Enzyme screening and engineering for N- and O-demethylation: key steps in the synthesis of buprenorphine

Carvalho, A. T. P., Dourado, D. F. A. R., Spratt, J., Caswell, J. M., Skvortsov, T., Quinn, D. J., Carey, J. S., & Moody, T. S. (2024). Enzyme screening and engineering for N- and O-demethylation: key steps in the synthesis of buprenorphine. *Organic Process Research & Development*. Advance online publication. https://doi.org/10.1021/acs.oprd.3c00417

Published in:
Organic Process Research & Development

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Enzyme Screening and Engineering for N- and O-Demethylation: Key Steps in the Synthesis of Buprenorphine

Alexandra T.P. Carvalho[1], Daniel F.A.R. Dourado[1], Jenny Spratt[1], Jill Caswell[1], Timofey Skvortsov[1†], Derek J. Quinn[1], John S. Carey[2], Thomas S. Moody[1,3*]*

1 Almac Sciences, Department of Biocatalysis and Isotope Chemistry, Almac House, 20 Seagoe Industrial Estate, Craigavon, BT63 5QD, Northern Ireland, UK.

2Indivior UK Limited, Henry Boot Way, Hull, HU4 7DY, United Kingdom.

3Arran Chemical Company, Unit 1 Monksland Industrial Estate, Athlone, Co. Roscommon, Ireland.

Keywords: Benzylisoquinoline alkaloids, buprenorphine, rational design, directed evolution, demethylation, oxygenase enzymes.

Abstract

Benzylisoquinoline alkaloids are valuable active ingredients in medicines that are typically extracted from plants and subsequently derivatized. Buprenorphine is a member of this family of compounds and is an effective analgesic and is also used for the treatment of opioid use disorder.

The commercial route of synthesis for buprenorphine starts from thebaine and uses toxic reagents and harsh reaction conditions for the N- and O-demethylation steps. Here we propose an alternative approach for buprenorphine synthesis via enzymatic N- and O-demethylation reactions. Utilizing rational enzyme design and directed evolution, two oxygenase enzymes were identified and engineered. For the N-demethylation reaction, the best variant achieved a cumulative improvement in conversion of 567-fold, while for the O-demethylation the best variant achieved 22-fold cumulative improvement in conversion. A separate variant was able to efficiently catalyse both the N-demethylation and the O-demethylation reactions.

Introduction

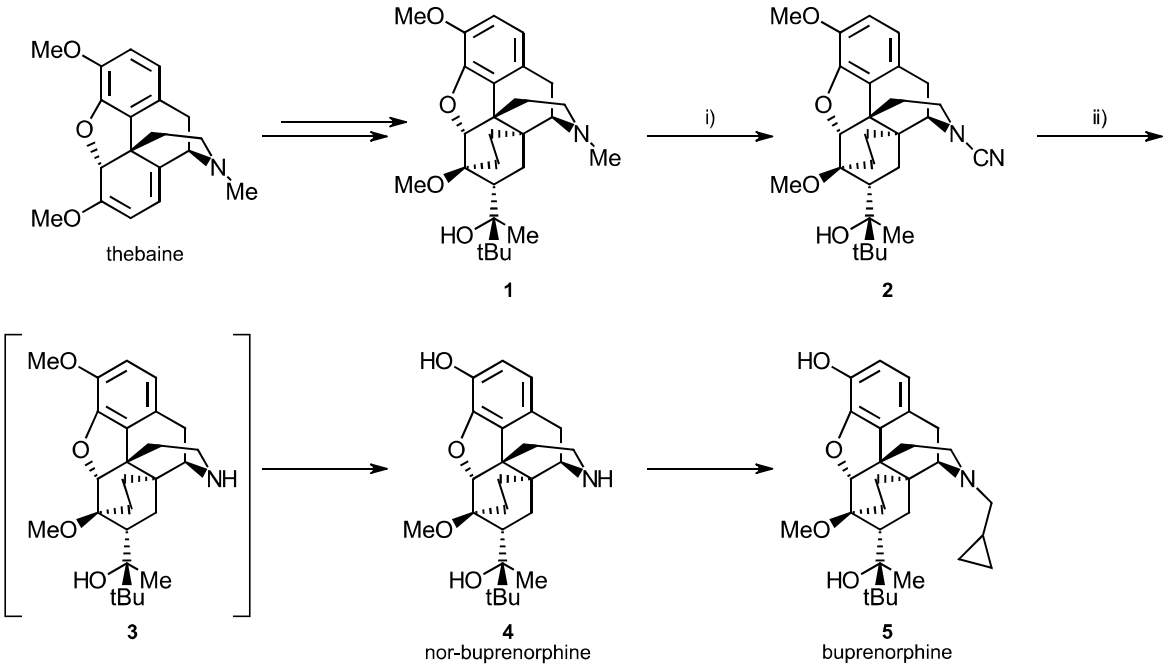
Buprenorphine **5** in low dosages is used to treat severe to chronic pain while at higher dosages it is used for the treatment of opioid use disorder.[1-3] It functions as a partial μ-opioid receptor (MOR) agonist and antagonist of the κ- and δ-opioid receptors, and as an agonist with low binding affinity for the opioid receptor-like 1 receptor.[4]

The commercial synthesis of buprenorphine uses thebaine as the starting material and requires 6 chemical steps, Scheme 1. A consequence of using thebaine as a starting material is the need to perform both N- and O-demethylation reactions. N-demethylation is achieved using the von Braun reaction whereby compound **1** is converted to cyanamide **2** using cyanogen bromide. In the second step, cyanamide **2** is heated with excess potassium hydroxide to give nor-buprenorphine **4**. Clean hydrolysis of the cyanamide occurs to produce intermediate **3** at temperatures of approximately 130-140°C. Temperatures in excess of 200°C are required to achieve complete O-demethylation.

Conversion of **1** to cyanamide **2** is a clean, high yielding reaction, however the reaction uses the highly toxic reagent cyanogen bromide and the highly toxic, volatile methyl bromide is produced

as a by-product. The reaction time for the hydrolysis step needs to be carefully controlled, as multiple low-level impurities are formed with increasing time and the yield for the step is only moderate.[5]

Scheme 1. Synthetic Route to Buprenorphine.



[a]Reagents and Conditions: (i) CNBr; (ii) KOH, >200°C.

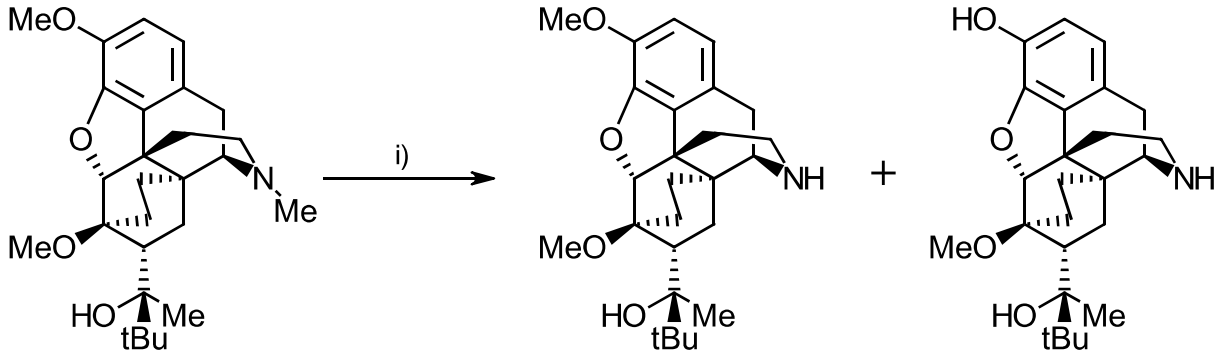
Alternative routes to buprenorphine starting from oripavine have been published.[6-8] Starting from oripavine obviates the need to perform the difficult C-3 O-demethylation. However, in order to obtain a productive yield in the Grignard reaction required to introduce the *tert*-butyl group a protection/deprotection strategy for the C-3 phenolic group needs to be employed, thus reducing the attractiveness of this approach.

