# Biochemical characterization and substrate specificity of the gene cluster for biosyntheses of K-252a and its analogs by *in vitro* heterologous expression system of *Escherichia coli*†

**Hsien-Tai Chiu,* Yu-Chin Lin,‡ Meng-Na Lee,‡ Yi-Lin Chen, Mei-Sin Wang and Chia-Chun Lai**

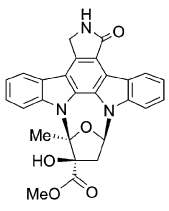
***Received 13th May 2009, Accepted 24th June 2009***
***First published as an Advance Article on the web 4th August 2009***
**DOI: 10.1039/b912395b**



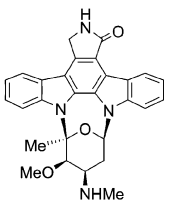
The indolocarbazole family of natural products has attracted great attention because of their unique structural features and potential therapeutic applications. Structurally distinct in the family, K-252a is characterized by an unusual dihydrostreptose moiety cross-bridged to K-252c aglycone with two C–N linkages. K-252a has served as a valuble lead for treatments of various cancers and neurodegenerative disorders. Recent cloning of the *nok* gene cluster for biosyntheses of K-252a and its analogs from *Nocardiopsis sp*. K-252 (NRRL15532) has revealed the *nokABCD* genes indispensible for K-252c biosynthesis and the key gene (*nokL*) coding for N-glycosylation. Herein, we report the first, successful demonstration of *in vitro* sugar transferase activity of indolocarbazole N-glycosyltransferase (NokL) by use of soluble protein expressed from *Escherichia coli*. Notably, NokL was found to exhibit peculiar mode of substrate promiscuity. Moreover, NokA and NokB reactions were biochemically characterized thoroughly by natual and alternative (*e.g*. fluoro-) substrates and by ammonium hydroxide ($NH_4OH$). Interestingly, the *in vitro* expression of NokA revealed high substrate stereoselectivity, giving several indole-3-pyruvic acid-derived compounds, including indol-3-carboxaldehyde (ICA) and indole-3-acetic acid. The use of $NH_4OH$ successfully dissected the *in vitro* NokA/NokB coupled reaction, revealing mechanistic insight into the enzymes and their cross-talking relationship. Also, a simple, useful method to synthesize K-252d, ICA and chromopyrrolic acid (the NokB product) was developed by the *E. coli* expression systems of NokL, NokA and NokA/NokB, respectively. Together with NokA and NokB, NokL may serve as a useful tool for combinatorial engineering of K-252a and its analogs for improved therapeutic values.

## Introduction

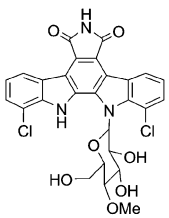
The indolocarbazole alkaloids are an emerging class of natural products that are attracting great attention because of their unique structural features, potential therapeutic applications and mechanisms of action.[1–3] The indolocarbazoles (*e.g*., staurosporine, rebeccamycin and K-252a) displayed broad range of potent biological activities, where antitumor, anticancer and neuroprotective applications have recently been of intense focus.[4–6] Mechanistically, they act through several possible mechanisms, including inhibition of DNA topoisomerases or protein kinases, and DNA intercalation.[7–9] Development of indolocarbazole derivatives has been actively pursued for enhancement of specificity and potency, and some (*e.g*. midostaurin) have, in fact, entered clinical trials for treatments of various cancer and diseases.[10] Among the indolocarbazole family, K-252a is, especially, novel in structure by possessing a special furanose moiety double cross-linked to a K-252 aglycone core. Since its discovery in 1986 from *Nocardiopsis sp*. K-252 (NRRL15532),[11,12] K-252a has drawn an increasing level of interest and attention for its potential therapeutic uses against various cancers and neurodegenerative disorders.[13,14] Lestaurtinib (CEP-701, KT-5555), CEP-2536 and CEP-751 (KT-6587) are examples of K-252a derivatives in several clinical trials.[15–18]



K-252a



Staurosporine



Rebeccamycin

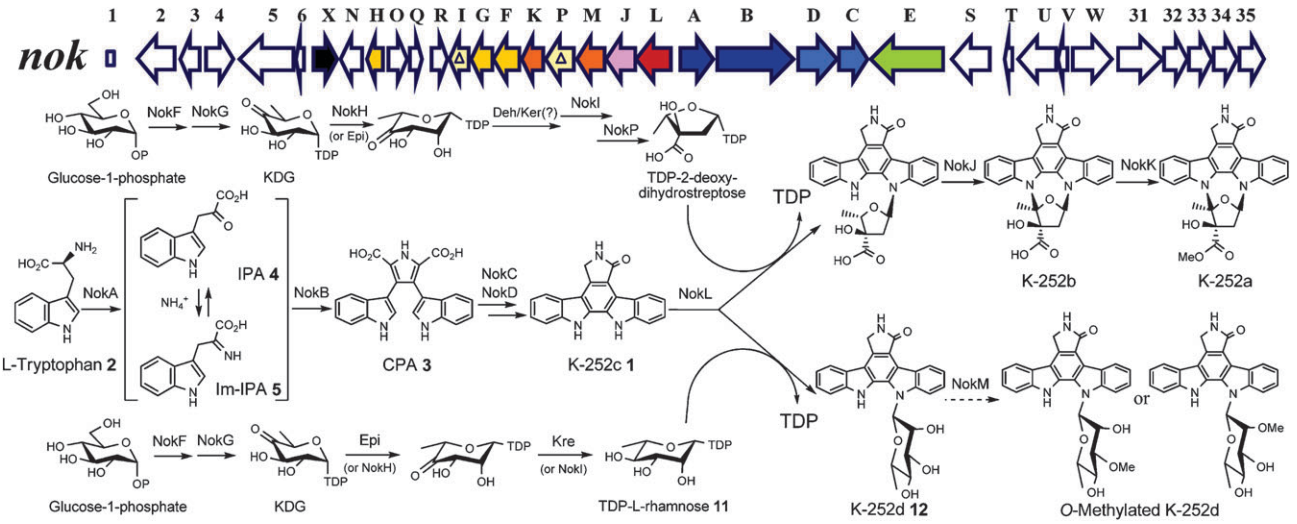
Recently, great efforts have been made to discover and manipulate biosynthetic genes of bioactive natural products for combinatorial biosynthesis of structurally diverse analogs

*Department of Biological Science and Technology, National Chiao Tung University, 75 Po-Ai Street, Hsinchu 300, Taiwan. E-mail: Chiu@mail.nctu.edu.tw; Fax: +886-3-5719605; Tel: +886-3-5131595*

† Electronic supplementary information (ESI) available: Experimental procedures for *in vitro* K-252d biosynthesis and NokL pH-profile; NokB reaction pathways; PCR primers; structures of NokL substrate analogs; [$NH_4OH$]/time correlation plots of NokA/NokB; MS and NMR data. See DOI: 10.1039/b912395b

‡ The equal contribution of these authors to this work is acknowledged.




**Scheme 1** The *nok* gene cluster and proposed biosynthetic pathways for biosyntheses of indolocarbazole metabolites in *Nocardiopsis sp.* K-252. The colored arrows indicate the genes presumably critical in biosynthesis, regulation and resistance of K-252a (see the companion paper).[20]

for improved or altered bioactivities.[19] In light of the prospects, we have recently cloned and sequenced the gene cluster responsible for biosyntheses of K-252a and its analogs from *Nocardiopsis sp.* K-252.[20] As described in the companion paper,[20] the discovery of the *nok* gene cluster allowed us to propose reaction pathways leading to their biosyntheses (Scheme 1), where the *nokABCD* genes indispensible for K-252c biosynthesis and the *nokL* gene coding for key enzyme responsible for N-glycosylation of K-252c were identified. Also, analysis of the biosynthetic gene clusters of indolocarbazoles revealed a conserved and coupled arrangement in gene order of L-amino acid oxidase (LAAO) and chromopyrrolic acid synthase (CPAS) genes, suggesting a close coupling mechanism likely adapted by the enzymes. The observation intrigued us to investigate here the possibility to dissect and/or improve the LAAO and CPAS reactions by use of the *in vitro* co-expression system of *Escherichia coli* established in our earlier study,[20] which may lead to a better understanding of the enzyme catalysis and/or efficient production of CPA **3** (or K-252c **1**) biosynthesis, respectively. On the other hand, N-glycosyltransferases (N-Gtfs) play a key role in the biosynthetic maturation of many bioactive indolocarbazole glycosides, *e.g.*, staurosporine and rebeccamycin.[21,22] The key enzyme responsible for N-glycosylation in the family has never been functionally characterized *in vitro* and at the molecular level. In K-252a, N-Gtf (NokL) catalyzes the bridging between the asymmetric aglycone (K-252c) and the unique dihydrostreptose moiety, serving as an essential element for its reported biological activities.[11,23] NokL may thus dictate an unique rational and feature of substrate specificity to harbor the furanose donor. In addition to its critical values in molecular enzymology, NokL can also be a useful tool for combinatorial biosynthesis to make novel K-252a analogs for therapeutic applications. However, recent efforts in N-Gtf studies have failed to either obtain the soluble protein in a heterologous expression system or demonstrate N-Gtf activity *in vitro*.[24] This situation would greatly hamper the discovery of unique features of N-Gtfs, especially NokL, with regards to its molecular mechanism and structure–activity relationship.

