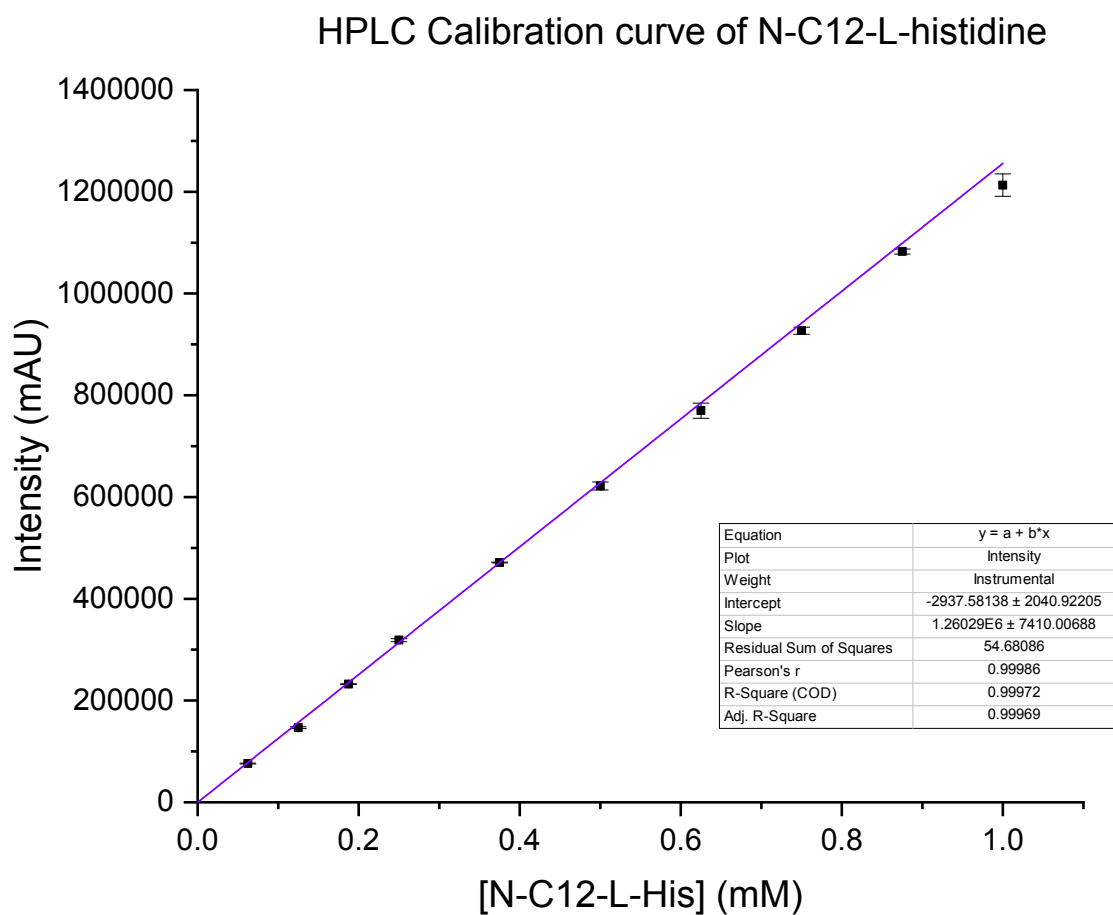


(a) ^1H NMR, (b) ^{13}C NMR, (c) Positive mode ESI Mass Spectrometry (ESI-MS).

Figure S9: Characterisation of N-C12-L-histidine synthetic standard

HPLC Data

(a)



(b)

Time	Conversion
10 min	1.5 %
30 min	5.6 %
1 hour	9.6 %
2 hours	12.6 %
4 hours	12.7 %
24 hours	12.7 %

Figure S10: a) HPLC calibration of synthetic N-C12-L-histidine at various concentrations (0.0625-1.0 mM) prepared from a 2mM stock in 4% DMSO/96% AcCN. Each measurement was done in triplicate. **b)** Product conversion determined using the calibration at different time points taken during the reaction.