The use of nucleophilic thiols in dipolar aprotic solvents at elevated temperatures has been reported for the demethylation of buprenorphine precursors.[6,8,9] However, the noxious nature of

the thiols coupled with the use of dipolar aprotic solvents makes this approach unattractive. Other chemical methods for the cleavage of phenolic methyl ethers tend to use acidic conditions, these conditions are not compatible with buprenorphine.[10]

Over 20 years ago, the use of filamentous fungi for the demethylation of buprenorphine synthetic intermediate **1** was reported.[11-13] The main products from these reactions were the N-demethylated compound **3**, obtained in isolated yields of ca 60% and the N- and O-demethylated compound **4** obtained in isolated yields of ca 20%, Scheme 2. The principal limitation of this chemistry was that the main product **3** obtained still contained the difficult to cleave C-3 O-methyl group.

Scheme 2. Demethylation of compound **1** using filamentous fungi.



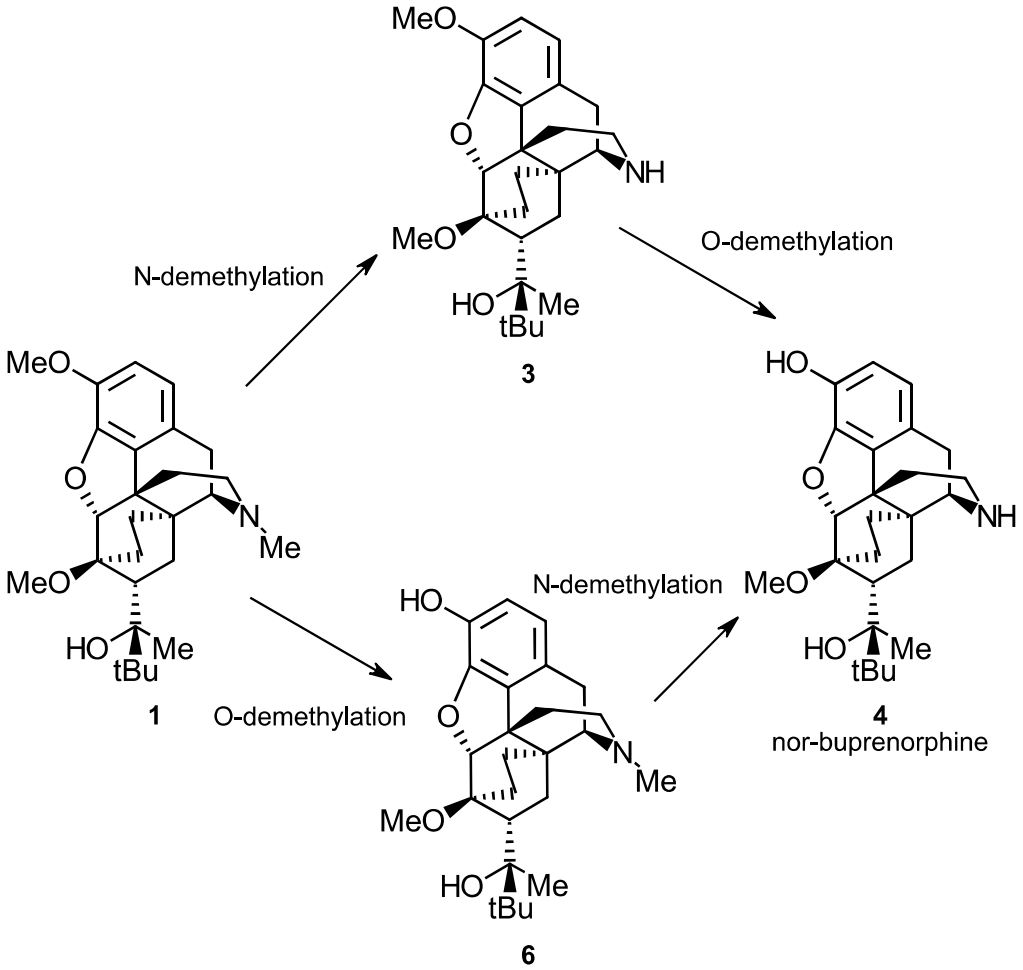
1 **3** **4**

[a]Reagents and Conditions: (i) filamentous fungi.

Given the rapid advances in enzyme engineering and biocatalysis[14-16] since these reports over 20 years ago it was decided to revisit this area and look to apply rational design and directed evolution methods to the problem of N- and O-demethylation in the synthesis of buprenorphine. In this publication we report our preliminary findings from these investigations which highlight the potential of this approach.

In principle there are two potential pathways to achieve the conversion of intermediate **1** to nor-buprenorphine **4,** Scheme 3. Either N-demethylation would occur first to produce intermediate **3** or O-demethylation would occur first to produce intermediate **6** with the roles reversed in the second step of the sequence.

Scheme 3. Potential biocatalytic pathways for the conversion of intermediate **1** to nor-buprenorphine **4.**



Almac panels of wild-type and recombinant enzymes were screened for the desired reactions and two enzymes were selected based on conversion for rational design and directed evolution. The engineered enzymes were able to catalyse N- and O-demethylation as two consecutive chemical steps. Interestingly, a mutant was also found that performed both the N- and O-

demethylation concurrently. A biocatalytic approach for the synthesis of nor-buprenorphine has been demonstrated.

Methods

Computational Modelling

3D models of WT BM3 variants were based on the initial WT BM3 crystal (pdb: 1BVY)[17]

For WT BM3 and IND-78 the 3D model was further sampled with molecular dynamics simulations. The final model of both enzymes is represented by a molecular Dynamics reference structure which corresponds to the structure with the lowest RMSD (α-C atoms) relative to the average structure of the simulation.

A 3D model of CODM was built using homology modeling. As templates crystals 1GP5 and 4XAE[18] were used that share 34 % and 33 % of sequence identity with 88% and 99% of CODM sequence coverage, respectively.

Molecular Docking

Molecular docking was performed using Autodock4.2[19] with the Lamarckian genetic algorithm (LGA) using a grid. A total of 1000 LGA runs were carried out per system. The population was 300, the GA elitism=1, the maximum number of generations was 27000 and the maximum number of energy evaluations was 2500000. The top ranked structure corresponds to lowest binding energy structure of the most populated cluster with the lowest mean binding energy.

Molecular Dynamics Simulations

Molecular dynamics (MD) simulations of WT BM3 and IND-78 were run to further sample the conformational space of those enzymes. The simulations were performed with GROMACS[20] with the amber parm99SB[21] force field. One initial energy minimization was performed, followed by two equilibration steps to slowly heat the system from 0 to 300 K. The first equilibration was performed in a canonical ensemble and the second in an isothermal-isobaric ensemble. Temperature and pressure coupling were 300 K and 1 bar, respectively and periodic boundary conditions (PBC) were used. Production runs were performed in an isothermal-isobaric ensemble. For each enzyme 4 replicas of the production simulations (60 ns) were carried out at 300 K The time step was set to 2 fs and LINCS[22] constraints were applied to all bonds involving hydrogen atoms. The particle mesh Ewald (PME) method[23] was used to calculate electrostatic interactions.

