Catalysis Science & Technology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Catal. Sci. Technol.*, 2021, 11, 7359

Structure of the fungal hydroxylase, CYP505A30, and rational transfer of mutation data from CYP102A1 to alter regioselectivity†

Jasmin C. Aschenbrenner, [ab] Ana C. Ebrecht, [a] Carmien Tolmie, [a] Martha S. Smit [ab] and Diederik J. Opperman *[a]

Received 27th July 2021,
Accepted 10th October 2021

DOI: 10.1039/d1cy01348c

rsc.li/catalysis

CYP505A30 is a fungal, self-sufficient cytochrome P450 monooxygenase that can selectively oxyfunctionalise *n*-alkanes, fatty alcohols, and fatty acids. From alkanes, it produces a mixture of non-vicinal diols by two sequential hydroxylation reactions. Here we report the structure of the haem domain of CYP505A30, the first structure for a member of the CYP505 family, with dodecanoic acid bound within the active site. Overall, a high structural similarity to the related bacterial CYP102A1 was observed, despite low sequence identity (<40%). Comparison of the active sites, however, showed a high degree of conservation with only two amino acid differences close to the haem. Stabilisation of the fatty acid substrate in CYP505A30 also occurs, as in CYP102A1, *via* an arginine residue. However, compared to R47, which is situated in the β1 region of CYP102A1, R358 is located in the β3 region of CYP505A30. We furthermore created mutants to test if it is possible to rationally transfer the knowledge on active site mutations in CYP102A1 to change the regioselectivity of CYP505A30. The introduction of F93V, I334F mutations resulted in increased ω-1 (C2) regioselectivity, similar to CYP102A1 87-328, of more than 80% for *n*-octane and 90% for *n*-decane. Changing residues to resemble the CYP102A1 wildtype increased the regioselectivity towards ω-2 (C3) to over 60% for both substrates. The knowledge gained from this study unlocks a more selective production of symmetrical non-vicinal diols from *n*-alkanes.

Introduction

Cytochrome P450 monooxygenases (CYP450s) are powerful biocatalysts that are widely distributed in nature. These enzymes are mostly known for using a haem prosthetic group to activate molecular oxygen to insert a single oxygen atom into the chemically inert C–H bonds of their substrates. However, other reactions including epoxidations, C–C bond cleavage, sulfoxidation and dealkylation reactions have also been reported.[1]

CYP102A1 (BM3) is an extensively studied bacterial CYP450 fatty acid hydroxylase, whose high activity has, in part, been attributed to the CYP450 domain being fused to an FAD–FMN-containing cytochrome P450 reductase (CPR) domain.[2] As no additional redox partner proteins are required for transferring reducing equivalents, it is classified as a "self-sufficient" CYP450. Recently, the structure of the full-length enzyme showing the haem and reductase domain interactions was solved using cryo-electron microscopy.[3] X-ray crystal structures of its haem domain with the active site have, however, been available for many years now,[4,5] allowing structure-guided mutagenesis studies to not only explore the catalytic mechanism, but to also produce new biocatalysts with changed regioselectivity or substrate scope or to even engineer its activity towards using H_2O_2 as oxidant.[6–9]

These mutagenesis and structural studies have revealed critical residues in CYP102A1 influencing its regioselectivity towards aliphatic substrates. Wildtype CYP102A1 produces a mixture of ω-1 to ω-3 hydroxylated products from fatty acids. As *n*-alkanes pose an even bigger challenge to hydroxylate with high regioselectivity, CYP102A1 mutants with increased activity towards alkanes have been explored. The 139-3 variant, with 11 amino acid changes generated by multiple rounds of error-prone PCR, showed a vastly improved activity towards C3–C8 *n*-alkanes with a shift in activity towards the ω-1 (C2) position (66% *vs.* 17% by WT).[10] Using a similar approach, combined with saturation

[a] *Department of Microbiology and Biochemistry, University of the Free State, 205 Nelson Mandela Drive, Bloemfontein, 9300, South Africa.*
E-mail: opperdj@ufs.ac.za
[b] *South African DST-NRF Centre of Excellence in Catalysis, c*change, University of Cape Town, South Africa*

† Electronic supplementary information (ESI) available: Details on mutant generation, additional figures of crystal structure and GC analysis, including chromatograms. Table of data collection and refinement statistics of crystal structure (7P6L). See DOI: 10.1039/d1cy01348c

mutagenesis, a 53-5H mutant was created, which was shown to have high turnover numbers with *n*-octane, producing mainly (*S*)-2-octanol (89% rr, 65% ee). Two different mutants of CYP102A1 were also created by combining directed evolution and site-directed mutagenesis starting from 139-3: 9-10A-A328V and 1-12G, both of which showed a high regioselectivity for the production of 2-octanol (>80%). While 9-10A-A328V predominantly produced the *S*-enantiomer of 2-octanol, the addition of a single amino acid change A82L in 1-12G altered it to primarily produce the *R*-enantiomer.[11] By targeting only two amino acid positions in the active site of CYP102A1, an F87V/A328F mutant was generated, with a high regioselectivity for the ω-1 (C2) position of *n*-octane (85%).[12]

Related to the bacterial CYP102s is the eukaryotic CYP505 family, also containing self-sufficient CYP450s. The first of these enzymes, CYP450foxy, was found in *Fusarium oxysporum* more than ten years after the description of CYP102A1.[13] Compared to bacterial CYP450s, little biochemical and structural information is available on fungal CYP450s, because they are more difficult to purify as they often are membrane-associated proteins which are not as easily expressed in *E. coli*.[14] The characterised CYP505s, similar to CYP102s, have been shown to hydroxylate alkanes, alcohols and fatty acids sub-terminally. Notable exceptions are CYP505E3 and CYP505D6 that also perform deeper in-chain hydroxylation of these substrates and CYP505B1 (Fum6), which is involved in fumonisin mycotoxin production.[15–19] Deviations from the typical sub-terminal hydroxylation pattern have also more recently been found in certain CYP102 members. BAMF2522 from *Bacillus amyloliquefaciens* DSM7 hydroxylates palmitic acid at positions ω-1 to ω-7. This regioselectivity seems to be chain-length dependent, as longer and shorter substrates than palmitic acid are mainly converted to ω-1 to ω-5 hydroxylated products.[20] Krac9955 from *Ktedonobacter racemifer* forms a mixture of ω-1 to ω-8 hydroxylated fatty acids with favoured hydroxylation at positions ω-3 to ω-6, especially at chain lengths between C13 and C15.[21]

We recently reported an in-depth characterisation of the products produced by CYP505A30 from *Myceliophthora thermophila*.[22] This self-sufficient CYP450 formed a mixture of non-vicinal diols through sequential hydroxylation of *n*-octane and *n*-decane. Here, we report the X-ray crystal structure of the haem domain of CYP505A30. This is the first structure of a member of the CYP505 family and the CYP505 family is now only the third fungal P450 family for which an X-ray structure is available. Through structural comparison with CYP102A1, amino acids involved in the different regioselectivities found between CYP102A1 and CYP505A30 were identified. Using structure-guided site-directed mutagenesis based on mutation data of CYP102A1, we were able to change the regioselectivity of CYP505A30 with *n*-alkanes. This, in turn, allowed us to fine-tune diol production to achieve high selectivity towards 2,7-octanediol and 2,9-decanediol synthesis.

