Tandem Friedel-Crafts-Alkylation-Enantioselective-Protonation by Artificial Enzyme Iminium Catalysis

*Reuben B. Leveson-Gower, Ruben M. de Boer and Gerard Roelfes**

Stratingh Institute for Chemistry, University of Groningen,

Nijenborgh 4, 9747 AG, Groningen, The Netherlands

KEYWORDS Artificial Enzymes; Iminium Catalysis; Enantioselective Protonation; Non-canonical Amino Acids; Friedel-Crafts Alkylation.

ABSTRACT The incorporation of organocatalysts into protein scaffolds, i.e. the production of organocatalytic artificial enzymes, holds the promise of overcoming some of the limitations of this powerful catalytic approach. In particular, transformations for which good reactivity or selectivity is challenging for organocatalysts may find particular benefit from translation into a protein scaffold so that its chiral microenvironment can be utilised in catalysis. Previously, we showed that incorporation of the non-canonical amino acid para-aminophenylalanine into the non-enzymatic protein scaffold LmrR forms a proficient and enantioselective artificial enzyme (LmrR_pAF) for the Friedel-Crafts alkylation of indoles with enals. The unnatural aniline side-chain is directly involved in catalysis, operating via a well-known organocatalytic iminium-based mechanism. In this study, we show that LmrR_pAF can enantioselectively form tertiary carbon centres not only during C-C bond formation, but also by enantioselective protonation. Control over

this process is an ongoing challenge for small-molecule catalysts for which general solutions do not exist. LmrR_pAF can selectively deliver a proton to one face of a prochiral enamine intermediate delivering product enantiomeric excesses and yields that rival the best organocatalyst for this transformation. The importance of various side-chains in the pocket of LmrR is distinct from the Friedel-Crafts reaction without enantioselective protonation, and two particularly important residues were probed by exhaustive mutagenesis. This study shows how organocatalytic artificial enzymes can provide solutions to transformations which otherwise require empirical optimisation and design of multifunctional small molecule catalysts.

INTRODUCTION

The quest to broaden the catalytic repertoire of enzymes is born out of a societal need for greener methods of chemical manufacture[1,2]. Enzymes' mild operating conditions and high efficiencies are of great appeal, yet the chemistries that they can catalyse are predominantly limited to those that are important for organismal fitness, and not necessarily those which are useful for humankind[3,4]. Strategies to this end include enhancing promiscuous activities of natural enzymes through directed evolution, computational design of 'de novo' enzymes from scratch, as well as the construction of hybrid catalysts known as artificial enzymes[1,2,5–8]. This last approach involves the situation of catalytic chemical moieties not exploited by natural enzymes into biomolecular scaffolds such as proteins. Of the many methods to combine these two components, the use of non-canonical amino acids (ncAAs) whose side-chains have inherent catalytic properties has recently emerged as an elegant and effective strategy which can reduce the handling steps required for artificial enzyme preparation[9–12]. In this method, amber-stop-codon-suppression is used to site-selectively incorporate the ncAA during protein biosynthesis in response to the amber (TAG)

codon[13,14]. The choice of biomolecular scaffold is paramount to success and in this work, as in our previous studies, we employed the homodimeric Lactococcal multi-drug resistance regulatory protein (LmrR) which has the unusual feature of a large hydrophobic pocket at its dimer interface[15,16]. This protein has proven the perfect catalytic pocket in which to conduct a variety of chemical transformations with rate acceleration and enantio-induction provided by this protein environment[17].

Target transformations for artificial enzymes are typically selected on the basis of their omission amongst Nature's catalytic reactions, as well as their proven synthetic utility. Amino-catalytic chemistry (often simply referred to as organocatalysis) is a highly powerful set of methodologies, many of which were first demonstrated in the past two decades, and whose remarkable contribution to the field of asymmetric synthesis was acknowledged with the 2021 Nobel Prize in Chemistry. It presents many different transformations worthy of translation into n biocatalytic setting with the use of artificial enzymes, some of which are already demonstrated in aqueous environments[18–20]. Of the many activation modes demonstrated in amino-catalysis, the electrophilic activation of enals via the formation of unsaturated iminium ions and their subsequent nucleophilic attack caught our attention due to the diversity of reaction pathways that it allows[21]. Recently we demonstrated that LmrR, with the ncAA para-aminophenyl alanine (pAF) incorporated at position 15, makes a competent and enantioselective catalyst for the Friedel-Crafts alkylation of indoles with aliphatic enal substrates, which are activated for nucleophilic attack at the β-position by iminium ion formation at the catalytic pAF residue (henceforth referred to as FC-reaction, Figure 1(a))[22]. This transformation, which was first demonstrated with organocatalysis by Austin and MacMillan in 2002 (Figure 1(b)), creates the chiral centre during the C-C bond forming step, and thus stereoselective formation of the iminium ion and controlled approach of indole are important

for good enantioselectivity[23,24]. In this study, we turned our attention to the tandem-Friedel-Crafts-alkylation-enantioselective-protonation (henceforth FC-EP reaction) of indoles with α-substituted acroleins (Figure 1(c)). Enantio-induction in this transformation eluded organocatalysis until 2011 when Fu et al. demonstrated good enantioselectivity with a bifunctional amino-catalyst (Figure 1(d))[25]. The greatest challenge for enantioselectivity with α-substituted acrolein substrates is that the chiral centre is formed by protonation of the enamine intermediate formed after the C-C bond formation step (Figure 1(c)). The controlled delivery of a proton to one prochiral face of a substrate is notoriously difficult, yet it is a feat achieved by several natural, artificial, and engineered enzymes[26–33]. Here, we show that LmrR_pAF performs this transformation with good yields and enantioselectivities with a variety of enal and indole substrates. Furthermore, we investigate how pH as well as mutations in the catalytic pocket of LmrR_pAF affect reaction outcomes of both the FC and FC-EP reactions. The steric demands of α-substituted aldehydes/enals make them challenging substrates for conventional secondary-amine-containing organocatalysts and thus this transformation with LmrR_pAF represents an important step forward for artificial enzymes and demonstrates the board catalytic potential of the primary-amine containing catalytic ncAA employed[34,35].

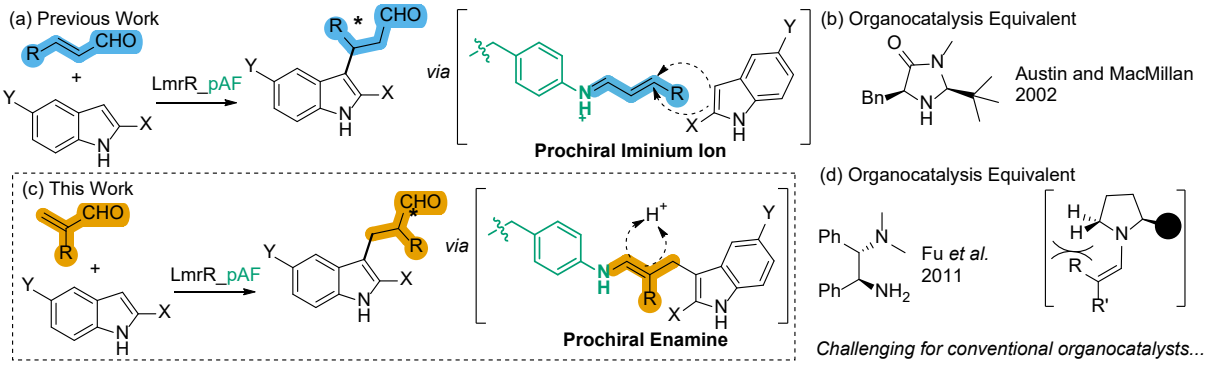


Figure 1. (a) our previous work on the Friedel-Crafts alkylation of indoles with β-substituted enals using LmrR_pAF as catalyst, which takes place *via* a prochiral iminium-ion intermediate[22]. (b) An

organocatalyst for this transformation demonstrated by Austin and MacMillan[23]. (c) This work – tandem-Friedel-crafts-alkylation-enantioselective-protonation of indoles employing α-substituted acroleins as substrates via protonation of a prochiral enamine intermediate. (d) Organocatalysts for this transformation require primary-amine moieties for iminium activation and tertiary-amine moieties for enantioselective proton delivery[25], steric constraints make α-substituted acroleins challenging substrates for conventional secondary-amine-containing organocatalysts[34,35].

RESULTS AND DISCUSSION

Our initial efforts in this study focussed on conducting a series of control experiments to both establish the activity and selectivity of LmrR_pAF for the FC-EP reaction between methacrolein **2a** and 2-methyl-indole **1a** and to rule out the efficacy of LmrR mutants with canonical amino acids at position 15 in place of pAF (Scheme 1, Table 1). Indeed 16 hours reaction time with just 2 mol% of LmrR_pAF afforded essentially quantitative yield of product **3a** with an enantiomeric excess of 74% (as with our previous study, reduction of the aldehyde product to the alcohol **3a** was necessary for normal-phase HPLC analysis of the reaction mixture). Replacing the pAF residue with either lysine, tyrosine or valine (which is present at this position in the wild-type sequence) abrogated activity and selectivity for this transformation. Likewise, incorporation of the pAF aniline sidechain into the protein backbone was also essential, since the combination of aniline and LmrR in ratios typically used for supramolecular catalysis with LmrR[36] failed to produce appreciable yields or enantiomeric excess. LmrR_pAF also significantly outperforms aniline itself for this transformation, which affords only 25% yield of **3a** even at equimolar catalyst loadings.

Scheme 1. The reaction between 2-methyl-indole (**1a**) and methacrolein (**2a**) produces **3a** after reduction via a tandem enantioselective protonation process (FC-EP reaction), whilst substitution of methacrolein with crotonaldehyde (**2b**) produces **4** after reduction (FC reaction).

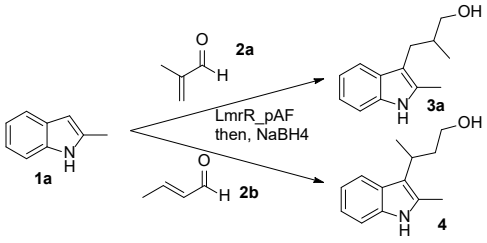


Table 1. Initial results of LmrR_pAF catalysed production of **3a** and control experiments.[a]

Catalyst[b]	Yield **3a** (%)[c]	ee (%)[d]
LmrR_pAF (20 μM)	95 ± 1	74 ± 2
LmrR_V15K (20 μM)	13 ± 2	5 ± 0
LmrR_V15Y (20 μM)	11 ± 1	-6 ± 0
LmrR (20 μM)	10 ± 0	6 ± 0
LmrR (20 μM) + aniline (16 μM)	12 ± 0	-8 ± 0
Aniline (1 mM)[e]	25 ± 1	N.D.

[a]Reaction conditions: [**1a**] = 1 mM; [**2a**] = 6 mM; 300 μL volume reaction in phosphate buffer (50 mM , pH 6.5) containing NaCl (150 mM) and DMF (8 vol %). Reactions conducted for 16 hours, followed by reduction to form **3a** by addition of $NaBH_4$ (60 μL, 20 mg mL^{-1} in 0.5 w/v% NaOH) for analysis. Errors given represent the standard deviation from two experiments with independently produced batches of protein, each conducted in duplicate. [b]Concentrations of LmrR dimer. [c]Analytical yields of **3a** determined by chiral normal-phase HPLC with the use of a calibration curve. [d]Enantiomeric excess determined by chiral normal-phase HPLC. [e]Error given is standard deviation from an experiment conducted in triplicate. N.D. = not determined.

Next, we investigated the effect of pH on the catalytic production of **3a** and **4**, to assess whether the abundance of solvent protons is important for catalysis in the enantioselective protonation process required for the production **3a**. We also chose a shorter reaction time and lower enzyme loading for these experiments in order to better observe any effects present. At all pH measured,

the yield of **3a** and **4** is very similar suggesting that positioning the methyl group in either the α- or β-positions of the substrate has little effect on the activity. Both reactions showed a pronounced effect from pH in the region of 6 to 7 (Figure 2(a)). Changing the pH from 6 to 6.5 and to 7 results in a significant loss in product yield, which is likely indicative of the pK_a of the iminium ion, whose protonation is crucial for effective catalysis hence the widespread use of acid-cocatalysts in iminium catalysis[21]. Much higher enantiomeric excesses are obtained for product **3a** than for **4**, which gives the somewhat surprising conclusion that LmrR_pAF can better control the enantioselective delivery of a proton than of the indole substrate. Loss in enantiomeric excess with increasing pH was more pronounced in the case of **3a** than **4**, however correcting for the background reaction (which is higher than for **4**, Supporting Table 2), finds that the enantioselectivity of the catalysed reaction is unaffected. We were interested to find that much higher enantiomeric excesses were obtained for **3a** (up to 88% at pH 6) than under the first conditions we tested (with longer reaction times and higher catalyst loadings). This lead us to suspect that racemisation may be affecting the ultimate enantio-enrichment in the product, as is well known to occur in water with compounds with stereo-centres in the α- position to a carbonyl functionality. We monitored the production of **3a** over 48 hours and found that whilst full conversion occurs after approximately 12 hours, the enantiomeric excess erodes steadily over the whole period (Figure 2(b)). The enantiomeric excess decreased in a near perfect linear manner, allowing us to determine a rate of decrease of approximately 0.5% per hour. We previously demonstrated that both the buffer and protein scaffold produced racemisation in another reaction involving enantioselective protonation with LmrR_pAF, and thus a variety of processes are presumably also involved here too. Consequently, shorter reaction times are desirable for obtaining the highest enantiomeric excesses[31]. We explored both lowering the enzyme loading, as well as

increasing the substrate concentration obtaining TONs over 200 in some cases. However in each case there was a concomitant drop in enantiomeric excess (Supporting Table 1). Nevertheless, the TON, yield and enantiomeric excess obtained (i.e. Figure 2(a) at pH 6) are all higher than the best performing organocatalyst reported for this transformation[25].

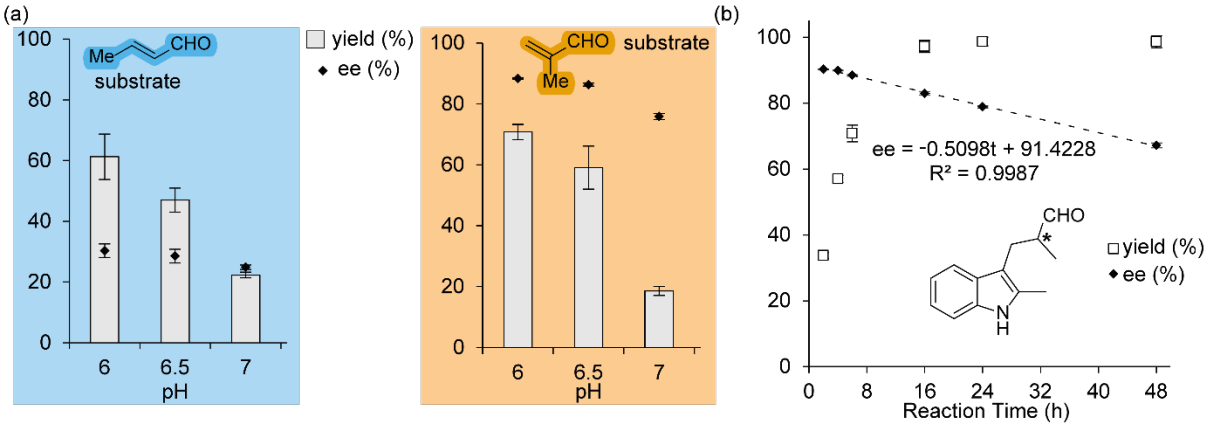


Figure 2. (a) Effect of reaction pH on analytical yields and enantiomeric excesses from the formation of 3a (left, blue) and 4 (right, orange) by LmrR_pAF. Reaction conditions [LmrR_pAF] = 10 μM (dimer concentration); [**1a**] = 1 mM; [**2a**] = 6 mM or [**2b**] = 5 mM; 300 μL volume reaction in phosphate buffer (50 mM) containing NaCl (150 mM) and DMF (8 vol %). Reactions conducted for 6 hours at 4 °C, followed by reduction to form 3a or 4 for analysis by normal-phase HPLC. (b) LmrR_pAF catalysed production of **3a** monitored over 48 hours, revealing product racemisation. Reaction conditions as in (a), pH = 6. In both (a) and (b), errors given represent the standard deviation from two experiments with independently produced batches of protein, each conducted in duplicate.

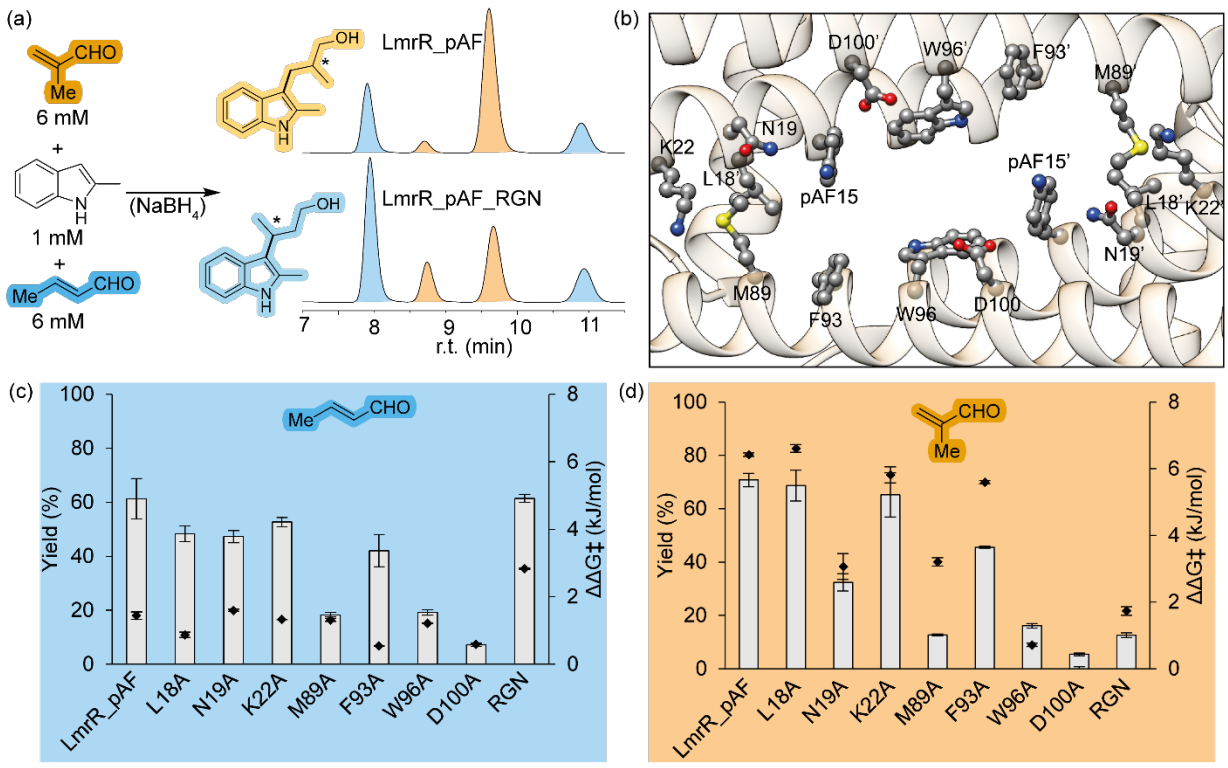


Figure 3. (a) representative chiral normal-phases HPLC traces obtained in competition experiments employing both substrates **2a** and **2b** together with indole **1a** to produce mixtures of products **3a** (orange) and **4** (blue) (Table 2). (b) Positions in LmrR_pAF subject to mutagenesis (PDB: 6I8N). Effect of various mutants on reaction outcomes producing product **4** (c) in blue and **3a** (d) in orange. $\Delta\Delta G^{\ddagger}$ (the difference in the Gibbs' free energy of activation for the production of the two product enantiomers) was calculated from the enantiomeric ratio (e.r.) according the equation $\Delta\Delta G^{\ddagger} = RT\ln(e.r.)$. In (c) and (d), errors given represent the standard deviation from two experiments with independently produced batches of protein, each conducted in duplicate.

In competition experiments employing equal concentrations of crotonaldehyde (**2b**) and methacrolein (**2a**) with 2-methylindole (**1a**) LmrR_pAF shows almost no preference for production of either **3a** or **4**, albeit with the far higher enantioselectivity for **3a** already noted

(Figure 3(a), Table 2). However, when we tested LmrR_pAF_RGN, which was previously subject to directed evolution for the Friedel-Crafts reaction with the linear trans-2-hexenal as screening substrate, we found that this triple mutant has a twofold preference for production of **4** over **3a**. **4** is produced by LmrR_pAF_RGN with a higher enantiomeric excess than by LmrR_pAF, with an overall lower conversion in the same time frame and large loss in enantiomeric excess for **3a**.

Table 2. Reaction outcomes of competition experiments with equal concentrations of **2a** and **2b** with **1a**, catalysed by LmrR_pAF(mutants). [a]

Catalyst	Yield **4a** (%)[b]	*ee* **4a** (%)[c]	Yield **3a** (%)[b]	*ee* **3a** (%)[c]	**4a**:**3a**
LmrR_pAF	46 ± 1	22 ± 1	55 ± 1	86 ± 0	1:1.2
LmrR_pAF_RGN	38 ± 2	50 ± 0	20 ± 1	37 ± 1	2:1

[a]Reaction conditions: [LmrR_pAF] or [LmrR_pAF_RGN] = 10 μM (dimer concentration); [**1a**] = 1 mM; [**2a**] = [**2b**] = 6 mM; 300 μL volume reaction in phosphate buffer (50 mM, pH 6) containing NaCl (150 mM) and DMF (8 vol %). Reactions conducted for 6 hours at 4 °C, followed by reduction to form **3a** or **4** for analysis by chiral normal-phase HPLC. [b]Yields determined by normal-phase HPLC with use of calibration curves. [c]Enantiomeric excesses determined by chiral normal-phase HPLC. Errors given represent the standard deviation from two experiments with independently produced batches of protein, each conducted in duplicate.

Noting this large divergence in substrate preference and enantioselectivity engendered by the mutations in LmrR_pAF_RGN we hypothesised that the two reaction pathways to produce either **3a** or **4** utilise the pocket of LmrR to promote catalysis in different manners. To test this, we performed alanine-scanning at 7 positions inside the pocket of LmrR encompassing both polar and apolar residues (Figure 3(b)-(d)). Since the enantiomeric excesses produced in each reaction are so different, and to present mutational effects on this parameter on a linear scale we show $\Delta\Delta G^{\ddagger}$ i.e. the level of energy discrimination that a particular mutant provides between the transition states leading to either enantiomer of **3a** or **4**. Alanine mutations at positions L18, K22 and F93 produced

only minor effects on catalysis outcomes for the production of **3a** and **4**. At positions W96 and D100, large detrimental effects were observed for both yield and enantioselectivity of **3a** and **4** (although LmrR_pAF_W96A has similar enantioselectivity for production of **4**), in line with previous results highlighting the importance of these residues in the majority of LmrR-based artificial enzymes[30,36–38]. Substitution of alanine at positions N19 and M89, however, showed distinctly different effects on the outcomes of the FC and FC-EP reactions. Whilst for the FC-EP reaction both of these mutants show severely reduced yields and enantioselectivities, the only significant effect on the FC reaction is a reduced yield in the case of LmrR_pAF_M89A.

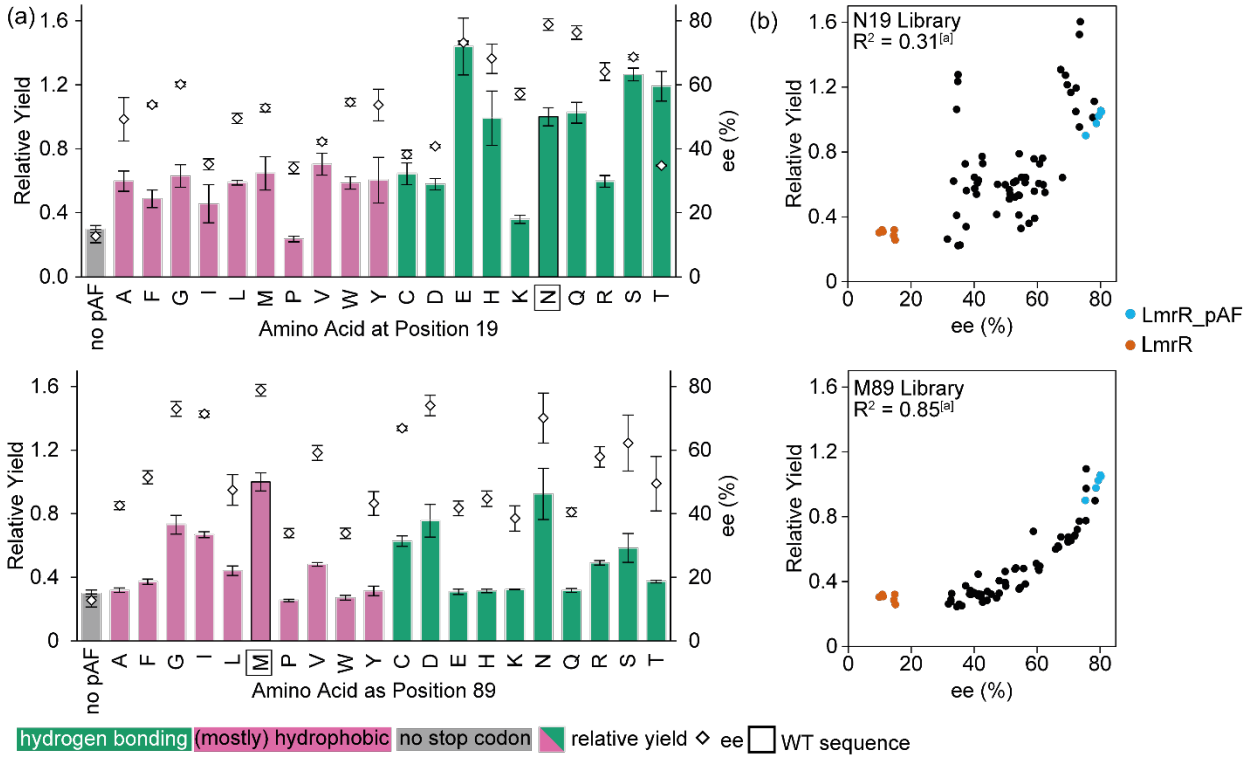


Figure 4. (a) Catalysis results for the FC-EP reaction producing **3a** using cell-free extract libraries with every mutant at the N19 (top) and M89 (bottom) positions. Results are an average of a triplicate experiment, and error bars shown reflect the standard deviation of those experiments, except for the results for LmrR (no pAF) which is six repeated experiments (data from both

libraries combined) and LmrR_pAF which is five repeated experiments (data from both libraries combined, one sample was calculated to be an outlier by the interquartile method). Analytical yields and enantiomeric excesses were determined by SFC with the use of an internal standard, and are given relative to the mean of the LmrR_pAF samples. (b) Results for each library: N19 shows a weak correlation between yield and ee for the pAF containing samples, whilst M89 shows a strong correlation. LmrR (no pAF) samples shown in orange and LmrR_pAF samples shown in blue. [a] Value obtained by performing a linear fit of the pAF containing members of the library, i.e. LmrR without pAF was not included in the fit.

In order to learn more about the roles the side-chains at N19 and M89 play in the FC-EP reaction, we prepared every mutant at each of these positions by QuikChange® PCR and expressed them together with LmrR with the wild-type valine at position 15 instead of pAF, and LmrR_pAF as negative and positive controls, respectively. We then lysed the cells and used the cell-free extract directly for catalysis, analysing the enantioselectivity and yield of reactions rapidly with super-critical fluid chromatography (SFC). We also analysed the soluble fraction of the cultures by SDS-PAGE to qualitatively inspect the relative expression levels of the 38 mutants (see Supporting Figure 1). Only the proline mutants showed particularly poor soluble expression, which is unsurprising given that both positions 19 and 89 are situated in α-helices. In general, mutations at position 19 have significant effects on soluble protein production, whilst the library of mutants at position 89 showed relatively uniform expression.

At position N19, the most functional replacements for this asparagine side-chain are other hydrogen-bonding side-chains (although notably the cysteine and aspartic acid mutants do not perform well) (Figure 4(a), top). The fact that mutants at position 19 show a poor correlation in yield and enantiomeric excess suggests that sidechains in this position can affect the rate-

determining or chiral-centre-forming steps via separate mechanisms. This is well illustrated by LmrR_pAF_N19T, which affords **3a** with increased yield, but much lower enantioselectivity, than LmrR_pAF. Taken together, the results from the N19 library suggest that side-chains at position 19 may be directly involved in the enantioselective protonation step by either shuttling protons or else stabilising ordered water molecules in the active site. Conversely, at position 89, functional replacements for the methionine side-chain can be found with a variety of sizes and polarities, suggesting a less direct role for side-chains at this position (Figure 4(a), bottom). Unlike the N19 library, this library shows a strong correlation between the yield and enantioselectivity obtained with a given mutant. Any effect that a side-chain at position 89 has on the yield of **4a** is also reflected proportionally in the enantiomeric excess, implying that sidechains at this position affect both the rate-determining and enantioselective protonation steps via a single mechanism, perhaps promoting active-site preorganisation for catalysis (Figure 4(b), bottom).

