JOC The Journal of Organic Chemistry

Note

Enzymatic reconstitution and biosynthetic investigation of the bacterial carbazole neocarazostatin A

Yating Liu, Li Su, Qing Fang, Jioji Tabudravu, Xiaohui Yang, Kirstie Rickaby, Laurent Trembleau, Kwaku Kyeremeh, Zixin Deng, Hai Deng, and Yi Yu

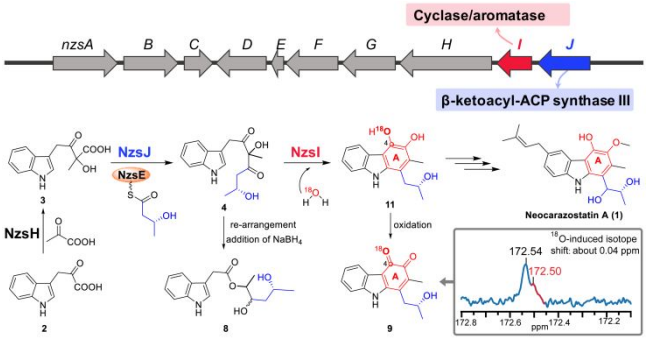
J. Org. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.joc.9b02688 • Publication Date (Web): 15 Nov 2019

Enzymatic reconstitution and biosynthetic investigation of the bacterial carbazole neocarazostatin A

Yating Liu[1#], Li Su[1#], Qing Fang[2], Jioji Tabudravu[3], Xiaohui Yang[4], Kirstie Rickaby[2], Laurent Trembleau[2], Kwaku Kyeremeh[5], Zixin Deng[1], Hai Deng[2*], Yi Yu[1*]

1. Institute of TCM and Natural Products, Key Laboratory of Combinatorial Biosynthesis and Drug Discovery (Ministry of Education), School of Pharmaceutical Sciences, Wuhan University, 185 East Lake Road, Wuhan 430071, P. R. China.
2. Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Aberdeen AB24 3UE, Scotland, UK.
3. School of Forensic & Applied Sciences, Faculty of Science & Technology, University of Central Lancashire, Preston, Lancashire, PR1 2HE, England, UK.
4. School of Chemistry, Xiamen University, Xiamen 361005, P. R. China
5. Department of Chemistry, University of Ghana, P.O. Box LG56, Legon-Accra, Ghana

[#] These authors contributed equally to this work



ABSTRACT: Tricyclic carbazole is an important scaffold in many naturally occurring metabolites, as well as valuable building blocks. Here we report reconstitution of the ring A formation of bacterial neocarazostatin A carbazole metabolite. We provide evidence of involvement of two unusual aromatic polyketide proteins. This finding suggests how new enzymatic activities can be recruited to specific pathways to expand biosynthetic capacities. Finally, we leveraged our bioinformatics survey to identify the untapped capacity of carbazole biosynthesis.

Neocarazostatin A **1** is a bacterial metabolite first isolated from the culture of *Streptomyces* sp. GP38 in 1991 and later from the soil bacterium *Streptomyces* sp. MA37, which consist of a tricyclic nucleus with two benzene rings (ring A and C) fused with a pyrrol ring (ring B) (Figure 1).[1-2] This compound exhibits potent antioxidant activity and displays considerably lower IC_{50} values for inhibition of lipid peroxidation than those of the free radical scavengers, such as butylhydroxytoluene and the drug flunarizine.[2] There have been considerable interests among medicinal chemists to develop **1** due to its pharmacological potential. **1** belongs to a group of simple carbazole alkaloids (CAs) with aliphatic side chains. Considering the oxidative status of ring A, this group of CAs can be categorized into four subgroups: indole-fused dihydroxyl-type CAs including **1** as representative, indole-fused monohydroxyl-type CAs such as carazostatin, indole-fused *ortho*-quinone CAs such as carquinonstatin A, and non-aromatic type CA (Figure 1).

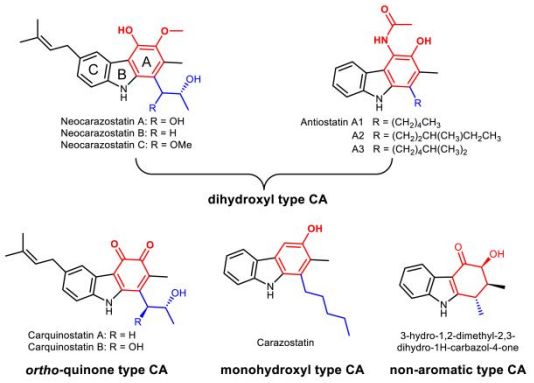


Figure 1. Representatives of four subgroups of bacterial tricyclic carbazole alkaloids.

In our previous studies, we have delineated the functions of five key enzymes encoded in the biosynthetic gene cluster (BGC) of **1** (*nzs*), including the phytoene-synthase-like prenyltransferase NzsG and the P450 hydroxylase NzsA,[1] a thiamine diphosphate-dependent enzyme NzsH,[3] a free-standing acyl carrier protein (ACP) NzsE and a classical β-ketoacyl-acyl carrier protein synthase III NzsF.[4]

Here, we demonstrated that *in vitro* reconstitution of NzsJ and I, together with NzsH and other necessary substrates and cofactors, enables the formation of the A ring of **1**, which is spontaneously oxidized into *ortho* quinone-containing carbazole, a similar observation discovered by Kobayashi and co-workers in parallel to our study.[5] Isotopic labelling studies demonstrate that one of oxygens in the A ring is derived from water. Moreover, comparative genomics and a global network analysis of sequence similarity of *nzsH*, *J* and *I* suggested that these gene homologues are clustered, and the occurrence of these BGCs in bacterial kingdom suggests important biological functions in these organisms.

