Preparation of new halogenated diphenyl pyrazine analogs in Escherichia coli by a mono-module fungal nonribosomal peptide synthetase from Penicillium herquei

Can Yang, Yuanyuan Xu, Kangping Xu, Guishan Tan, Xia Yu

PII: S0040-4039(18)30844-X
DOI: https://doi.org/10.1016/j.tetlet.2018.06.065
Reference: TETL 50103

To appear in: *Tetrahedron Letters*

Received Date: 10 May 2018
Revised Date: 20 June 2018
Accepted Date: 28 June 2018

Please cite this article as: Yang, C., Xu, Y., Xu, K., Tan, G., Yu, X., Preparation of new halogenated diphenyl pyrazine analogs in Escherichia coli by a mono-module fungal nonribosomal peptide synthetase from Penicillium herquei, *Tetrahedron Letters* (2018), doi: https://doi.org/10.1016/j.tetlet.2018.06.065

Preparation of new halogenated diphenyl pyrazine analogs in *Escherichia coli* by a mono-module fungal nonribosomal peptide synthetase from *Penicillium herquei*

Can Yang [a], Yuanyuan Xu [a], Kangping Xu [a], Guishan Tan [a,b], Xia Yu [a]*

[a] *Xiangya School of Pharmaceutical Sciences, Central South University, Changsha, Hunan, P.R. China*

[b]*Xiangya Hospital of Central South University, Changsha, P.R. China*

* Corresponding author to: X. Yu, Xiangya School of Pharmaceutical Sciences, Central South University, Changsha, Hunan, P.R. China

E-mail address: xyu226@csu.edu.cn (X. Yu).

Abstract:

Pyrazines are important structures widely found in many known drugs. The biological approaches for their synthesis were poorly applied. Herein, microbial production of several halogenated diphenyl pyrazines is reported. These compounds are accumulated via feeding corresponding precursor analogs to *Escherichia coli* expressing a fungal non-ribosomal peptide synthetase HqlA. Substrate specificity of HqlA was also determined by comparing substrate incorporation efficiencies. HqlA requires a C4-hydroxyl in the substrate and can tolerate certain degrees of size change on the substitution at the carbon next to the hydroxyl group.

Keywords:

diphenyl pyrazine, biosynthesis, NRPS, tyrosine recognition, enzyme promiscuity

Highlights

The biocatalytic approaches to obtain diphenyl pyrazine analogs

Exploration of the catalytic ability of HqlA towards tyrosine derivatives

The potential usage of HqlA as biocatalyst to synthase substituted pyrazines

Introduction

Pyrazines are an important class of heterocyclic compounds containing the 1,4-diazine core. The pyrazine derivatives not only exhibit characteristic reactivities, such as acting as bridging ligand to combine two molecules, but also possess ideal pharmacological activities.[1, 2] Many of them have been developed into clinical drugs, for example, the proteasome inhibitor Bortezomib, anti-diabetic agent Glipizide, anti-tubercular agents Morinamide and Pyrazinamide (Figure 1).[3-6] Moreover, pyrazine derivatives are used as flavoring additives in food.[7] The synthesis of pyrazines has been studied for years. Although a number of chemical routes for substituted pyrazines have been designed, the biological approaches were poorly applied.

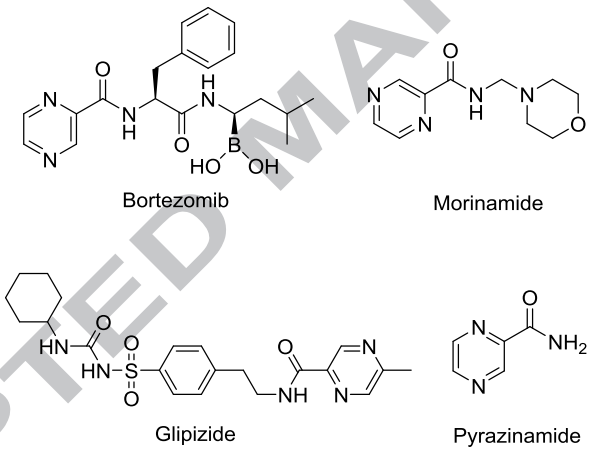

Figure 1. Examples of pyrazine containing drugs.

