

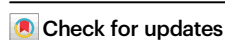


# O-methyltransferase-like enzyme catalyzed diazo installation in polyketide biosynthesis

Received: 20 December 2022

Accepted: 17 August 2023

Published online: 04 September 2023



Yuchun Zhao<sup>1,6</sup>, Xiangyang Liu<sup>1,6</sup>, Zhihong Xiao<sup>1</sup>, Jie Zhou<sup>1</sup>, Xingyu Song<sup>2</sup>, Xiaozheng Wang<sup>1</sup>, Lijun Hu<sup>3</sup>, Ying Wang<sup>3</sup>, Peng Sun<sup>4</sup>, Wenning Wang<sup>2</sup>, Xinyi He<sup>1</sup>, Shuangjun Lin<sup>1</sup>, Zixin Deng<sup>1</sup>, Lifeng Pan<sup>5</sup> & Ming Jiang<sup>1</sup>✉

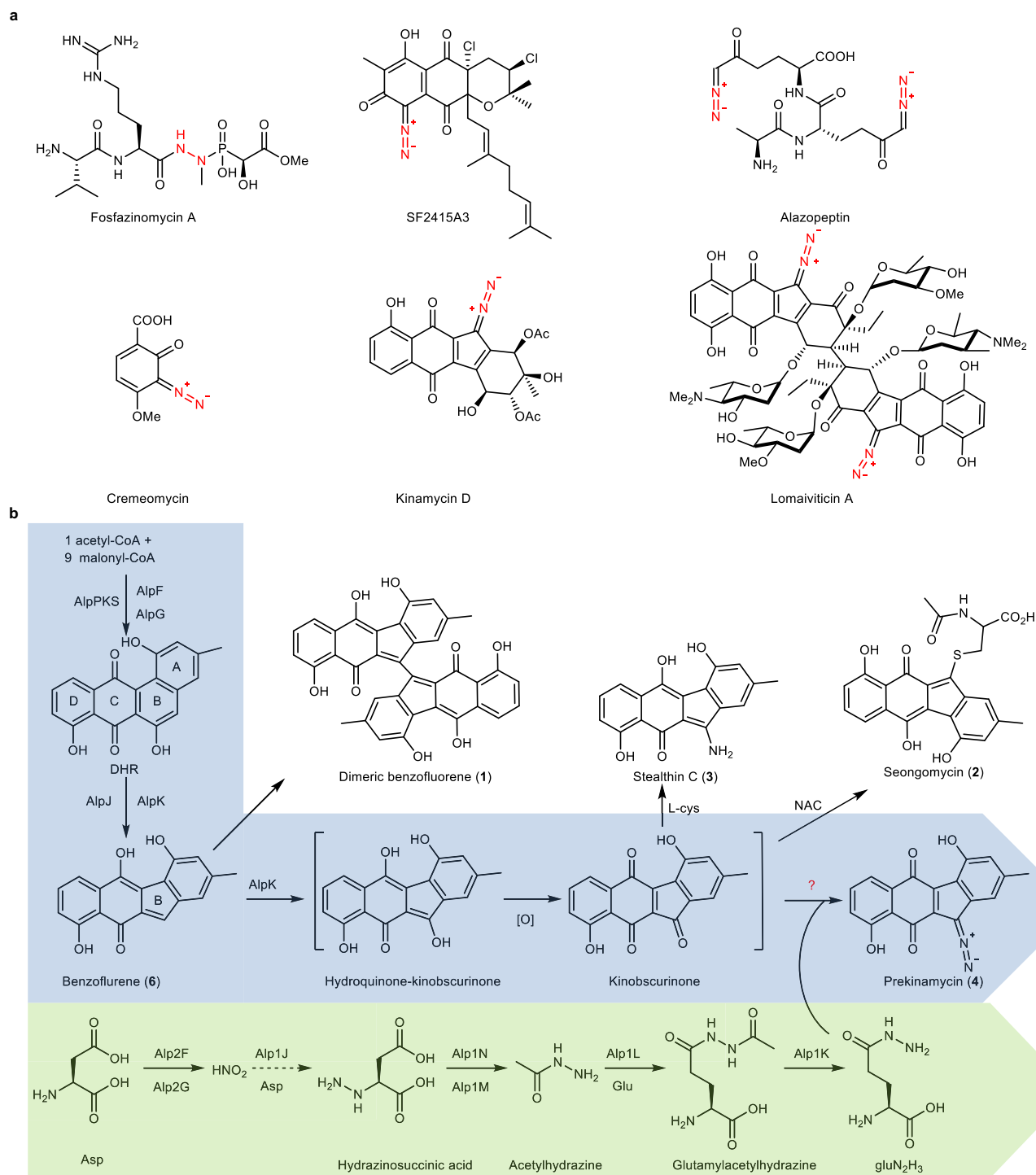
Diazo compounds are rare natural products possessing various biological activities. Kinamycin and lomaiviticin, two diazo natural products featured by the diazobenzofluorene core, exhibit exceptional potency as chemotherapeutic agents. Despite the extensive studies on their biosynthetic gene clusters and the assembly of their polyketide scaffolds, the formation of the characteristic diazo group remains elusive. L-Glutamylhydrazine was recently shown to be the hydrazine donor in kinamycin biosynthesis, however, the mechanism for the installation of the hydrazine group onto the kinamycin scaffold is still unclear. Here we describe an *O*-methyltransferase-like protein, AlpH, which is responsible for the hydrazine incorporation in kinamycin biosynthesis. AlpH catalyses a unique SAM-independent coupling of L-glutamylhydrazine and polyketide intermediate via a rare Mannich reaction in polyketide biosynthesis. Our discovery expands the catalytic diversity of *O*-methyltransferase-like enzymes and lays a strong foundation for the discovery and development of novel diazo natural products through genome mining and synthetic biology.

N–N bond-containing natural products (NPs) are a group of important molecules with enormous structural diversities. More than 300 such NPs with different N–N bonds, including diazo, hydrazide, azoxy, hydrazine, hydrazone, and pyridazine, have been discovered and characterized in the past few decades<sup>1</sup>. These NPs attract numerous research attentions due to their structural diversities and broad spectrum of biological activities, including antifungal, antiviral, anti-tumor, and other activities<sup>2</sup>. In contrast, the biosynthetic machineries of N–N bond formation remain to be elucidated until recently. Over the past ten years, our understanding of enzymatic N–N bond formation has significantly improved due to the uncovering of an increasing number of biosynthetic gene clusters (BGCs) for N–N bond containing NPs<sup>3–11</sup>. The nitrogen donors and the mechanisms for the installation of

the N–N bonds in those NPs have been elucidated through detailed genetic and biochemical studies. To overcome the challenges in generating N–N bond chemically, nature has evolved several different biosynthetic strategies, including the activation of amine through an electrophilic intermediate, utilizing nitrite as a nitrogen donor, and radical-mediated nonenzymatic mechanism<sup>1,12</sup>.

The diazo group is an important functional group in organic synthesis but rarely presents in NPs (Fig. 1a). Previous biosynthetic studies in diazo group formation suggested a stepwise nitrogen incorporation involving a diazotization process<sup>13</sup>. Nitrate has long been proposed to be involved in quite a few N–N bond-containing NP biosynthetic processes. Recent studies in the diazoquinone NP, cremeomycin, revealed an unprecedented nitrous acid biosynthetic pathway

<sup>1</sup>State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic & Developmental Sciences, and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 200030 Shanghai, P. R. China. <sup>2</sup>Ministry of Education Key Laboratory of Computational Physical Sciences, Department of Chemistry, Institutes of Biomedical Sciences, Fudan University, 200438 Shanghai, China. <sup>3</sup>Guangdong Province Key Laboratory of Pharmacodynamic Constituents of TCM and New Drugs Research, Center for Bioactive Natural Molecules and Innovative Drugs Research, Jinan University, 510632 Guangzhou, P. R. China. <sup>4</sup>School of Pharmacy, Second Military Medical University, 325 Guo-He Road, 200433 Shanghai, P. R. China. <sup>5</sup>State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, Chinese Academy of Sciences, 200032 Shanghai, China. <sup>6</sup>These authors contributed equally: Yuchun Zhao, Xiangyang Liu. ✉e-mail: [jiangming9722@sjtu.edu.cn](mailto:jiangming9722@sjtu.edu.cn)



**Fig. 1 | Selected natural products containing nitrogen–nitrogen bonds and key steps of the prekinamycin biosynthesis. a** Selected natural products containing nitrogen–nitrogen bonds. The nitrogen–nitrogen bonds are highlighted in red.

**b** Proposed pathway for the biosynthesis of prekinamycin. Proposed pathways for the biosynthesis of polyketide and gluN<sub>2</sub>H<sub>3</sub> are highlighted in blue and green.

responsible for the formation of the diazo group<sup>14,15</sup>. Nitrous acid is generated from aspartic acid catalyzed by a flavin-dependent monooxygenase (CreE) and a nitrosuccinate lyase (CreD), and subsequently reacts with the primary aromatic amine of 3-amino-2-hydroxy-4-methoxybenzoic acid to form the diazo group spontaneously in an acidic condition or enzymatically by an ATP-dependent diazo-forming protein, CreM.

The kinamycins and lomaiviticins form a small group of NPs featuring diazobenzofluorene core structures<sup>16</sup>. Their biosynthesis is of

intense interest due to their impressive biological activities<sup>17</sup>. The BGCs of kinamycin<sup>18–20</sup> and lomaiviticin<sup>21</sup> have been characterized, and the heterologous production of kinamycin propels the genetic and biochemical studies of the biosynthesis of kinamycin family NPs. The biosynthetic machinery for constructing the benzofluorene skeleton has been proposed and biochemically verified<sup>22</sup> (Fig. 1b). During the assembly of the benzofluorene core, two FAD-dependent enzymes, AlpJ and AlpK, were identified to be involved in the contraction of the B ring in the kinamycin biosynthesis<sup>22, 23</sup>. AlpJ belongs to a unique family of

FAD-dependent oxygenase and has been shown to catalyze a Bayer-Villiger-type oxidative C-C bond cleavage of dehydrabelomycin (DHR), forming of an unstable aldehyde/acid intermediate, which undergoes decarboxylative ring closure and dehydration to give the benzofluorene ring. AlpK could possibly supply necessary FADH<sub>2</sub> for AlpJ catalyzed reaction. And it was also proposed that AlpK might catalyze the C5-hydroxylation reaction of the benzofluorene intermediate. Meanwhile, the enzymes involved in the late stage tailoring steps, such as the acetylation<sup>18</sup> and epoxide hydrolyzation<sup>20</sup> on A ring, have been identified and characterized. However, there is still a knowledge gap in the biosynthesis of the diazo group in kinamycin and lomaiviticin. Previous feeding experiments suggested that kinobscurinone<sup>13</sup> and stealthin C<sup>24</sup> isolated from *Streptomyces murayamaensis* mutants are the biosynthetic intermediates of kinamycin, indicating a step-by-step nitrogen incorporation in diazo formation analogous to the biosynthetic scenario of cremeomycin<sup>15</sup> and SF2415A3<sup>25</sup>. However, this hypothesis is challenged by recent studies. The biosynthetic studies of the N-N bond formation in two structurally unrelated NPs, fosfazinomycin and kinamycin, provided more information about the details of the N-N bond formation in kinamycin family NPs. The N-N bond in kinamycin was confirmed to be constructed separately, followed by incorporation into the polyketide scaffold through a convergent biosynthetic process reminiscent of the strategy used in organic synthesis<sup>26</sup>. Particularly, L-glutamylhydrazine (gluN<sub>2</sub>H<sub>3</sub>) possessing the N-N bond synthon was identified as the essential biosynthetic intermediate in kinamycin biosynthesis through in-depth biochemical assays and feeding experiments.

