

Cite this: *RSC Chem. Biol.*, 2024,  
5, 1177

# Characterization of the Baeyer–Villiger monooxygenase in the pathway of the bacterial pyrrolizidine alkaloids, legonmycins†

Shan Wang,<sup>id</sup>\*<sup>ab</sup> Fleurdeliz Maglangit,<sup>id</sup><sup>c</sup> Qing Fang,<sup>b</sup> Kwaku Kyeremeh<sup>id</sup><sup>d</sup> and Hai Deng<sup>id</sup>\*<sup>b</sup>

The Baeyer–Villiger monooxygenase (BVMO), LgnC, plays a crucial role in the biosynthesis of bacterial pyrrolizidine alkaloids, legonmycins. It processes bicyclic indolizidine substrates generated from the coordinative action of two non-ribosomal peptide synthetases (LgnB and LgnD) and the standalone type II thioesterase-like enzyme (LgnA). It has been demonstrated that the enzyme selectively inserts molecular oxygen into the carbon–carbon bond adjacent to the carbonyl group in legonindolizidines to form bicyclic 1,3-oxazepine carbamate intermediates. After ring opening and contraction, the most advanced products, prelegonmycins, are formed. However, factors controlling the final hydroxylation step and how the enzyme handles the substrates have remained elusive. In this study, we show that the final hydroxylation at the activated carbon of the electron-rich pyrrole system is attributed to either spontaneous oxidation or the action of an endogenous redox reagent. Substrate docking on the structural model of LgnC combined with site-directed mutagenesis allows the identification of several key amino acids that are essential for substrate/intermediate binding and a mechanism of LgnC-catalysed transformation is proposed.

Received 6th August 2024,  
Accepted 27th September 2024

DOI: 10.1039/d4cb00186a

rsc.li/rsc-chembio

## Introduction

The Baeyer–Villiger (BV) oxidation is a valuable organic transformation that converts ketones and lactones into linear or cyclic esters, respectively, by introducing an oxygen atom, resulting in a carbonyl C–O linkage.<sup>1</sup> In nature, a class of enzymes known as Baeyer–Villiger monooxygenases (BVMOs) also catalyse this reaction.<sup>1–3</sup> These flavin-dependent enzymes facilitate the insertion of an oxygen atom from molecular oxygen. They do this precisely between a carbonylic carbon (Csp<sup>2</sup>) and the adjacent carbon (Csp<sup>3</sup>) of ketones and aldehydes. Moreover, BVMOs exhibit the ability to oxidize molecules containing heteroatoms, such as sulphides, amines, and boron compounds.<sup>4,5</sup> Extensive studies have shown that these enzymes can oxidize a diverse array of substrates with remarkable chemo-

regio-, and enantioselectivity. As a result, they can create intriguing chemical moieties, including carbonates, twisted spiroketal functionalities, and dearomatized products, which are prevalent in various natural product biosynthetic pathways.<sup>6–13</sup>

Pyrrolizidine alkaloids (PAs) are a group of heterocyclic specialised metabolites, featured by two-fused 5-membered rings with a nitrogen atom at the bridgehead. While PAs have been mainly found as plant metabolites, less than 40 PAs have been discovered from bacterial origins, many of which exhibit potent anticancer and antimicrobial activities.<sup>14</sup> The representatives include clazamycins **1**,<sup>15,16</sup> jenamidine **2**,<sup>17</sup> legonmycins **3**,<sup>18</sup> pyrrolizixenamide **4**<sup>19</sup> and azetidomonoamide **5**<sup>20,21</sup> (or azabicyclene) **5** (Fig. 1(A)). Some of these PAs (*i.e.* **1–3**) exist as isomers as they possess an achiral hydroxyl group at the C7 position on the carbon bridgehead of their bicyclic rings (Fig. 1(A)). It has been proposed that, although the hydroxylation at C7 may occur through enzymatic reactions in a stereo-specific manner, the products undergo spontaneous racemization at neutral pH.<sup>15,18</sup>

Recent biosynthetic studies demonstrated that the biosynthesis of bacterial PAs starts with the production of bicyclic [5+6] indolizidine intermediates **6**, resulting from the action of bimodular non-ribosomal peptide synthetases (NRPSs) in most bacterial PA pathways<sup>18–22</sup> (Fig. 1(B)). These indolizidine intermediates are further modified by a group of pathway-specific

<sup>a</sup> State Key Laboratory of Microbial Technology, Shandong University, Qingdao 266237, China. E-mail: shan.wang@sdu.edu.cn

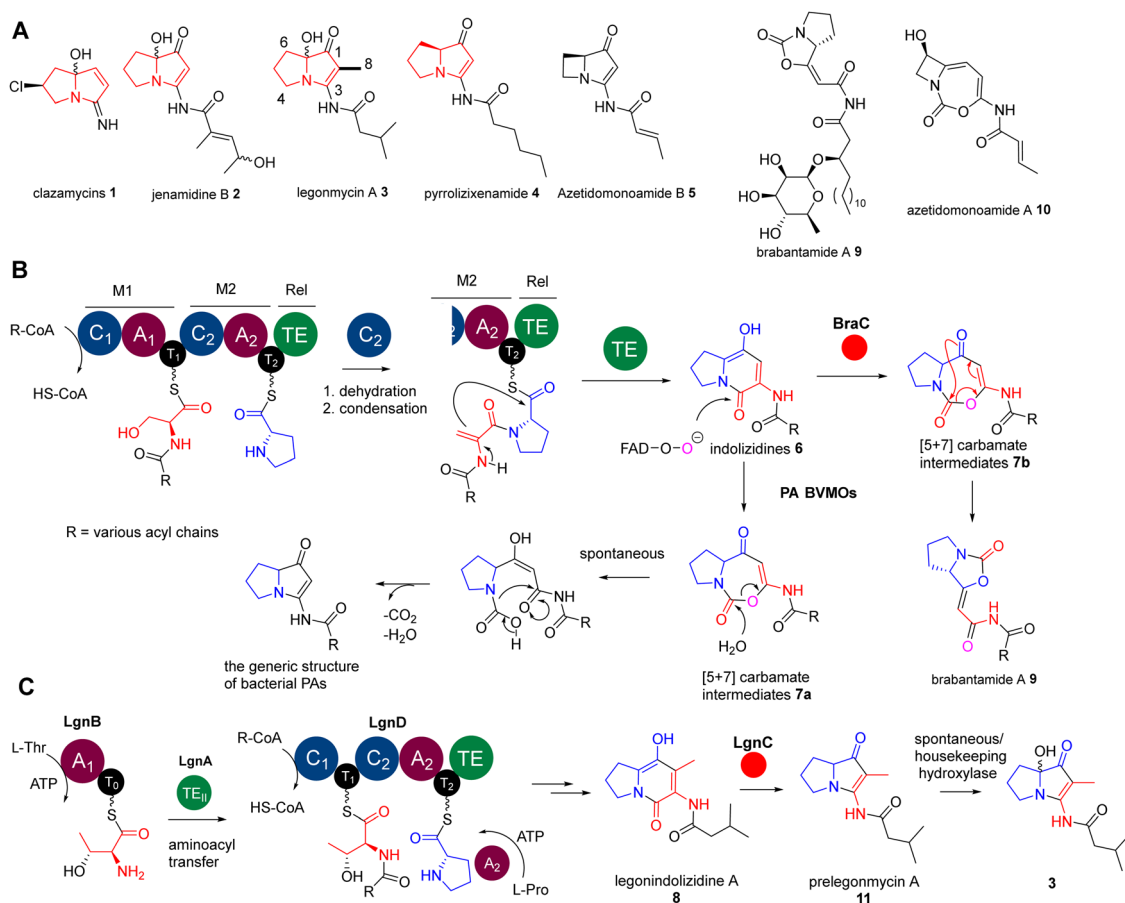
<sup>b</sup> Department of Chemistry, School of Natural and Computing Sciences, University of Aberdeen, Aberdeen AB24 3UE, UK. E-mail: h.deng@abdn.ac.uk