Biosynthesis of 2,5-substituted pyrazines had been proposed to use amino acids as building blocks and corresponding diketopiperazines as intermediates. However, research from Wittmann *et al*. on biosynthesis of some highly substituted pyrazines in *Corynebacterium glutamicum* supported that those pyrazines are synthesized from intermolecular condensation of alpha-amino-ketones.[8] Feeding experiments by Schulz and co-workers confirmed amino acid as precursor and excluded involvement of diketopiperazine as intermediate, they also proposed that pyrazines are probably dimerization products from the corresponding α-minocarbonyl intermediates.[9]

A few enzymatic/genetic studies had been reported on synthesis of pyrazines. Tang *et*

al. produced a series of thiopyrazines by using a truncated nonribosomal peptide synthetase (NRPS325) module from a PKS-NRPS hybrid protein ATEG00325,[10] which is involved in the biosynthesis of isoflavipucine and dihydroisoflavipucine from *Aspergillus terreus*.[11] As another example, Keller *et al.* correlated the formation of diphenylpyrazines with the redundant clusters including the NRPS genes *lnaA* and *lnbA* in *Aspergillus flavus*.[12] In the previous study for biosynthesis of piperazine alkaloid Herquline A in *Penicillium herquei*, we found that a NRPS enzyme HqlA could efficiently produce the diphenyl pyrazine **2** *in vitro*.[13]

HqlA is a single module NRPS-like protein consisting of an adenylation domain (A), a thiolation domain (T) and a reductase domain (R) (Figure 2).[13] Unlike a typical NRPS, HqlA lacks the key condensation domain (C) that is responsible for the peptide bond formation. However, HqlA still follows the same thioester-tethered mechanism as in NRPSs mediated biosynthesis, where each module contributes a single amino acid to the assembling intermediate tethered on the carrier proteins.[14-16] Specifically, the A domain in HqlA recognizes and activates L-tyrosine to from the tyrosyl-T thioester intermediate, then the R domain reduces the intermediate into the corresponding amino aldehyde. Dimerization between two aldehydes yields the diphenyl pyrazine **2** (Figure 2). In recent years, efforts were performed for the manipulations of recombinant NRPSs to produce novel peptides.[17, 18] For HqlA, elucidation of the substrate promiscuity of this protein is the first step to manipulate it for the biocombinatorial synthesis of substituted pyrazines.

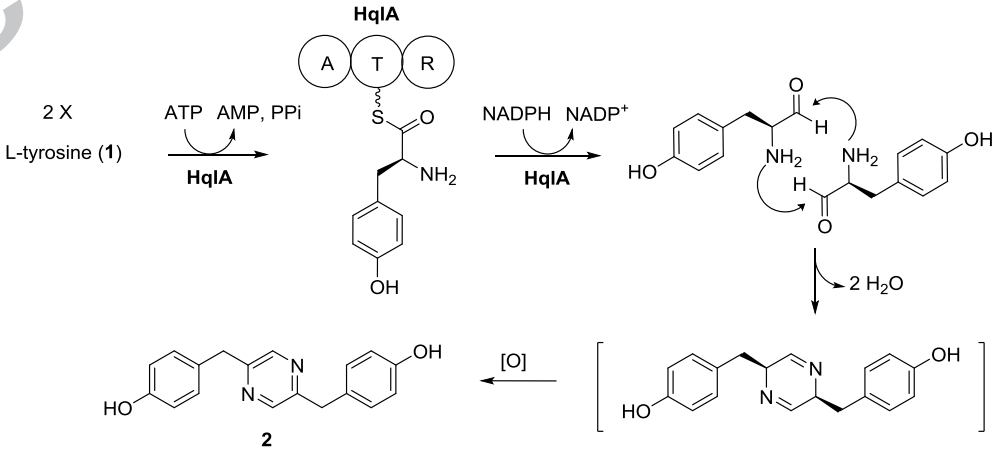

Figure 2. Reaction of HqlA by using the natural substrate L-tyrosine.