Herein, we report the first, successful demonstration of the *in vitro* sugar transferase activity of indolocarbazole N-Gtf (NokL). Strikingly, NokL was found to utilize TDP-L-rhamnose as an alternative substrate. Moreover, NokL was examined with a series of analogs of sugar donors and acceptors to further explore its substrate specificity. Together with the gene cluster sequence information from the companion paper, the *in vitro* functional characterization of NokL here allowed us to propose that the gene cluster served not only for biosynthesis of K-252a but also for those of its analogs. In addition, a heterologous expression system of *E. coli* was developed here and allowed us to efficiently produce K-252d *in vitro* by NokL with dTDP-D-glucose as a precursor. For resolution of the LAAO and CPAS reactions, *nokA* was cloned and overexpressed in *E. coli*, subsequently leading to discovery of several indole compounds derived from NokA reaction. Furthermore, NokA and NokB, together as an *in vitro* couple, were biochemically characterized, where they also displayed interesting substrate flexibility and metabolic relationship.

## Results and discussion

### *In vitro* biochemical characterization of NokA

As shown in Scheme 1, LAAO is the first committed enzyme for the biosynthesis of K-252c. However, the presumed product of LAAO reaction, the imine form (Im-IPA, **5**) of indole-3-pyruvic acid (IPA, **4**), has never been isolated or characterized in indolocarbazole biosynthesis.[25] To resolve the situation and the coupling relationship between LAAO and CPAS (NokA and NokB, respectively, in K-252a biosynthesis), the NokA reaction alone was biochemically characterized as an initial attempt in this study. To achieve the goal, an *in vitro E. coli* expression system of NokA was adapted. Hence, a wild-type expression plasmid (pJZ22) of

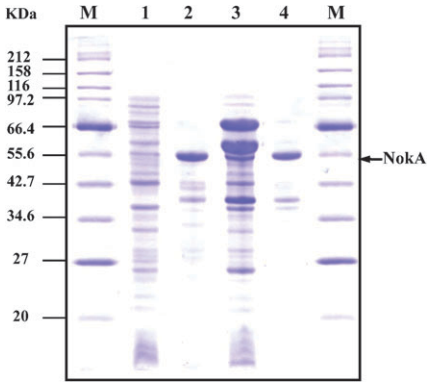



**Fig. 1** SDS-PAGE analysis of NokA protein from the heterologous expression. Proteins from the *E. coli* expression system (pJZ22) of NokA were subjected to SDS-PAGE and visualized by staining with Coomassie Brilliant Blue. Lanes M, molecular weight standards; lane 1, soluble fraction; lane 2, insoluble fraction; lane 3, soluble fraction (with chaperones); lane 4, insoluble fraction (with chaperones). The numbers to the side refer to the molecular mass of the standards in kilodaltons.

NokA, obtained as described in the companion paper,[20] was transformed into *E. coli* BL21 (DE3) for heterologous expression. Consequently, the IPTG-induced expression of pJZ22 resulted in mostly insoluble aggregates of NokA, despite the various induction temperatures and times attempted. The solubility of NokA was, however, greatly improved by co-expression of pJZ22 with pG-KJE7,[26] which can express chaperone teams of DnaK-DnaJ-GrpE and GroEL-GroES under regulation of *araC* gene for proper folding. As shown in Fig. 1, soluble form of NokA protein was clearly increased as co-expressed with the chaperones. Subsequently, the cell-free crude extract of NokA was incubated with L-tryptophan (L-Trp **2**) to examine its *in vitro* enzymatic activity. As a result, RP-HPLC analysis of the NokA reaction revealed a major new product peak (at retention time ($R_t$) 23.7 min) and two minor ones ($R_t$ 26.3 min and $R_t$ 27.4 min), appearing and growing along with a time-dependent consumption of L-Trp **2**, as shown in Fig. 2. Strikingly, upon product isolation and structural characterization by NMR and mass spectroscopy, the major product was found to be indole-3-carboxaldehyde (ICA, **6**) showing a characteristic aldehyde signal at 9.822 ppm

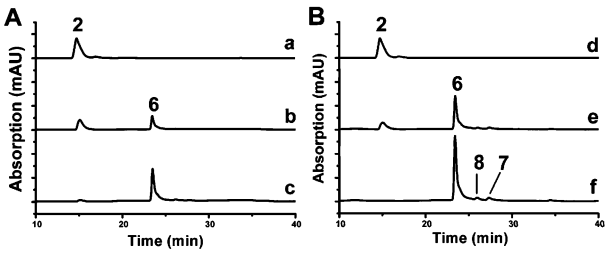



**Fig. 2** HPLC trace profiles of the NokA reactions with L-Trp. (A) The reactions without exogenous ammonium hydroxide; (B) the reactions in the presence of 30 mM ammonium hydroxide. The reactions were analyzed by detection at 300 nm at the indicated time points: (a) 0 h, (b) 6 h, (c) 12 h, (d) 0 h, (e) 3 h, and (f) 6 h. The HPLC peaks were identified to be L-Trp (**2**), ICA (**6**), an unknown IPA derivative (**8**), and IAA (**7**).

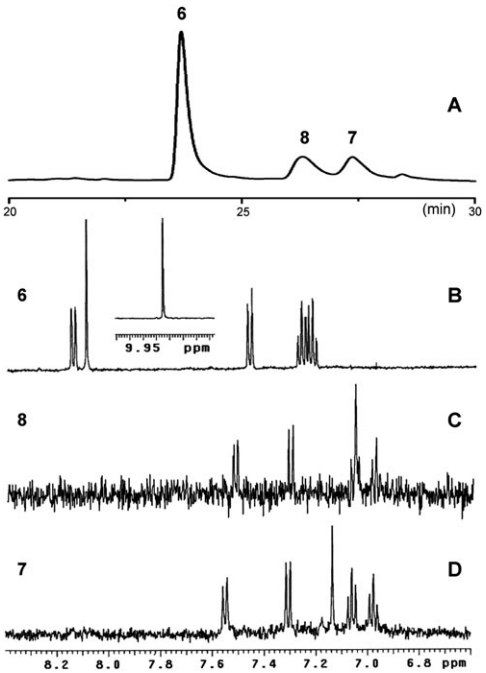
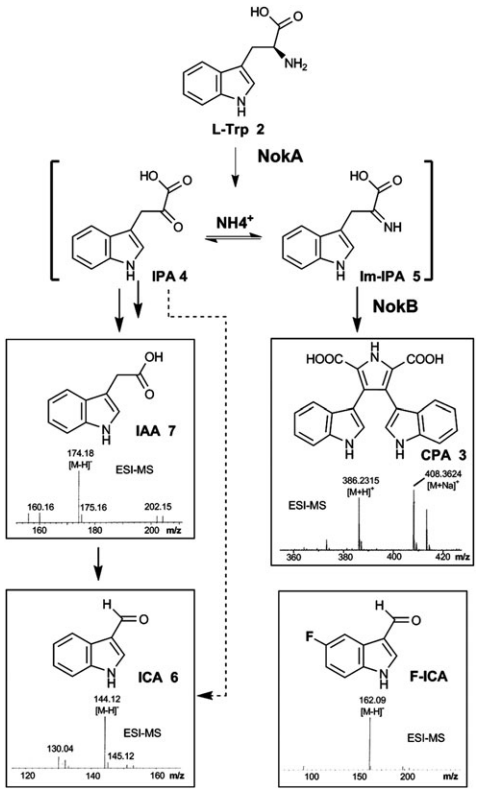



**Fig. 3** Spectrometric identification of the IPA-derived products from the NokA reaction by HPLC and NMR ($CD_3OD$) analysis. (A) HPLC trace monitored at 275 nm for the NokA reaction. (B) $^1$H NMR spectrum of ICA **6**. In the inset is shown the aldehyde signal. (C) $^1$H NMR spectrum of the unknown indole compound **8**. (D) $^1$H NMR spectrum of IAA **7**. All three IPA-derivatives show characteristic signals of aromatic indole ring.

(see Fig. 3). The minor product peak at $R_t$ 27.4 min was identified to be indole-3-acetic acid (IAA, **7**), as also verified by co-elution with an authentic standard; whereas the other at $R_t$ 26.3 min to be an unknown indole derivative **8**, as shown in Fig. 2 and 3. In rebeccamycin biosynthesis, Im-IPA has been proposed to serve as the oxidation product of RebO (a NokA homolog), but the Im-IPA could never be found and isolated.[25] In principle, Im-IPA could readily be hydrolyzed to give IPA, due to its inherent instability. Based on our experimental results, the three indole products may thus be derived from metabolic consequence of IPA by *E. coli* enzymes in the cell-free system, or alternatively from chemical degradation of IPA,[27] as illustrated in Scheme 2. The former proposition has been supported by several previous studies, where IPA could be transformed to IAA and ICA by microbial and plant enzymes.[28–33] Moreover, control experiments with L-Trp incubated with the *E. coli* cell-free crude extract lacking NokA did not give appreciable degradation products, suggesting these indole compounds must be derived from the NokA reaction product, Im-IPA. Therefore, above interesting findings have firmly verified the catalytic function of NokA to be an L-amino acid oxidase, as proposed, required for K-252c biosynthesis (Scheme 1).

In an attempt to improve NokA activity in the cell-free system, we also examined possible effects of exogenous ammonium hydroxide ($NH_4OH$) on the NokA activity. Very interestingly, the experiments led to the same production of the three indole products, but the catalytic rate of NokA was greatly improved by at least 4 fold in the presence of




**Scheme 2** The proposed reaction pathways of NokA and NokB with L-Trp in the *E. coli* cell-free expression system. The F-ICA was generated from an incubation of cell-free crude NokA with F-L-Trp **9**.