Results and discussion

Crystal structure of CYP505A30 haem domain

Numerous attempts at the crystallisation of the full-length CYP505A30 were unsuccessful. Crystals of the CYP505A30 haem domain (HD), constituting residues 1–464, were, however, obtained by co-crystallisation with a substrate. Not only did the presence of substrate increase the diffraction resolution, but also the ease of obtaining crystals, with CYP505A30 HD crystallising more readily with fatty acid substrates than fatty alcohols. The crystal structure of CYP505A30 HD was solved at a resolution of 2.3 Å by molecular replacement using the HD of CYP102A1 as an initial search model (39% sequence identity between the haem domains). The asymmetric unit (ASU) contained 4 copies with a space group of $P4_22_12$. Electron density was observed for residues 7–464, with chain A showing the best electron density of the 4 copies in the ASU. Poor electron density was observed for chains C and D, with difficult to interpret electron density for flexible loop residues of the FG-loop (residues 196–205). Furthermore, no clear electron density was observed for the substrate in chain D. Images and discussions will therefore refer to chain A.

The overall structure of CYP505A30 HD adopts the general fold typically observed for CYP450s (Fig. 1A). Comparisons to haem domains of related CYP450s showed a higher structural similarity to CYP102A1 (bound to palmitoleic acid) with an rmsd of 0.753 Å, than to other available fungal CYP450 structures, such as P450nor (rmsd 2.460 Å) or CYP51 (rmsd 3.067 Å). As the CYP505 family constitutes the eukaryotic equivalent of the bacterial CYP102 family and are evolutionarily related, a certain similarity between members of these families is expected.[23] It is, however, important to note that this is the first X-ray structure of a member of the CYP505 family and that the only fungal P450 families for which X-ray structures have hitherto been available are the CYP51 and CYP55 families.

Electron density for dodecanoic acid bound in the substrate channel was observed, with the overall structure in the resulting closed conformation. Similar to the crystal structure of CYP102A1 in complex with palmitoleic acid,[4] a distance of approximately 8 Å between the haem and the substrate indicates that this does not, however, constitute a productive binding orientation. With the exception of the catalytic acid–alcohol pair (T274, E273)[24] (Fig. S18†), the substrate channel is lined mainly by non-polar, aliphatic amino acids such as valine, (iso)leucine and phenylalanine (Fig. S20†). Despite the low overall sequence identity between CYP505A30 and CYP102A1, the active site channel leading to the haem domain is highly conserved, with only 3 amino acid differences, where alanine residues of CYP102A1 are substituted by bigger amino acids I334, I336 and V88 in CYP505A30 (Fig. 2).

More pronounced variation is observed at the substrate entry site. The carboxyl group of dodecanoic acid is within 3.2 Å of the guanidinium group of R358, allowing "anchoring" *via*

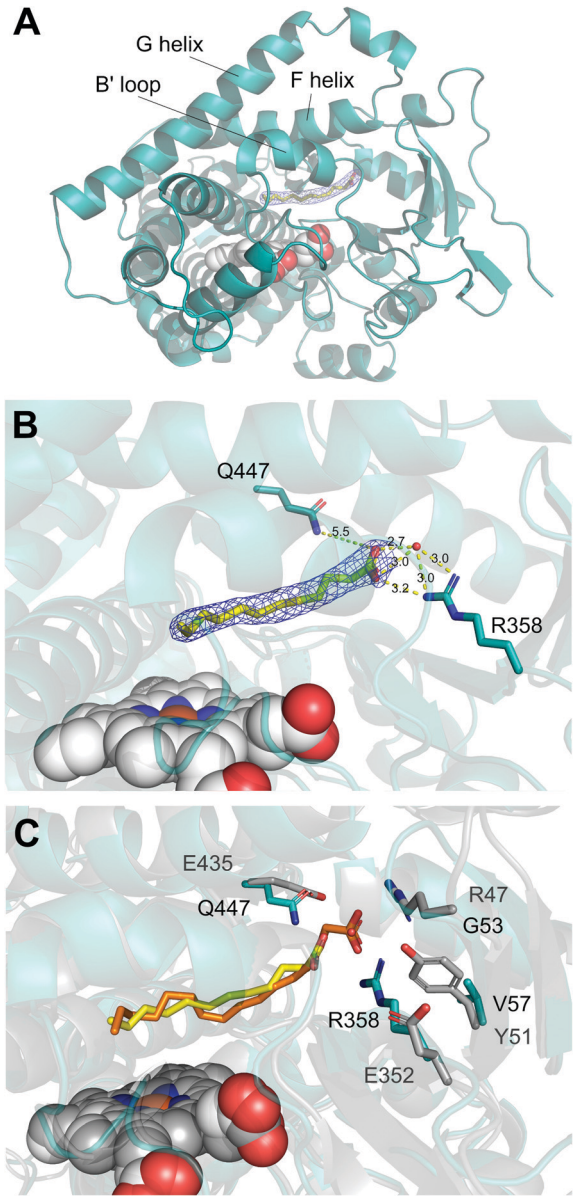

Fig. 1 (A) Overall structure of CYP505A30 HD (cyan) with 1-dodecanoic acid (yellow) bound in the active site. Electron density is shown as blue mesh from omit map ($2F_o - F_c$) contoured at σ = 1 Å. (B) Active site of CYP505A30 HD showing the 1-dodecanoic acid anchored by R358, Q447 and a water molecule. Distances are measured in Å. (C) Overlay of haem domains of CYP505A30 and CYP102A1 (grey) with palmitoleic acid (orange) showing the similarities in substrate orientation and the differences in substrate anchoring.

ionic interaction in addition to indirect hydrogen-bonding *via* a water molecule. Unlike CYP102A1 where substrate stabilisation occurs by interaction with R47 located near the N-terminal region, R358 is located on a beta-strand in the β3 region. Glutamine Q447, located on beta-strand β5, is also found in close proximity to the carboxylic acid group of the fatty acid substrate (Fig. 1B). Within CYP102A1, glutamic acid residues occupy the equivalent positions in both these positions, but with the sidechains adopting an orientation rotated away from the substrate (Fig. 1C).