<sup>c</sup> Department of Biology and Environmental Science, College of Science, University of the Philippines Cebu, Lahug, Cebu City 6000, Philippines

<sup>d</sup> Marine and Plant Research Laboratory of Ghana, Department of Chemistry, University of Ghana, P.O. Box LG56, Legon-Accra, Ghana

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4cb00186a>





**Fig. 1** (A) Representative structures of naturally occurring bacterial pyrrolizidine alkaloids (PAs) and biosynthetically related lipocarbamates and cyclocarbamates. PAs **1–3** contain a hydroxyl group at the C7 position. (B) The biosynthesis of bacterial PAs and lipocarbamates only requires two essential enzymes, bimodular NRPS enzymes (C<sub>1</sub>-A<sub>1</sub>-T<sub>1</sub>-C<sub>2</sub>-A<sub>2</sub>-T<sub>2</sub>-Te) and diverged BVMOs, to provide these two distinct scaffolds. (C) A rather unusual pathway of legonmycins which rather require the presence of four enzymes, LgnA-D, to complete the biosynthesis. The atypical TE<sub>II</sub> enzyme domain, LgnA, is essential to catalyse the aminoacyl transfer between two T domains of two NRPSs, LgnB and LgnD.

BVMO enzymes, which catalyse a Baeyer-Villiger ring expansion from **6** to [5+7] 1,3-oxazepine carbamates **7<sup>18</sup>** (Fig. 1(B)). Subsequent non-enzymatic hydrolysis leads to ring opening and decarboxylative contraction, ultimately yielding [5+5] PA ring systems, as demonstrated in a recent report (Fig. 1(B)).<sup>23</sup> In the case of legonmycins, their biosynthesis, however, requires four enzymes, where LgnB and LgnD are NRPSSs, possessing domain arrangements of A<sub>1</sub>-T<sub>0</sub> and C<sub>1</sub>-T<sub>1</sub>-C<sub>2</sub>-A<sub>2</sub>-T<sub>2</sub>-TE, respectively (Fig. 1(C)). It has been demonstrated that LgnA, an unusual type II thioesterase (TE<sub>II</sub>), catalyses the essential aminoacyl translocation from LgnB-T<sub>0</sub> to LgnD-T<sub>1</sub>, facilitating downstream reactions<sup>24</sup> (Fig. 1(C)). Biochemical analysis also showed that the LgnD-C<sub>2</sub> domain, a member of unusual C domains,<sup>25</sup> catalyses the essential dehydration reaction on LgnD-T<sub>1</sub>-tethered acyl-Thr to provide the corresponding acyl-dehydrobutyryne intermediates<sup>24</sup> (Fig. 1(C)).

Interestingly, the pathways of bacterial PAs share the same bimodular NRPS architectures with the lipocyclocarbamates, brabantamides **9**, to provide intermediates with identical [5+6] indolizidine ring systems **7b** (Fig. 1(B)).<sup>26,27</sup> However, a homologue of PA BVMOs, BraC, encoded in the corresponding biosynthetic gene cluster, catalyses a diverse ring re-arrangement

of the [5+7] 1,3-oxazepine carbamates to provide **9** (Fig. 1(B)).<sup>23</sup> In some cases, the [5+7] 1,3-oxazepine carbamates produced by these BVMOs could be trapped or further modified by other enzymes to form stable products, which are then released into the fermentation broth. This can be exemplified in clipibicyclenes **10**,<sup>28</sup> cyclocarbamides,<sup>29</sup> legoncarbamate,<sup>30</sup> and SB-315021<sup>31</sup> (Fig. S1, ESI<sup>†</sup>), which were found to be potent antimicrobial agents. Recently, it has been shown that **10** is a potent covalent inhibitor of bacterial caseinolytic proteases (ClpPs),<sup>28</sup> which are part of bacterial highly conserved proteolytic complexes that are, in many cases, essential for growth.

In this study, we report that the recombinant LgnC can efficiently mediate the biotransformation of legonindolizidine **8** to prelegonmycin **11** (Fig. 1(C)). However, the hydroxylation at the C7 position of **11** to form **3** in our *in vitro* assays likely results from spontaneous oxidations. Assays of legonindolizidine **8** or purified prelegonmycin **11** with *E. coli* whole cells where the recombinant LgnC was overexpressed along with the necessary cofactors, demonstrate that legonmycin **3** is efficiently produced, strongly indicating that the hydroxylation at the C7 position of **11** is mediated by an endogenous