Compounds

Compound **1** is an intermediate in the commercial synthesis of buprenorphine. For this study, recrystallised **1** with a purity of >99% was used. Authentic samples of compounds **3** and **6** were available. A sample of compound **3** was prepared by reacting cyanamide **2** with excess KOH at 130-140°C. An authentic sample of compound **6** was produced by alkylating nor-buprenorphine **4** with methyl iodide. Nor-buprenorphine **4** is an intermediate in the commercial synthesis of buprenorphine.

Panel screening

Screening of selectAZyme™ wild-type microbial panel

The wild-type enzymes were grown in 10 ml cultures of TB growth medium. Following inoculation of media from microbial glycerol stocks, the cultures were incubated for a week at 28°C with shaking (200 rpm), which allowed for the activation of secondary metabolic pathways. After that, 1 ml of each microbial culture was transferred into the wells of deep 96 well plates. The plates were incubated at 28° C with agitation for 48 hours before addition of compound **1** dissolved in DMSO (to the final concentration of 1 mg/ml). The plates were incubated for a further 36 hours before addition of 1 ml of 100% acetonitrile, after which the plates were incubated at room temperature before centrifugation at 9,000×g for 15 mins. 100 µL of supernatant were transferred to a clean 96 well plate containing 100 µl of 50% acetonitrile before analysis by UHPLC.

Screening of STREPTOMET™ panel

The cultures for the screening were prepared by inoculating 25 mL of GYM Streptomyces medium with 10 mg of freeze-dried powder of each STREPTOMET™ strain. The cultures were grown for a week at 28°C with shaking (100 rpm). Prior to addition of compound **1** in DMSO the cultures were centrifuged. The cell pellets were resuspended in 5 ml of 0.1 M potassium phosphate buffer (pH 7.4) and 200 µl of 20 g/L compound **1** in DMSO were added. The cultures were incubated horizontally at 28 °C and 200 rpm for three days. To quench the reaction and disrupt the cells each culture was mixed with acetonitrile in a ratio 1:1 (v/v). The suspensions were then centrifuged for 10 minutes to remove the cell debris. The supernatants were used for UHPLC analysis.

Screening of selectAZyme™ recombinant microbial P450 enzymes

The enzymes were screened in 96 well plates containing 10 mg of freeze dried P450 enzyme. To each well, NADPH (1.53 mg), glucose (3.6 mg) and GDH (2 mg) were added as solutions in 0.1 M potassium phosphate buffer (500 µL, pH 7.4) supplemented with magnesium chloride (10 mM). After that, 10 µl of 20 g/L compound **1** in DMSO were added to each well. The plates were incubated with shaking at 37 ºC overnight. The reactions were stopped by addition of 500 µL of acetonitrile. The plates were centrifuged (9,000×g, 5 min), and the supernatants transferred to a clean 96 well plate and analysed by UHPLC.

Enzyme expression and screening

The genes were cloned into pET28(a)+ and transformed into E. coli expression strain BL21 (DE3) with the resulting transformants plated on selective solid media. For primary cultures a single colony was transferred to 500 µL of LB supplemented with 50 µg/mL kanamycin and incubated at 37 °C with shaking at 200 rpm overnight. Production of screening plates by inoculation of 20 µL from primary plate into a fresh 96-well plate containing 480 µL of the Terrific Broth (TB) broth (24 g yeast extract, 12 g tryptone and 4 g glycerol in 900 mL of water, sterilize and 100 mL of phosphate buffer added, pH 7.4) supplemented with 50 µg/mL kanamycin. The plates were incubated at 37 °C with agitation until OD600 reached 0.6−0.8, after which protein expression was induced by addition of IPTG at a final concentration of 1 mM. For the BM3 variant 1 hour prior to induction $FeCl_3$, $MgSO_4$ and the heme precursor, δ-aminolevulinic acid were added to a final concentration of 0.5 mM, 1 mM, and 1 mM, respectively and temperature reduced to 25 °C. For the CODM enzyme, the cultures were supplemented with Fe_2SO_4 (0.5 mM final concentration) 1 hour before induction and the temperature was lowered down to 10 °C. The plates were incubated for a further 18 hours with

agitation, after which the bacterial cell cultures were centrifuged (9,000×g, 15 min), supernatants were discarded and the plates with bacterial cell pellets were stored at −80 °C.

For the BM3 variant, the screening was carried out in 96 well plates containing the dealkylation enzymes prepared as bacterial cell pellets that were subjected to 3 cycles of freeze/thaw. To each well, NADPH (1.5 mg), glucose (3.6 mg) and GDH (2 mg) were added as solutions in 0.1 M potassium phosphate buffer (500 µL, pH 7.4) supplemented with magnesium chloride (10 mM). Followed by addition of 10 µL compound **1** in DMSO (final concentration 1 g/L) to each well. The plates were incubated with agitation at 37 ºC overnight. The reactions were quenched by adding 500 µL of acetonitrile. The plates were centrifuged (9,000×g, 5 min), and the supernatants transferred to a clean 96 well plate and analysed by UHPLC, using a Kinetex 2.6 µm EVO C18 100 Å 50×2.1 mm column (Phenomenex) with the mobile phase A: 0.1% H_3PO_4 in H_2O and mobile phase B: 0.1% H_3PO_4 in MeOH, a flow rate of 0.5 mL/min and column temperature of 40 °C.

For the CODM screening, the screening was carried out in 96 well plates containing the dealkylation enzymes prepared as bacterial cell pellets that were subjected to 3 cycles of freeze/thaw. The reactions were performed in 0.1 M Tris buffer pH 7.4 with 10 mM ketoglutaric acid, 10 mM sodium ascorbate and 0.5 mM iron sulphate at 30 °C overnight with agitation. The reactions were quenched by adding 500 µL of acetonitrile. The plates were centrifuged (9,000×g, 5 min), and the supernatants transferred to a clean 96 well plate and analysed by UHPLC, using a Kinetex 2.6 µm EVO C18 100 Å 50×2.1 mm column (Phenomenex) with the mobile phase A: 0.1% H_3PO_4 in H_2O and mobile phase B: 0.1% H_3PO_4 in MeOH, a flow rate of 0.5 mL/min and column temperature of 40 °C.

Error-prone PCR

A mutant library with low error rate was generated for P450 BM3 enzyme. The plasmid prep DNA concentration was measured, and the volume calculated to add the required amount of DNA (500 ng). The error prone PCR was set up using GeneMorph II Random Mutagenesis Kit (Agilent Technologies cat # 200550), with the template DNA of the gene of interest (500ng), gene specific PCR primers (1 µL each, 10 µM), 1 µL Mutazyme II polymerase, 5µL of 10xMutazyme II buffer, 1 µL dNTPs (40 mM) and water. The PCR reaction was analysed by agarose gel electrophoresis stained with ethidium bromide and visualised by UV light.

For CODM, error-prone PCR was used to introduce mutations with a medium to high frequency error rate of 7 mutations per Kb and the library was obtained following a Megaprimer protocol. DNA was extracted and transformed into BL21 (DE3) *E.coli*. A total of 2784 clonal mutants were screened in 96 well plates and the mutants with improved conversion were regrown, expressed and rescreened.

Megaprimer mutant library generation and expression

An epPCR was performed using GeneMorph II Random Mutagenesis Kit (Agilent Technologies cat #200550), template DNA of the gene of interest and gene specific PCR primers. Error rate is controlled via the amount of template DNA included in the PCR reaction and the number of PCR cycles performed. Once amplified, the epPCR product was cloned into a suitable expression vector using the "megaprimer" method. This method involves using the epPCR product as large primers and performing mutagenesis to the parental template. The large primers, each complementary to opposite strands of the vector, were extended during PCR temperature cycling by Phusion HF DNA polymerase. Extension of the oligonucleotide primers

generates a mutated plasmid containing staggered nicks. Following the mutagenesis PCR step, the product was treated with Dpn I to remove all WT template which has not been mutated using PCR.