Our previous studies suggested that both *nzsJ* and *nzsI* genes are essential for the production of **1** and no obvious intermediates were accumulated in the both *nzsJ* and *I* knockout mutants.[1] Bioinformatics analysis suggested that NzsJ is a putative FabH-like 3-ketoacyl-ACP synthase (KAS) III and NzsI is a structural homologue of aromatase/cyclases (ARO/CYCs) that involve in type II aromatic polyketide biosynthesis (PKS II), strongly suggesting that *Streptomyces* sp. MA37 recruits new enzyme activities from PKS II assembly line to the CA biosynthetic gene clusters to enable evolution of biosynthetic capacity. During the manuscript preparation, reconstitution of the biosynthetic pathway of carquinostatin A, a structural homolog of **1**, was reported[5] with the structural insights of one of the key biosynthetic enzymes, CqsB2, the homologue of NzsI with high sequence identity (80%).

In this study, we provide a full account of our work and show that NzsJ and NzsI catalyze the formation of the A ring of **1**. To this end, we set out *in vitro* enzymatic assays. Overexpression of NzsJ and NzsI in *S. lividans* allowed us to isolate and purify the recombinant protein to near homogeneity with the estimated molecular weight of 35 kDa and 28 kDa, respectively, as determined by SDS-PAGE analysis (Figure S1). Considering the instability of α-hydroxyl acyloin **3** from NzsH (the structure of **3** has been revised here compared to our previous structural interpretation[3]), we first performed one pot reaction by incubating NzsJ (1 μM) with the enzymatic systems of NzsH and NzsE we've established before to generate the acyloin and 3-hydroxybutyryl-NzsE *in situ*, respectively, two new compounds with mass-to-charge ratios (*m/z*) of 290.1382 was formed as observed at LC-HR-ESIMS analysis. *N*-acetylcysteamine (SNAC) are commonly used as a simplified synthetic mimic of the reactive biosynthetic intermediates, such as acyl-ACP in the biosynthetic studies.[6-7] To isolate new compounds for structural elucidation, we chemically prepared *R*-3-hydroxybutyryl-SNAC (Figure S2-4). Upon incubation of NzsJ with *R*-3-hydroxybutyryl-SNAC and the enzymatic system of NzsH, the target molecules with identical molecular weight and UV absorption but different retention time (compound **5** and **6**) were observed in the HR-ESIMS and HPLC analyses (Figure 2A, S5). Interpretation of NMR spectral and HR-ESIMS data suggested that **5** is likely to be an indole-acetyl ester despite of the presence of the impurity (Figure S6-8). **6** was proposed to be the isomers of **5**.

To further confirm the structures of **5** and **6**, we added $NaBH_4$ into the enzymatic mixture at the end of assay, resulting in appearance of two new ions, **7** and **8**, with both of which have *m/z* of 292.1540 as observed in the HR-ESIMS (Figure S9), further confirming the presence of only one ketone functional group. HR-ESLMS data for compound **7** and **8** indicated a molecular formula of $C_{16}H_{21}NO_4$ suggesting 7 degrees of unsaturation. The structure of **8** were proposed to be the reduced product of **5** by the interpretation of NMR spectral data (Figure S10-13, Table S1) coupled with theoretical calculation (Figure S14-16). **7** was deduced to be the isomer of **8**.

Analysis of the structure of **5, 6** and their derivatives, **7** and **8**, led to speculate that **5** or **6** may be the substrate of NzsI. Surprisingly, an enzymatic assay of NzsI with the isolated **5**

and **6** resulted in no new product, strongly suggesting that **5** and **6** is a re-arranged enzymatic product, a similar observation of recent parallel report.[5]

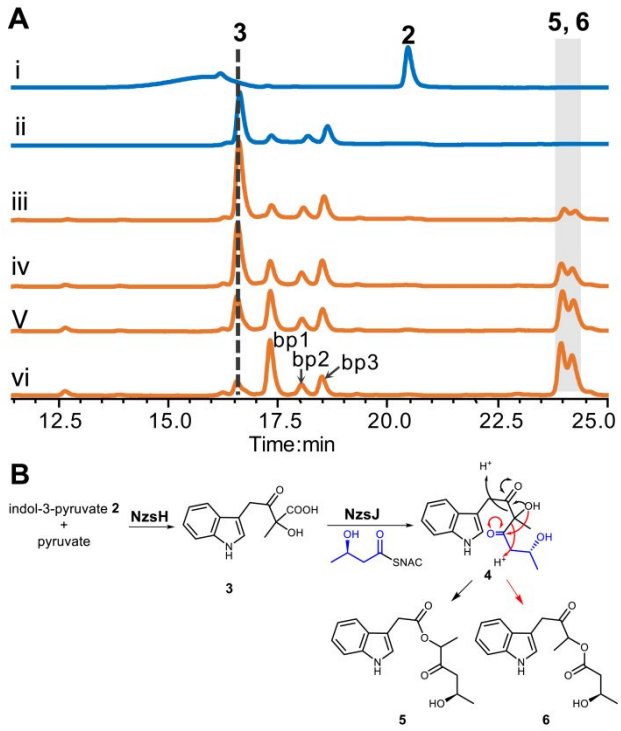


Figure 2. Functional dissection of NzsJ. (**A**) HPLC analysis of the one-pot enzymatic reaction catalysed by NzsH/J. (i) Control reaction using pyruvate, indol-3-pyruvate (**2**) and boiled NzsH/J; (ii) Control reaction using pyruvate, **2**, 3-hydroxybutyryl-SNAC and boiled NzsJ; (iii) – (vi) Time course of the NzsH/J one-pot reactions containing pyruvate, **2**, 3-hydroxybutyryl-SNAC, NzsH and NzsJ, terminated at 10, 30, 60 and 120 min, respectively. UV was monitored at 280 nm. It is worth to note that three HPLC peaks at the retention times of 17.5-18.5 min were putative byproducts (bp) during the biotransformation. The bp1 has the *m/z* of 378.1576 as observed in HR-ESIMS. However, we were unable to determine the m/z of bp2 and 3. (**B**) The proposed reaction of NzsJ.