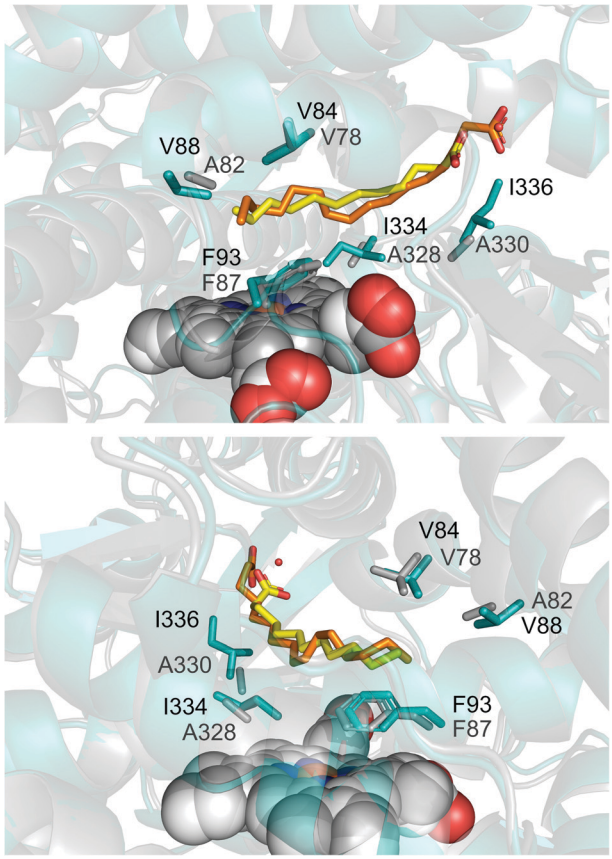

Fig. 2 The active site of CYP505A30 HD (cyan) with bound 1-dodecanoic acid (yellow) overlaid with CYP102BM3 (grey) with palmitoleic acid (orange) showing the amino acid residues V84, V88, F93 and I334 close to the haem in CYP505A30 HD that were mutated to change the regioselectivity.

The I helix contains the oxygen-binding groove leading to a slight bend in the helix.[25] Located between amino acids A270 to T274 in CYP505A30, its architecture resembles the groove found in CYP102A1 which also shares the same amino acid motif AGHET. The F- and G-helices, including the connecting FG-loop, as well as the BC-loop, including the B′ helix, are generally involved in substrate recognition and undergo conformational changes upon substrate binding, closing off the substrate channel.[26–28] In CYP505A30, part of the FG-loop appears ordered into a short 3_{10} helix (Fig. S19†). The B′ helix (residues 80–88) is located near the F- and G-helices and is lying nearly parallel to the 1-dodecanoic acid, as is seen in CYP102A1. The conserved F93 is similarly positioned to the F87 residue in CYP102A1 which has been determined to be critical for efficient catalytic turnover and regiospecificity for dodecanoic acid oxidation.[7]

Structural determinants of regioselectivity

Although the substrate scope of CYP102A1 and CYP505A30 is comparable, a difference in regioselectivity is seen in the product profile. CYP102A1 produces a majority of ω-3 (C4) hydroxylated products from *n*-octane and *n*-decane, whereas

CYP505A30's regioselectivity favours hydroxylation at the ω-1 (C2) and ω-2 (C3) positions (even though regioselectivity is also dependent on substrate length).[22,29] We anticipated these differences in regioselectivity to arise due to variations in amino acids in the active site. Comparing the active sites of both enzymes, we noticed two distinct alanine residues in close proximity to the haem in CYP102A1 (A82 and A328) that are replaced by valine (V88) and isoleucine (I334), respectively, in CYP505A30 (Fig. 2). Isoleucine and valine, as well as alanine, are non-polar amino acids and are not surprisingly found in the active site and the substrate channel, as these CYPs accept hydrophobic substrates. They, however, vary in size and potentially coordinate the substrate into different positions. Changes in A328, an amino acid that is conserved in most CYP102s and located close to the haem, have been reported to influence substrate binding.[11,21] To test our hypothesis, we created a CYP505A30 V88A, I334A mutant to mimic the active site of CYP102A1 more closely (Table 1).

We additionally looked into previous studies in which the regioselectivity of CYP102A1 had been changed. Peters and co-workers described a CYP102A1 mutant, 1-12G, that showed high selectivity for the ω-1 (C2) position when tested on C6 to C10 *n*-alkanes.[11] 1-12G differs from the wildtype in 13 residues, but notably, the A82L and A328V mutations were stated as being responsible for the shift in regioselectivity towards the ω-1 (C2) position, while the A82L mutation also changed the enantioselectivity of the products from *S*- to *R*-2-alcohol.[11]

When F87V and A328F mutations were introduced to CYP102A1, the regioselectivity also shifted towards the 2-alcohol and favours the *R*-enantiomer.[12] A bulky amino acid at either the F87 or A328 position has been implicated in the more tightly bound octane to be preferentially hydroxylated at the second carbon.[12] To establish if the same mutations have a similar effect in CYP505A30, site-directed mutagenesis was performed at the respective corresponding amino acids (Table 1).