A portion of DNA amplified using the megaprimer method was transformed into Top10 electrocompetent *E.coli* and plated onto agar plates containing appropriate antibiotics. Colony counting was performed to allow estimation of library size. A number of colonies were selected at random and grown overnight in 3 mL LB broth. Plasmid DNA was extracted using QIAprep Spin Miniprep Kit and DNA sequencing performed in order to determine the error rate of the library.

Mutant colonies were harvested from multiple agar plates and pooled and DNA was extracted using QIAprep Spin Miniprep Kit and transformed into the BL21 (DE3) expression strain of *E.coli*. This generated the mutant library which was plated onto LB-agar plates and then individual colonies were picked into 96 well plates for expression analysis as detailed.

Site directed mutagenesis

Site directed mutagenesis is used to make specific and intentional changes to the DNA sequence of a gene. This method was utilized to allow the introduction of further amino acid changes to mutants already known to show enhanced activity. The basic procedure utilizes a supercoiled double-stranded DNA vector (pET28a) with an insert of interest (IND-78 mutant) and two synthetic oligonucleotide primers, both containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, are extended during PCR temperature cycling by Phusion HF DNA polymerase. Extension of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. Following the

mutagenesis PCR step, the product was treated with Dpn I. The Dpn I endonuclease (target sequence:5′-Gm6ATC-3′) is specific for methylated and hemimethylated DNA and was used to digest the parental DNA template and to select for mutation-containing PCR synthesized DNA. The nicked vector DNA containing the desired mutations is then transformed into Top10 electrocompetent cells and plated on LB-agar plates. Colonies are selected for plasmid preparation and sequence verification. Once the sequence was verified with the desired mutation the plasmid was transformed into *E.coli* expression strain for expression and screening.

Results and Discussion

Almac enzyme panels were initially screened for the N- or O-demethylation of compound **1**. The initial panels contained 136 wild-type microorganisms from the selectAZyme™ microbial collection, 21 Streptomyces strains from the Streptomet™ panel, and 390 recombinant P450 enzymes from the standard and extended selectAZyme™ panels. 34 hits were found for the N-demethylation of compound **1** to produce compound **3,** with validated conversions up to 15% for biotransformation by microorganisms and 19% for recombinant enzymes. No hits for the O-demethylation of compound **1** to produce compound **6** were observed. In a second round of screening, 96 designed P450 enzymes derived from plant and microbial sources were profiled for the N-demethylation of compound **1** and the O-demethylation of compound **3**. For the N-demethylation reaction 32 hits were confirmed with conversions up to 11% at 0.4 mg/mL. Only one hit from this panel was identified for the O-demethylation of **3**. In a further screening exercise, a panel of mammalian P450 were screened for the O-demethylation of compounds **1** and **3** to produce compounds **6** and **4** respectively. In both cases hits were identified but these

were not pursued because mammalian P450s are known to be expensive to produce and sensitive to reaction conditions. The best performing enzyme for each transformation was selected for enzyme engineering.

For the N-demethylation of compound **1** the cytochrome P450 BM3 (BM3) from *Bacillus megaterium*, was selected. BM3 is a heme-dependent, soluble and self-sufficient enzyme containing a diflavin reductase fused to a P450 domain. BM3 has been extensively studied and variants are known to catalyse the hydroxylation of compounds like steroidal C7β alcohols[24], sugars and alkaloids[25], the cyclopropanation of olefins[26], the carbene transfer to olefins[27] and the amination of intramolecular C-H[28]. .

For the O-demethylation of compound **3** codeine *O*-demethylase (CODM) was selected. CODM is a 2-oxo-gluturate-dependent dioxygenase (2OG)-Fe(II) from *Papaver somniferum*.[29] It uses 2-oxoglutarate and dioxygen to form formaldehyde, succinate and carbon dioxide products. It has a catalytic triad with two histidine residues and one aspartate residue, which are coordinated to the Fe(II) ion. The active centre further contains an YX82RX2S motif that stabilizes 2-oxoglutarate. Industrially, this class of enzyme was tested for the hydroxylation of amino acids and in the biosynthesis of antibiotics.[30]

Improving N-demethylation with an engineered P450-BM3

BM3 was engineered for the conversion of compound **1** into compound **3** (Scheme 3) using both rational enzyme design and directed evolution approaches. Our rational approach was based on a detailed analysis of the reaction mechanism. The N-demethylation reaction involves an initial hydroxylation by the porphyrin cofactor oxo-iron intermediate into the Cα-H bond of the substrate generating a carbinolamine intermediate. Two hypotheses have been proposed for this

mechanism that differ in the first step. In the single-electron-transfer (SET) mechanism an electron from the N-atom of the substrate is transferred to the cofactor oxo-iron center and then a proton is released from the Cα. The alternative mechanism, hydrogen atom transfer (HAT), is the most probable and starts with the homolytic cleavage of the Cα-H bond. In a subsequent step, the carbinolamine intermediate decomposes into the products (demethylated product and formaldehyde). The oxo-iron intermediate was parameterized and compound **1** was docked to the active site of BM3. New BM3 variants were later designed and then compound **1** was docked into their active centres.

The distance and orientation between the methyl amine carbon atom of the substrate and the oxygen atom coordinated to the heme (Cα- heme O) as well as the predicted binding interactions at the active site were used as the selection criteria in accordance with the mechanism described above (Figure 1, Table 1, Table S1). The variants with the best poses and that make more interactions at the active site than the wild type (WT) enzyme are reported in Table 1.

Table 1. Selected variants and distance between the substrate (alpha-carbon of the amine) and the oxygen of the heme co-factor.

Enzyme	**Sequence changes**	**distance $C\alpha_{amine}$ - O_{heme}/ Å**	**Fold improvement**
WT		4.9 Å	1.0
IND-73	F88V_L438A	6.3 Å	19.5x
IND-74	L76A_F88V_L182A	3.6 Å	25.8x
IND-57	L438A_N320G	6.3 Å	14.6x
IND-65	F88V_L76S	3.2 Å	56.2x

IND-78	F88I_L182A	3.7 Å	**44.3x**

Substrate loading: 0.5g/L

In the selected variants, the N-methyl group of compound **1** is directed to the heme group (Figure 1). For the double mutants IND-78 (F88I_L182A), IND-65 (F88V_L76S) and the triple mutant IND-74 (L76A_F88V_L182A) the reaction coordinate distance (amine Cα- heme 0 distance) is closer than what was observed for the WT BM3. *In vitro* screening indicated that IND-78 (F88I_L182A) and IND-65 (F88V_L76S) had the best fold improvements in relation to the wild-type enzyme (44.3- and 56.2-fold, respectively) (Table 1). IND-78 (F88I_L182A) was chosen as the parental clone for directed evolution. Although this variant has a slightly lower experimental conversion, its two mutations are further apart than in the F88V_L76S variant and was considered a better starting point for directed evolution.