Results

We overexpressed the recombinant *P. herquei* HqlA in its *holo* form with a C-terminal His_6-tag in *Escherichia coli* BAP1 cells.[19] Since most NRPSs show low enzymatic efficiencies on *in vitro* reconstituted assays and only trace amounts of products can be detected, we chose a whole-cell biotransformation approach to characterize the substrate promiscuity of HqlA. The *in vivo* biotransformation of *Escherichia coli* BAP1 cells harboring *hqlA* was performed with different substrates at 16°C for 2 days (Figure 3). Biotransformation by *E. coli* BAP1 cells with the empty vector pCDFDuet-1 were used as negative controls.

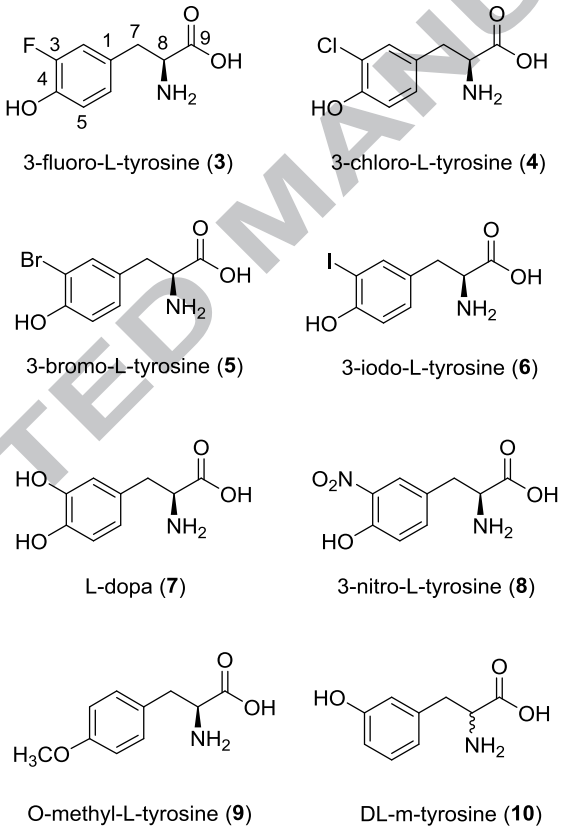

Figure 3. Tyrosine derivatives investigated in this study.

The first group of substrates was O-methyl-L-tyrosine (**9**) and DL-m-tyrosine (**10**). Since L-tyrosine (**1**) is a proteinogenic amino acid synthesized in *E. coli*, the corresponding product **2** was also detected in the biotransformation. However, when the hydroxyl group at C4 was replaced with a methoxy group or shifted to C3, no

other product was detected, indicating that there is probably a key hydrogen binding between the C4-OH and the enzyme, therefore the presence of C4-OH is necessary for the substrate recognition in HqlA.

To further probe the substrate specificity of HqlA, we tested six tyrosine derivatives with C4-OH and different C3 substitution groups, including 3-fluoro-L-tyrosine (**3**), 3-chloro-L-tyrosine (**4**), 3-bromo-L-tyrosine (**5**), 3-iodo-L-tyrosine (**6**), L-dopa (**7**) and 3-nitro-L-tyrosine (**8**). The results indicated that the substrate acceptance in HqlA is related to the size of the substitution groups at C3. HqlA utilized tyrosine derivatives with smaller groups (F-, Cl- or Br-, *i.e.* compounds **3**-**5**) on C3, but not those with larger substitution groups (I- and NO_2-, *i.e.* compounds **6** and **8**). Detailed inspection of the HPLC chromatograms for biotransformation of **3**-**5** revealed that two new products were detected in the extracts (Figure 4). Reduced total conversion yield was found for **3**, **4** and **5** in corresponding to the larger C3- substitution group. Specific product yield from **3**, **4** and **5** by HqlA was 3.7, 0.9 and 0.4, respectively, which are represented by the ratio between the amount of the new compounds to the amount of **2** in the background (Figure 4). These results suggest the inverse relation of the sizes of the C3 substitution groups to the HqlA activities. In the case of L-dopa (**7**), a C3-OH substituted tyrosine, although the group OH- is smaller than Br-, no product was detected. This is probably due to the nonactive 3-OH mentioned above interfere the hydrogen binding formation between the active 4-OH and HqlA.