Except for the N-N bond formation step, gluN<sub>2</sub>H<sub>3</sub> biosynthesis has been well characterized<sup>26</sup> (Fig. 1b). Briefly, aspartic acid is oxidized by FzmM/Alp2F to generate nitrosuccinate, which is cleaved by FzmL/Alp2G to release nitrous acid. Subsequently, the nitrous acid is proposed to be incorporated into aspartic acid through the N-N bond to generate hydrazinosuccinic acid by FzmP/AlpI which still lacks biochemical and genetic evidence. The hydrazinosuccinic acid is acetylated by FzmQ/AlpIN to produce N-acetylhydrazinosuccinic acid, which is cleaved by FzmR/AlpIM to release acetylhydrazine. Eventually, the acetylhydrazine is transferred onto glutamic acid by FzmN/AlpIL to produce glutamylacetylhydrazine, which is further deacetylated to generate gluN<sub>2</sub>H<sub>3</sub> by FzmO/AlpIK. However, the enzyme(s) and mechanism involved in the installation of gluN<sub>2</sub>H<sub>3</sub> or the hydrazine onto the polyketide scaffold remain unknown. The putative enzyme responsible for the coupling reaction is ambiguous in kinamycin BGC. A conserved amidotransferase-like enzyme, Alp1W, was proposed to transfer the hydrazine synthon to the polyketide scaffold without any genetic or biochemical evidence<sup>26</sup>.

In this work, we reconstitute the production of prekinamycin (4), the known inline diazo-containing intermediate in kinamycin biosynthesis, in vitro successfully for the first time and characterize an important benzofluorene intermediate, compound 5, a condensation product of gluN<sub>2</sub>H<sub>3</sub> and the polyketide intermediate (7). Particularly, the O-methyltransferase (OMT)-like enzyme, AlpH, from the kinamycin BGC is determined to mediate the C-N bond formation through loading gluN<sub>2</sub>H<sub>3</sub> to the polyketide scaffold. In vitro studies show that compound 5 is produced from DHR through a cooperative catalysis by AlpJ, AlpK, and AlpH. This study provides significant insights into the installation of the diazo group in kinamycin family NPs, a biomarker for in silico genome mining of putative diazo NPs and a biobrick to generate novel diazo NPs through synthetic biology.

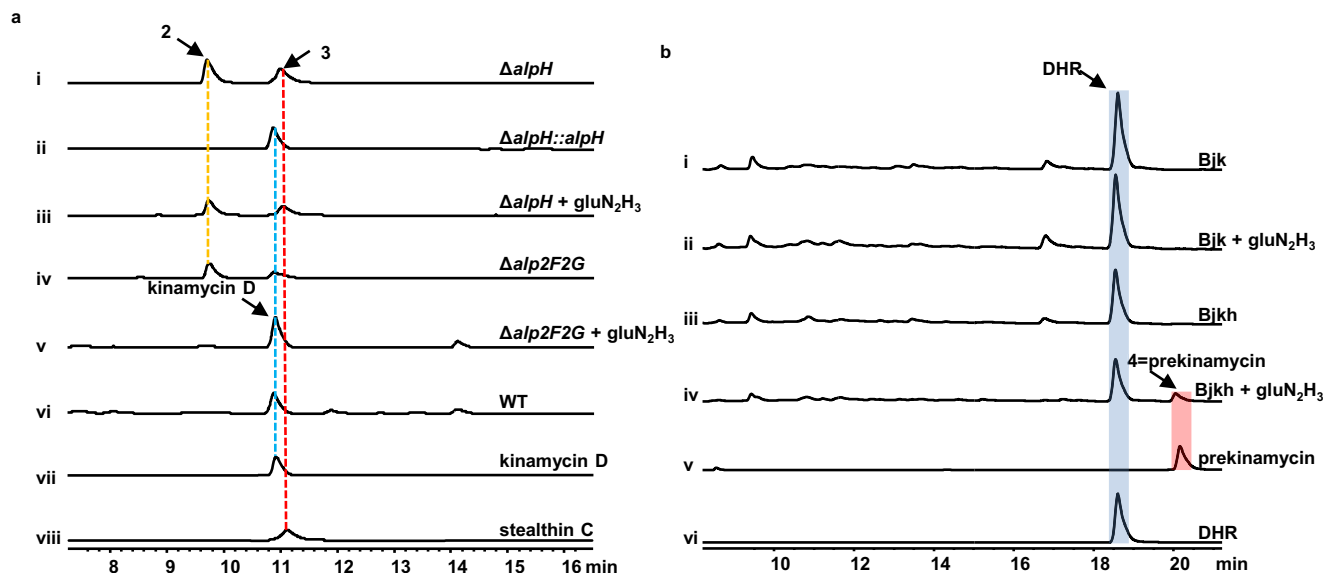
## Results

### In vivo genetic verification of *alpH*'s role in kinamycin biosynthesis

Although the BGC for kinamycin has been identified by us and other groups separately, the putative enzyme for the N-N bond installation is still obscure. A previous study predicts that Alp1W (a putative

glutamine amidotransferase) as the putative enzyme to transfer hydrazine onto kinobscurinone, which is produced from DHR by AlpJ and AlpK, with no experimental evidence<sup>26</sup>. Alp1W is conserved among kinamycin, lomaiviticin, and nenestatin biosynthetic pathways, which all include diazo incorporation. However, there is no *alp1W* homologous gene in the BGC of fluostatin, whose biosynthesis also includes a similar diazo incorporation (Supplementary Fig. 1). To test the function of Alp1W in vivo, we generated an *alp1W* in-frame null mutant strain, *S. albus* J1074/W-3C2Δ*alp1W*. The Δ*alp1W* strain lost most of its ability to produce kinamycin and no other related compounds were accumulated according to the high-performance liquid chromatography (HPLC) profile (Supplementary Fig. 2), indicating the essential role of Alp1W in kinamycin biosynthesis. The production of kinamycin was restored to a comparable level to the wild-type strain by ex-situ complementation of *alp1W* driven by its own promoter. However, we can't tell whether Alp1W is involved in the gluN<sub>2</sub>H<sub>3</sub> loading onto the polyketide scaffold. To further test the aforementioned hypothesis in vitro, Alp1W, AlpJ, and AlpK were expressed, and purified from *Streptomyces* or *Escherichia coli* and incubated with DHR, NADH, and gluN<sub>2</sub>H<sub>3</sub> in a one-pot reaction (Supplementary Figs. 3, 4). As seen in a previous AlpJK one-pot reaction, the dimeric benzofluorene product (1) was generated and confirmed by liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) (Supplementary Fig. 5). Unfortunately, we could not detect any other new products after elaborate attempts. Apparently, our data challenges the hypothesis of Alp1W as the N-N bond installation enzyme in the kinamycin biosynthesis.

However, there is an OMT family protein coding gene, *alpH*, present in all kinamycin and lomaiviticin family NP BGCs with unassigned function in their biosynthesis (Supplementary Fig. 1). AlpH appears highly conserved in kinamycin, lomaiviticin, fluostatin, and nenestatin biosynthetic pathways (Supplementary Figs. 1 and 6, Supplementary Table 3). It possesses an N-terminal glycine rich motif, DFCGGQG, suggesting a putative SAM binding motif (DxGxGxG)<sup>27</sup> (Supplementary Fig. 6). AlpH also exhibits moderate homology to the OMT, MmcR<sup>28</sup> (89% coverage/55% identity), from the mitomycin BGC. Interestingly, O-methylation seems to be unnecessary for the kinamycin biosynthesis (Supplementary Fig. 7). However, we could not exclude the possibility of a cryptic methylation in the kinamycin biosynthesis, which has been reported in other NPs' biosynthesis such as the C-terminal methylation in thioestrepton biosynthesis<sup>29</sup>. Additionally, kinafluorenone<sup>30</sup>, an O-methylated shunt product from the kinamycin biosynthesis was also isolated from the *S. murayamaensis* mutant strain MCl. To verify the function of AlpH in kinamycin biosynthesis, we generated an *alpH* in-frame deletion mutant strain, *S. albus* J1074/W-3C2Δ*alpH*. Kinamycin production was completely abolished in this Δ*alpH* strain and was restored to a comparable level to the wild-type strain by ex-situ complementation of *alpH* driven by the *kasOp*\* promoter (Fig. 2a). Despite the disappearance of kinamycin D peak from the HPLC chromatogram of Δ*alpH* strain, two new peaks corresponding to seongomycin (2) and stealthin C (3) were observed, which were further confirmed by comparison with the authentic standards using HPLC and HRMS (Supplementary Figs. 8, 9). 2 and 3 are both shunt products from the kinamycin biosynthesis derived from DHR and converted by AlpJ (AlpK may also be involved)<sup>23</sup>. Accumulation of 2 and 3 in Δ*alpH* strain indicates that AlpH is not required for the construction of the polyketide scaffold and may function after the AlpJK-mediated ring contraction in kinamycin biosynthesis. As observed in Δ*alpH* strain, the accumulation of 2 was also seen in our previous Δ*alp2F-2G* strain. Alp2F and Alp2G are homologs of CreE and CreD, catalyzing the generation of nitrous acid from aspartic acid<sup>18</sup>. In kinamycin biosynthesis, Alp2F and Alp2G are required for the biosynthesis of gluN<sub>2</sub>H<sub>3</sub>, which is the N-N bond carrier. However, the biosynthesis of gluN<sub>2</sub>H<sub>3</sub> has not been fully characterized, and the key enzymes involved in the N-N bond formation have not been



**Fig. 2 | In vivo functional analysis of AlpH.** **a** HPLC profiles of the crude extracts from different strains: (i)  $\Delta alpH$  strain; (ii)  $\Delta alpH::alpH$  strain; (iii)  $\Delta alpH$  strain fed with  $gluN_2H_3$ ; (iv)  $\Delta alp2F2G$  strain; (v)  $\Delta alp2F2G$  strain fed with  $gluN_2H_3$ ; (vi) wild-type strain; (vii) kinamycin D standard; (viii) stealthin C standard. **b** HPLC profiles of

the crude extracts from the *E. coli* BL21(DE3) derivatives fed with  $gluN_2H_3$ ; (i) DHR producing strain plus *alpJ-K* (Bjk); (ii) Bjk fed with  $gluN_2H_3$ ; (iii) DHR producing strain plus *alpJ-K-H* (Bjkh); (iv) Bjkh fed with  $gluN_2H_3$ ; (v) prekinamycin standard; (vi) DHR standard.