Regioselectivity of CYP505A30 mutants

In total, four mutants of CYP505A30 were generated (Table 1) that would either mimic the active site of wildtype BM3 or the 87-328 and 1-12G variants. Expression of CYP505A30 and the mutants was performed in *E. coli* BL21-Gold(DE3) cells in auto-induction medium. Our analysis of the CYP content in the cell-free extract (CFE) *via* CO-difference spectra showed that all mutants were expressed, with the characteristic peak observed at 450 nm for all (Fig. S4†). The 505A30_BM3wt mutant showed a similar expression level to the wildtype CYP505A30, whereas the mutations introduced to create the 1-12G and the 87-328_2 mutants resulted in lower expression levels. This was comparable to what was seen before for expression levels of CYP102A1 (BM3) and its mutants.[29]

Good CYP recovery was obtained with most enzymes, as the amount of CYP measured in whole cells (WC) was similar to the amount of CYP measured in the CFE after disruption and removal of cell debris. The 505A30_1-12G and the 505A30_87-328_2 mutants, however, showed poor recovery, as only 80% of CYP was recovered from WC for A30_1-12G and more than 70% of CYP was lost for the 505A30_87-328_2 mutant (Fig. S3A†). The volume of CFE was kept constant in each biotransformation reaction mixture and contained on average between 1 and 5 μM of CYP (Fig. S3B†).

The activity of all tested enzymes was generally higher towards *n*-octane than towards *n*-decane (Table 2). The achieved turnover numbers of CYP505A30wt are, however, lower than the activities we have reported previously (reduction by *ca.* 13–27%),[22] which can be due to the different conditions we used in this set of experiments (*e.g.* lower temperature of 20 °C *vs.* 30 °C, Tris buffer *vs.* potassium phosphate buffer). In the biotransformations with *n*-octane, a reduced TOF and TTN for all mutated enzymes compared to the 505A30wt were seen (Table 2). For *n*-decane, the 505A30_87-328 and 87-328_2 variants, however, showed an increased TOF and TTN, with the 505A30_87-328 mutant producing more than *ca.* 10 mM total products within 24 h (Fig. S13†). Compared to the TTNs achieved by Pennec *et al.* who performed *n*-octane and *n*-decane biotransformations with the CYP102A1 wildtype and variants under similar conditions,[29] higher TTNs were achieved with CYP505A30 wildtype and all its derived variants (Table S4†).

The introduced mutations did indeed affect the regioselectivity of CYP505A30 (Fig. 3). An increased percentage of ω-1 (C2) hydroxylation of *n*-octane (>75%) and

Table 1 CYP505A30 mutants created based on CYP102A1wt and mutant counterparts that have been described

CYP505 variants	Mutations introduced	Corresponding CYP102A1 variants and reported regioselectivity[a,b]
A30_BM3wt	V88A, I334A	A1wt (A82, A328) >40% 4-octanol
A30_87-328	F93V, I334F	87-328 (F87V, A328F) >90% *R*-2-octanol
A30_87-328_2[c]	V88A, F93V, I334F	87-328 (A82, F87V, A328F) >90% *R*-2-octanol
A30_1-12G	V84A, V88L, I334V	1-12G (V78A, A82L, A328V) >75% *R*-2-octanol

[a] Corresponding amino acids and mutations in the active site of CYP102A1 given in parenthesis. [b] Regioselectivity of CYP102A1wt and variants reported on *n*-octane as substrate.[11,12,29] [c] Two mutants of CYP505A30 for 87-328 were generated: a double mutant CYP505A30_87-328 (F93V, I334F) and a triple mutant CYP505A30_87-328_2 (V88A, F93V and I334F), effectively combining BM3wt with 87-328.

Table 2 Total turnover numbers (TTN) and turnover frequencies (TOF) for CYP505A30wt and mutants. 1 h and 2 h data points of biotransformations were used to calculate TOF, while 8 h and 24 h data points were used to calculate TTN. The data are given as the mean of duplicates from at least 2 independent experiments ± standard deviation (SD)

	n-Octane[a]		*n*-Decane[a]	
Enzyme	TTN	TOF (min^{-1})	TTN	TOF (min^{-1})
505A30wt	7860 ± 1580 (10 720[b])	34.9 ± 2.8 (73.8[b])	2730 ± 480 (3130[b])	11.4 ± 3.2 (8.8[b])
505A30_BM3wt	4390 ± 1060	18.3 ± 1.4	2300 ± 230	7.5 ± 1.7
505A30_1-12G	4720 ± 1520	14.8 ± 5.9	1500 ± 310	6.0 ± 1.8
505A30_87-328	3670 ± 910	23.6 ± 7.5	3160 ± 960	11.2 ± 1.4
505A30_87-328_2	4660 ± 960	22.7 ± 5.1	4120 ± 350	17.5 ± 2.3

[a] Alkanes were used in excess (two-phase systems), resulting in substrate concentrations of *ca.* 1.2 M for *n*-octane biotransformations and *ca.* 1.0 M for *n*-decane biotransformations. [b] TTN and TOF values in parenthesis were previously reported for CYP505A30wt by Ebrecht *et al.*[22]

n-decane (>90%) was observed for both the 87-328 and 87-328_2 mutants. The higher preference for hydroxylating at ω-1 (C2) also significantly reduced the production of 4-octanol and 4-decanol by both 87-328 variants (Fig. 3). F87 is a conserved amino acid in the CYP102 family, regulating access to the haem iron and therefore regioselectivity.[30] Changing it to a valine in CYP102A1 or BAMF2522 has been shown to increase in-chain hydroxylation of dodecanoic acid and only minimally increased 2-octanol production from *n*-octane compared to the CYP102A1wt.[12,31,32]

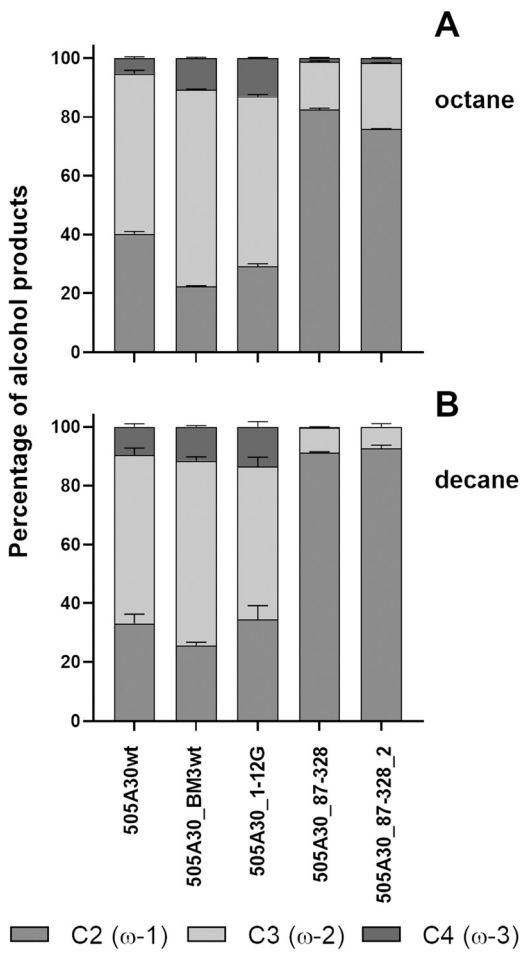

Fig. 3 Regioselectivity of CYP505A30wt and mutants for *n*-octane (A) and *n*-decane (B). The 1 h data points of the CFE biotransformations were used to determine the regioselectivities. Data does not include the formed diols (<20% for octane and <50% for decane). Results are the averages for duplicate reactions of at least two independent experiments.

