Tetrahedron Letters 54 (2013) 1380–1383

Contents lists available at SciVerse ScienceDirect

Tetrahedron Letters

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

Tataricins A and B, two novel cyclotetrapeptides from *Aster tataricus*, and their absolute configuration assignment

Hui-Min Xu [a,b,†], Hua Yi [a,†], Wen-Bing Zhou [a,b], Wen-Jun He [a], Gang-Zhi Zeng [a], Wen-Yan Xu [a], Ning-Hua Tan [a,*]

[a] State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, PR China
[b] University of Chinese Academy of Sciences, Beijing 100049, PR China

ARTICLE INFO

Article history:
Received 29 October 2012
Revised 14 December 2012
Accepted 25 December 2012
Available online 3 January 2013

Keywords:
Aster tataricus
Cyclotetrapeptides
Tataricin A
Tataricin B

ABSTRACT

Two novel cyclotetrapeptides, tataricins A and B, with a unique cyclopeptide backbone and a $^{\triangle 2,4}$Pro side chain, were isolated from the traditional Chinese medicine *Aster tataricus*. Their structures and absolute configurations were determined using a combination of spectroscopic data, the advanced Marfey's method, and a total synthesis.

The roots and rhizomes of *Aster tataricus* L. (Compositae) are a traditional Chinese medicine and have long been used for relieving cough and eliminating phlegm.[1] Previously it has been reported that the main components of the plant were terpenoids, sterols, flavonoids, and cyclopeptides.[2] It is noteworthy that this species is the only source of cyclopeptides in the family of Compositae.[2,3] Cyclopeptides in this plant named Compositae-type cyclopeptides are chlorinated homocyclopentapeptides and characterized by the presence of four nonproteinogenic amino acids (L-β-Phe, L-Abu, L-*allo*-Thr, and chlorinated L-Pro derivatives) and one proteinogenic amino acid (L-Ser).[3] So far, only nine cyclopentapeptides (astins A–I)[4] and eight acyclic pentapeptides (astin J,[5] asterinins A–F,[6] and aurantiamide acetate[7]) have been isolated from *A. tataricus*. Previous studies have shown that astins A–C exhibit antitumor activity in the Sarcoma 180A ascites in mice.[4a,8] Recently astin C was reported to possess an immunosuppressant property which is due to its induction of activated T-cell apoptosis by us.[9] The distinctive chlorinated pentacyclic feature and potential activities have attracted great interest.[10] Aiming at identifying structurally interesting cyclopeptides and bioactive metabolites from traditional Chinese medicines,[11] two novel cyclopeptides, tataricins A (**1**) and B (**2**) (Fig. 1), with an unique skeleton, were isolated from the titled plant.

The plants were commercially purchased from the Yunnan Lv-Sheng Pharmaceutical Co. Ltd., China and identified by Professor Xi-Wen Li at the Kunming Institute of Botany (voucher No. 200704). The methanol extract (13 kg) of the roots and rhizomes (50 kg) was suspended in water and partitioned successively with ethyl acetate (EtOAc) and *n*-butanol. The EtOAc fraction (2 kg) was subjected to silica gel column chromatography (CC) and eluted with gradient $CHCl_3$/MeOH to yield six fractions (Fr.1–Fr. 6). Fr.4 (85 g) was repeatedly submitted to silica gel CC, then purified by Sephadex LH-20 and semi-HPLC to afford tataricin A (**1**) (20 mg). Fr.6 (174 g) was fractionated repeatedly by silica gel CC, and then by semi-HPLC to obtain tataricin B (**2**) (15 mg).

Tataricin A (**1**)[12] was obtained as a white powder. Its molecular formula $C_{25}H_{31}N_5O_7$ was assigned by positive HRESIMS [(M+H)$^+$

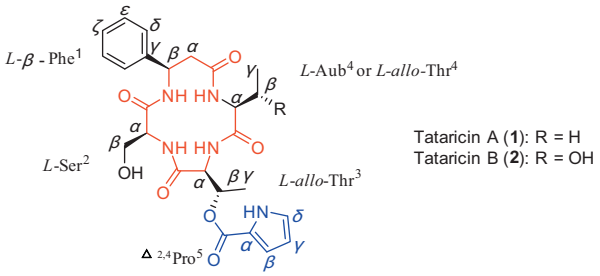


Figure 1. Structures of **1** and **2**.

* Corresponding author. Tel./fax: +86 871 5223800.
E-mail address: nhtan@mail.kib.ac.cn (N.-H. Tan).
† These authors contributed equally to this Letter.