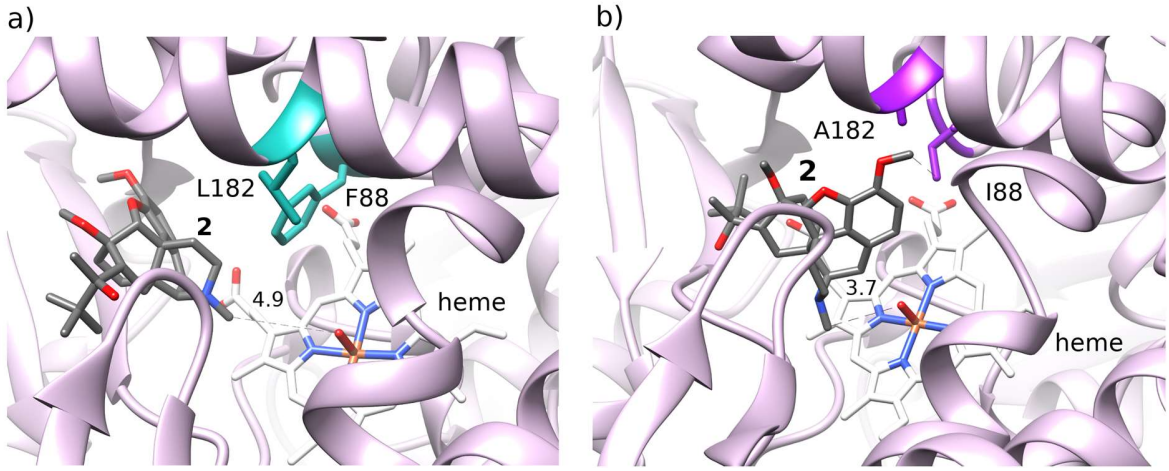


Figure 1. The substrate **1** docked to the active site of a) a MD reference structure of the WT enzyme; b) a MD reference structure of the IND-78 variant. Relevant distances are shown (Å).

Using IND-78 as a parental clone, error-prone PCR was conducted with a mutation rate of 1-5 base changes per kilobase. A total of 2304 colonies were picked (24 x 96 plates) and grown,

expressed and screened with 0.5 mg/mL of the substrate **1**. Overall, 50 % of the mutants were active with conversions similar or higher than IND-78 and 28 % of mutants were inactive. Mutants were then selected for further analysis at an increased substrate loading of 1.5 g/L. The sequencing data and fold improvement for the best variants are summarized in (Table 2). The variant F88I_L182A_V212A_Q513R_R680H showed a 4-fold improvement compared to IND-78 and the variant F88I_L182A_G936C_G993D showed a 2.3-fold improvement (Table 2). These two variants were used as parental clones for the second round of directed evolution. A similar protocol was followed, with a substrate loading of 1 g/L, 3 variants showed improvements. The most notable improvements for F88I_L182A_V212A_Q513R_R680H were reported for variants R2_P3_C5 and R2_P10_H1 with 2.3- and 3.2-fold improvements (Table 2). No mutants with improved activity were generated from the F88I_L182A_G936C_G993D parental clone.

In summary, with just 1 round of rational enzyme design and two rounds of directed evolution 10 mutations were added (R2_P10_H1, Table 2), and the best variant showed at least 567 (44.8 x 4.0 x 3.2)-fold higher conversion of compound **1** into compound **3** compared to the wild-type enzyme (Figure 2).

Table 2. Fold improvement for the best variants and for conversion of compound **1** to compound **3**.

		Fold improvement	Sequence Changes
DE Round 1	**IND-78**	1.0	F88I_L182A
	R1_P6_H7	2.3x	F88I_L182A_G936C_G993D
	R1_P7_H7	1.8x	F88I_L182A_K450I
	R1_P9_H1	1.9	F88I_L182A_M238R_F276L_N523D_D572G_K658E_S681T_E44V_S55N
	R1_P10_G1	4.0x	F88I_L182A _V212A_Q513R_R680H
	R1_P10_H1	1.8	F88I_L182A _N187Y_E353G_P659S
	R1_P16_H4	2.5x	F88I_L182A _Q920*
	R1_P21_E2	1.8x	F88I_L182A _N164K_K508R_P969L
DE Round 2	**R1_P10_G1**	1.0	F88I_L182A _V212A_Q513R_R680H
	R2_P3_C5	2.3	F88I_L182A_V212A_Q513R_R680H_ E840V
	R2_P10_H1	**3.2x**	F88I_L182A_V212A_E268E*_Q513R_E527E*_D587D*_R680H_E840V_Q1005Q
	R2_P10_H2	1.3x	F88I_L182A_V212A_ H389Q_Q513R_S556T_R680H_E840V

DE- Directed evolution. Substrate loading: Round 1:1.5g/L; Round 2:1g/L. *Silent mutations

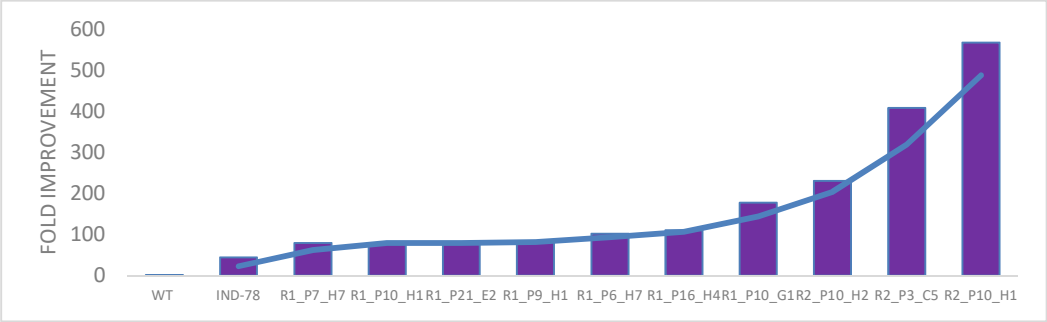


Figure 2. Fold improvements of the variants in relation to the WT enzyme.

Conversion of compound 1 into compound 4 can be achieved with a BM3 variant

For mutant F88V_L182A_S73T generated by rational design (Table 1), molecular docking analysis identified not only a substrate pose described above that is associated with converting compound **1** into compound **3**, but also a new binding pose where the C-3-methoxy is turned to the oxo-iron (the distance is 3.0 Å, Figure 3) which indicates that O-demethylation could also be catalysed. The T73 sidechain is improving the binding affinity of this new pose by establishing a hydrogen bond with the N-H group. So, the molecular docking results seem to indicate that this variant can catalyse not only the N-demethylation of compound **1**, but also the O-demethylation of the *in situ* generated compound **3**, thus producing directly nor-buprenorphine **4** (Figure 3).

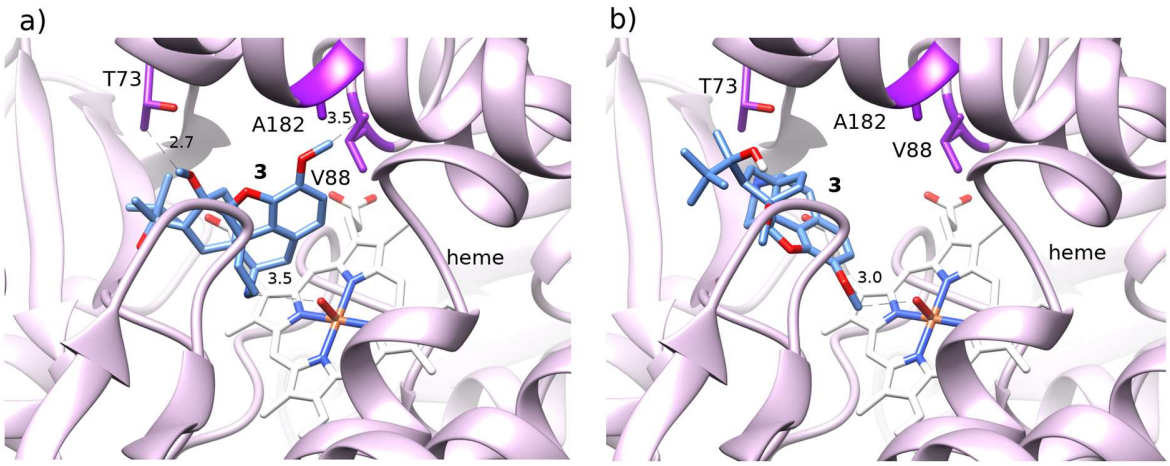


Figure 3. Variant F88V_L182A_S73T docked with **3**: a) most stable pose; b) second most stable pose. Relevant distances are shown in (Å).