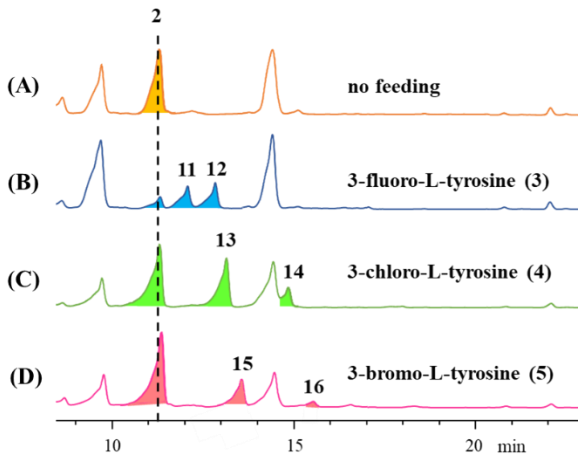

Figure 4. Biotransformation of **3**-**5** by *E. coli* BAP1 cells harboring *hqlA*. Detection was carried out on the HPLC with a Photodiode array detector and illustrated for absorption at 254 nm.

For structure elucidation, the products **11** and **12**, **13** and **14**, **15** and **16** were isolated from the biotransformation of **3**, **4**, **5** by HqlA, respectively. All six compounds were subjected to MS and NMR analyses (Figures S1-S13 and Tables S1-S4). The appearance of typical pyrazine protons (δ_H 8.44-8.49, s) revealed clearly the formation of a pyrazine ring in **11-16**.[13] A singlet signal with corresponding to two protons was observed at δ_H 3.94-3.97 (H-7), suggesting a standalone CH_2 connecting the phenyl and pyrazine rings. Combinational analysis of the HR-ESI-MS and ^{1}H-NMR spectra of **12**, **14** and **16** revealed symmetric structures in them (Figures S2, S4, S6, S9, S11, S13 and Tables S1, S4). The signals in the ^{1}H-NMR spectrum appeared as two identical sets of protons, each containing three coupling aromatic protons for the C3-substituted phenyl ring and two typical singlets related to the pyrazine ring. Therefore, **12**, **14** and **16** were identified as 3,3′-fluoro-diphenyl pyrazine, 3,3′-chloro-diphenyl pyrazine, 3,3′-bromo-diphenyl pyrazine, respectively (Figure 5). The ^{1}H-NMR spectrum of **11** showing the same signals as a sum of those observed for **12** and **2** (Figure S14), then **11** is identified as a pyrazine produced from one molecule of the substrate 3-fluoro-L-tyrosine (**3**) and one molecule of the proteinogenic amino acid L-tyrosine, *i.e.* 3-fluoro-diphenyl pyrazine (Figure 5). The structure of **11** was furtherly confirmed by HR-ESI-MS analysis. Similar phenomena were also found for **13** and **15**, which were also identified as pyrazine produced from one molecule of substrate and one molecule of L-tyrosine, *i.e.* 3-chloro-diphenyl pyrazine and 3-bromo-diphenyl pyrazine (Figure 5). **14** was reported in a literature as product from the bimodular NRPS AusA, however, no NMR data was revealed there.[20] We give the first reported NMR data of **14**. The structures of **11**-**13**, **15** and **16** have not been described previously.