Finally, we assessed the scope of the FC-EP reaction by LmrR_pAF with regards to both enal α-substituents as well as indole substituents in the 2- and 5-positions (Figure 5). Methacrolein could be substituted with 2-ethyl-acrolein to produce **3b** whilst maintaining good yields and enantioselectivity. Even the bulky 2-benzyl-acrolein could be employed successfully as substrate, requiring slightly increased enzyme loadings and reactions times to give good yields and enantioselectivities. The doubly substituted tiglic aldehyde could also be employed to afford product **3d**, however higher catalyst loadings and longer reaction times were required to afford modest yields and enantio- and diastereoselectivities. When LmrR_pAF_RGN was used to catalyse the conversion of this substrate, a similar diastereomeric ratio but different enantiomeric excesses were obtained, which may reflect the propensity of this mutant for enantioselective C-C bond formation, rather than enantioselective protonation. Similarly to our previous work, the 2-

methyl substituent has a significant effect on the yields obtained, and thus products **3e-h** required increased catalyst loadings and reaction times to accumulate good yields[22]. Erosion of enantiomeric excess was also observed for these products, with shorter reaction times giving higher selectivities but lower yields, and longer reaction times improving yields at the expense of selectivity. Electron donating substituents on the indole ring proved beneficial for activity (products **3f** and **3g**); the electron withdrawing 5-chloro substituent in product **3h** did not negatively affect the yield obtained, but did give more rapid erosion of enantiomeric excess. The highly electron-withdrawing methyl-ester substituent in **3i**, however, reduced activity to the point that no product was detected. These substituent effects are consistent with the nucleophilic role of indole in the reaction pathway. Finally, product **3j**, with both 2-methyl and 5-methoxy substituents, was obtained with good yields and enantiomeric excess whilst still using a low catalyst loading and short reaction time.

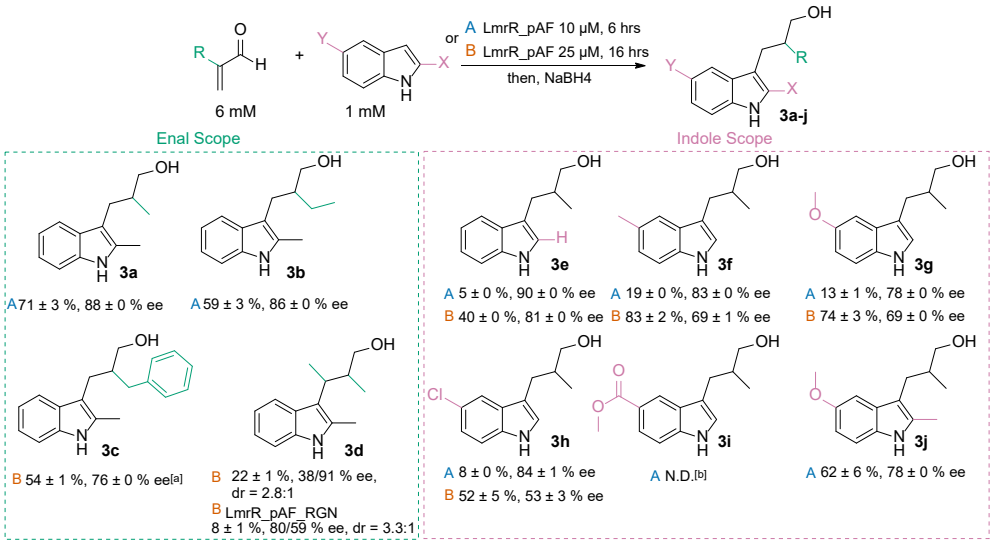


Figure 5. Substrate scope of FC-EP reaction catalysed by LmrR_pAF. Analytical yields were determined by HPLC or SFC with the use of a calibration curve. [a]Enal concentration of 1.5 mM was used due to low solubility of this substrate under reaction conditions. [b]No product was

detected by HPLC. Errors given represent the standard deviation from two experiments with independently produced batches of protein, each conducted in duplicate. N.D. = not determined.

CONCLUSION

In this work we have shown how a challenging enantioselective protonation process can be achieved in a protein scaffold by using an organocatalytic mechanism mediated by a non-canonical amino acid side-chain. This reactivity responds to mutations in a markedly different manner than the Friedel-Crafts reaction which does not involve enantioselective protonation. This suggests that the LmrR scaffold acts as a 'blank canvas' where the amino-acid sidechains in the pocket can be utilised in different manners to promote the different catalytic reactions which can be conducted there. In particular, the N19 and M89 positions play crucial roles in the activity shown herein, and the patterns in reactivity and selectivity of mutants at these positions suggest that side-chains here play roles in both the C-C bonding forming, and enantioselective protonation steps. The FC-EP reaction is another promiscuous activity of LmrR_pAF building on the hydrazone formation, Friedel-Crafts and synergistically catalysed processes that we have already demonstrated[11,22,31,39]. We anticipate that LmrR_pAF will find application in yet further useful and challenging reactions operating *via* organocatalytic processes, realizing the benefits of biocatalyatic processes for a broad array of transformations[1,2].

ASSOCIATED CONTENT

Supporting Information. Supporting figures and tables. Full description of methods employed and characterization of reference compounds. List of primers used. HPLC and SFC chromatograms and NMR and Mass Spectra.

AUTHOR INFORMATION

Corresponding Author

* E-mail: j.g.roelfes@rug.nl

Funding Sources

This work was supported by the Netherlands Ministry of Education, Culture and Science (Gravitation program no. 024.001.035) and the European Research Council (ERC advanced grant 885396).

ACKNOWLEDGMENT

Our thanks to Dr. H van Beek of GECCO Biotech B.V. and Dr. F Aalbers for advice on library preparation and screening, to Dr. Z Zhou and C da Settimo Passetti for preparing 2-benzyl-acrolein[40], and to R Sneep for assistance with SFC analysis.

REFERENCES

(1) Chen, K.; Arnold, F. H. Engineering New Catalytic Activities in Enzymes. *Nat. Catal.* **2020**, *3*, 203–213.
(2) Leveson-Gower, R. B.; Mayer, C.; Roelfes, G. The Importance of Catalytic Promiscuity for Enzyme Design and Evolution. *Nat. Rev. Chem.* **2019**, *3*, 687–705.
(3) Sheldon, R. A.; Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **2018**, *118*, 801–838.
(4) Sheldon, R. A.; Brady, D.; Bode, M. L. The Hitchhiker's Guide to Biocatalysis: Recent Advances in the Use of Enzymes in Organic Synthesis. *Chem. Sci.* **2020**, *11*, 2587–2605.
(5) Kiss, G.; Çelebi-Ölçüm, N.; Moretti, R.; Baker, D.; Houk, K. N. Computational Enzyme Design. *Angew. Chem. Int. Ed.* **2013**, *52*, 5700–5725.
(6) Zeymer, C.; Hilvert, D. Directed Evolution of Protein Catalysts. *Annu. Rev. Biochem.* **2018**, *87*, 131–157.
(7) Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Köhler, V.; Lewis, J. C.; Ward, T. R. Artificial Metalloenzymes: Reaction Scope and Optimization Strategies. *Chem. Rev.* **2018**, *118*, 142–231.
(8) Rosati, F.; Roelfes, G. Artificial Metalloenzymes. *ChemCatChem* **2010**, *2*, 916–927.
(9) Zhao, J.; Burke, A. J.; Green, A. P. Enzymes with Noncanonical Amino Acids. *Curr. Opin. Chem. Biol.* **2020**, *55*, 136–144.
(10) Drienovská, I.; Roelfes, G. Expanding the Enzyme Universe with Genetically Encoded Unnatural Amino Acids. *Nat. Catal.* **2020**, *3*, 193–202.
(11) Drienovská, I.; Mayer, C.; Dulson, C.; Roelfes, G. A Designer Enzyme for Hydrazone and Oxime Formation Featuring an Unnatural Catalytic Aniline Residue. *Nat. Chem.* **2018**, *10*, 946–952.

(12) Burke, A. J.; Lovelock, S. L.; Frese, A.; Crawshaw, R.; Ortmayer, M.; Dunstan, M.; Levy, C.; Green, A. P. Design and Evolution of an Enzyme with a Non-Canonical Organocatalytic Mechanism. *Nature* **2019**, *570*, 219–223.
(13) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the Genetic Code of Escherichia Coli. *Science* **2001**, *292*, 498–500.
(14) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. Designing Logical Codon Reassignment – Expanding the Chemistry in Biology. *Chem. Sci.* **2014**, *6*, 50–69.
(15) Madoori, P. K.; Agustiandari, H.; Driessen, A. J. M.; Thunnissen, A.-M. W. H. Structure of the Transcriptional Regulator LmrR and Its Mechanism of Multidrug Recognition. *EMBO J.* **2009**, *28*, 156–166.
(16) Jeong, W. J.; Yu, J.; Song, W. J. Proteins as Diverse, Efficient, and Evolvable Scaffolds for Artificial Metalloenzymes. *Chem. Commun.* **2020**, *56*, 9586–9599.
(17) Roelfes, G. LmrR: A Privileged Scaffold for Artificial Metalloenzymes. *Acc. Chem. Res.* **2019**, *52*, 545–556.
(18) MacMillan, D. W. C. The Advent and Development of Organocatalysis. *Nature* **2008**, *455*, 304–308.
(19) Melchiorre, P.; Marigo, M.; Carlone, A.; Bartoli, G. Asymmetric Aminocatalysis-Gold Rush in Organic Chemistry. *Angew. Chem. Int. Ed.* **2008**, *47*, 6138–6171.
(20) van der Helm, M. P.; Klemm, B.; Eelkema, R. Organocatalysis in Aqueous Media. *Nat. Rev. Chem.* **2019**, *3*, 491–508.
(21) Erkkilä, A.; Majander, I.; Pihko, P. M. Iminium Catalysis. *Chem. Rev.* **2007**, *107*, 5416–5470.
(22) Leveson-Gower, R. B.; Zhou, Z.; Drienovská, I.; Roelfes, G. Unlocking Iminium Catalysis in Artificial Enzymes to Create a Friedel–Crafts Alkylase. *ACS Catal.* **2021**, *11*, 6763–6770.
(23) Austin, J. F.; MacMillan, D. W. C. Enantioselective Organocatalytic Indole Alkylations. Design of a New and Highly Effective Chiral Amine for Iminium Catalysis. *J. Am. Chem. Soc.* **2002**, *124*, 1172–1173.
(24) Gordillo, R.; Carter, J.; Houk, K. N. Theoretical Explorations of Enantioselective Alkylation Reactions of Pyrroles and Indoles Organocatalyzed by Chiral Imidazolidinones. *Adv. Synth. Catal.* **2004**, *346*, 1175–1185.
(25) Fu, N.; Zhang, L.; Li, J.; Luo, S.; Cheng, J. P. Chiral Primary Amine Catalyzed Enantioselective Protonation via an Enamine Intermediate. *Angew. Chem. Int. Ed.* **2011**, *50*, 11451–11455.
(26) Mohr, J. T.; Hong, A. Y.; Stoltz, B. M. Enantioselective Protonation. *Nat Chem* **2009**, *1*, 359–369.
(27) Matsumoto, K.; Tsutsumi, S.; Ihori, T.; Ohta, H. Enzyme-Mediated Enantioface-Differentiating Hydrolysis of .Alpha.-Substituted Cycloalkanone Enol Esters. *J. Am. Chem. Soc.* **1990**, *112*, 9614–9619.
(28) Matoishi, K.; Ueda, M.; Miyamoto, K.; Ohta, H. Mechanism of Asymmetric Decarboxylation of α-Aryl-α-Methylmalonate Catalyzed by Arylmalonate Decarboxylase Originated from Alcaligenes Bronchisepticus. *J. Mol. Catal. B* **2004**, *27*, 161–168.
(29) Miyamoto, K.; Hirokawa, S.; Ohta, H. Conversion of α-Methyltropate to Optically Active α-Phenylpropionate by Tropate-Degrading Rhodococcus Sp. KU1314. *J. Mol. Catal. B* **2007**, *46*, 14–19.
(30) Villarino, L.; Chordia, S.; Alonso-Cotchico, L.; Reddem, E.; Zhou, Z.; Thunnissen, A. M. W. H.; Maréchal, J.-D.; Roelfes, G. Cofactor Binding Dynamics Influence the Catalytic

Activity and Selectivity of an Artificial Metalloenzyme. *ACS Catal.* **2020**, *10*, 11783–11790.
(31) Zhou, Z.; Roelfes, G. Synergistic Catalysis of Tandem Michael Addition/Enantioselective Protonation Reactions by an Artificial Enzyme. *ACS Catal.* **2021**, *11*, 9366–9369.
(32) Steck, V.; Carminati, D. M.; Johnson, N. R.; Fasan, R. Enantioselective Synthesis of Chiral Amines via Biocatalytic Carbene N–H Insertion. *ACS Catal.* **2020**, *10*, 10967–10977.
(33) Liu, Z.; Calvó-Tusell, C.; Zhou, A. Z.; Chen, K.; Garcia-Borràs, M.; Arnold, F. H. Dual-Function Enzyme Catalysis for Enantioselective Carbon–Nitrogen Bond Formation. *Nat. Chem.* **2021**.
(34) Melchiorre, P. Cinchona-Based Primary Amine Catalysis in the Asymmetric Functionalization of Carbonyl Compounds. *Angew. Chem. Int. Ed.* **2012**, *51*, 9748–9770.
(35) Xu, L.-W.; Luo, J.; Lu, Y. Asymmetric Catalysis with Chiral Primary Amine-Based Organocatalysts. *Chem. Commun.* **2009**, 1807.
(36) Bos, J.; Browne, W. R.; Driessen, A. J. M.; Roelfes, G. Supramolecular Assembly of Artificial Metalloenzymes Based on the Dimeric Protein LmrR as Promiscuous Scaffold. *J. Am. Chem. Soc.* **2015**, *137*, 9796–9799.
(37) Villarino, L.; Splan, K. E.; Reddem, E.; Alonso-Cotchico, L.; Lledós, A.; Gutiérrez de Souza, C.; Thunnissen, A.-M. W. H.; Maréchal, J.-D.; Roelfes, G. An Artificial Heme Enzyme for Cyclopropanation Reactions. *Angew. Chem. Int. Ed.* **2018**, *57*, 7785–7789.
(38) Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. Enantioselective Artificial Metalloenzymes by Creation of a Novel Active Site at the Protein Dimer Interface. *Angew. Chem. Int. Ed.* **2012**, *51*, 7472–7475.
(39) Zhou, Z.; Roelfes, G. Synergistic Catalysis in an Artificial Enzyme by Simultaneous Action of Two Abiological Catalytic Sites. *Nat. Catal.* **2020**, *3*, 289–294.
(40) Erkkilä, A.; Pihko, P. M. Mild Organocatalytic α-Methylenation of Aldehydes. *J. Org. Chem.* **2006**, *71*, 2538–2541.

Supporting Information

Tandem Friedel-Crafts-Alkylation-Enantioselective-Protonation by Artificial Enzyme Iminium Catalysis

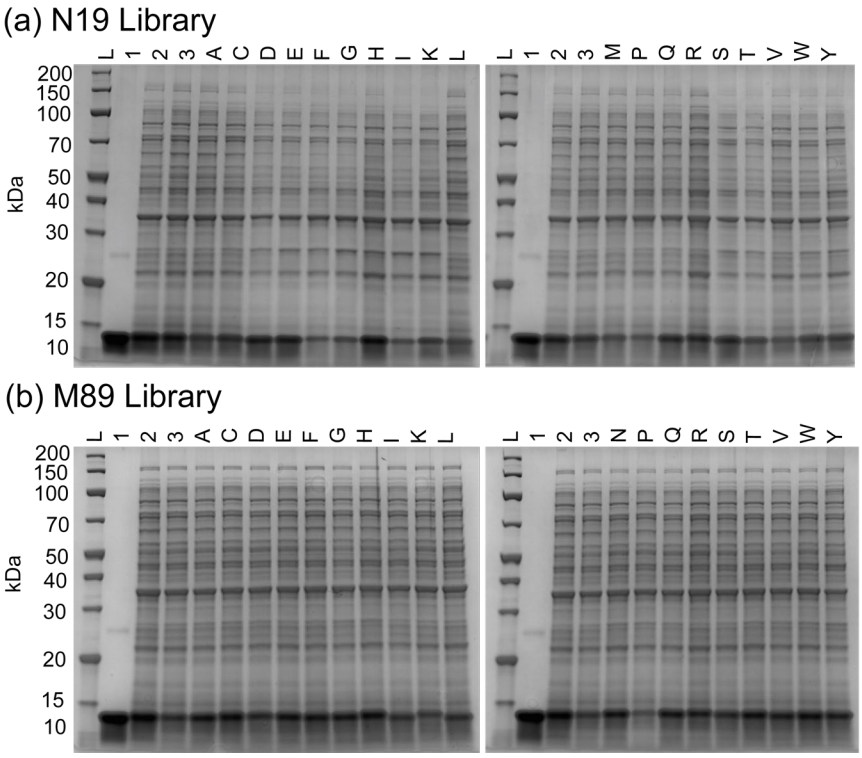
Reuben B. Leveson-Gower, Ruben M. de Boer, Gerard Roelfes*

Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands

*j.g.roelfes@rug.nl

Table of Contents

1. Supporting Figures



Supporting Figure 1: SDS-PAGE (ExpressPlus™ 12%, GenScript) analysis of soluble fraction of libraries expressing (a) LmrR_pAF_N19X mutants and (b) LmrR_pAF_M89X mutants. L = Thermo-Fisher broad-range unstained protein ladder; 1 = LmrR_pAF (purified) 5 μM; 2 = LmrR; 3 = LmrR_pAF.

2. Supporting Tables

Supporting Table 1. Optimisation of enzyme loading and substrates concentration for LmrR_pAF catalysed production of **3a**.[a]

[LmrR_pAF] μM[b]	[1a] mM	[2a] mM	Time (h)	Yield (%)[c]	ee (%)[d]	TON[e]
10	6	1	6	71 ± 3	88 ± 0	71 ± 3
5	6	1	16	69 ± 4	84 ± 0	138 ± 8
2	6	1	40	70 ± 14	76 ± 3	275 ± 70[f]
1	6	1	40	28 ± 9	40 ± 21	130 ± 90[f]
0	6	1	40	15 ± 0	-	-
10	1	6	16	97 ± 2	82 ± 2	97 ± 2
10	2	12	16	87 ± 8	78 ± 3	171 ± 15
10	3.5	21	16	61 ± 4	63 ± 2	214 ± 15
10	5	30	16	61 ± 11	54 ± 12	304 ± 57

[a]Reaction conditions: 300 μL volume reaction in phosphate buffer (50 mM , pH 6.5) containing NaCl (150 mM) and DMF (8 vol %). Reactions conducted at 4 °C for specified time. [b]LmrR_pAF concentration refers to that of the dimer. [c]Analytical yield calculated using normal-phase HPLC with a calibration curve. [d]Enantiomeric excess determined with chiral normal-phase HPLC. [e]TON = [**3a**]/[LmrR_pAF]. [f]40 hours background reaction subtracted from these entries. Experiments were conducted with two batches of protein, and in duplicate, of the four values obtained the results presented are the mean, whilst the errors given are the standard deviation.

Supporting Table 2. Background reactions for FC and FC-EP reactions[a]

pH	Yield **3a** (%)	Yield **4** (%)
6	3 ± 0	<1
6.5	3 ± 0	<1
7	3 ± 0	<1

[a]Reaction conditions: 300 μL volume reaction in phosphate buffer (50 mM , pH 6.5) containing NaCl (150 mM) and DMF (8 vol %) [2-methyl-indole] = 1 mM and [methacrolein] = 6 mM or [crotonaldehyde] = 5 mM. Reactions conducted at 4 °C for 6 hours, followed by reduction with $NaBH_4$. Analytical yield calculated using normal-phase HPLC with a calibration curve. Experiments were conducted in triplicate, the results are the mean of the values obtained and the errors are the standard deviation.

3. Materials and Equipment

Chemicals were purchased from commercial suppliers (Sigma (UK), Acros (Germany), TCI (Belgium/Japan) and Fluorochem (UK)) and used without further purification unless specified. Flash column chromatography was performed on silica gel (Silica-P flash silica gel from Silicycle, 0.040-0.063 mm, 230-400 mesh). The unnatural amino acid pAzF was purchased as racemic mixture from Bachem (Switzerland) or as the enantiopure hydrochloride salt from Iris-Biotech (Germany). NMR ^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Varian 400 (400 MHz) spectrometer in $CDCl_3$ or $(CD_3)_2SO$. Chemical shifts values (δ) are denoted in ppm using residual solvent peaks as the internal standard ($CHCl_3$: δ 7.26 for ^{1}H; 77.16 for ^{13}C. $(CD_3)_2SO$: δ 2.50 for ^{1}H; 39.52 for ^{13}C). HPLC analysis was conducted using a Shimadzu LC-10ADVP HPLC equipped with a Shimadzu SPD-M10AVP diode array detector. Plasmid pEVOL-pAzF was obtained from Addgene (pEvol-pAzF was a gift from Prof. Peter Schultz (The Scripps Research Institute))[1]. Plasmid pEVOL_pAzFRS.2.t1 was obtained from Addgene (pEVOL_pAzFRS.2.t1 was a gift from Prof. Farren Isaacs, Yale University)[2]. *E. coli* strains, NEB5-alpha, NEB10-beta and BL21(DE3) (New England Biolabs) were used for cloning and expression. Primers were synthesized by Eurofins MWG Operon (Germany) and Sigma-Aldrich (UK). Plasmid Purification Kits were obtained from QIAGEN (Germany) and DNA sequencing carried out by GATC-Biotech (Germany). Phusion polymerase and DpnI were purchased from New England Biolabs. Strep-tactin columns (Strep-Tactin® Superflow® high capacity) and Desthiobiotin were purchased from IBA-Lifesciences (Germany). Concentrations of DNA and protein solutions were determined based on the absorption at 260 nm or 280 nm on a Thermo Scientific Nanodrop 2000 UV-Vis spectrophotometer. Molar extinction coefficients were approximated using the ProtParam Expasy web server https://web.expasy.org/protparam/. UPLC/MS analysis was performed on Waters Acquity Ultra Performance LC with Acquity TQD detector. Water (solvent A) and acetonitrile (solvent B) containing 0.1% formic acid by volume, were used as the mobile phase at a flow rate of 0.3 mL/min. Gradient: 90% A for 2 min, linear gradient to 50% A in 2 min, linear gradient to 20% A in 5 min, followed by 2 min at 5% A. Re-equilibration of the column with 2 min at 90% A. SFC analysis was performed using a Water Acquity UPC^2 system. High-Resolution Mass Spectrometry (HRMS) measurements were performed using a Thermo LTQ Orbitrap XL. Low-Resolution Mass Spectrometry (LRMS) measurements were performed using a Waters Acquity H-class UPLC with Waters Xevo G2 QTOF with the column set to bypass (direct-injection).

Methods

3.1 Protein Production and Purification

LmrR_pAF variants were produced and purified as previously described[3]. The identity and purity of proteins and the successful reduction of pAzF were determined by mass spectrometry. Protein concentration was determined by correcting the calculated extinction coefficients for LmrR variants for the absorbance of pAF (ε_{280} = 1333 M^{-1} cm^{-1})

3.2 Construction of Mutants

LmrR_V15TAG_L18A, LmrR_V15TAG_N19A, LmrR_V15TAG_K22A, LmrR_V15TAG_M89A, LmrR_V15TAG_F93A, LmrR_V15TAG_W96A, LmrR_V15TAG_D100A and LmrR_V15TAG_RGN were prepared and characterised previously[4,5]. The 18 remaining possible mutants at positions N19 and M89 were constructed by site-directed mutagenesis (QuikChange, Agilent Technologies). Primers are described in supporting information section 5. 25 µL reactions were set-up using Phusion polymerase (New England Biolabs) according to the manufacturer protocol with pET17b+_LmrR_V15TAG as template. The following PCR protocol was used: (1) initial denaturation at 98 °C for 1 min, (2) 20 cycles of denaturation at 98 °C for 30 s, annealing at 56-68 °C for 30 s, and extension at 72 °C for 2 min, (3) a final extension at 72 °C for 10 min. The crude reaction mixture was transformed into chemically competent *E. coli* cells of either NEB5α or NEB10β strains and spread onto LB agar plates containing ampicillin (100 µg/mL). Single colonies were sent for sequencing (GATC Biotech.). The isolated plasmid was then co-transformed with pEVOL_pAzFRS.2.t1 or pEVOL_pAzF into chemically-competent *E. coli* BL21(DE3) cells or transformed into chemically-competent *E. coli* BL21(DE3) cells containing pEVOL_pAzF_RS2.t1 or pEVOL_pAzFRS.2.t1 which were spread onto LB agar plates containing ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL). Single colonies from these plates were grown overnight in 5 mL LB containing ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) and these culture were used to prepare glycerol stocks which were stored at -70 ºC until further use. pEVOL_pAzF or pEVOL_pAzFRS.2.t1 were used in the preparation of purified LmrR_pAF mutants, whilst only pEVOL_pAzFRS.2.t1 was used for the preparation of cell-free extracts in deep-well format, owing to the higher activity of cell-free extracts prepared using this OTS plasmid. These glycerol stocks were used for protein production.

3.3 Preparation of Cell-Free Extracts in Deep Well Format

Conducted in a similar manner to our previous works[4,6]. Glycerol stocks of the relevant mutants (along with LmrR_WT and LmrR_pAF as controls) were used to inoculate 1.5 mL deep well plates containing 500 µL LB media and appropriate antibiotics in triplicate. The resulting deep well plates were incubated overnight at 37 °C while shaking at 950 rpm (Titramax 1000 & Incubator 1000, Heidolph). The next morning, 50 µL of the densely grown overnight cultures were transformed into fresh 96-deep well plates containing 1150 µL LB media and appropriate antibiotics. Glycerol (500 µL, 50 % with miliQ water) was added to the remaining overnight culture, mixed thoroughly and stored at - 70 °C. Bacteria were cultured at 37 °C for 5 - 6 hours while shaking at 950 rpm. Subsequently, protein production was induced by addition of 50 µL LB media, containing IPTG (1.2 µL of a 1 M stock solution), arabinose (1.2 µL of a 20% arabinose stock solution) and p-azidophenylalanine at a concentration of 30 mM (final concentrations: IPTG = 1 mM, arabinose = 0.02%, pAzF = 1.2 mM). To avoid precipitation of the unnatural amino acid, pAzF was dissolved by addition of base (1 M NaOH) prior to addition to the LB media. Plates were then incubated at 30 °C for 16 hours while shaking (950 rpm) and then 50 µL was removed from each triplicate and combined to give 150 µL for each distinct mutant. This was used for SDS-PAGE analysis using the BugBuster® (primary amine-free) Extraction Reagent (Millipore). The remaining culture was harvested by centrifugation (3,500 rpm at 4 °C for 15 minutes). After removing the supernatant, cells were washed by addition of 500 µL of buffer A (50 mM sodium phosphate, 150 mM NaCl, pH = 6), and the supernatant was again discarded after centrifugations (3,500 rpm for 10 minutes). For the preparation of cell-free extracts, bacteria were resuspended in 300 µL buffer A, containing protease inhibitor (Roche cOmplete), lysozyme (1 mg/mL). DNase I (0.1 mg/mL) and $MgSO_4$ (10 mM) to assist in cell lysis and prevent protein degradation. Resuspended cells were incubated for 2 hours at 30 °C at 800 rpm and then stored until further use at -20 °C. The lysates were defrosted and 30 µL of a TCEP stock solution (100 mM in buffer A, adjusted to pH 6 by addition of 6 M NaOH) was added to individual wells. The reduction was initially performed for 2 hours at 30 °C, after which incubation was continued overnight at 4 °C. Subsequently, cell debris was removed by centrifugation (4,000 rpm, 1 hour, 4 °C) and 276 µL of cell-free lysate was transferred into 2 mL microcentrifuge tubes for catalysis, stored at 4 °C and used within 8 hours.

3.4 Catalysis with Cell-Free Extracts and Purified Protein

Reactions were conducted in 300 µL total volume in a 2 mL microcentrifuge tube. Stock solutions of protein in PBS buffer (50 mM NaCl, 150 mM NaH_2PO_4, pH as specified) to give the specified final concentration and the same buffer was added to make up 276 µL volume. For screening of N19 and M89 mutant libraries, 276 µL of cell-free lysate was used instead. Stock solutions of indole (25 mM in DMF, 12 µL added, final concentration 1 mM) and enal (150 mM or 750 mM when using cell-free lysate, 12 µL added to give final concentrations of 6 mM or 18 mM with cell free lysate) substrates were added. The microcentrifuge tubes were then mixed by continuous inversion in a cold room as 4 °C for the specified reaction time. After the reaction time had elapsed, $NaBH_4$ solution (60 µL, 20 mg/mL in 0.5 w/v % NaOH) and 3-(3-hydroxypropyl)indole internal standard solution (12 µL, 5 mM in DMF) were added. The micro-centrifuge tubes were mixed by continuous inversion for a further 30 minutes.

For HPLC analysis (Products 3a, 3b and 3d-i): the reaction products and internal standard were then extracted by vortex mixing with EtOAc (1 mL) and the organic extract was dried over Na_2SO_4, filtered and evaporated to dryness. The residue thus obtained was redissolved by vortex mixing with HPLC grade solvent (heptane:isopropanol 4:1, 90 µL) and analysed by normal phase HPLC to determine yield and enantioselectivity with a 20 µL injection volume (injection volumes were reduced proportionately when indole concentrations above 1 mM were used).

For SFC analysis (cell free extract libraries with product 3a, purified protein with product 3c) 400 µL n-butanol was added to the reactions and vortexed for one minute. The layers were separated with the aid of centrifugation (14,500 rpm, 5 minutes) and 150 µL of the organic layer was taken for SFC analysis, using a 10 µL injection volume.