In fact, the screening results showed a peak that corresponded to the same retention time as **4** for this mutant. A UPLC chromatogram showing the appearance of compound **4** peak in F88V_L182A_S73T can be found in the Supporting Information (Figure S1).

The presence of compound **4** was further confirmed by LC-MS. By comparing the sample from the triple variant F88V_L182A_S73T with the analytical standard of **4**, a peak was identified with m/z of 414 and consistent fragmentation patterns. All spectra are in the supporting information (Figures S2 and S3).

Improving O-demethylation with an engineered CODM

In the multiple enzymatic sequential routes, the second step for buprenorphine synthesis requires the O-demethylation of **3** to generate nor-buprenorphine **4** (Scheme 3). For this reaction CODM enzyme was re-engineered (Scheme 4 and Figure 4). The reactive Fe(IV)-oxo intermediate for this 2-oxoglutarate (2OG)-dependent dioxygenase is formed after the binding and reaction of Fe(II), 2-oxoglutarate (2OG) and molecular oxygen[31] and the release of succinate

and CO_2[32] The reactive Fe(IV)-oxo intermediate can then abstract a hydrogen atom from **3** forming a radical and a Fe(III)-hydroxide species. Then the radical attacks Fe(III), which results in the restoration of the Fe(II) state, the hydroxylation of the substrate to **4** and the release of formaldehyde.[32]

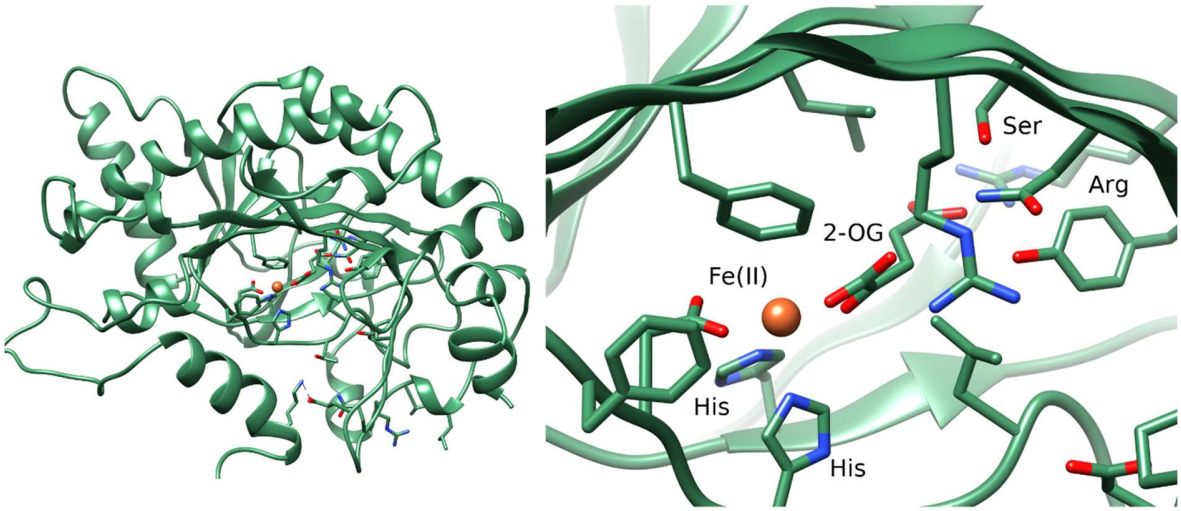
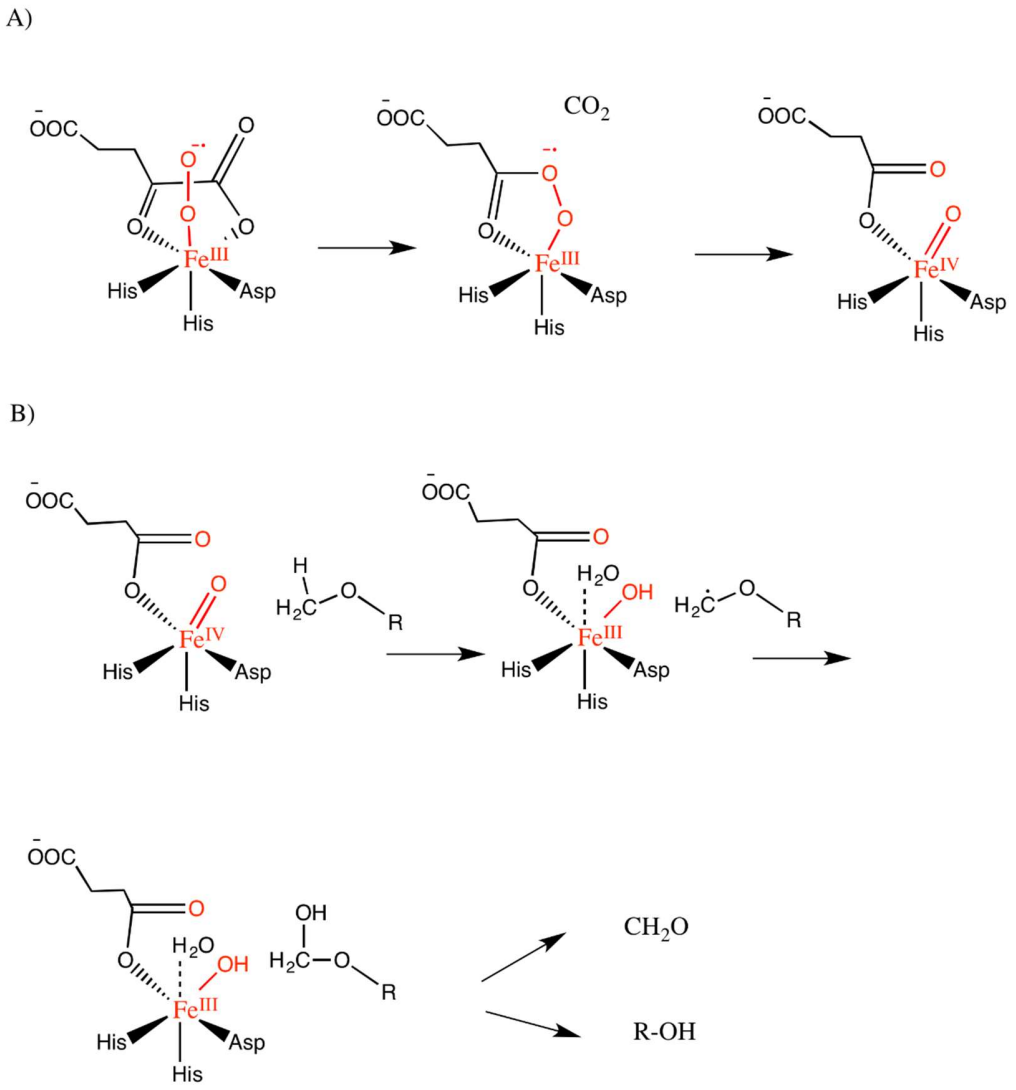


Figure 4. CODM active centre. The catalytic residues are depicted in addition to Fe(II) and 2-oxyglutarate (2-OG).



Scheme 4. Most supported catalytic mechanism hypothesis of CODM, including IND-3. Part A) formation of the Fe-oxo intermediate; Part B) hydroxylation at the O-methyl group by the Fe-oxo intermediate.