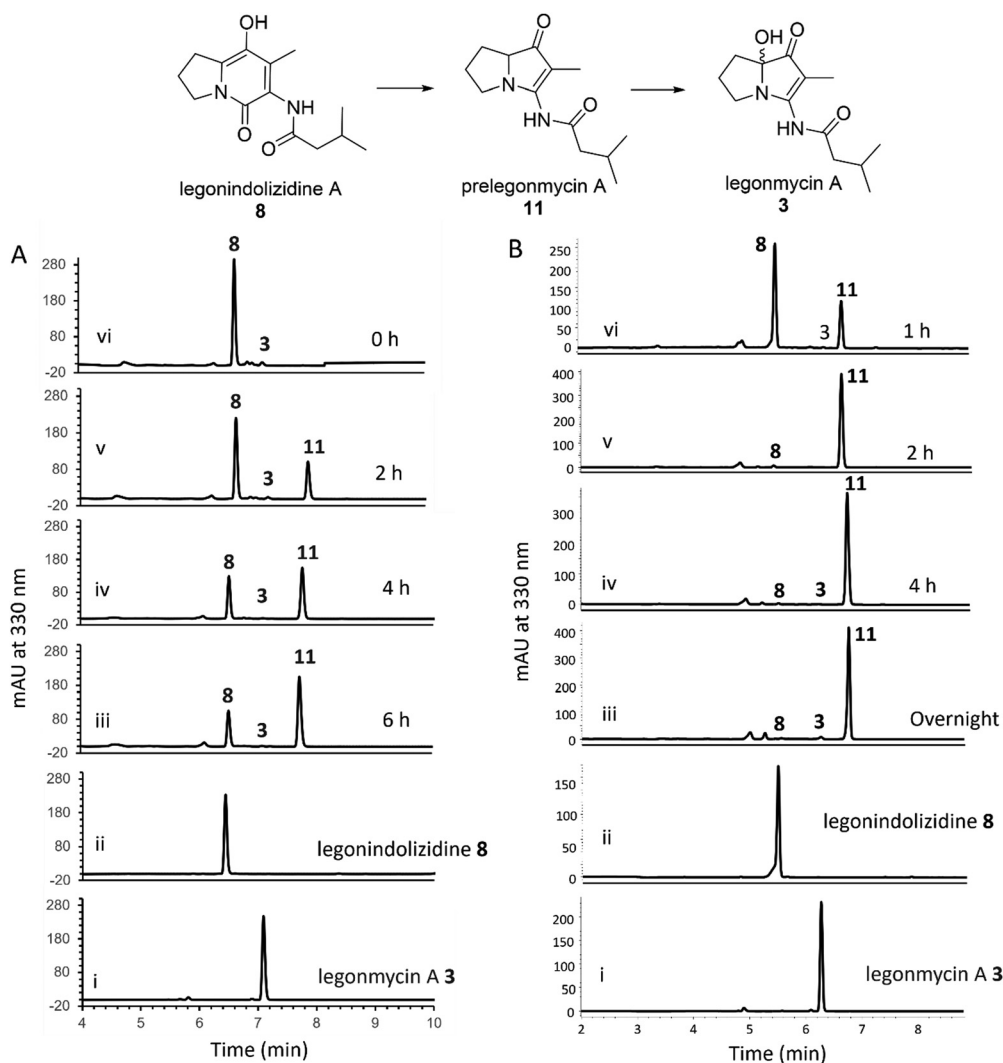
hydroxylase. Structural modelling of LgnC with its cofactor FAD and subsequent substrate docking analysis allow the identification of key amino acid residues in its putative active site. Site-directed mutagenesis (SDM) experiments coupled with bioinformatic analysis indicate that H200, Y219, and W221, which are highly conserved in PA BVMOs, play important roles in the BV biotransformation from **8** to **11**. Finally, a mechanism of LgnC-mediated transformation is proposed.

## Results and discussion

It was demonstrated that the recombinant LgnC with a trigger factor (TF) tag was able to convert legonindolizidine **8** to legonmycin A **3** (Fig. 1). Extracted ion chromatography (EIC) of our LC-MS analysis showed the presence of proposed intermediates during the biotransformation catalysed by LgnC.

However, only a trace amount of legonmycin A **3** was observed.<sup>18</sup>

Efforts were then made to optimize the reaction conditions to improve the efficient biotransformation from **8** to **3**. Incubation with the recombinant LgnC-TF in the presence of **8** (acquired from our previous studies<sup>18</sup>) together with FAD (0.1 mM), NADP<sup>+</sup> (1 mM), and the NADPH recycling system<sup>18</sup> has only resulted in a trace amount of **3** (by HPLC) (Fig. 2(A)) as also observed in MS and MS fragmentation analysis (Fig. S2, ESI<sup>+</sup>). However, a major compound with the same UV absorption has emerged in our HPLC analysis, as indicated in our time course assays (Fig. 2(A)). LC-MS and MS fragmentation analysis indicated that this accumulated compound is prelegonmycin A **11** that lacks the hydroxyl group at the C7 position (Fig. S3, ESI<sup>+</sup>). It is worth noting that there was a large portion of substrate **8** that remained even after extensive reaction time (6 h) (Fig. 2(A)). Although, to the best of our knowledge, there is



**Fig. 2** HPLC traces of the biotransformation from **8** to **3**. (A) HPLC analysis of the conversion of **8** to **11** and **3** using purified recombinant LgnC-TF. Trace i. standard legonmycin A **3**; ii. The standard **8**; iii–vi. The time course assays showing the conversion of **8** to **11** and **3** (0 h, 2 h, 4 h, 6 h). (B) HPLC analysis of the conversion of **8** to **11** and **3** using cell-free extract of M1152-IgnC containing overexpressed recombinant His<sub>6</sub>-LgnC. Trace i. the standard of **3**; ii. The standard of **8**; iii–vi. The time course assay shows the conversion of **8** to **11** and **3** (1 h, 2 h, 4 h, overnight).