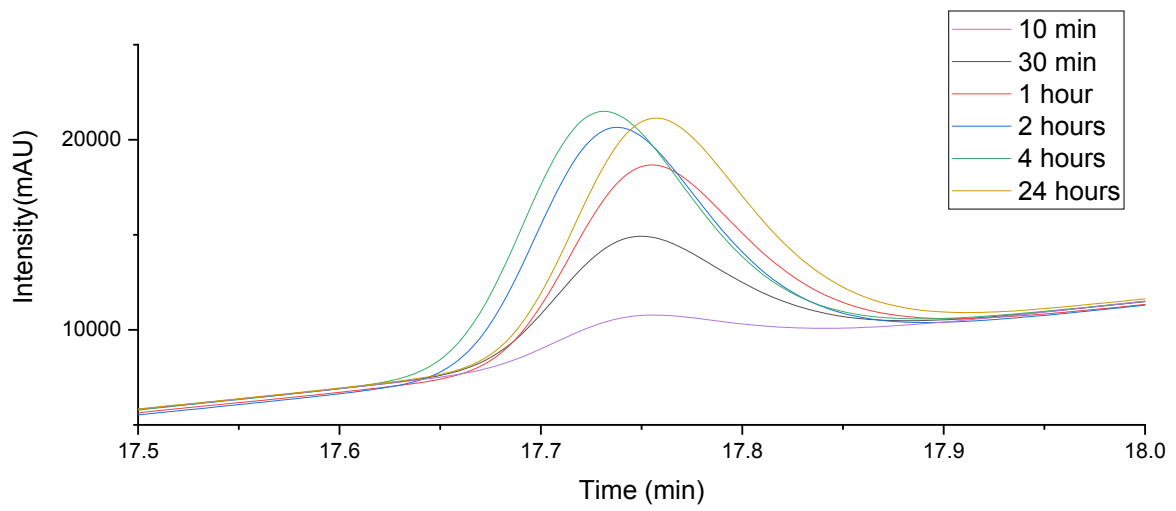


Figure S11: HPLC analysis of TamA ANL reaction stopped at different timepoints.