The CODM enzyme was successfully cloned into pET28a and optimal growth and expression conditions were investigated. The induction temperature was investigated over a range of 10 °C, 15 °C, 20 °C and 28 °C as well as the 3 induction stages early, normal and late. The determined optimal growth and expression conditions were obtained using LB supplemented with the iron sulphate medium, optical density of 0.6-0.8 and expression at 10 °C (Table S2). These conditions

were used as standard for growth and expression of the libraries. Error-prone PCR was used to introduce mutations with a medium to high frequency error rate of 7 mutations per kb. Following one round of error-prone PCR, three enzymatic variants were selected (R1_P9_G4, R1_P6_G11 and R1_P8_F3), (Table 3) along with the rational variants Mut9 and Mut12 (*vida infra*) and used as parents for a further round of evolution. Of particular interest was the variant R1_P9-G4, which has the mutations S149G and R225H that showed 11.2-fold improvement over the parent enzyme.

In the second round only two variants, based upon S149G_R225H mutant showed improved conversion of approximately two-fold in relation to R1_P9-G4, which equated to about 22-fold improvement in relation to the WT enzyme (Table 3). R2_P11_F7 has a silent mutation that might affect protein folding.[33]

Table 3. Fold improvement for the best variants from directed evolution for conversion of compound **3** to nor-buprenorphine **4**.

		Fold improvement	**Sequence Changes**
DE Round 1	WT	1	
	R1_P6-G11	4.2x	V54A_T145I_E314D
	R1_P8-F3	3.9x	L26L*_E138G_S234N
	R1_P9-G4	**11.2x**	**S149G_R225H**
	R1_P13-B9	3.0x	-
	R1_P13-C9	3.2x	-
	R1_P13-D11	2.2x	S64N
	R1_P14-D2	2.6x	K180E

DE Round 2	R1_P9-G4	1.0	S149G_R225H
	R2_P2-B2	**1.7x**	**H161H*, F189I**
	R2_P11-F7	**2.1x**	**I135I***

* silent mutation

Rational enzyme design of CODM

Molecular docking of compound **3** to the WT CODM, showed that this substrate can bind in a pose, where the methoxy group is relatively close to the iron-center. The distance between the oxygen atom of the methoxy group and the active centre Fe(II) is 5.0 Å. Compound **3** makes a H bond with K348 and a pi-stacking interaction with H235 (Figure 5). Furthermore, there is a flexible loop close to the iron-center. A strong ion-ion interaction between K348 and the loop residue E136 seems to be responsible for keeping the loop in the closed conformation.

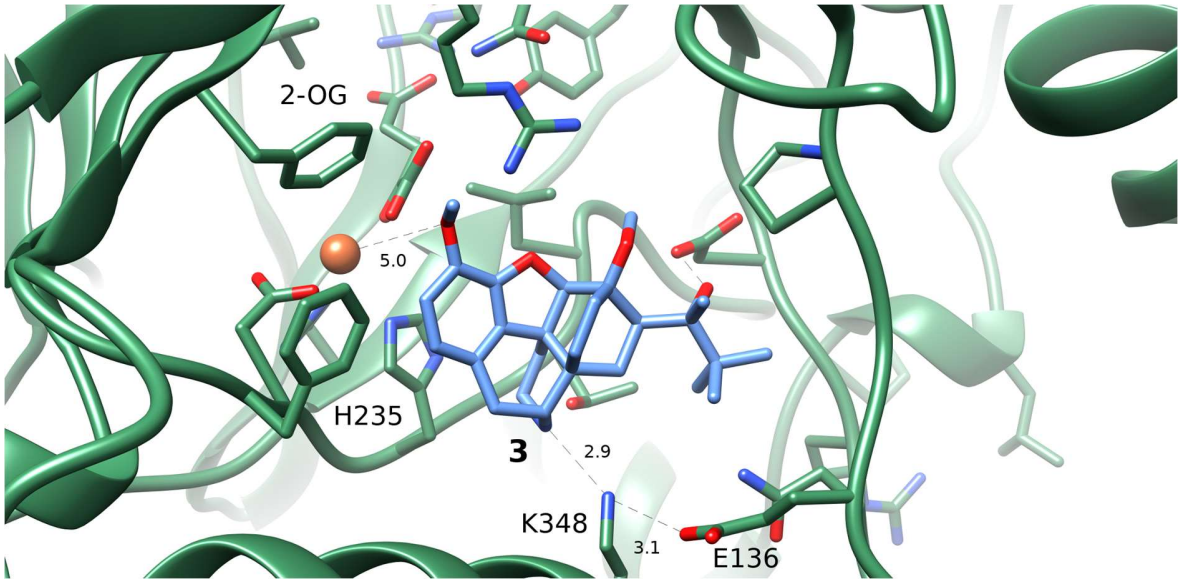


Figure 5. Compound **3** docked to the active site of CODM. Relevant distances (Å) are shown.

Table 4. Fold improvement for the best variants from rational design for conversion of compound **3** to **4** in relation to the WT enzyme.

Rational variants	Fold improvement	Sequence Changes
Mut9	5.5x	E136V
Mut12	5x	E136V_K348A

In parallel to the directed evolution approach, a programme of rationale enzyme design was undertaken. The strategy for the rational mutagenesis involved creating new interactions between compound **3** and CODM active site to improve the binding affinity, but also to promote binding poses that allowed for a good approximation to the iron-centre. There were 22 residues identified as hotspots for mutagenesis and testing with molecular docking. In total 67 variants were gene synthetized, transformed, grown and expressed, and finally tested for conversion of compound **3** into nor-buprenorphine **4**. Most mutants showed modest improvements, with the variants with better fold improvement (above 5) reported on Table 4. The mutations E136V and K348A remove charged residues from the active site. The substrate interacts with London dispersion forces with E136V maintaining a good catalytic pose, with the methoxy group oriented towards the iron cofactor.

In summary, with two rounds of directed evolution 3 mutations were added (S149G_R225H_I135I, Table 3) and the best variant showed 23.5 (11.2 x 2.1) fold higher conversion of compound **3** into nor-buprenorphine **4** when compared to the wild type enzyme.

Conclusions

In this work two enzymes were selected from Almac enzyme panels and engineered for the synthesis of the BIA buprenorphine. The enzymes under study were the P450s BM3 and CODM, both of which have a reactive oxo-iron (IV) intermediate that can abstract a hydrogen atom either from the N-methyl or O-methyl group from buprenorphine.

Enzyme engineering was based on rational enzyme design[15,16] and directed evolution approaches. For BM3, variants were first selected based on the distance between the N-methyl group of compound **1** and the oxygen bond to the heme group of the enzyme. IND-78, which was one of the best variants, had a 44.3-fold improvement in conversion in relation to the WT enzyme. In the first round of directed evolution the best variant improved 4.0-fold for the conversion of compound **1** into compound **3** in relation to IND-78. In the second round a further 3.2-fold improvement was obtained. This corresponds to a cumulative improvement of more than 567-fold. With the rational enzyme design approach, it was possible to design a mutant that can catalyse both the N-demethylation and the O-demethylation converting compound **1** into nor-buprenorphine **4**. Finally, for the CODM enzyme in two rounds of directed evolution the best variant showed a 23.5-fold improvement for conversion of compound **3** into nor-buprenorphine **4**.

Corresponding Authors

john.carey@indivior.com, tom.moody@almacgroup.com

Present Addresses

† School of Pharmacy, Queen's University Belfast, Belfast, UK

Author Contributions

DFARD and ATPC performed the computational calculations, JC, JS and TS carried out the experiments. DJQ, JSC and TSM contributed to the design of the project. ATPC, DFARD, JC, JS, TS, DJQ, JSC and TSM contributed to the manuscript writing, editing and revision. All authors have given approval to the final version of the manuscript.

References

1. *Buprenorphine: Combatting Drug Abuse with a Unique Opioid*; Cowan, A., Lewis, J. W., Eds.; Wiley-Liss: New York, 1995.

