Bioorganic & Medicinal Chemistry Letters 28 (2018) 1071–1076

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Promiscuous (+)-γ-lactamase activity of an amidase from nitrile hydratase pathway for efficient synthesis of carbocyclic nucleosides intermediate

Hongxia Li, Shaozhou Zhu *, Guojun Zheng *

State Key Laboratory of Chemical Resources Engineering, Beijing University of Chemical Technology, Beijing 100029, People's Republic of China

ARTICLE INFO

Article history:
Received 21 November 2017
Revised 4 February 2018
Accepted 9 February 2018
Available online 10 February 2018

Keywords:
Carbocyclic nucleosides
2-azabicyclo[2.2.1]hept-5-en-3-one
(+)-γ-Lactamase
Nitrile hydratase
Enzymatic resolution

ABSTRACT

Based on bioinformatics analysis, the promiscuous (+)-γ-lactamase activity of an amidase was identified in *Rhodococcus erythropolis* PR4 and found to be involved in the nitrile hydratase pathway. The amidase is highly enantioselective and can be used in the kinetic resolution of the Vince lactam. The known structure provides a rare insight into the catalytic mechanism of (+)-γ-lactamase with absolute chiral selectivity. This lactamase was cloned, purified, biochemically characterized, and demonstrated to be an ideal catalyst for the preparation of carbocyclic nucleosides of pharmaceutical interest. The chiral selectivity of this enzyme was investigated by molecular docking and site-specific mutagenesis, which provides a foundation for further engineering of these versatile biocatalysts.

Carbocyclic nucleosides are a group of nucleoside analogs with a diverse set of pharmacological activities, including antimicrobial, antiviral, and antitumor activities.[1,2] Since the isolation of aristeromycin, a unique carbocyclic nucleoside antibiotic produced by *Streptomyces citricolor*, these compounds have attracted considerable attention and some drugs are currently available on the pharmaceutical market as effective treatments for several viral infections.[3] For example, entecavir (Baraclude) is an antiviral medication used to treat hepatitis B virus (HBV) infection. Abacavir (Ziagen) is widely used as a treatment for human immunodeficiency virus infection. Alogliptin (Nesina) was developed as an orally administered anti-diabetic drug in the DPP-4 inhibitor class.[4]

One common feature of these carbocyclic nucleosides is the methylene unit which replaces oxygen in the sugar portion of the nucleoside. This structural motif confers extraordinary stability against nucleoside phosphorylases.[5] Moreover, its conformational similarity with tetrahydrofuran endows these analogs very interesting biological activities. While the replacement of furanosyl oxygen by a methylene brings several advantages, it also makes the synthesis of these compounds very challenging. Thus, a general synthon to access these drugs is necessary. Over the past decades, several synthetic routes for carbocyclic nucleoside analogs have been developed, such as oxirane ring cleavage, ring-closing metathesis (RCM) reaction from D-ribose, transition metal mediated reactions were tried.[5,6] Among all these progresses, one breakthrough was made by Robert Vince who developed 2-azabicyclo [2.2.1]hept-5-en-3-one (Vince lactam) as a key synthon for the preparation of enantiopure carbocyclic nucleosides.[5] Generally, hydrolysis of the amide bond of Vince lactam can provide a cyclopentane ring template and the product can be further modified for the synthesis of various carbocyclic nucleosides.

To introduce stereospecificity, an enzyme of microbial origin related to Vince lactam resolution was intensively screened, leading to the development of (+)-γ-lactamase.[7] γ-Lactamases are named based on their abilities to catalyze the enantioselective hydrolysis of Vince lactam. This synthetic route is simple, feasible, highly efficient, inexpensive, and ecologically friendly. Because of these advantages, γ-lactamase-mediated strategies are standard procedures for preparing enantiomerically pure carbocyclic nucleosides.[8] With the development of novel screening and genome mining methods, several different classes of γ-lactamases have been identified and the number is increasing.[8–13] Despite these advances, several important issues have not been addressed. Most importantly, the detailed structures of most (+)-γ-lactamases are unknown. Therefore, the catalytic mechanisms of these versatile enzymes remain largely speculative, limiting further engineering efforts. Second, because γ-lactamases are named according to their industrial activities, their natural functions remain unclear.