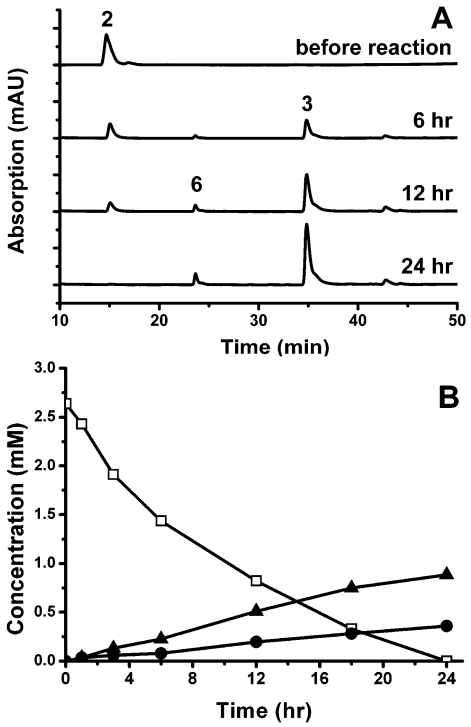



**Fig. 4** RP-HPLC analyses of the reactions catalyzed by the NokA and NokB encoded by pCY20. (A) HPLC traces monitored at 300 nm for the reactions (without exogenous ammonium hydroxide) quenched at 0, 6, 12 and 24 h. (B) Complete time courses of L-Trp **2** (□), CPA **3** (▲), the NokA IPA-derived products (●) from the reactions analyzed at different reaction time points (0, 3, 6, 12, 18 and 24 h).

exogenous 30 mM $NH_4OH$ as shown in Fig. 2B. Further experiments on pH-dependent examination of NokA activity revealed that the observed improvement in NokA activity was mainly attributed to the resulting higher pH of reaction solution caused by $NH_4OH$ (data not shown). Based on above results, our study has demonstrated a facile production of IAA (a key auxin in most plants)[29–31,34] and ICA (an antibacterial agent),[35,36] for the first time, by the simple *E. coli* cell-free expression system carrying *nokA*. Since the use of $NH_4OH$ also increases solubility of L-Trp, the above finding could also facilitate larger-scale production of the compounds.

### *In vitro* biochemical characterization of NokA and NokB by tandem enzymatic reactions

Upon successful characterization of NokA by the heterologous expression system, we further pursued to examine the NokA and NokB coupled reactions by *in vitro* tandem reactions of the enzymes with L-Trp **2**. In principle, the NokB reaction would have to compete with the degradation pathway of IPA/Im-IPA described earlier for availability of the substrates, Im-IPA and/or IPA. For efficient synthesis of CPA or K-252c by the tandem enzymatic reactions, one must figure out a simple way to monitor the coupled reactions of NokAB or NokABCD, respectively, so as to avoid the IPA/Im-IPA degradation. We have, in the companion paper, reported the production of CPA by *in vitro* heterologous co-expression of the NokABCD enzymes encoded by pCY20.[20] To thoroughly characterize the NokA/NokB coupled reactions, the *E. coli* expression system of pCY20 was thus adapted here for further experiments. Therefore, time-dependent assays of the NokABCD with L-Trp were carried out to analyze the reaction progress of NokA/NokB. Consequently, the incubation of L-Trp with the cell-free crude extract of NokABCD indeed resulted in formation of CPA ($R_t$ 34.8 min), as expected, in a time-dependent manner, which also correlated well with consumption of L-Trp as revealed by RP-HPLC (Fig. 4). Nonetheless, the CPA formation was found to be also accompanied by concurrent production of the IPA-derivatives (**6**, **7** and **8**) corresponding to the formation of Im-IPA by NokA. Based on RP-HPLC analyses shown in Fig. 4, the tandem enzymatic reactions of NokA and NokB led to about quantitative conversion of two molecules of L-Trp to one molecule of CPA, where Im-IPA could be detected as a reaction intermediate in the form of the IPA-derived products.

In light of the improved catalytic efficiency on NokA by the ammonium hydroxide, we were prompted to investigate its possible effects on the tandem reactions of NokA and NokB. This may also help to resolve the actual substrate for NokB by monitoring the balance in equilibrium among IPA, Im-IPA and an IPA enamine (Ea-IPA, an Im-IPA tautomer) with ammonium hydroxide. Hence, the tandem NokA/NokB reactions were treated with L-Trp in the presence of various concentrations of $NH_4OH$ (0–100 mM), and the reactions for different time peroids were subjected to RP-HPLC analysis. As a consequence, the catalytic efficiency of the NokA/NokB reactions was also greatly improved and the CPA production

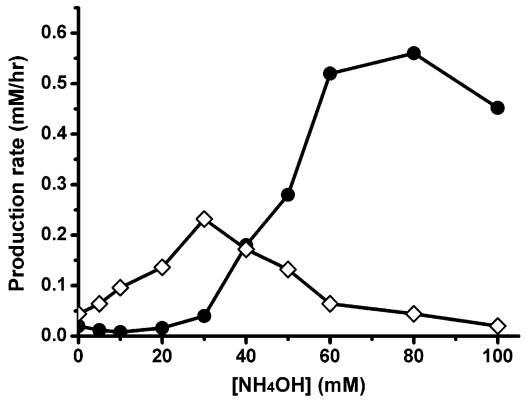



**Fig. 5** The effect of exogenous ammonium hydroxide concentration on observed production rates of CPA (◇) and the NokA IPA-derived products (●) by HPLC. The reaction mixture containing L-Trp (4 mM), aqueous $NH_4OH$ (0–100 mM) and the crude extract of the *E. coli*/pCY20 was incubated at 30 °C for 3 h and subsequently quenched for RP-HPLC analysis.

rate reached the maxmium also at *ca.* 30 mM of exogenous ammonium hydroxide, as shown in Fig. 5. In Fig. S1 (see ESI†) are also shown the [$NH_4OH$] and time correlation plots of the tandem NokA/NokB enzymatic activity. Notably, the overall catalytic efficiency gradually decreased as the exogenous $NH_4OH$ concentration increased from 30 mM to 100 mM, where the HPLC peaks derived from Im-IPA continued to grow and reached the maxmium at *ca.* 80 mM of $NH_4OH$ (Fig. 5).

Walsh *et al.* previously proposed that Im-IPA could couple with Ea-IPA or another molecule of Im-IPA to serve as the co-substrates of RebD (a NokB homolog) in rebeccamycin biosynthesis, as illustrated in Scheme S1 (see ESI†).[37,38] Our experiments also showed that, in the absence of exogenous $NH_4OH$, an exogenous addition of IPA into the NokA/NokB reactions with L-Trp did not alter the observed catalytic rate, indicating that IPA could not act as a direct substrate for NokB (data not shown). In principle, an increase in the $NH_4OH$ concentration could result in a higher molar ratio of [Im-IPA]/[IPA] in solution. The observation of the improved catalytic rate on the NokA/NokB reactions may thus be explained by the increase of Im-IPA in concentration, thereby supporting that Im-IPA acts as a direct substrate for NokB. On the other hand, the gradually reduced NokA/NokB activity in the higher concentration of $NH_4OH$ than 30 mM could be attributed to two possible causes. In the case of Im-IPA and Ea-IPA (1 : 1 equiv.) acting as co-substrates for NokB, the higher concentration of $NH_4OH$ would result in loss of the balance between Im-IPA and Ea-IPA, thereby leading to a gradual reduction in the overall production of CPA by the NokA/NokB reactions. In the instance of two Im-IPA molecules serving as co-substrates for NokB, the observed drop in CPA production in 30–100 mM $NH_4OH$ could simply be caused by inhibition of NokB reaction by high concentration of $NH_4OH$. In addition, at higher concentration (30–80 mM) of $NH_4OH$ the observed increase in overall production of the IPA-derived products from NokA reaction may be attributed to the improved catalytic efficiency of NokA by $NH_4OH$ as addressed earlier, and/or the gradual shutdown of the NokB reaction (Scheme 2). It should also be noted that, as exogenous $NH_4OH$ concentration increased from 0 to 30 mM, the NokA/NokB reactions accumulated little IPA-derived (NokA) products, whereas CPA, however, increased accordingly in production. This result may suggest that the NokA reaction be rate-limiting in the cell-free coupled expression system (Scheme 2) under the described condition.

Since the addition of ammonium hydroxide also caused variation in final pH of the reactions, we conducted, in parallel, the pH-dependent profile of NokA/NokB activity as a control to rule out possible interference from the pH variation. As shown in Fig. 6 the variation in pH, ranging from pH 7.6 to pH 9.7, did not cause the observed dramatic effect on the NokA/NokB activity in association with the exogenous ammonium hydroxide. In fact, the activity displayed by the NokA/NokB reaction showed only slight dependence on the pH. It is therefore concluded, for the first time, that the exogenous addition of ammonium hydroxide can greatly promote the catalytic efficiency of the tandem enzymatic reactions of NokA and NokB, and the coupled enzyme reactions can be successfully dissected by the simple use of $NH_4OH$. Consequently, the experiment has led to a better understanding of NokA and NokB catalyses and their cross-talking in K-252c biosynthesis.