2. Mattick, R. P.; Ali, R.; Lintzeris, N. *Pharmacotherapies for the Treatment of Opioid Dependence: Efficacy, Cost-Effectiveness, and Implementation Guidelines*; Informa Healthcare: New York, 2009.

3. Husbands, S. M. Buprenorphine and Related Orvinols. In *Res. Dev. Opioid-Relat. Ligands*; American Chemical Society, 2013; pp 127–144.

4. Gudin, J.; Fudin, J. A Narrative Pharmacological Review of Buprenorphine: A Unique Opioid for the Treatment of Chronic Pain. *Pain Ther* **2020**, *9* (1), 41–54.

5. Carey, J. 32nd SCI Process Development Symposium; 2015.

6. Hudlicky, T. Recent Advances in Process Development for Opiate-Derived Pharmaceutical Agents. *Can. J. Chem.* **2015**, *93* (5), 492–501.

7. Werner, L.; Machara, A.; Adams, D. R.; Cox, D. P.; Hudlicky, T. Synthesis of Buprenorphine from Oripavine via N-Demethylation of Oripavine Quaternary Salts. *J. Org. Chem*., **2011**, *76*, 4628-4634.

8. Machara, A.; Werner, L.; Endoma-Arias, M. A.; Cox, D. P.; Hudlicky, T. Improved Synthesis of Buprenorphine from Thebaine and/or Oripavine *via* Palladium-Catalyzed N-Demethylation/Acylation and/or Concomitant O-Demethylation. *Adv, Synth. Catal*., **2012**, *354*, 613-626.

9. Archer, N.; August, D.; Bease, M.; Jamieson, B.; Marmor, R. S. Method of Preparing Buprenorphine. US 10,208,054 B2, 2019.

10. Grivas, K.; Breeden, S. W.; Ganter, C.; Husbands, S. M.; Lewis, J. W. Acid Catalysed Rearrangements of the Thevinols: The Mechanism of Furanocodide Formation. *Tetrahedron Letters* **1999**, *40* (9), 1795–1798.

11. Abel, A. M.; Allan, G. R.; Carnell, A. J.; Davis, J. A. A Novel Regiospecific N to O-Methyl Transferase Activity in the Biotransformation of a Thebaine Derivative with Cunninghamella Echinulata NRRL 1384. *Chem. Commun.* **2002**, No. 16, 1762–1763.

12. Abel, A. M.; Carnell, A. J.; Davis, J. A.; Paylor, M. Synthesis of Potential Buprenorphine Intermediates by Selective Microbial N- and O-Demethylation. *Biotechnology Letters* **2002**, *24* (15), 1291–1294.

13. Abel, A. M.; Carnell, A. J.; Davis, J. A.; Paylor, M. The Synthesis of Buprenorphine Intermediates by Regioselective Microbial N- and O-Demethylation Reactions Using Cunninghamella Echinulata NRRL 1384. *Enzyme and Microbial Technology* **2003**, *33* (5), 743–748.

14. Hanefeld, U.; Hollmann, F.; Paul, C. E. Biocatalysis Making Waves in Organic Chemistry. Chem. Soc. Rev. 2022, 51 (2), 594–627.

15. Dourado, D. F. A. R.; Pohle, S.; Carvalho, A. T. P.; Dheeman, D. S.; Caswell, J. M.; Skvortsov, T.; Miskelly, I.; Brown, R. T.; Quinn, D. J.; Allen, C. C. R.; Kulakov, L.; Huang, M.; Moody, T. S. Rational Design of a (S)-Selective-Transaminase for Asymmetric Synthesis of (1S)-1-(1,1′-Biphenyl-2-Yl)Ethanamine. *ACS Catal.* **2016**, 6 (11), 7749–7759.

16. Scott, M. E.; Wang, X.; Humphreys, L. D.; Geier, M. J.; Kannan, B.; Chan, J.; Brown, G.; Dourado, D. F. A. R.; Gray, D.; Mix, S.; Pukin, A. Enzyme Optimization and Process Development for a Scalable Synthesis of (R)-2-Methoxymandelic Acid. *Org. Process Res. Dev.* **2022**, 26 (3), 849–858.

17. Sevrioukova, I. F.; Li, H.; Zhang, H.; Peterson, J. A.; Poulos, T. L. Structure of a Cytochrome P450–Redox Partner Electron-Transfer Complex. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (5), 1863–1868.

18. Sun, X.; Zhou, D.; Kandavelu, P.; Zhang, H.; Yuan, Q.; Wang, B.-C.; Rose, J.; Yan, Y. Structural Insights into Substrate Specificity of Feruloyl-CoA 6'-Hydroxylase from Arabidopsis Thaliana. *Sci Rep* **2015**, *5* (1), 10355.

19. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19* (14), 1639–1662.

20. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

21. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65* (3), 712–725.

22. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.

23. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \cdot \log(N)$ Method for Ewald Sums in Large Systems. *The Journal of Chemical Physics* **1993**, *98* (12), 10089–10092.

24. Li, A.; Acevedo-Rocha, C. G.; D'Amore, L.; Chen, J.; Peng, Y.; Garcia-Borràs, M.; Gao, C.; Zhu, J.; Rickerby, H.; Osuna, S.; Zhou, J.; Reetz, M. T. Regio- and Stereoselective Steroid Hydroxylation at C7 by Cytochrome P450 Monooxygenase Mutants. *Angew. Chem.* **2020**, *132* (30), 12599–12605.

25. Lewis, J. C.; Mantovani, S. M.; Fu, Y.; Snow, C. D.; Komor, R. S.; Wong , C.-H.; Arnold, F. H. Combinatorial Alanine Substitution Enables Rapid Optimization of Cytochrome P450BM3 for Selective Hydroxylation of Large Substrates. *Chem. Eur. J. of Chem. Bio.* **2010**, *11* (18), 2502–2505.

26. Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin Cyclopropanation via Carbene Transfer Catalyzed by Engineered Cytochrome P450 Enzymes. *Science* **2013**, *339* (6117), 307–310.

27. Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H.; Brustad, E. M. A Serine-Substituted P450 Catalyzes Highly Efficient Carbene Transfer to Olefins in Vivo. *Nat Chem Biol* **2013**, *9* (8), 485–487.

28. McIntosh, J. A.; Farwell, C. C.; Arnold, F. H. Expanding P450 Catalytic Reaction Space through Evolution and Engineering. *Current Opinion in Chemical Biology* **2014**, *19*, 126–134.

29. Hagel, J. M.; Facchini, P. J. Dioxygenases Catalyze the O-Demethylation Steps of Morphine Biosynthesis in Opium Poppy. *Nat Chem Biol* **2010**, *6* (4), 273–275.

30. Peters, C.; Buller, R. Industrial Application of 2-Oxoglutarate-Dependent Oxygenases. *Catalysts* **2019**, *9* (3), 221.

31. Wilmouth, R. C.; Turnbull, J. J.; Welford, R. W. D.; Clifton, I. J.; Prescott, A. G.; Schofield, C. J. Structure and Mechanism of Anthocyanidin Synthase from Arabidopsis Thaliana. *Structure* **2002**, *10* (1), 93–103.

32. Hagel, J. M. Biochemistry and Occurrence of O-Demethylation in Plant Metabolism. *Front. Physio.* **2010**, *1*.

33. Kimchi-Sarfaty, C.; Oh, J. M.; Kim, I.-W.; Sauna, Z. E.; Calcagno, A. M.; Ambudkar, S. V.; Gottesman, M. M. A "Silent" Polymorphism in the *MDR* 1 Gene Changes Substrate Specificity. *Science* **2007**, *315* (5811), 525–528.