This led us to speculate that **4** is the bona fide product of NzsJ that underwent a spontaneous α-ketol rearrangement reaction (Figure 2B). Such a rearrangement is also observed in the interconversion from dibenzo[*b*]fluorene skeleton to a benzo[*g*]chromene in the synthetic experiments[8] and the biogenetic interconversions between prekinamycin and isoprekinamycin[9] (Figure S17). Density functional theory (DFT) calculation also confirmed that the structure of **4** has higher energy than that of **5** or **6** (Table S2). It is worth to note that no ions with 4 Da increase were observed, strongly suggesting that the predicted intermediate **4** was not accumulated and underwent immediate re-arrangement.

To investigate the roles of both NzsJ and NzsI, we performed one pot reaction of NzsI in the presence of NzsJ, 3-hydroxybutyryl-SNAC, and the NzsH system. A new product **9** were generated with *m/z* of 270.1125 as observed in the HR-ESIMS analysis (Figure 3A, S18). In a control assay with boiled NzsI and in other control assays lacking either NzsH system or NzsJ or SNAC substrate, the formation of **9** was not observed.

To confirm the structure of **9**, a large-scale biotransformation reaction was carried out to afford the isolation of **9** (8 mg). The inspection of 1D and 2D NMR spectral data demonstrated that **9** is an *ortho*-quinone carbazole (Figure S19-22). Intriguingly, bioinformatics analysis suggested that NzsI is not an oxidative enzyme, leading us to speculate **10** is the bona-fide enzymatic product of NzsI but was oxidized spontaneously (Figure S23A). The presence of dithiothreitol (DTT, 1 mM) in the biotransformation of NzsH, NzsJ and NzsI indeed extended the existence of the *ortho*-dihydroxyl compounds **10** as observed in HPLC and HR-MS analyses (Figure S23B, C).

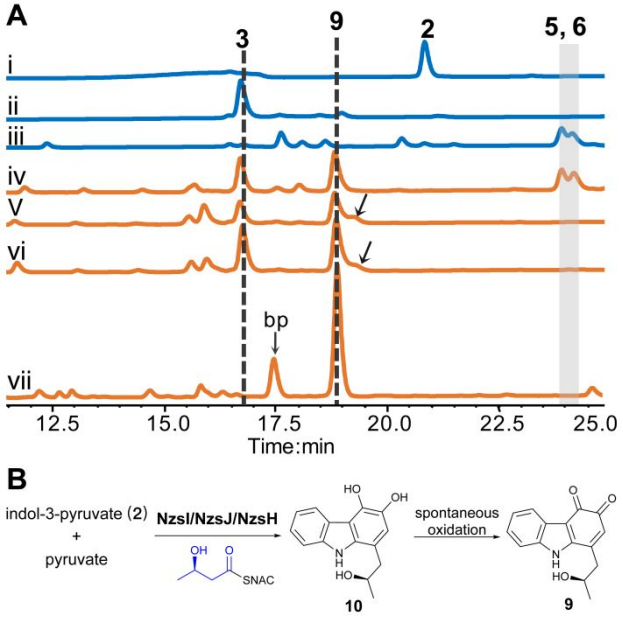


Figure 3. Functional dissection of NzsI. (**A**) HPLC analysis of the one-pot enzymatic reactions catalyzed by NzsH/I/J. (i) Control reaction using pyruvate, **2** and boiled NzsH/I/J; (ii) Control reaction using pyruvate, **2**, 3-hydroxybutyryl-SNAC, NzsH and boiled NzsJ/I; (iii) NzsH/J one-pot reactions plus boiled NzsI for 30 min incubation. (iv) NzsI was added into reaction iii and incubated for another 30 min; (v) – (vii) Time course of one-pot reactions containing NzsH/I/J, pyruvate, **2**, 3-hydroxybutyryl-SNAC, terminated at 30, 60 and 120 min, respectively. The tiny peak following **9** corresponds to a tautomer (**10** or **10a**). UV was monitored at 280 nm. (**B**) The proposed reaction of NzsI.

This led us to propose that NzsI may install the hydroxyl group at C4 of **4** from one water molecule. To verify this hypothesis, we carried out the biotransformation using isotopically enriched water. An assay of NzsI with the substrates in the buffer system enriched with $H_2{}^{18}O$ (15%) was performed. Indeed, isotopically enriched **9a** ($[M+2]^+$:18%) was observed (Figure 4A, B), as determined by HR-ESIMS analysis and by comparison with the MS pattern of **9** (Figure S24). To further confirm that the hydroxyl group is indeed located at C4 position in the A ring, a large-scale of biotransformation reaction was performed to afford the isolation the mixture of **9** and isotopically enriched **9a**. The inspection of $^{13}C\{^1H\}$ NMR of the resultant product demonstrated the ^{18}O induced isotope shifts at the ketone of C4 of **9** (Figure 4C, S25). The magnitude of this isotope shift is 0.04 ppm for the C4 attached to ^{18}O, consistent with the previous observation.[10]

Taken together, these experiments confirmed that NzsI utilizes the acid-base chemistry to catalyze the cyclization, followed by nucleophilic aromatization of water to afford the catechol moiety of carbazole metabolites, strongly suggesting that NzsI is a new type of cyclases/aromatases.