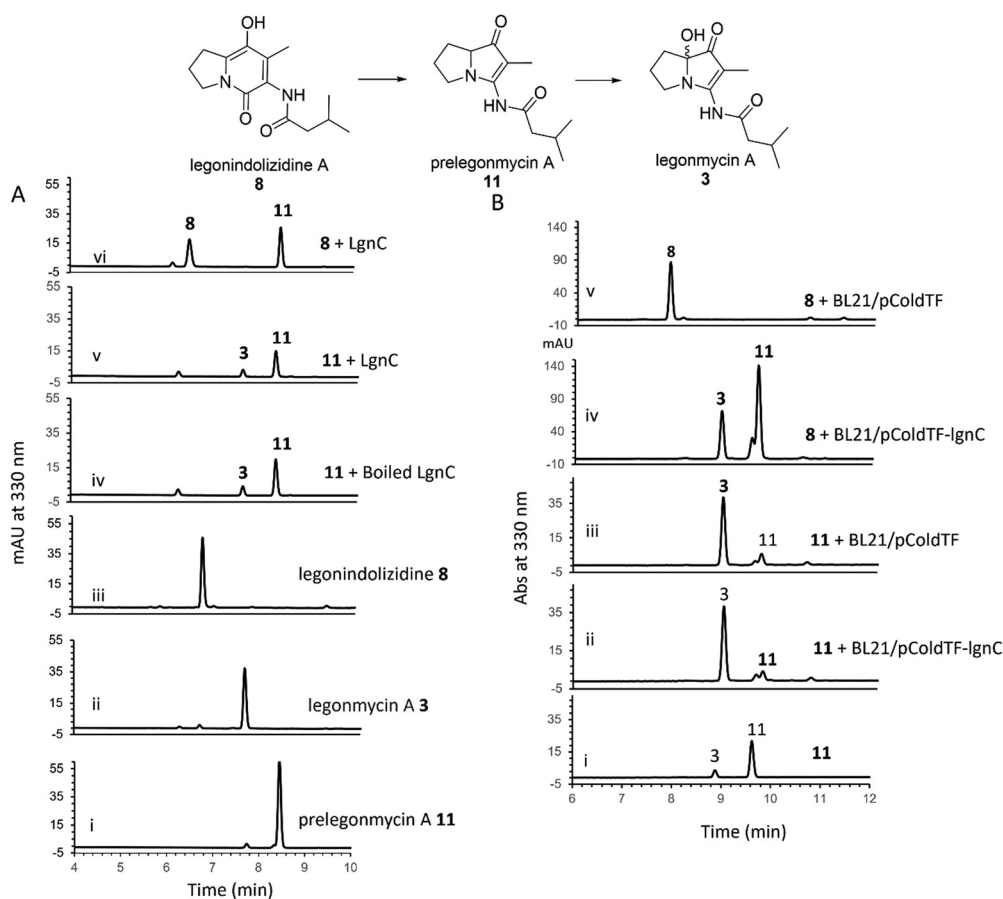
no literature supporting that recombinant proteins with TF tags affect the overall enzymatic activities,<sup>32</sup> attempts were made to remove the TF tag from the recombinant LgnC. However, such cleavage resulted in the complete loss of soluble proteins and enzymatic activities (data not shown).

To investigate this further, we cloned the *lgnC* into pGM1190, a self-replicative shuttle vector with a thiostrepton-inducible promoter PtipA used for gene expression in *Streptomyces*.<sup>33</sup> The subsequently isolated pGM1190-*lgnC* was directly introduced into *Streptomyces coelicolor* M1152, a heterologous host for recombinant protein production,<sup>34</sup> via conjugation for heterologous expression without further modification. The protein extract containing overexpressed recombinant LgnC was used for enzyme assays. Addition of **8** into this protein extract, together with exogenous FAD (0.1 mM) and NADPH (1 mM), resulted in the accumulation of **11** with only a trace amount of **3** in our time course experiments (Fig. 2(B)). It appears that all of **8** was consumed after a 4-h assay as observed in HPLC traces. Interestingly, the NADPH recycling system was no longer required to complete the transformation from **8** to **11**. Similar to what was

observed in the assays of LgnC-TF, only a trace amount of **3** was produced during these time-course experiments. However, purification of recombinant LgnC from the *Streptomyces* host proved to be difficult due to the low expression level, often resulting in protein contamination.

Taken together, it is likely that, while the recombinant LgnC is able to efficiently catalyse the conversion of **8** to **11**, the final hydroxylation in our *in vitro* assays may result from spontaneous oxidation on the activated C7 position of the electron-rich pyrrole scaffolds, similar to the observation found during the synthesis of legonmycins.<sup>35</sup>

To further investigate the hydroxylation, we set out to isolate compound **11** via large-scale biotransformation and multi-rounds of semi-preparative HPLC separations. MS and MS<sup>2</sup> fragmentation analysis confirmed the molecular identity of **11**, displaying a similar UV profile to the ones of **3** (Fig. S4, ESI†). In a controlled experiment where the boiled recombinant LgnC (inactivated enzyme) was used, **11** remained unchanged with a small amount of **3** presented, which likely originated from the spontaneous oxidation of **11** during this overnight incubation (Fig. 3(A) trace iv).



**Fig. 3** (A) HPLC analysis of the assays containing recombinant LgnC-TF with purified **11**. Traces i-iii, standard compounds **11**, **3**, and **8**, respectively. Trace iv. The negative control assay by incubating **11** with boiled LgnC-TF (inactivated). Trace v. The assay by incubating **11** with LgnC-TF. Trace vi. The positive control assay by incubating **8** with LgnC-TF. (B) HPLC analysis of the whole cell assays with either **8** or **11**. Trace i, standard compound **11** with a trace amount of **3**, possibly coming from spontaneous oxidation. Trace ii. The assay of incubating **11** with whole cells containing overexpressed LgnC-TF. Trace iii. The assay of incubating **11** with whole cells only containing empty plasmid. Trace iv. The assay of incubating **8** with whole cells containing overexpressed LgnC-TF. Trace v. The assay of incubating **8** with whole cells containing an empty plasmid.



An assay with recombinant LgnC-TF and **8** showed a similar level of **11** produced compared to the control experiment with a small amount of **3** (Fig. 3(A) trace v). This strongly indicated that LgnC is only responsible for the conversion of **8** to **11** (Fig. 3(A), trace vi), but it is unable to install the hydroxyl group on C7 of **11** to provide **3**. However, the slow spontaneous oxidation rate on **11** observed in our *in vitro* assays was only able to result in a trace amount of **3** and may not be able to account for our previous observations where only the hydroxylated products were isolated in the original producing isolate, *Streptomyces* sp. MA37 (MA37) and the heterologous expression strain, *S. albus*:7G1.<sup>18</sup> This suggests that the final hydroxylation at the C7 position in some bacterial PAs results from an endogenous redox reagent present (*i.e.* a promiscuous hydroxylase or even hydrogen peroxide) in the host cells.

To confirm this, we set out to investigate whether we are able to observe efficient hydroxylation on **3** in the *E. coli* whole cell biotransformation. The benefit of cell-based biotransformation lies in the ability to utilize multistep cascade reactions linked to oxidative pathways, as many biochemicals require consecutive enzymatic modifications for effective synthesis and functionality.<sup>36</sup> Another benefit during the synthesis is that coenzymes can be provided through intracellular metabolism, or reducing potential can be supplied through cofactor generation in cells with high volumetric activity.<sup>36</sup> To this end, we first utilized *E. coli* BL21(DE3) harbouring the overexpression construct pColdTF-LgnC as the catalyst.<sup>18</sup> Overnight incubation of compound **8** in the cell-based system led to its complete conversion to compounds **11** and **3**, as confirmed by HPLC analysis (Fig. 3(B), trace iv). In a control experiment, where BL21(DE3) cells harboring the empty plasmid (BL21/pCold-TF) were used, no conversion of **8** was observed, indicating that the transformation to **11** and **3** requires recombinant LgnC (Fig. 3(B), trace v). Interestingly, when **11** was added to both the BL21/pCold-TF-LgnC (Fig. 3(B), trace ii) and BL21/pCold-TF (Fig. 3(B), trace iii) systems, it was efficiently converted to compound **3**. This demonstrates that the final hydroxylation of **11** to **3** is independent of recombinant LgnC but can be mediated by endogenous redox reagents. The **11**-to-**3** ratio in the cell-based system was approximately 3:1, suggesting that the conversion to **3** is more efficient in the whole-cell system compared to *in vitro* assays.