4. Primer List

Primer	Sequence
M89C_fw	TGAAAACTGTCGCCTGGCGTTCGAAT
M89C_rv	CCAGGCGACAGTTTTCATGGCCGATT
M89D_fw	TGAAAACGATCGCCTGGCGTTCGAAT
M89D_rv	CCAGGCGATCGTTTTCATGGCCGATT
M89E_fw	ATCGGCCATGAAAACGAGCGCCTG
M89E_rv	TCGAACGCCAGGCGCTCGTTTTCA
M89F_fw	TGAAAACTTTCGCCTGGCGTTCGAAT
M89F_rv	CCAGGCGAAAGTTTTCATGGCCGATT
M89G_fw	TGAAAACGGTCGCCTGGCGTTCGAAT
M89G_rv	CCAGGCGACCGTTTTCATGGCCGATT
M89H_fw	TGAAAACCACCGCCTGGCGTTCGAAT
M89H_rv	CCAGGCGGTGGTTTTCATGGCCGATT
M89I_fw	TGAAAACATTCGCCTGGCGTTCGAAT
M89I_rv	CCAGGCGAATGTTTTCATGGCCGATT
M89K_fw	TGAAAACAAGCGCCTGGCGTTCGAAT
M89K_rv	CCAGGCGCTTGTTTTCATGGCCGATT
M89L_fw	TGAAAACCTGCGCCTGGCGTTCGAAT
M89L_rv	CCAGGCGCAGGTTTTCATGGCCGATT
M89N_fw	TGAAAACAATCGCCTGGCGTTCGAAT
M89N_rv	CCAGGCGATTGTTTTCATGGCCGATT
M89P_fw	TGAAAACCCGCGCCTGGCGTTCGAAT
M89P_rv	CCAGGCGCGGGTTTTCATGGCCGATT
M89Q_fw	TGAAAACCAGCGCCTGGCGTTCGAAT
M89Q_rv	CCAGGCGCTGGTTTTCATGGCCGATT
M89R_fw	TGAAAACCGTCGCCTGGCGTTCGAAT
M89R_rv	CCAGGCGACGGTTTTCATGGCCGATT
M89S_fw	TGAAAACTCTCGCCTGGCGTTCGAAT
M89S_rv	CCAGGCGAGAGTTTTCATGGCCGATT
M89T_fw	TGAAAACACCCGCCTGGCGTTCGAAT
M89T_rv	CCAGGCGGGTGTTTTCATGGCCGATT
M89V_fw	TGAAAACGTGCGCCTGGCGTTCGAAT
M89V_rv	CCAGGCGCACGTTTTCATGGCCGATT
M89W_fw	TGAAAACTGGCGCCTGGCGTTCGAAT
M89W_rv	CCAGGCGCCAGTTTTCATGGCCGATT
M89Y_fw	TGAAAACTATCGCCTGGCGTTCGAAT
M89Y_rv	CCAGGCGATAGTTTTCATGGCCGATT
N19C_fw	CTGCTGTGCGTCCTGAAACAAGGC
N19C_rv	TCAGGACGCACAGCAGGATCTAATTGGT
N19D_fw	CTGCTGGATGTCCTGAAACAAGGC
N19D_rv	TCAGGACATCCAGCAGGATCTAATTGGT
N19E_fw	CTGCTGGAAGTCCTGAAACAAGGC
N19E_rv	TCAGGACTTCCAGCAGGATCTAATTGGT
N19F_fw	CTGCTGTTTGTCCTGAAACAAGGC
N19F_rv	TCAGGACAAACAGCAGGATCTAATTGGT
N19G_fw	CTGCTGGGCGTCCTGAAACAAGGC
N19G_rv	TCAGGACGCCCAGCAGGATCTAATTGGT
N19H_fw	CTGCTGCATGTCCTGAAACAAGGC
N19H_rv	TCAGGACATGCAGCAGGATCTAATTGGT
N19I_fw	CTGCTGATTGTCCTGAAACAAGGC
N19I_rv	TCAGGACAATCAGCAGGATCTAATTGGT
N19K_fw	CTGCTGAAAGTCCTGAAACAAGGC
N19K_rv	TCAGGACTTTCAGCAGGATCTAATTGGT
N19L_fw	CTGCTGCTGGTCCTGAAACAAGGC
N19L_rv	TCAGGACCAGCAGCAGGATCTAATTGGT
N19M_fw	CTGCTGATGGTCCTGAAACAAGGC
N19M_rv	TCAGGACCATCAGCAGGATCTAATTGGT
N19P_fw	CTGCTGCCGGTCCTGAAACAAGGC
N19P_rv	TCAGGACCGGCAGCAGGATCTAATTGGT
N19Q_fw	CTGCTGCAGGTCCTGAAACAAGGC
N19Q_rv	TCAGGACCTGCAGCAGGATCTAATTGGT

N19R_fw	CTGCTGCGCGTCCTGAAACAAGGC
N19R_rv	TCAGGACGCGCAGCAGGATCTAATTGGT
N19S_fw	CTGCTGAGCGTCCTGAAACAAGGC
N19S_rv	TCAGGACGCTCAGCAGGATCTAATTGGT
N19T_fw	CTGCTGACCGTCCTGAAACAAGGC
N19T_rv	TCAGGACGGTCAGCAGGATCTAATTGGT
N19V_fw	CTGCTGGTGGTCCTGAAACAAGGC
N19V_rv	TCAGGACCACCAGCAGGATCTAATTGGT
N19W_fw	CTGCTGTGGGTCCTGAAACAAGGC
N19W_rv	TCAGGACCCACAGCAGGATCTAATTGGT
N19Y_fw	CTGCTGTATGTCCTGAAACAAGGC
N19Y_rv	TCAGGACATACAGCAGGATCTAATTGGT

5. Preparation and Characterisation of Reference Products

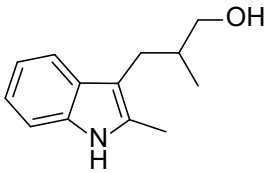
5.1 General Procedure for Preparation of Reference Products

The relevant indole (0.5 or 1 mmol) and enal (1.5 or 3 mmol) were dissolved in hexafluoroisopropanol (2 or 4 mL) and benzylamine (12 µL, 0.11 mmol or 24 µL, 0.22 mmol) was added. The solution was stirred at ambient temperature until consumption of the indole as evidenced by TLC. Methanol (2 or 4 mL) was added, followed by portionwise additions of $NaBH_4$ (113 mg, 3 mmol or 227 mg, 3 mmol). The reaction was stirred for 30 mins at ambient temperature and then partitioned between EtOAc (~20 mL) and brine:water 1:1 (~10 mL) and shaken. The organic phase was dried over Na_2SO_4 and concentrated *in vacuo* and then purified by silica-gel flash chromatography (pentane:EtOAc 4:1). Fractions containing the product were combined and concentrated in vacuo to afford the title compounds.

Note on mass spectrometric analysis: attempts to obtain a HRMS of the molecular ion or its adducts failed despite the use of several ionisation techniques, instead we observed the loss of hydroxide for all reference compounds when using positive electrospray ionisation. This is likely because loss of hydroxide results in a stabilised tertiary carbocation after 1,2-H migration. The molecular proton adduct could be observed without fragmentation when we used a lower resolution technique with a different instrument, for all reference products. The abundance of this ion was very low in all cases, supporting the propensity of these compounds to lose hydroxide under mass spectrometry conditions.

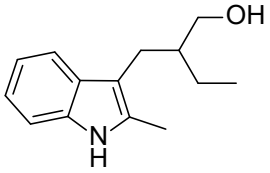
5.2 Characterisation of Reference Products

2-methyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3a)



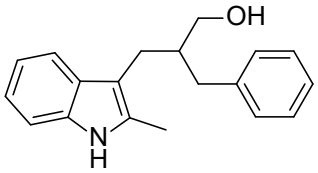
Prepared via the general procedure outlined above from 2-methyl-indole (66 mg, 0.5 mmol) and methacrolein (114 µL, 1.5 mmol). Product obtained as a yellow oil (36 mg, 35 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.74 (s, 1H), 7.50 (d, *J* = 7.6 Hz, 1H), 7.27 (m, 1H), 7.18 – 7.02 (m, 2H), 3.59 (dd, *J* = 10.5, 5.7 Hz, 1H), 3.50 (dd, *J* = 10.5, 5.9 Hz, 1H), 2.79 (dd, *J* = 14.3, 6.9 Hz, 1H), 2.55 (dd, *J* = 14.3, 7.5 Hz, 1H), 2.38 (s, 3H), 2.09 – 1.96 (m, 1H), 0.97 (d, *J* = 6.8 Hz, 3H). The spectral data are in accordance with the literature[7]. LRMS calc'd for $C_{13}H_{18}NO$ ($[M+H]^+$) 204.1; measured 204.2. HRMS: calc'd for $C_{13}H_{16}N$ ($[M-OH]^+$) 186.1277; measured 186.1274. HPLC analysis: Chiracel OJ-H heptane:isopropanol 80:20 1 mL/min retention times 9.8 min and 10.8 min. SFC analysis: super-critical CO_2 (A) and methanol (B) were used for the mobile phase at a flow rate of 1.8 mL/min and a Trefoil AMY1 column was used for chiral separation. Program: 97% A with a linear gradient to 59% A in 3 min, 40% A for 1 min, 97% A for 1 min. Retention times of the enantiomers – 2.37 min and 2.48 min.

2-((2-methyl-1H-indol-3-yl)methyl)butan-1-ol (3b)



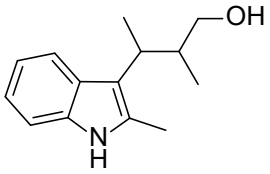
Prepared via the general procedure outlined above from 2-methyl-indole (66 mg, 0.5 mmol) and 2-ethyl-acrolein (150 µL, 1.5 mmol). Product obtained as a yellow oil (41 mg, 38 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.75 (s, 1H), 7.51 (d, *J* = 7.5 Hz, 1H), 7.26 (d, *J* = 8.4 Hz, 1H), 7.14 – 7.04 (m, 1H), 3.63 – 3.52 (m, 2H), 2.75 – 2.64 (m, 2H), 2.38 (s, 3H), 1.86 – 1.75 (m, 1H), 1.54 – 1.37 (m, 2H), 1.00 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 135.5, 131.6, 129.2, 121.2, 119.4, 118.4, 110.8, 110.3, 65.4, 43.8, 26.1, 24.2, 12.0, 11.8. LRMS calc'd for $C_{14}H_{20}NO$ ($[M+H]^+$) 218.2; measured 218.3. HRMS calc'd for $C_{14}H_{18}N$ ($[M-OH]^+$) 200.1434; measured 200.1433. HPLC analysis Chiracel AS-H heptane:isopropanol 90:10 0.5 mL/min retention times of the enantiomers 11.8 and 12.7 min.

2-benzyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3c)



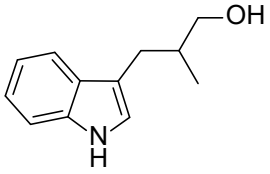
Prepared via the general procedure outlined above from 2-methyl-indole (66 mg, 0.5 mmol) and 2-benzyl-acrolein (114 µL, 1.5 mmol). Product obtained as a yellow oil (20 mg, 14 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.75 (s, 1H), 7.40 (d, *J* = 7.8 Hz, 1H), 7.34 – 7.27 (m, 3H), 7.22 (d, *J* = 7.1 Hz, 3H), 7.11 (td, *J* = 7.5, 1.3 Hz, 1H), 7.06 (td, *J* = 7.4, 1.2 Hz, 1H), 3.58 – 3.48 (m, 2H), 2.85 – 2.67 (m, 4H), 2.35 (s, 3H), 2.24 – 2.14 (m, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 141.1, 135.4, 131.7, 129.3, 129.1, 128.5, 126.1, 121.1, 119.3, 118.2, 110.4, 110.3, 64.9, 44.2, 38.3, 26.0, 11.9. LRMS calc'd for $C_{19}H_{22}NO$ ([M+H]$^+$) 280.2; measured 280.3. HRMS calc'd for $C_{19}H_{20}N$ ([M-OH]$^+$) 262.1590; measured 282.1591. SFC analysis: super-critical CO_2 (A) and methanol (B) were used for the mobile phase at a flow rate of 1.8 mL/min and a Chiralcel OJ-3 column was used for chiral separation. Program: 99% A with a linear gradient to 90% A in 8.5 min, linear gradient to 50% A in 0.5 min, 50% A for 1 min, 99% A for 1 min. Retention times of the enantiomers – 9.43 min and 9.61 min.

2-methyl-3-(2-methyl-1H-indol-3-yl)butan-1-ol (3d)



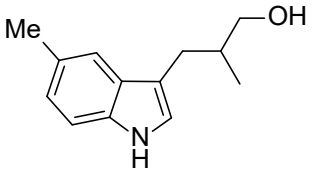
Tiglic aldehyde was added to a solution of 2-Me-indole (250 mg) in 8 mL DCM thereafter catalysts tosylic acid (31 mg) and benzylamine (19 µL) were added. The reaction mixture was stirred at 30 °C for 70 h. The solvent was removed *in vacuo* after which the intermediate was dissolved in 8 mL MeOH. $NaBH_4$ (400 mg, 6 eq) was added and the mixture was stirred at room temperature for 60 minutes. The solvent was removed *in vacuo* and the crude product was dissolved in 20 mL EtOAc thereafter it was washed with brine (3x 10 mL) in a separation funnel. The organic layer was dried over Na_2SO_4. The crude was purified by silica column (6:1, pentane:EtOAc). The product was obtained as a yellow oil in a mixture of diastereomers (44.8 mg, 10.7 %). First diastereomer: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (s, 1H), 7.64 – 7.58 (m, 1H), 7.29 – 7.24 (m, 1H), 7.13 – 7.07 (m, 1H), 7.07 – 7.00 (m, 1H), 3.82 (dd, *J* = 10.5, 3.8 Hz, 1H), 3.66 (dd, *J* = 10.5, 6.3 Hz, 1H), 3.00 – 2.87 (m, 1H), 2.38 (s, 3H), 2.22 – 2.09 (m, 1H), 1.45 (d, *J* = 7.2 Hz, 3H), 0.84 (d, *J* = 6.8 Hz, 3H). Second diastereomer: ^{1}H NMR (400 MHz, cdcl$_3$) δ 7.71 (s, 1H), 7.64 – 7.58 (m, 1H), 7.29 – 7.24 (m, 1H), 7.13 – 7.07 (m, 1H), 7.07 – 7.00 (m, 1H), 3.45 (dd, *J* = 10.5, 4.1 Hz, 1H), 3.33 – 3.22 (m, 1H), 2.86 – 2.74 (m, 1H), 2.38 (s, 3H), 2.22 – 2.09 (m, 1H), 1.42 (d, *J* = 7.1 Hz, 3H), 1.15 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 135.6, 130.7, 121.0, 120.7, 119.8, 119.6, 119.1, 118.9, 115.6, 110.5, 110.4, 67.7, 66.8, 41.3, 40.9, 34.0, 33.3, 18.8, 18.8, 16.0, 15.9, 12.5, 12.4. (Peaks for both diastereomers). LRMS calc'd for $C_{14}H_{20}NO$ ([M+H]$^+$) 218.2; measured 218.3; HRMS calc'd for $C_{14}H_{18}N$ ([M-OH]$^+$) 200.1432; measured 200.1429. HPLC analysis Chiralcel OD-H heptane:isopropanol 90:10 0.5 mL/min retention times 25.9 and 27.2 min (enantiomers of first diastereomer) and 32.6 and 33.9 min (enantiomers of second diastereomer)

3-(1H-indol-3-yl)-2-methylpropan-1-ol (3e)



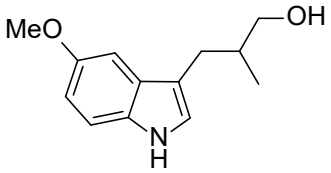
Prepared via the general procedure outlined above from indole (118 mg, 1 mmol) and methacrolein (228 µL, 3 mmol). Product obtained as a slightly turbid colourless oil (108 mg, 53 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.97 (s, 1H), 7.63 (d, J = 7.8 Hz, 1H), 7.36 (d, J = 8.1 Hz, 1H), 7.20 (t, J = 7.5 Hz, 1H), 7.12 (t, J = 7.5 Hz, 1H), 7.00 (s, 1H), 3.60 (dd, J = 10.5, 5.9 Hz, 1H), 3.52 (dd, J = 10.5, 6.0 Hz, 1H), 2.87 (dd, J = 14.4, 6.6 Hz, 1H), 2.63 (dd, J = 14.4, 7.4 Hz, 1H), 2.15 – 2.03 (m, 1H), 0.99 (d, J = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 136.5, 128.0, 122.2, 122.1, 119.4, 119.3, 114.9, 111.3, 68.3, 36.9, 29.1, 17.2. LRMS calc'd for $C_{12}H_{16}NO$ ([M+H]$^+$) 190.1; measured 190.2. HRMS calc'd for $C_{12}H_{14}N$ ([M-OH]$^+$) 172.1121; measured 172.1121. HPLC analysis Chiralcel OD-H heptane:isopropanol 85:15 1 mL/min retention times of the enantiomers 14.4 and 15.5 min.

2-methyl-3-(5-methyl-1H-indol-3-yl)propan-1-ol (3f)



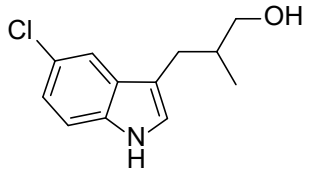
Prepared via the general procedure outlined above from 5-methyl-indole (66 mg, 0.5 mmol) and methacrolein (114 µL, 1.5 mmol). Product obtained as a colourless oil (31 mg, 31 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.87 (s, 1H), 7.39 (s, 1H), 7.25 (d, J = 8.6 Hz, 1H), 7.02 (d, J = 8.3 Hz, 1H), 6.96 (s, 1H), 3.60 (dd, J = 10.5, 5.8 Hz, 1H), 3.52 (dd, J = 10.5, 6.0 Hz, 1H), 2.84 (dd, J = 14.3, 6.6 Hz, 1H), 2.60 (dd, J = 14.3, 7.5 Hz, 1H), 2.46 (s, 3H), 2.16 – 2.04 (m, 1H), 0.99 (d, J = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 134.9, 128.7, 128.2, 123.8, 122.4, 118.9, 114.4, 110.9, 68.4, 36.8, 29.2, 21.7, 17.3. LRMS calc'd for $C_{13}H_{18}NO$ ($[M+H]^+$) 204.1; measured 204.2. HRMS calc'd for $C_{13}H_{16}N$ ($[M-OH]^+$) 186.1277; measured 186.1278. HPLC analysis Chiralcel OJ-H heptane:isopropanol 80:20 retention time of the enantiomers 8.6 and 10.9 min.

3-(5-methoxy-1H-indol-3-yl)-2-methylpropan-1-ol (3g)



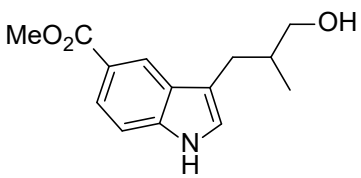
Prepared via the general procedure outlined above from 5-methoxy-indole (74 mg, 0.5 mmol) and methacrolein (114 µL, 1.5 mmol). Product obtained as a yellow oil (40 mg, 37 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.87 (s, 1H), 7.25 (d, J = 9.3 Hz, 1H), 7.05 (d, J = 2.4 Hz, 1H), 6.98 (d, J = 2.4 Hz, 1H), 6.86 (dd, J = 8.8, 2.3 Hz, 1H), 3.87 (s, 3H), 3.60 (dd, J = 10.5, 6.1 Hz, 1H), 3.52 (dd, J = 10.5, 6.0 Hz, 1H), 2.83 (dd, J = 14.4, 6.6 Hz, 1H), 2.60 (dd, J = 14.4, 7.4 Hz, 1H), 2.16 – 2.01 (m, 1H), 0.99 (d, J = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.1, 131.7, 128.5, 123.1, 114.6, 112.2, 112.0, 101.3, 68.3, 56.2, 36.8, 29.2, 17.3. LRMS calc'd for $C_{13}H_{18}NO_2$ ($[M+H]^+$) 220.1; measured 220.2. HRMS calc'd for $C_{13}H_{16}NO$ ($[M-OH]^+$) 202.1226; measured 202.1225. HPLC analysis Chiracel OJ-H heptane:isopropanol 80:20 1 mL/min retention time of the enantiomers 11.0 and 12.1 min.

3-(5-chloro-1H-indol-3-yl)-2-methylpropan-1-ol (3h)



Prepared via the general procedure outlined above from 5-chloro-indole (76 mg, 0.5 mmol) and methacrolein (114 µL, 1.5 mmol). 1 mL methanol was added to the reaction mixture to ensure complete dissolution of the indole starting material. Product obtained as a slightly turbid colourless oil (21 mg, 19 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 (s, 1H), 7.58 (s, 1H), 7.27 (d, J = 7.9 Hz, 1H), 7.13 (dd, J = 8.6, 1.7 Hz, 1H), 7.03 (d, J = 1.7 Hz, 1H), 3.64 – 3.47 (m, 2H), 2.84 (dd, J = 14.4, 6.4 Hz, 1H), 2.57 (dd, J = 14.4, 7.6 Hz, 1H), 2.11 – 1.99 (m, 1H), 0.97 (d, J = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 134.8, 129.2, 125.3, 123.7, 122.4, 118.8, 114.8, 112.3, 68.1, 36.8, 28.9, 17.1. LRMS calc'd for $C_{12}H_{15}{}^{35}ClNO$ ($[M+H]^+$) 224.1; measured 224.2; calc'd for $C_{12}H_{15}{}^{37}ClNO$ ($[M+H]^+$) 226.1; measured 226.2. HRMS calcd for $C_{12}H_{13}{}^{35}ClN$ ($[M-OH]^+$) 206.0731; measured 206.0729; calc'd $C_{12}H_{13}{}^{37}ClN$ ($[M-OH]^+$) 208.0702; measured 208.0702. HPLC analysis Chiralcel OJ-H heptane:isopropanol 76:24 1 mL/min retention time of the enantiomers 6.0 and 6.6 min.

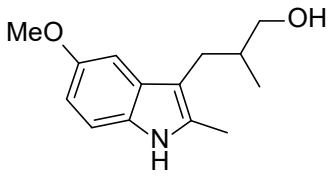
methyl 3-(3-hydroxy-2-methylpropyl)-1H-indole-5-carboxylate (3i)



Prepared via the general procedure outlined above from methyl indole-5-carboxylate (88 mg, 0.5 mmol) and methacrolein (114 µL, 1.5 mmol). Product obtained as a white solid (43 mg, 35 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.38 (s, 1H), 8.16 (s, 1H), 7.90 (d, *J* = 8.6 Hz, 1H), 7.36 (d, *J* = 8.6 Hz, 1H), 7.07 (s, 1H), 3.94 (s, 3H), 3.63 – 3.50 (m,

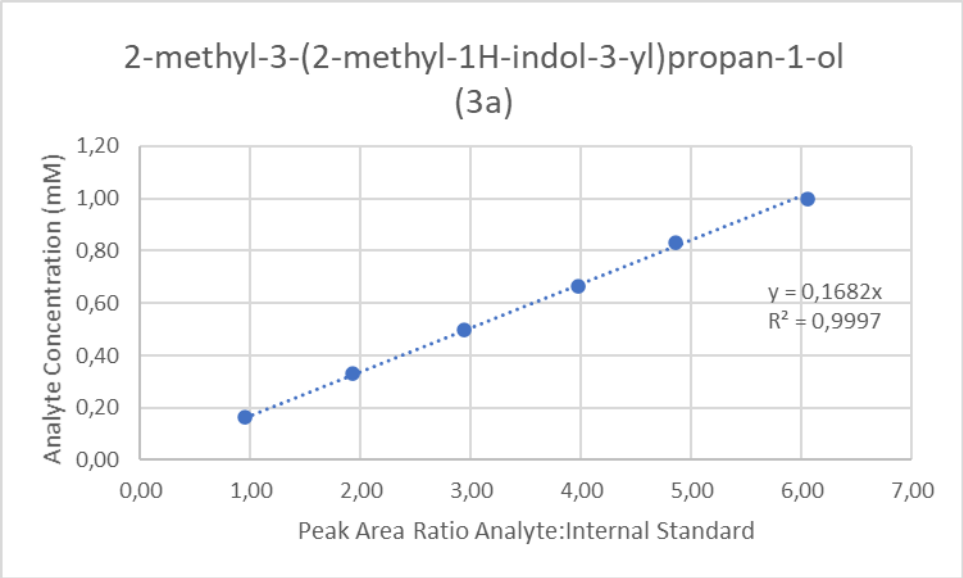
2H), 2.91 (dd, *J* = 14.4, 6.5 Hz, 1H), 2.65 (dd, *J* = 14.4, 7.7 Hz, 1H), 2.20 – 2.04 (m, 1H), 0.99 (d, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 168.4, 139.1, 127.7, 123.6, 123.5, 122.3, 121.6, 116.5, 110.9, 68.1, 52.1, 36.9, 28.9, 17.1.. LRMS calc'd for $C_{14}H_{18}NO_3$ ($[M+H]^+$) 248.1; measured 248.2. HRMS calc'd for $C_{14}H_{16}NO_2$ ($[M-OH]^+$) 230.1176; measured 230.1176. HPLC analysis Chiralcel AS-H heptane:isopropanol 90:10 0.5 mL/min retention time of the enantiomers 18.9 and 20.2 min.

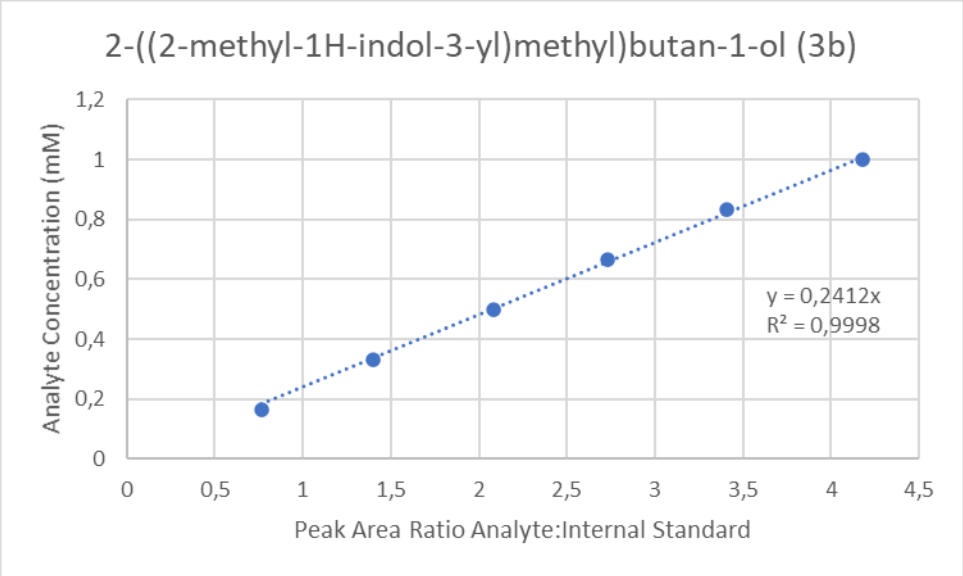
3-(5-methoxy-2-methyl-1H-indol-3-yl)-2-methylpropan-1-ol (3j)

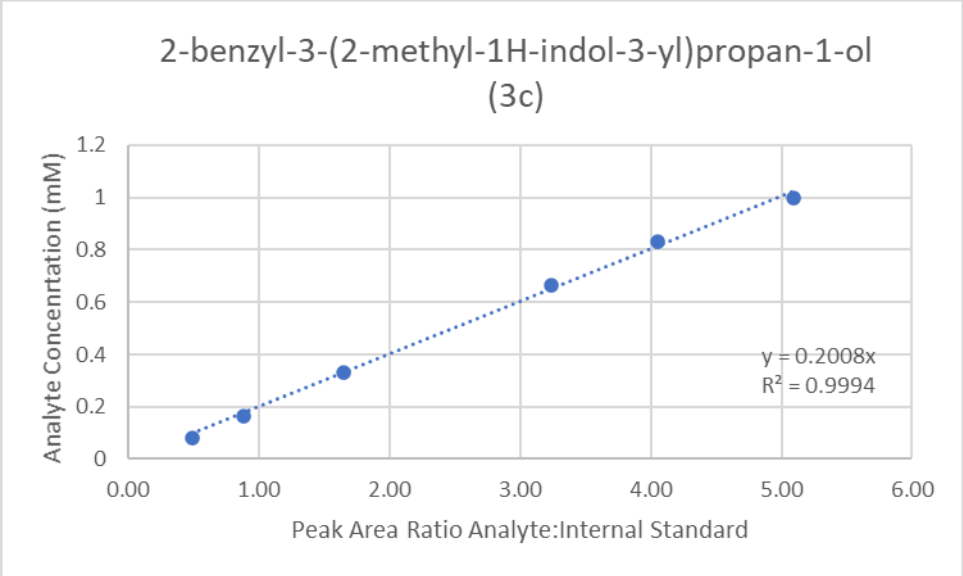


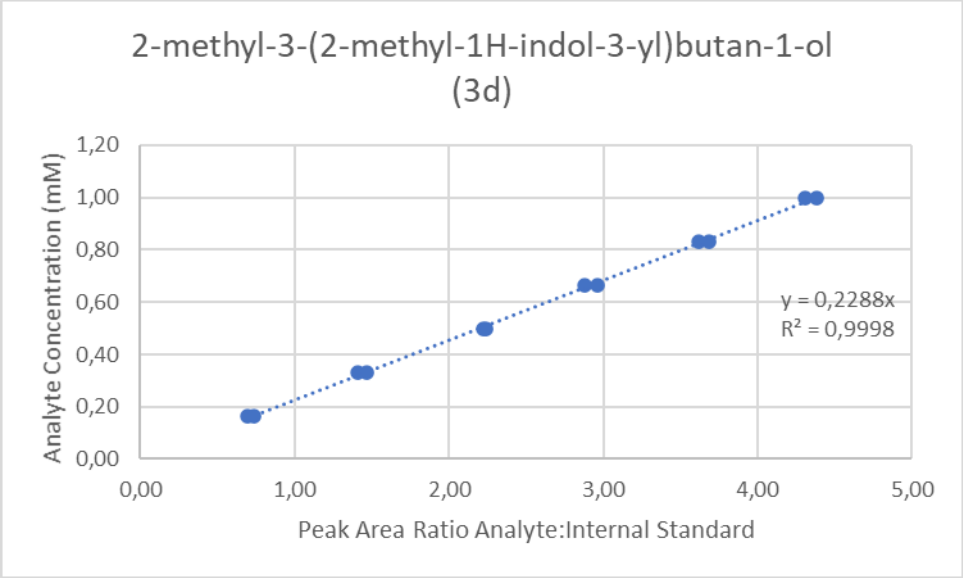
Prepared via the general procedure outlined above from 5-methoxy-2-methly-indole (81 mg, 0.5 mmol) and methacrolein (114 µL, 1.5 mmol). Product obtained as a yellow oil (48 mg, 41 % yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.64 (s, 1H), 7.15 (d, J = 8.7 Hz, 1H), 6.97 (d, J = 2.1 Hz, 1H), 6.77 (dd, J = 8.7, 2.1 Hz, 1H), 3.86 (s, 3H), 3.59 (dd, J = 10.4, 5.7 Hz, 1H), 3.50 (dd, J = 10.4, 6.0 Hz, 1H), 2.74 (dd, J = 14.3, 6.9 Hz, 1H), 2.52 (dd, J = 14.3, 7.5 Hz, 1H), 2.36 (s, 3H), 2.13 – 1.95 (m, 1H), 0.97 (d, J = 6.7 Hz, 3H) ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.1, 132.7, 130.6, 129.7, 110.9, 110.5, 110.5, 101.2, 68.4, 56.3, 37.4, 28.2, 17.3, 12.1. LRMS calc'd for $C_{14}H_{20}NO_2$ (M+H) 234.1; measured 234.2. HRMS calc'd for $C_{14}H_{18}NO$ (M-OH) 216.1383; measured 216.1382. HPLC analysis Chiralcel OD-H heptane:isopropanol 85:15 1 mL/min retention time of the enantiomers 14.2 and 15.3 min.