However, its combinatorial effect with other residues resulted in more drastic changes in regioselectivity. A328 is conserved in many CYP102s,[21] whereas in CYP505A30 an isoleucine occupies the equivalent position (I334). In the 87-328 and 87-328_2 variants, the even bulkier amino acid (I334F) directly next to the haem appears to limit the movement of the substrate, thus leading to a preferred orientation of the second carbon in close proximity to the haem, resulting in a majority of ω-1 (C2) hydroxylated products. A similar restriction in the variation of substrate binding has been described for an A330V mutant of CYP102A3 (corresponding to A328 in CYP102A1).[33] The additional mutation V88A in 505A30_87-328_2 only had a slight additive influence on the regioselectivity. A reduction in ω-1 (C2) hydroxylation and an increase in ω-2 (C3) and ω-3 (C4) hydroxylation was observed for the 505A30_BM3wt and 1-12G mutants (Fig. 3). The mutations introduced to 505A30_1-12G (V84A, V88L, I334V) only had a marginal effect on the regioselectivity compared to the wildtype, especially regarding the alcohol products from *n*-decane. The 505A30_BM3wt mutant's selectivity for ω-2 (C3) and ω-3 (C4) increased (Fig. 3). The preference for hydroxylating further in-chain was expected for the BM3wt mutant but surprising for 505A30_1-12G as its CYP102A1 counterpart mainly produces 2-octanol. Looking at the mutations V84A, V88L and I334V, and V88A and I334A, it appears that the substrate is able to enter deeper into the cavity to position the C3 and C4 of the substrate closer to the haem. This can be due to the smaller size of the amino acids alanine and valine that replaced I334 in CYP505A30wt.

To determine if the mutations had a similar effect on the enantioselectivity of octanol products formed by CYP505A30 as seen with CYP102A1 mutants, we analysed the 2-octanol enantiomers after 1 hour of biotransformations when only limited amounts of diols had been produced (Fig. S14†). The wildtype predominantly produced (*R*)-2-octanol. Both 505A30_87-328 mutants and the 1-12G mutant also produced

(*R*)-2-octanol as the main enantiomer, which we expected as it is the primary product of CYP102 87-328 and CYP102 1-12G.[11,12] However, the selectivity for the production of the *R*-enantiomer (74% ee) was reduced compared to 505A30wt and 505A30_BM3wt (both >94% ee).

Diol production by CYP505A30 and mutants

CYP505A30 produces diols by performing a second round of hydroxylation on the alcohol produced from the alkane without over-oxidising the alcohol to a ketone. This behaviour was also seen in other members of the CYP505 family, which use this sequential hydroxylation activity for mycotoxin production.[19] Diol formation was already detected in the biotransformation reaction after 1 h. The total concentration of diols produced by CYP505A30 reactions was on average 5.6 mM of octanediol and 5.3 mM of decanediol after 24 h. Overall, the relative percentage of diols to all products formed in *n*-decane biotransformations was higher than in *n*-octane biotransformations (Fig. S12†), possibly due to the low solubility of *n*-decane compared to the higher solubility of the *n*-decanol products.[22] Control reactions with CYP102A1 showed that it also produces diols but to a lesser percentage of overall products formed (Fig. S14†).

A similar pattern of diol formation was observed for both *n*-octane and *n*-decane biotransformations. Both 87-328 mutants predominantly produced diols by two sequential ω-1 hydroxylation steps, resulting in 2,7-octanediol and 2,9-decanediol being formed as the major products (*ca.* 72% and 88%) (Fig. 4). A higher regioselectivity was seen with the longer substrate, which was also previously described for CYP102A1 mutants.[29] The mutants with higher regioselectivity for the ω-2 position and the 505A30wt produced a similar amount of 2,8-decanediol of about 50%, while an increase in 3,8-decanediol and 3,6-octanediol production was evident for the 505A30_BM3wt mutant.

Nevertheless, the regioselectivity for diol formation depends not only on the regioselectivity of the first hydroxylation step but also on the preference to use the produced alcohol for the second round of hydroxylation. We previously reported that CYP505A30wt accepts 3- and 4-octanol at lower turnover frequencies compared to 2-octanol.[22] This might be the reason for the low formation of 2,5-octanediol and 2,7-decanediol, even by the 505A30_BM3wt and 1-12G mutants, which showed an increased preference for the ω-3 position compared to the wildtype.

Increasing octanediol production

Previously, biotransformations were done with the addition of *n*-alkane substrates in excess, as their low solubility in water hinders an efficient turnover.[29] Using this two-phase system, higher TOF and TTN of CYP505A30 were achieved, but it resulted in large amounts of alcohols present after 24 h. Per the calculated concentration of products, CYP505A30 can convert more than 20 mM *n*-octane in 24 h. Intrigued by

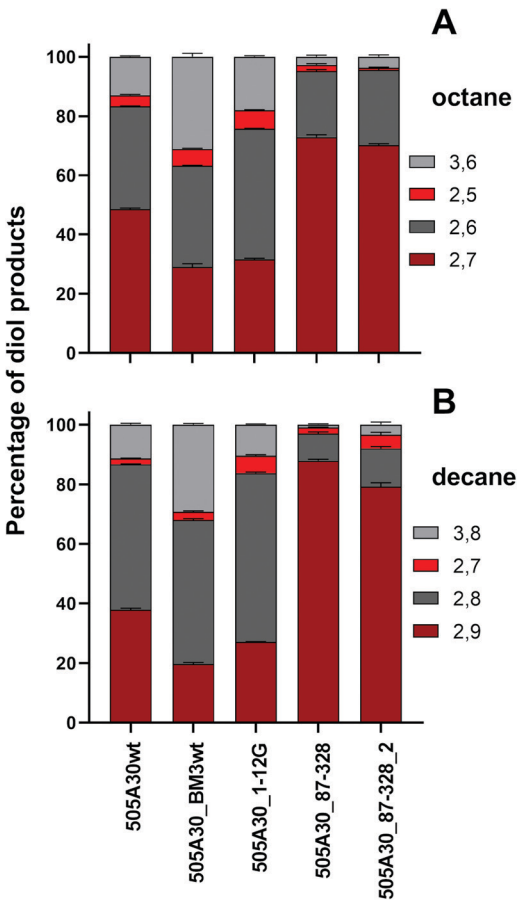

Fig. 4 Diols produced by CYP505A30wt and mutants from *n*-octane (A) and *n*-decane (B). The 8 h and 24 h data points of the CFE biotransformations were used to determine the ratios. Similar trends were observed for the 1 h and 2 h data points. Results are the averages for duplicate reactions of at least two independent experiments.