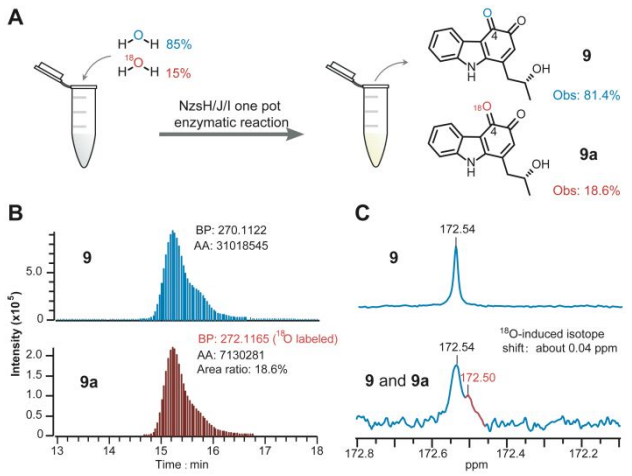


Figure 4. Isotope-labelling strategy and the representative mass chromatograms and NMR spectra. (**A**) Schematic depiction of the $H_2{}^{18}O$ isotope substitution. (**B**) LC-HR-ESIMS analysis of the enzymatic mixture of NzsH/I/J one-pot reaction. Extracted ion chromatograms representing the off-labelled **9** and on-labelled **9a**. BP: base peak mass, AA: peak area. (**C**) ^{18}O causes an upfield shift (0.04 ppm) of attached C4 NMR signals in compound **9a**.

Based on the oxygen pattern in the A ring, this group of bacterial carbazole metabolites can be categorized into four subgroups, mono-hydroxyl type, dihydroxyl type, ortho-quinone type and non-aromatic type (Figure S26). We propose that the A ring of the carbazole is assembled by NzsH, NzsJ and NzsI or corresponding enzyme homologs. The assembly of the A ring is initiated by the carboligation of **2** and pyruvate in a reaction catalyzed by NzsH or its homologs to afford the corresponding α-keto acid **3**. Depending on the bioavailability of the Acyl-CoA species in the producing strains, an unidentified acyltransferase will catalyze the transthiolation reaction to load Acyl-CoA into NzsE or its homologs to generate acyl-tethered thioesters. NzsJ or its homologs will utilize both **3** and acyl-tethered thioesters to mediate the decarboxylation-driven retro-aldol reaction to generate the intermediate, followed by cyclization via abstraction of the acidic proton at C4 position, spontaneous dehydration and water-based nucleophilic aromatization to yield the catechol motif. Methylation or transamination followed by acylation would result in dihydroxyl type carbazoles. Reduction on key catechol-type intermediate will generate the non-aromatic type. Currently only one metabolite without aromatic feature in the A ring has been isolated from *Streptomyces* sp. BCC26924.[11] If a reduction is involved in the NzsI-mediated reaction, the hydride-based nucleophilic aromatization would afford the mono-hydroxyl type of carbazole metabolites.

The presence of NzsI homologues in the bacterial kingdom was investigated. A blast search using NzsI as the query in the NCBI database allowed identification of a collection of NzsI-like homologues that share high amino acid sequence identity (43%-80%). The corresponding genes are found in the genome of microorganisms, including the gram-positive actinomycetes, the well-studied *S. cattleya*, the gram-negative bacteria including the thermophilic bacterium *Legionella gratiana*, the soil-dwelling myxobacterium *Sorangium cellosum* and the cyanobacterium *Scytonema tolypothrichoides* (Figure 5, Table S3). Strikingly, adjacent to these *nzsI*-like genes are *nzsH*-like and *nzsJ*-like genes in some cases (Table S3). *In silico* analysis also indicated that these NzsJ-like open reading frames (ORFs) also share high amino acid sequence identity (24%-73%) to NzsJ (Table S3). Furthermore, MEME prediction[12] demonstrated that all of the NzsI-like ORFs share highly conserved motifs with NzsI (Figure S27) whereas all of the NzsJ-like ORFs share the conserved catalytic triad C-H-N with NzsJ and other KAS III proteins (Figure S28), suggesting that these ORFs should process the same chemical reactions as those for NzsI and NzsJ, respectively. However, analysis of the genes in the close proximity of these *orf*s suggested that these identified gene clusters encode completely different sets of auxiliary enzymes that were predicted to modify the CA ring system or the fatty acid chains. It is likely that these strains have the potential to produce not-yet-discovered new CA-like metabolites.

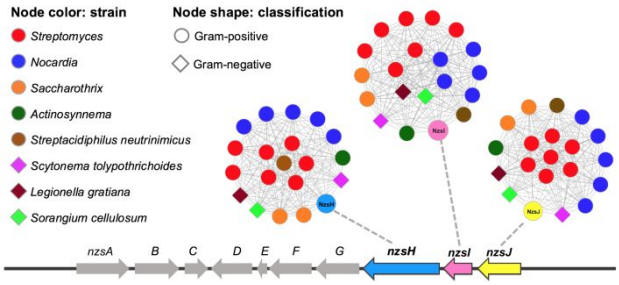


Figure 5. Sequence similarity network analysis of homologs of NzsH, NzsI, and NzsJ. Each node in the network represents a homologous protein and each edge represents the pairwise connection between two proteins with a blastP E value <1E-10.

