OPEN ACCESS

UNIVERSITY OF BRISTOL

Stenner, R. A., & Anderson, J. L. R. (2020). Chemoselective N-H insertion catalysed by a de novo carbene transferase. *Biotechnology and Applied Biochemistry*. https://doi.org/10.1002/bab.1924

Peer reviewed version

Link to published version (if available):
10.1002/bab.1924

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Chemoselective N-H insertion catalysed by a *de novo* carbene transferase

Stenner, Richard[1,2] and Anderson, John Leslie Ross[1,3*].

[1]School of Biochemistry, University of Bristol, University Walk, Bristol, BS8 1TD, UK.

[2]Bristol Centre for Functional Nanomaterials, HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, UK.

[3]BrisSynBio Synthetic Biology Research Centre, Life Sciences Building, University of Bristol, Tyndall Avenue, Bristol BS8 1TQ, UK.

Email: ross.anderson@bristol.ac.uk

Running title: *Chemoselective N-H insertion*

ABSTRACT

The ability to perform organic reactions with chemoselectivity is of critical importance in synthesis. Recently we reported that a *de novo* carbene transferase, a tetra-α-helical *c*-type heme-containing protein, C45, is proficient at N-H insertion reactions, proceeding *via* the intermolecular transfer of a metallocarbenoid intermediate into the N-H σ-bond to form a new N-C σ-bond. Here we demonstrate that C45 can also catalyse N-H insertion reactions chemoselectively, even when the substrate contains an unprotected hydroxyl group.

KEY WORDS:

- **Biocatalysis**
- **Carbene transferase**
- *De novo* **protein design**

INTRODUCTION

Organic molecules often contain more than one type of functional group. When such molecules are used in organic transformations, a fundamental problem that is frequently encountered is a lack of chemoselectivity,[1,2] often resulting in undesirable and uncontrolled modification of off-target functional groups. To circumvent such chemoselectivity issues, it is often necessary to employ protecting group methodologies, which add additional steps to the overall synthetic pathway.[3–7] Therefore, the ability to chemoselectively transform substrates is of high importance both technologically and, where enzymes are involved, biologically. In addition to stereoselectivity, enzymes often exhibit significant chemoselectivity during biosynthesis, selectively transforming one particular functional group in the presence of others.[1,8–13] The recent explosion of engineered carbene transferases has afforded novel biosynthetic strategies for cyclopropanation,[14,15,24,25,16–23] X-H insertion (X = N, S, B, Si)[26–33], carbonyl olefination reactions[34,35] and ring expansion reactions.[36] While engineered proteins have proven versatile scaffolds for catalysing such transformations, multiple rounds of directed evolution are usually required before a desired activity is achieved.[37,38]

As a viable alternative to using engineered natural proteins, *de novo* protein design has afforded simplified, novel protein scaffolds capable of reproducing natural protein and enzyme functions.[39–45] In particular, we have demonstrated that the tetra-α-helical, *c*-type heme containing maquette C45 is a proficient and promiscuous peroxidase catalysing the peroxidation of numerous substrates with rates similar to, and even surpassing, natural peroxidases such as horseradish peroxidase.[46] More recently, we have demonstrated that C45 is also a proficient and stereoselective carbene transferase, capable of catalysing cyclopropanations, N-H insertions, carbonyl olefinations and ring expansion reactions with good yields and tuneable stereoselectivity.[36] As C45 has already demonstrated proficient N-H carbene insertion chemistry towards *para*-chloroaniline and piperidine,[36] the question of whether C45 would catalyze the X-H carbene insertion reaction when alkanolamines are employed as substrates was raised (**Figure 1**). The three regioisomers of aminophenol (2-aminophenol (2AP), 3-aminophenol (3AP), and 4-aminophenol (4AP)) were identified as suitable substrates for exploring the chemoselectivity of biocatalytic carbene X-H insertions. Here, we report the chemoselective carbene N-H insertion reaction of the three regioisomers of aminophenol catalysed by C45.

RESULTS AND DISCUSSION

We initially tested the ability of C45 to catalyze carbene insertion into the O-H bond of phenol, using ethyl diazoacetate (EDA) as the carbene precursor. The normalized C18-HPLC chromatogram for phenol (red) and for the C45 O-H insertion assay product mixture (blue) are presented in **Figure 2**.

Analysis of the product mixture *via* HPLC and LC-MS revealed no appreciable O-H insertion product, indicating that C45 does not catalyze the insertion of carbene into the O-H of phenol.

We then explored the potentially chemoselective C45-catalyzed insertion of carbenes into the N-H bonds of the three regioisomers of aminophenol **Figure 3**. **Figure 4 (top)** presents the C18-HPLC chromatograms and the LC-MS mass spectra acquired for the C45-catalysed reactions between 2-aminophenol and i) ethyl diazoacetate (EDA), ii) *tert*-butyl diazoacetate (ᵗBuDA), and iii) benzyl diazoacetate (BnDA). The chromatogram for the reaction between EDA and 2AP is dominated by a major peak at 16.25 minutes, preluded by a split peak in the range 15.1-15.5 minutes which is attributed to the various ionization states of unreacted 2AP **(SI Figure 1)**. The emergence of the new peak at 16.25 minutes corresponds exceedingly well to the retention time reported for *N*-phenylglycine ethyl ester (16.6 minutes, **SI Figure 1)**, with the deviations probably accounted for by the introduction of a hydroxy group in the *ortho*-position of the product. Subsequent LC-MS analysis of the peak at 16.25 minutes generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact mass expected for the iminium fragment of the 2AP/EDA single-insertion product ethyl (2-hydroxyphenyl)glycinate. The mass spectrum also possesses a small peak at 195 m/z, corresponding to the expected mass of the [M+H] parent ion. The chromatogram for the reaction between ᵗBuDA and 2AP exhibits a single peak at 15.70 minutes, preluded by a dominant single peak at 15.1 minutes which corresponds to unreacted 2AP **(SI Figure 1)**. Subsequent LC-MS analysis of the peak at 15.70 minutes generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact mass expected for the iminium fragment of the 2AP/ᵗBuDA single-insertion product *tert*-butyl (2-hydroxyphenyl)glycinate. No parent ion could be detected for *tert*-butyl (2-hydroxyphenyl)glycinate, probably on account of the increased probability of fragmentation of a *tert*-butyl group relative to an ethyl group. The chromatogram for the reaction between BnDA and 2AP exhibits a dominant single peak at 16.01 minutes, preluded by a single smaller peak at 15.1 minutes which corresponds to unreacted 2AP **(SI Figure 1)**. Subsequent LC-MS analysis of the peak at 16.01 minutes generated a mass spectrum dominated by three peaks: i) a peak at 91 m/z corresponding to the tropylium ion fragment previously observed for ESI-MS analysis of benzyl substituents, ii) a peak at 212 m/z corresponding to the disodium carboxylate fragment resulting from loss of the tropylium fragment, and iii) a peak at 258 m/z which corresponds to the exact mass of the [M+H]⁺ parent ion for benzyl (2-hydroxyphenyl)glycinate.

The C18-HPLC chromatograms acquired for the C45-catalysed X-H insertion assays employing 2-aminophenol as the substrate revealed the presence of a new, single species being formed in the reaction. The di-insertion product was not detected in the LC-MS analysis, and the fragmentation patterns for all three diazo substrates are consistent with the patterns observed for aniline insertion.[36]

The data suggests that C45 is catalyzing the carbene insertion reaction of 2-aminophenol chemoselectivity, inserting the carbene exclusively into a N-H bond, even in the presence of unprotected O-H groups. From the chromatograms, average products yields of 90.78, 59.61 and 75.26% and ($N_{insertion}/O_{insertion}$) ratios of >99.9% were calculated for ethyl-, *tert*-butyl-, and benzyl (2-hydroxyphenyl)glycinate respectively (**Table 1**). These values are consistent with the product yields reported in the literature using copper-based catalysts.[36] The relative trends observed for product yield differences between the three diazo compounds (EDA > BnDA > [t]BuDA) for the previously reported cyclopropanation and N-H insertion reactions is also observed for 2-aminophenol.[36] It should be explicitly noted that the calculated yields are estimations based on comparisons with the unhydroxylated *N*-phenylglycine ethyl ester. We were unable to perform the synthesis of the three hydroxylated standards chemoselectively and so *N*-phenylglycine ethyl ester was selected as the closest representation for establishing an external calibration. We calculated the ratios of the extinction coefficients for aniline vs 2AP, 3AP, and 4AP as 1.07, 1.10 and 1.12 respectively, meaning the employment of *N*-phenylglycine ethyl ester should allow reasonable estimations to be made. In addition, the initial reactions were performed with a 1:1 substrate/diazo ratio, and although the chemoselectivity of the reaction was not affected the product yields were low; a 3:1 ratio has been previously demonstrated to engender consistently higher product yields[36] and so was employed in the assays reported here.

Figures 5 and **6** present the C18-HPLC chromatograms and the LC-MS mass spectra acquired for the C45-catalysed reaction between 3-aminophenol/4-aminophenol and EDA, [t]BuDA, and BnDA respectively. The trend of chemoselective mono-N-H-insertion is continued across all six reactions explored. The chromatograms for the reactions between EDA, [t]BuDA, and BDA and 3AP are all dominated by a single major peak at between 15.10-15.60 minutes, with subsequent LC-MS analysis confirming the identity of the single-insertion products for each carbene precursor. From the chromatograms, products yields of 61.50, >99.9 and 44.36% and ($N_{insertion}/O_{insertion}$) ratios of >99.9% were calculated for ethyl-, *tert*-butyl-, and benzyl (3-hydroxyphenyl)glycinate respectively (**Table 1**). The chromatogram for the reaction between EDA and 4AP is dominated by a major peak at 15.10 minutes, with no unreacted 4AP peak detectable (it should be noted that 3 M HCl was employed in the purification of assays conducted with 4AP instead of acetic acid. The increased acidity of HCl likely converted most of the unreacted 4-aminophenol to the hydrochloride salt, enhancing its removal from the product mixture upon aqueous-organic work up). Subsequent LC-MS analysis confirmed the identity of the 4AP/EDA single-insertion product. The chromatogram for the reactions between [t]BuDA/BDA and 4AP exhibits single peaks at 15.42 and 15.50 minutes respectively, with subsequent LC-MS analysis confirming the identity of the single insertion products. From the chromatograms,

product yields of 47.35, 19.37 and 34.34% and ($N_{insertion}$/$O_{insertion}$) ratios of >99.9% were calculated for ethyl-, *tert*-butyl-, and benzyl (4-hydroxyphenyl)glycinate respectively (**Table 1**).