this, we tested the substrate turnover and product profile ratio of alcohols and diols with a reduced substrate concentration (between 10 and 100 mM *n*-octane) (Fig. 5).

Reducing the *n*-octane concentration in the BRM changed the ratio between alcohol and diol production. CYP505A30 produces mainly alcohols if *n*-octane is added in excess, and diols are only a small percentage of the total products. With a lower substrate concentration, CYP505A30 increasingly hydroxylated the produced alcohols, leading to over 80% diols relative to all products (Fig. 5A).

Furthermore, a higher total diol concentration was obtained after 24 h in contrast to biotransformations with excess *n*-octane. Reducing the substrate concentration from excess to 40 and 50 mM of *n*-octane more than doubled the produced diol concentration in the BRM to over 13 mM (Fig. 5B). The 1 mL reactions with excess octane resulted, on average, in space–time yields (STY) of 4.4 mg mL^{-1} per day while reactions containing 50 mM octane translated to STY of 2.8 mg mL^{-1} per day which was overall in a similar range of what we previously reported.[22]

Lowering the substrate available to CYP505A30, however, significantly reduced its TTN (Fig. 5B). With all substrate

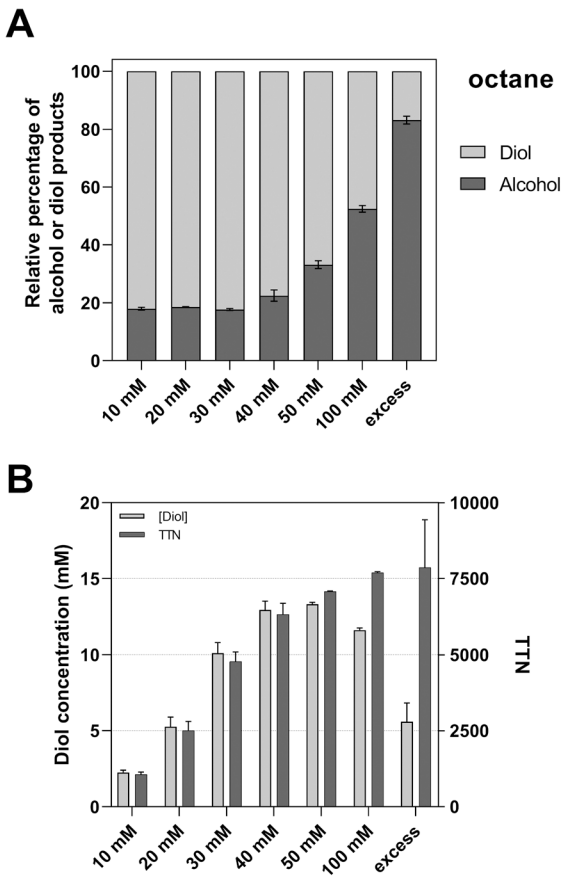

Fig. 5 Relative percentage of alcohol and diol products (A) and total diol concentration and TTN (B) of CFE biotransformations with excess *n*-octane or 10 mM to 100 mM *n*-octane. Biotransformations performed in duplicates at 20 °C.

concentrations, more than 60% of *n*-octane was still unused after 24 h (*ca.* 40% conversion for 50 mM *n*-octane), highlighting the low solubility, and showing the advantage of adding *n*-octane in excess.

Conclusions

In this study, we presented the structure of CYP505A30, which has a high structural similarity to CYP102A1. We showed that it is possible to rationally transfer the knowledge we have on CYP102A1 to evolve the regioselectivity of the related CYP505A30. Introducing the same amino acid mutations into the active site that are found in CYP102 87-328, we increased the regioselectivity for ω-1 (C2) to over 80% and 90% for *n*-octane and *n*-decane, respectively. This was also seen in the formed diols, with 2,7-octanediol and 2,9-decanediol constituting 72% and 88% of the diol products.

Materials and methods

Cloning

The gene encoding CYP505A30 from *Myceliophthora thermophila* was purchased from GenScript (USA). The constructs pET28:CYP505A30 HD and pET22:CYP505A30 with a C-terminal His-Tag were created as previously described.[22] CYP505A30 mutants were obtained using the Megaprimer method[34] with the primers stated in Table S1.† The template DNA was removed with *Dpn*I digestion at 37 °C overnight before the PCR products were purified from a DNA agarose gel and transformed into *E. coli* TOP10 cells (Invitrogen). Single colonies were selected and grown overnight in LB medium containing 100 μg mL^{-1} ampicillin. Plasmid DNA was isolated and verified by DNA sequencing.

Enzyme expression

E. coli BL21-Gold(DE3) (Stratagene) was transformed with the constructs. Positive colonies were selected on LB-agar plates supplemented with the appropriate antibiotic (30 μg mL^{-1} kanamycin or 100 μg mL^{-1} ampicillin) and used for a pre-inoculum that was incubated overnight at 37 °C and 200 rpm. Cells were grown in ZYP-5052 auto-induction medium,[35] containing antibiotic, 0.5 mM *δ*-aminolevulinic acid hydrochloride and 0.05 mM $FeCl_3{\cdot}6H_2O$ (48 h, 25 °C, 200 rpm). The gene encoding GDH from *Bacillus megaterium* (*Bm*GDH) was expressed and purified as previously described.[36] For control reactions, an empty pET22 or pET28: CYP102A1 were used.