* Corresponding author.
E-mail addresses: zhusz@mail.buct.edu.cn (S. Zhu), zhenggj@mail.buct.edu.cn (G. Zheng).

https://doi.org/10.1016/j.bmcl.2018.02.019
0960-894X/

In this study, we evaluated the promiscuous and enantioselective (+)-γ-lactamase activity on the Vince lactam of a signature amidase from *R. erythropolis* PR4. The amidase was genetically coupled with nitrile hydratase and involved in the nitrile degradation pathway, indicating the function of *in vivo* function of (+)-γ-lactamase.[14] The determined 3-dimensional structure also provided an ideal model for investigating the catalytic mechanism of (+)-γ-lactamase, which will facilitate further engineering efforts. Moreover, our data showed that this thermostable (+)-γ-lactamase has high-resolution activity towards Vince lactam with excellent enantioselectivity, which provides another efficient and environmentally friendly biocatalyst for the green manufacturing of carbocyclic nucleoside medicines.

Currently, the *in vivo* function of (+)-γ-lactamase is unknown. We investigated the natural function of type I (+)-γ-lactamases, which are classified into the signature amidase superfamily. Only two types of I (+)-γ-lactamases have been reported. The first is from *Sulfolobus solfataricus* whose activity was identified in 2004 to be promiscuous.[10] The second is from *Bradyrhizobium japonicum* USDA 6 which was identified by a genome mining strategy.[11] Initially, bioinformatic analysis was performed to probe the potential pathway in which they are involved. For the amidase from *S. solfataricus*, bioinformatic studies showed that it is surrounded by two transposases and one peroxiredoxin, indicating that the amidase was acquired via horizontal gene transfer. In contrast, the (+)-γ-lactamase from *B. japonicum* USDA 6 was surrounded by one potential nitrile hydratase regulator located directly upstream and five potential ABC transporters related to amino acid transport. These genes strongly indicate that the amidase is involved in a nitrile degradation pathway.[14] We then used these two (+)-γ-lactamases as queries to perform BLAST searches in the protein data bank dadabase. This survey might provide two important information: 1) the potential pathway in which the (+)-γ-lactamases are involved, 2) the structural information for (+)-γ-lactamases. Indeed, our bioinformatic survey led us to the amidases from *Rhodococcus*. Interestingly, the enzyme from *Rhodococcus sp*. N-771 whose crystal structure was solved in 2010 attracted our attention (44% identities with the (+)-γ-lactamase from *S. solfataricus* and 46% identities with the (+)-γ-lactamase from *B. japonicum* USDA 6).[15] The amidase from *Rhodococcus sp*. N-771 is the same as that from *R. erythropolis* PR4 (100% identities). These amidases were shown to be involved in the nitrile degradation pathway (Fig. 1a).[15,16] Unlike the gene cluster from *B. japonicum* USDA 6, the enzyme contains an aldoxime dehydratase upstream, a nitrile hydratase regulator, an amidase, and two nitrile hydratase subunits downstream that can degrade aldoxime to the corresponding organic acid (Fig. 1b). The amidases may hydrolyze the amide to the corresponding acid and might show promiscuous (+)-γ-lactamase activity similarly to those from *B. japonicum* USDA 6 and *S. solfataricus* (Fig. 1c).

To demonstrate that the amidase from the nitrile degradation pathway is indeed a (+)-γ-lactamase, amidase from *R. erythropolis* PR4 (ReGL) was cloned into the pET41a vector and overexpressed in *Escherichia coli* BL21(DE3). A His_8 tag was introduced at the C terminus of the recombinant protein. The enzyme was successfully overexpressed in the soluble form. The recombinant protein (predicted mass: 56 KDa) was purified firstly by Ni-NTA affinity purification, followed by a second size exclusion chromatography step to yield a very pure product (Fig. 2a). The purified enzyme was then assayed using Vince lactam as a substrate. The results revealed high activity towards (+)-γ-lactam and specific enantioselectivity (ee > 99%), demonstrating that it is indeed a (+)-γ-lactamase (Fig. 2b, Fig. S1-S3).