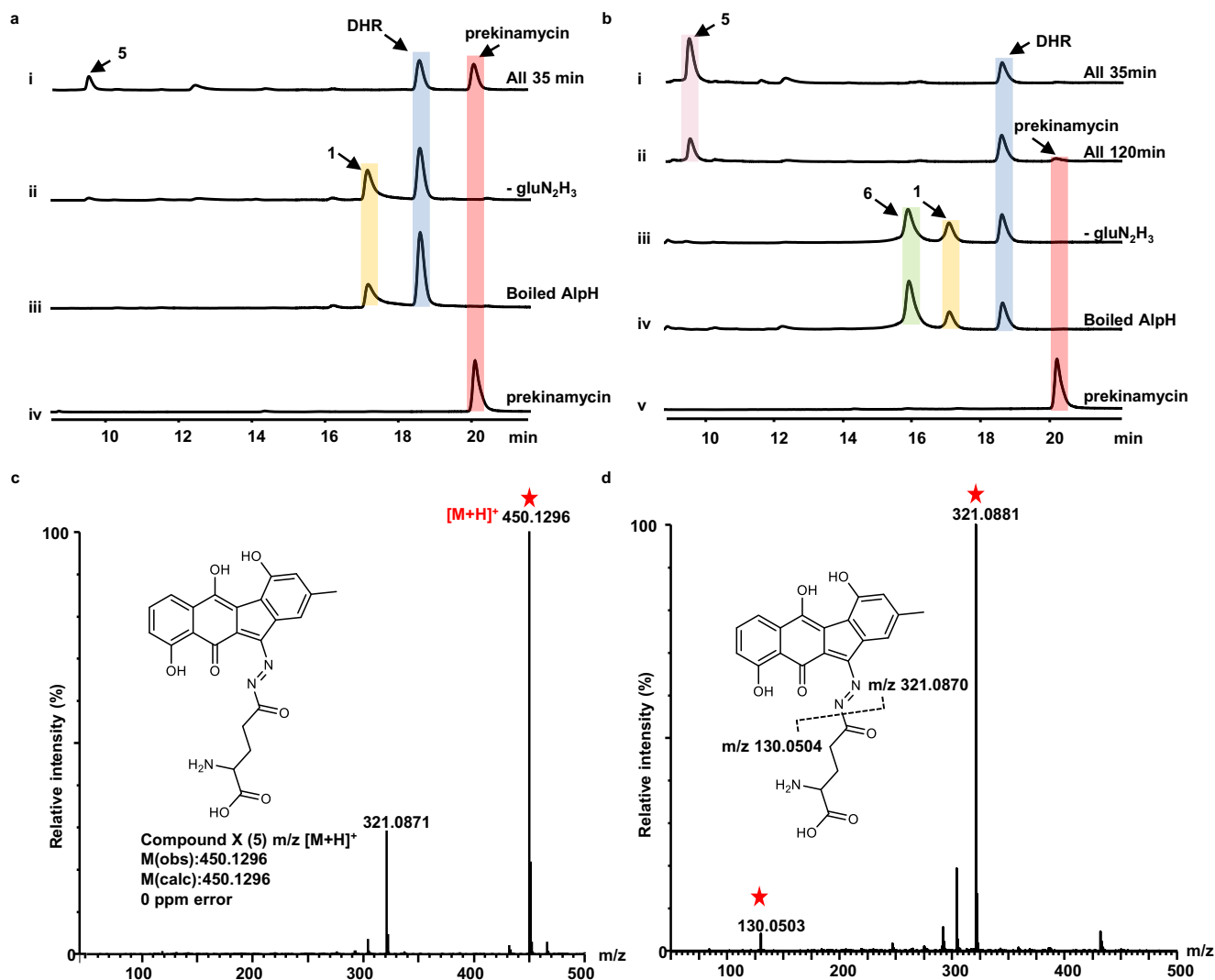
determined either. To test if AlpH is involved in  $gluN_2H_3$  biosynthesis, synthetic  $gluN_2H_3$  was fed to  $\Delta alpH$  strain. In contrast to  $\Delta alp2F2G$  strains,  $gluN_2H_3$ -fed  $\Delta alpH$  strain couldn't restore the production of kinamycin D, suggesting that AlpH is independent from  $gluN_2H_3$  biosynthesis (Fig. 2a). It is noteworthy that both mutant strains exhibited similar HPLC metabolic profile. Collectively, AlpH may play a role in the biosynthesis steps between the benzofluorene intermediate (6) and prekinamycin and be involved in the installation of the N–N bond to the polyketide scaffold. Since AlpH is predicted to be an OMT, it may catalyze a cryptic methylation, which is essential for subsequent installation of the N–N bond on the polyketide intermediate generated by AlpJ and AlpK. The methyl group may be removed during the maturation of prekinamycin. Alternatively, AlpH could probably function as a  $gluN_2H_3$  transferase instead of an OMT to install the N–N bond synthon onto kinamycin polyketide scaffold. Overall, the in vivo data suggests that AlpH plays an important role in kinamycin biosynthesis.

### Reconstitution of diazo formation in kinamycin biosynthesis

We've developed an *E. coli*-based expression system for producing DHR in previous studies, providing a robust heterologous platform to investigate the function of AlpH in kinamycin biosynthesis<sup>31</sup>. Given that AlpH may participate in the biosynthesis after AlpJ and AlpK, it was co-expressed in the DHR-producing strain with AlpJ and AlpK, resulting in Bjkh strain. Unfortunately, we did not see any differences between the control strain and the Bjkh strain from the HPLC metabolic profile (Fig. 2b). We speculated that the AlpH-catalyzed product may be unstable. Subsequently, we fed  $gluN_2H_3$  into Bjkh culture and observed the appearance of a new product (4) from the HPLC chromatogram. 4 exhibited an identical UV absorbance spectrum to prekinamycin, indicating the formation of a diazo product (Supplementary Fig. 10a). Large scale fermentation and purification yielded analytically pure 4 with a mass/charge ratio of 317.0568 ( $[M-H]^+$ ) determined by HRMS, identical to that of prekinamycin (Supplementary Fig. 10b). 4 was proven to be prekinamycin through further comparison to the authentic standard using HPLC (Fig. 2b). The production of 4 in Bjkh strain fed with  $gluN_2H_3$  suggests that AlpH is involved in the diazo formation in prekinamycin and can convert DHR to prekinamycin in cooperation with AlpJ and AlpK in the presence of  $gluN_2H_3$ .

We next sought to reconstitute the function of AlpH in vitro. AlpH was cloned, expressed, and purified into homogeneity from *E. coli* as N-terminal His<sub>6</sub>-tagged protein (Supplementary Fig. 3). The above in vivo data suggest that the substrate for AlpH is the product of DHR catalyzed by AlpJ and AlpK, which has been well-known to be unstable. Therefore, we designed a one-pot coupled assay using DHR as the substrate in the presence of AlpJ, AlpK,  $gluN_2H_3$ , FAD, NADH, and S-adenosylmethionine (SAM). HPLC monitoring of this one-pot reaction revealed the production of prekinamycin (Fig. 3a). Meanwhile, prekinamycin was absent in the control reactions using boiled AlpH or without  $gluN_2H_3$ , indicating that the diazo installation is AlpH and  $gluN_2H_3$ -dependent. Taken together, AlpJ, AlpK, and AlpH were sufficient to convert DHR to prekinamycin using  $gluN_2H_3$  as the N–N bond donor.

To further ascertain the details of the transformation of DHR into prekinamycin, we analyzed the reaction carefully and monitored the one-pot reaction in a time-course manner, which led to the discovery of a new product (compound X) beside prekinamycin (Fig. 3b). Since X is unstable during solvent extraction, the one-pot reaction was directly injected into HPLC for analysis without further clean-up. X was massively produced in 35 min and then gradually decreased with the appearance of prekinamycin after 120 min in the reaction, suggesting X as the true turnover product and prekinamycin as the decomposition product in the one-pot reaction. X exhibited a distinct UV absorbance spectrum from prekinamycin and DHR, and its molecular weight was determined to be 450.1296 ( $[M+H]^+$ ) through LC-HRMS, corresponding to a molecular formula of  $C_{23}H_{19}N_3O_7$  (Supplementary Fig. 11a, Fig. 3c). During the purification of X through HPLC, we noticed its spontaneous conversion to prekinamycin and a compound with the same molecular weight of dehydrated glutamic acid in DMSO (Supplementary Fig. 11b–d). MS/MS fragmentation of X yielded two major daughter ions,  $m/z$  321.0881 and  $m/z$  130.0503, which are consistent with the products of "dihydroprekinamycin" and glutamyl ions (Fig. 3d). Therefore, X was determined to be 5, which is the adduct of  $gluN_2H_3$  onto the polyketide scaffold. When any one of these three enzymes and NADH was removed from the assay, 5 was no longer produced, indicating all three enzymes and NADH are essential for the conversion of DHR to 5 (Supplementary Fig. 12). In addition, we noticed the production of benzofluorene (6) and dimeric



**Fig. 3 | In vitro functional characterization of AlpH.** **a** HPLC analysis of the organic extracts from the enzymatic reactions: (i) DHR + AlpJ + AlpK + AlpH + NADH + gluN<sub>2</sub>H<sub>3</sub>; (ii) DHR + AlpJ + AlpK + AlpH + NADH; (iii) DHR + AlpJ + AlpK + boiled AlpH + NADH + gluN<sub>2</sub>H<sub>3</sub>; (iv) prekinamycin standard. **(b)** Direct HPLC analysis of the enzymatic reactions: (i) DHR + AlpJ + AlpK + AlpH + NADH + gluN<sub>2</sub>H<sub>3</sub>

(35 min of reaction time); (ii) DHR + AlpJ + AlpK + AlpH + NADH + gluN<sub>2</sub>H<sub>3</sub> (120 min of reaction time); (iii) DHR + AlpJ + AlpK + AlpH + NADH; (iv) DHR + AlpJ + AlpK + boiled AlpH + NADH + gluN<sub>2</sub>H<sub>3</sub>; (v) prekinamycin standard. **c** LC-HRMS data (positive mode) of **5**. **d** MS/MS (positive mode) spectrum and proposed fragmentation of **5**.

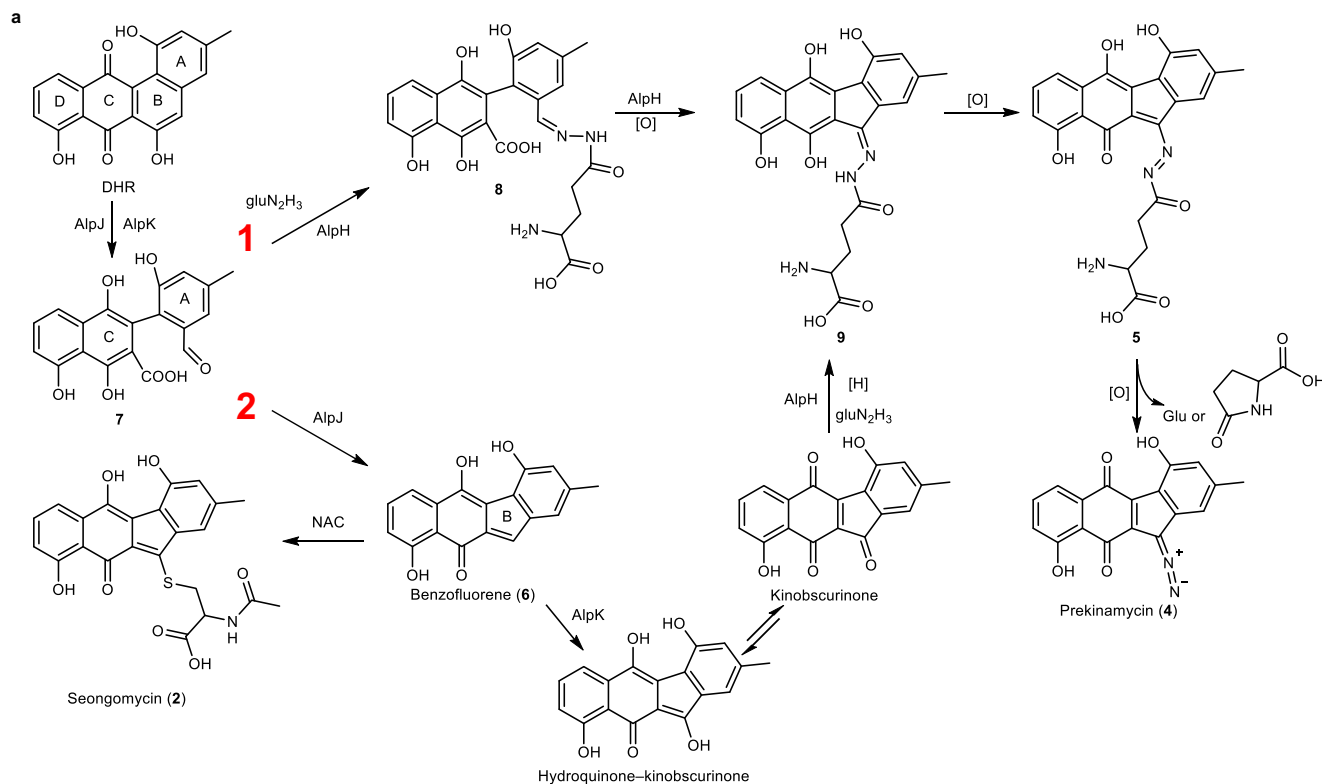
benzofluorene (**1**) in the control reactions, which were also observed by Yang and co-authors<sup>22</sup> (Supplementary Fig. 13). Our results unambiguously rule out the possibility of hydrazine, the hydrolyzed product of gluN<sub>2</sub>H<sub>3</sub>, as the N–N bond donor in kinamycin biosynthesis. Instead of hydrazine, gluN<sub>2</sub>H<sub>3</sub> was directly loaded onto the kinamycin biosynthetic intermediate, which may possess a carbonyl group to accept the gluN<sub>2</sub>H<sub>3</sub> adduct, generating **5**. **5** can be further hydrolyzed and oxidized into prekinamycin spontaneously at room temperature.