One point of interest from these results is the C45-catalysed N-H insertion reaction between ^{t}BuDA and 3-aminophenol, which consistently affords high product yields. The origin of the apparent activity exhibited by C45 towards this reaction is currently unknown. The bulky *tert*-butyl group possibly disrupts the active-site of C45 upon metallocarbenoid formation causing temporary deviation in conformation which could facilitate the entry of a substrate (X-aminophenol) into the active site. Such deviations favouring the accommodation of the bulky *tert*-butyl-substituted metallocarbenoid intermediate could also engender stabilizing interactions for 3-aminophenol, and/or interactions which modify the *pK$_a$* of the -NH group, rendering the amine significantly more nucleophilic. Alternatively, it is also possible that such interactions could orientate the 3-aminophenol amine group suitably for nucleophilic attack on the metallocarbenoid carbon. The exact dynamics and electronics governing the reaction between ^{t}BuDA and 3-aminophenol remain unelucidated and require additional investigation to be fully understood.

From the data, it can be concluded that C45 is proficient at catalyzing the chemoselective N-H carbene insertion of unprotected alkanolamines. Notwithstanding the reaction between ^{t}BuDA and 3AP, two general trends can be identified with respect to the estimated product yields: i) the reaction yields follow the trend EDA > BnDA > ^{t}BuDA, and ii) the reaction yields seemingly increase as the molecular volume of the substrate decreases (1.38, 1.51 and 1.61 x10^{-22} cm^3 for 2AP, 3AP and 4AP respectively). The first trend has been observed previously for styrene cyclopropanation and *p*-chloroaniline insertion[36] and has been rationalized by the sterics of the three substituents, and possible π-π interactions between the phenyl group of benzyl diazoacetate and aromatic residues in the active site of C45.[36] The increase in product yield with respect to decreasing molecular volume suggests a key factor governing C45's functionality is entry and accommodation of a substrate into the heme-containing active site, a process easier for smaller molecules. The observed general trend in product yields (2AP > 3AP > 4AP) can therefore be accounted for, at least partially, by the molecular volume of the substrate (2AP < 3AP < 4AP). No observable trend was detected between product yields and *pK$_a$*(NH) of the substrates (*pK$_a$*(NH) = 4.37 (3AP) < 4.78 (2AP) < 5.48 (4AP)), indicating entry into the active site, and not the nucleophilicity of the amine, plays a more critical role in governing enzymatic activity.

CONCLUSIONS

C45, a tetra-α-helical c-type *de novo* hemeprotein, has previously been reported to be a proficient carbene transferase, catalysing the insertion of carbene intermediates into N-H σ-bonds to form new

N-C σ-bonds.[36] Here it was reported that C45's catalytic activity extends to amino substrates also possessing unprotected O-H σ-bonds. It has been demonstrated that C45 catalyses the N-H insertion reaction chemoselectively, exhibiting no reactivity towards unprotected O-H bonds in either phenol or any of the three regioisomers of aminophenol. Although chemoselective carbene N-H insertion by an enzyme has been reported for amine-containing substrates also containing alkenes,[47,48] thiols,[47] and silanes,[28] the work presented here is, to the best of the author's knowledge, the first demonstration of an enzymatic chemoselective N-H insertion in the presence of an unprotected hydroxy group. The estimated product yields followed the trends 2-aminophenol > 3-aminophenol > 4-aminophenol and EDA > BnDA > ᵗBuDA, which has been rationalised on account of sterics and substrate entry. Interestingly, the reaction between 3-aminophenol and ᵗBuDA exhibited remarkably high activity, marked by an apparent complete turnover of substrate into product. The reasons governing C45's predilection for this reaction are currently unclear and will be the subject of further investigations. In conclusion, it has been demonstrated that C45 shows proficient activity towards the carbene N-H insertion of aminophenols, marked by a chemoselectivity for amine groups in the presence of unprotected hydroxy groups.

METHODS AND MATERIALS

All chemicals were purchased from either Sigma or Fisher Scientific, and T7 Express competent cells and Q5 polymerase were purchased from NEB. All LC-MS mass spectra were acquired using positive electron-spray-ionization (ESI) mass spectrometry (Waters Xevo G2-XS QTof) attached to a C8 reverse column (Grace Vydac, 100 x 21 mm, 5 µm. All High Performance Liquid Chromatography (HPLC) chromatograms were acquired using an Agilent waters HPLC and a C18 HPLC reverse phase column (Phenomenex, 150 x 15 mm, 5 µm).

GENERAL & MOLECULAR BIOLOGY, PROTEIN EXPRESSION AND PURIFICATION

All transformations were performed using *E. coli* T7 Express competent cells (NEB) and by following a previously outlined protocol.[43] Briefly, a modified pMaI-p4x+ plasmid (the maltose binding protein (MBP) sequence is deleted), containing the sequence for C45 with a TEV-cleavable N-terminal His_6 tag, and pEC86 (containing the *ccm* machinery required for incorporation of the *c*-type heme into C45) expression vectors were used. 50% glycerol stocks for each cell-line were prepared and frozen for all proteins.

EXPRESSION PROTOCOL

C45 was prepared *via* by a protocol previously outlined.[43,46] All steps were performed under sterile conditions. Overnight starter cultures were prepared by adding 100 µL aliquots of carbenicillin (50 mg/mL), and 100 µL chloramphenicol (50 mg/mL) to 100 mL of LB before inoculating with a C45-containing glycerol stock. Starter cultures were incubated overnight at 37 °C and 180 rpm. To 1 L of LB was added 1mL of the appropriate antibiotics (50 mg/mL) and 50mL of an overnight starter culture, which were left to grow until an OD_{600} value between 0.6-0.8 was obtained. Once at the correct OD, 1 mL of the inducer IPTG (1 M stock, 1 mM induction concentration) was added and the induced solutions were left to express for an addition three hours. After three hours the cultures were collected, centrifuged (4000 rpm, 4 °C, 30 mins) and the cell pellets were isolated from the supernatant, resuspended in 10 mL lysis buffer (300 mM NaCl, 50 mM sodium phosphate, 20 mM imidazole, pH 8) and stored at -20 °C until purification.

PURIFICATION PROTOCOL

The lysis suspensions were defrosted at room temperature and sonicated multiple times in durations of 20 seconds. After sonication, the suspensions were transferred to 50 mL SS34 centrifuge tubes and centrifuged (SS34 centrifuge) at 4 °C and a rate of 18000 rpm for 30 minutes. The supernatant was subsequently transferred to a 50 mL microfuge tube and the cell pellets disposed of. The supernatant was filtered *via* 45 µm and 20 µm membrane filters and transferred onto a nickel column under lysis

buffer (300 mM NaCl, 50 mM Sodium phosphate (dibasic), 20 mM imidazole, pH 8). C45 contains a His$_6$-tag, which exhibits a high affinity for nickel, meaning the the His$_6$-tagged C45 remained attached to the nickel column while the unwanted components of the solution were filtered off. After the unwanted components had eluted into the waste, the column lines were switched to elution buffer (300 mM NaCl, 50 mM Sodium phosphate (dibasic), 250 mM imidazole, pH 8) which passed through the column, displacing His-tagged C45 and eluting it off the column; the His-tagged C45 was then re-suspend in TEV protease buffer (0.5 mM EDTA, 50 mM Tris, 5L Milli-Q, pH 8) *via* dialysis. Once in TEV buffer, excess DDT and TEV protease was added (under a nitrogen atmosphere) and the mixture was left to stir for 3-6 hours. The mixture was then filtered again (45 and 20um membrane filters) and transferred back onto a nickel column. After elution, the His-tag cleaved C45, the protein was concentrated down to 5 mL using a 10,000-spin-membrane microfuge and centrifuging at 4500 rpm and 4°C. The protein was then purified using a HighLoad S75 16/600 (GE Healthcare) gel filtration column employing CHES buffer (20 mM CHES, 100 mM KCl, pH 8.6) as the mobile phase. The progress of the protein along the column was monitored spectroscopically, at 280 nm, and the collected fractions containing the C45 were characterised *via* UV-spectroscopy and an SDS-PAGE gel. The best fractions were then combined, concentrated (if necessary) and aliquoted into 500 μL samples which were stored at -80 ^{o}C until required.

UV-VIS SPECTROSCOPY

All UV-VIS spectra were recorded on an Agilent Cary-60 UV-visible spectrophotometer using a 1 mm pathlength quartz cuvette. Reduced spectra for each sample were acquired by the addition of 10 μL of Na$_2$S$_2$O$_4$ (1 mM in de-ionised water). All UV-VIS spectra were recorded across a wavelength range of 300-750 nm. The concentrations were determined using the extinction coefficients for the ferric and ferrous samples respectively. C45: ε_{406} = 147,433 M^{-1} cm^{-1} and ε_{417} = 119,900 M^{-1} cm^{-1}.

CARBENE TRANSFER CHEMISTRY

Unless stated otherwise, all assays were conducted under scrubbed nitrogen inside an anaerobic glovebox ([O$_2$] < 5 ppm; Belle Technology). The assays were conducted inside 1.5 mL screw top vials sealed with a silicone-septum containing cap, in accordance with a previously reported protocol.[36] All assays were conducted in CHES buffer (100 mM KCl, 20 mM CHES, pH 8.6). The final reaction volumes for all assays were 400 μL unless otherwise stated.

Chemoselective N-H insertion assays

The assay were conducted in accordance with a previously established protocol.[36] Generally, to 370 μL of a 10 μM C45 solution was added 10 μL of Na$_2$S$_2$O$_4$ (400 mM stock; deionized water) to afford

complete reduction of the C45. 10 μL of the selected aminophenol (1.2 M stock in DMSO) was the added and the reaction left to mix for 30 seconds. Another vial with the selected diazo compound (400 mM stock in EtOH) was deoxygenated together with the assay vial. The vials were tightly sealed, transferred out of the glovebox and cooled in an ice bath. Once the samples had sufficiently cooled, 10 μL of the diazo compound was added *via* gastight syringe into the reaction vials to initiate the reaction. Final reaction concentrations were 10 μM enzyme (0.1% mol%), 10 mM sodium dithionite, 10 mM diazo compound, and 30 mM aminophenol. Once mixed, the vials were stirred on a roller at room temperature. After 2 hours, the reaction was quenched by the addition of 20 μL 10% (w/v) acetic acid (2-aminophenol and 3-aminophenol) or 3 M HCl (4-aminophenol). The vials were the unscrewed, and 1 mL of ethyl acetate was added. The solution was pipetted into a 1.5 mL microfuge tube, vortexed and centrifuged for 1 minute at 13,500 rpm. The upper organic layer was collected (and if necessary, dried with $MgSO_4$) for analysis. All the assays were analysed by C18-HPLC and LC-MS as described below. The product yields, ($N_{insertion}/O_{insertion}$) ratios and total turnover numbers (TTN; concentration of product formed/concentration of enzyme) were calculated *via* external calibration with commercial *N*-phenylglycine ethyl ester and starting material.