In conclusion, *in vitro* reconstitution of two unusual enzymes, NzsJ and NzsI together with NzsH and other necessary substrates and cofactors, demonstrate the assembly of the A ring moiety of the bacterial carbazole metabolite **1**. Isotopic labelling studies demonstrated that the hydroxyl group at C4 position of the A ring is originated from one water molecule. Bioinformatics analysis further uncovered that the homologues of NzsI are widespread in bacteria associated with NzsH and NzsJ homologues, suggesting that carbazole-type metabolites are ubiquitous in the bacterial kingdom.

EXPERIMENTAL SECTION

General materials and methods. All chemicals and solvents were obtained from Sigma-Aldrich except where noted. Oligonucleotide primers were synthesized by GenScript (Nanjing, China). DNA sequencing of PCR products and all constructive plasmids were performed by TSINGKE Biological Technology (Wuhan, China). Restriction endonucleases and T4 ligase were purchased from NEB. High-fidelity DNA polymerase KOD purchased from TOYABO. A typical PCR reaction contained 10-100 ng of DNA template, 0.5 μM of each primer, and 2 μL 10 x KOD buffer, 0.3 μL high-fidelity KOD polymerase. Thermocycling was carried out in Bio-Rad C1000 thermocycler. *E. coli* DH10B was used as cloning host. *E. coli* ET12567/pUZ8002 was used for intergeneric conjugation between *E. coli* and *Streptomyces*. *E. coli* DH10B and *E. coli* ET12567/pUZ8002 were cultured in Luria-Bertani (LB, tryptone 1%, yeast extract 0.5%, NaCl 1%) or LB agar medium at 37 °C. *S. lividans* TK24 and *S. lividans* 1326 were grown on MS agar plates (soybean 2%, D-mannitol 2%, Agar 2%) at 28 °C for sporulation and in YEME medium (glucose 1%, tryptone 0.5%, yeast extract 0.3%, malt extract 0.3%, sucrose 10.3%) for protein overexpression. MS media with additional 10 mM $MgCl_2$ was used for intergeneric conjugation. HPLC analysis

was were carried out on a Shimadzu (Kyoto, Japan) HPLC instrument equipped with a degasser (DGU-20A3), an auto sampler (SIL-20A), a column oven (CTO-20A) and two pumps (LC-20AT) and a Phenomenex columns (C18, 5 μm) 250 × 4.6 mm. HPLC conditions for analysis of the one-pot reaction: C18 column was pre-equilibrated with 20% B and developed at a flow rate of 0.8 mL/min, 0-20 min, a linear gradient from 80% A to 20% A; 20-25 min, a linear gradient from 80% A to 5% A; 25-27 min, constant with 5% A; 27-30 min, a linear gradient to 80% A; 30-35 min, constant with 80% A; UV monitored at 280 nm. Solvent A was 0.1% formic acid in H_2O and solvent B was 0.1% formic acid in CH_3CN. HR-ESI-MS analysis was carried out in positive ion mode by using a Thermo Scientific LTQ XL Orbitrap mass spectrometer equipped with a Thermo Scientific Accela 600 pump (Thermo Fisher Scientific Inc.). Each LC conditions were described as above. All MS analysis parameters were set as 45 V capillary voltage, 45 °C capillary temperature, auxiliary gas flow rate 10 arbitrary units, sheath gas flow rate 40 arbitrary units, 3.5 kV spray voltage, and 50-1000 Amu mass range (maximum resolution 30,000).

Chemical synthesis of 3-hydroxybutyryl-SNAC. 3-hydroxybutyryl-SNAC was synthesized as a simplified mimic of the reactive biosynthetic intermediate, which can be recognized by NzsJ *in vitro*. (*R*)-(-)-3-hydroxybutanoyl sodium salt (630 mg, 5 mmol) was added to a solution of CF_3SO_3H (750 mg, 5 mmol) in 25 mL dichloromethane (DCM), and the mixture was stirred at 0 °C for 1h. Then the *N*-acetylcysteamine (NAC, 500 mg, 4.2 mmol, 1 eq), N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDCI, 800 mg, 1 eq) and N, N-Dimethylpyridin-4-amine (DMAP, 52 mg, 0.1 eq) were sequentially added and the reaction mixture was stirred at 22 °C for 16h. After removing the solid by filtration, the filtrate was poured into 10 ml cold water and extracted with ethyl acetate. Then the ethyl acetate layer was concentrated to dryness. The residue was dissolved in 2 ml DCM and the reaction was monitored by TLC (DCM/methanol = 15/1). The solvent was removed and the residue was purified on a silica gel flash column (DCM/methanol = 100/1) to yield 3-hydroxybutyryl-SNAC (0.12 g, 12%). ^{1}H NMR, δ_H ($CDCl_3$, 400 MHz): 1.21 (d, J = 6.3 Hz, 3H), 1.95 (s, 3H), 2.70 (m, 2H), 3.03 (m, 2H), 3.43 (m, 2H), 4.24 (m, 1H). ^{13}C{^{1}H} NMR, δ_C (101 MHz, $CDCl_3$): 199.6 (CO, C-7), 170.8 (CO, C-2), 65.3 (COH, C-9), 52.7 (CH2, C-8), 39.5 (CH2, C-4), 29.0 (CH2, C-5), 23.4 (CH3, C-1), 22.9 (CH3, C-10). HR-ESIMS m/z: $[M + H]^+$ Calcd for $C_8H_{16}NO_3S^+$ 206.0845; Found 206.0843.