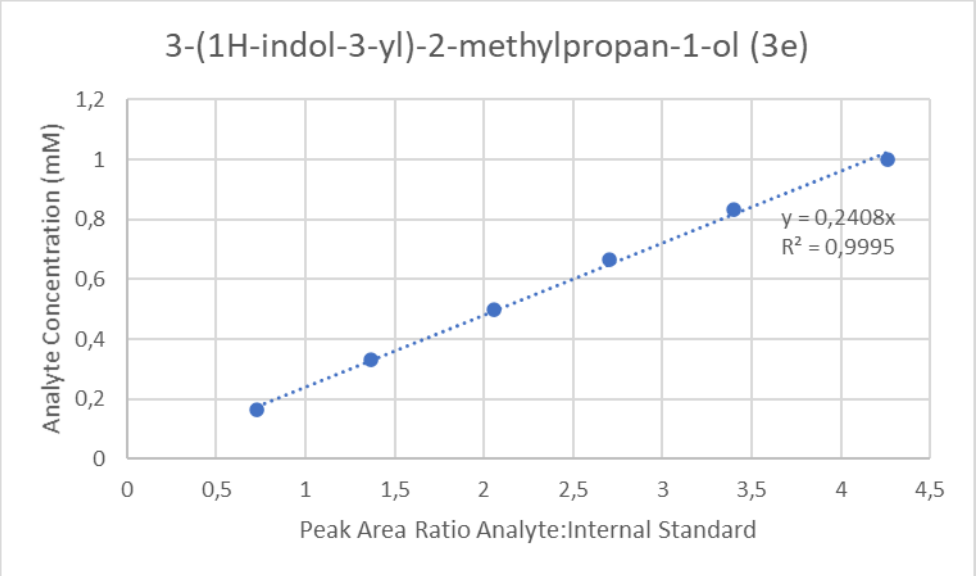
6. Calibration Curves

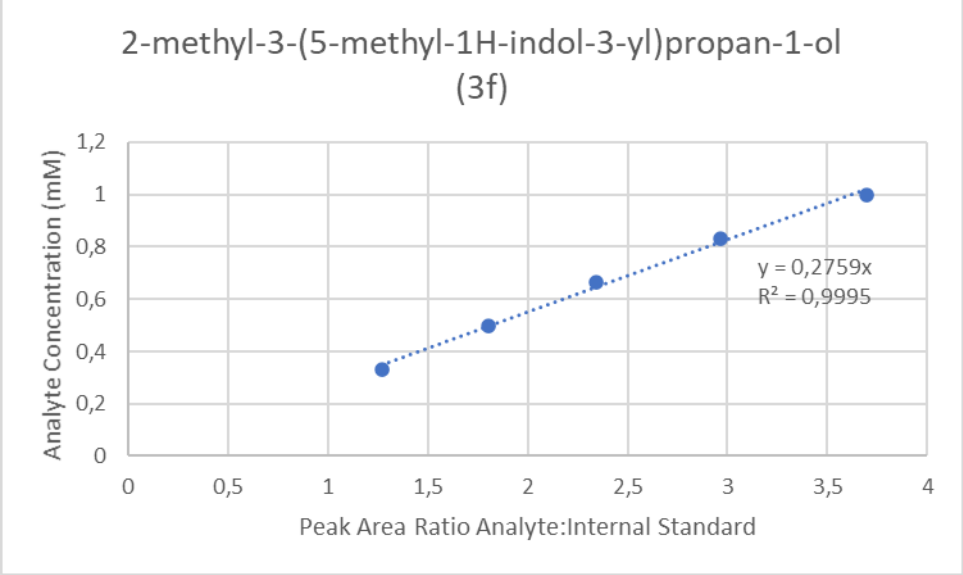


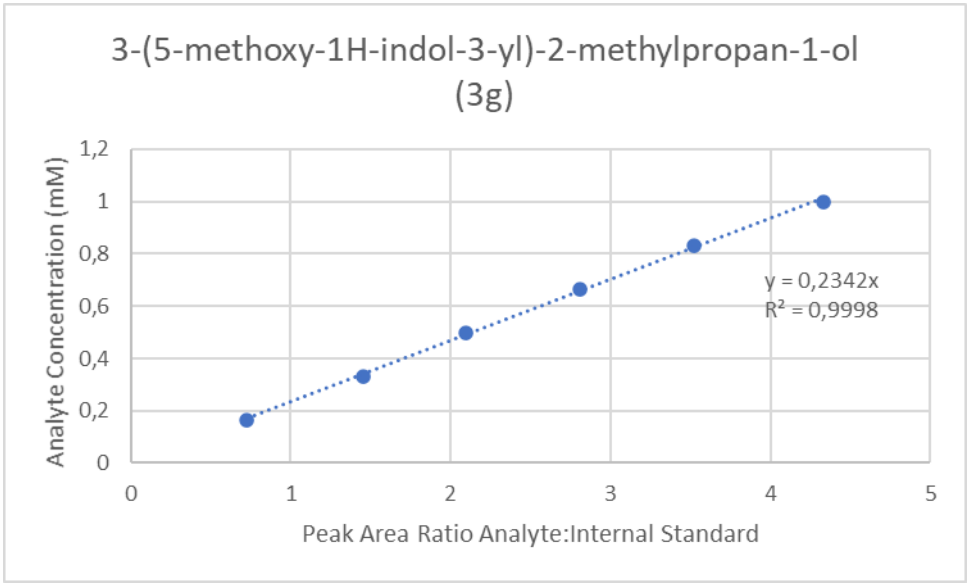


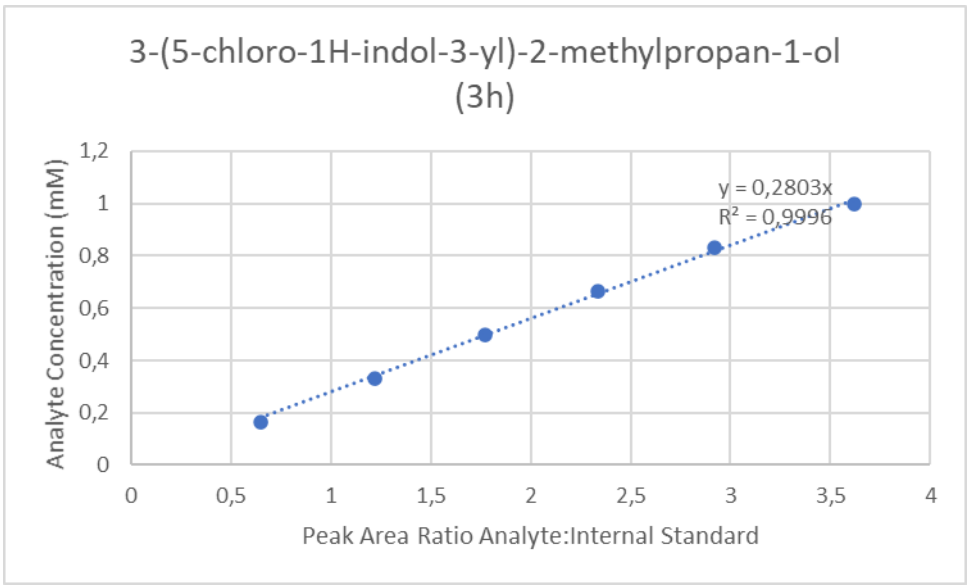


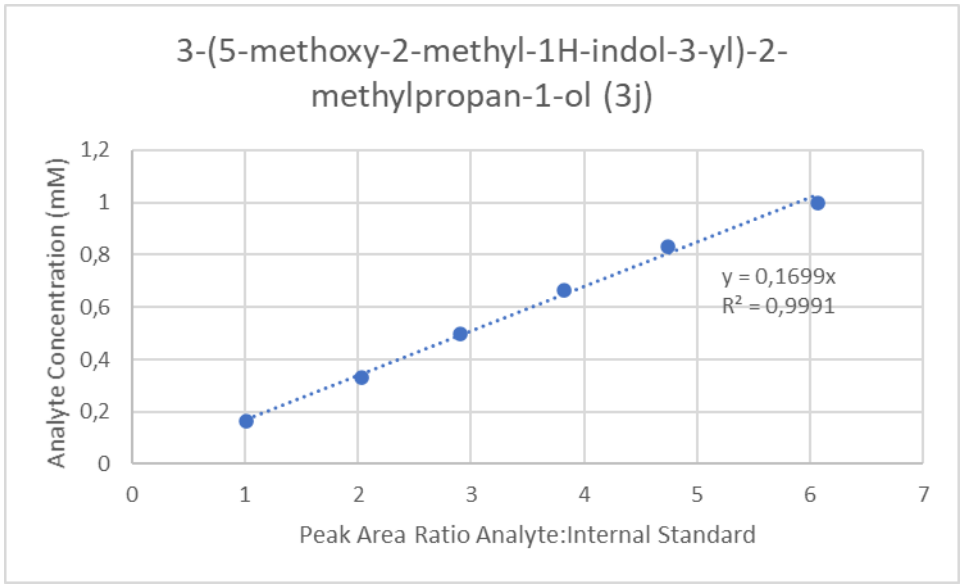

2-methyl-3-(2-methyl-1H-indol-3-yl)butan-1-ol (3d)
Analyte Concentration (mM)
1,20
1,00
0,80
0,60
0,40
0,20
0,00
0,00
1,00
2,00
3,00
4,00
5,00
y = 0,2288x
R² = 0,9998
Peak Area Ratio Analyte:Internal Standard


3-(1H-indol-3-yl)-2-methylpropan-1-ol (3e)
Analyte Concentration (mM)
1,2
1
0,8
0,6
0,4
0,2
0
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
y = 0,2408x
R² = 0,9995
Peak Area Ratio Analyte:Internal Standard


2-methyl-3-(5-methyl-1H-indol-3-yl)propan-1-ol (3f)
Analyte Concentration (mM)
1,2
1
0,8
0,6
0,4
0,2
0
0
0,5
1
1,5
2
2,5
3
3,5
4
y = 0,2759x
R² = 0,9995
Peak Area Ratio Analyte:Internal Standard


3-(5-methoxy-1H-indol-3-yl)-2-methylpropan-1-ol (3g)
Analyte Concentration (mM)
1,2
1
0,8
0,6
0,4
0,2
0
0
1
2
3
4
5
y = 0,2342x
R² = 0,9998
Peak Area Ratio Analyte:Internal Standard

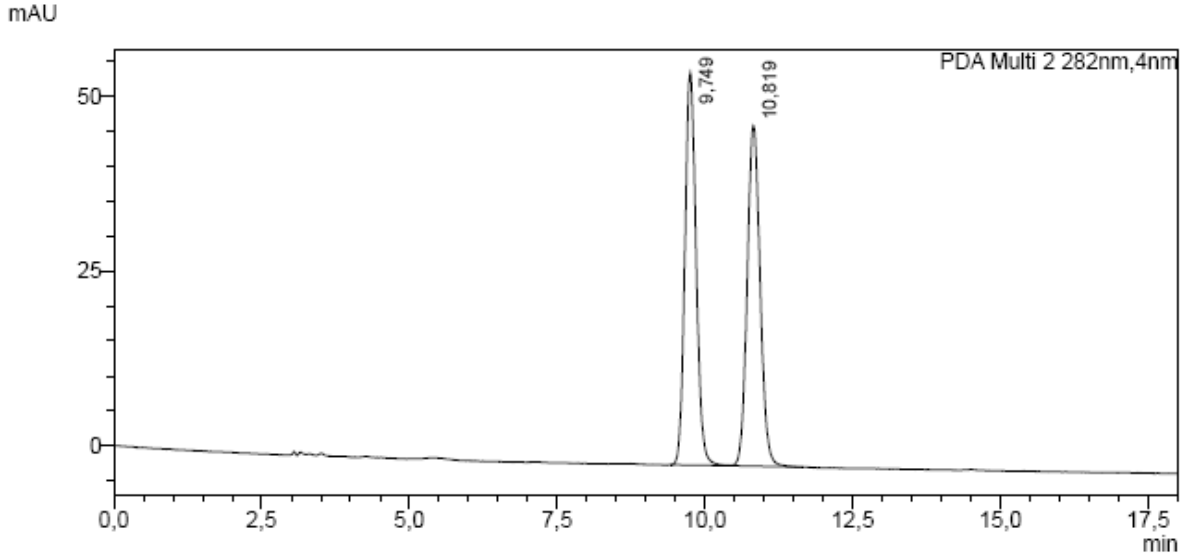

3-(5-chloro-1H-indol-3-yl)-2-methylpropan-1-ol (3h)
Analyte Concentration (mM)
1,2
1
0,8
0,6
0,4
0,2
0
0
0,5
1
1,5
2
2,5
3
3,5
4
y = 0,2803x
R² = 0,9996
Peak Area Ratio Analyte:Internal Standard


3-(5-methoxy-2-methyl-1H-indol-3-yl)-2-methylpropan-1-ol (3j)
Analyte Concentration (mM)
1,2
1
0,8
0,6
0,4
0,2
0
0
1
2
3
4
5
6
7
y = 0,1699x
R² = 0,9991
Peak Area Ratio Analyte:Internal Standard

7. HPLC Chromatograms

2-methyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3a)

Reference product:

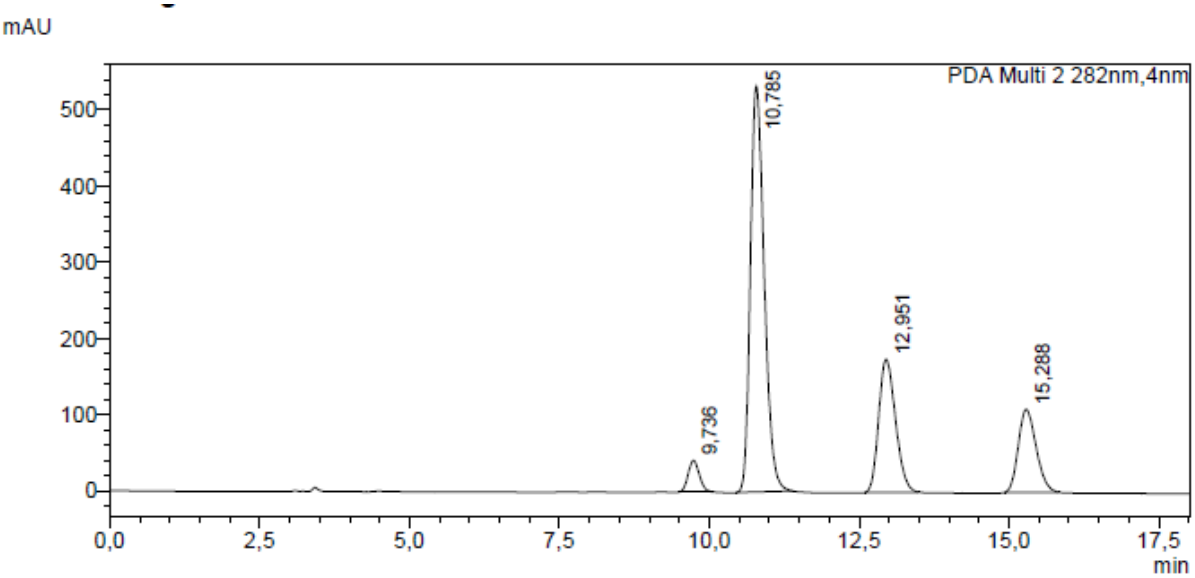


<Peak Table>

PDA Ch2 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	9,749	739769	56018	49,970			
2	10,819	740666	48623	50,030			
Total		1480435	104641				

Catalysis sample: 10 µM LmrR_pAF pH 6. 6 hours reaction time at 4°. 13.0 min = 2-methyl-indole. 15.3 min = 3-(3-hydroxypropyl)-indole (internal standard).



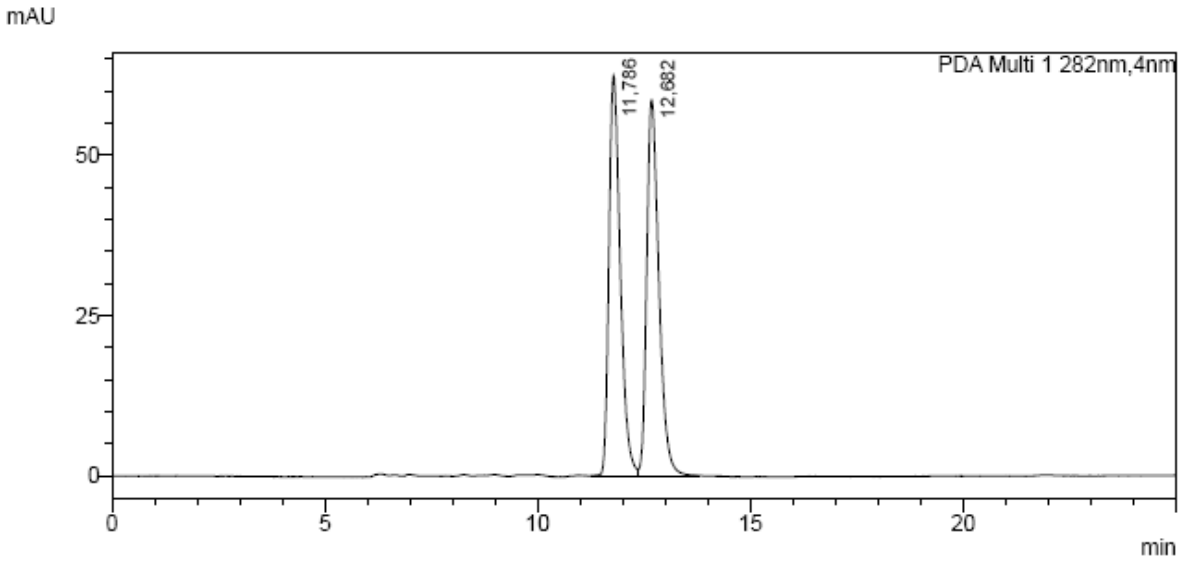
<Peak Table>

PDA Ch2 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	9,736	536858	41397	0,000			
2	10,785	8547512	532680	0,000			
3	12,951	3412447	174615	0,000			
4	15,288	2250590	108690	0,000			
Total		14747407	857382				

2-((2-methyl-1H-indol-3-yl)methyl)butan-1-ol (3b)

Reference product:

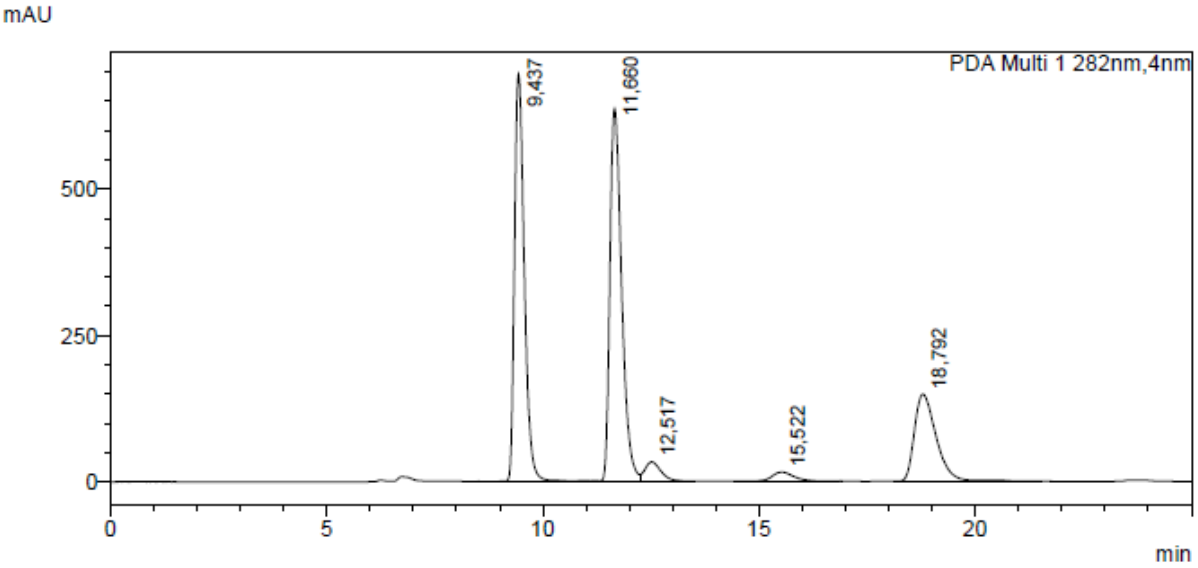


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	11,786	1156851	62322	0,000			
2	12,682	1163857	58529	0,000		SV	
Total		2320708	120850				

Catalysis sample: 10 µM LmrR_pAF pH 6. 6 hours reaction time at 4°. 9.4 min = 2-methyl-indole. 18.8 min = 3-(3-hydroxypropyl)-indole (internal standard).



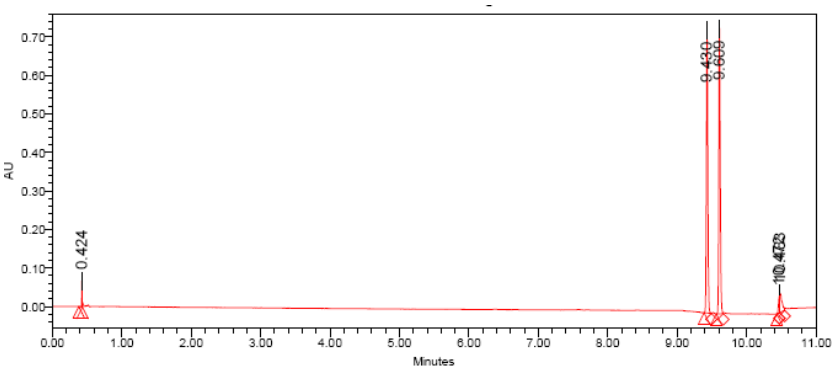
<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	9,437	11082272	696206	37,062		SV	
2	11,660	12047177	637707	40,289		V	
3	12,517	909755	33637	3,042		V	
4	15,522	599441	16129	2,005		V	
5	18,792	5263220	149035	17,602		S	
Total		29901865	1532714				

2-benzyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3c)

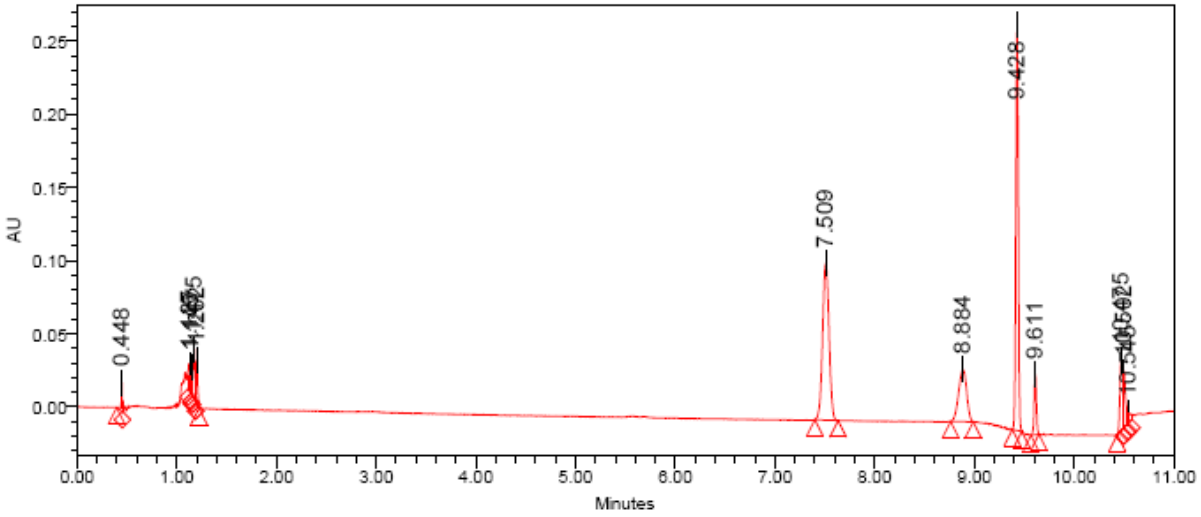
Reference product:



Peak Results

	Name	RT	Area	Height	% Area
1		0.424	28233	68531	1.23
2		9.430	1053073	735221	45.79
3		9.609	1060366	741880	46.10
4		10.472	40160	43148	1.75
5		10.483	118177	49805	5.14

Catalysis sample: 25 µM LmrR_pAF pH 6. 16 hours reaction time at 4°. 7.5 min = 2-methyl-indole. 8.9 min = 3-(3-hydroxypropyl)-indole (internal standard).

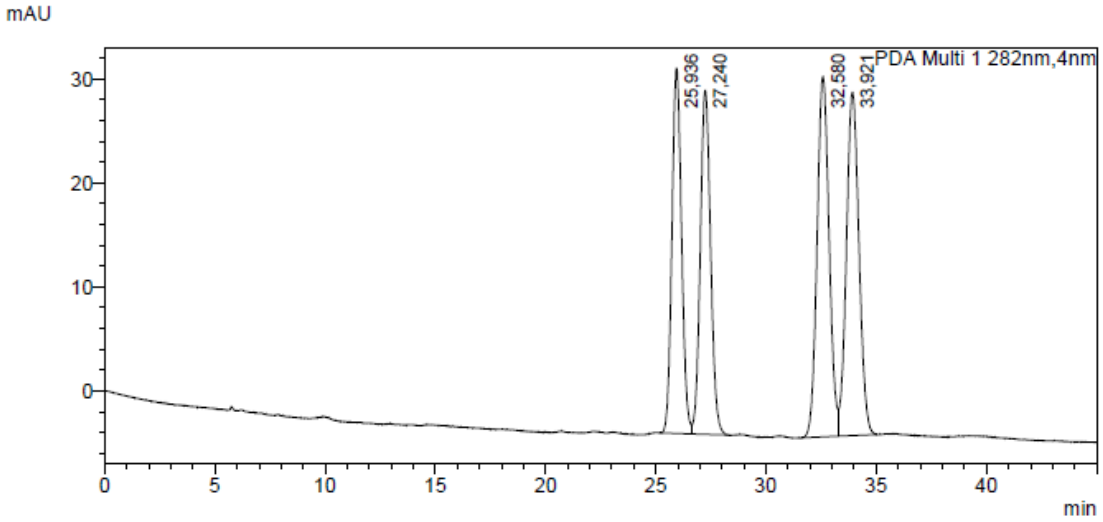


Peak Results

	Name	RT	Area	Height	% Area
1		0.448	11020	19628	0.73
2		1.127	25665	18509	1.70
3		1.145	12729	18633	0.84
4		1.175	49773	34558	3.30
5		1.202	29346	30191	1.95
6		7.509	480325	106813	31.85
7		8.884	199165	35517	13.20
8		9.428	469351	276902	31.12
9		9.611	63101	41056	4.18
10		10.475	105837	48099	7.02
11		10.502	51123	38082	3.39

2-methyl-3-(2-methyl-1H-indol-3-yl)butan-1-ol (3d)

Reference product:

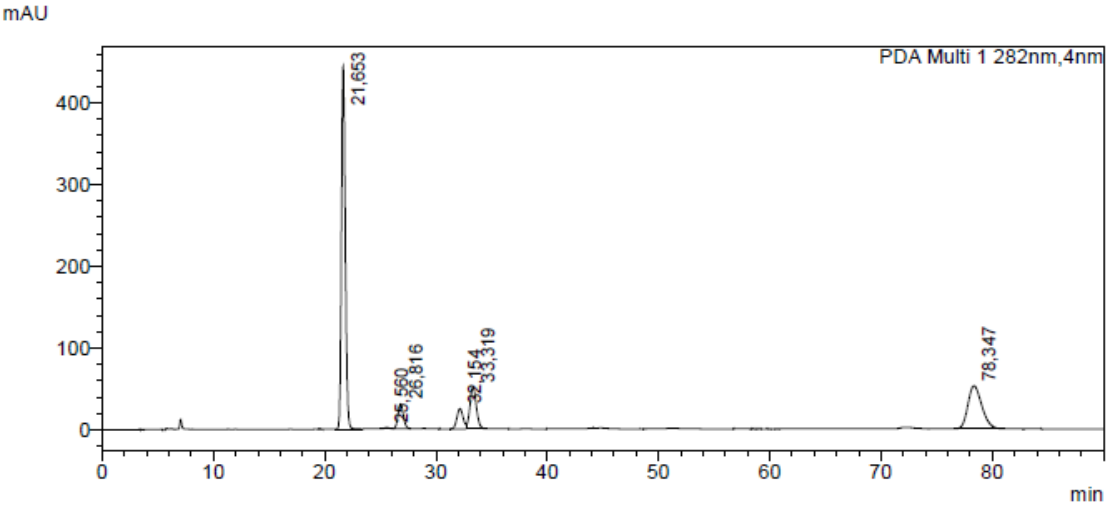


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	25,936	1064487	35056	22,342			
2	27,240	1078263	33040	22,631		V	
3	32,580	1312015	34614	27,538			
4	33,921	1309699	32971	27,489		V	
Total		4764463	135681				

Catalysis sample: 25 µM LmrR_pAF pH 6. 16 hours reaction time at 4°. 21.7 min = 2-methyl-indole. 78.3 min = 3-(3-hydroxypropyl)-indole (internal standard).