### Substrate specificity of the NokA and NokB reactions

To explore the combinatorial potential of the K-252c biosynthesis, we also carried out a pilot experiment to examine substrate tolerance of NokA and NokB towards 5-fluoro-L-tryptophan (F-Trp, **9**). Fluorine has become a common substitute of hydrogen atoms in pharmacologically active agents,[39,40] and an incorporation of F at various positions may improve structural diversity of K-252a and its analogs. In principle, the van der Waals radius of F is similar to that of H, and the substitution of H with F usually would not disturb proper positioning of a fluorinated substrate analogue in an enzyme active site. In light of the above prospects, we conducted an enzymatic incubation of the NokA and NokB with F-Trp **9** (4 mM) in the absence or presence of exogenous 30 mM $NH_4OH$. As a consequence, the incubation led to

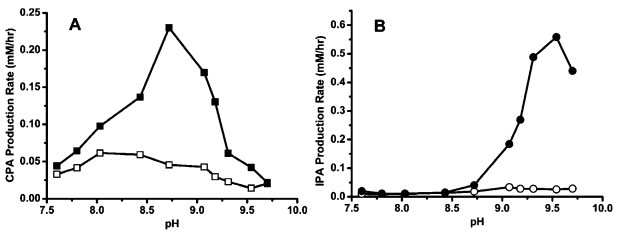
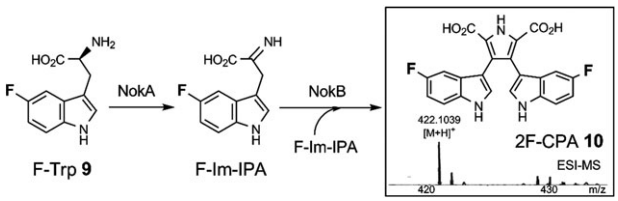



**Fig. 6** Comparison of the effects caused by exogenous ammonium hydroxide and pH on the NokA/NokB activities. Panel A: the production rates of CPA at different concentrations (0, 5, 10, 20, 30, 40, 50, 60, 80 and 100 mM) of $NH_4OH$ (■) and at different pH (□). Panel B: the production rate of IPA/Im-IPA, indicated as the IPA-derived products, in the NokA/NokB reactions at the same corresponding concentrations (0–100 mM) of $NH_4OH$ (●) and at different pH (○). The above different concentrations (0–100 mM) of exogenous $NH_4OH$ resulted in final pH of 7.6, 7.8, 8.0, 8.4, 8.7, 9.1, 9.2, 9.3, 9.5 and 9.7, respectively.




**Scheme 3** Enzymatic conversion of 5-fluoro-L-tryptophan (F-Trp) to 9,9′-difluoro-chromopyrrolic acid (2F-CPA) by NokA and NokB.

production of 9,9′-difluoro-chromopyrrolic acid (2F-CPA, **10**) as an end product ($R_t$ 36.0 min), as shown in Scheme 3, at a comparable catalytic rate as that with L-Trp. Moreover, there appeared to be no measurable difference in the consumption of the substrate between the reactions with L-Trp and with F-Trp, as judged by RP-HPLC analysis. The chemical structure of 2F-CPA **10** was further confirmed by extensive NMR analyses (*e.g.*, $^{19}$F, gCOSY and gHMQC) and high resolution ESI-MS spectroscopy. This intriguing observation suggests that, under the experimental conditions, the strong electron-withdrawing feature of F seems to neither disturb chemical events in the catalysis through electronic effect, nor reduce the binding affinity of F-Trp to the active sites of both NokA and NokB. In a similar fashion, an incubation of NokA alone with F-Trp also resulted in isolation of 5-fluoro-indole-3-carboxaldehyde (F-ICA), as verified by NMR and Mass spectroscopy. Therefore, our study with F-Trp has clearly demonstrated the possibility of utilizing the enzymes in K-252c biosynthesis, which may couple with the N-Gtf (NokL), to expand the combinatorial potential of K-252a biosynthesis.

Furthermore, we have also examined the stereospecificity of NokA by incubation of the cell-free crude NokA, encoded by pJZ22, with D-tryptophan (D-Trp, 4 mM) in the presence of exogenous $NH_4OH$ (30 mM). Consequently, the experiment resulted in little activity of NokA catalysis, as compared to the reaction with L-Trp under the same condition. The result indicated that NokA, previously proposed to be an L-amino acid oxidase, indeed exhibited high stereoselectivity in substrate recognition.

### *In vitro* functional characterization of the N-glycosyltransferase gene

The N-Gtf gene (*nokL*) of the gene cluster plays a central role in determining the glycosyl pattern of the indolocabazole metabolites in *Nocardiopsis sp*. K-252. Hence, characterization of the functional role and substrate specificity of NokL represents an important issue of this study. Sequence analysis of *nokL* revealed a suitable ribosome binding site (*i.e.* GGAGG) *ca.* 7 nucleotides upstream of TTG, a start codon used in *Streptomyces*, and gave a deduced product of 436 amino acids bearing a conserved N-terminus His-14 in alignment with the catalytic base proposed in some O-Gtfs, *e.g.*, UGT71G1 and OleI.[41] To functionally characterize NokL *in vitro*, *nokL* was amplified from pJC3B5 by PCR and ligated into pET28a vector for heterologous expression in *E. coli* BL21(DE3).[20] The derived construct failed to express N-terminal $His_6$-tag NokL after IPTG induction. Another attempt with a pET21b-based construct coding for C-terminal

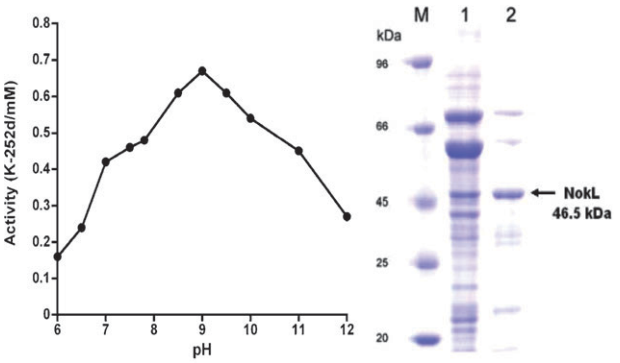
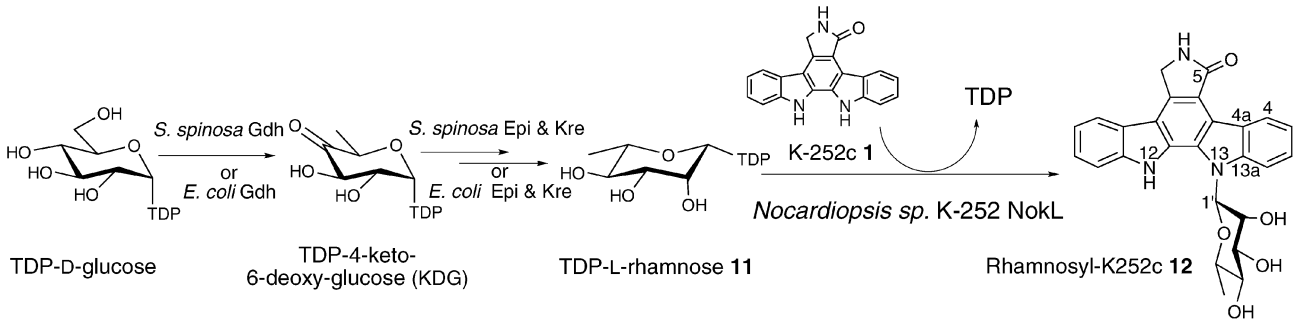



**Fig. 7** (A) The pH-dependent profile of NokL activity. (B) SDS-PAGE analysis of NokL expression in *E. coli*. The wild-type *nokL* and pG-KJE7 were co-expressed in *E. coli*. The protein was analyzed on a 10% SDS-PAGE with Coomassie Blue staining. Lane M: molecular weight standard; lane 1: cell-free crude extract of NokL; lane 2: insoluble proteins.

$His_6$-tag NokL gave fair amount of soluble protein, which however could not to be purified by Ni-NTA affinity chromatography. Hence, *nokL* was cloned for expression as a wild-type recombinant protein to avoid possible interference from poly-His tag. As a result, the extremely low yield of soluble NokL was found even under the optimized condition (1 mM IPTG, overnight at 30 °C). The problem was finally resolved by co-expression of the wild-type construct with pG-KJE7. Consequently, the wild-type NokL (*ca.* 46.5 kDa) was obtained in sufficient quantity in the form of soluble cell-free crude extract for further experiments, as revealed by SDS-PAGE (Fig. 7B).

Thus far, the dihydrostreptose biosynthetic pathway has not yet been well characterized; hence, the natural sugar donor was not available for NokL activity assays. We attempted to utilize an alternative substrate to functionally characterize the N-Gtf activity, with an additional advantage of exploring combinational potential and substrate specificity of NokL. Interestingly, Yasuzawa reported production of K-252a, K-252b, K-252c and K-252d in culture broths of *Nocardiopsis* strains K-252 and K-290,[12] where both rhamnose and dihydrostreptose moieties were incorporated into the indolocarbazole core. This finding may suggest the possibility that the N-Gtf may be able to utilize NDP-rhamnose as an alternative substrate. A common N-Gtf might possess substrate flexibility for efficient synthesis of indolocarbazole compounds in *Nocardiopsis*. To examine this possibility, we therefore carried out the *in vitro* NokL enzymatic assays with K-252c **1** and TDP-L-rhamnose (TDP-Rha, **11**) as possible substrates (Scheme 4). Synthesis of TDP-Rha was accomplished, according to our recent report,[42] by *in vitro* tandem enzymatic conversion of TDP-D-glucose with gene products of *gdh*, *epi* and *kre* from *Saccharopolyspora spinosa*. Consequently, the incubation of NokL crude extract with K-252c **1** and TDP-Rha **11** (1 : 1 equiv.) led to production of a new product in a time-dependent manner, as analyzed by RP-HPLC (Fig. 8). High-resolution MALDI-TOF analysis of the product revealed $[M]^+$ at $m/z$ 457.177 and $[M + H]^+$ at $m/z$ 458.197, as expected for rhamnosyl-K252c **12** (calcd. M.W. 457.163). Further tandem mass spectrometry (MS/MS)