Purification of CYP505A30 HD

Harvested cells were resuspended in 50 mM Tris HCl buffer, pH 7.4, 500 mM NaCl, 40 mM imidazole and broken at 30 kpsi by single passage through a One-Shot Cell disrupter system (Constant Systems Ltd). Triton X-100 was added to a concentration of 1% to the crude broken cells to solubilise the membrane-associated protein. The cell debris was removed by centrifugation before the cell-free extract was loaded onto a HisTrap FF 5 ml column. The protein was eluted using a linear gradient of increasing imidazole concentration (40–500 mM). Pure fractions were pooled, concentrated, and desalted by size exclusion chromatography using Sephacryl S100 HR equilibrated with 10 mM Tris, pH 8.0, 100 mM NaCl. The purification of CYP505A30 has been described elsewhere.[22] The method of Omura and Sato was used to determine the concentration of CYP.[37] Protein concentration was quantified with the Pierce BCA assay (Thermo Fisher Scientific, USA) using bovine serum albumin as a standard.

Crystallisation

Screening for suitable crystallisation conditions for purified CYP505A30HD was performed with the Oryx Nano (Douglas Instruments). Sitting drop crystallisation screens from Hampton Research were set up in 96 well plates. 6 and 8 mg mL^{-1} of CYP505A30 HD with 10 mM of C12 fatty acid or alcohol in 100% ethanol was used. Droplets contained a 1 : 1 volume ratio of protein and reservoir solution (50 μL) with a total volume of 2 μL. Screens were incubated at 16 °C. Initial

hits were obtained in various conditions, including 0.1 M BICINE, pH 8.5, 15% PEG1500.

Optimisation of crystallisation hits was done by setting up hanging drops. The best diffracting crystals formed using 6 mg mL^{-1} of CYP505A30HD with 5 mM dodecanoic acid in 0.1 M BICINE, pH 8.0, 10% PEG1500 with a 2 : 1 volume ratio of protein to reservoir solution and a total volume of 3 μL. Crystals were briefly soaked in reservoir solution containing 20% PEG400 before flash-cooling in liquid nitrogen.

Data collection and structure determination

Data was collected at the Diamond Light Source, Didcot, Oxfordshire, UK, beamline i04. The obtained data was processed using the CCP4 suite.[38] Xia2[39] and DIALS[40] were used to auto-index and integrate the diffraction data. For molecular replacement with Phaser,[41] the structure of a CYP450BM3 mutant (PDB code 1ZOA) was used. Model building was done in COOT.[42] Refinement was performed using REFMAC[43] and BUSTER.[44] Improvement of the model and validation was done by using PDB-REDO[45] and the PDB validation server. Data collection and refinement statistics are summarised in Table S2.† The coordinates and structure factors can be accessed in the Protein Data Bank (7P6L). Images were produced using Phenix[46] and PyMOL. RMSD values between the haem domains were calculated while superimposing the protein backbones in PyMOL.

Biotransformations

Cells were harvested, resuspended in 200 mM Tris HCl, pH 8.0, 100 mM glucose, 100 mM glycerol in a ratio of 5 mL per 1 g of cells (wet weight) and broken at 30 kpsi by single passage through the One-Shot Cell disrupter system (Constant Systems Ltd) before centrifuging (12 000 × *g*, 30 min, 4 °C) to obtain the cell-free extract (CFE). Concentration of active CYPs in whole cells and CFE was determined by CO-difference spectra. Biotransformations were performed in 40 mL amber glass vials containing 900 μL CFE, 0.2 U purified *Bm*GDH, 0.3 mM $NADP^+$ in a total reaction volume of 1 mL. The reactions were started with the addition of the substrates (250 μL alkane if added in excess, or 10 to 100 mM of *n*-octane) and incubated shaking at 20 °C, 200 rpm. The reactions were extracted by addition of 2 × 500 μL ethyl acetate containing 2 mM internal standard (1-undecanol or 1-dodecanol). The organic phase was analysed *via* GC-FID and GC–MS analysis (Table S3†). For silylation, 100 μL of the organic phase from the biotransformation extracts were dried under N_2. 50 μL *N*,*O*-bis(trimethylsilyl)acetamide containing 2% (v/v) trimethylchlorosilane was added to the dried samples, mixed and incubated at 70 °C for 3 h and analysed. Alcohol and diol concentrations were derived from standard curves of authentic samples.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research was supported by the DST-NRF Centre of Excellence in Catalysis (c*change, Grant PAR-02) and the National Research Foundation (NRF-SA, Grant 112094). This work was supported through the Global Challenges Research Fund (GCRF) through Science & Technology Facilities Council (STFC), grant number ST/R002754/1: Synchrotron Techniques for African Research and Technology (START). The authors would like to thank Mr. Sarel Marais for his assistance with GC-(MS) analysis. The authors would also like to thank Diamond Light Source for beamtime (proposal mx20303), and the staff of beamline i04 for assistance with data collection. Furthermore, we would like to express our gratitude to the lecturers of the CCP4 Crystallographic School in South Africa – Data Collection to Structure Refinement and Beyond, 2021, for assistance with refining the structure of CYP505A30 HD.

Notes and references

1 R. Bernhardt and V. B. Urlacher, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 6185–6203.
2 W. C. Huang, A. C. G. Westlake, J. D. Maréchal, M. G. Joyce, P. C. E. Moody and G. C. K. Roberts, *J. Mol. Biol.*, 2007, **373**, 633–651.
3 M. Su, S. Chakraborty, Y. Osawa and H. Zhang, *J. Biol. Chem.*, 2020, **295**, 1637–1645.
4 H. Li and T. L. Poulos, *Nat. Struct. Biol.*, 1997, **4**, 140–146.
5 K. Ravichandran, S. Boddupalli, C. Hasermann, J. Peterson and J. Deisenhofer, *Science*, 1993, **261**, 731–736.
6 O. Shoji, T. Fujishiro, K. Nishio, Y. Kano, H. Kimoto, S.-C. Chien, H. Onoda, A. Muramatsu, S. Tanaka, A. Hori, H. Sugimoto, Y. Shiro and Y. Watanabe, *Catal. Sci. Technol.*, 2016, **6**, 5806–5811.
7 M. A. Noble, C. S. Miles, S. K. Chapman, D. A. Lysek, A. C. Mackay, G. A. Reid, R. P. Hanzlik and A. W. Munro, *Biochem. J.*, 1999, **339**, 371.
8 C. J. C. Whitehouse, S. G. Bell and L. L. Wong, *Chem. Soc. Rev.*, 2012, **41**, 1218–1260.
9 A. J. Warman, O. Roitel, R. Neeli, H. M. Girvan, H. E. Seward, S. A. Murray, K. J. McLean, M. G. Joyce, H. Toogood, R. A. Holt, D. Leys, N. S. Scrutton and A. W. Munro, *Biochem. Soc. Trans.*, 2005, **33**, 747–753.
10 A. Glieder, E. T. Farinas and F. H. Arnold, *Nat. Biotechnol.*, 2002, **20**, 1135–1139.
11 M. W. Peters, P. Meinhold, A. Glieder and F. H. Arnold, *J. Am. Chem. Soc.*, 2003, **125**, 13442–13450.
12 E. Weber, A. Seifert, M. Antonovici, C. Geinitz, J. Pleiss and V. B. Urlacher, *Chem. Commun.*, 2011, **47**, 944–946.
13 T. Kitazume, N. Takaya, N. Nakayama and H. Shoun, *J. Biol. Chem.*, 2000, **275**, 39734–39740.
14 P. Durairaj, J. S. Hur and H. Yun, *Microb. Cell Fact.*, 2016, **15**, 125.
15 M. J. L. J. Fürst, B. Kerschbaumer, C. Rinnofer, A. K. Migglautsch, M. Winkler and M. W. Fraaije, *Adv. Synth. Catal.*, 2019, **361**, 2487–2496.