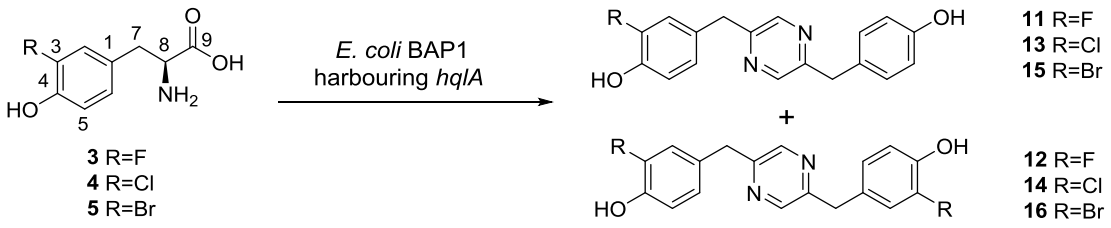

Figure 5. Reactions of **3**-**5** catalyzed by *E. coli* BAP1 cells harboring *hqlA*.

Discussion

HqlA shares about 59% sequence identity at amino-acid level to LnaA/LnbA.[13] All of them activate L-tyrosine and reduce it into the corresponding alpha-amino aldehyde followed by intermolecular condensation of two aldehydes. Both LnaA and LnbA can adenylate L-tyrosine (**1**) specifically, while LnaA can somewhat also adenylate D-tyrosine and L-histidine.[12] A slightly different aerobic oxidation process was as proposed for the formation of thiopyrazines by the truncated NRPS protein NRPS325, involves only one reduction step and an additional nucleophilic attack on the thioester intermediate by the thiol substrate.[10] The domain architecture of HqlA (A-T-R) resembles that of the well-known carboxylic acid reductases (CARs), which catalyze the reduction of carboxylic acids to their corresponding aldehydes.[21] Similar domain organization could also be found within the ending module of some NRPSs or PKS-NRPS hybrids, where the peptidyl chain is extended by one amino acid moiety and released by the R domain as an aldehyde, an alcohol or a tetramic acid (when the R domain is a Dieckmann-condensation domain).[22]

Many nonribosomal peptide synthetases show high degree of substrate specificity with only one known amino acid substrate *in vivo*.[23, 24] HqlA selectively activates and incorporates L-tyrosine into the biosynthetic product, among the naturally occurring amino acids. However, our above-mentioned experiments show that some 3-substutited derivatives of tyrosine can be efficiently utilized by HqlA *in vivo*. HqlA is a multi-domain enzyme which involves substrate recognition twice during each catalytic cycle, the adenylation on A domain and the reduction on R domain. Observation of a pyrazine product from this enzyme will require acceptance of the substrate on both the A and R domains. The A domain recognizes and activates the substrate amino acid, via consuming an ATP to form an aminoacyl-AMP intermediate, and subsequently loads the activated amino acids to the T domain to form an aminoacyl-*S*-T thioester. The R domain recognizes the aminoacyl-S-T and reduces it into the corresponding aldehyde, with the cost of an NADPH cofactor. The gate-keeping function on A domain should be stricter than that on the R domain, since

it will avoid unnecessary hijacking on the T domain, which will result in non-reacting enzyme.

Prediction of adenylation domain substrate specificity based on sequence analysis for fungal NRPSs is not as successful as for bacterial counterparts. Although the 10-AA code,[23] which is based on the structure of a bacteria phenylalanine activating domain PheA, doesn't conclusively explain substrate specificities of HqlA, HqlA still have the conserved aspartic acid (D235 in PheA) and lysine (K517 in PheA) residues which are proposed to interact with the carboxylic acid and the alpha-amino group.[23] In contrast, the CARs usually don't show a conserved D235 in their sequences.[25] The failure to accept O-methyl-L-tyrosine and m-tyrosine for HqlA indicates that this enzyme has a strong recognizing factor towards the para-hydroxyl group in tyrosine, likely being through a hydrogen bond forming mechanism on A domain. The reducing acceptance in accordance with the size on 3-substitution is probably due to steric effects between substrate and the enzyme.