O-H insertion assays

The C45-catalysed O-H insertion assays were conducted identically to the N-H insertion assays, with the only difference phenol substituting for an aminophenol substrate.

PRODUCT CHARACTERIZATION BY REVERSE PHASE AND C18-HPLC

All the reactions performed were quantified by High Performance Liquid Chromatography (HPLC). A C18 HPLC reverse phase column (Phenomenex, 150 x 15 mm, 5 μm) was used to quantify the chemoselective N-H insertion and O-H insertions assays. An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO, 0.5 mL min^{-1} flow rate and 20 μL injection volume) was used. All elution traces were monitored spectroscopically at 220, 254 and 280 nm. The C18 column allowed the retention times of the starting materials and the reaction products to be determined and allowed the ($N_{insertion}/O_{insertion}$) ratio to be quantified. If appreciable product formation could be detected after the initial HPLC experiments, the products were subsequently analysed using Liquid chromatography-Mass spectrometry (LC-MS) to assist in product identification (described below).

External calibrations

An external calibration of *N*-phenylglycine ethyl ester was obtained using HPLC with a commercially pure sample of *N*-phenylglycine ethyl ester. The conditions and mobile phases employed in the product analysis for each assay were also employed for the calibration. *N*-phenylglycine ethyl ester

was prepared at multiple concentrations (i.e. 50-500 µM, 1 mM, 2 mM, 5 mM, 7.5 mM, 10 mM, and 20 mM) and the peak height response in the chromatogram was recorded as a function of concentration. A plot of [*N*-phenylglycine ethyl ester] vs peak height engendered a straight line which could be used to determine the product yields and TTNs for each assay. Injection volumes were 8 µL to account for the 2.5:1 dilution factor in the product extraction phase. The following equations were used to calculate total turnover numbers (TTN) and ($N_{insertion}$/$O_{insertion}$) ratios:

$$TTN = \frac{[product]}{[enzyme]}$$

Equation 1

$$\frac{N_{insertion}}{O_{insertion}} = \frac{[N_{insertion}\ product]}{[N_{insertion}\ product] + [O_{insertion}\ product]} \times 100$$

Equation 2

PRODUCT CHARACTERIZATION BY LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY

Liquid chromatography-Mass spectrometry (LC-MS, C8) was employed to identify the products formed in each assay. A C8 reverse column (Grace Vydac, 100 x 21 mm, 5 µm) was used with a 10-minute gradient mobile phase (95:5% H_2O:CH_3CN to 10:90% H_2O:CH_3CN; 0.1% v/v formic acid, 0.25 mL min^{-1}). Injection volumes were 20 µL and the chromatogram was recorded using UV-VIS spectroscopy in the wavelength range 240-300 nm. After eluting from the column, the mixture entered an isocratic solvent chamber where a 1:100 dilution preceded the injection of the sample into a positive electron-spray-ionization (ESI) mass spectrometer (Waters Xevo G2-XS QTof). The mass spectrum recorded peaks screened across a m/z range of 70-300 (depending on the molecular mass of the product being screened). Retention times and fragmentation patterns were compared using commercial samples of *N*-phenylglycine ethyl ester and to previously reported LC-MS data.[36]

REFERENCES

1.   Shenvi, R. A., O'Malley, D. P. & Baran, P. S. Chemoselectivity: The mother of invention in total synthesis. *Acc. Chem. Res.* **42**, 530–541 (2009).

2.   Afagh, N. A. & Yudin, A. K. Chemoselectivity and the curious reactivity preferences of functional groups. *Angewandte Chemie - International Edition* **49**, 262–310 (2010).

3.   Jarowicki, K. & Kocienski, P. Protecting groups. *Journal of the Chemical Society. Perkin Transactions 1* **1**, 2109–2135 (2001).

4.   Isidro-Llobet, A., Álvarez, M. & Albericio, F. Amino acid-protecting groups. *Chem. Rev.* **109**, 2455–2504 (2009).

5.   Klán, P. *et al.* Photoremovable protecting groups in chemistry and biology: Reaction mechanisms and efficacy. *Chemical Reviews* **113**, 119–191 (2013).

6.   Rücker, C. The Triisopropylsilyl Group in Organic Chemistry: Just a Protective Group, or More? *Chem. Rev.* **95**, 1009–1064 (1995).

7.   Sartori, G. *et al.* Protection (and Deprotection) of Functional Groups in Organic Synthesis by Heterogeneous Catalysis. *Chemical Reviews* **104**, 199–250 (2004).

8.   Natoli, S. N. & Hartwig, J. F. Noble-Metal Substitution in Hemoproteins: An Emerging Strategy for Abiological Catalysis. *Acc. Chem. Res.* (2019). doi:10.1021/acs.accounts.8b00586

9.   Liang, D. M. *et al.* Glycosyltransferases: Mechanisms and applications in natural product development. *Chemical Society Reviews* **44**, 8350–8374 (2015).

10.   Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).

11.   Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N. & Dwek, R. A. Emerging principles for the therapeutic exploitation of glycosylation. *Science* **343**, (2014).

12.   Frear, D. S., Swanson, H. R. & Tanaka, F. S. Herbicide metabolism in plants. in *Recent Advances in Phytochemistry* **5**, 225–246 (Pergamon Press, 1972).

13.   Cleland, W. W. & Kennedy, E. P. The enzymatic synthesis of psychosine. *J. Biol. Chem.* **235**, 45–51 (1960).

14.   Key, H. M., Dydio, P., Clark, D. S. & Hartwig, J. F. Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **534**, 534–537 (2016).

15.   Renata, H., Wang, Z. J., Kitto, R. Z. & Arnold, F. H. P450-catalyzed asymmetric cyclopropanation of electron-deficient olefins under aerobic conditions. *Catal. Sci. Technol.* **4**, 3640–3643 (2014).

16.   Sreenilayam, G., Moore, E. J., Steck, V. & Fasan, R. Stereoselective Olefin Cyclopropanation under Aerobic Conditions with an Artificial Enzyme Incorporating an Iron-Chlorin e6 Cofactor. *ACS Catal.* **7**, 7629–7633 (2017).

17.   Tinoco, A., Steck, V., Tyagi, V. & Fasan, R. Highly Diastereo- and Enantioselective Synthesis of Trifluoromethyl-Substituted Cyclopropanes via Myoglobin-Catalyzed Transfer of Trifluoromethylcarbene. *J. Am. Chem. Soc.* **139**, 5293–5296 (2017).

18.   Knight, A. M. *et al.* Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. *ACS Cent. Sci.* **4**, 372–377 (2018).

19.   Heel, T., McIntosh, J. A., Dodani, S. C., Meyerowitz, J. T. & Arnold, F. H. Non-natural olefin cyclopropanation catalyzed by diverse cytochrome P450s and other hemoproteins. *ChemBioChem* **15**, 2556–2562 (2014).

20.   Coelho, P. S. *et al.* A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. *Nat. Chem. Biol.* **9**, 485–487 (2013).

21.   Wang, Z. J. *et al.* Improved cyclopropanation activity of histidine-ligated cytochromeP450 enables the enantioselective formal synthesis of levomilnacipran. *Angew. Chemie - Int. Ed.* **53**, 6810–6813 (2014).

22.   Key, H. M. *et al.* Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. *ACS Cent. Sci.* **3**, 302–308 (2017).

23.   Dydio, P. *et al.* An artificial metalloenzyme with the kinetics of native enzymes. *Science (80-.).* **354**, 102–106 (2016).

24.   Bordeaux, M., Tyagi, V. & Fasan, R. Highly diastereoselective and enantioselective olefin cyclopropanation using engineered myoglobin-based catalysts. *Angew. Chemie - Int. Ed.* **54**, 1744–1748 (2015).

25.   Wei, Y., Tinoco, A., Steck, V., Fasan, R. & Zhang, Y. Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. *J. Am. Chem. Soc.* **140**, 1649–1662 (2018).

26.   Sreenilayam, G., Moore, E. J., Steck, V. & Fasan, R. Metal Substitution Modulates the Reactivity and Extends the Reaction Scope of Myoglobin Carbene Transfer Catalysts. *Adv. Synth. Catal.* **359**, 2076–2089 (2017).

27.   Huang, X. *et al.* A Biocatalytic Platform for Synthesis of Chiral α-Trifluoromethylated Organoborons. *ACS Cent. Sci.* **5**, 270–276 (2019).

28.   Kan, S. B. J., Lewis, R. D., Chen, K. & Arnold, F. H. Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. *Science (80-.).* **354**, 1048–1051 (2016).

29.   Jennifer Kan, S. B., Huang, X., Gumulya, Y., Chen, K. & Arnold, F. H. Genetically programmed chiral organoborane synthesis. *Nature* **552**, 132–136 (2017).

30.   Vargas, D. A., Tinoco, A., Tyagi, V. & Fasan, R. Myoglobin-Catalyzed C–H Functionalization of Unprotected Indoles. *Angew. Chemie - Int. Ed.* **57**, 9911–9915 (2018).

31.   Moore, E. J., Steck, V., Bajaj, P. & Fasan, R. Chemoselective Cyclopropanation over Carbene Y-H Insertion Catalyzed by an Engineered Carbene Transferase. *J. Org. Chem.* **83**, 7480–7490 (2018).

32.   Sreenilayam, G. & Fasan, R. Myoglobin-catalyzed intermolecular carbene N-H insertion with arylamine substrates. *Chem. Commun.* **51**, 1532–1534 (2015).

33.   Tyagi, V., Bonn, R. B. & Fasan, R. Intermolecular carbene S-H insertion catalysed by engineered myoglobin-based catalysts. *Chem. Sci.* **6**, 2488–2494 (2015).

34.   Weissenborn, M. J. *et al.* Enzyme-Catalyzed Carbonyl Olefination by the E. coli Protein YfeX in the Absence of Phosphines. *ChemCatChem* **8**, 1636–1640 (2016).

35.   Tyagi, V. & Fasan, R. Myoglobin-catalyzed olefination of aldehydes. *Angew. Chemie - Int. Ed.* **55**, 2512–2516 (2016).

36.   Stenner, R., Steventon, J. W., Seddon, A. & Anderson, J. L. R. A de novo peroxidase is also a

promiscuous yet stereoselective carbene transferase. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 1419–1428 (2020).

37. Schoemaker, H. E., Mink, D. L. & WubboLts, M. G. Dispelling the myths - Biocatalysis in industrial synthesis. *Science* **299**, 1694–1697 (2003).