In contrast, incubation of **8** in cell-free lysates containing overexpressed LgnC-TF resulted primarily in the accumulation of **11**, with only trace amounts of **3** (Fig. S5, ESI<sup>†</sup>). Additionally, no conversion of **8** to **11** or **3** was observed in control lysates from BL21(DE3) containing the empty pColdTF plasmid, in line with previous *in vitro* observations (Fig. 3(A)).

These results suggest that the complete conversion of **8** to **11** is catalyzed by recombinant LgnC, whereas the final hydroxylation step to form **3** is likely driven by an endogenous hydroxylase or hydrogen peroxide produced by *E. coli* BL21(DE3) under aerobic conditions.<sup>37</sup> The diminished production of hydroxylated products in cell-free lysates may be due to reduced activity of this endogenous redox system.

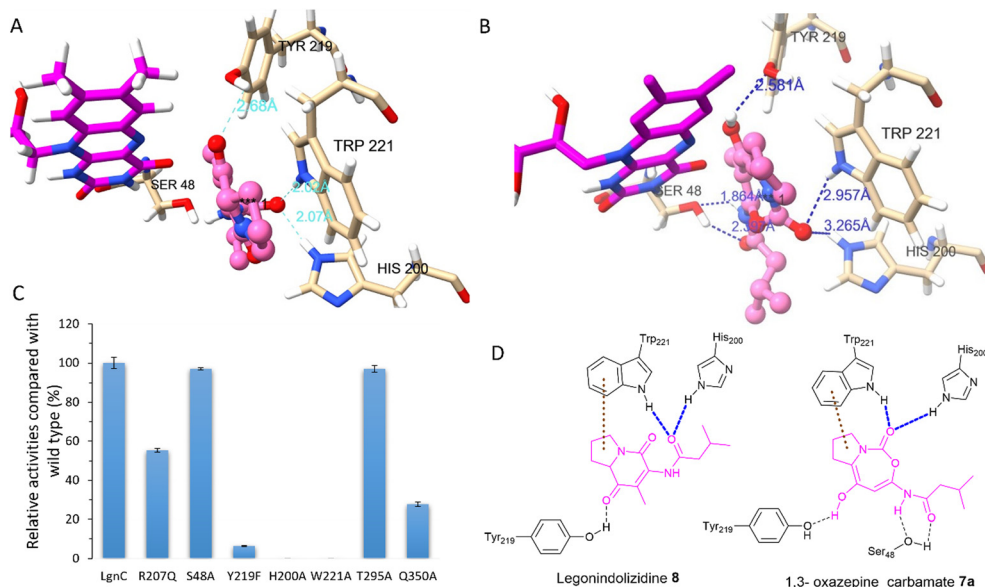
Attempts to crystallize recombinant LgnC-TF were made owing to its unusual multistep catalysis. Despite repeated trials,

we were unable to obtain a crystal structure of recombinant LgnC-TF. Nonetheless, we sought to understand the ability of LgnC to have a bound FAD cofactor by using the recently developed AlphaFold 3 algorithm.<sup>38</sup> The predicted structure of LgnC with a bound FAD cofactor was then submitted to the Dali server<sup>39</sup> for 3D structural comparison. The predicted LgnC structure displays 26% structural similarity (Z-scores of 35.2, rmsd value of 2.8) to AbsH3, an FAD-dependent reductase from the abyssomicin biosynthetic pathway,<sup>40</sup> the highest %id among other structures of FAD-dependent oxidoreductases. Similar to AbsH3, LgnC appears to have two “domains” (Fig. S6, ESI<sup>†</sup>), site A containing residues 5–73, 90–176, 270–345, and 370–381, and site B containing residues 70–90, 177–269 and 346–369 (Fig. S6, ESI<sup>†</sup>). The arrangement of bound FAD in site A is similar to AbsH3 (Fig. S7, ESI<sup>†</sup>). One cavity was predicted by CAVER software<sup>41</sup> in close proximity to the quinone of the FAD cofactor (Fig. S8, ESI<sup>†</sup>). This void appears to link channels to facilitate the transport of substrates and products in site B (Fig. S8, ESI<sup>†</sup>). We then conducted docking studies on legonin-dolizidine **A 8** and the transient 5+7 **7a** to gain insights into how LgnC handles the substrate and intermediate.<sup>42</sup> In substrate **8**, the hydrophobic 5+6 ring system appears to be inserted into a pocket between the indole ring of W221 and the quinone moiety of the FAD cofactor through hydrophobic interaction with W221 (Fig. 4(A)). The oxygen amide at the isovaleryl motif of **8** also engages in hydrogen bonding interactions with H200 and NH of W221 (Fig. 4(A)). Such an insertion of **8** between FAD and W221 is likely to be further supported by the hydrogen bonding interactions between the oxygen at C1 of **8** and the hydroxyl group of Y219 (Fig. 4(A)). While the overall position of 1,3 oxazepine carbamate intermediate **7a** remains similar to **8**, and Y219 also forms H-bonding with the oxygen at C1 of **7a**, but there was variability observed in the docking results in the initial trajectory and binding position of **7a** (Fig. 4(B)). This variability could account for the ring expansion from the original **8** to the transient 1,3-oxazepine carbamate intermediate **7a**. We observed that, while the amide oxygen of the isovaleryl motif is likely to form the H-bonding interaction with S48, NH of W221 and H200 instead forms the H-bonding interactions of the carbonyl oxygen of the carbamate motif (Fig. 4(D)). This suggests that NHs of H200 and W221 may form an oxyanion hole to stabilise the transient intermediate after the nucleophilic attack initiated by the FAD-C4a-hydroperoxide ion for the subsequent structural rearrangement (Fig. S9, ESI<sup>†</sup>).

Interestingly, we also observed that residues R207, T295, and Q350 are in close proximity to the exit tunnel and within 6 Å distance of the H200 and W221 (Fig. S8D, ESI<sup>†</sup>). Such hydrophilic surroundings may position one water molecule close to the carbonyl carbon of 1,3 oxazepine carbamate intermediate **7a** to facilitate the ring opening, followed by the final decarboxylative ring contraction.