Conclusion

We demonstrated the acceptance of several tyrosine derivatives by a mono-module NRPS HqlA from *P. herquei*, and therefore the biocatalytic formation of six pyrazines. This work not only gives key information on how tyrosine-activating enzyme recognizes substrate, but also shows potential on synthesizing pyrazines using engineered microorganism.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 21708049) and Natural Science Foundation of Hunan Province (No. 2018JJ3630). We thank Dr. Guogen Liu and Dr. Hongping Long for taking NMR and HR-ESI-MS spectra.

References

1. Dolezal M, Zitko J. *Expert Opin Ther Pat* **2015**; 25: 33-47.

2. Miniyar PB, Murumkar PR, Patil PS, Barmade MA, Bothara KG. *Mini Rev Med Chem* **2013**; 13: 1607-1625.

3. Adams J, Kauffman M. *Cancer Investigation* **2004**; 22: 304-311.

4. Prendergast BD. *Clin Pharm* **1984**; 3: 473-485.

5. Zimhony O, Cox JS, Welch JT, Vilcheze C, Jacobs WR, Jr. *Nat Med* **2000**; 6: 1043-1047.

6. Bonanni G, Ciccariello M, Mancini P, Pace V, Sagliaschi G. *Riv Eur Sci Med Farmacol* **1993**; 15: 171-174.

7. Adams TB, Doull J, Feron VJ, Goodman JI, Marnett LJ, Munro IC, Newberne PM, Portoghese PS, Smith RL, Waddell WJ, Wagner BM. *Food Chem Toxicol* **2002**; 40: 429-451.

8. Dickschat JS, Wickel S, Bolten CJ, Nawrath T, Schulz S, Wittmann C. *Eur J Org Chem* **2010**: 2687-2695.

9. Nawrath T, Dickschat JS, Kunze B, Schulz S. *Chem Biodivers* **2010**; 7: 2129-2144.

10. Qiao K, Zhou H, Xu W, Zhang W, Garg N, Tang Y. *Org Lett* **2011**; 13: 1758-1761.

11. Gressler M, Zaehle C, Scherlach K, Hertweck C, Brock M. *Chemistry & Biology* **2011**; 18: 198-209.

12. Forseth RR, Amaike S, Schwenk D, Affeldt KJ, Hoffmeister D, Schroeder FC, Keller NP. *Angew Chem Int Ed Engl* **2013**; 52: 1590-1594.

13. Yu X, Liu F, Zou Y, Tang M-C, Hang L, Houk KN, Tang Y. *J Am Chem Soc* **2016**; 138: 13529-13532.

14. Hur GH, Vickery CR, Burkart MD. *Nat Prod Rep* **2012**; 29: 1074-1098.

15. Sussmuth RD, Mainz A. *Angew Chem Int Ed Engl* **2017**; 56: 3770-3821.

16. Marahiel MA, Essen LO. *Methods Enzymol* **2009**; 458: 337-351.

17. Niquille DL, Hansen DA, Mori T, Fercher D, Kries H, Hilvert D. *Nat Chem* **2018**; 10: 282-287.

18. Cai X, Nowak S, Wesche F, Bischoff I, Kaiser M, Furst R, Bode HB. *Nat Chem* **2017**; 9: 379-386.

19. Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, Khosla C. *Science* **2001**; 291: 1790-1792.

20. Wyatt MA, Mok MC, Junop M, Magarvey NA. *Chembiochem* **2012**; 13: 2408-2415.

21. Winkler M. *Curr Opin Chem Biol* **2018**; 43: 23-29.

22. Du L, Lou L. *Natural Product Reports* **2010**; 27: 255-278.

23. Stachelhaus T, Mootz HD, Marahiel MA. *Chem Biol* **1999**; 6: 493-505.

24. Strieker M, Tanovic A, Marahiel MA. *Curr Opin Struct Biol* **2010**; 20: 234-240.

25. Wang M, Zhao H. *ACS Catalysis* **2014**; 4: 1219-1225.

Graphical abstract