38. Strohmeier, G. A., Pichler, H., May, O. & Gruber-Khadjawi, M. Application of designed enzymes in organic synthesis. *Chemical Reviews* **111**, 4141–4164 (2011).

39. Grayson, K. J. & Anderson, J. R. The ascent of man(made oxidoreductases). *Current Opinion in Structural Biology* **51**, 149–155 (2018).

40. Lin, Y. W. Rational design of metalloenzymes: From single to multiple active sites. *Coordination Chemistry Reviews* **336**, 1–27 (2017).

41. Barber, J. *et al.* Maquette Strategy for Creation of Light- and Redox-Active Proteins. *Photosynth. Bioenerg.* 1–33 (2017). doi:10.1142/9789813230309_0001

42. Yu, F. *et al.* Protein design: Toward functional metalloenzymes. *Chemical Reviews* **114**, 3495–3578 (2014).

43. Watkins, D. W. *et al.* A suite of de novo c-type cytochromes for functional oxidoreductase engineering. *Biochim. Biophys. Acta - Bioenerg.* (2016). doi:10.1016/j.bbabio.2015.11.003

44. Robertson, D. E. *et al.* Design and synthesis of multi-haem proteins. *Nature* **368**, 425–432 (1994).

45. Koder, R. L. & Dutton, P. L. Intelligent design: the de novo engineering of proteins with specified functions. *Dalton Trans.* 3045–3051 (2006). doi:10.1039/b514972j

46. Watkins, D. W. *et al.* Construction and in vivo assembly of a catalytically proficient and hyperthermostable de novo enzyme. *Nat. Commun.* **8**, (2017).

47. Moore, E. J., Steck, V., Bajaj, P. & Fasan, R. Chemoselective Cyclopropanation over Carbene Y-H Insertion Catalyzed by an Engineered Carbene Transferase. *J. Org. Chem.* **83**, 7480–7490 (2018).

48. Wang, Z. J., Peck, N. E., Renata, H. & Arnold, F. H. Cytochrome P450-catalyzed insertion of carbenoids into N-H bonds. *Chem. Sci.* **5**, 598–601 (2014).

FIGURE LEGENDS

Figure 1: The three possible reaction products following the C45-catalysed O/N-H insertion reaction between a regioisomer of aminophenol and EDA.

Figure 2: C18-HPLC chromatograms for phenol (red) and the C45-catalysed (10 μM, 0.1 % catalyst loading) O-H insertion assay between phenol (30 mM) and EDA (10 mM) (blue). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 μl.

Figure 3: The C45-catalysed chemoselective N-H insertion reaction between 2-/3-/4-aminophenol and the diazo compounds ethyl/*tert*-butyl/benzyl diazoacetate.

Figure 4: (**top**) C18-HPLC chromatograms for the C45-catalysed (10 μM, 0.1 % catalyst loading) N-H insertion assays between 2-aminophenol (30 mM) and (**A**) EDA (10 mM), (**B**) ᵗBuDA (10 mM), and (**C**) BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 μl. (**bottom**) LC-MS spectra for the C45-catalysed (10 μM, 0.1 % catalyst loading) N-H insertion assays between 2-aminophenol (30 mM) and (**D**) EDA (10 mM), (**E**) ᵗBuDA (10 mM), and (**F**) BnDA (10 mM). All spectra were recorded in ES+ mode

and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

**Figure 5:** (**top**) C18-HPLC chromatograms for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 3-aminophenol (30 mM) and (**A**) EDA (10 mM), (**B**) ^tBuDA (10 mM), and (**C**) BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). An isocratic mobile phase (100% CH$_3$OH: 0.1% w/v NH$_4$AcO) was employed and injection volumes were 20 µl. (**bottom**) LC-MS spectra for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 3-aminophenol (30 mM) and (**D**) EDA (10 mM), (**E**) ^tBuDA (10 mM), and (**F**) BnDA (10 mM). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

Figure 6: (**top**) C18-HPLC chromatograms for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 4-aminophenol (30 mM) and (**A**) EDA (10 mM), (**B**) ^tBuDA (10 mM), and (**C**) BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). An isocratic mobile phase (100% CH$_3$OH: 0.1% w/v NH$_4$ACO) was employed and injection volumes were 20 µl. (**bottom**) LC-MS spectra for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 4-aminophenol (30 mM) and (**D**) EDA (10 mM), (**E**) ^tBuDA (10 mM), and (**F**) BnDA (10 mM). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

Table 1: The retention times, average product yields, and nitrogen/oxygen insertion ratios for the C45-catalysed N-H insertion reactions between the three regioisomers of aminophenol and the diazo compounds EDA, ^tBuDA and BnDA (all reactions formed in triplicate). The reaction yields are all the closest possible estimations that can be made because *N*-phenylglycine ethyl ester was employed as the standard for establishing an external calibration.
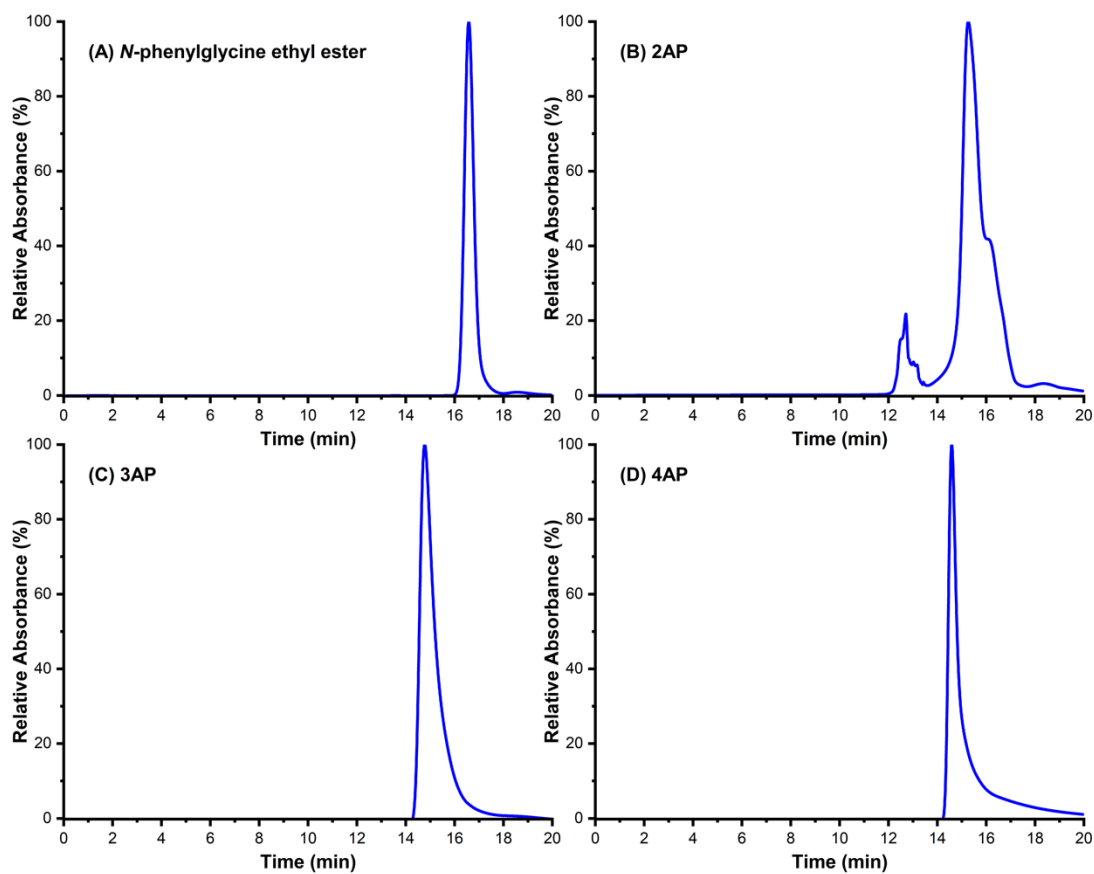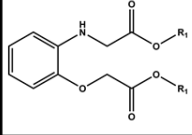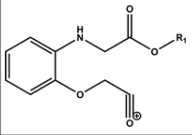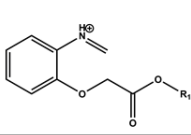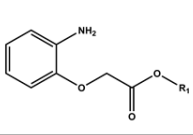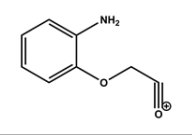
SI Figure 1: C18-HPLC chromatogram for (**A**) *N*-phenylglycine ethyl ester (10 mM), (**B**) 2-aminophenol (10 mM), (**C**) 3-aminophenol (10 mM), and (**D**) 4-aminophenol (10 mM). All standards were prepared in in EtOH. An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4ACO) was employed and traces were recorded at 254 nm; all injection volumes were 20 µl.
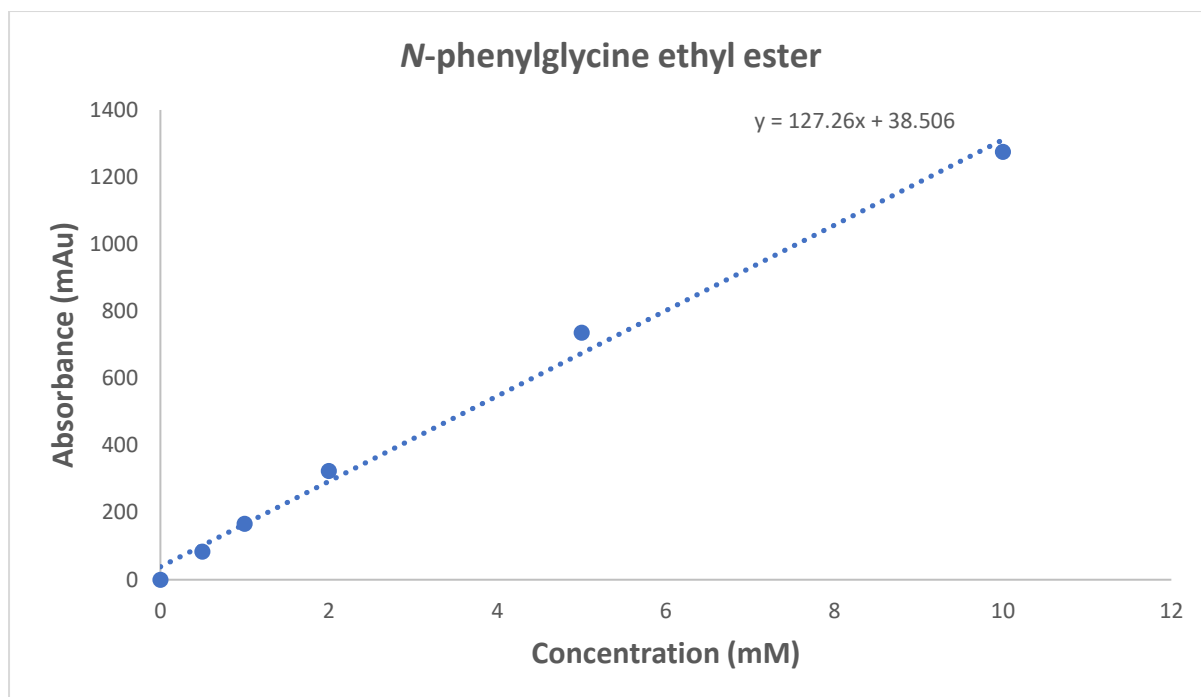
**SI Figure 2:** C18-HPLC external calibrations for *N*-phenylglycine ethyl ester at 254 nm; An isocratic mobile phase (100% CH₃OH: 0.1% w/v NH₄ACO) was employed and traces were recorded at 254 nm; all injection volumes were 8 µl.
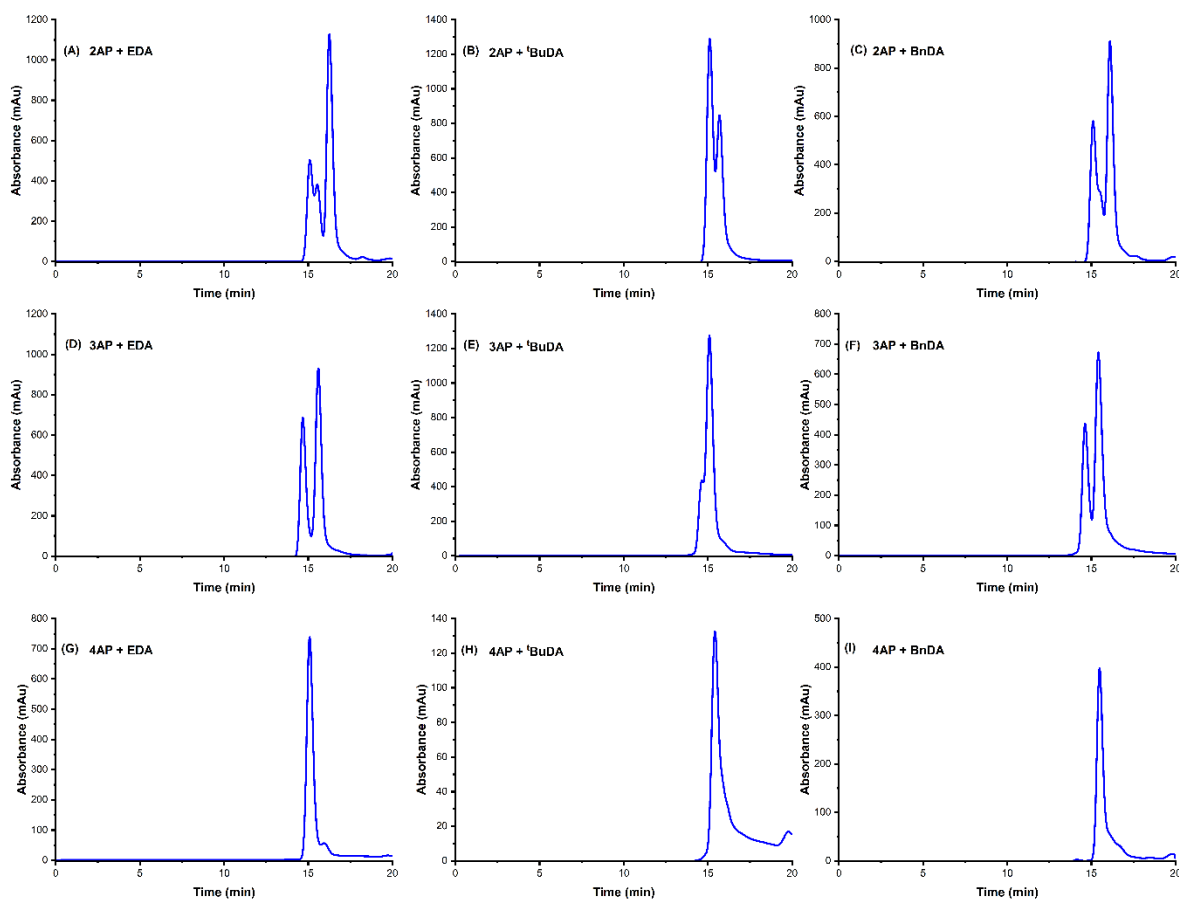
SI Figure 3: C18-HPLC chromatogram for the C45 (10 μM) catalyzed N-H insertion assays. (**A**) 2-aminophenol (30 mM) and ethyl diazoacetate (EDA; 10 mM); (**B**) 2-aminophenol (30 mM) and *tert*-butyl diazoacetate (ᵗBuDA; 10 mM); (**C**) 2-aminophenol (30 mM) and benzyl diazoacetate (BnDA; 10 mM); (**D**) 3-aminophenol (30 mM) and ethyl diazoacetate (10 mM); (**E**) 3-aminophenol (30 mM) and *tert*-butyl diazoacetate (10 mM); (**F**) 3-aminophenol (30 mM) and benzyl diazoacetate (10 mM); (**G**) 4-aminophenol (30 mM) and ethyl diazoacetate (10 mM); (**H**) 4-aminophenol (30 mM) and *tert*-butyl diazoacetate (10 mM); (**I**) 4-aminophenol (30 mM) and benzyl diazoacetate (10 mM). All assays were performed in CHES buffer (pH 8.6) with 10 μM C45 (0.1 % catalyst loading). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4ACO) was employed and traces were recorded at 254 nm; all injection volumes were 20 μl.