Intriguingly, we didn't identify any methylated intermediates in either the *E. coli* in vivo reconstitution system or the in vitro one-pot reaction system, which is contradictory to the predicted function of AlpH as a methyltransferase. Therefore, we removed SAM from the in vitro one-pot reaction for further analysis. Surprisingly, the production of **5** was significantly improved without SAM (Supplementary Fig. 14). In other words, the production of **5** was significantly inhibited by exogenously added SAM. Similarly, the production of **5** was significantly decreased by two SAM analogues, SAH and sinefungin (Supplementary Fig. 14). To rule out the carryover of SAM from protein purification, AlpH was denatured, and no SAM was detected in the supernatant by LC-MS (Supplementary Fig. 15). To further investigate whether SAM is required for AlpH activity, G203A mutation was

introduced into its putative SAM binding motif. AlpH G203A can still mediate the production of **5** efficiently in the in vitro one-pot reaction (Supplementary Fig. 16). Therefore, the results hereby indicate that, although AlpH shows significant homology to OMT, SAM is not required for its functionalization. Phylogenetic analysis of AlpH with methyltransferases from other NP BGCs showed that AlpH and its homologs form a separate clade, indicating their distinct functions (Supplementary Fig. 17, Supplementary Table 4).

#### AlpH catalyzes N–N incorporation in kinamycin biosynthesis

Since AlpH is a SAM-independent enzyme, we propose that AlpH may catalyze the transfer of gluN<sub>2</sub>H<sub>3</sub> onto the polyketide scaffold. However, we can't rule out the possibility of AlpH catalyzing an unknown modification on the polyketide product generated by AlpJ and AlpK from DHR, which can facilitate the transfer of gluN<sub>2</sub>H<sub>3</sub>. It has been reported that seongomycin (**2**) was produced by a spontaneous nucleophilic attack of N-acetylcysteine (NAC) to an electrophilic intermediate, putatively hydroquinone-kinobscurinone, converted from DHR by AlpJ and AlpK<sup>23</sup>. We hypothesize that the production of **2** will be impaired if AlpH modifies the product/intermediates generated by AlpJ and AlpK. Therefore, AlpH was added to the one-pot reaction mixture for



**Fig. 4 | Mechanisms of  $\text{gluN}_2\text{H}_3$  and NAC installations onto the polyketide skeleton. a** Mechanistic alternatives for the coupling of  $\text{gluN}_2\text{H}_3$  and polyketide intermediate. **b** HPLC analysis of the competition reactions: (i) DHR + AlpJ + AlpK + NADH; (ii) DHR + AlpJ + AlpK + NADH +  $\text{gluN}_2\text{H}_3$ ; (iii) DHR + AlpJ + AlpK + NADH + NAC; (iv) DHR + AlpJ + AlpK + NADH +  $\text{gluN}_2\text{H}_3$  + NAC; (v) DHR + AlpJ + AlpK + AlpH + NADH; (vi) DHR + AlpJ + AlpK + AlpH + NADH +  $\text{gluN}_2\text{H}_3$ ; (vii) DHR + AlpJ + AlpK + AlpH + NADH + NAC; (viii) DHR + AlpJ + AlpK + AlpH + NADH +  $\text{gluN}_2\text{H}_3$  + NAC. **c** HPLC analysis of the two-step reactions: (i) DHR + AlpJ + AlpK + AlpH +

NADH, DHR was consumed with the formation of **1** and **6**; (ii) DHR + AlpJ + AlpK + NADH, DHR was consumed with the formation of **1** and **6**; (iii)  $\text{gluN}_2\text{H}_3$  was added to reaction (i) when DHR was totally consumed; (iv) NAC was added to reaction (i) when DHR was totally consumed; (v) AlpH and  $\text{gluN}_2\text{H}_3$  was added to reaction (ii) when DHR was totally consumed; (vi) AlpH and NAC was added to reaction (ii) when DHR was totally consumed; (vii) DHR + AlpJ + AlpK + AlpH + NADH +  $\text{gluN}_2\text{H}_3$ ; (viii) DHR + AlpJ + AlpK + NADH + NAC.

producing **2** including AlpJ, AlpK, DHR, and NAC. There was no difference in the production of **2** between the control reaction and the AlpH-added reaction from the HPLC chromatograms (Supplementary Fig. 18), indicating that AlpH can't modify the product/intermediates generated by AlpJ and AlpK. Collectively, AlpH is most likely catalyzing the transfer of  $\text{gluN}_2\text{H}_3$  directly onto the polyketide scaffold, producing **5**, in kinamycin biosynthesis.

There are two different mechanisms to generate **5** from DHR, 1) a Mannich-like reaction-mediated  $\text{gluN}_2\text{H}_3$  condensation with an aldehyde intermediate (**7**) similar to the isoleucine coupling in jadomycin

biosynthesis and ring closure through an enolate anion or hydroquinone attack on the iminium<sup>32,33</sup>; 2) coupling of  $\text{gluN}_2\text{H}_3$  to the polyketide scaffold after ring contraction, featuring a putative benzofluorene intermediate, which is in alignment with the formation of **2** (Fig. 4a). In mechanism 2, **2** and **5** compete for the same polyketide substrate. To further investigate the underlying mechanism of AlpH, competition assays were performed using the one-pot reaction mixture including AlpJ, AlpK, AlpH, and DHR in the presence of  $\text{gluN}_2\text{H}_3$  and NAC. When equal amounts of  $\text{gluN}_2\text{H}_3$  and NAC (molar ratio 1:1) were added, only **5** but not **2** was detected in the one-pot reaction (Fig. 4b). The

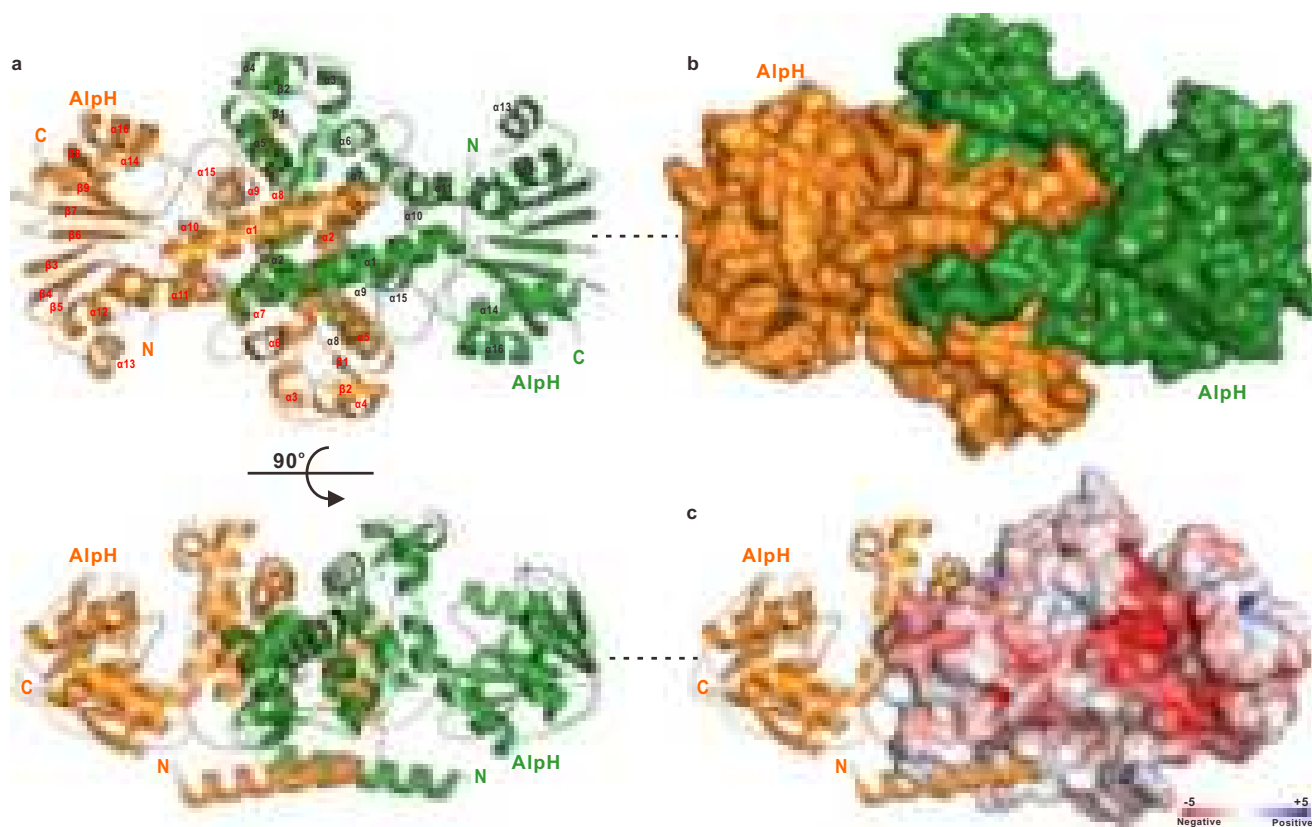
supplementation of 10 times more NAC over gluN<sub>2</sub>H<sub>3</sub> still did not result in the production of **2** (Supplementary Fig. 19), suggesting that AlpH may utilize a distinct polyketide substrate to generate **5**. We also noticed that **1** and **6** are only produced in the one-pot reaction for producing **2** but not **5**, indicating that the coupling of gluN<sub>2</sub>H<sub>3</sub> may proceed before the closure of the five-membered ring B (mechanism 1). To verify this hypothesis, we tested the activity of AlpH towards the putative five-membered ring intermediates, including **6** and its hydroxylated product hydroquinone-kinobscurinone. Due to their well-known instability, **6** and hydroquinone-kinobscurinone were prepared in situ by incubating AlpJ and AlpK with DHR and NADH as previously reported. AlpH and gluN<sub>2</sub>H<sub>3</sub> were added to the reaction after complete conversion of DHR into **1** and **6** ( $t=50$  min). However, no production of **5** was observed. In contrast, **2** was produced to a level similar to the control reactions with NAC added at  $t=0$  min and  $t=50$  min (Fig. 4c). Furthermore, when isoleucine was added to the reaction mixture without AlpH instead of gluN<sub>2</sub>H<sub>3</sub>, jadomycin A could be detected by LC-MS indicating that **7** is the potential product of AlpJ and AlpK catalyzed reaction (Supplementary Fig. 20). Therefore, although **2** and **5** share the same precursor, DHR, and possess identical benzofluorene scaffold, their direct precursors are different. Taken together, we propose that AlpH exploits a decarboxylative Mannich reaction to incorporate the N–N bond and construct the five-membered ring B in kinamycin biosynthesis (Fig. 4a). After the formation of **5**, a hydrolytic release of glutamic acid followed by oxidation completes the installation of the diazo group in kinamycin, resulting in prekinamycin.