**Scheme 4** The *in vitro* enzymatic synthesis of rhamnosyl-K252c (K-252d) by NokL.

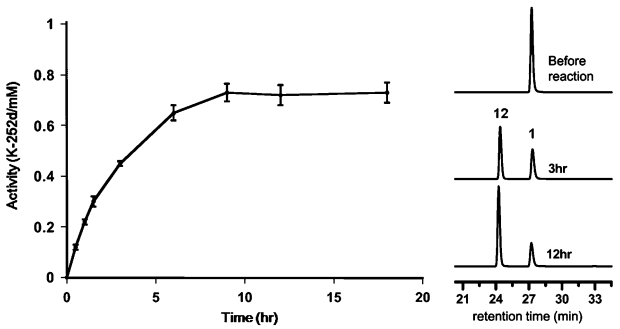



**Fig. 8** Time-dependent HPLC traces of *in vitro* NokL reactions.

found the aglycone $[M]^+$ ion at $m/z$ 311.193 from fragment patterns of the product (Fig. S5, ESI†). Upon $^1$H-NMR, $^{13}$C-NMR and COSY, the rhamnosylated product was fully assigned with chemical shifts, in excellent agreement with those of K-252d reported by Yasuzawa.[12]

Notably, the unique asymmetrical structural feature of **1** could result in two isomeric forms of the product, derived from nucleophilic attack of either one of indole nitrogen atoms at C-1′ of TDP-Rha **11**. The situation was resolved by gHMBC correlation experiments, where carbon 13a ($\delta_c$ 142.24 ppm) was correlated to both protons at H-4 ($\delta_H$ 9.40 ppm) and H-1′ ($\delta_H$ 6.54 ppm).[12] The result distinctly indicated that the rhamnose moiety was specifically attached to the N-13 atom near to the C-5 carboxyl group. As shown in Fig. 8, the HPLC profile demonstrated successful rhamnosylation of K-252c **1** to afford the expected product **12**, consistent with our proposition that NokL might accept TDP-Rha. Moreover, the enzymatic conversion reached maximum after overnight incubation. NokL exhibited optimum sugar transferase activity at pH 9.0 (Fig. 7A, see also ESI), similarly observed in some *O*-Gtfs, such as GtfA.[43] Interestingly, an exogenous $Mg^{2+}$ ion (10–15 mM) was found to enhance the activity by at least 2 fold.

### Substrate specificity of the K-252a *N*-glycosyltransferase

In light of striking substrate flexibility displayed by NokL to utilize TDP-Rha, we were prompted to examine its degree of substrate tolerance towards various sugar donors and acceptors. Interestingly, further incubations of NokL with TDP-D-galactose, TDP-D-mannose and UDP-D-glucose did not show detectable activity. This intriguing observation indicates that the active site of NokL may present a peculiar 3-dimensional configuration to selectively accommodate a sugar donor of certain conformation like TDP-Rha **11**, which is very different from the natural donor (NDP-2-deoxy-dihydrostreptose, Scheme 1) in terms of sugar structure and ring size. This statement is also consistent with the observation that TDP-6-deoxy-D-glucose was found to be a possible substrate for NokL as judged by RP-HPLC, although its conversion was not high enough for product isolation and characterization. To probe the substrate specificity of NokL towards sugar acceptors, several indole structural analogs were used to examine NokL substrate specificity: L-tryptophan **2**, chromopyrrolic acid (CPA, **3**), carbazole, carprofen, pindolol and 2-hydroxybenzo[a] carbazole-3-carboxylic acid (Fig. S7, ESI†). Consequently, NokL showed remarkably stringent specificity towards the aglycone, in that RP-HPLC revealed no product formation for those tested. Based on these results, NokL appeared to preferentially act on a bisindole or "closed" planar aglycone, instead of the open one like CPA, as similarly observed in *in vivo* experiments with RebG.[44]

To further dissect the contribution of the NDP-sugar structure to molecular recognition adapted by NokL, TDP and UDP were examined as potential inhibitors against TDP-Rha under the assay condition. Consequently, TDP (0.6 mM) inhibited *ca.* 20–40% of NokL total activity on TDP-Rha within 12 h, whereas UDP (0.6 mM) did not show obvious inhibition. The result may suggest that the enzymes of the biosynthetic gene cluster utilize TDP as a definite structural counterpart of NDP-sugar, and, in the TDP-sugar structure, TDP alone donates a substantial contribution to the molecular recognition.

### Heterologous production of K-252d by the *E. coli* expression system of NokL

In analogy to K-252a, K-252d (rhamnosyl-K252c, **12**) is also a potent protein kinase C (PKC) inhibitor.[23] The PKC family plays important roles in cellular proliferation and signal transduction. Hence, specific inhibitors against PKC are promising antitumor drugs for cancer chemotherapy.[45,46] Recent efforts in medicinal chemistry have aimed to make derivatives of this class of compounds, *e.g.*, UCN-01, GGP-41251 and CEP-751, with higher target specificity.[47–49] Therefore, an efficient method to produce the compounds is also very important. In this study, we have successfully provided an alternative method to make K-252d by an heterologous expression system, which utilized NokL from the

K-252a biosynthetic pathway and the three TDP-Rha biosynthetic enzymes (Gdh, Epi and Kre) from spinosyn biosynthetic pathway (Scheme 4). Nonetheless, in several strains of *E. coli* homologous genes of *gdh*, *epi* and *kre* have been found.[50–53] To more efficiently generate K-252d or its derivatives in the *E. coli* host, we further investigated the possibility of utilizing native homologs of the three TDP-Rha biosynthetic enzymes from the *E. coli* for *in situ* generation of TDP-Rha **11** as shown in Scheme 4. As a result, an incubation of TDP-D-glucose alone with the NokL cell-free crude extract and K-252c **1** indeed, but surprisingly, led to formation of K-252d **12** as a sole product at a comparable rate similarly as demonstrated in Fig. 8, indicating the N-glycosylation by NokL was a rate-limiting step in the *in situ* tandem enzymatic reactions (Scheme 4). Very interestingly, the same observation was also obtained with another incubation of the NokL crude extract and K-252c directly with TDP-4-keto-6-deoxy-glucose (KDG), the product of Gdh. Together, the results evidenced an efficient, simple approach successfully accomplished to synthesize K-252d, whereas NokL clearly showed much higher substrate specificity towards TDP-Rha than TDP-D-glucose and KDG.

## Materials and methods

### Materials

Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs (NEB). *E. coli* and *pfu* DNA polymerase were purchased from Stratagene. The pUC19 cloning vector and the pET expression vectors were purchased from NEB and Novagen, respectively. DNA primers were purchased from Operon Biotech Inc. Unless specified otherwise, all chemicals were purchased from Sigma.

### Bacterial strains, culture conditions and DNA manipulations

The pET21b was used to overexpress the genes. The *pfu* DNA polymerase was routinely utilized in polymerase chain reaction (PCR). The mini-preparation of DNA was performed on QIAprep spin miniprep kit (Qiagen). The pJC3B5 fosmid, pJZ22 plasmid and pCY20 plasmid were obtained as described in the companion paper for the cloning and sequence analysis of the gene cluster for biosyntheses of K-252a and its analogs in *Nocardiopsis sp.* K-252 (NRRL15532, *Nonomuraea longicatena* K-252T).[20] The *E. coli* XL1-Blue and *E. coli* BL21-Codon Plus (DE3)-RP served as hosts for routine subcloning and protein expression, respectively, under standard culture conditions as described by Sambrook *et al.*[54] Sequences of PCR primers are listed in Tables S1 (see ESI†).

### Heterologous expression of wild-type recombinant NokA and NokA/NokB in *E. coli*

To overexpress *nokA*, *E. coli* BL21 (DE3) cells, transformed with pJZ22, were grown at 37 °C in Luria-Bertani (LB) medium with 100 μg ml$^{-1}$ ampicillin and 30 μg ml$^{-1}$ kanamycin until $OD_{600}$ reached 0.5. After induction with 250 μM isopropyl β-D-l-thiogalacto-pyranoside (IPTG), the culture grew at 25 °C for 10 h. The cells were then harvested by centrifugation at 4 °C (1902 g, 20 min). Cells resuspended in 30 ml of Tris buffer (104 mM Tris-HCl, 10% (v/v) glycerol, pH 7.6) were subsequently disrupted by sonication at 4 °C. Cell debris was removed by centrifugation (15 700 g, 20 min) and the resulting cell-free crude extract was used for enzyme assays. For expression of NokA in the presence of the chaperones, the cells co-transformed with pJZ22 and pG-KJE7 were induced with 250 μM IPTG and 0.2% (w/v) L-arabinose. To co-express NokA and NokB for biochemical characterization of the NokA/NokB coupled reactions, the co-expression construct (pCY20) of NokABCD and pG-KJE7 were co-transformed into *E. coli* BL21 (DE3) cells and the same expression condition was adapted as described for NokA with the chaperones.

### *In vitro* time-dependent enzymatic assays of NokA and NokA/NokB

The *in vitro* time-dependent enzymatic assays of NokA or NokA/NokB were carried out similarly with the cell-free crude extracts of NokA or NokABCD, respectively, prepared as described above in the presence of the chaperones encoded by pG-KJE7. For a typical assay, the reaction contained 2.6 mM substrate (L-tryptophan **2**) and 40 μl cell-free crude extract in a reaction buffer (80 mM Tris-HCl, pH 7.8, 7.6% (v/v) glycerol) was incubated in a total volume of 52 μl. The reaction was quenched at specific time points by an equal volume of ice-cold MeOH, and was then subjected to reverse phase (RP-) HPLC analysis using by Agilent 1100 HPLC series equipped with quaternary pump and diode-array detector. For the NokA reaction, the cell-free crude extract of NokA was used in the absence or presence of exogenous ammonium hydroxide ($NH_4OH$, 30 mM). To examine the effect of exogenous $NH_4OH$ on the NokA/NokB coupled reactions, the same reaction mixture was co-incubated with a series of concentrations (0–100 mM) of exogenous $NH_4OH$. The [$NH_4OH$]/time correlation plots for the NokA/NokB activity measurements are shown in Fig. S1 in ESI.† As described for the time-dependent assays of NokA/NokB, the pH-dependent measurements of the NokA/NokB activities were conducted at a reaction time of 3 h over a pH range from 7.6 to 9.7 without exogenous $NH_4OH$.