R = Et	281	236	208	195	150
R = ^{t}Bu	337	264	236	223	150
R = Bz	433	312	284	271	150

SI Table 1: List of potential side products and the anticipated ESI-MS fragments following the reaction of 2-AP and xDA (x = E, ^{t}B, Bn) (fragments for 3-aminophenol and 4-aminophenol would have identifcal masses).
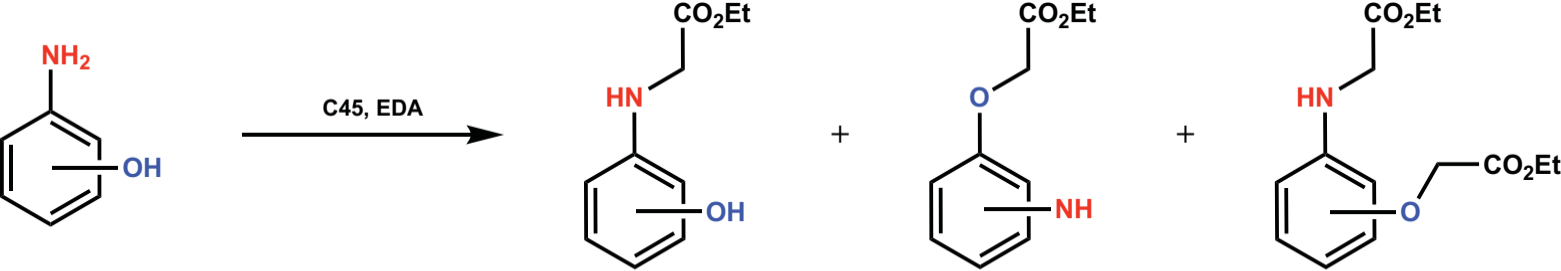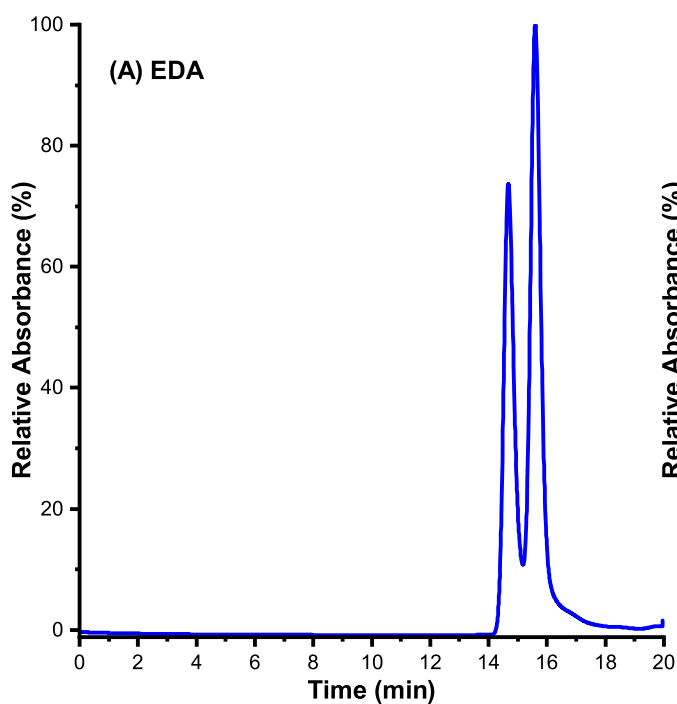
Three possible X-H insertions

R = Et, ^tBu, Bn

(A) EDA

(B) ᵗBuDA

(C) BnDA

(D) EDA (122.07)

(E) ᵗBuDA (122.07)