To investigate the catalytic roles of the above identified residues, we performed site-directed mutagenesis (SDM). Changing His200 and Trp221 to Ala, resulting in protein variants H200A and W221A (Fig. S10, ESI<sup>†</sup>), respectively, completely abolished the enzymatic activities of LgnC from **8** to **11**





**Fig. 4** A model of LgnC active site bound with FAD cofactor (pink) was generated by AlphaFold 3. (A) Docked legonindolizidine **8** (salmon) interacting with key residues in the active site of LgnC. (B) Docked 1,3-oxazepine carbamate **7a** (salmon) interacting with key residues in the active site of LgnC. (C) The relative enzyme activities of protein variants compared to wild-type LgnC. The biotransformation was measured from legonindolizidine A **8** to prelegonmycin A **11**. (D) A plain structural elucidation of the interactions between LgnC key amino acid residues and legonindolizidine **8** (left) and 1,3-oxazepine carbamate **7a** (right). Blue dashed lines represent hydrogen bonding interactions and brown dashed lines represent hydrophobic interactions.

(Fig. 4(C)). The protein variant, Y219F, lost 95% activity compared to the wild-type (WT) LgnC. While these three residues are absent in other BVMOs, they are highly conserved in other LgnC BVMO homologues (Fig. S11, ESI<sup>†</sup>), strongly suggesting that they indeed are key to the biotransformation of oxygen insertion from [5+6] indolizidines to [5+7] 1,3 oxazepine carbamate intermediates. While protein variants S48A and T295A (Fig. S10, ESI<sup>†</sup>) retained almost the same activity as the WT LgnC, the enzyme activities were reduced to approximately 50% and 30% compared to the WT LgnC in the assays of protein variants R207Q and Q350A (Fig. S10, ESI<sup>†</sup>), respectively. Partial loss of enzymatic activities in protein variants, R207Q and Q350A, suggested that these residues may have some impact on the enzyme activity (possibly water activation).

In conclusion, we characterized the enzyme activity of the BVMO, LgnC, that catalyses the conversion of legonindolizidine A **8** to prelegonmycin A **11**, the last intermediate in the biosynthesis of legonmycin A. The final hydroxylation at the C7 position of **11** is likely to originate from either spontaneous oxidation by hydrogen peroxide or an endogenous hydroxylase in bacterial strains. Structural modelling of LgnC with FAD cofactor and subsequent substrate docking enabled the identification of key amino acid residues in the putative active sites. Our SDM experiments revealed that Y219F, H200A, and W221A play important roles in biotransformation, which are likely to position legonindolizidine A **8** to facilitate the proposed Baeyer–Villiger ring expansion. Further insights into the molecular mechanism will require protein structures of LgnC or its homologues and the protein complexes with the corresponding substrates.

## Materials and methods

### Protein expression and purification

The pCold-TF-lgnC plasmid and its derivatives were individually transformed into *E. coli* BL21 (DE3). Single colonies from each transformation were grown overnight in LB media (5 mL) containing ampicillin (50  $\mu\text{g mL}^{-1}$ ). The overnight culture was transferred to fresh LB medium (500 mL) supplemented with ampicillin (50  $\mu\text{g mL}^{-1}$ ) and cultivated at 37 °C until the cell density reached an OD<sub>600</sub> of 0.6. IPTG was added to a final concentration of 0.1 mM to induce protein expression. Cells were grown for 16–20 h at 16 °C and then harvested by centrifugation at 4 °C. The cell pellets were resuspended in ice-cold lysis buffer (20 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, pH 8.0), and further disrupted by an Ultrasonic Homogenizer JY92-IIN. Then, the supernatant of cell debris was loaded onto the Ni-NTA-affinity column. Bound proteins were eluted with the same Tris-HCl buffer containing different concentrations of imidazole. The desired elution fractions were combined and concentrated using a Centrifugal Filter Unit (Millipore). The purified proteins were stored at –80 °C in storage buffer (100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10% (w/v) glycerol, 1 mM DTT). The protein concentrations were determined on a NanoDrop 2000 by using the corresponding extinction coefficient.

### Activity assay for LgnC and its mutants

Activity of LgnC and its mutant was assayed by monitoring the conversion of substrates into products as analyzed by HPLC. Assays were set up containing 10  $\mu\text{M}$  purified protein, 0.3 mM



substrate, 0.1 mM FAD, 1 mM NADPH, and 15  $\mu\text{L}$  NADPH recycling system<sup>18</sup> in a final volume of 50  $\mu\text{L}$ . After incubating at 28 °C over time, the reactions were quenched by the addition of two equal volumes of methanol, and protein precipitates were removed by centrifugation. The supernatants were then analyzed using an Agilent Technologies 1260 infinity HPLC system equipped with an ACE C18 10  $\mu\text{m}$  10  $\times$  250 mm column and PDA detector monitoring absorbance at 330 nm. Samples were separated on a linear gradient of 5% to 80% methanol (v/v) in water over 20 min at a flow rate of 1 mL min<sup>-1</sup>.

### Streptomyces cell-free assay for LgnC

*S. coelicolor* M1152/pGM1190-lgnC was inoculated into 50 mL of 10.3% YEME with 10  $\mu\text{g mL}^{-1}$  apramycin and cultivated at 30 °C and 180 rpm for 2 days. 50  $\mu\text{L}$  of 25  $\mu\text{g mL}^{-1}$  thiostrepton was then supplemented into the culture to induce the expression of LgnC. After 3-day fermentation, 2 mL of cell culture was collected by centrifugation, resuspended in 800  $\mu\text{L}$  tris-HCl buffer (50 mM tris-HCl, 125 mM NaCl, pH 8.0), and lysed with sonication on ice. The cellular debris was removed from the lysate by centrifugation. A typical 50- $\mu\text{L}$  assay solution contained 40  $\mu\text{L}$  of the above cell-free extract, 0.3 mM substrate, 1 mM NADPH and 0.1 mM FAD. The reactions were performed at 28 °C during the time course experiments. At each time point, the reaction was quenched by the addition of two equal volumes of methanol, and protein precipitates were removed by centrifugation prior to HPLC detection.

### Preparation of prelegonmycin A

Large scale reaction using M1152-lgnC cell-free extract was conducted to obtain a sufficient amount of prelegonmycin A. To initiate the reaction, legonindolizidine (1.7 mg) was added to 20 mL of freshly prepared M1152-lgnC cell-free extract supplemented with 0.5 mM NADPH and 50  $\mu\text{M}$  FAD. After incubating at 28 °C for about 16 h, the reaction was terminated by the addition of an equal volume of methanol, and the protein precipitate was removed by centrifugation. The supernatant was dried and redissolved in methanol. The crude extract was then purified by reverse-phase HPLC (ACE 5  $\mu\text{m}$ , 250  $\times$  10 mm) on a linear gradient of 40% to 76% methanol over 18 min at a flow rate of 2 mL min<sup>-1</sup>, to give prelegonmycin A ( $t_{\text{R}}$  = 14.7 min, 1.2 mg).