### HPLC analyses of the NokA and NokA/NokB reactions

The *in vitro* enzymatic assays of NokA or NokA/NokB coupled reactions were analyzed by an ODS-C18 RP-HPLC analytical column (4.6 × 250 mm, 5 μm, Zorbax, Agilent) eluted with mobile solution consisted of methanol (solvent A), 2.5 mM aqueous potassium phosphate at pH 3.5 (solvent B) and acetonitrile (solvent C). The RP-HPLC elution was conducted with 8% solvent A at a flow rate of 1.0 ml min$^{-1}$, whereas the elution gradient was programmed as follows: 92% solvent B in 5 min, 0–16% solvent C over 10 min, 16–28% solvent C over 15 min, 28–70% solvent C over 15 min, 70–80% solvent C over 2 min, 80% solvent C held for 5 min, 80% to 0% solvent C over 3 min, and then 92% solvent B held for 5 min. The elution profiles were monitored within full-range UV-VIS wavelength, where the specific wavelength for optimal detection was found at 275 nm for L-Trp **2**, CPA **3** and their analogs, and 300 nm for ICA **6** and its analogs.

### Substrate specificity of the NokA/NokB reactions

The time-dependent assays of the NokA/NokB reactions were carried out with L-tryptophan **2** or 5-fluoro-L-tryptophan **9**, using the described conditions for the enzymatic assays of the NokABCD proteins. The reactions were conducted with or without exogenous $NH_4OH$ to measure production of enzymatic intermediates or products by analytical RP-HPLC as described. Similar reaction and analytical conditions were also applied on the enzymatic assays of the cell-free crude extract of NokA with D-tryptophan.

### Preparation and characterization of the IPA-derived products from NokA reaction

To prepare the three IPA-derived products (ICA **6**, IAA **7** and the indole compound **8**) from the NokA reaction, a reaction mixture (104 μl) containing 4 mM L-tryptophan **2**, and 80 μl cell-free crude extract of NokA in a reaction buffer (80 mM Tris-HCl, 7.6% (v/v) glycerol, pH 7.8) was incubated at 30 °C for 24 h. After removal of the protein by centrifugation, the reaction mixture was extracted with an equal volume of ethyl acetate to yield a crude mixture of the IPA-derived products upon removal of ethyl acetate by evaporation. The crude mixture, prepared from 19.8 mL of reaction, was further purified by a semi-preparative C18 RP-HPLC column (5C18-AR-I, 8.0 × 250 mm, 5 μm, Cosmosil, Nacalai tesque, Japan) eluted with the mobile phase consisting of methanol (solvent A), acetonitrile (solvent B) and water (solvent C). The HPLC gradient for elution at 3 ml $min^{-1}$ was programmed as follows: 8% solvent A held over the entire gradient; 92% solvent C in 5 min, 0–23% solvent B over 5 min, 23% solvent B held for 5 min, 23–25% solvent B over 5 min, 25–26% solvent B over 5 min, 26–80% solvent B over 2 min, 80% solvent B held for 5 min, and then 80% to 0% solvent B over 3 min. Finally, the column was equilibrated with 92% solvent C for 5 min. The eluted fraction corresponding to each of the IPA-derived products was then lyophilized to yield **6**, **7** or **8**. The structural elucidation of the products was performed with NMR and mass spectroscopy. NMR spectra of the enzymatic products were taken on a Varian INOVA-500 (500 MHz) NMR spectrometer. ICA (**6**): $^1$H-NMR ($CD_3OD$, 500 MHz), $\delta_H$ 7.170 (1H, ddd, $J$ = 1, 7.5 Hz), 7.212 (1H, ddd, $J$ = 1.5, 8 Hz), 7.410 (1H, d, $J$ = 8 Hz), 8.031 (1H, s), 8.091 (1H, d, $J$ = 8 Hz), and 9.822 (1H, s) ppm. $^{13}$C-NMR ($CD_3OD$, 125 MHz), $\delta_C$ 113.122, 120.132, 122.384, 123.611, 124.998, 125.722, 138.940, 139.673, and 187.406 ppm. ESI-MS calculated for $C_9H_7NO$ $[M–H]^-$ 144.05; found 144.12. The gCOSY and gHMQC 2D-NMR results are shown in Fig. S2.† IAA (**7**): $^1$H-NMR ($CD_3OD$, 500 MHz), $\delta_H$ 3.647 (2H, s), 6.976 (1H, dd, $J$ = 7.5 Hz), 7.060 (1H, dd, $J$ = 7.5 Hz), 7.136 (1H, s), 7.306 (1H, d, $J$ = 8 Hz), and 7.551 (1H, dd, $J$ = 8.5 Hz) ppm. ESI-MS calculated for $C_{10}H_9NO_2$ $[M–H]^-$ 174.06; found 174.18.

### Enzymatic preparation and structural characterization of 9,9′-difluoro-chromopyrrolic acid

To prepare 9,9′-difluoro-chromopyrrolic acid (2F-CPA, **10**) from the NokA/NokB reactions, a reaction mixture (104 μl) containing 5-fluoro-D,L-tryptophan (4 mM), ammonium hydroxide (30 mM) and the NokABCD cell-free crude extract (80 μL) in a reaction buffer (80 mM Tris-HCl, 7.6% (v/v) glycerol, pH 7.8), was incubated at 30 °C for 24 h. Subsequent purification and analysis to obtain 2F-CPA were similarly as described for CPA **3** in the companion paper.[20] The crude 2F-CPA from 86.3 mL of reaction was purified by semi-preparative RP-HPLC using the 5C18-AR-II column (8.0 × 250 mm, 5 μm, Cosmosil, Nacalai tesque, Japan) and the mobile phase consisting of methanol (solvent A), acetonitrile (solvent B) and deionized water (solvent C). The product was purified by a HPLC elution gradient employed at a flow rate of 3.0 ml $min^{-1}$ with 8% solvent A as follows: 0% solvent B in 5 min, 0–16% solvent B over 5 min, 16–26% solvent B over 5 min, 26–27% solvent B over 25 min, 27–80% solvent B over 5 min, 80% solvent B held for 5 min, 80% to 0% solvent B over 5 min, and then 0% solvent B for 5 min. After removing solvents, 2F-CPA **10** was obtained as a light yellow powder (3.5 mg, *ca.* 99% purity as judged by RP-HPLC analysis). $^1$H-NMR ($CD_3OD$, 500 MHz), $\delta_H$ 6.766 (2H, ddd, $J$ = 2.0, 10 Hz), 6.776 (2H, d, $J$ = 9.5 Hz), 7.028 (2H, s), and 7.221 (2H, dd, $J$ = 4.0, 8.5 Hz) ppm. $^{13}$C-NMR ($CD_3OD$, 125 MHz), $\delta_C$ 105.348, 109.824, 112.443, 124.320, 126.191, 128.023, 129.504, 133.939, 157.916, 159.763, and 163.900 ppm. $^{19}$F-NMR ($CD_3OD$, 470 MHz), $\delta_F$ −124.425. High resolution ESI-MS calculated for $C_{22}H_{13}F_2N_3O_4$ $[M+H]^+$ 422.095; found 422.104. The gCOSY and gHMQC 2D-NMR results are shown in Fig. S3.†

### Enzymatic preparation and structural characterization of 5-fluoro-indole-3-carboxaldehyde

For enzymatic synthesis of 5-fluoro-indole-3-carboxaldehyde (F-ICA), a reaction mixture (104 μl), containing 5-fluoro-D,L-tryptophan (4 mM), ammonium hydroxide (80 mM) and the NokA cell-free crude extract (80 μl) in a reaction buffer (80 mM Tris-HCl, 7.6% (v/v) glycerol, pH 7.8), was incubated at 30 °C for 24 h. The crude F-ICA from 94.6 ml of reaction was subjected to subsequent preparation and analysis similarly as described for ICA **6**. After RP-HPLC purification, F-ICA (5.1 mg, *ca.* 99% purity as judged by RP-HPLC) was obtained. $^1$H-NMR ($CD_3OD$, 500 MHz), $\delta_H$ 6.947 (1H, ddd, $J$ = 2.5, 9 Hz), 7.353 (1H, dd, $J$ = 4.5, 9 Hz), 7.708 (1H, dd, $J$ = 2.5, 9 Hz), 8.040 (1H, s), and 9.771 (1H, s) ppm. $^{13}$C-NMR ($CD_3OD$, 125 MHz), $\delta_C$ 107.435, 113.072, 114.244, 120.107, 126.400, 135.343, 140.671, 161.030, and 187.244 ppm. $^{19}$F-NMR ($CD_3OD$, 470 MHz), $\delta_F$ −119.095 ppm. ESI-MS calculated for $C_9H_6FNO$ $[M–H]^-$ 162.04; found 162.09. The gCOSY and gHMQC 2D-NMR results are shown in Fig. S4.†

### Construction of the NokL expression plasmid

The *nokL* gene was amplified on pJC3B5 by PCR with a forward primer with an *Nde*I site and a reverse primer with a stop codon (TGA) followed by an *Eco*RI site near 5′-end. The amplified PCR product was ligated with blunt-ended pUC19 at *Sma*I to generate pCY15. After digestion of pCY15 with *Nde*I and *Eco*RI, the digestion fragment carrying *nokL* was cloned into pET21b vector to give pCY16 for the wild-type NokL expression experiments. For plasmid construction of

N-terminal $His_6$-tagged NokL, the same digestion fragment was cloned into pET28a to afford pCY17. For *C*-terminal $His_6$-tagged NokL expression, the *nokL* gene was amplified on pCY16 by PCR with primer pairs of NKLNdF1 (forward, with *Nde*I) and NKLXR1 (reverse, with *Xho*I). The resulting PCR product was subsequently cloned into pET21b at the corresponding sites to yield pMS4.