(F) BnDA (91.0...
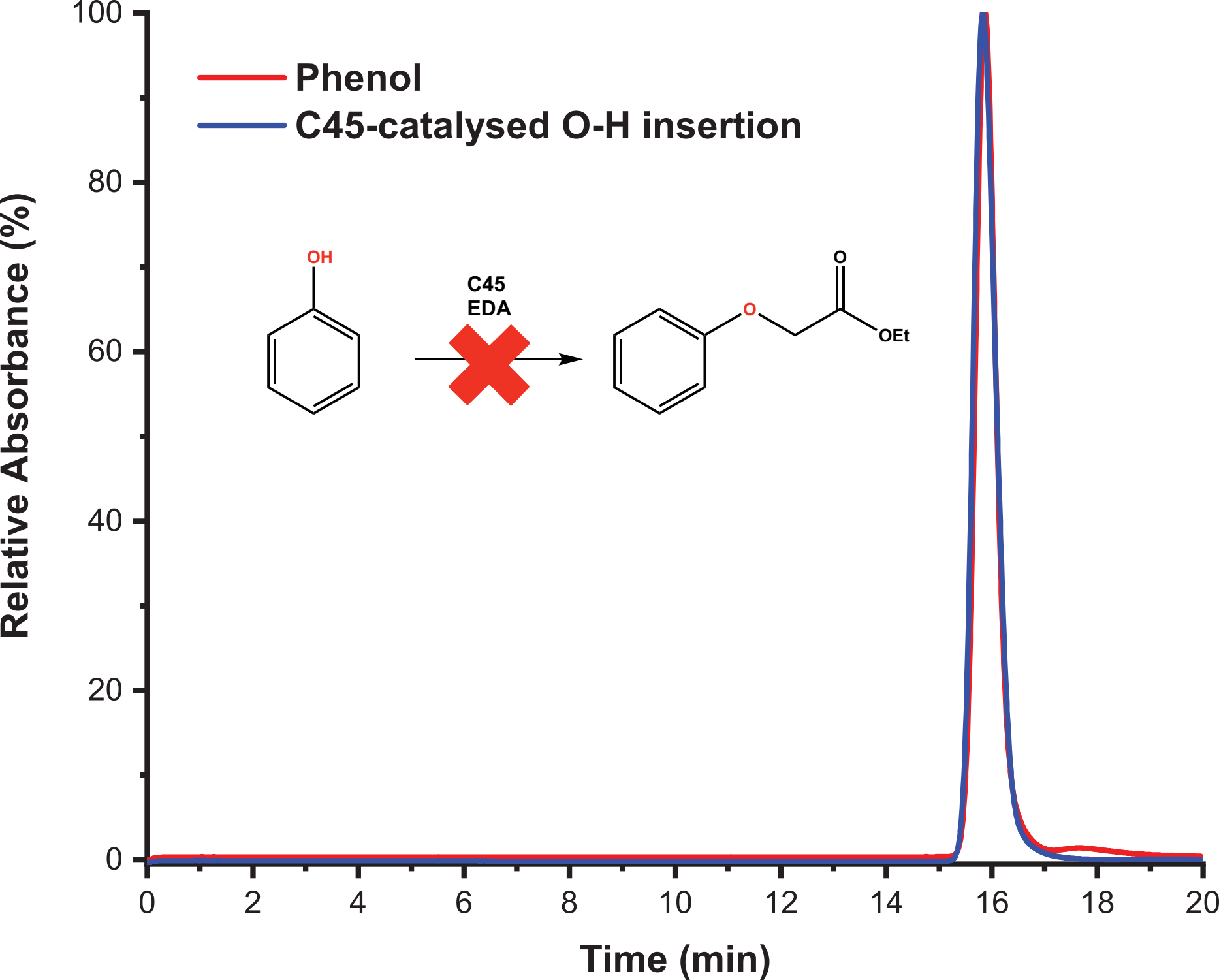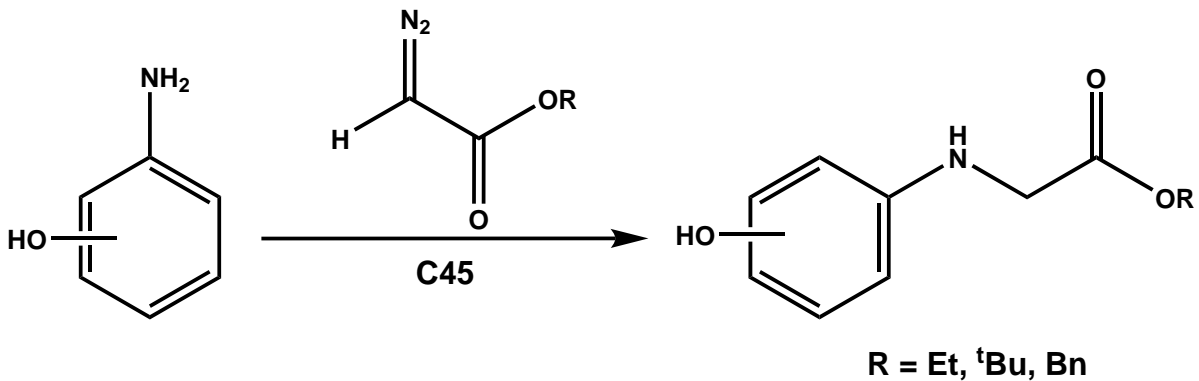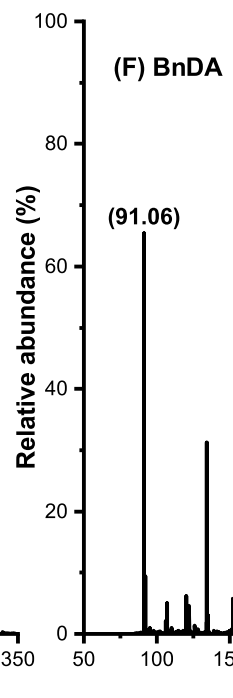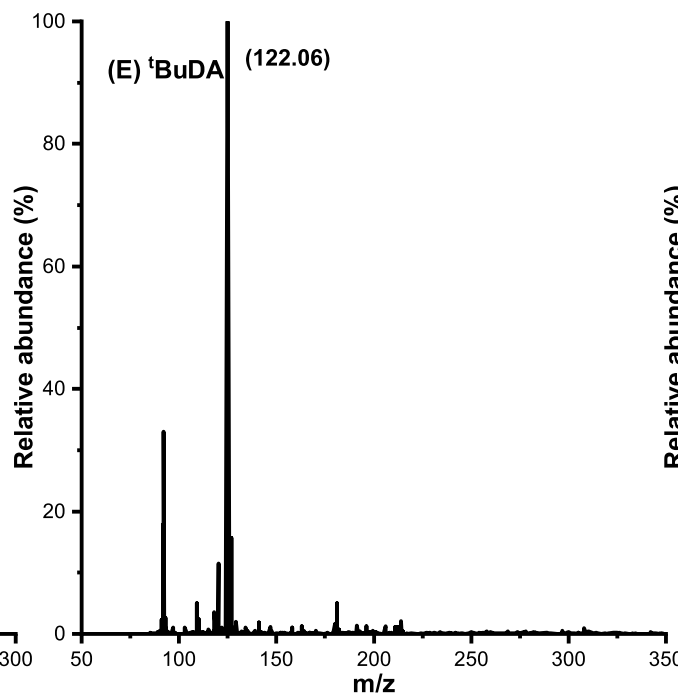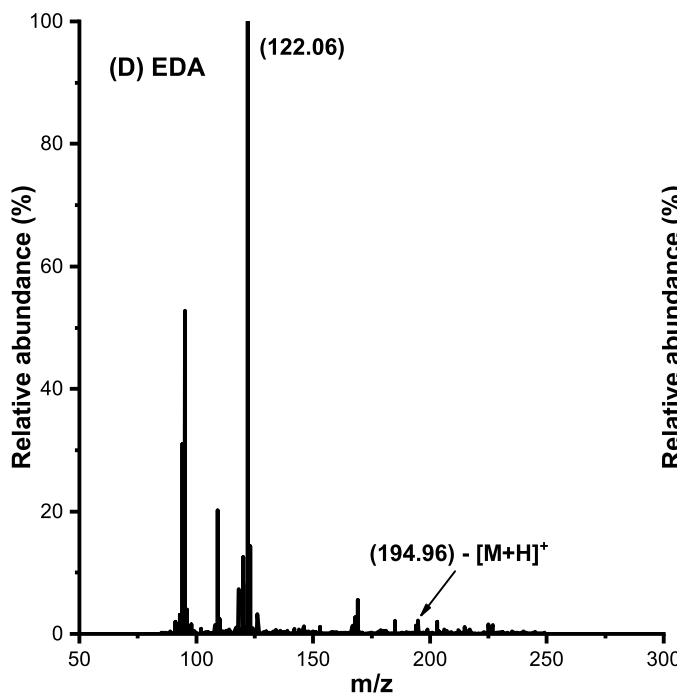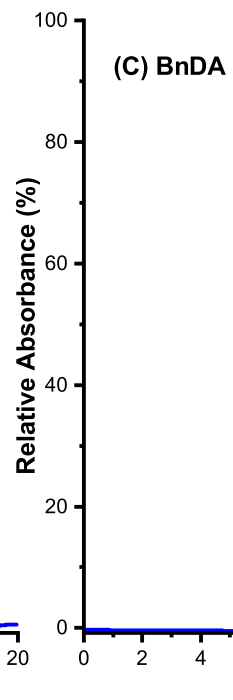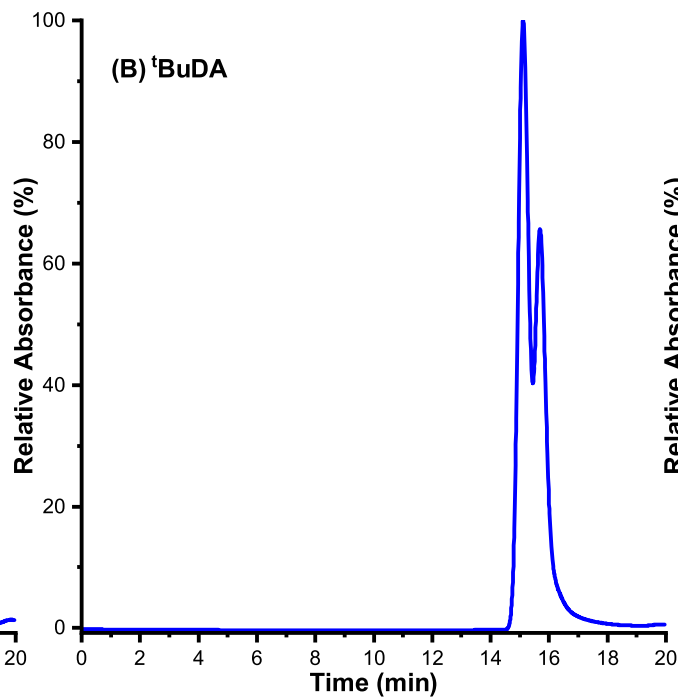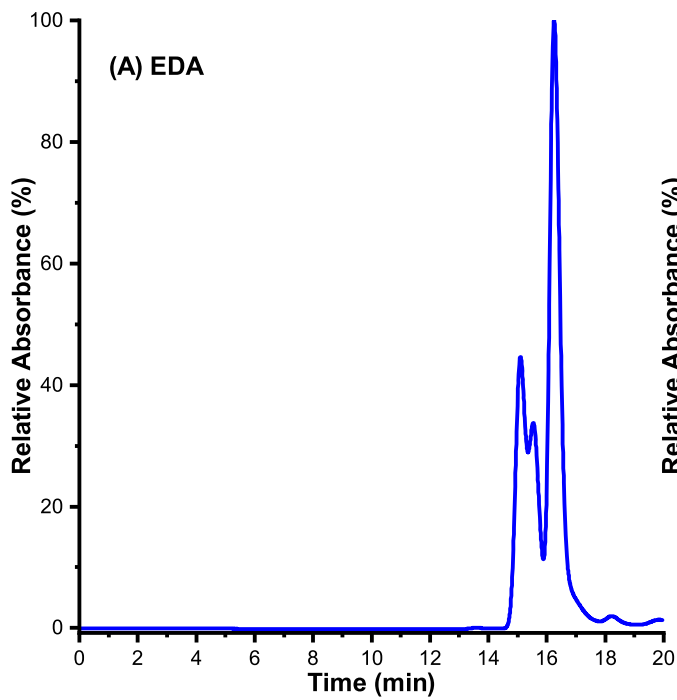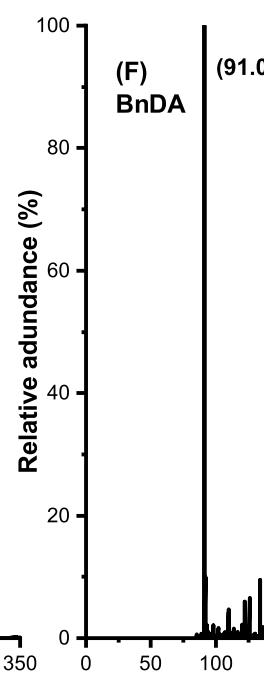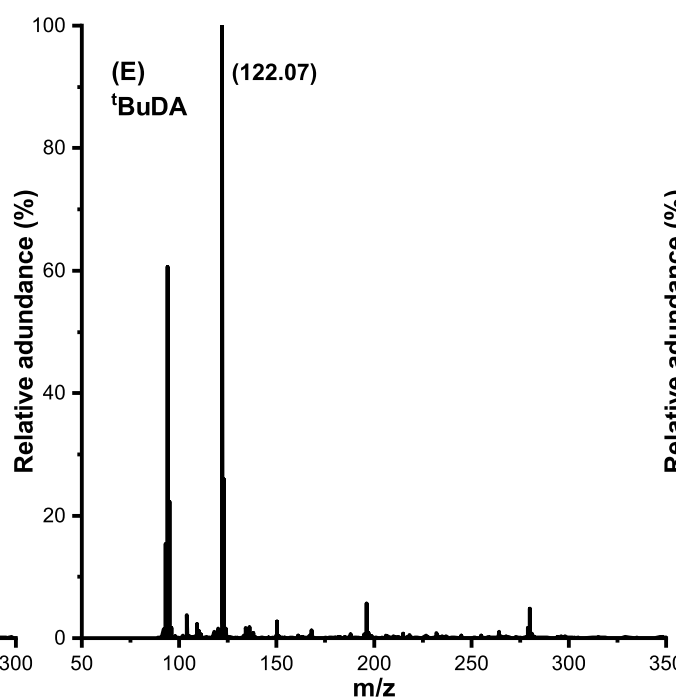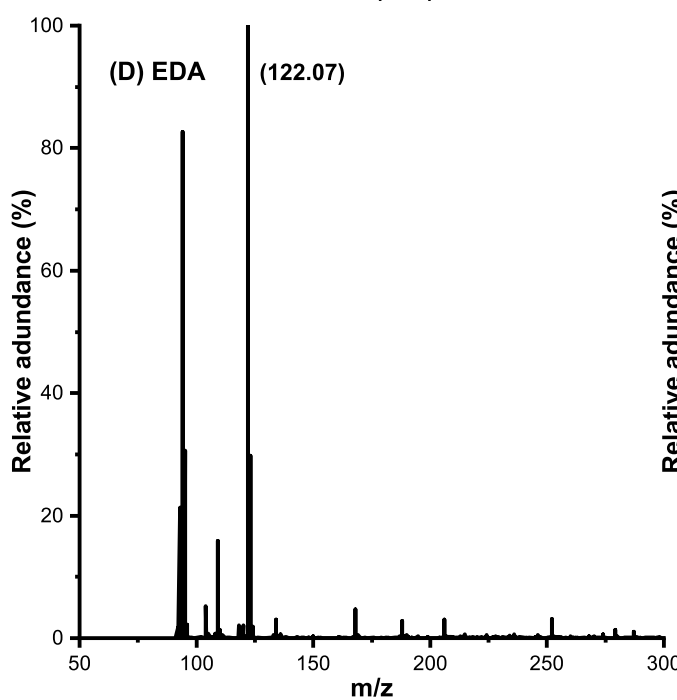