N-C12-L-His TamA ANL Biotransformation

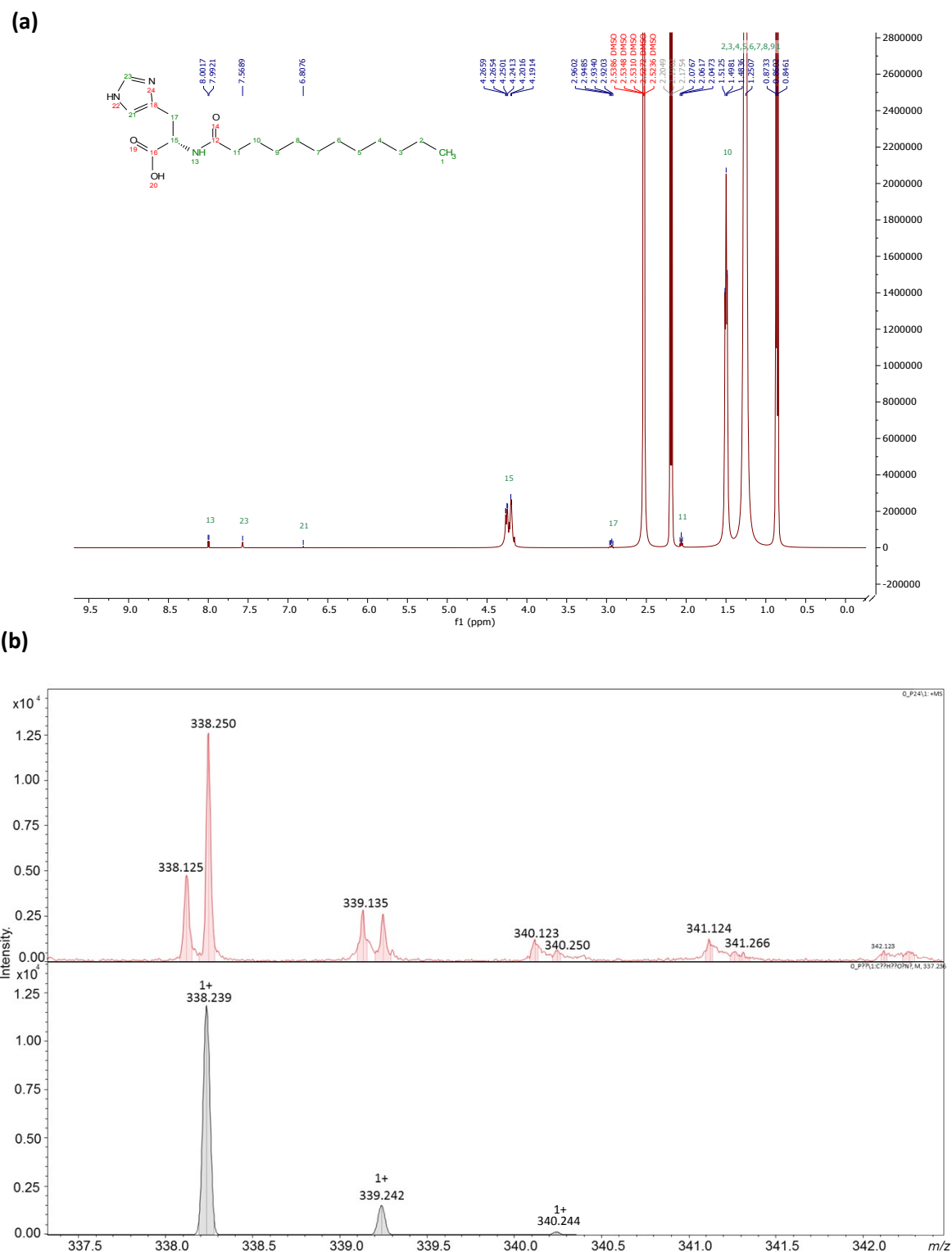


Figure S12: Characterisation of biocatalytically formed N-C12-L-histidine by (a) ^1H NMR and (b) MALDI-MS chromatogram of biocatalytic N-C12-L-histidine product (red) compared with the simulated chromatogram (grey).

Histidine Amide Products ESI-MS

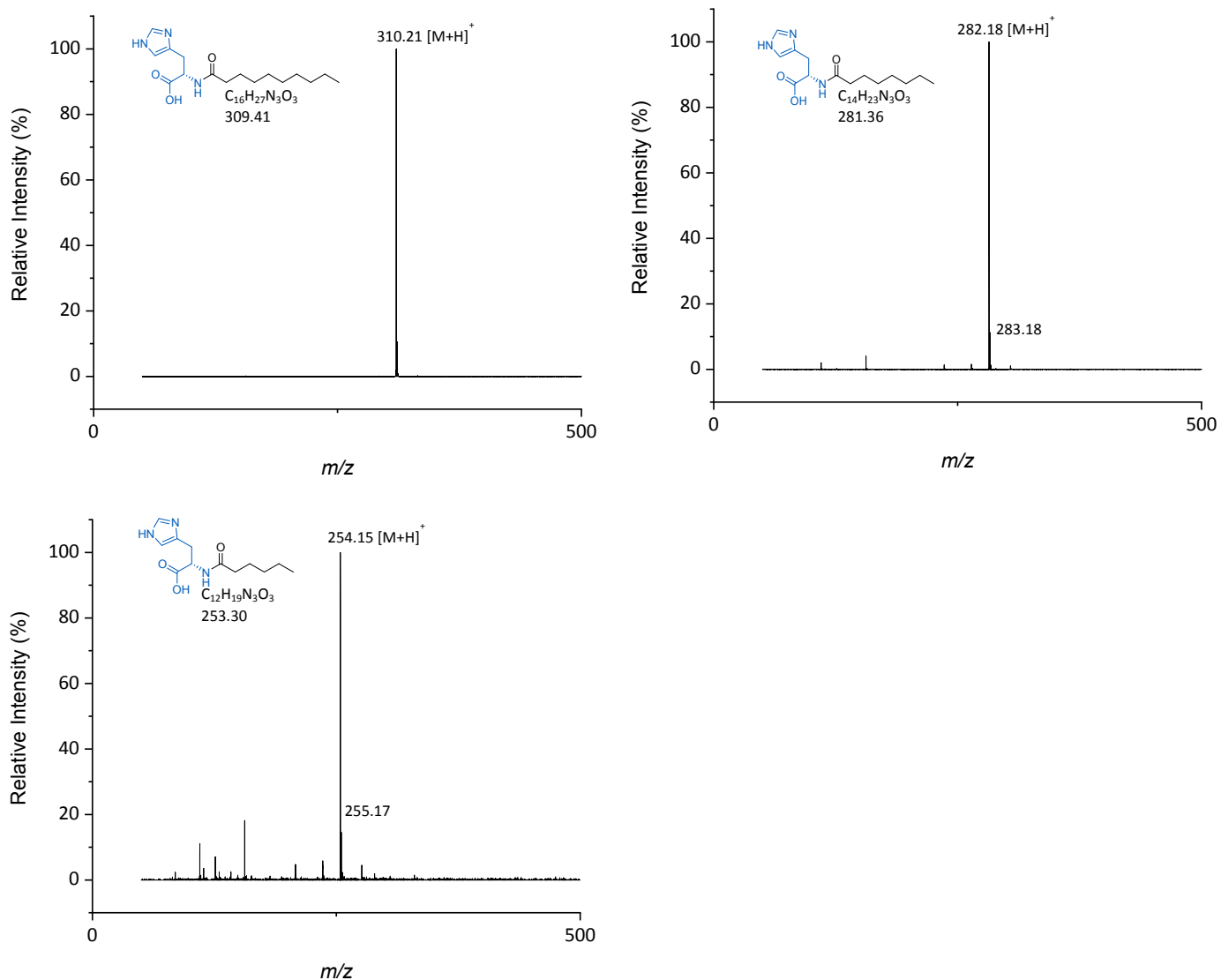


Figure S13: ESI-MS chromatograms of TamA ANL catalysed N-acyl Histidine synthesis with fatty acid lengths C6-C12.