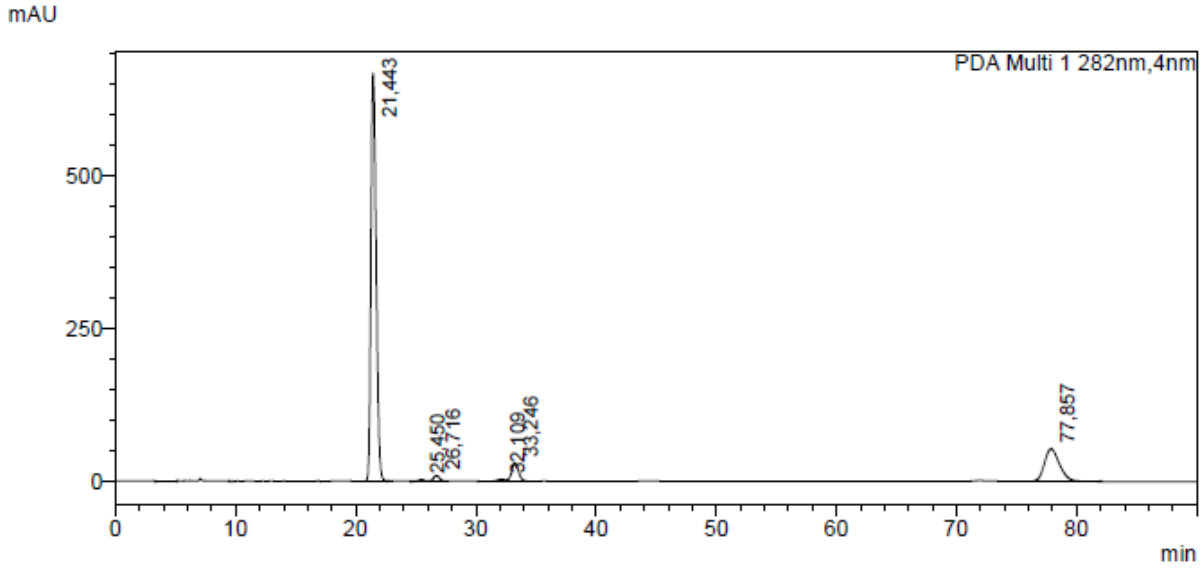


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	21,653	10659024	443565	55,058			
2	25,560	46761	1493	0,242			
3	26,816	1029775	30523	5,319		V	
4	32,154	958332	24775	4,950			
5	33,319	2054049	51266	10,610		V	
6	78,347	4611513	52665	23,820			
Total		19359454	604288				

Catalysis sample: 25 µM LmrR_pAF_RGN pH 6. 16 hours reaction time at 4°. 21.4 min = 2-methyl-indole. 77.9 min = 3-(3-hydroxypropyl)-indole (internal standard).



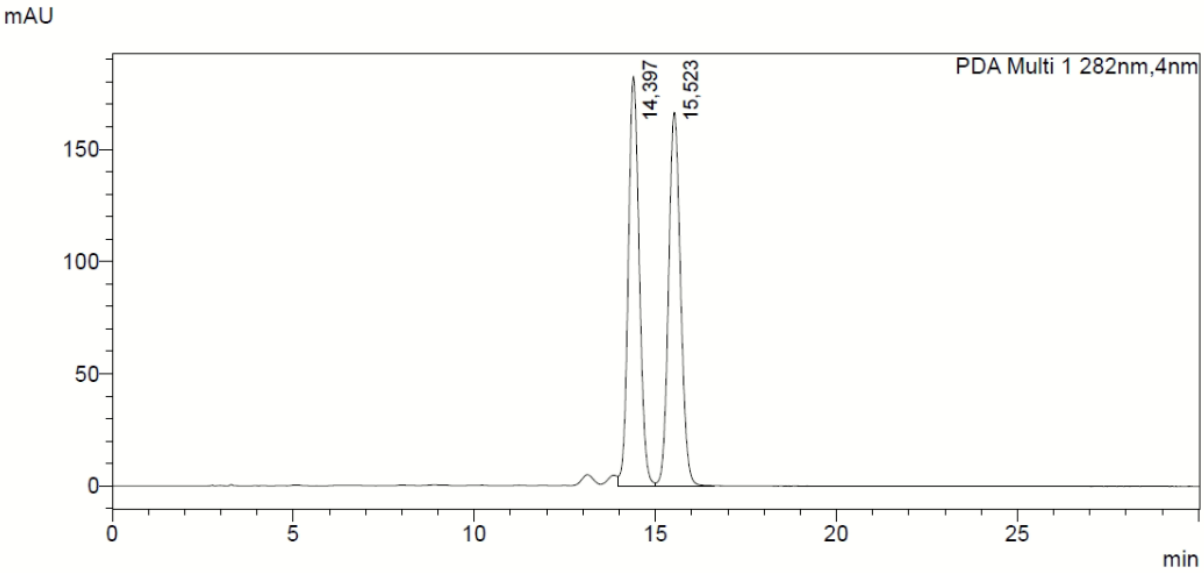
<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	21,443	20051618	667605	75,783			
2	25,450	84052	2597	0,318			
3	26,716	322964	9377	1,221		V	
4	32,109	135608	3536	0,513			
5	33,246	1206446	30227	4,560		V	
6	77,857	4658434	53273	17,606		S	
Total		26459123	766615				

3-(1H-indol-3-yl)-2-methylpropan-1-ol (3e)

Reference product:

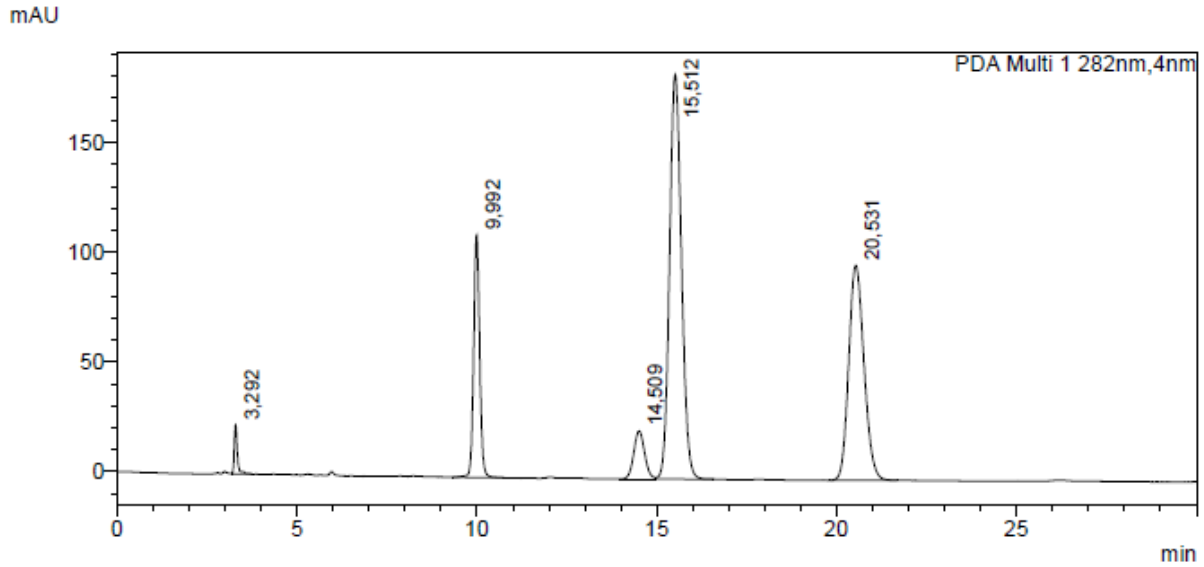


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	14,397	3841134	182233	50,091		V	
2	15,523	3827240	166374	49,909		V	
Total		7668374	348607				

Catalysis sample: 25 µM LmrR_pAF pH 6. 16 hours reaction time at 4°. 9.9 min = indole. 20.5 min = 3-(3-hydroxypropyl)-indole (internal standard).



<Peak Table>

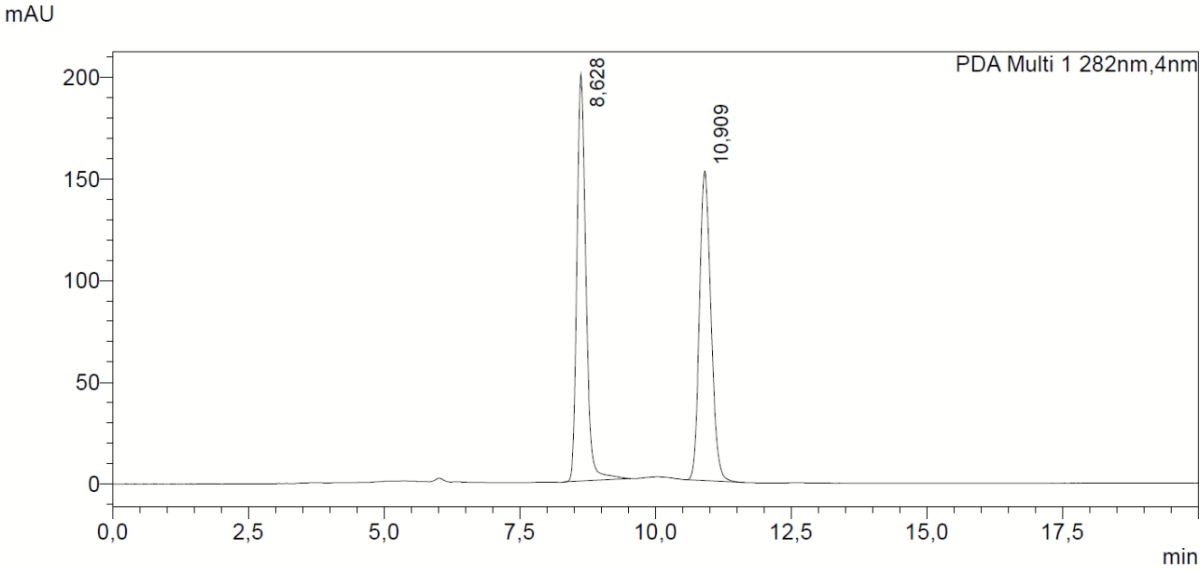
PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	3,292	134660	22413	0,000		SV	
2	9,992	1287082	110493	0,000			
3	14,509	457834	21866	0,000			
4	15,512	4240937	184187	0,000		V	
5	20,531	2895440	97955	0,000			
Total		9015953	436914				

2-methyl-3-(5-methyl-1H-indol-3-yl)propan-1-ol (3f)

Reference product:

<Chromatogram>

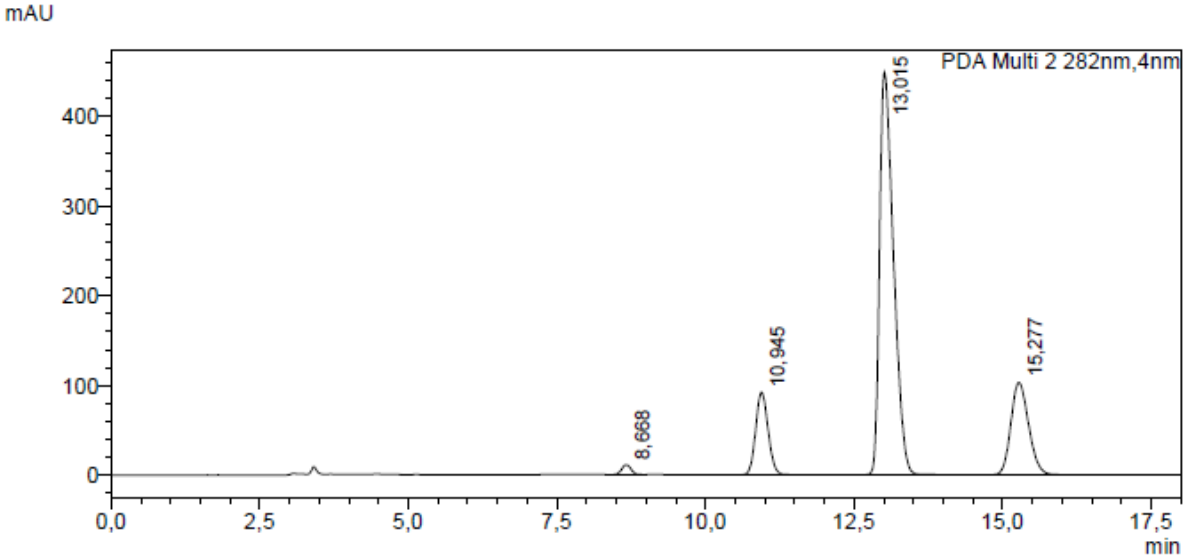


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	8,628	2313819	199883	50,658			
2	10,909	2253692	152113	49,342			
Total		4567511	351996				

Catalysis sample: 10 µM LmrR_pAF pH 6. 6 hours reaction time at 4°. 13.0 min = 5-methyl-indole. 15.3 min = 3-(3-hydroxypropyl)-indole (internal standard).



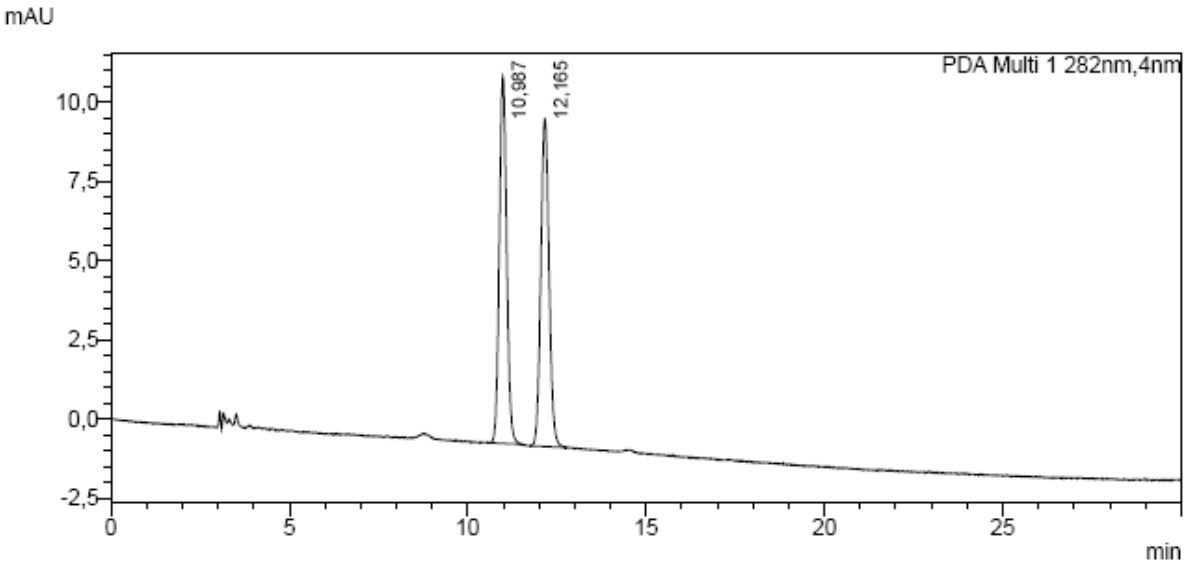
<Peak Table>

PDA Ch2 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	8,668	124316	11260	0,000			
2	10,945	1345720	92429	0,000			
3	13,015	7346666	449543	0,000			
4	15,277	2145366	103057	0,000			
Total		10962068	656289				

3-(5-methoxy-1H-indol-3-yl)-2-methylpropan-1-ol (3g)

Reference product:

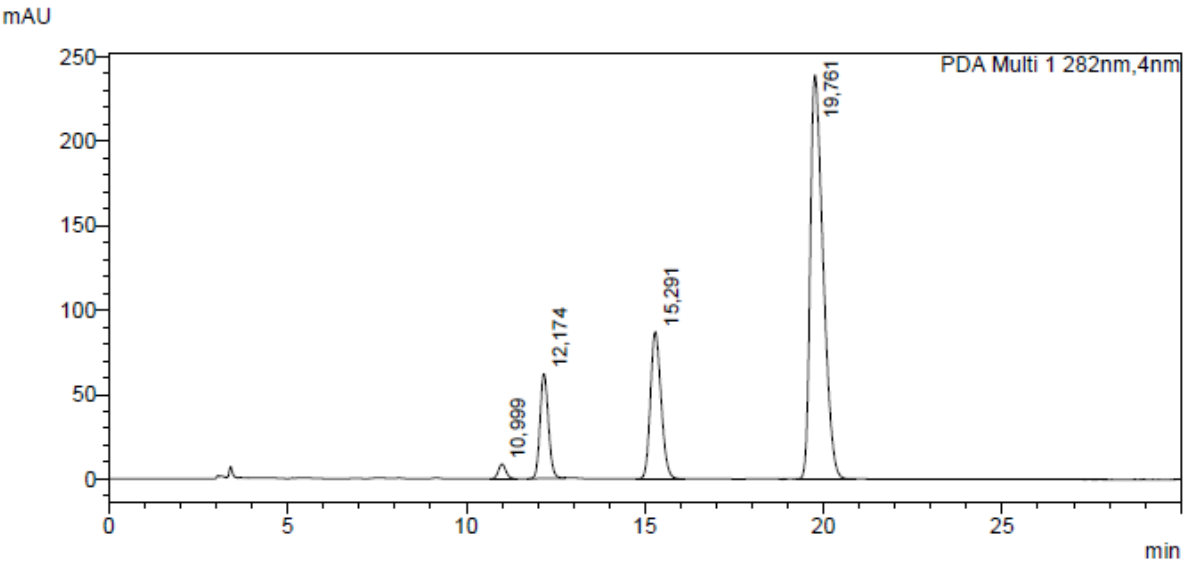


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	10,987	165688	11581	0,000			
2	12,165	166202	10334	0,000			
Total		331889	21915				

Catalysis sample: 10 µM LmrR_pAF pH 6. 6 hours reaction time at 4°. 15.3 min = 3-(3-hydroxypropyl)-indole (internal standard). 19.8 = 5-methoxy-indole.



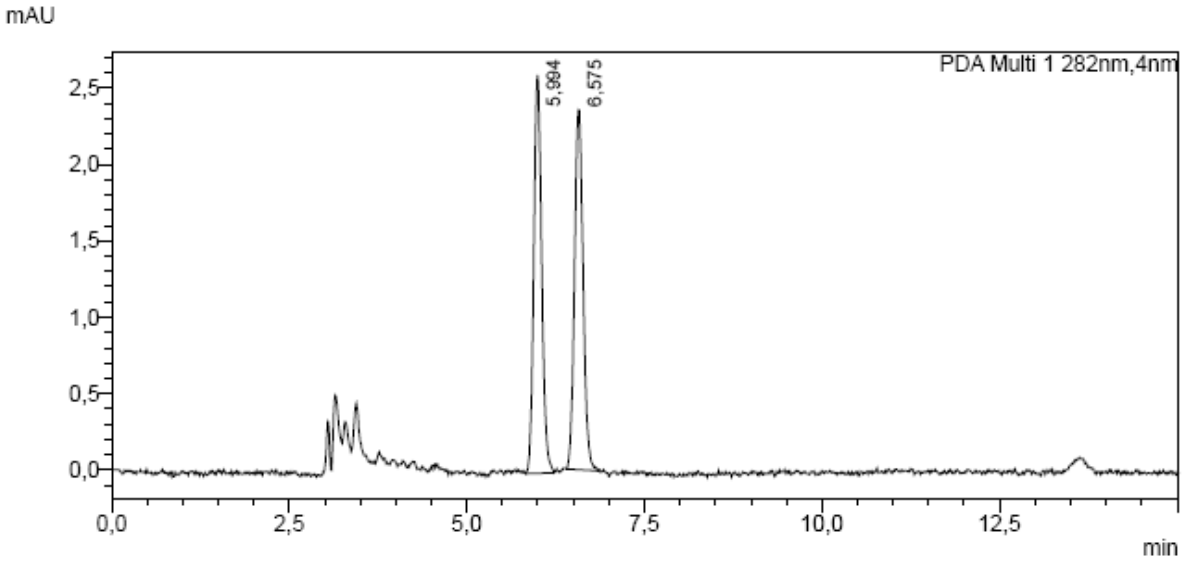
<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	10,999	126744	8633	0,000			
2	12,174	1007665	62018	0,000			
3	15,291	1812097	87076	0,000			
4	19,761	5996364	239017	0,000			
Total		8942870	396745				

3-(5-chloro-1H-indol-3-yl)-2-methylpropan-1-ol (3h)

Reference product:

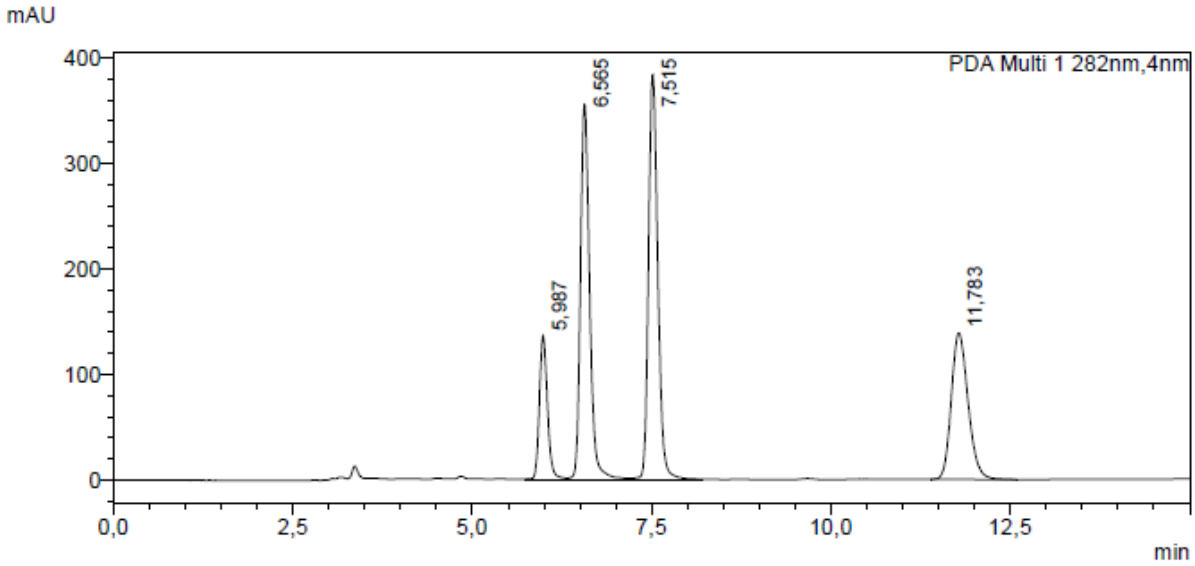


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	5,994	20099	2607	50,272			
2	6,575	19881	2357	49,728			
Total		39980	4964				

Catalysis sample: 25 µM LmrR_pAF pH 6. 16 hours reaction time at 4°. 7.5 min = 5-chloro-indole. 11.8 min = 3-(3-hydroxypropyl)-indole (internal standard).



<Peak Table>

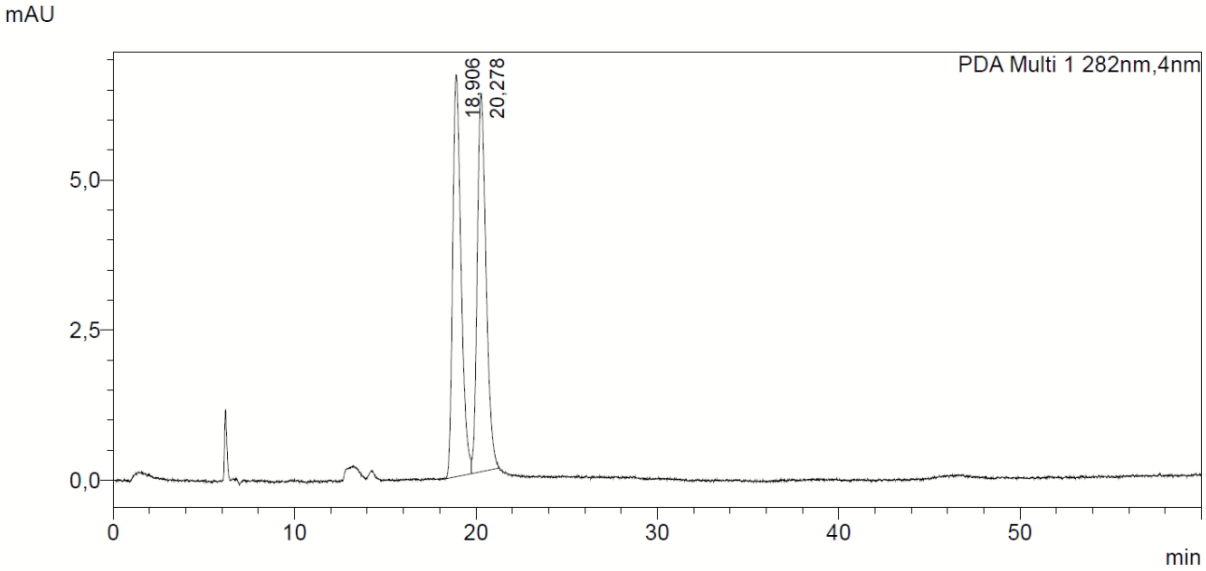
PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	5,987	1056534	135794	10,662			
2	6,565	3161184	355043	31,901		V	
3	7,515	3438740	383302	34,702		V	
4	11,783	2252776	138563	22,734			
Total		9909233	1012702				

methyl 3-(3-hydroxy-2-methylpropyl)-1H-indole-5-carboxylate (3i)

Reference product:

<Chromatogram>



<Peak Table>

PDA Ch1 282nm

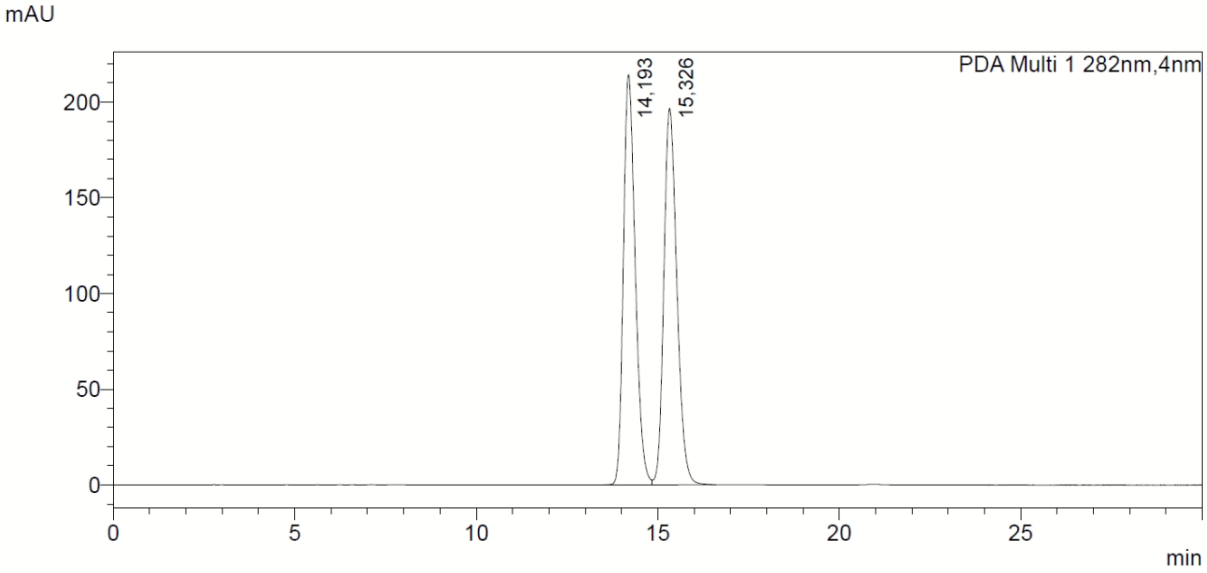
Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	18,906	215153	6694	50,257			
2	20,278	212956	6292	49,743		V	
Total		428108	12986				

No product peaks found in the catalysis sample.