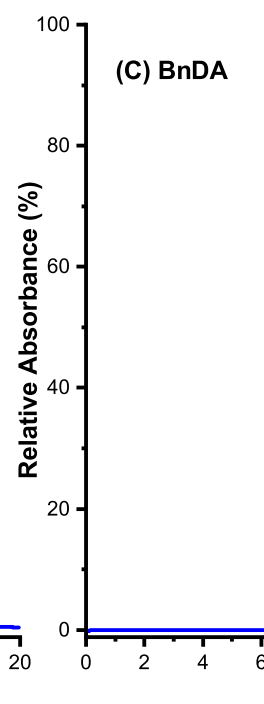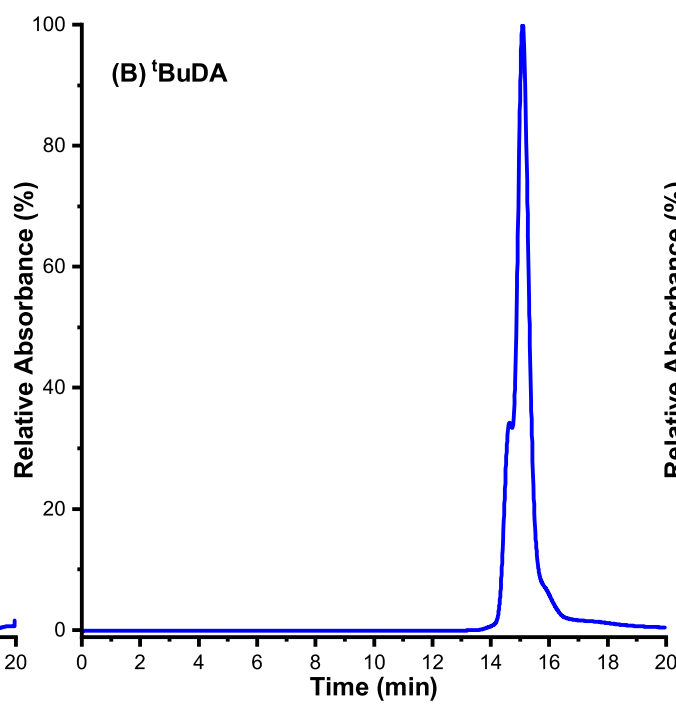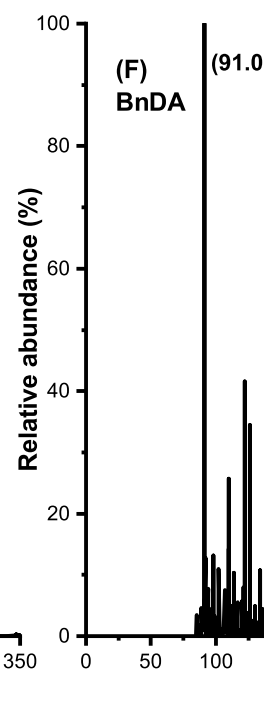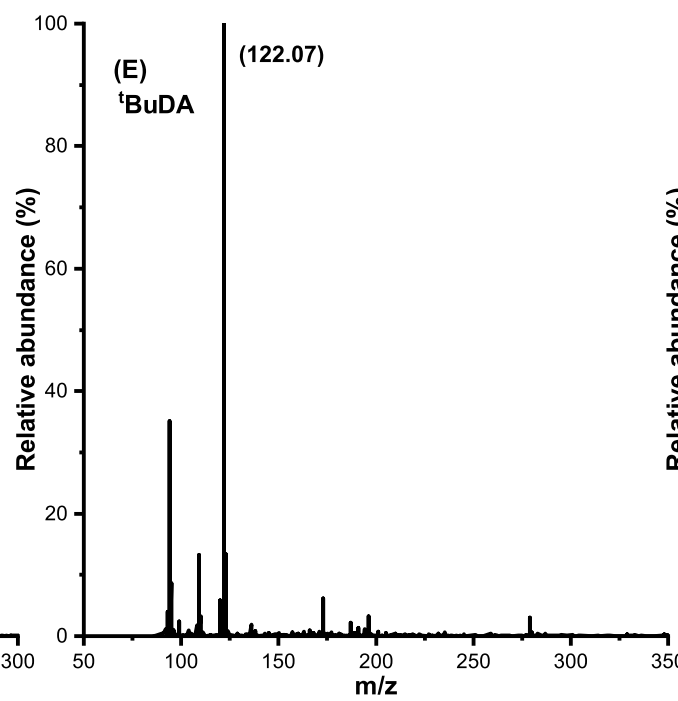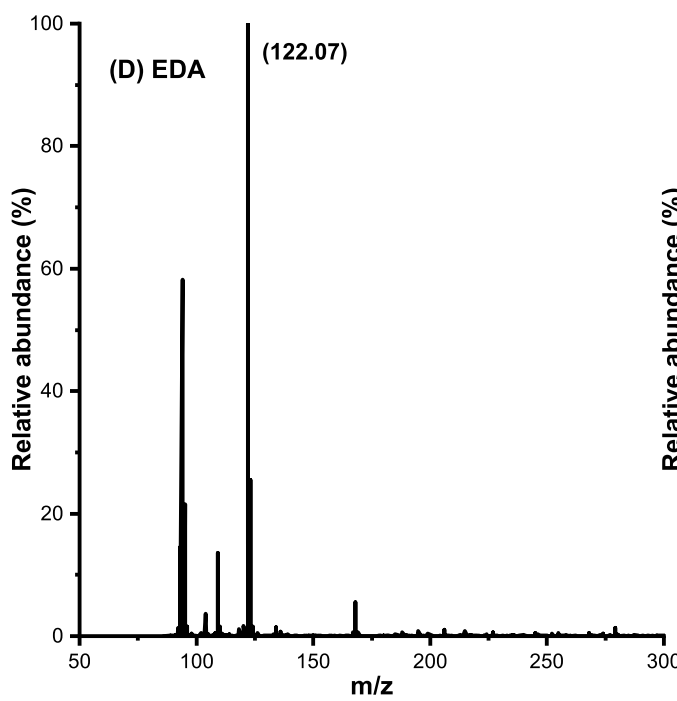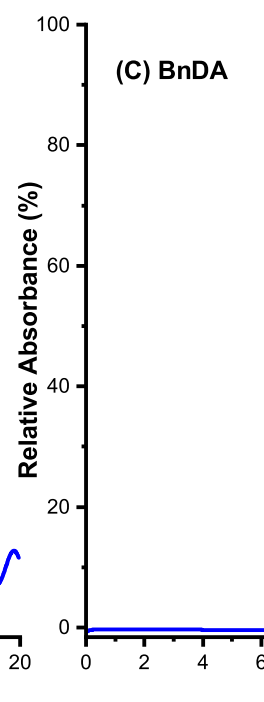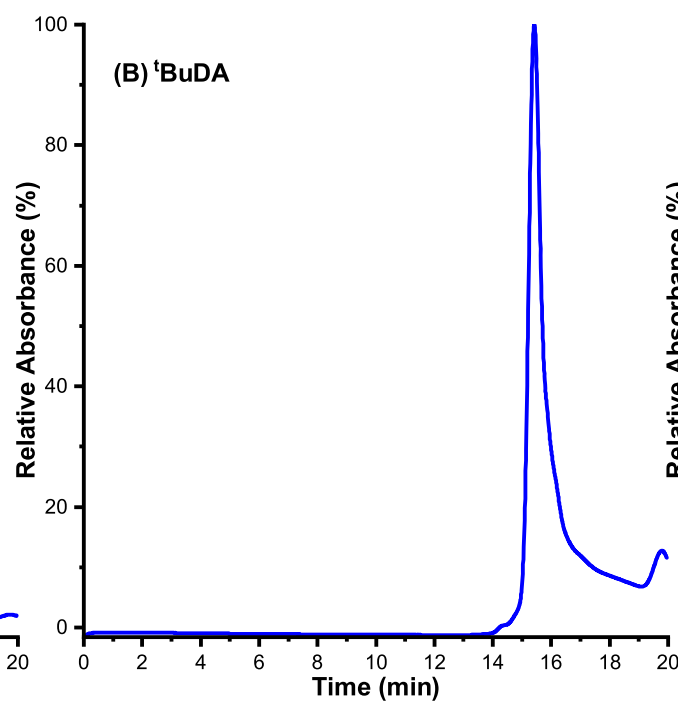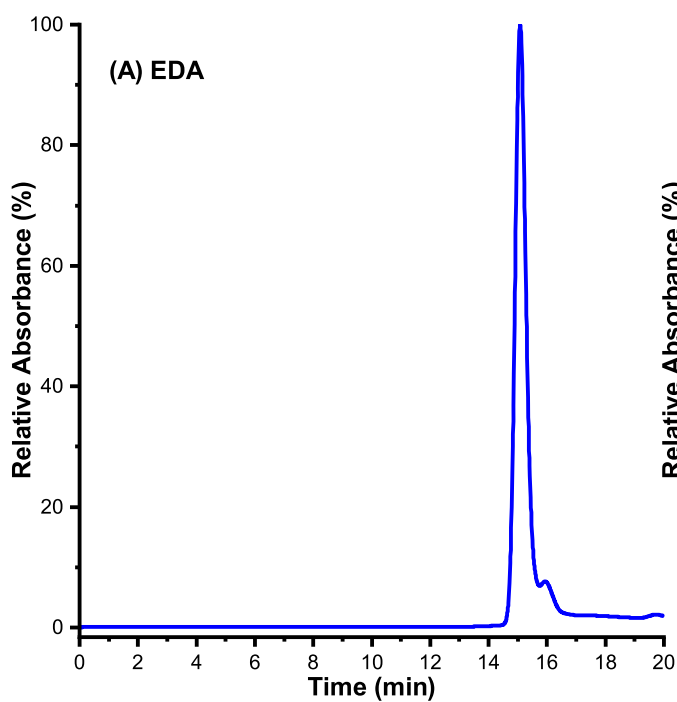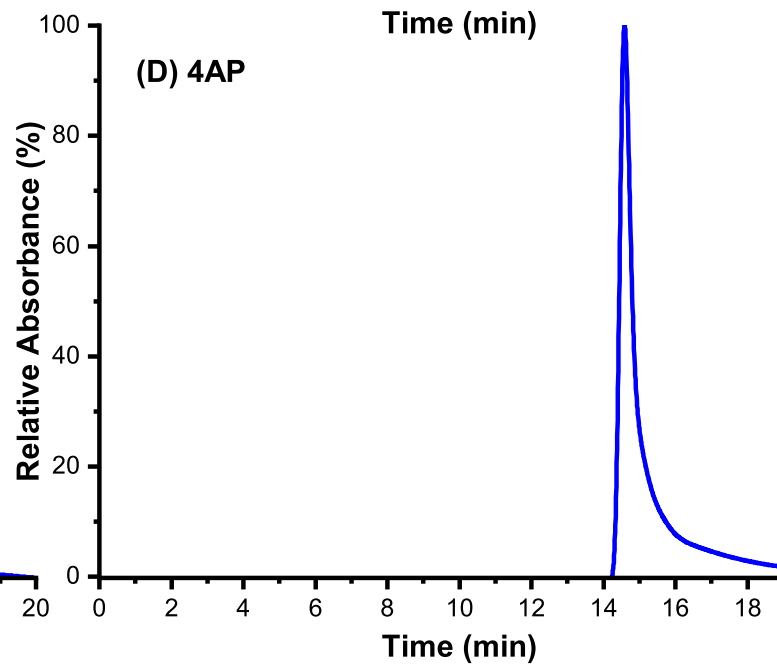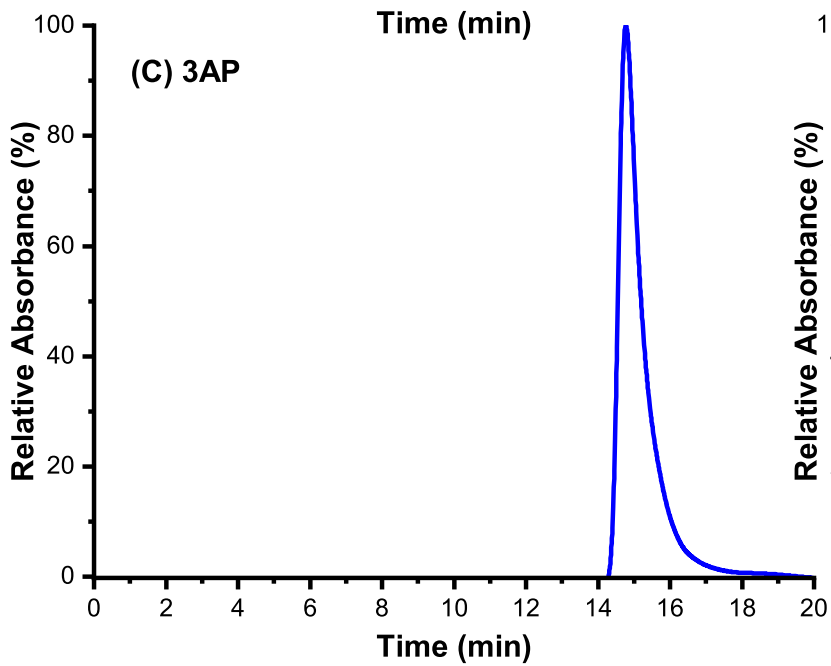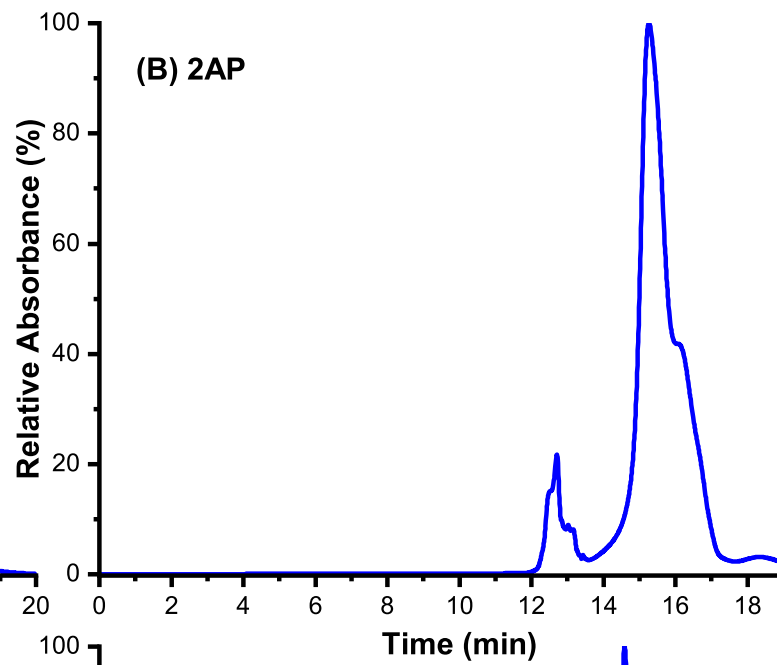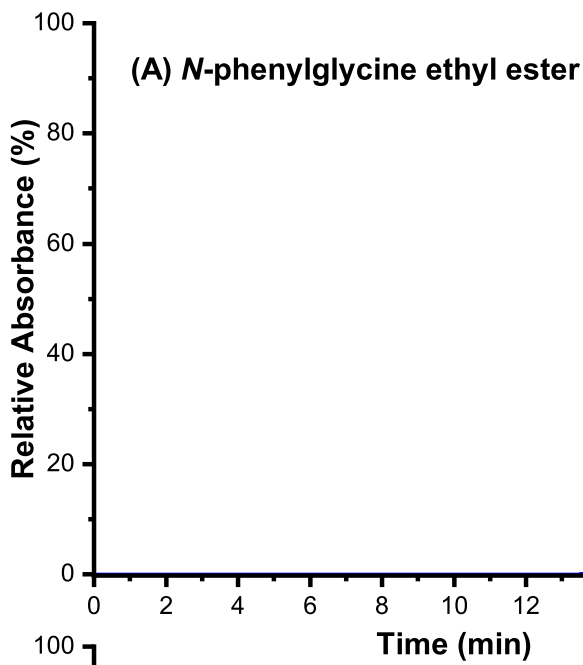
Table 1

	2-aminophenol				3-aminophenol				4-aminophenol			
	t_r (min)	Yield (%)	N/O (%)	TTN	t_r (min)	Yield (%)	N/O (%)	TTN	t_r (min)	Yield (%)	N/O (%)	TTN
AP	15.27	-	-	-	14.77	-	-	-	14.60	-	-	-
EDA	16.25	90.78 ± 2.98	>99.9	908	15.60	61.50 ± 8.52	>99.9	615	15.10	47.35 ± 7.69	>99.9	474
ᵗBuDA	15.70	59.61 ± 6.87	>99.9	596	15.10	>99.9 ± 0.1	>99.9	999	15.42	19.37 ± 6.37	>99.9	194
BnDA	16.01	75.26 ± 3.51	>99.9	753	15.41	44.36 ± 6.26	>99.9	444	15.50	34.34 ± 3.17	>99.9	343