### Crystal structure of AlpH

We determined the crystal structure of apo-AlpH to 1.87 Å resolution using the molecular replacement method (Supplementary table 5),

which exhibits a symmetric dimer conformation (Fig. 5a). Each AlpH monomer possesses 16  $\alpha$ -helices ( $\alpha$ 1– $\alpha$ 16) and 9  $\beta$ -strands ( $\beta$ 1– $\beta$ 9), forming two distinct sub-domains: an N-terminal dimerization domain and a C-terminal canonic  $\alpha/\beta$  Rossmann fold catalytic domain. The swapped dimer conformation is mediated by the N-terminal dimerization domain of each AlpH monomer, which interlocks with each other and forms an extensive dimer interface, covering a total surface area of 3947 Å<sup>2</sup> (Fig. 5b). Except for the protruding N-terminal  $\alpha$ -helix region, the overall architecture of AlpH monomer/dimer is highly similar to that of the SAM-dependent OMT LaPhzM<sup>34</sup> (PDB ID: 6C5B; RMSD, 2.155 Å) and MmcR<sup>28</sup> (PDB ID: 3GWZ; RMSD, 2.188 Å) (Supplementary Fig. 21a–d), and the SAM-independent OMT-like pericyclase PdxI<sup>35</sup> (PDB ID: 7BQK; RMSD, 2.512 Å) (Supplementary Fig. 21e, f) that catalyzes a unique Alder-ene reaction in fungal NP biosynthesis, as revealed by a structural similarity search using the program Dali<sup>36</sup>. Structural comparison analyses of AlpH with LaPhzM and MmcR revealed an extra protruding N-terminal  $\alpha$ -helix in AlpH that can further stabilize the interlocked dimer structure (Supplementary Fig. 21a, c). However, PdxI possesses one more N-terminal  $\alpha$ -helix comparing to AlpH for the formation of the swapped dimer (Supplementary Fig. 21e). Notably, although the overall structure of the catalytic domain of AlpH shows high similarity to the SAM-dependent OMT LaPhzM and MmcR, there is no incorporation of SAM in the determined AlpH structure, in agreement with the previous finding that purified AlpH is free of SAM or SAH.

Further structural comparison identified a unique solvent-exposed and negative charge enriched pocket between the dimerization domain and the catalytic domain in AlpH, which corresponds to the SAM/SAH binding pocket in MmcR<sup>28</sup> (Fig. 5c, Supplementary Fig. 22a–c). Sequence alignment and structural comparison of the

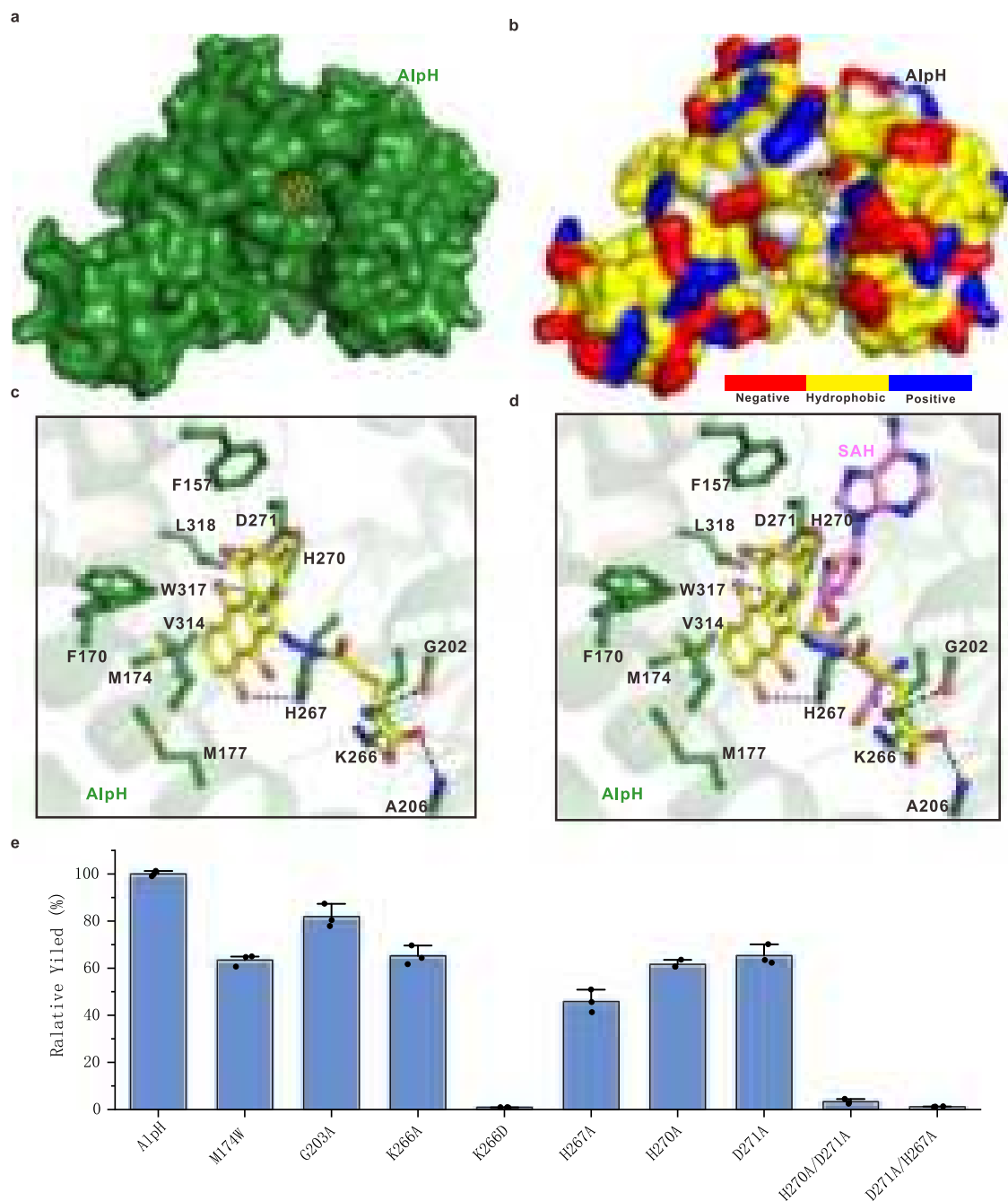


**Fig. 5 | The crystal structure of AlpH dimer.** **a** Ribbon diagram showing the overall structure of AlpH dimer. In this drawing, the two monomeric AlpH are colored in orange and forest green, respectively. **b** Surface representation showing the overall structure of AlpH dimer with the same orientation and color scheme as in panel A.

**c** The combined ribbon diagram and surface charge potential representation (contoured at  $\pm 5$  kT/eV; blue/red) indicates the putative negatively charged substrate-binding pocket of AlpH.

SAM/SAH binding pocket revealed that the critical glycine-rich GxGxG motif and F241 residue for SAH binding in MmcR/SAH complex are replaced by CxGxG and L252 residue in AlpH (Supplementary Figs. 6, 22d, e). Apparently, the bulky side chain of Cys introduces steric hindrance for SAM/SAH binding and the replacement of F241 by L252 also eliminates the putative  $\pi$ - $\pi$  stacking interaction between SAM/SAH and the aromatic side chain of F241, leading to the reduced binding ability of AlpH for SAM/SAH.

Given the uniqueness of the solvent-exposed pocket in AlpH, we propose that this pocket might be the substrate-binding pocket to generate **9**. We performed a molecular docking analysis of AlpH with **9** due to its relatively rigid structure comparing to the two flexible substrates. A search for the best complementary shape and the lowest binding energy led to a reasonable conformer of **9**, which accommodates the pocket very well and forms reasonable hydrophobic and polar interactions with the surrounding residues in AlpH (Fig. 6a–c)



**Fig. 6 | The docking model of AlpH in complex with the product **9**.**

**a** Combination of a surface representation and the stick-ball model showing the overall docking model of a monomeric AlpH in complex with the product **9**. The fitted product **9** is shown in stick model (orange) and is embedded in the pocket between the N-terminal dimerization domain and the C-terminal catalytic domain. **b** Surface representation showing the overall docking model of a monomeric AlpH in complex with the product **9** with the same orientation as in panel A. Specifically, the amino acid residues were colored as yellow (hydrophobic amino acid), blue (positively charged amino acid), red (negatively charged amino acid), and gray

(uncharged polar amino acid). **c** The combined ribbon diagram and the stick-ball model showing the detailed interactions between AlpH and the fitted product **9**. The related hydrogen bonds and salt bridges involved in the binding are shown as dotted lines. **d** A structural modeling analysis of the product **9** and SAH in binding to AlpH revealing that AlpH is unable to simultaneously interact with the product **9** and SAH. **e** Relative activities of AlpH mutants. The columns represent the average values of the products (the highest mean value was set 100%); bars indicate  $\pm$  SD (standard deviation) of  $n = 3$  independent replicates. Source data are provided as a Source Data file.



and was confirmed by MD simulations (Supplementary Fig. 23). In particular, the benzofluorene ring of **9** fits into a hydrophobic pocket formed by the side chains of F157, F170, M174, M177, V314, W317, and L318 (Fig. 6c). Hydrogen bonds are also observed between the three hydroxy groups of ring A, B, and C of **9** and the side chains of D271, H270 and H267. In addition, the glutamyl side chain of **9** forms a salt bridge with the side chain of K266 and a hydrogen bond with the carbonyl group of G202. Sequence alignment analysis indicate all these amino acid residues potentially involved in the substrate binding in the active site of AlpH are conserved through AlpH and its homologous proteins (Supplementary Fig. 6). A further molecular docking attempt using SAM/SAH revealed that AlpH cannot adopt **9** and SAM/SAH simultaneously, owing to putative steric hindrances (Fig. 6d). To further probe their functions in catalysis, we used site-directed mutagenesis to introduce M174W, K266A, K266D, H267A, H270A, and D271A mutations into AlpH separately which were overexpressed and purified (Supplementary Fig. 24). The M174W, K266A, K266D, H267A, H270A, and D271A mutations impair the catalytic activity of AlpH to various extent (Fig. 6e, Supplementary Fig. 16). Interestingly, the K266D mutant almost completely lost its catalytic activity. Double mutations of H270A/D271A and D271A/H267A further decreased the AlpH activity to around 5% of the wild-type AlpH. These experimental results verified the importance of the residues identified from the docking analysis.

## Discussion

Diazo NPs are hot spots in NP biosynthesis studies due to their structural rarity and broad spectrum of bioactivities<sup>37</sup>. Our understanding of the biosynthesis of the diazo NPs has been significantly improved owing to recent biochemical studies in cremeomycin<sup>14</sup>, alazopeptin<sup>8</sup>, and others<sup>1,6,38</sup>. Specifically, the secondary metabolism-specific nitrous acid biosynthetic pathway was identified to contribute a nitrogen atom for diazo group formation, which is catalyzed by an ATP-dependent ligase-CreM in cremeomycin biosynthesis and a transmembrane protein-AzpL in alazopeptin biosynthesis. Because of the identification of stealthin C from the mutant strains of the kinamycin producer, the diazo group formation in kinamycin has long been considered as a stepwise installation of the two nitrogen atoms<sup>13</sup>. This hypothesis is challenged by more recent studies; instead, it is speculated that the N–N bond is pre-formed in gluN<sub>2</sub>H<sub>3</sub> and then incorporated into the polyketide scaffold. In kinamycin biosynthesis, the two FAD-dependent monooxygenases AlpJ and AlpK have been reported to catalyze the ring contraction reaction to generate hydroquinone-kinobscurinone. The hydroquinone-kinobscurinone was believed to be decorated by gluN<sub>2</sub>H<sub>3</sub> or hydrazine to introduce the N–N bond into to polyketide scaffold. However, the critical step for the transfer of the N–N bond onto the polyketide scaffold is still missing. Herein, we identified an OMT enzyme, AlpH, which catalyzes a SAM-independent installation of gluN<sub>2</sub>H<sub>3</sub> onto the polyketide scaffold.