### Whole cell reaction assay for LgnC

1 mL overnight culture of BL21/pCold-TF-lgnC was transferred into 50 mL SOB with 50  $\mu\text{g mL}^{-1}$  Amp and cultivated at 37 °C, 180 rpm. When the OD<sub>600</sub> of the medium reached 0.6, IPTG was added to a final concentration of 0.1 mM, and the culture was further incubated at 16 °C for 18 h. 2 mL of the culture was collected by centrifugation for each reaction and washed with 1 mL of potassium phosphate buffer (100 mM, pH 7.4) three times. The washed cells were resuspended in 500  $\mu\text{L}$  of potassium phosphate buffer and supplemented with legonindolizidine prior to the incubation at 28 °C. At the end of the incubation period, cells were removed by centrifugation. The supernatant was extracted twice with ethyl acetate, the

combined organic layers of which were dried and dissolved in 200  $\mu\text{L}$  of methanol.

### Site-directed mutagenesis of recombinant LgnC-TF

Site-directed mutagenesis of LgnC-TF was performed using an In-Fusion<sup>®</sup> HD Cloning kit (Takara Bio). PCR primer couples containing the desired mutant site were designed to amplify two DNA fragments that overlap each other by 15–20 bp from the template pColdTF-lgnC, respectively (see ESI,<sup>†</sup> Table S1 for primer sequences), and cloned into the BamHI/HindIII site of pET-28a (+) vector. All constructed plasmids were verified by DNA sequencing.

### HR-MS analysis

High-resolution electrospray ionisation mass spectrometry (HR-ESIMS) was determined using LC MS Thermo Scientific MS system (LTQ Orbitrap) coupled to a Thermo Instrument HPLC system (Accela PDA detector, Accela PDA autosampler, and Accela Pump, C18 Sunfire 150  $\times$  46 mm Waters<sup>®</sup>). The following parameters were used: capillary voltage 45 V, capillary temperature 320 °C, auxiliary gas flow rate 10–20 arbitrary units, sheath gas flow rate 40–50 arbitrary units, spray voltage 4.5 kV, mass range 100–2000 amu (maximum resolution 30 000). Additionally, the samples were run on an Agilent 6200 series TOF/6500 series Q-TOF instrument with a scan rate of 1.5 Hz, scan range from 100–1000, variable CID energy, 3.5 kV source voltage, fragmentor 175 V ( $\pm 200\%$ ) and reference masses (121.05087 and 922.00979) enabled. The instruments were optimised as required.

### LgnC modelling and substrate docking

The protein sequence of LgnC (UniPort: A0A0U1XUB1) was supplied to the AlphaFold 3 server to generate a structural prediction bound with one molecule of FAD cofactor using the default parameters. Legonindolizidine A and 1,3 oxazepine carbamate models were generated from SMILES strings at the CACTUS server (<https://cactus.nci.nih.gov/translate/>). Substrate/intermediate and protein models were prepared for docking and used in docking experiments with AutoDock Vina<sup>43</sup> using the default parameters.

### Data availability

Data for this manuscript is available within the text or the ESI.<sup>†</sup>

### Conflicts of interest

There are no conflicts to declare.

### Acknowledgements

This work received financial support from Biotechnology and Biological Sciences Research Council UK (S. W. and H. D., BB/P00380X/1 and BB/R00479X/1), Leverhulme Trust-Royal Society Africa award (AA090088) and the jointly funded UK Medical



Research Council-UK Department for International Development (MRC/DFID) Concordat agreement African Research Leaders Award (MR/S00520X/1 to KK and HD), and National Natural Science Foundation of China (31929001 to HD), and the Royal Society-NSFC Newton Mobility Grant Award (IEC\NSFC\ 211349 to HD). HD and SW thank Dr Juan-Pablo Gomez-Escribano and Professor Mervyn Bibb in John Innes Centre, UK for the gift of *Streptomyces coelicolor* M1152. S. W. thanks Qilu Youth Scholar Startup Funding of Shandong University and the National Natural Science Foundation of China (No. 22307066).