We then characterized the intrinsic properties of purified ReGL using Vince lactam as a substrate (details in Supplementary Material). We first investigated the optimum reaction temperature for ReGL. The reactions were performed at various temperature in the universal buffer (pH 7.0) for 5 min. Results showed that ReGL was highly active from 40 to 70 °C, and maximum activity was observed at 60 °C. Thermostability analysis revealed that ReGL could maintain 100% of its original activity when incubated at 50 °C for 1 h (Fig. 3a). Compared with other thermolabile (+)-γ-lactamases such as the one from *Microbacterium hydrocarbonoxydans* and extremely thermostable (+)-γ-lactamases such as the one from *S. solfataricus* which requires high temperature to perform the reaction, this moderate thermostable (+)-γ-lactamase might be more suitable for industrial application.[9,12] We then investigated the optimum reaction pH for ReGL. The optimal pH of ReGL was evaluated at 60 °C in the universal buffer with different pH ranging from 3.0 to 12.0. Purified ReGL could hydrolyze Vince lactam between pH 6.0 and 10.0, with pH 9.0 as the optimal pH for its activity. Moreover, pH stability studies showed that ReGL was stable from pH 6.0 to 9.0 (Fig. 3b). The effects of metal ions were also studied. Similar activity in the presence of most metal ions compared with the control indicates that this amidase is metal-independent (Fig. S4).

The kinetic parameters of ReGL were shown to be *k*cat = 78 s^{-1} and *km* = 53.9 mM, suggesting that the lactamase has a low binding affinity to Vince lactam. However, the turnover rate is relatively high. The km value of Vince lactam (Table. 1) was relatively high compared to that of four amide compounds (acetamide, propionamide, acrylamide and benzamide) [15]. The *k*cat for Vince lactam was higher than that of acetamide and acrylamide and lower than that of propionamide and benzamide, indicating that Vince lactam is a moderate preferable substrate for ReGL. The *k*cat/*km* value was 1.45 $mM^{-1}s^{-1}$. Thus, benzamide is the most preferable substrate, and Vince lactam could only be moderately hydrolyzed.

The catalytic mechanism of (+)-γ-lactamases remains unclear. To date, >20 γ-lactamases have been reported. However, only one crystal structure of (−)-γ-lactamase has been solved.[17] The (−)-γ-lactamase identified from an *Aureobacterium* strain has been characterized and a detailed mechanism for (−)-γ-lactam hydrolysis was proposed. However, few crystal structures of (+)-γ-lactamase have been solved. Although the crystal structure of a (+)-γ-lactamase from *Comamonas acidovorans* was solved in 2009 (PDB: 2WKN), its catalytic mechanism underlying the hydrolysis of (+)-γ-lactam remains elusive.[18] Recently, the crystal structure of a new family of γ-lactamase from *M. hydrocarbonoxydans* was solved.[19] The lactamase could hydrolyze both (+)- and (−)-γ-lactam, but with apparently different specificities. Based on structural studies, a mechanism for its different catalytic specificities of the enantiomers of γ-lactam was suggested.[19]

Except for these two examples, no other catalytic mechanism for (+)-γ-lactamase have been proposed. Here, the successful discovery of promiscuous (+)-γ-lactamase activity of the known signature amidase (PDB: 3A1I) provides an ideal model for investigating the mechanism of its absolute selectivity towards (+)-γ-lactam.[15] Based on the PDB model, this lactamase has a very interesting fold (Fig. 4a). Generally, the structure can be roughly separated into two domains: the core domain and α-helical *N*-terminal domain, which is far from the core domain. The α-helical *N*-terminal domain is mainly used to form a dimer.[15] This domain is isolated from the core structure, but directly interacts with the core domain of the other monomer in the dimer. The unusual "hook" structure may contribute to the excellent thermostability of this enzyme. Moreover, this *N*-terminal domain is directly located above the active site of its dimeric partner, which might be used to cover the substrate pathway and may be involved in substrate recognition.[15] The active center is in the center of the core domain. Previous studies showed that the signature amidase superfamily typically uses three residues (K-S-S) for catalysis.[15,20] In ReGL, these three residues are K96, S171, and S195 (Fig. 4b).