Construction of NzsJ and NzsI overexpression plasmid. The genes *nzsJ* were amplified from *S.* sp. MA37 genomic DNA using primer pair His-NzsJ-F and His-NzsJ-R (Table S4) by high fidelity PCR. The PCR-amplified products were purified and then cloned into a NdeI/HindIII cleaved expression vector pWDYHS01 by using an In-fusion HD Cloning Kit (Clontech) to yield pWDY900. pWDY900 was then transformed into *S. lividans* TK24 and the resulting colonies were selected by *aac(3)IV* resistant marker and confirmed by PCR to obtain the recombinant strain WDY900. The strain WDY900 was used for overexpression His-tagged NzsJ in *Streptomyces*. The genes *nzsI* were PCR amplified from MA37 genomic DNA using the primers His-NzsJ-F/R and His-NzsI-F/R (Table S5). The PCR products were purified from an agarose gel and then cloned into an NdeI/HindIII linearized Vector pWDYHS01 using an InFusion HD cloning kit (Clontech). The ligation mixture was transformed into competent cells *E. coli* DH10B by heat shock. The recovery culture was plated on LB agar containing Apramaycin (50 μg/mL) to screen the positive clones, and then the positive assembly was confirmed by digestion and sequencing.

Expression and purification of NzsI and NzsJ. *E. coli* ET12567/pUZ8002 was firstly transformed with the verified plasmids and then conducted general intergeneric conjugation with the *S. lividans* TK24 or 1326 to yield the protein overexpressed strain WDY900 and WDY901 (Table S6), respectively. The overexpression strain WDY900 or WDY901 was inoculated into 50 mL TSBY medium supplemented with 50 μg/mL apramycin and grown at 28 °C with shaking at 200 rpm for 2 days. The preculture was transferred to 500 mL YEME medium supplemented with 50 μg/mL apramycin at 28 °C with shaking at 200 rpm for 2 days. Protein expression was induced with addition of 0.5 mM (final concentration) thiostrepton. After additional 60 h of incubation, the cells were harvested by centrifugation at 5000 g for 20 min at 4 °C. The cell pellets collected by centrifugation were re-suspended in ice-cold lysis buffer (20 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, and 10% (V/V) glycerol, pH 8.0) and disrupted by a high-pressure Homogenizer Machine (Avestin, EmulsiFlex-C3). Cell debris was removed via centrifugation (4 °C, 12000 g, 30min) and the supernatant was filtered before loaded onto 5 mL HisTrapHP column (GE Healthcare). The HisTrapHP column was washed by six concentration-step elution buffers with 15 mL for each step (20 mM Tris-HCl, 300 mM NaCl, 10% (V/V) glycerol, along with 20 mM, 50mM, 100 mM, 150 mM, 200 mM or 300 mM imidazole, pH 8.0). The elution fractions were checked by SDS-PAGE and then the desired elution fractions were concentrated using a Centrifugal Filter Units (Milipore, 10,000 MWCO, Merck). The concentrated protein solution was subsequently desalted using a PD-10 Column (GE Healthcare) pre-equilibrated with the elution buffer (20 mM Tris-HCl, 100 mM NaCl, and 10% (v/v) glycerol, pH 8.0). The desalted effluent was centrifuged again before aliquoted, flash frozen in liquid nitrogen and stored at -80 °C. SDS-PAGE and spectrophotometric analysis (Bradford assay kit, Promega, BSA used as standard, absorbance at 595 nm) were used to check the purity and concentration of the protein.

In vitro enzymatic assay of NzsH/J one-pot reaction. The NzsH/J one-pot reaction was performed by incubating 1 μM NzsJ and 1 mM 3-hydroxybutyryl-SNAC with the enzyme systems of NzsH. The reaction mixture was incubated at 30 °C for scheduled time and quenched by equal volume of ice-cold methanol. The clarified supernatant was obtained by centrifugation at 13, 000 g for 30min and then subjected to HPLC and HR-ESIMS analysis. The NzsH and boiled NzsH/J added mixture were used as control reactions.

In vitro enzymatic assay of NzsH/J/I one-pot reaction. The NzsH/J/I one-pot reaction was performed by adding 1 μM NzsI in the mixture of NzsH/J one-pot system as described above. The mixture was incubated at 30 °C for scheduled time and quenched by equal volume precooled MeOH. The clarified supernatant was obtained by centrifugation at 13, 000 g for 30min and then subjected to HPLC and HR-ESIMS analysis. The NzsH, boiled NzsH/J/I added and NzsH/J one-pot mixture were all used as control reactions. The NzsH/J/I reactions with DTT was performed by adding 1 mM DTT in the one-pot system.

Isotopic labelling experiments. To verify the oxygen attached on C4 of **9** (1-(2-hydroxypropyl)-2-methyl-carbazole-3,4(9H)-dione) was derived from water, the NzsH/J/I one-pot reaction was

performed by incorporating 15% ^{18}O-water. The labelled compound **9a** was obtained with theoretically consistent mass data on the HR-ESIMS analysis. We then scale up the enzymatic reaction to 50 mL with 15% incorporation ratio of ^{18}O-water in order to accumulate **9a** for ^{13}C{^{1}H} NMR spectra analysis. The resulting incorporation ratio was calculated by integrating the peak area of **9a** on mass chromatogram as follows: ^{18}O incorporation ratio % = peak area **9a**/ (peak area **9a** + peak area **9**) × 100.