3-(5-methoxy-2-methyl-1H-indol-3-yl)-2-methylpropan-1-ol (3j)

Reference product:

<Chromatogram>

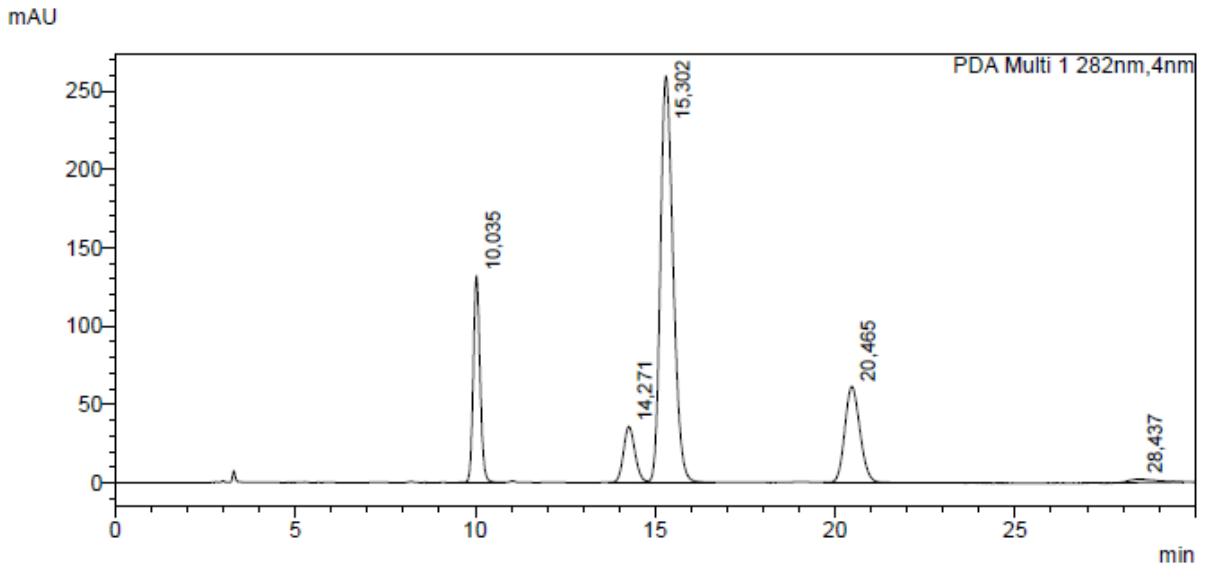


<Peak Table>

PDA Ch1 282nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	14,193	4865836	214184	49,894			
2	15,326	4886432	196494	50,106		V	
Total		9752267	410677				

Catalysis sample: 10 µM LmrR_pAF pH 6. 6 hours reaction time at 4°. 10.0 min = 5-methoxy-2-methyl-indole. 20.5 min = 3-(3-hydroxypropyl)-indole (internal standard).



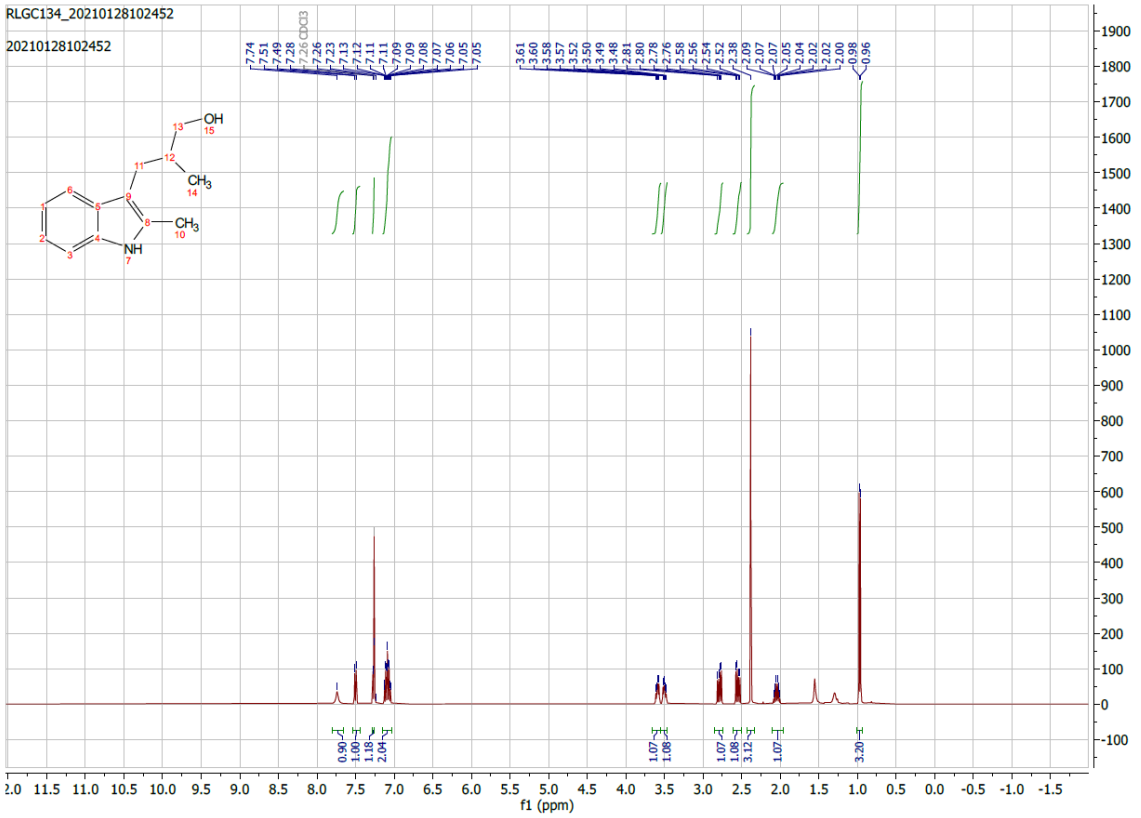
<Peak Table>

PDA Ch1 282nm

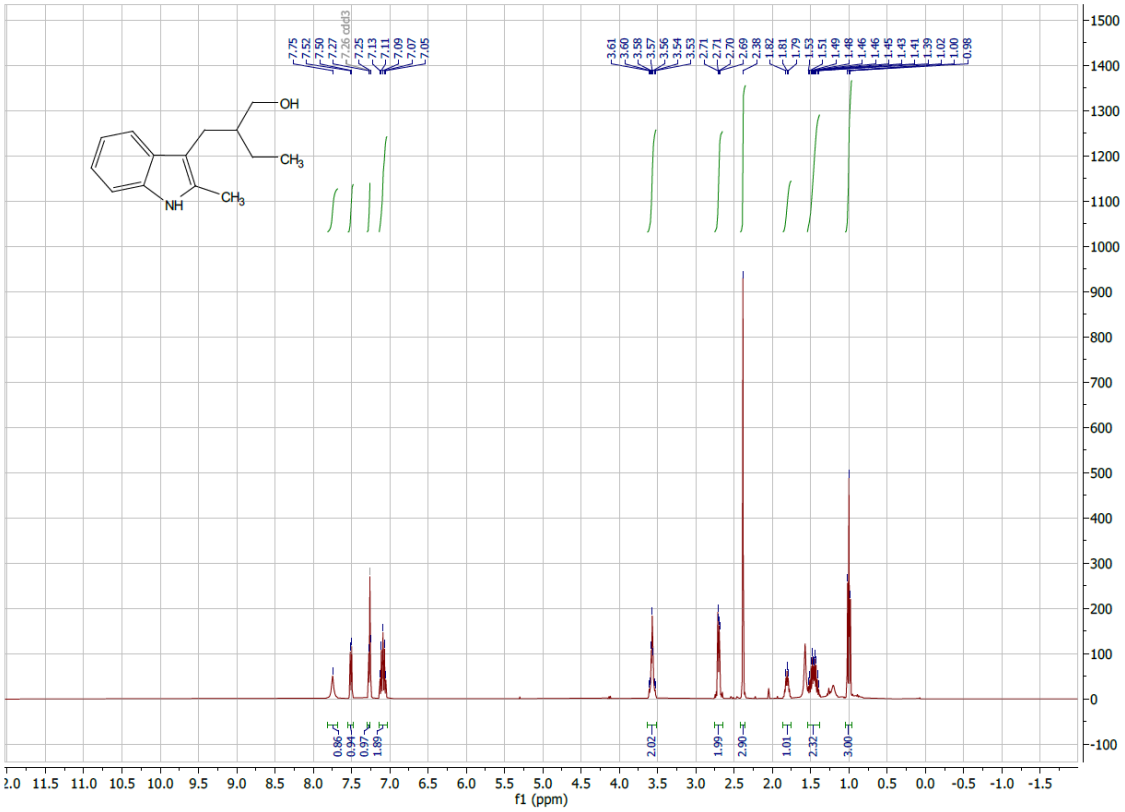
Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	10,035	1737204	131556	0,000			
2	14,271	796343	35656	0,000			
3	15,302	6394485	259084	0,000		V	
4	20,465	1796785	61284	0,000		V	
5	28,437	114997	2208	0,000			
Total		10839813	489786				

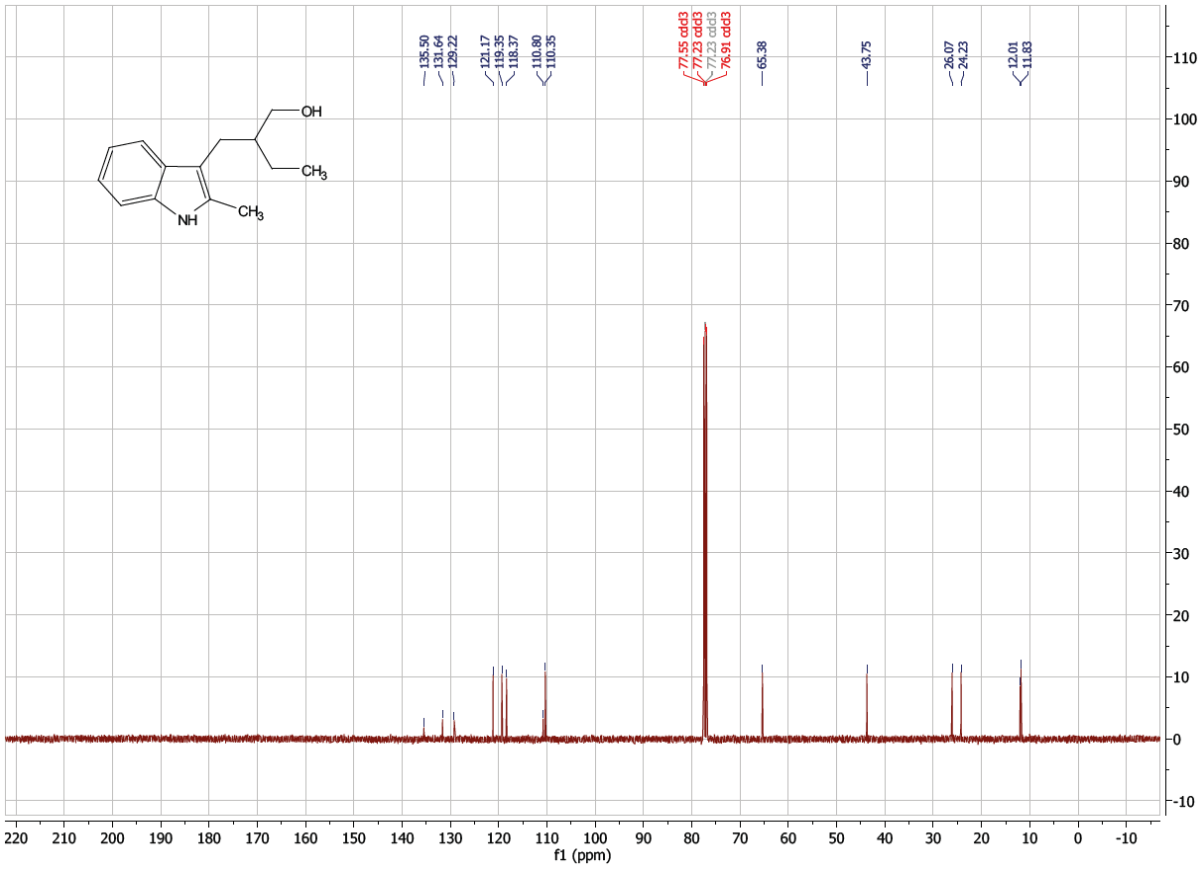
8. NMR Spectra

2-methyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3a)

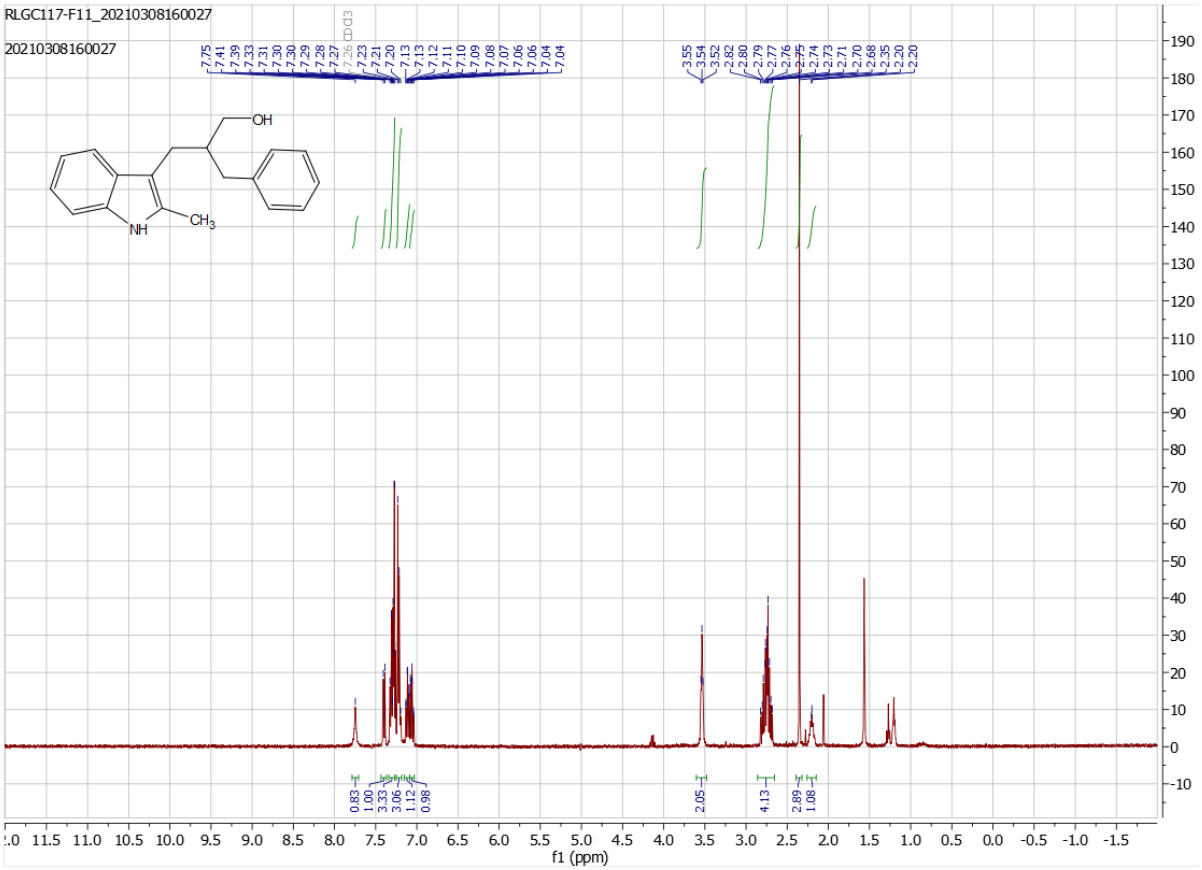


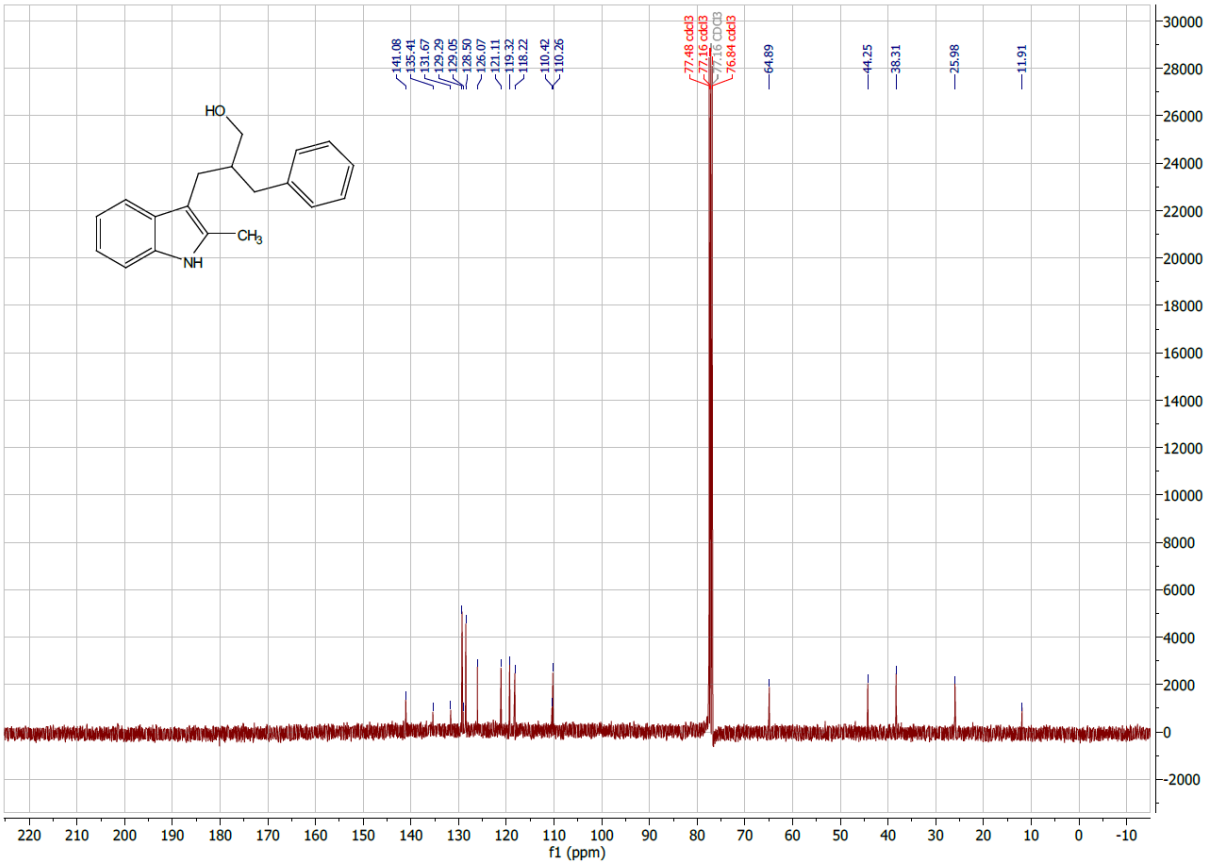
2-((2-methyl-1H-indol-3-yl)methyl)butan-1-ol (3b)



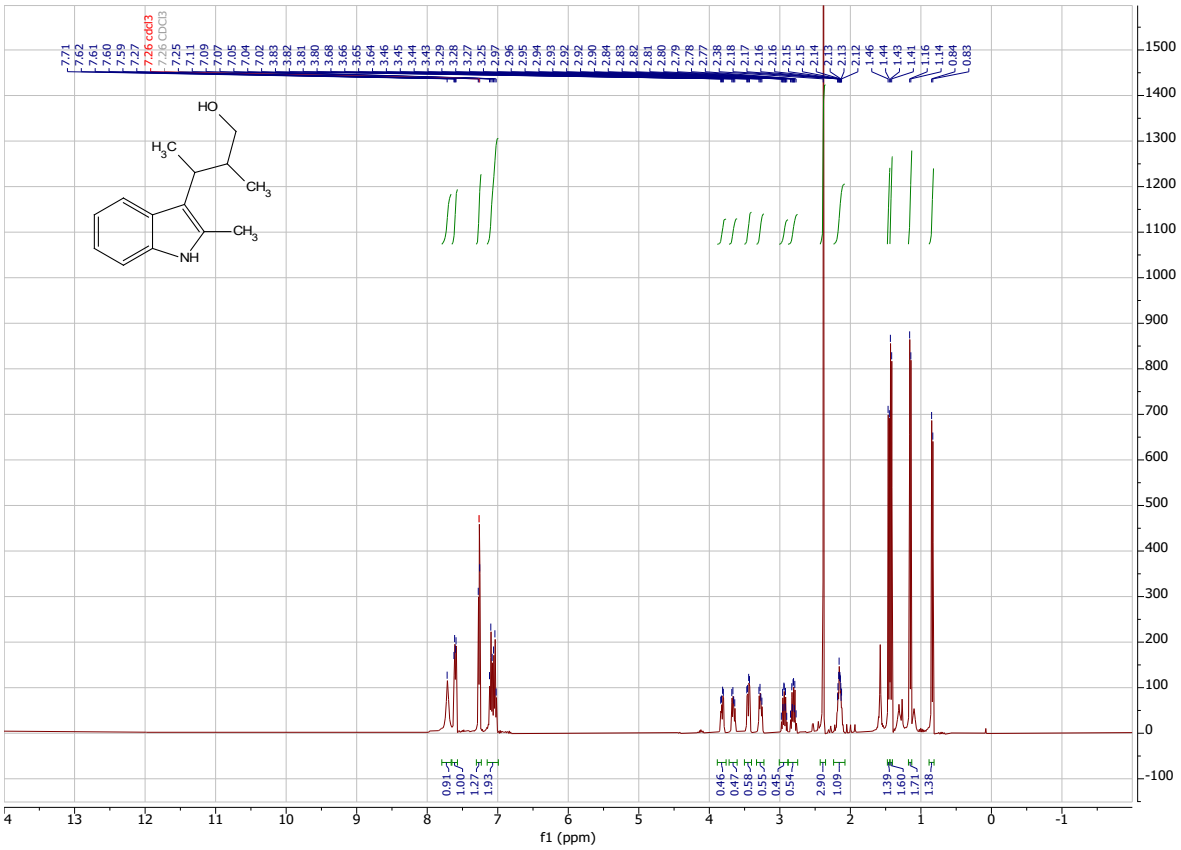


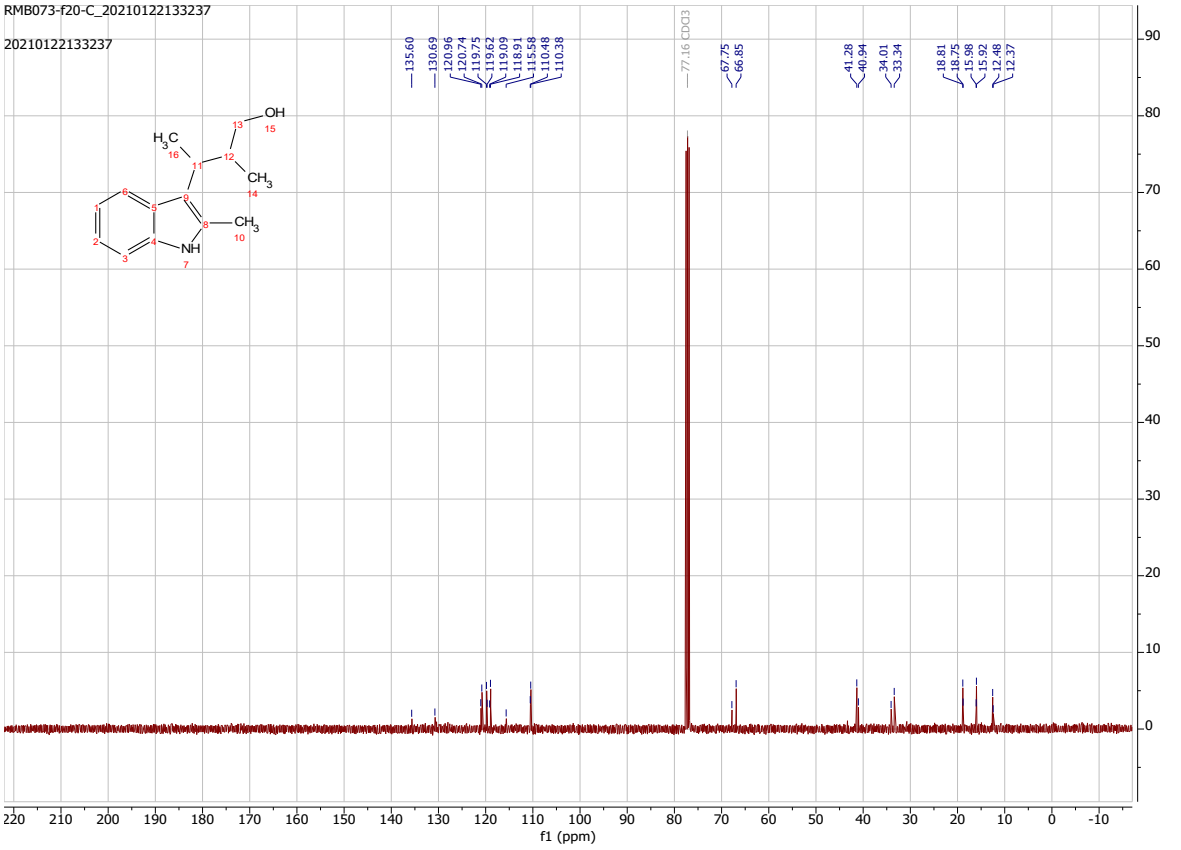
2-benzyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3c)



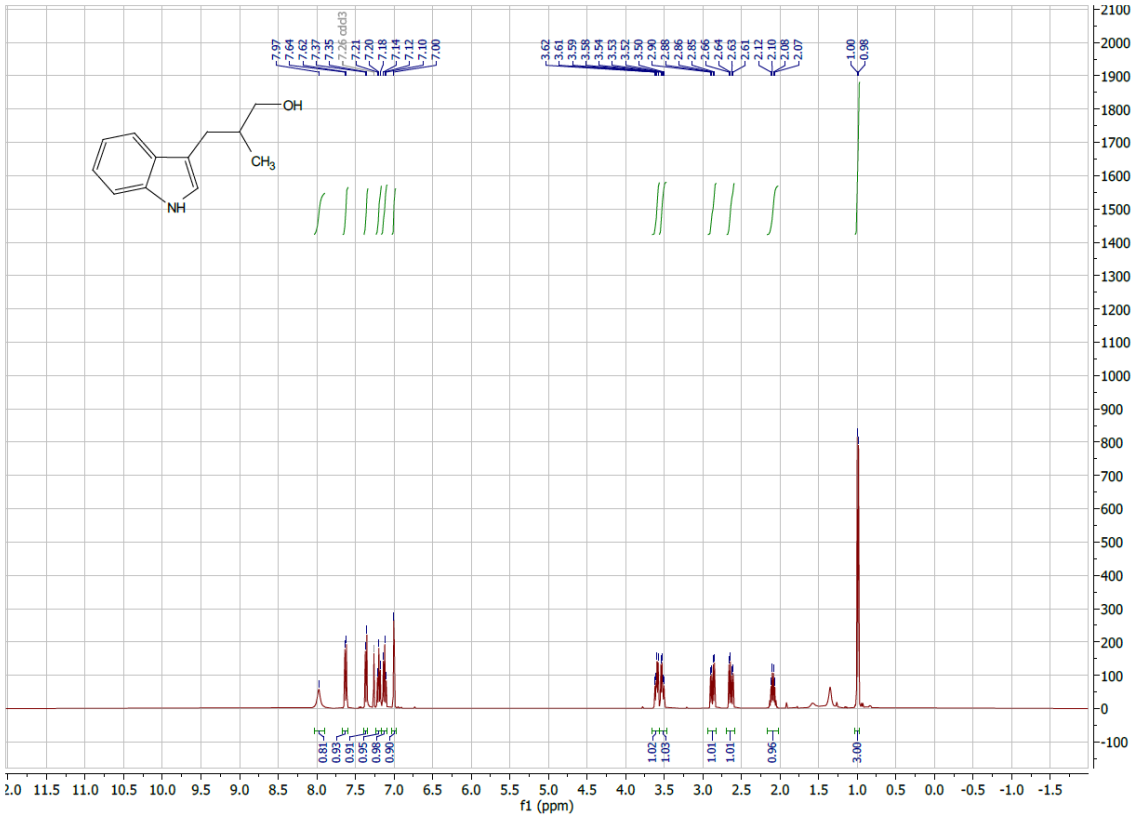


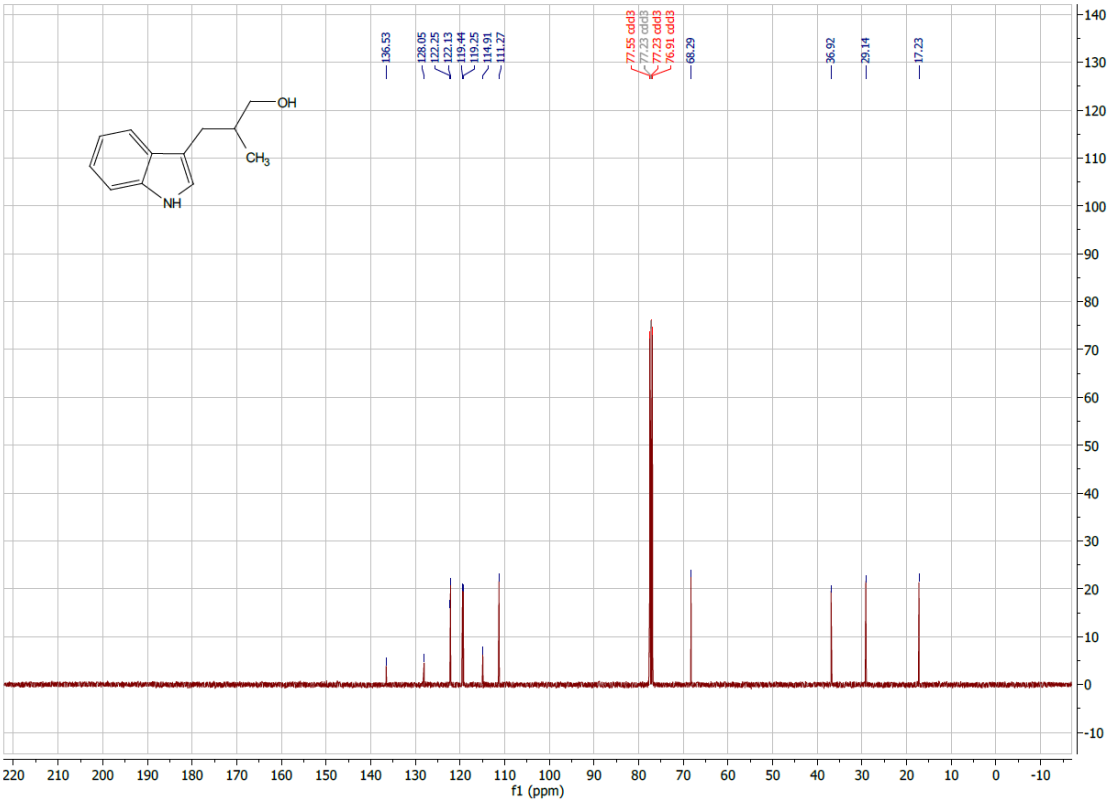
2-methyl-3-(2-methyl-1H-indol-3-yl)butan-1-ol (3d)