## References

- M. J. L. J. Fürst, A. Gran-Scheuch, F. S. Aalbers and M. W. Fraaije, *ACS Catal.*, 2019, **9**, 11207–11241.
- H. Leisch, K. Morley and P. C. K. Lau, *Chem. Rev.*, 2011, **111**, 4165–4222.
- C. Tolmie, M. S. Smit and D. J. Opperman, *Nat. Prod. Rep.*, 2019, **36**, 326–353.
- Y. Li, X. Yang, Z. Deng and D. Zhu, *Chin. J. Biotechnol.*, 2019, **35**, 351–362.
- K. Balke, A. Beier and U. T. Bornscheuer, *Biotechnol. Adv.*, 2018, **36**, 247–263.
- M. C. Tang, H. Y. He, F. Zhang and G. L. Tang, *ACS Catal.*, 2013, **3**, 444–447.
- Y. Matsuda, T. Awakawa, T. Wakimoto and I. Abe, *J. Am. Chem. Soc.*, 2013, **135**, 10962–10965.
- N. Tibrewal, P. Pahari, G. Wang, M. K. Kharel, C. Morris, T. Downey, Y. Hou, T. S. Bugni and J. Rohr, *J. Am. Chem. Soc.*, 2012, **134**, 18181–18184.
- Y. Hu, D. Dietrich, W. Xu, A. Patel, J. A. J. Thuss, J. Wang, W. B. Yin, K. Qiao, K. N. Houk, J. C. Vederas and Y. Tang, *Nat. Chem. Biol.*, 2014, **10**, 552–554.
- Y. Matsuda, T. Iwabuchi, T. Fujimoto, T. Awakawa, Y. Nakashima, T. Mori, H. Zhang, F. Hayashi and I. Abe, *J. Am. Chem. Soc.*, 2016, **138**, 12671–12677.
- M. Gibson, M. Nur-E-Alam, F. Lipata, M. A. Oliveira and J. Rohr, *J. Am. Chem. Soc.*, 2005, **127**, 17594–17595.
- Y. Katsuyama, K. Sone, R. Satou, M. Izumikawa, M. Takagi, M. Fujie, N. Satoh, K. Shin-Ya and Y. Ohnishi, *ChemBioChem*, 2016, **17**, 1021–1028.
- D. R. Jackson, X. Yu, G. Wang, A. B. Patel, J. Calveras, J. F. Barajas, E. Sasaki, M. Metsä-Ketelä, H. W. Liu, J. Rohr and S. C. Tsai, *ACS Chem. Biol.*, 2016, **11**, 1137–1147.
- J. Robertson and K. Stevens, *Nat. Prod. Rep.*, 2017, **34**, 62–89.
- D. D. Buechter and D. E. Thurston, *J. Nat. Prod.*, 1987, **50**, 360–367.
- Y. Horiuchi, S. Kondo, T. Ikeda, D. Ikeda, K. Mirura, M. Hamada, T. Takeuchi and H. Umezawa, *J. Antibiot.*, 1979, **32**, 762.
- B. B. Snider and J. R. Duvall, *Org. Lett.*, 2005, **7**, 4519–4522.
- S. Huang, J. Tabudravu, S. S. Elsayed, J. Travert, D. Peace, M. H. Tong, K. Kyeremeh, S. M. Kelly, L. Trembleau, R. Ebel, M. Jaspars, Y. Yu and H. Deng, *Angew. Chem., Int. Ed.*, 2015, **127**, 12888–12892.
- O. Schimming, V. L. Challinor, N. J. Tobias, H. Adihou, P. Grün, L. Pöschel, C. Richter, H. Schwalbe and H. B. Bode, *Angew. Chem., Int. Ed.*, 2015, **127**, 12893–12896.
- Z. Hong, A. Bolard, C. Giraud, S. Prévost, G. Genta-Jouve, C. Deregnacourt, S. Häussler, K. Jeannot and Y. Li, *Angew. Chem., Int. Ed.*, 2019, **58**, 3178–3182.
- J. B. Patteson, A. R. Lescalette and B. Li, *Org. Lett.*, 2019, **21**, 4955–4959.
- L. Liu, S. Li, R. Sun, X. Qin, J. Ju, C. Zhang, Y. Duan, Y. Duan, Y. Duan, Y. Huang and Y. Huang, *Org. Lett.*, 2020, **22**, 4614–4619.
- M. Einsiedler, K. Lamm, J. F. Ohlrogge, S. Schuler, I. J. Richter, T. Lübken and T. A. M. Gulder, *J. Am. Chem. Soc.*, 2024, **146**, 16203–16212.
- S. Wang, W. D. G. Brittain, Q. Zhang, Z. Lu, M. H. Tong, K. Wu, K. Kyeremeh, M. Jenner, Y. Yu, S. L. Cobb and H. Deng, *Nat. Commun.*, 2022, **13**, 62, DOI: [10.1038/s41467-021-27512-0](https://doi.org/10.1038/s41467-021-27512-0).
- S. Wang, K. Wu, Y. J. Tang and H. Deng, *Nat. Prod. Rep.*, 2023, **41**, 273–297.
- Y. Schmidt, M. Van Der Voort, M. Crüsemann, J. Piel, M. Josten, H. G. Sahl, H. Miess, J. M. Raaijmakers and H. Gross, *ChemBioChem*, 2014, **15**, 259–266.
- C. W. Johnston, R. Zvanych, N. Khyzha and N. A. Magarvey, *ChemBioChem*, 2013, **14**, 431–435.
- E. J. Culp, D. Sychantha, C. Hobson, A. C. Pawlowski, G. Prehna and G. D. Wright, *Nat. Microbiol.*, 2022, **7**, 451–462.
- A. Isogai, S. Sakuda, K. Shindo, S. Watanabe, A. Suzuki, S. Fujita and T. Furuya, *Tetrahedron Lett.*, 1986, **27**, 1161–1164.
- Q. Fang, L. Wu, C. Urwald, M. Mugat, S. Wang, K. Kyeremeh, C. Philips, S. Law, Y. Zhou and H. Deng, *Synth. Syst. Biotechnol.*, 2021, **6**, 12–19.
- J. J. Thirkettle, E. Alvarez, H. Boyd, M. Brown, E. Dlez, J. Hueso, S. Elson, M. Fulston, C. Gershtater, M. L. Morata, P. Perez, S. Ready, J. Maria Sanchez-Puelles, R. Sheridan, A. Stefanska and S. Warr, *J. Antibiot.*, 2000, **53**, 664–669.
- K. Wu, T. C. Minshull, S. E. Radford, A. N. Calabrese and J. C. A. Bardwell, *Nat. Commun.*, 2022, **13**, 4126.
- M. P. Andreas and T. W. Giessen, *MethodsX*, 2022, **16**, 101787.
- J. P. Gomez-Escribano and M. J. Bibb, *Microb. Biotechnol.*, 2011, **4**, 207–215.
- W. J. M. Lewis, D. M. Shaw and J. Robertson, *Beilstein J. Org. Chem.*, 2021, **17**, 334–342.
- B. Lin and Y. Tao, *Microb. Cell Fact.*, 2017, **16**, 106.
- L. C. Seaver and J. A. Imlay, *J. Bacteriol.*, 2001, **183**, 7171–7181.
- J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstern, D. A. Evans, C. C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen,



- E. D. Zhong, M. Zielinski, A. Židek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper, *Nature*, 2024, **630**, 493–500.
- 39 L. Holm, A. Laiho, P. Törönen and M. Salgado, *Protein Sci.*, 2023, **32**, e4519, DOI: [10.1002/pro.4519](https://doi.org/10.1002/pro.4519).
- 40 J. A. Clinger, X. Wang, W. Cai, Y. Zhu, M. D. Miller, C. G. Zhan, S. G. Van Lanen, J. S. Thorson and G. N. Phillips, *Proteins: Struct., Funct., Bioinf.*, 2021, **89**, 132–137.
- 41 E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor and J. Damborsky, *PLoS Comput. Biol.*, 2012, **8**, e1002708, DOI: [10.1371/journal.pcbi.1002708](https://doi.org/10.1371/journal.pcbi.1002708).
- 42 J. Lyu, N. Kapolka, R. Gumpper, A. Alon, L. Wang, M. K. Jain, X. Barros-Álvarez, K. Sakamoto, Y. Kim, J. DiBerto, K. Kim, I. S. Glenn, T. A. Tummino, S. Huang, J. J. Irwin, O. O. Tarkhanova, Y. Moroz, G. Skiniotis, A. C. Kruse, B. K. Shoichet and B. L. Roth, *Science*, 2024, **384**, eadn6354, DOI: [10.1126/science.adn6354](https://doi.org/10.1126/science.adn6354).
- 43 J. Eberhardt, D. Santos-Martins, A. F. Tillack and S. Forli, *J. Chem. Inf. Model.*, 2021, **61**, 3891–3898, DOI: [10.1021/acs.jcim.1c00203](https://doi.org/10.1021/acs.jcim.1c00203).