Structural characterization of reaction products. The compound **8** (3,5-dihydroxyhexan-2-yl 2-(1H-indol-3-yl) acetate) was isolated and purified from enzymatic reaction, evaporated to dryness and resolved in CD_3OD for structure characterization on NMR analysis. The compound **9** was isolated and purified from enzymatic reaction, evaporated to dryness and resolved in DMSO-d_6 for structure characterization on NMR analysis. The NMR data of the compound was collected on a Bruker Avance 600 MHz NMR spectrometry.

Compound **8**: ^{1}H NMR, δ_H (600 MHz, CD_3OD), 3.89 (m, H-1), 1.58 (m, H-2a), 1.36 (m, H-2b), 3.44 (m, H-3), 3.49 (q, H-4), 3.67 (s, H-6), 7.41 (d, H-8), 6.91 (dd, J = 7.7, 8.1 Hz, H-9), 6.99 (dd, J = 7.7, 8.0 Hz, H-10), 7.25 (d, J = 8.0 Hz, H-8), 7.06 (s, H-13), 1.06 (d, J = 6.4 Hz, H-14), 1.09 (d, J = 6.1 Hz, H-15). HR-ESIMS m/z: [M + H] $^{+}$ Calcd for $C_{16}H_{22}NO_4^{+}$ 292.1543; Found 292.1540.

Compound **9**: ^{1}H NMR, δ_H (600 MHz, DMSO-d_6), 7.83-7.86 (m, H-5), 7.50-7.55 (m, H-4), 7.22 (dd, J = 5.6, 2.9 Hz, H-6, H-7), 3.97 (q, J = 6.0 Hz, H-11), 2.79 (m, CH_2), 1.93 (s, CH_3), 1.24 (d, J = 6.0 Hz, CH_3). ^{13}C{^{1}H} NMR, δ_C (150 MHz, DMSO-d_6), 183.5 (CO, C-3), 172.54 (CO, C-4), 146.3 (C, C-9a), 139.6 (C, C-1), 136.9 (C, C-8a), 134.5 (C, C-2), 125.5 (C, C-4b), 123.9 (C, C-6), 123.7 (C, C-7), 120.0 (C, C-5), 113.3 (C, C-8), 110.7 (C, C-4a), 65.7 (CH, C-11), 37.6 (CH_2, C-10), 23.5 (CH_3, C-12), 12.0 (CH_3, C-13). HR-ESIMS m/z: [M + H] $^{+}$ Calcd for $C_{16}H_{16}NO_3^{+}$ 270.1125; Found 270.1125.

Compound **9a**: ^{13}C{^{1}H} NMR, δ_C (150 MHz, DMSO-d_6), 184.0 (CO, C-3), 172.54 and 172.50 (CO and C^{18}O, C-4), 147.1 (C, C-9a), 140.2 (C, C-1), 137.8 (C, C-8a), 134.6 (C, C-2), 126.0 (C, C-4b), 123.9 (C, C-6), 123.8 (C, C-7), 120.2 (C, C-5), 113.8 (C, C-8), 110.9 (C, C-4a), 66.0 (CH, C-11), 37.9 (CH_2, C-10), 23.7 (CH3, C-12), 12.2 (CH_3, C-13). HR-ESIMS m/z: [M + H] $^{+}$ Calcd for $C_{16}H_{16}NO_2{}^{18}O^{+}$ 272.1167; Found 272.1166.

Sequence similarity network analysis. The homologs of NzsH, NzsI, and NzsJ from different microorganisms were obtained by blastp search by using NzsH, NzsI, and NzsJ as query sequences, respectively. Adjacent to these *nzsI*-like genes, there are *nzsH*-like and *nzsJ*-like genes located in some gene clusters (Table S3). The proteins of these *nzsI*-like, *nzsH*-like and *nzsJ*-like genes were selected for further network analysis. The network was constructed by an all-by-all blastp comparison of each sequence against each other sequence, and was generated using the EFI-Enzyme Similarity Tool.[13] Sequence similarity networks were visualized in Cytoscape 3.6.1[14] to illustrate the distribution of NzsH, NzsI, and NzsJ cluster in microorganisms. The nodes were arranged by using the yFiles organic layout with manual adjustment. Accession numbers were listed in Supplementary Table 3. The alignment identity of NzsI homologs was visualized with heat map.

Sequence alignment analysis. Multiple sequence alignment of NzsJ, KAS III proteins, KAS III-like proteins and NzsJ homologs were performed by ClustalX2[15] and visualized by ESPript 3.0.[16] To determine conserved motifs in 20 NzsI-like sequences analysed by blast and cluster mining, MEME[12] search was performed with the following parameters: mod = zoops, nmotifs = 30, minw = 6 and maxwidth = 50.

Computational Methods for Theoretical Calculation. HR-ESIMS data for compound **8** indicated a molecular formula of $C_{16}H_{21}NO_4$ suggesting 7 degrees of unsaturation in the structure of compound **8**. Detailed analysis of ^{1}H, ^{13}C{^{1}H}, COSY, HSQC, and HMBC data enabled the construction of the spin systems. A good correlation (r^2=0.9984) of predicted ^{13}C NMR versus experimental ^{13}C{^{1}H} NMR data suggested the structure is possibly correct[17]. 1D and 2D NMR data were inputted to the ACD/Labs Structure Elucidator and all possible structures calculated within an average difference of 4 ppm difference between calculated and experimental chemical shifts yielding 48 possible structures. Top 5 calculated structures calculated by the Structure Elucidator with ^{13}C chemical shift deviations between experimental and predicted of the HOSE-code (dA),[18] Artificial Neural Net (dN), and Incremental Method. After ranking using the Neural Network Match Factor, the resulted number one candidate is in consistent with the predicted structure of **8**.