16 G. J. Baker, H. M. Girvan, S. Matthews, K. J. McLean, M. Golovanova, T. N. Waltham, S. E. J. Rigby, D. R. Nelson, R. T. Blankley and A. W. Munro, *ACS Omega*, 2017, **2**, 4705–4724.
17 M. J. Maseme, A. Pennec, J. van Marwijk, D. J. Opperman and M. S. Smit, *Angew. Chem., Int. Ed.*, 2020, **59**, 10359–10362.
18 K. Sakai, F. Matsuzaki, L. Wise, Y. Sakai, S. Jindou, H. Ichinose, N. Takaya, M. Kato, H. Wariishi and M. Shimizu, *Appl. Environ. Microbiol.*, 2018, **84**, 1–15.
19 S. Uhlig, M. Busman, D. S. Shane, H. Rønning, F. Rise and R. Proctor, *J. Agric. Food Chem.*, 2012, **60**, 10293–10301.
20 L. Zong, R. Gao, Z. Guo, Z. Shao, Y. Wang and B. E. Eser, *Biochem. Eng. J.*, 2021, **166**, 107871.
21 S. D. Munday, N. K. Maddigan, R. J. Young and S. G. Bell, *Biochim. Biophys. Acta, Gen. Subj.*, 2016, **1860**, 1149–1162.
22 A. C. Ebrecht, J. C. Aschenbrenner, M. S. Smit and D. J. Opperman, *Org. Biomol. Chem.*, 2021, **19**, 439–445.
23 H. M. Girvan, T. N. Waltham, R. Neeli, H. F. Collins, K. J. McLean, N. S. Scrutton, D. Leys and A. W. Munro, *Biochem. Soc. Trans.*, 2006, **34**, 1173–1177.
24 T. Coleman, J. E. Stok, M. N. Podgorski, J. B. Bruning, J. J. De Voss and S. G. Bell, *J. Biol. Inorg. Chem.*, 2020, **25**, 583–596.
25 T. L. Poulos, *Drug Metab. Rev.*, 2007, **39**, 557–566.
26 O. Gotoh, *J. Biol. Chem.*, 1992, **267**, 83–90.
27 S. Stoll, Y.-T. Lee, M. Zhang, R. F. Wilson, R. D. Britt and D. B. Goodin, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 12888–12893.
28 A. W. Munro, H. M. Girvan, A. E. Mason, A. J. Dunford and K. J. McLean, *Trends Biochem. Sci.*, 2013, **38**, 140–150.
29 A. Pennec, C. L. Jacobs, D. J. Opperman and M. S. Smit, *Adv. Synth. Catal.*, 2015, **357**, 118–130.
30 P. K. Chowdhary, M. Alemseghed and D. C. Haines, *Arch. Biochem. Biophys.*, 2007, **468**, 32–43.
31 M. Dietrich, T. A. Do, R. D. Schmid, J. Pleiss and V. B. Urlacher, *J. Biotechnol.*, 2009, **139**, 115–117.
32 L. Zong, Y. Zhang, Z. Shao, Y. Wang, Z. Guo, R. Gao and B. E. Eser, *Catalysts*, 2021, **11**, 665.
33 K. A. Feenstra, E. B. Starikov, V. B. Urlacher, J. N. M. Commandeur and N. P. E. Vermeulen, *Protein Sci.*, 2007, **16**, 420–431.
34 J. Sanchis, L. Fernández, J. D. Carballeira, J. Drone, Y. Gumulya, H. Höbenreich, D. Kahakeaw, S. Kille, R. Lohmer, J. J. P. Peyralans, J. Podtetenieff, S. Prasad, P. Soni, A. Taglieber, S. Wu, F. E. Zilly and M. T. Reetz, *Appl. Microbiol. Biotechnol.*, 2008, **81**, 387–397.
35 F. W. Studier, *Protein Expression Purif.*, 2005, **41**, 207–234.
36 F. M. Ferroni, C. Tolmie, M. S. Smit and D. J. Opperman, *PLoS One*, 2016, **11**(7), e0160186.
37 T. Omura and R. Sato, *J. Biol. Chem.*, 1964, **239**, 2370–2378.
38 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 235–242.
39 G. Winter, *J. Appl. Crystallogr.*, 2010, **43**, 186–190.
40 G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter and G. Evans, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2018, **74**, 85–97.
41 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.
42 P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.
43 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1997, **53**, 240–255.
44 G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorec, P. Roversi, A. Sharff, O. Smart, C. Vonrhein and T. Womack, in *BUSTER Version 2.10.3*, Global Phasing Ltd., Cambridge, United Kingdom, 2017.
45 R. P. Joosten, J. Salzemann, V. Bloch, H. Stockinger, A.-C. Berglund, C. Blanchet, E. Bongcam-Rudloff, C. Combet, A. L. Da Costa, G. Deleage, M. Diarena, R. Fabbretti, G. Fettahi, V. Flegel, A. Gisel, V. Kasam, T. Kervinen, E. Korpelainen, K. Mattila, M. Pagni, M. Reichstadt, V. Breton, I. J. Tickle and G. Vriend, *J. Appl. Crystallogr.*, 2009, **42**, 376–384.
46 D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2019, **75**, 861–877.