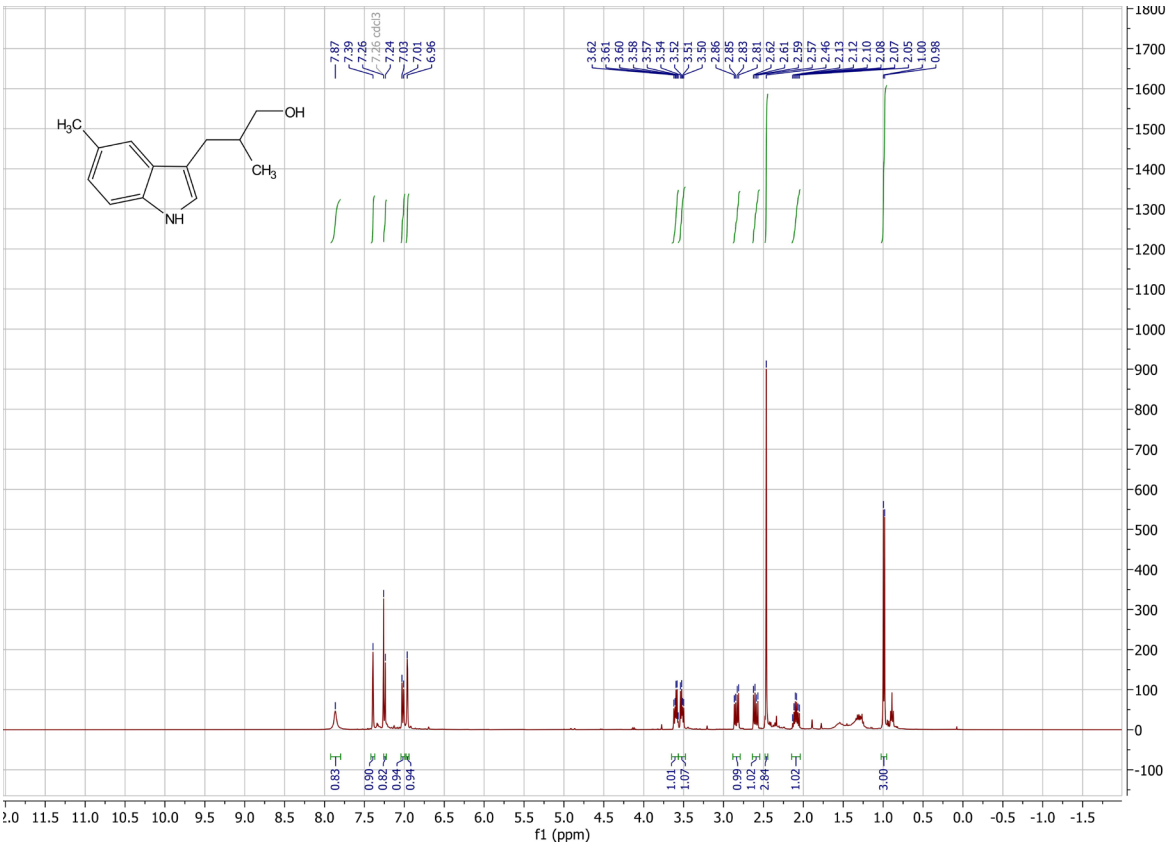


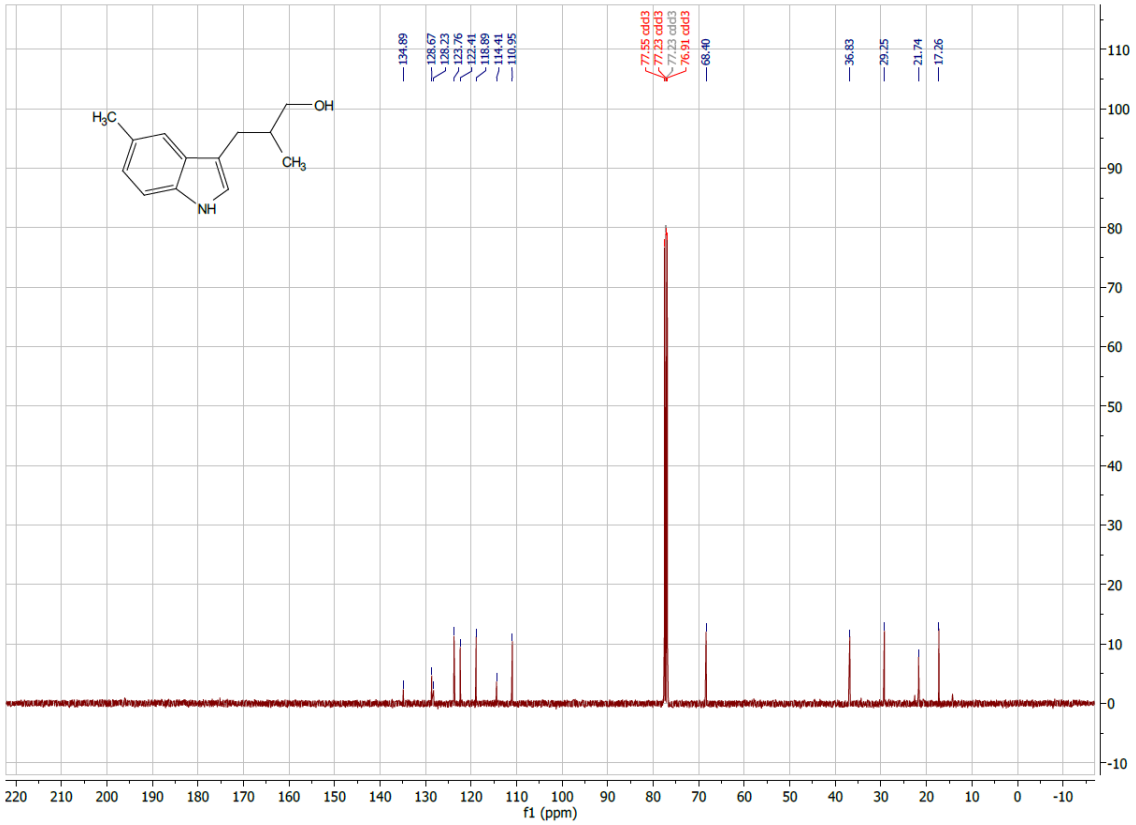
3-(1H-indol-3-yl)-2-methylpropan-1-ol (3e)



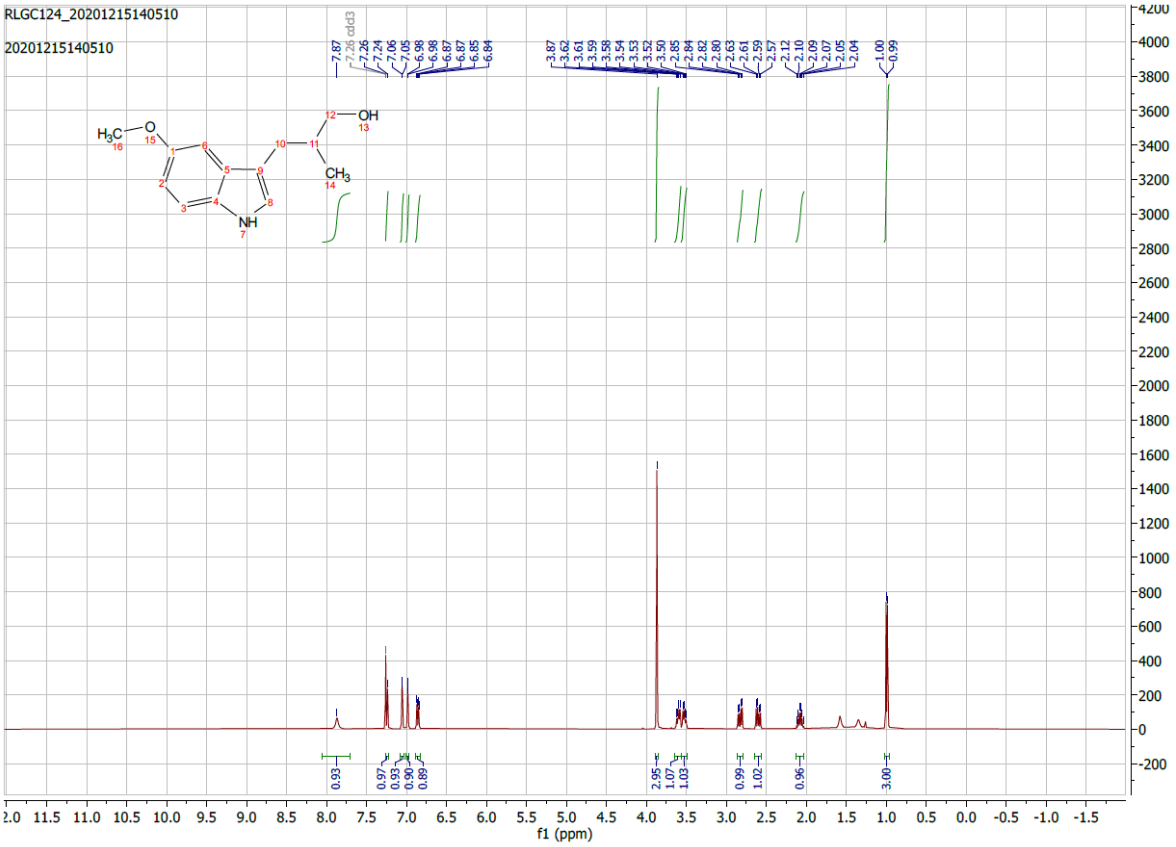


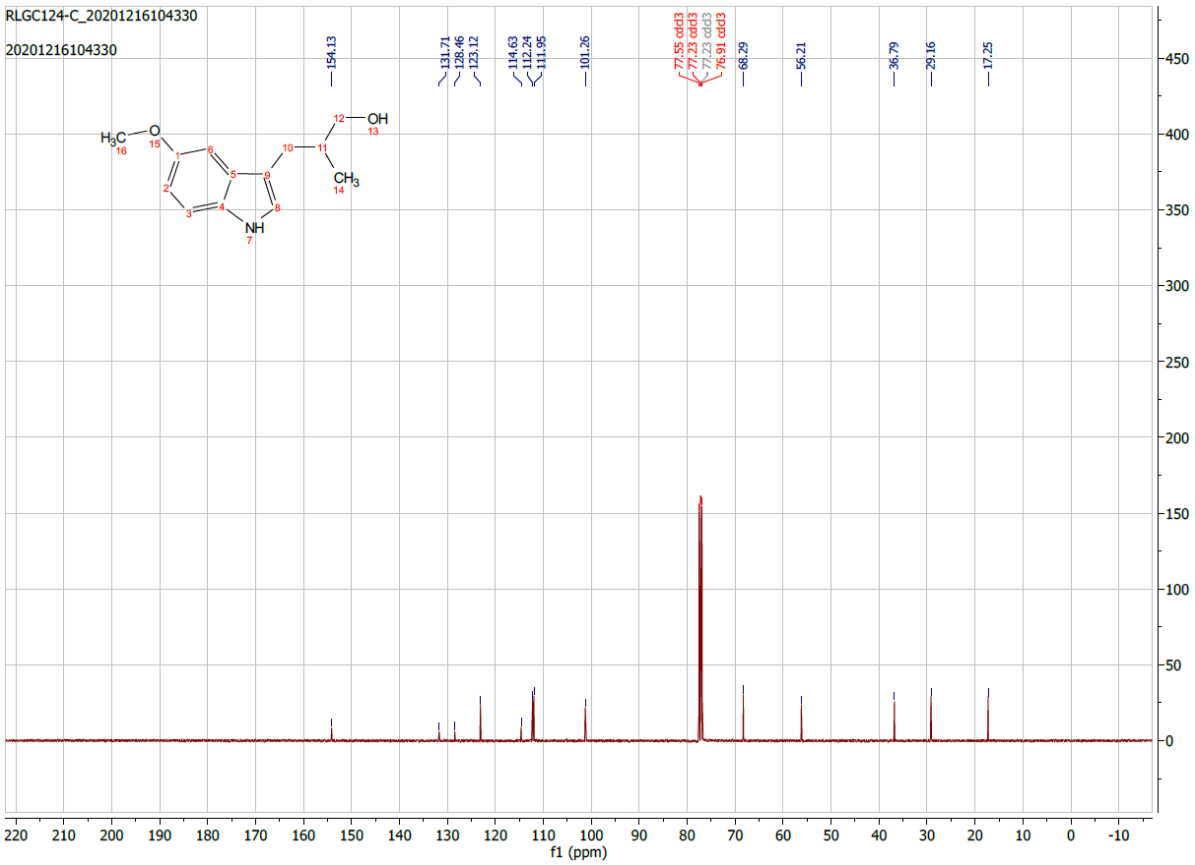
2-methyl-3-(5-methyl-1H-indol-3-yl)propan-1-ol (3f)



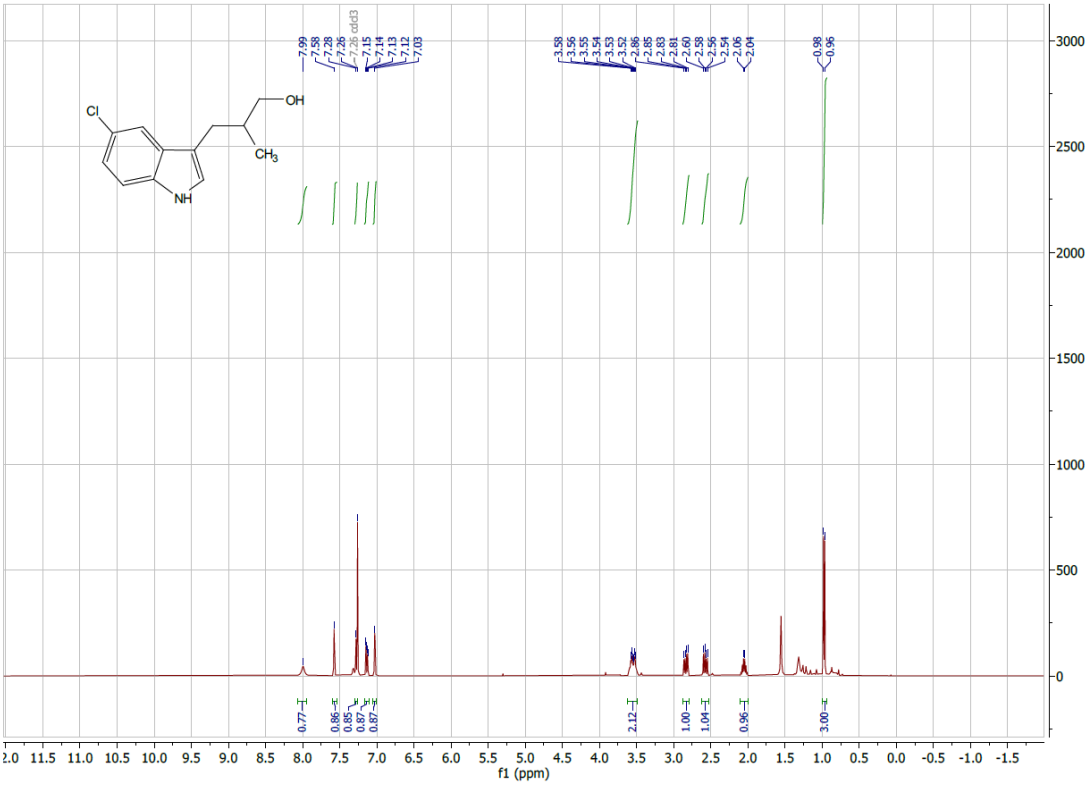


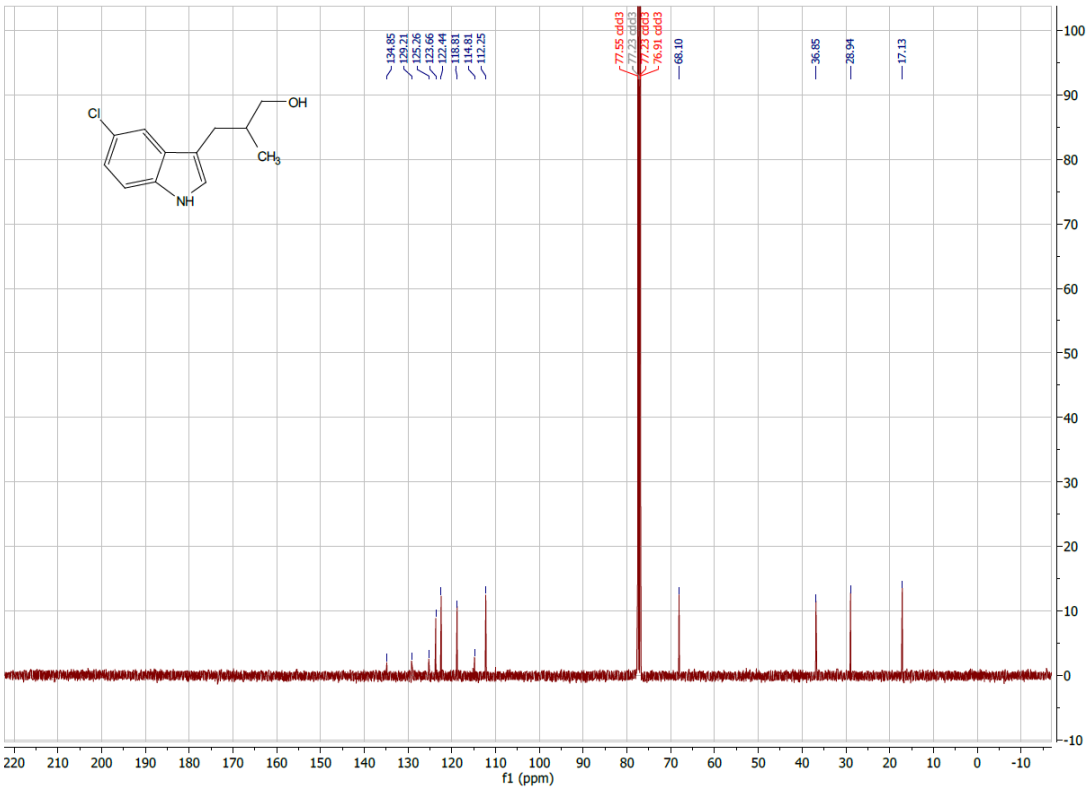
3-(5-methoxy-1H-indol-3-yl)-2-methylpropan-1-ol (3g)



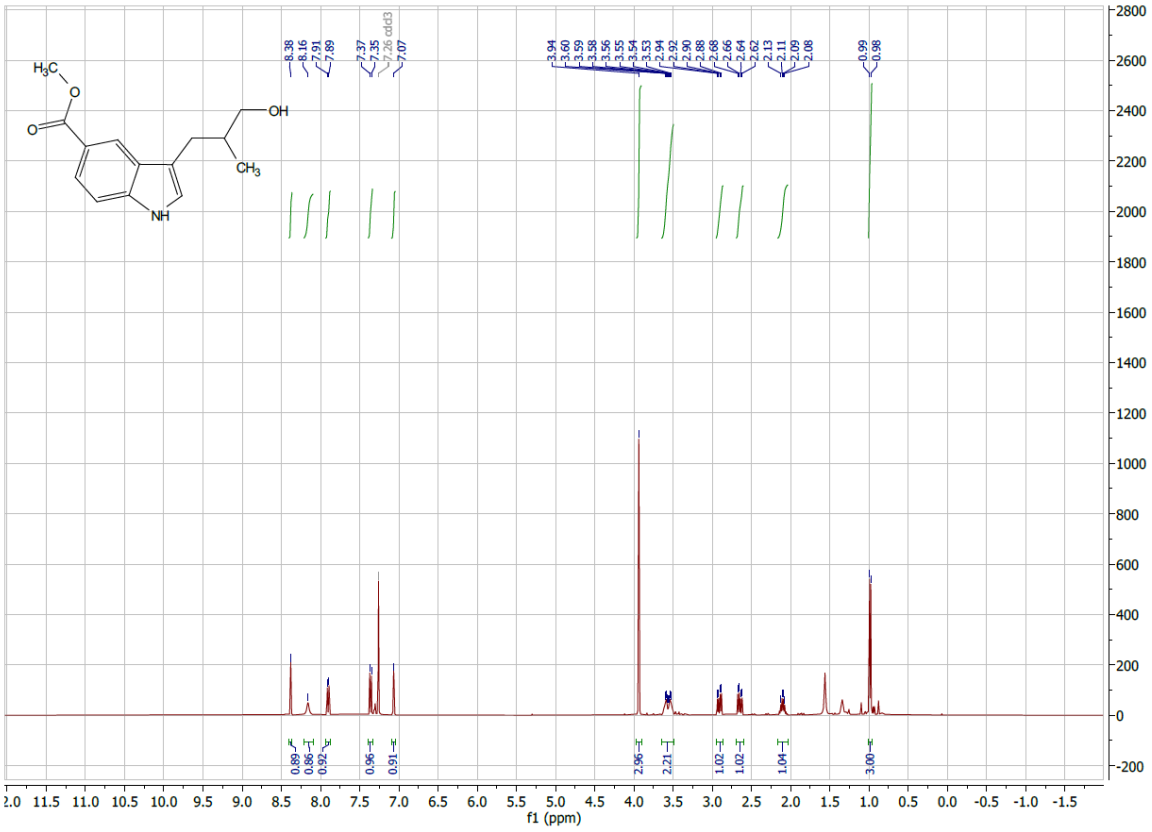


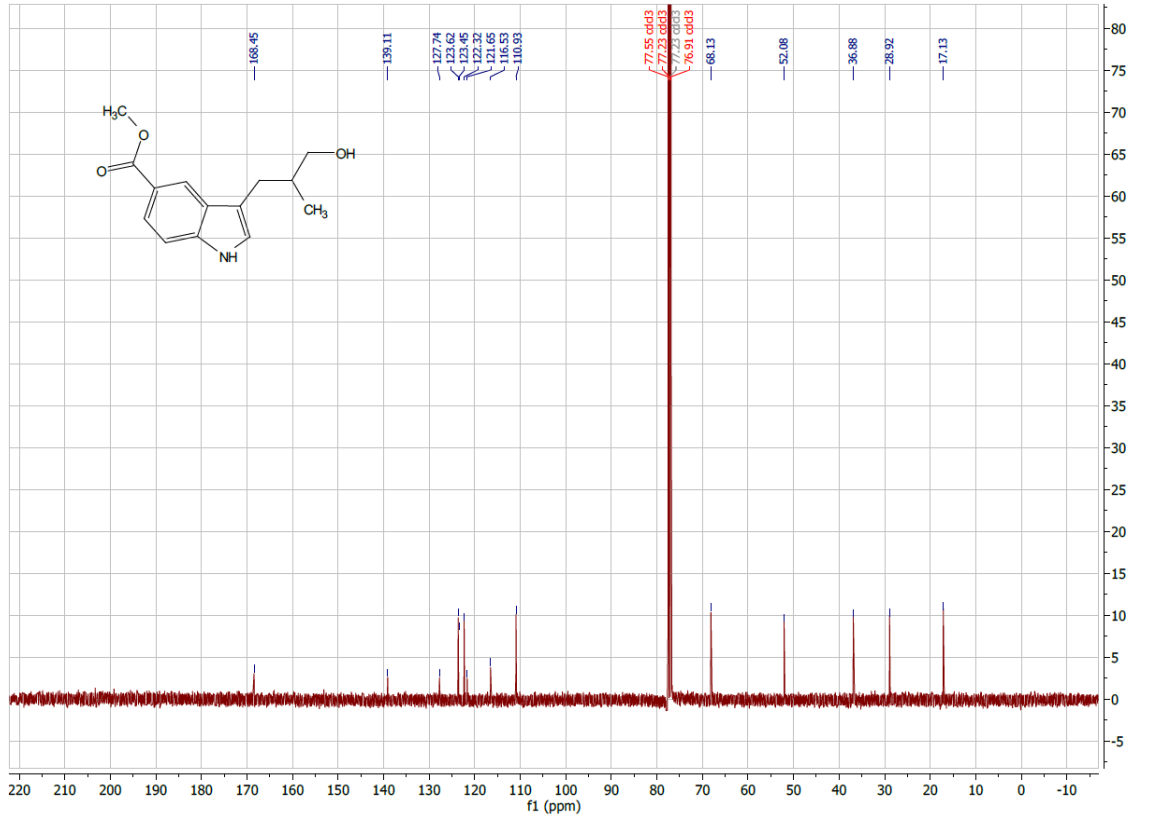
3-(5-chloro-1H-indol-3-yl)-2-methylpropan-1-ol (3h)