ASSOCIATED CONTENT

Supporting Information

Additional tables of primers, strains and protein accession numbers; NMR, HR-ESIMS and UV spectra of reactions and products; Sequence analysis of NzsJ and NzsI; DFT computation data and parameters.

AUTHOR INFORMATION

Corresponding Author

*E-mail: yu_yi@whu.edu.cn.

*E-mail: h.deng@abdn.ac.uk.

ORCID

Hai Deng: 0000-0002-0047-0622

Yi Yu: 0000-0001-8590-0574

ACKNOWLEDGEMENTS

We thank the financial supports from the National Key Research and Development Program of China (2018YFA0900400), the National Natural Science Foundation of China (31570033, 31811530299, and 31870035 to YY), the Royal Society-NSFC Newton Mobility Grant Award (IEC\NSFC\170617 to HD and YY), Biotechnology and Biological Sciences Research Council UK (HD, KR and LT, BB/R00479X/1), the Elphinstone Scholarship of University of Aberdeen (QF), Leverhulme Trust-Royal Society Africa award (AA090088) and Medical Research Council UK Africa Future leader (KK and HD, MR/S00520X/1)

REFERENCES

(1) Huang, S.; Elsayed, S. S.; Lv, M.; Tabudravu, J.; Rateb, M. E.; Gyampoh, R.; Kyeremeh, K.; Ebel, R.; Jaspars, M.; Deng, Z.; Yu, Y.; Deng, H. Biosynthesis of neocarazostatin A reveals the sequential carbazole prenylation and hydroxylation in the tailoring steps. *Chem. Biol.* **2015,** *22*, 1633-1642.

(2) Kato, S.; Shindo, K.; Kataoka, Y.; Yamagishi, Y.; Mochizuki, J. Studies on free radical scavenging substances from microorganisms. *J. Antibiot.* **1991,** *44*, 903-907

(3) Su, L.; Lv, M.; Kyeremeh, K.; Deng, Z.; Deng, H.; Yu, Y. A ThDP-dependent enzymatic carboligation reaction involved in Neocarazostatin A tricyclic carbazole formation. *Org. Biomol. Chem.* **2016,** *14*, 8679-8684.

(4) Su, L.; Zhang, R.; Kyeremeh, K.; Deng, Z.; Deng, H.; Yu, Y. Dissection of the neocarazostatin: a C4 alkyl side chain biosynthesis by in vitro reconstitution. *Org. Biomol. Chem.* **2017,** *15*, 3843-3848.

(5) Kobayashi, M.; Tomita, T.; Shin-Ya, K.; Nishiyama, M.; Kuzuyama, T. Discovery of unprecedented carbazole alkaloid-synthesizing machinery in *Streptomyces*. *Angew. Chem. Int. Edit.* **2019,** *58*, 13349-13353.

(6) Prasad, G.; Amoroso, J. W.; Borketey, L. S.; Schnarr, N. A. N-activated β-lactams as versatile reagents for acyl carrier protein labeling. *Org. Biomol. Chem.* **2012,** *10*, 1992-2002

(7) Storm, P. A.; Townsend, C. A. In trans hydrolysis of carrier protein-bound acyl intermediates by CitA during citrinin biosynthesis. *Chem. Commun.* **2017,** *54*, 50-53

(8) Kawamura, M.; Kamo, S.; Azuma, S.; Kubo, K.; Sasamori, T.; Tokitoh, N.; Kuramochi, K.; Tsubaki, K. Skeletal rearrangements of polycyclic α-ketols. *Org. Lett.* **2017,** *19*, 301-303.

(9) Proteau, P. J.; Li, Y.; Chen, J.; Williamson, R. T.; Gould, S. J.; Laufer, R. S.; Dmitrienko, G. I. Isoprekinamycin is a diazobenzo[*a*]fluorene rather than a diazobenzo[*b*]fluorene. *J. Am. Chem. Soc.* **2000,** *122*, 8325-8326

(10) Diakur, J.; Nakashima, T. T.; Vederas, J. C. Magnitudes of ^{18}O isotope shifts in ^{13}C nuclear magnetic resonance spectra of ketones and alcohols. *Can. J. Chem.* **1980,** *58*, 1311-1315

(11) Intaraudom, C.; Rachtawee, P.; Suvannakad, R.; Pittayakhajonwut, P. Antimalarial and antituberculosis substances from *Streptomyces* sp BCC26924. *Tetrahedron* **2011,** *67*, 7593-7597.

(12) Bailey, T. L.; Johnson, J.; Grant, C. E.; Noble, W. S. The MEME Suite. *Nucleic Acids Res.* **2015,** *43*, W39-49.

(13) Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta.* **2015**, *1854* (8), 1019-1037.

(14) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13* (11), 2498-2504.

(15) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Clustal W and clustal X version 2.0. *Bioinformatics* **2007**, *23* (21), 2947-2948.

(16) Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42* (W1), W320-W324.

(17) Williams, A. J.; Elyashberg, M. E.; Blinov, K. A.; Lankin, D. C.; Martin, G. E.; Reynolds, W. F.; Porco, J. A.; Singleton, C. A.; Su, S. Applying computer-assisted structure elucidation algorithms for the purpose of structure validation: Revisiting the NMR assignments of hexacyclinol. *J. Nat. Prod.* **2008**, *71* (4), 581-588.

(18) Bremser, W. Hose — a novel substructure code. *Analytica Chimica Acta* **1978**, *103* (4), 355-365