Our biochemical data suggested a Mannich reaction-mediated gluN<sub>2</sub>H<sub>3</sub> transfer and five-membered ring closure in the kinamycin biosynthesis. The putative key aldehyde/acid intermediate **7**, generated from DHR by AlpJ and AlpK, condenses with gluN<sub>2</sub>H<sub>3</sub> to form a hydrazone intermediate **8**, which subsequently undergoes an enolate anion attack to the iminium to close the five-membered ring, resulting in intermediate **9** (Fig. 4a). Therefore, kinobscurinone, a long believed kinamycin biosynthetic intermediate, might be an artifact. The intermediacy of kinobscurinone is mainly proposed from early feeding experiments, although with a very low incorporation rate<sup>13</sup>. AlpK is proposed to hydroxylate the benzofluorene intermediate to give hydroquinone-kinobscurinone due to its importance in seongomycin production<sup>22</sup>. However, no other direct biochemical evidence for this hydroxylation step was provided. Recently, Wang et al. reported that seongomycin could also be produced without AlpK, which is in contradiction to previous results<sup>23</sup>. After the ring closure, **9** is oxidized to

form a relatively stable intermediate, **5**. The identification of **5** clearly demonstrated that gluN<sub>2</sub>H<sub>3</sub> is directly transferred to the polyketide scaffold and ruled out the possibility of the hydrazine transfer suggested in previous studies. **5** could be gradually converted to pre-kinamycin, confirming it as a genuine intermediate in kinamycin biosynthesis. Mannich reactions have mainly been implicated in alkaloid NPs biosynthesis and are rarely described in other NPs biosynthesis<sup>39</sup>. Also, there are only limited examples of enzymes that can catalyze Mannich reactions. To our knowledge, LkcE is the only enzyme (except for promiscuous enzymes) to catalyze an enzymatic intramolecular Mannich reaction in LC-KA05 biosynthesis<sup>40</sup>. In light of the data in this study, AlpH represents the second enzyme catalyzing a Mannich reaction in the biosynthesis of aromatic polyketide kinamycin. Considering the docking and site-directed mutation results, we propose a mechanism for the AlpH catalyzed reaction (Supplementary Fig. 25). The Mannich reaction may be initiated through the protonation of the carbonyl group of the aldehyde intermediate (**7**) by a general acid (H267) to generate an iminium intermediate. Subsequently, the enol/enolate attack the iminium followed by decarboxylation to form the five-membered ring (Supplementary Fig. 25).

SAM-dependent methyltransferase family proteins from various NPs BGCs mainly function as methyltransferases that enriched their structures<sup>41</sup>. Intriguingly, there is a group of SAM-dependent methyltransferase family enzymes catalyzing diverse non-methylation reactions such as hydroxylation, decarboxylation, cyclization, and others<sup>42</sup>. Although they catalyze methylation-unrelated reactions, SAM is usually a required cofactor, such as SlnM<sup>43</sup> in salinomycin biosynthesis and LepI<sup>44</sup> in leporin C biosynthesis. AlpH is evolutionarily related to a family of small molecule OMT as indicated by their high amino acid sequence similarity. HHpred analysis also indicated the most similar structure of AlpH is OMT-LaPhzM<sup>44</sup> which agrees with an overall similar crystal structure. However, AlpH doesn't catalyze a methyl transfer reaction and purified AlpH is free of SAM. Biochemical assays and mutagenesis experiments demonstrated that AlpH catalyzes a SAM-independent Mannich reaction including hydrazine-carbonyl condensation and the following ring closure. In the AlpH-catalyzed reaction, the two small molecule substrates, gluN<sub>2</sub>H<sub>3</sub> and an aromatic polyketide intermediate **7**, may need a relatively large active site. We propose that an empty pseudo-SAM binding site may be suitable for accommodating gluN<sub>2</sub>H<sub>3</sub> and the putative aromatic polyketide intermediate **7**. The crystal structure of AlpH together with our docking analyses is consistent with this hypothesis. The condensation product **9** could be properly docked into the active site, suggesting the pseudo-SAM binding site may be suitable for gluN<sub>2</sub>H<sub>3</sub> binding. Surface plasmon resonance analysis of AlpH and gluN<sub>2</sub>H<sub>3</sub> also identified a specific interaction between them (Supplementary Fig. 26). Recently, two homologous groups of pericyclases were characterized as SAM-independent OMT-like enzymes<sup>35</sup>. Our phylogenetic analysis revealed that AlpH and its homologs form a separate clade from these pericyclases and their evolution distance is not close compared with other OMT family enzymes (Supplementary Fig. 17). Therefore, AlpH may represent a new functional subtype of OMT enzymes.

Cofactors are key components of many enzymes and possess extensive catalytic potential for chemical transformations<sup>45,46</sup>. Many enzymes gain new functions through binding of new cofactors<sup>47</sup>. Our results suggest an interesting evolution process for AlpH. The protein shares a very similar 3D structure to the OMT family proteins. Intriguingly, the loss of SAM binding is critical for AlpH, providing an additional binding site for gluN<sub>2</sub>H<sub>3</sub>. During the evolution of AlpH, the putative OMT ancestral protein seems to undergo a series of mutations to decrease its binding affinity to SAM but increase its binding affinity to gluN<sub>2</sub>H<sub>3</sub>. Overall, we described the preformed N–N bond incorporation onto the polyketide scaffold and completed the long missing steps for diazo installation in kinamycin biosynthesis. The recruitment of an OMT-like enzyme, AlpH, for gluN<sub>2</sub>H<sub>3</sub> transfer may

provide a new strategy of protein engineering for novel functions in the future.

## Methods

Bacterial strains, plasmids, and BACs used in this study are listed in Supplementary table 1, and PCR primers are listed in Supplementary Table 2. General enzymes, chemicals, Kits, media, and molecular biological reagents were purchased from standard commercial sources. The *E. coli* BAP1<sup>48</sup> was fermented in a production medium at 18 °C. The *E. coli* strains were cultured in LB at 37 °C or 30 °C supplemented with appropriate antibiotics as needed. *E. coli* DH5 $\alpha$  was used for plasmid propagation, BL21(DE3) for protein expression, and *E. coli* BW25113/pIJ790<sup>49</sup> was used for PCR-targeted mutagenesis. *E. coli* ET12567 bearing the RK2-derived helper plasmid pUB307<sup>50</sup> was used to facilitate the intergeneric triparental conjugation from *E. coli* ET12567/BACs to *S. albus* J1074<sup>51</sup> and their derivatives were grown on soya flour medium agar for sporulation and conjugation, in R2 liquid medium containing 5% HP-20 resin for fermentation.

## Sequence alignment and phylogenetic analysis

The phylogenetic tree was constructed by MEGA7<sup>52</sup> using a neighbor joining method. Multiple-sequence alignment was constructed by ClustalW<sup>53</sup> and ESPrit 3.0<sup>54</sup>. Protein sequences used for bioinformatic analysis were listed in the Supplementary table 4.

## Heterologous expression of BACs

The transfer of 3C2 and its derivatives from *E. coli* to *S. albus* J1074 was accomplished using *E. coli* ET12567/pUB307-mediated triparental conjugation. Integration of the BACs was confirmed by apramycin resistance and diagnostic PCR.

## Production and analysis of kinamycin and related metabolites

For the productions of kinamycin and related metabolites, *Streptomyces* spore suspensions were inoculated in liquid trypticase soy broth (TSB) and incubated at 30 °C for 24–48 h as seed culture. A total of 1 mL of seed culture was then transferred into 35 mL of R2 fermentation medium and incubated with shaking (220 rpm) for an additional 2–3 days at 30 °C.

For the reconstitution of diazo formation, *E. coli* BAP1/pGro7/pXY-2/pXY-3/pXY-6 was used as the starting strain for the strain construction of this work<sup>31</sup>. The gene *alpJ* and *alpK* fragments excised by *Xba* I and *Hind* III were sequentially ligated into the *Spe* I and *Hind* III sites of the pXY-2 just behind MCAT (malonyl-CoA: acyl carrier protein transacylase) to form pXY-2JK, the gene *alpH* was inserted into pXY-6 behind *alpG* to form pXY-6H. The BAP1 strains Bjk and BjkH were inoculated into 3 mL LB liquid medium with required antibiotics and cultured overnight at 37 °C with shaking at 220 rpm. Then 500  $\mu$ L of overnight culture was transferred to 50 mL fresh production medium with required antibiotics and cultured at 37 °C until the OD<sub>600</sub> reached between 0.4 and 0.6. IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) (final concentration 500  $\mu$ M) and arabinose (final concentration 3 mM) were added and the cultures were grown at 18 °C for an additional 2–3 days.

For metabolite analysis, the culture broth was extracted with an equal volume of ethyl acetate after adding 1% glacial acetic acid. The organic extract was dried by rotary evaporation, and the dried material was dissolved in methanol for subsequent analysis. HPLC analysis was performed on an Agilent 1200 HPLC system using an Eclipse XDB-C18 column (5  $\mu$ m, 4.6  $\times$  150 mm). In each experiment, the following condition was used: 75% solvent A (water with 0.1% trifluoroacetic acid [TFA]) to 100% solvent B over 20 min and 100% solvent B for 10 min at a flow rate of 1 mL/min.

## Isolation of DHR and prekinamycin

For the preparation of DHR, a 3 L liquid fermentation of the *E. coli* BAP1/pGro7/pKM-1/pXY-3/pXY-6 was extracted with an equal volume

of ethyl acetate after adding 1% glacial acetic acid<sup>55</sup>. The ethyl acetate was removed by vacuum evaporation. The ethyl acetate extract was loaded onto the silica gel column and flushed with chloroform. DHR was eluted at 100% chloroform. The fractions containing DHR were collected and pooled, and further purified by reverse-phase semi-preparative HPLC. At last, 1.1 g of DHR was obtained. For preparation of prekinamycin, a 5 L liquid fermentation broth of the mutant  $\Delta$ *alpIS-IV* was extracted by ethyl acetate after adding 1% glacial acetic acid. The ethyl acetate was removed by vacuum evaporation, and the extract was dissolved in methanol, and then applied to reverse phase C18 column chromatography, eluted with a gradient elution of H<sub>2</sub>O: MeOH mixture. The fractions were further purified by semi-preparative HPLC to afford pure prekinamycin.

## Chemical synthesis of L-glutamylhydrazine

16 mL hydrazine hydrate (90%) and 8 g L-glutamic acid-5-methyl ester were mixed and stirred for 5–7 h at 30 °C. Methanol was then added and the reaction mixture was kept at 4 °C overnight. The solids were filtered out and washed with methanol and dried to obtain 6 g gluN<sub>2</sub>H<sub>3</sub>. The yield of L-glutamylhydrazine was 75%. In the end, 2.0 mg of gluN<sub>2</sub>H<sub>3</sub> was dissolved in D<sub>2</sub>O for NMR experiments.

## Feeding experiments in *E. coli* BAP1

The BAP1 strains Bjk and BjkH were inoculated into 3 mL of LB liquid medium with required antibiotics and cultured overnight at 37 °C with shaking at 220 rpm. Then 500  $\mu$ L of overnight culture was transferred to 50 mL fresh production medium with required antibiotics and cultured at 37 °C. When OD<sub>600</sub> reached 0.4–0.6, IPTG (500  $\mu$ M) and arabinose (3 mM) were added. After growing at 18 °C for one day, gluN<sub>2</sub>H<sub>3</sub> was fed into the culture at a final concentration of 1 mM. The culture was grown for another two days at 18 °C.