### Preparation of the NokL cell-free crude extract

The pCY16 (NokL) and pG-KJE7 (chaperones) plasmids were co-transformed into *E. coli* BL21 (DE3). Cells were grown at 37 °C in LB medium with antibiotics (100 μg $ml^{-1}$ ampicillin and 30 μg $ml^{-1}$ kanamycin) until $OD_{600}$ reached 0.5. After induction with 0.1% (w/v) L-arabinose and 1 mM IPTG, the culture was allowed to grow at 30 °C for additional 20 h. All procedures for the preparation of cell-free crude extract were carried out on ice or at 4 °C. The cells were harvested by centrifugation (3200 g, 15 min), followed by resuspension with potassium phosphate buffer (20 mM $K_2HPO_4$, pH 7.8, 15% glycerol). Cells were broken and disrupted by two passages through a French press cell (Spectronic Instruments) at 16 000 psi. After removal of cell debris by centrifugation at 16 000 g for 20 min, the desired cell-free crude extract was obtained for further enzymatic reactions.

### HPLC methods for analyses of NokL reactions

The NokL assays were analyzed by using an analytical C18 column (5C18-ARII, 4.6 × 250 mm, 5 μm, Cosmosil, Nacalai tesque, Japan) with two stages of linear gradient: (1) 0–36% acetonitrile, 8% methanol and 92-56% potassium phosphate (2.5 mM, pH 3.5) in 10 min; (2) 36–40% acetonitrile, 8% methanol and 56–52% potassium phosphate (2.5 mM, pH 3.5) in 15 min, and the elution was monitored at 290 nm. In some cases, K-252c **1** (retention time at 27.2 min) and rhamnosyl-K252c **12** (retention time at 24.2 min) were recorded with full range of UV-Visible wavelengths, where both displayed essentially identical absorbance spectra with $\lambda_{max}$ at *ca.* 290 nm. Purification of rhamnosyl-K252c (K-252d, **12**) was carried out on a semi-preparative RP-HPLC column (5C18-AR, 8 × 250 mm, 5 μm, Cosmosil) eluted at a flow rate of 3 ml $min^{-1}$, and the HPLC condition was as described. On the other hand, the analytical SAX HPLC experiments were carried out on a SAX column (ZORBAX SAX, 4.6 × 250 mm, 5 μm, Agilent) with a linear gradient of potassium phosphate (60–90 mM in 30 min, pH 3.5), and the detection of NDP-sugars was monitored at 267 nm.

### Time-dependent activity measurements of NokL

The NokL assay was performed in a reaction of 15 μl containing TDP-Rha **11** (1.27 mM, 3 μl, 12 mM Tris-HCl, pH 7.6, 9% glycerol), K-252c **1** (1.27 mM, 3 μL, 50% DMSO), $MgCl_2$ (12 mM, 3 μl in $H_2O$) and NokL crude extract (3 μl, 20 mM $K_2HPO_4$ at pH 7.8, 15% glycerol). The final reaction condition was adjusted to the optimal condition (pH 9.0, 12 mM $MgCl_2$, 50 mM $K_2HPO_4$) described for NokL. The reaction was incubated at 30 °C and quenched at different time points with an equal volume of ice–cold alcohol solution (MeOH–EtOH, 1 : 1 equiv.). The precipitated protein was removed by centrifugation (16 000 g, 4 °C) and the supernatant was analyzed by HPLC. As a control, TDP-Rha **11** was quantitatively recovered in above reaction condition without NokL, as confirmed by SAX HPLC. K-252c is commercially available. TDP-Rha was generated by an *in situ* tandem enzymatic reaction method and was used without further purification. In brief, TDP-D-glucose was incubated with *S. spinosa* Gdh (TDP-glucose 4,6-dehydratase) in 20 mM Tris-HCl (pH 7.8) at 30 °C till completion of the conversion, followed by addition of NAD(P)H, *S. spinosa* Epi (TDP-4-keto-6-deoxyglucose 3,5-epimerase; in 20 mM Tris-HCl/pH 7.3) and *S. spinosa* Kre (TDP-4-ketorhamnose reductase; in 20 mM Tris-HCl/pH 7.3).[42] TDP-Rha **11** was obtained after removal of proteins by heating at 95 °C for 3 min and subsequent centrifugation. For evaluation of other sugar donors and acceptors as possible substrates, the NokL reactions were conducted using the same condition described above, except the final concentrations of both were varied up to 1.5 mM and 33 mM, respectively.

### Inhibition of NokL by TDP or UDP

The inhibition assays of NokL were conducted, similarly as described for the time-dependent experiments of NokL with K-252c **1** and TDP-Rha **11**, by measuring the formation of rhamnosyl-K252c **12** in the presence of TDP or UDP (0.6 mM). The inhibition reactions were quenched and analyzed also as described. The percentage of inhibition of NokL total activity against TDP-Rha was estimated by comparing the formation of **12** in the presence of TDP (or UDP) with that in the absence of TDP (or UDP) at the same time points. TDP (thymidine diphosphate) and UDP (uridine diphosphate) were commercially available.

### Preparation of NokL enzymatic product (rhamnosyl-K252c)

A reaction mixture in a total volume of 5 ml, containing TDP-Rha **11**, K-252c **1** and the NokL crude extract, was carried out as described for the NokL time-dependent experiments. The reaction was incubated at 30 °C for 24 h, and then quenched by 5 ml of the ice-cold alcohol solution (MeOH–EtOH). The resulting mixture was subsequently subjected to centrifugation (16 000 g, 4 °C) for 2 h to remove precipitated proteins. The supernatant was purified by semi-preparative RP-HPLC as described. Fractions containing rhamnosyl-K252c **12** were pooled and evaporated to remove organic solvents, followed by extraction with ether. The ether layer containing **12** was evaporated to remove ether and then vacuumed to gain the compound **12** (2.0 mg). The purity of **12** was greater than 95% as judged by analytical RP-HPLC. The NMR spectrum of **12** dissolved in $D_4$-methanol ($CD_3OD$) was recorded at 500 MHz. For NMR analysis: $^1$H-NMR ($CD_3OD$, 500 MHz) $\delta_H$ 1.80 (3H, d, $J$ = 7.0 Hz), 4.15 (1H, dd, $J$ = 4.0 Hz), 4.33 (1H, td, $J$ = 4.0 Hz), 4.55 (1H, m), 4.70 (1H, d, $J$ = 4.0 Hz) 5.05 (2H, d, $J$ = 4.0 Hz), 6.54 (1H, d, $J$ = 10.0 Hz), 7.26 (1H, t, $J$ = 7.0 Hz), 7.30 (1H, t, $J$ = 8.5 Hz), 7.45 (1H, t, $J$ = 8.0 Hz), 7.48 (1H, t, $J$ = 8.5 Hz), 7.61 (1H, d, $J$ = 8.5 Hz), 7.72 (1H, d, $J$ = 8.5 Hz), 8.02 (1H, d, $J$ = 8.5 Hz), 9.40 (1H, d, $J$ = 8.0 Hz) ppm. $^{13}$C-NMR ($CD_3OD$, 125 MHz) $\delta_C$ 15.8,

47.0, 68.7, 73.3, 73.6, 78.5, 78.6, 110.5, 112.3, 116.5, 118.9, 119.7, 120.7, 121.0, 121.9, 123.6, 124.1, 126.2, 126.5, 126.8, 129.0, 129.7, 134.5, 141.0, 142.2, 175.8 ppm. For high resolution MALDI-TOF spectrometric analysis: $C_{26}H_{23}N_3O_5$ molecular weight calculated as 457.163, and found $m/z$ of 457.177 $[M]^+$.

## Conclusions

In the family of indolocarbazole antitumor antibiotics, K-252a and its derivatives possess unique structural characteristics and great potential as therapeutic agents against various cancers and neurodegenerative disorders. Study of the enzymes responsible for biosynthesis of K-252a, especially the N-glycosylation and the aglycone formation, may greatly facilitate the development of more specific and potent analogs of K-252a for clinical treatments. As a significant breakthrough, our work described here on the functional expression and characterization of the N-glycosyltransferase (NokL) serves as the first example ever demonstrated *in vitro* in biosyntheses of indolocarbazole glycosides. Notably, NokL was found to display striking substrate promiscuity. Along with biochemical characterization and inhibition of the enzyme, the use of various substrate analogs to explore the substrate specificity has provided further insight into the peculiar mode of molecular recognition adapted by NokL. Moreover, the two enzymes, NokA and NokB, in K-252c biosynthesis were thoroughly characterized by the use of an *in vitro* expression system of *E. coli*, unveiling their enzyme functions, substrate specificity and interactions. For the first time, the enzyme product of NokA was trapped as derivatives of indole-3-pyruvic acid (IPA), including indole-3-carboxaldehyde (ICA) and indole-3-acetic acid (IAA). Most notably, by use of ammonium hydroxide, the NokA/NokB coupled reactions were successfully controlled to partition into either the NokA/IPA degradation pathway or the formation of chromopyrrolic acid (CPA), thereby resolving their cross-talking relatioship. In addition, the use of L-tryptophan analogs revealed that NokA exhibited high stereoselectivity in substrate recognition, whereas the enzymes (NokA/NokB) were able to utilize 5-fluoro-L-tryptophan with comparable catalytic efficiency, thereby generating corresponding fluorinated products. The study also provided a simple, useful method to generate K-252d, ICA/IAA and CPA by NokL, NokA and NokA/NokB, respectively, using *E. coli* expression systems. Together, the information gained from this study may thus be useful for future manipulation (*e.g.*, protein engineering) of NokL or NokA/NokB to make, in a combinatorial fashion, unnatural analogs of K-252a or K-252d distinct from others of the family in both structure and biological function.