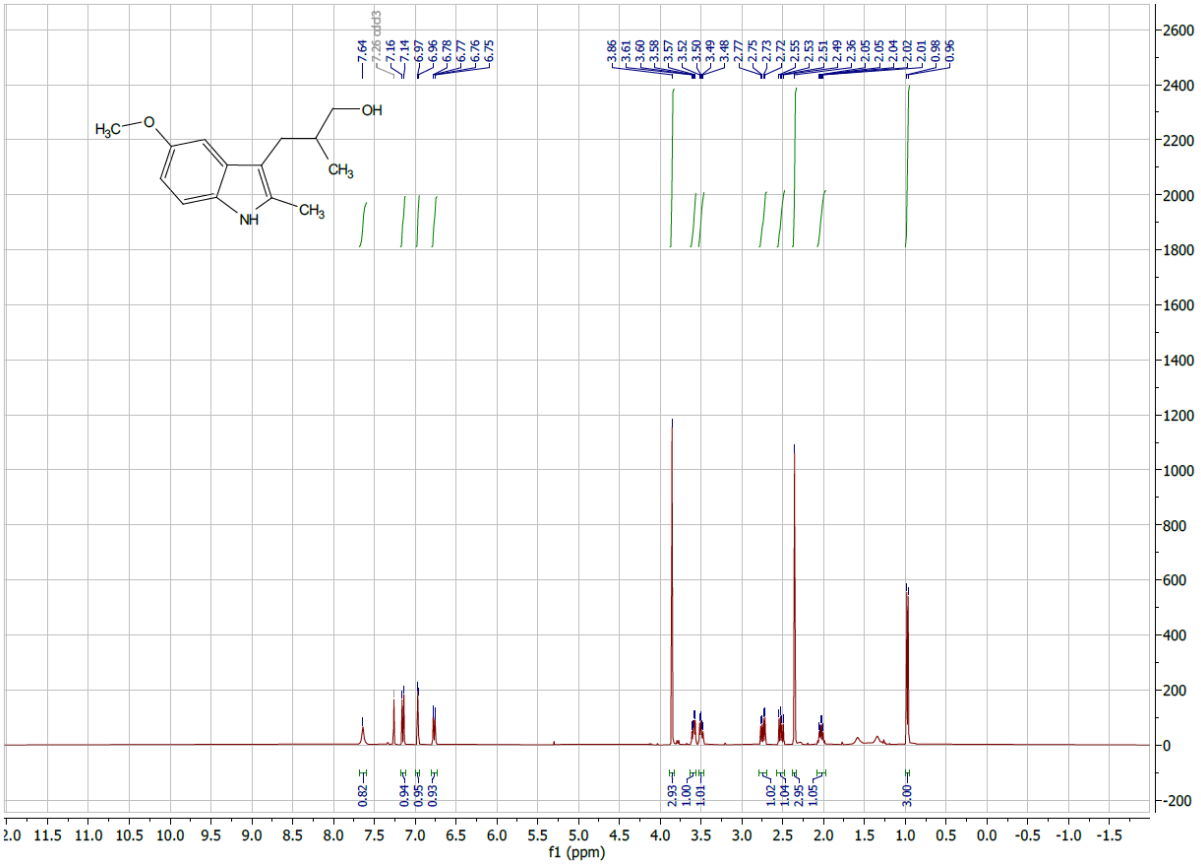


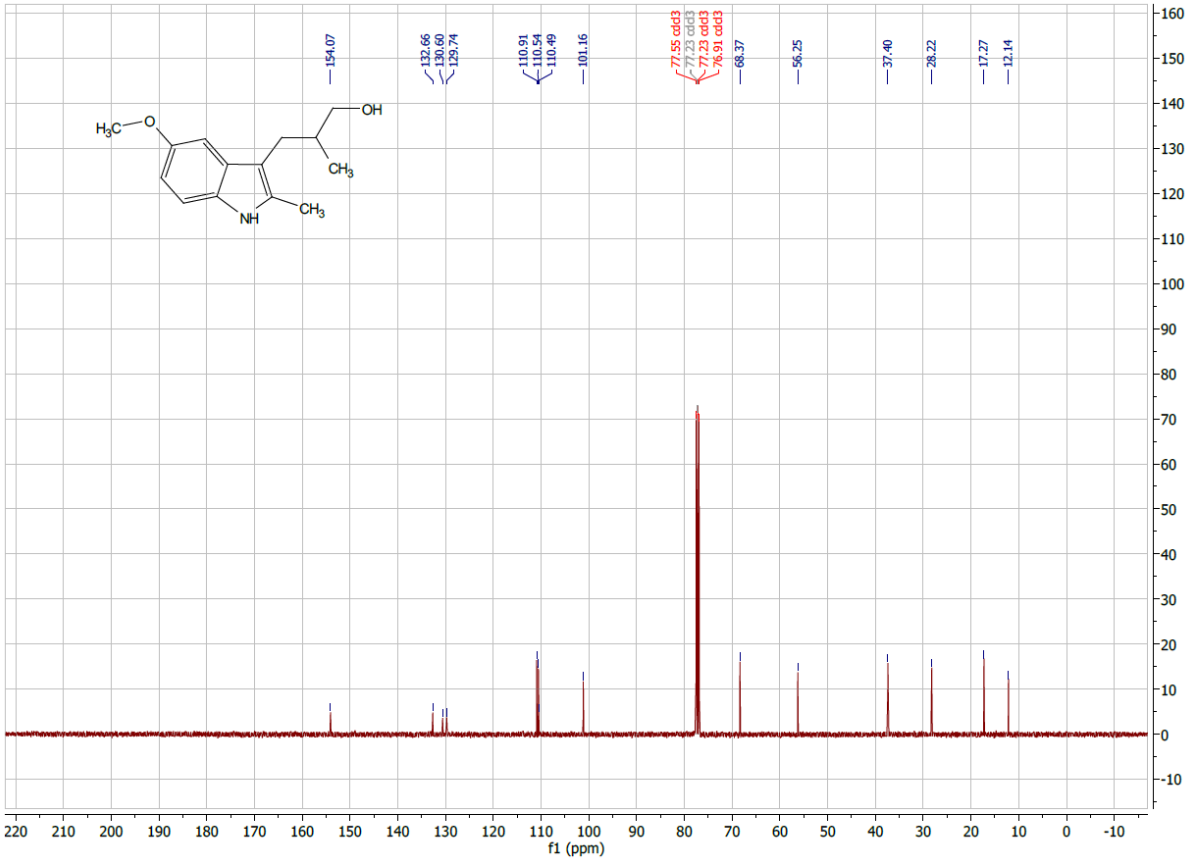
methyl 3-(3-hydroxy-2-methylpropyl)-1H-indole-5-carboxylate (3i)





3-(5-methoxy-2-methyl-1H-indol-3-yl)-2-methylpropan-1-ol (3j)

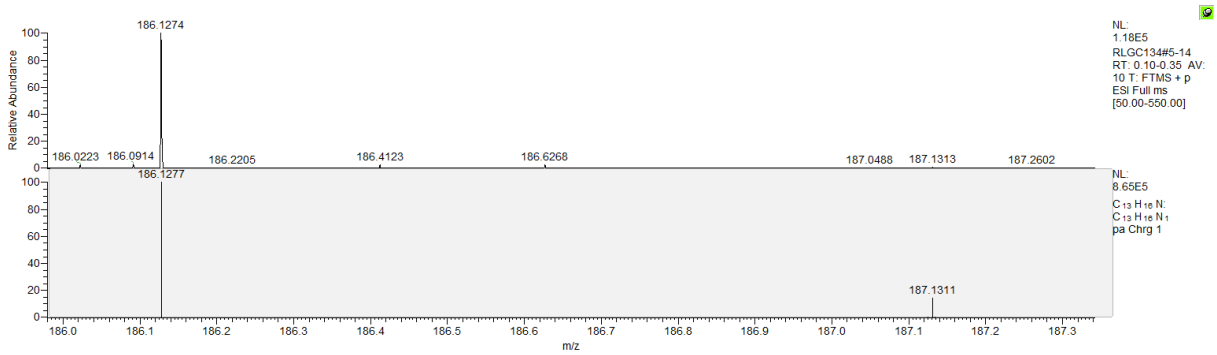




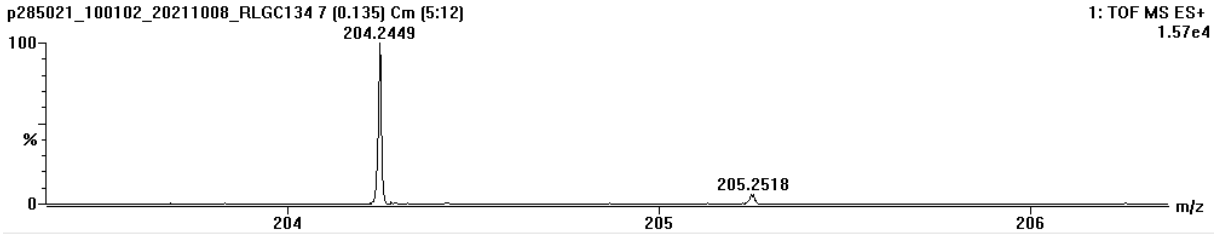
9. Mass Spectra of Reference Compounds

2-methyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3a)

HRMS ([M-OH]$^+$)

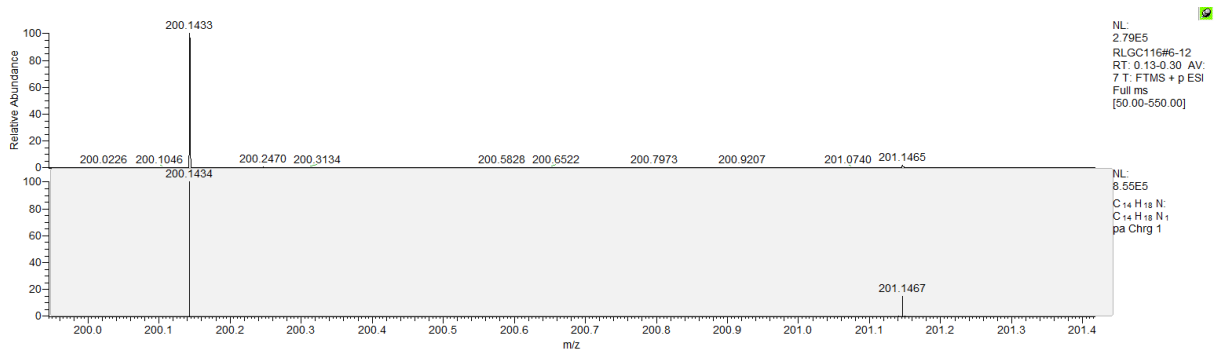


LRMS ([M+H]$^+$)



2-((2-methyl-1H-indol-3-yl)methyl)butan-1-ol (3b)

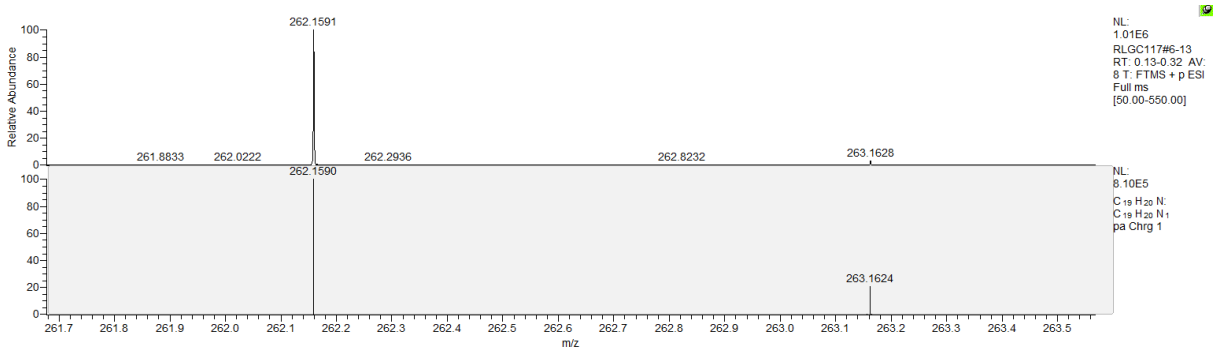
HRMS ([M-OH]$^+$)



LRMS ([M+H]$^+$)

2-benzyl-3-(2-methyl-1H-indol-3-yl)propan-1-ol (3c)

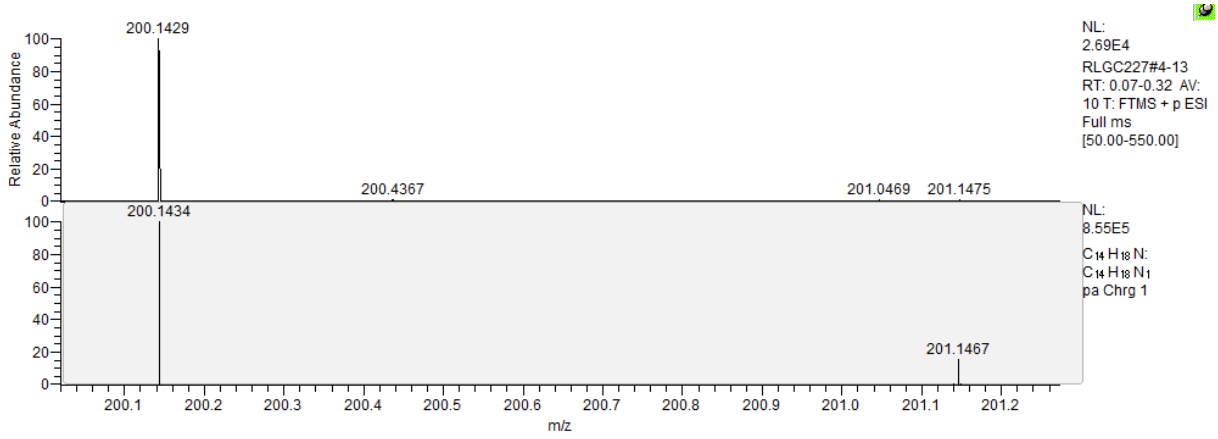
HRMS ([M-OH]$^+$)



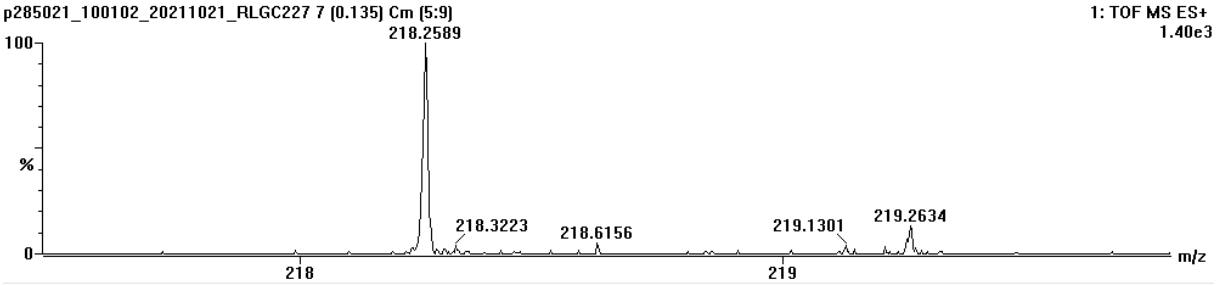
LRMS ([M+H]$^+$)

2-methyl-3-(2-methyl-1H-indol-3-yl)butan-1-ol (3d)

HMRS ([M-OH]$^+$)

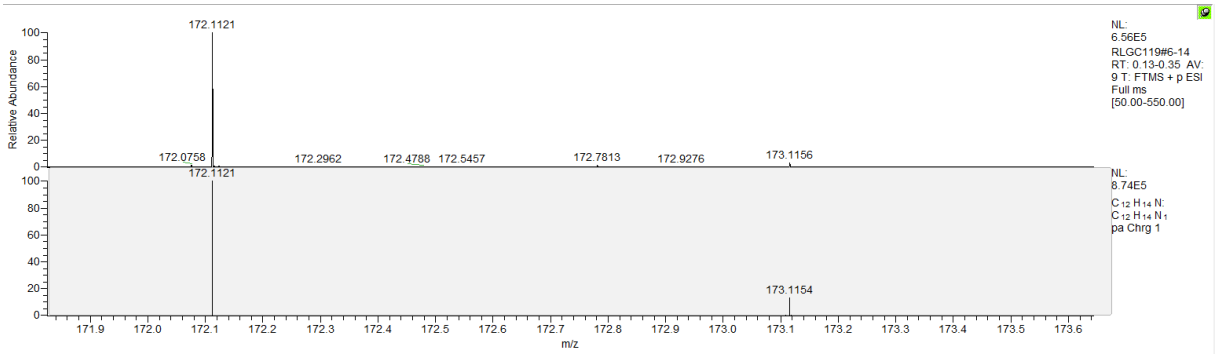


LRMS ([M+H]$^+$)

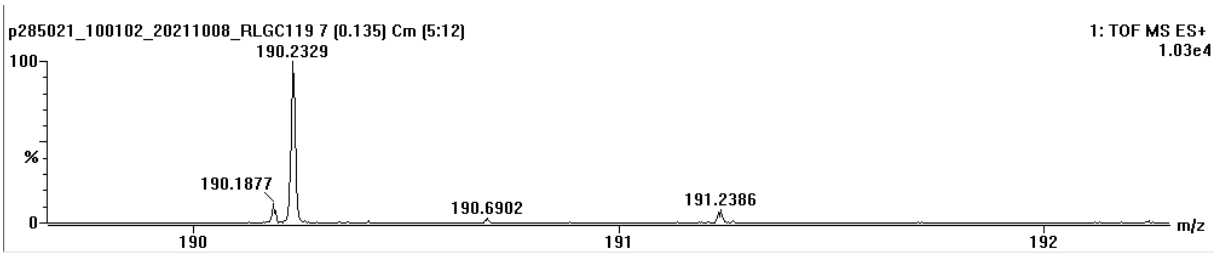


3-(1H-indol-3-yl)-2-methylpropan-1-ol (3e)

HRMS ([M-OH]$^+$)

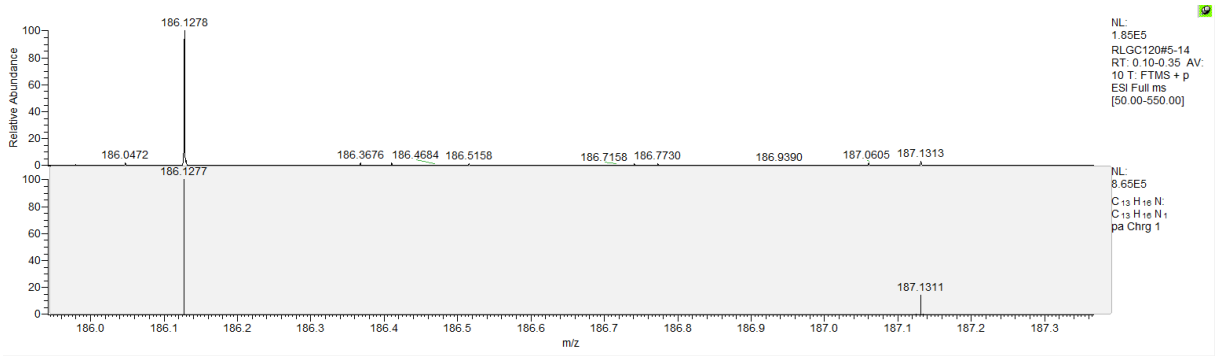


LRMS ([M+H]$^+$)

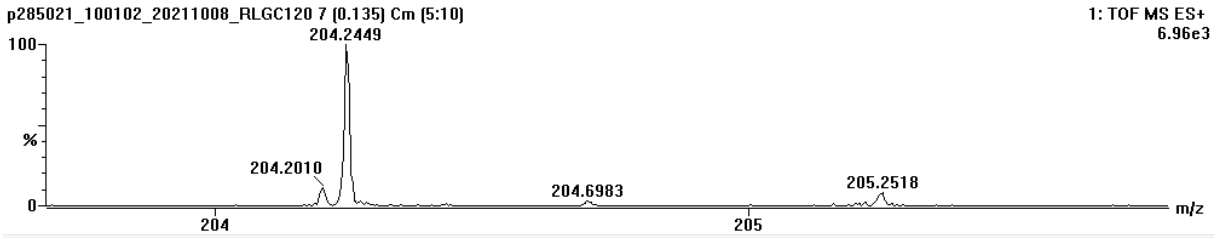


2-methyl-3-(5-methyl-1H-indol-3-yl)propan-1-ol (3f)

HRMS ([M-OH]$^+$)

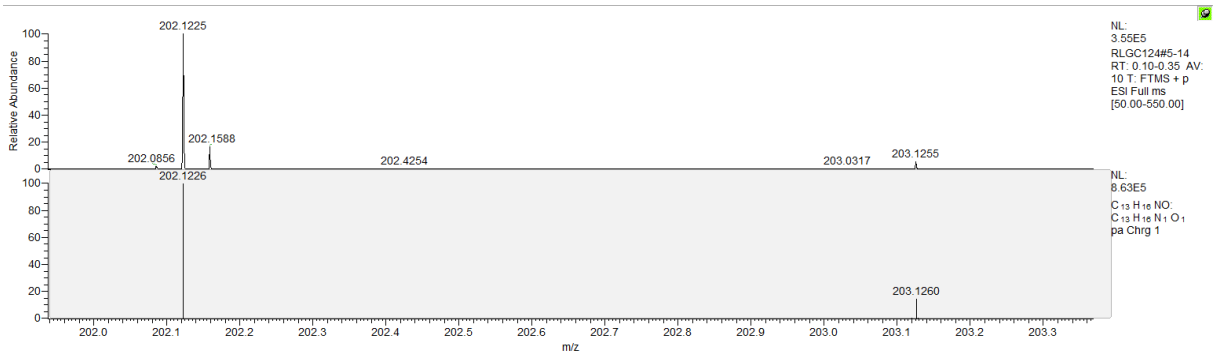


LRMS ([M+H]$^+$)

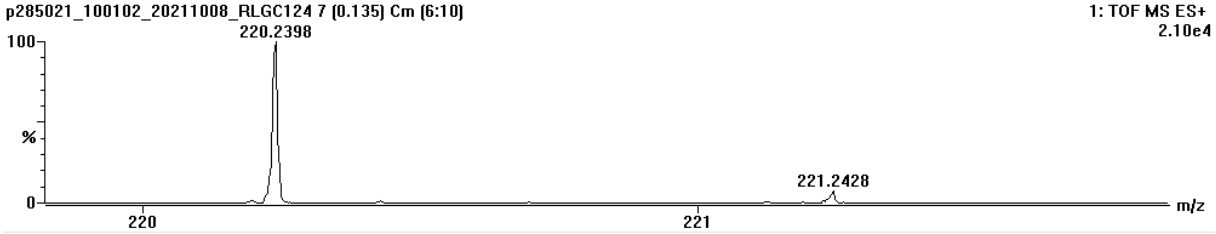


3-(5-methoxy-1H-indol-3-yl)-2-methylpropan-1-ol (3g)

HRMS ([M-OH]$^+$)

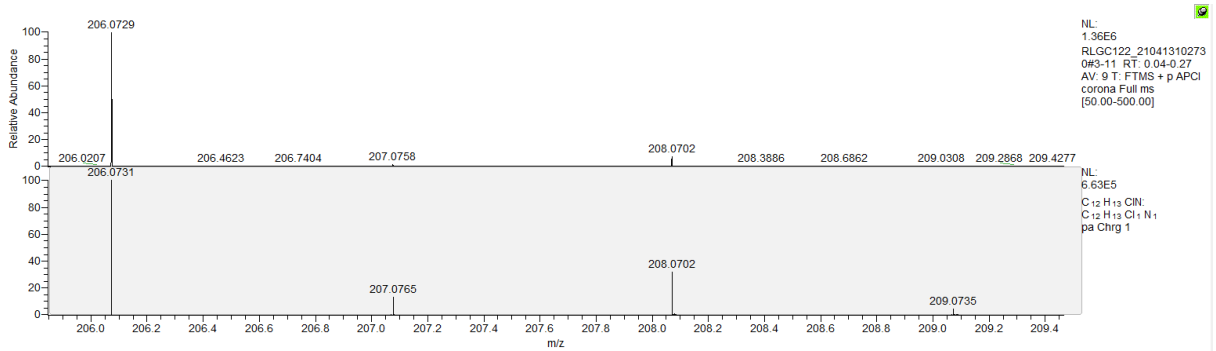


LRMS ([M+H]$^+$)

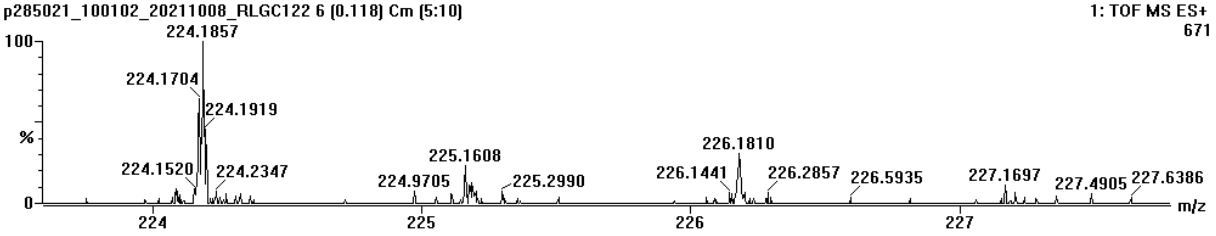


3-(5-chloro-1H-indol-3-yl)-2-methylpropan-1-ol (3h)

HRMS ([M-OH]$^+$)

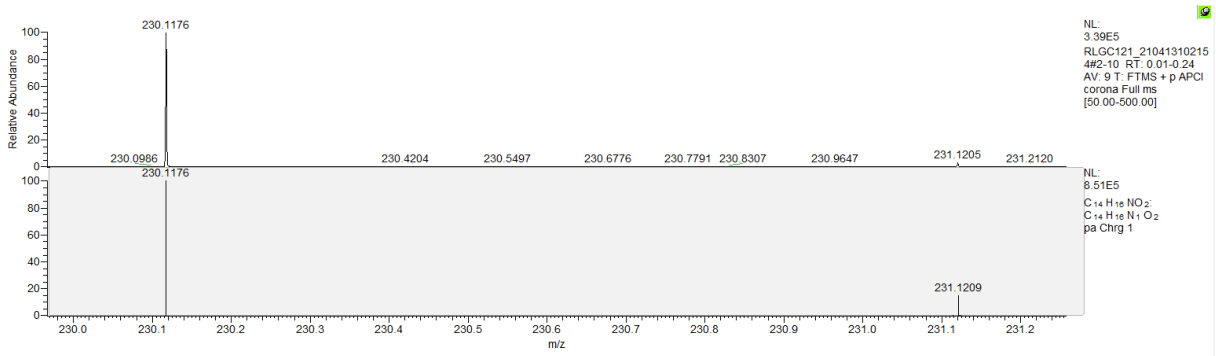


LRMS ([M+H]$^+$)

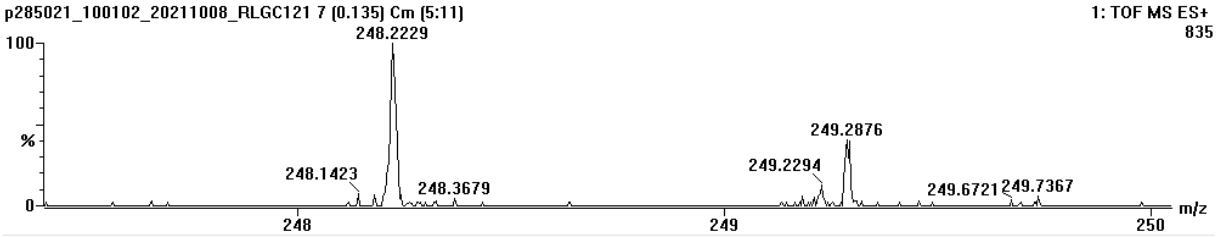


methyl 3-(3-hydroxy-2-methylpropyl)-1H-indole-5-carboxylate (3i)

HRMS ([M-OH]$^+$)

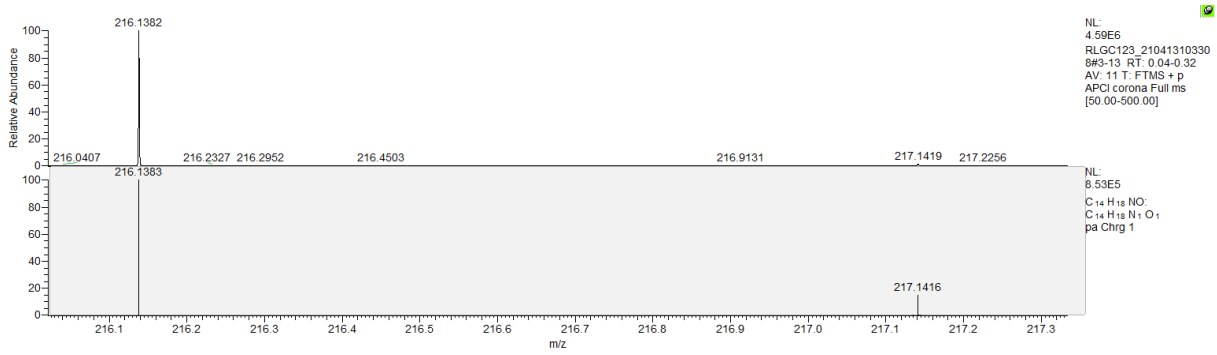


LRMS ([M+H]$^+$)

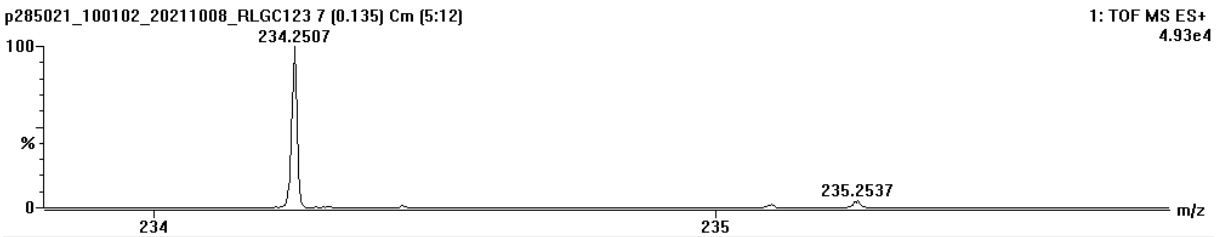


3-(5-methoxy-2-methyl-1H-indol-3-yl)-2-methylpropan-1-ol (3j)

HRMS ([M-OH]$^+$)



LRMS ([M+H]$^+$)



References

(1) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of P-Azido-l-Phenylalanine to the Genetic Code of Escherichia Coli. *J. Am. Chem. Soc.* **2002**, *124*, 9026–9027.

(2) Amiram, M.; Haimovich, A. D.; Fan, C.; Wang, Y.-S.; Aerni, H.-R.; Ntai, I.; Moonan, D. W.; Ma, N. J.; Rovner, A. J.; Hong, S. H.; Kelleher, N. L.; Goodman, A. L.; Jewett, M. C.; Söll, D.; Rinehart, J.; Isaacs, F. J. Evolution of Translation Machinery in Recoded Bacteria Enables Multi-Site Incorporation of Nonstandard Amino Acids. *Nat Biotechnol* **2015**, *33*, 1272–1279.

(3) Drienovská, I.; Mayer, C.; Dulson, C.; Roelfes, G. A Designer Enzyme for Hydrazone and Oxime Formation Featuring an Unnatural Catalytic Aniline Residue. *Nat. Chem.* **2018**, *10*, 946–952.

(4) Leveson-Gower, R. B.; Zhou, Z.; Drienovská, I.; Roelfes, G. Unlocking Iminium Catalysis in Artificial Enzymes to Create a Friedel–Crafts Alkylase. *ACS Catal.* **2021**, *11*, 6763–6770.

(5) Zhou, Z.; Roelfes, G. Synergistic Catalysis in an Artificial Enzyme by Simultaneous Action of Two Abiological Catalytic Sites. *Nat. Catal.* **2020**, *3*, 289–294.

(6) Mayer, C.; Dulson, C.; Reddem, E.; Thunnissen, A.-M. W. H.; Roelfes, G. Directed Evolution of a Designer Enzyme Featuring an Unnatural Catalytic Amino Acid. *Angew. Chem. Int. Ed.* **2019**, *58*, 2083–2087.

(7) Fu, N.; Zhang, L.; Li, J.; Luo, S.; Cheng, J. P. Chiral Primary Amine Catalyzed Enantioselective Protonation via an Enamine Intermediate. *Angew. Chem. Int. Ed.* **2011**, *50*, 11451–11455.