## Protein expression and purification

The *alpH*, *alpJ*, and *alpK* genes were PCR amplified using 3C2 as a template and cloned into pET28a vector individually using appropriate restriction sites (seen in Supplementary table 2) and then were transformed to DH5 $\alpha$ . The error-free plasmids were verified by sequencing and then transformed into *E. coli* BL21(DE3) for protein expression. Transformants were cultivated overnight at 37 °C in LB with 100  $\mu$ g/mL kanamycin, and 10 mL of the overnight culture was then used to inoculate 1 L of LB with 100  $\mu$ g/mL kanamycin. Cultivation continued at 37 °C with shaking at 220 rpm, until the OD<sub>600</sub> reached between 0.4 and 0.6, and IPTG was added to 500  $\mu$ M. The cultures continued to grow at 16 °C for an additional 18 h. The *alpIW* gene was amplified from 3C2 and cloned into pWY45 using *Nde* I and *EcoR* I restriction sites to generate pWY45-*alpIW*. The error-free plasmids were verified by sequencing and introduced into *S. lividans* SBT5 and plated on MS-agar medium containing apramycin (50  $\mu$ g/mL) and trimethoprim (50  $\mu$ g/mL) for 4 days at 30 °C. Positive exconjugants were cultivated for 3 days at 30 °C in 50 mL TSBY with apramycin (50  $\mu$ g/mL) and trimethoprim (50  $\mu$ g/mL), and 15 mL of the culture was then used to inoculate 250 mL YEME medium with apramycin (50  $\mu$ g/mL) for 3 days at 30 °C. Then thiostrepton (25  $\mu$ g/mL) was added to induce protein expression for another 24 h. For protein purification, cells were harvested by centrifugation (8,000 g) at 4 °C and re-suspended in the binding buffer (20 mM Tris-HCl, 500 mM NaCl, 5 mM imidazole, pH 8.0). Then cells were disrupted by sonication. Following centrifugation, the supernatant was loaded onto the prewashed Ni-NTA column. The crude proteins were washed with the washing buffer (20 mM Tris-HCl, 500 mM NaCl, 30 mM imidazole, pH 8.0) and then eluted with elution buffer (20 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, pH 8.0). The eluate was then concentrated to a volume of 2.5 mL using an Amicon spin filter with a 10 kDa molecular weight cutoff. Excess imidazole was removed using a PD-10 column (GE Healthcare) equilibrated using storage buffer (50 mM Tris-HCl, 300 mM NaCl, 10%

glycerol). Protein concentrations were determined by Bradford method and the protein purity was assessed by SDS-PAGE.

### In vitro enzyme reactions

In AlpH catalyzed reaction, 0.25 mM  $\text{gluN}_2\text{H}_3$  (1 mM L-Cys or 1 mM NAC was used when necessary), 1 mM TCEP, 1 mM SAM, 2 mM NADH, 5  $\mu\text{M}$  FAD, 40  $\mu\text{M}$  AlpJ, 25  $\mu\text{M}$  AlpK, 35  $\mu\text{M}$  AlpH, and 120  $\mu\text{M}$  DHR were incubated in pH 7.5 50 mM Tris-HCl buffer at 30 °C for 35 min. In the in vitro assay of Alp1W, 4.5  $\mu\text{M}$  Alp1W was used to replace AlpH. At the end of the reaction, an equal volume of ethyl acetate with 1% glacial acetic acid was used to extract the product twice. The organic extract was dried by rotary evaporation and the final extract was dissolved in 50  $\mu\text{L}$  DMSO. For direct analysis of compound **5** in the reaction mixture, half a volume of cold methanol was added to the reaction mixture at the end of the reaction and the supernatant was directly analyzed by HPLC after centrifugation. To prepare the standard seongomycin, AlpJ, AlpK, NADH, FAD, NAC, and DHR were incubated together in 50 mM Tris-HCl buffer (pH 7.5) at 30 °C for 2 h and quenched by adding 400  $\mu\text{L}$  cold methanol. The extract was dried by rotary evaporation, and the final extract was dissolved in 50  $\mu\text{L}$  methanol. To prepare the standard stealthin C, AlpJ, AlpK, NADH, FAD, L-Cys, and DHR were incubated in 50 mM MOPS buffer (pH 7.5) at room temperature for 2 h and quenched by adding 400  $\mu\text{L}$  cold methanol. The extract was dried by rotary evaporation, and the final extract was dissolved in 50  $\mu\text{L}$  methanol.

### Construction of gene disruption mutants and complementation mutants

All genetic manipulations on BAC 3C2 were conducted using  $\lambda$ -Red recombination-mediated PCR-targeted gene deletion<sup>56</sup>. Individual gene disruption cassettes were generated via PCR amplification of the kanamycin resistance gene cassette with flanking FRT sites from pJTU4659. Primers were designed with 39 nucleotides matching the adjacent sequences of the targeted gene. Gene replacement by  $\lambda$ -Red recombination was achieved in *E. coli* BW25113/pJ790/3C2. Mutated constructs were confirmed by PCR and then transferred to *E. coli* BT340 for flippase (FLP) recombinase-mediated excision of the kanamycin resistance gene. The corresponding final knockout BAC, with FLP scar, was verified by PCR before transfer into *S. albus* J1074. For complementation experiments, the integrative plasmid pPM927 was used. The *alpH* gene was amplified from BAC 3C2 using primers AlpH-Nde I-F and AlpH-EcoR I-R and was ligated to the pre-digested pJTU968 to generate the construct pJTU968-*alpH*. The fragment of *alpH* under the control of *kasOp*\* promoter was cut from pJTU968-*alpH* by digestion with *Mun* I and *Eco* R I and was then ligated to *Eco*RI single-digested and dephosphorylated plasmid pPM927 to generate plasmid pPM927-*alpH*. The resulting pPM927-*alpH* was introduced into the *S. albus* J1074/W-3C2 $\Delta$ *alpH* mutant to obtain the complementation strain  $\Delta$ *alpH*::*alpH*. Exconjugants were selected on MS agar plates with thioestrepton, apramycin, and trimethoprim. The *alp1W* gene was amplified from BAC 3C2 using primers Alp1W-onestep-F and Alp1W-onestep-R and was cloned into pPM927. The resulting pPM927-*alp1W* was introduced into the *S. albus* J1074/W-3C2 $\Delta$ *alp1W* mutant to obtain the complementation strain  $\Delta$ *alp1W*::*alp1W*.

### Measurement of the presence of SAM in AlpH

To detect whether SAM was present in AlpH, 360  $\mu\text{M}$  AlpH in 40  $\mu\text{L}$  storage buffer (50 mM Tris-HCl, 300 mM NaCl, 10% glycerol) was denatured by acetonitrile. 100  $\mu\text{M}$  SAM was treated by acetonitrile as a control. Then the solutions were centrifuged, and the supernatants were analyzed by LC-MS. The standard of SAM was also analyzed by LC-MS.

### Protein crystallization and structural elucidation

Crystals of AlpH were obtained by the hanging drop vapor diffusion technique at 16 °C. Freshly purified AlpH protein (10 mg/mL or

20 mg/mL in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM DTT, and 1 mM EDTA) was mixed with equal volumes of reservoir solution containing 0.03 M citric acid, 0.07 M Bis-Tris propane (pH 7.6), and 20% (w/v) PEG3350. Crystals were frozen using reservoir solution added 20% (v/v) glycerol as cryo-protectant. A 1.87 Å resolution X-ray data set was collected at the beamline BL10U2 of the Shanghai Synchrotron Radiation Facility (SSRF)<sup>57</sup>. The diffraction data was processed and scaled with XDS and autoPROC software suite<sup>58,59</sup> (Supplementary table 5).

The phase problem of the complex was solved by the molecular replacement method using the automatic molecular replacement pipeline MoRDA within the CCP4 suite<sup>60,61</sup>. The initial model was rebuilt manually using COOT<sup>62</sup>, and then refined using REFMAC<sup>63</sup>, and PHENIX<sup>64</sup>. The qualities of the final model were validated by MolProbity<sup>65</sup>. The final refinement statistics of solved structure in this study were listed in Supplementary table 5. All the structural diagrams were prepared using the program PyMOL (<http://www.pymol.org/>).

### Molecular docking of product to AlpH

Rigid molecular docking was performed using AutoDock Vina 1.2.0<sup>66</sup>. The ligand product **9** was generated using ChemBio3D Ultra 12.0, and then energy minimized with Molecular Mechanics (MM2) until a minimum root-mean-square (RMS) gradient of 0.100 was achieved (<https://www.chemdraw.com.cn/>). Water molecules were removed from the AlpH structure in preparation for docking. The Gasteiger charges and hydrogens were added to the receptor structure and ligand with AutoDocktools<sup>67</sup>. When docking, the explorable space for docking was defined as a cube 20 Å in length centered at -67.713, 18.156, and -36.454 including the entire active-site, other key parameters such as algorithm were set to default values. Finally, the complex structure with the best combination of low binding energy and favorable orientation was selected. PyMOL 2.5 (<http://www.pymol.org/>) was used for viewing the molecular interactions and image processing.

### Measurement of the binding affinity between $\text{gluN}_2\text{H}_3$ and AlpH

The binding affinity was determined by surface plasmon resonance (SPR) using the Biacore 8 K. AlpH was diluted to 20  $\mu\text{g}/\text{ml}$  in pH 4.5 sodium acetate buffer and immobilized onto a CM5 sensor chip.  $\text{GluN}_2\text{H}_3$  was diluted into 0.0975, 0.195, 0.39, 0.78, and 1.56  $\mu\text{M}$  using PBS-0.05% Tween20 buffer. The experiment was performed in PBS-0.05% Tween20 buffer at a flow rate of 30  $\mu\text{L}/\text{min}$  for 120 s (contact time) and another 120 s (dissociation time). The  $K_{\text{on}}$ ,  $K_{\text{off}}$ , and  $K_{\text{d}}$  values were determined using Biacore™ Insight Evaluation Software version 4.0 with a 1:1 binding model.

### Molecular dynamics (MD) simulations

Three independent 200 ns MD simulations for ligand-bound systems were performed by using Gromacs 2018.3<sup>68</sup>. The a99SB-ILDN force field<sup>69</sup> was utilized for the protein, while the TIP3P model<sup>70</sup> was employed for water molecules. For the ligand, AmberTools 20 antechamber<sup>71</sup> was used to process its topology using GAFF force field<sup>72</sup> and AM1-BCC charge type<sup>73</sup>. The net charge of the ligand was determined to be 0. The system was solvated in a dodecahedron box with dimension of 11.996  $\times$  11.996  $\times$  11.996 nm<sup>3</sup>, and it was filled with 36,212 water molecules. Na<sup>+</sup> and Cl<sup>-</sup> were added to obtain a salinity of 0.15 M and maintain neutral total charge, resulting in a total of 119753 atoms in the system.

Prior to commencing the production simulations, an energy minimization step was performed using the steepest descent method. The maximum force was set to 1000 kJ mol<sup>-1</sup> nm<sup>-1</sup>, and convergence was achieved within 50,000 steps. Subsequently, a 100 ps NVT equilibrium simulation was conducted, applying a position restraint on heavy atoms with a force constant of 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup>. The LINCS method was used to constrain all covalent bonds, allowing a time step of 2 fs<sup>74</sup>. The van der Waals interaction cutoff was also set to 1.0 nm.

Long-range electrostatic interactions were calculated using the Particle Mesh Ewald method with a cutoff of 1.0 nm<sup>75</sup>.

Temperature control was maintained around 310 K using the V-rescale coupling method<sup>76</sup>. Following the NVT equilibrium simulation, a 100 ps NPT equilibrium simulation was performed, still retaining the position restraints and using the Parrinello Rahaman method<sup>77</sup> for pressure coupling. Finally, 200 ns NPT production simulations with no restraints were carried out, the remaining parameters remained consistent with the NPT equilibrium simulations.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The coordinate and structure factor of AlpH solved in this study have been deposited in the Protein Data Bank under the accession code [8H3T](#). All other relevant data supporting the findings of this study are available in this manuscript and Supplementary Information. Source data are provided with this paper.