## Acknowledgements

This research was supported by National Science Council (Taiwan) grants (NSC-93-2113-M-009-002 and NSC-94-2113-M-009-007) to Hsien-Tai Chiu. We thank HSP Research Institute (Kyoto Research Park, Japan) for kindly providing pG-KJE7 as a gift.

## References

1 B. H. Long, W. C. Rose, D. M. Vyas, J. A. Matson and S. Forenza, *Curr. Med. Chem.: Anti-Cancer Agents*, 2002, **2**, 255–266.
2 A. Stolz, C. Vogel, V. Schneider, N. Ertych, A. Kienitz, H. Yu and H. Bastians, *Cancer Res.*, 2009, **69**, 3874–3883.
3 H. Nakano and S. Omura, *J. Antibiot.*, 2009, **62**, 17–26.
4 C. Sánchez, C. Méndez and J. A. Salas, *Nat. Prod. Rep.*, 2006, **23**, 1007–1045.
5 Y. Goodman and M. P. Mattson, *Brain Res.*, 1994, **650**, 170–174.
6 H. A. Mucke, *IDrugs*, 2003, **6**, 377–383.
7 B. L. Staker, M. D. Feese, M. Cushman, Y. Pommier, D. Zembower, L. Stewart and A. B. Burgin, *J. Med. Chem.*, 2005, **48**, 2336–2345.
8 A. Gescher, *Gen. Pharmacol.*, 1998, **31**, 721–728.
9 C. Carrasco, M. Facompré, J. D. Chisholm, D. L. Van Vranken, W. D. Wilson and C. Bailly, *Nucleic Acids Res.*, 2002, **30**, 1774–1781.
10 A. Tomillero and M. A. Moral, *Methods Find. Exp. Clin. Pharmacol.*, 2008, **30**, 231–251.
11 H. Kase, K. Iwahashi and Y. Matsuda, *J. Antibiot.*, 1986, **39**, 1059–1065.
12 T. Yasuzawa, T. Iida, M. Yoshida, N. Hirayama, M. Takahashi, K. Shirahata and H. Sano, *J. Antibiot.*, 1986, **39**, 1072–1078.
13 A. Morotti, S. Mila, P. Accornero, E. Tagliabue and C. Ponzetto, *Oncogene*, 2002, **21**, 4885–4893.
14 P. P. Roux, G. Dorval, M. Boudreau, A. Angers-Loustau, S. J. Morris, J. Makkerh and P. A. Barker, *J. Biol. Chem.*, 2002, **277**, 49473–49480.
15 C. Collins, M. A. Carducci, M. A. Eisenberger, J. T. Isaacs, A. W. Partin, R. Pili, V. J. Sinibaldi, J. S. Walczak and S. R. Denmeade, *Cancer Biol. Ther.*, 2007, **6**, 1360–1367.
16 S. D. Undevia, N. J. Vogelzang, A. M. Mauer, L. Janisch, S. Mani and M. J. Ratain, *Invest. New Drugs*, 2004, **22**, 449–458.
17 A. M. Camoratto, J. P. Jani, T. S. Angeles, A. C. Maroney, C. Y. Sanders, C. Murakata, N. T. Neff, J. L. Vaught, J. T. Isaacs and C. A. Dionne, *Int. J. Cancer*, 1997, **72**, 673–679.
18 M. S. Butler, *Nat. Prod. Rep.*, 2005, **22**, 162–195.
19 D. E. Cane, C. T. Walsh and C. Khosla, *Science*, 1998, **282**, 63–68.
20 H.-T. Chiu, Y.-L. Chen, C.-Y. Chen, C. Jin, M.-N. Lee and Y.-C. Lin, *Mol. BioSyst.*, 2009, **5**, DOI: 10.1039/b905293c, (the companion paper).
21 H. Onaka, S. Taniguchi, Y. Igarashi and T. Furumai, *J. Antibiot.*, 2002, **55**, 1063–1071.
22 C. Sánchez, I. A. Butovich, A. F. Braña, J. Rohr, C. Méndez and J. A. Salas, *Chem. Biol.*, 2002, **9**, 519–531.
23 S. Nakanishi, Y. Matsuda, K. Iwahashi and H. Kase, *J. Antibiot.*, 1986, **39**, 1066–1071.
24 C. Zhang, C. Albermann, X. Fu, N. R. Peters, J. D. Chisholm, G. Zhang, E. J. Gilbert, P. G. Wang, D. L. Van Vranken and J. S. Thorson, *ChemBioChem*, 2006, **7**, 795–804.
25 T. Nishizawa, S. Grschow, D.-H. E. Jayamaha, C. Nishizawa-Harada and D. H. Sherman, *J. Am. Chem. Soc.*, 2006, **128**, 724–725.
26 K. Nishihara, M. Kanemori, M. Kitagawa, H. Yanagi and T. Yura, *Appl. Environ. Microbiol.*, 1998, **64**, 1694–1699.
27 M. A. Bittinger, L. P. Nguyen and C. A. Bradfield, *Mol. Pharmacol.*, 2003, **64**, 550–556.
28 J. Koga, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 1995, **1249**, 1–13.
29 A. W. Woodward and B. Bartel, *Ann. Bot.*, 2005, **95**, 707–735.
30 Y. Furuya, H. Sawada, T. Hirahara, K. Ito, T. Ohshiro and Y. Izumi, *Biosci., Biotechnol., Biochem.*, 2000, **64**, 1486–1493.
31 Q.-T. Phi, Y. M. Park, C. M. Ryu, S. H. Park and S. Y. Ghim, *J. Microbiol. Biotechnol.*, 2008, **18**, 1235–1244.
32 V. Magnus, Š. Šimaga, S. Iskric and S. Kveder, *Plant Physiol.*, 1982, **69**, 853–858.
33 S. W. Yang and G. A. Cordell, *J. Nat. Prod.*, 1997, **60**, 44–48.
34 L. R. M. B. Mourão, R. S. S. Santana, L. M. Paulo, S. M. P. Pugine, L. M. Chaible, H. Fukumasu, M. L. Z. Dagli and M. P. de Melo, *Cell Biochem. Funct.*, 2009, **27**, 16–22.
35 E. Nakajima, H. Nakano, K. Yamada, H. Shigemori and K. Hasegawa, *Phytochemistry*, 2002, **61**, 863–865.
36 M.-T. Gutierrez-Lugo, G. M. Woldemichael, M. P. Singh, P. A. Suarez, W. M. Maiese, G. Montenegro and B. N. Timmermann, *Nat. Prod. Res.*, 2005, **19**, 645–652.

37 A. R. Howard-Jones and C. T. Walsh, *Biochemistry*, 2005, **44**, 15652–15663.
38 C. J. Balibar and C. T. Walsh, *Biochemistry*, 2006, **45**, 15444–15457.
39 J.-P. Bégué and D. Bonnet-Delpon, *J. Fluorine Chem.*, 2006, **127**, 992–1012.
40 H. Fujisawa, T. Fujiwara, Y. Takeuchi and K. Omata, *Chem. Pharm. Bull.*, 2005, **53**, 524–528.
41 H. Shao, X. He, L. Achnine, J. W. Blount, R. A. Dixon and X. Wang, *Plant Cell*, 2005, **17**, 3141–3154.
42 Y.-L. Chen, Y.-H. Chen, Y.-C. Lin, K.-C. Tsai and H.-T. Chiu, *J. Biol. Chem.*, 2009, **284**, 7352–7363.
43 A. M. Mulichak, H. C. Losey, W. Lu, Z. Wawrzak, C. T. Walsh and R. M. Garavito, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9238–9243.
44 C. Sánchez, C. Méndez and J. A. Salas, *J. Ind. Microbiol. Biotechnol.*, 2006, **33**, 560–568.
45 E. Rodrigues Pereira, L. Belin, M. Sancelme, M. Prudhomme, M. Ollier, M. Rapp, D. Sevère, J.-F. Riou, D. Fabbro and T. Meyer, *J. Med. Chem.*, 1996, **39**, 4471–4477.
46 C. A. Carter, *Curr. Drug Targets*, 2000, **1**, 163–183.
47 S. Akinaga, K. Sugiyama and T. Akiyama, *Anti-Cancer Drug Des.*, 2000, **15**, 43–52.
48 D. Fabbro, S. Ruetz, S. Bodis, M. Pruschy, K. Csermak, A. Man, P. Campochiaro, J. Wood, T. O'Reilly and T. Meyer, *Anti-Cancer Drug Des.*, 2000, **15**, 17–28.
49 R. L. Hudkins, M. Iqbal, C. H. Park, J. Goldstein, J. L. Herman, E. Shek, C. Murakata and J. P. Mallamo, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 1873–1876.
50 G. Stevenson, B. Neal, D. Liu, M. Hobbs, N. H. Packer, M. Batley, J. W. Redmond, L. Lindquist and P. Reeves, *J. Bacteriol.*, 1994, **176**, 4144–4156.
51 J. M. D'Souza, L. Wang and P. Reeves, *Gene*, 2002, **297**, 123–127.
52 G. Samuel and P. Reeves, *Carbohydr. Res.*, 2003, **338**, 2503–2519.
53 C. L. Marolda and M. A. Valvano, *J. Bacteriol.*, 1995, **177**, 5539–5546.
54 J. Sambrook and D. Russell, *Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor Laboratory Press, New York, 2001.