### References

1. He, H. Y., Niikura, H., Du, Y. L. & Ryan, K. S. Synthetic and biosynthetic routes to nitrogen-nitrogen bonds. *Chem. Soc. Rev.* **51**, 2991–3046 (2022).
2. Blair, L. M. & Sperry, J. Natural products containing a nitrogen-nitrogen bond. *J. Nat. Prod.* **76**, 794–812 (2013).
3. Katsuyama, Y. & Matsuda, K. Recent advance in the biosynthesis of nitrogen-nitrogen bond-containing natural products. *Curr. Opin. Chem. Biol.* **59**, 62–68 (2020).
4. Kawai, S. et al. Complete biosynthetic pathway of alazopeptin, a tripeptide consisting of two molecules of 6-diazo-5-oxo-L-norleucine and one molecule of alanine. *Angew. Chem. Int. Ed. Engl.* **60**, 10319–10325 (2021).
5. Kawai, S., Hagihara, R., Shin-Ya, K., Katsuyama, Y. & Ohnishi, Y. Bacterial avenalamic acid biosynthesis includes substitution of an aromatic amino group for hydride by nitrous acid dependent diazotization. *Angew. Chem. Int. Ed. Engl.* **61**, e202211728 (2022).
6. Ma, G. L. et al. Biosynthesis of tasikamides via pathway coupling and diazonium-mediated hydrazone formation. *J. Am. Chem. Soc.* **144**, 1622–1633 (2022).
7. Zhao, G. et al. Molecular basis of enzymatic nitrogen-nitrogen formation by a family of zinc-binding cupin enzymes. *Nat. Commun.* **12**, 7205 (2021).
8. Del Rio Flores, A. et al. Biosynthesis of triacsin featuring an N-hydroxytriazene pharmacophore. *Nat. Chem. Biol.* **17**, 1305–1313 (2021).
9. Matsuda, K. et al. A natural dihydropyridazinone scaffold generated from a unique substrate for a hydrazine-forming enzyme. *J. Am. Chem. Soc.* **144**, 12954–12960 (2022).
10. Du, Y. L., He, H. Y., Higgins, M. A. & Ryan, K. S. A heme-dependent enzyme forms the nitrogen-nitrogen bond in piperazate. *Nat. Chem. Biol.* **13**, 836–838 (2017).
11. Waldman, A. J., Ng, T. L., Wang, P. & Balskus, E. P. Heteroatom-heteroatom bond formation in natural product biosynthesis. *Chem. Rev.* **117**, 5784–5863 (2017).
12. Chen, L., Deng, Z. & Zhao, C. Nitrogen-nitrogen bond formation reactions involved in natural product biosynthesis. *ACS Chem. Biol.* **16**, 559–570 (2021).
13. Gould, S. J. & Melville, C. R. Kinamycin biosynthesis. Synthesis, detection, and incorporation of kinobscurinone, a benzo [b] fluorenone. *Bioorg. Med. Chem. Lett.* **5**, 51–54 (1995).
14. Waldman, A. J. & Balskus, E. P. Discovery of a diazo-forming enzyme in cremeomycin biosynthesis. *J. Org. Chem.* **83**, 7539–7546 (2018).
15. Sugai, Y., Katsuyama, Y. & Ohnishi, Y. A nitrous acid biosynthetic pathway for diazo group formation in bacteria. *Nat. Chem. Biol.* **12**, 73–75 (2016).
16. Herzon, S. B. & Woo, C. M. The diazofluorene antitumor antibiotics: structural elucidation, biosynthetic, synthetic, and chemical biological studies. *Nat. Prod. Rep.* **29**, 87–118 (2012).
17. Woo, C. M., Ranjan, N., Arya, D. P. & Herzon, S. B. Analysis of diazofluorene DNA binding and damaging activity: DNA cleavage by a synthetic monomeric diazofluorene. *Angew. Chem. Int. Ed. Engl.* **53**, 9325–9328 (2014).
18. Liu, X. et al. Reconstitution of Kinamycin Biosynthesis within the Heterologous Host *Streptomyces albus* J1074. *J. Nat. Prod.* **81**, 72–77 (2018).
19. Gould, S. J., Hong, S. T. & Carney, J. R. Cloning and heterologous expression of genes from the kinamycin biosynthetic pathway of *Streptomyces murayamaensis*. *J. Antibiot. (Tokyo)* **51**, 50–57 (1998).
20. Wang, B. et al. Identification of Alp1U and Lom6 as epoxy hydrolases and implications for kinamycin and lomaiviticin biosynthesis. *Nat. Commun.* **6**, 7674 (2015).
21. Kersten, R. D. et al. Bioactivity-guided genome mining reveals the lomaiviticin biosynthetic gene cluster in *Salinispora tropica*. *ChemBioChem* **14**, 955–962 (2013).
22. Wang, B. et al. Kinamycin biosynthesis employs a conserved pair of oxidases for B-ring contraction. *Chem. Commun. (Camb.)* **51**, 8845–8848 (2015).
23. Wang, P., Hong, G. J., Wilson, M. R. & Balskus, E. P. Production of Stealthin C involves an S-N-type smiles rearrangement. *J. Am. Chem. Soc.* **139**, 2864–2867 (2017).
24. Gould, S. J., Melville, C. R., Cone, M. C., Chen, J. & Carney, J. R. Kinamycin biosynthesis. synthesis, isolation, and incorporation of Stealthin C, an Aminobenzo[b]fluorene. *J. Org. Chem.* **62**, 320–324 (1997).
25. Winter, J. M., Jansma, A. L., Handel, T. M. & Moore, B. S. Formation of the pyridazine natural product azamerone by biosynthetic rearrangement of an aryl diazoketone. *Angew. Chem. Int. Ed. Engl.* **48**, 767–770 (2009).
26. Wang, K. A. et al. Glutamic acid is a carrier for hydrazine during the biosyntheses of fosfazinomycin and kinamycin. *Nat. Commun.* **9**, 3687 (2018).
27. Kagan, R. M. & Clarke, S. Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Arch. Biochem. Biophys.* **310**, 417–427 (1994).
28. Singh, S. et al. Structural characterization of the mitomycin 7-O-methyltransferase. *Proteins* **79**, 2181–2188 (2011).
29. Liao, R. & Liu, W. Thiostrepton maturation involving a deesterification-amidation way to process the C-terminally methylated peptide backbone. *J. Am. Chem. Soc.* **133**, 2852–2855 (2011).
30. Cone, M. C., Melville, C. R., Gore, M. P. & Gould, S. J. Kinamycin, a benzo [b] fluorenone isolated from the kinamycin producer *Streptomyces murayamaensis*. *J. Organic Chem.* **58**, 1058–1061 (1993).
31. Liu, X. et al. Heterologous biosynthesis of type ii polyketide products using *E. coli*. *ACS Chem. Biol.* **15**, 1177–1183 (2020).
32. Böhm, M., Proksch, K. & Mahrwald, R. Decarboxylative Mannich reactions. *European J. Org. Chem.* **2013**, 1046–1049 (2013).
33. Rix, U. et al. The dynamic structure of jadomycin B and the amino acid incorporation step of its biosynthesis. *J. Am. Chem. Soc.* **126**, 4496–4497 (2004).
34. Jiang, J. et al. Functional and structural analysis of phenazine O-methyltransferase LaPhzM from lysobacter antibioticus OH13 and one-pot enzymatic synthesis of the antibiotic myxin. *ACS Chem. Biol.* **13**, 1003–1012 (2018).
35. Ohashi, M. et al. An enzymatic Alder-ene reaction. *Nature* **586**, 64–69 (2020).

36. Holm, L. Dali server: structural unification of protein families. *Nucleic Acids Res.* **50**, W210–W215 (2022).
37. Nawrat, C. C. & Moody, C. J. Natural products containing a diazo group. *Nat. Prod. Rep.* **28**, 1426–1444 (2011).
38. Van Cura, D. et al. Discovery of the azaserine biosynthetic pathway uncovers a biological route for  $\alpha$ -diazooester production. *Angew. Chem. Int. Ed. Engl.* **62**, e202304646 (2023).
39. Lichman, B. R. The scaffold-forming steps of plant alkaloid biosynthesis. *Nat. Prod. Rep.* **38**, 103–129 (2021).
40. Dorival, J. et al. Insights into a dual function amide oxidase/macrocyclase from lankacidin biosynthesis. *Nat. Commun.* **9**, 3998 (2018).
41. Liscombe, D. K., Louie, G. V. & Noel, J. P. Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Nat. Prod. Rep.* **29**, 1238–1250 (2012).
42. Sun, Q., Huang, M. & Wei, Y. Diversity of the reaction mechanisms of SAM-dependent enzymes. *Acta Pharm. Sin. B* **11**, 632–650 (2021).
43. Jiang, C. et al. Formation of the  $\Delta(18,19)$  double bond and bis(spiroacetal) in Salinomycin is atypically catalyzed by SlnM, a methyltransferase-like. *Enzyme. Angew. Chem. Int. Ed. Engl.* **54**, 9097–9100 (2015).
44. Ohashi, M. et al. SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature* **549**, 502–506 (2017).
45. Richter, M. Functional diversity of organic molecule enzyme cofactors. *Nat. Prod. Rep.* **30**, 1324–1345 (2013).
46. Fischer, J. D., Holliday, G. L., Rahman, S. A. & Thornton, J. M. The structures and physicochemical properties of organic cofactors in biocatalysis. *J. Mol. Biol.* **403**, 803–824 (2010).
47. Chu, X. Y. & Zhang, H. Y. Cofactors as molecular fossils to trace the origin and evolution of proteins. *ChemBioChem* **21**, 3161–3168 (2020).
48. Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E. & Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291**, 1790–1792 (2001).
49. Gust, B. et al. Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Adv. Appl. Microbiol.* **54**, 107–128 (2004).
50. Flett, F., Mersinias, V. & Smith, C. P. High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol. Lett.* **155**, 223–229 (1997).
51. Chater, K. F. & Wilde, L. C. *Streptomyces albus* G mutants defective in the SalGI restriction-modification system. *J. Gen. Microbiol.* **116**, 323–334 (1980).
52. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
53. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
54. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDScript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
55. Hua, K. et al. Offloading role of a discrete thioesterase in type ii polyketide biosynthesis. *mBio* **11** (2020).
56. Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. USA.* **100**, 1541–1546 (2003).
57. Wang, Q. S. et al. The macromolecular crystallography beamline of SSRF. *Nucl. Sci. Tech.* **26**, 12–17 (2015).
58. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D: Biol. Crystallogr.* **67**, 293–302 (2011).
59. Kabsch, W. Xds. *Acta Crystallogr. D: Biol. Crystallogr.* **66**, 125–132 (2010).
60. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D: Biol. Crystallogr.* **67**, 235–242 (2011).
61. Vagin, A. & Lebedev, A. MoRDa, an automatic molecular replacement pipeline. *Acta Crystallogr. A* **71**, S19–S19 (2015).
62. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
63. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
64. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
65. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
66. Morris, G. M. et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
67. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J. Chem. Inf. Model.* **61**, 3891–3898 (2021).
68. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
69. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
70. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
71. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* **26** (2005).
72. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25** (2004).
73. Jakalian, A., Bush, B. L., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **21** (2000).
74. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18** (1997).
75. Darden, T. A., York, D. M. & Pedersen, L. G. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
76. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
77. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

## Acknowledgements

This work was supported by the National Key Research and Development Program of China (2021YFC2100600 and 2018YFA0900402), the National Science Foundation of China (32170076, 31870026, and 21621002), the Science and Technology Commission of Shanghai Municipality (20XD1425200) (to L.P.), and the open project of the State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences. We thank SSRF BL10U2 for X-ray beam time. We thank Prof. Min Xu of Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences and Prof. Keqiang Fan of the Institute of Microbiology, Chinese Academy of Sciences for helpful discussion.

## Author contributions

Y.C.Z., X.Y.L., L.J.H., Y.W., S.J.L., P.S., and X.Z.W. performed compound isolation and structure determination. Y.C.Z., X.Y.L., and J.Z. conducted the in vitro biochemical studies and analysis. Y.C.Z. and L.F.P. carried out the protein crystallization experiments, X-ray analysis, and the crystal structure determinations of the proteins. Z.H.X. performed molecular docking analysis. X.Y.S. and W.N.W. conducted the molecular dynamics simulations. Y.C.Z., X.Y.L., X.Y.H., and M.J. analyzed the data and wrote the manuscript. M.J. and Z.X.D. directed the research. Y.C.Z. and X.Y.L. contributed equally to this work.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-41062-7>.

**Correspondence** and requests for materials should be addressed to Ming Jiang.

**Peer review information** *Nature Communications* thanks Yohei Katsuyama and the other, anonymous, reviewer(s) for their contribution to

the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023