0040-4039/$ - see front matter Crown Copyright © 2012 Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tetlet.2012.12.111

Table 1
^{1}H and ^{13}C NMR data for compounds **1** and **2** in DMSO-d_6 (600 and 150 MHz)

		1				**2**	
		δ_C	δ_H, mult (J in Hz)			δ_C	δ_H, mult (J in Hz)
L-β-Phe^1	α_1	42.6	2.55 (1H, d, 14.1)	L-β-Phe^1	α_1	43.0	2.56 (1H, overlapped)
	α_2		2.65 (1H, dd, 14.1, 3.6)		α_2		2.63 (1H, m)
	β	51.1	4.97 (1H, m)		β	51.4	4.93 (1H, m)
	γ	142.9			γ	143.2	
	δ	125.7	7.32 (2H, overlapped)		δ	125.7	7.32 (2H, overlapped)
	ε	128.5	7.29 (2H, overlapped)		ε	128.4	7.30 (2H, overlapped)
	ζ	126.7	7.21 (1H, t, 6.6)		ζ	126.6	7.22 (1H, m)
	C=O	169.3			C=O	169.2	
	NH		7.65 (1H, d, 7.2)		NH		7.96 (1H, br s)
L-Ser^2	α	55.7	4.14 (1H, m)	L-Ser^2	α	56.2	4.17 (1H, m)
	β_1	61.3	3.60 (1H, m)		β_1	61.4	3.60 (1H, m)
	β_2		3.29 (1H, m)		β_2		3.36 (1H, overlapped)
	β_{OH}		5.42 (OH, br s)				
	C=O	170.7			C=O	170.5	
	NH		7.94 (1H, d, 9.6)		NH		8.15 (1H, d, 7.2)
L-*allo*-Thr^3	α	58.3	4.58 (1H, t, 10.8)	L-*allo*-Thr^3	α	58.2	4.61 (1H, t, 10.2)
	β	67.8	5.35 (1H, m)		β	68.2	5.34 (1H, overlapped)
	γ	17.4	1.23 (3H, d, 6.0)		γ	17.4	1.28 (3H, d, 6.6)
	C=O	167.8			C=O	168.1	
	NH		7.98 (1H, d, 10.2)		NH		8.22 (1H, overlapped)
L-Aub^4	α	56.6	3.91 (1H, dd, 12.6, 5.8)	L-*allo*-Thr^4	α	60.8	4.09 (1H, dd, 7.2, 3.6)
	β	24.1	1.67 (2H, m)		β	65.4	3.99 (1H, m)
	γ	10.7	0.96 (1H, t, 7.2)		γ	19.5	1.16 (1H, d, 6.6)
	C=O	171.7			C=O	169.1	
	NH		8.22 (1H, d, 5.8)		NH		8.28 (1H, overlapped)
$^{\triangle 2,4}$Pro^5	α	121.5		$^{\triangle 2,4}$Pro^5	α	121.5	
	β	124.7	7.03 (1H, br s)		β	124.5	7.04 (1H, m)
	γ	109.9	6.16 (1H, br s)		γ	109.7	6.18 (1H, m)
	δ	116.0	6.76 (1H, br s)		δ	115.8	6.78 (1H, m)
	C=O	159.5			C=O	159.4	
	NH		11.90 (1H, s)		NH		11.91 (1H, s)

m/z 514.2295, calcd for 514.2301], indicating 13° of unsaturation. Its UV spectrum at 204 and 268 nm was ascribed to phenyl and $^{\triangle 2,4}$Pro groups. It was assumed to be a peptide from the IR absorptions at 3424 (NH, OH), 1660 (CO), and 1533 (C=C) cm^{-1}. The appearance of five amino carbonyl signals at δ_C 159.5, 167.8, 169.3, 170.7, and 171.7 in ^{13}C NMR spectrum and five NH signals at δ_H 7.65, 7.94, 7.98, 8.22, and 11.90 in ^{1}H NMR spectrum, reflected characteristics of pentapeptide. Further analysis of ^{1}H, ^{13}C NMR, and HMQC resonance signals displayed two methyls, three methylenes [one was oxidized 61.3 (Ser-C$_\beta$)], 13 methines [eight olefinic carbons 128.5 (two β-Phe^1-C$_\varepsilon$), 126.7 (β-Phe^1-C$_\zeta$), 125.7 (two β-Phe^1-C$_\delta$), 124.7 ($^{\triangle 2,4}$Pro^5-C$_\beta$), 116.0 ($^{\triangle 2,4}$Pro^5-C$_\delta$), 109.9 ($^{\triangle 2,4}$Pro^5-C$_\gamma$)], and seven quaternary carbon atoms [two olefinic carbons 142.9 (β-Phe^1-C$_\gamma$), 121.5 ($^{\triangle 2,4}$Pro-C$_\alpha$) and five carbonyls 171.7 (Aub^4-CO), 170.7 (Ser^2-CO), 169.3 (β-Phe^1-CO), 167.8 (*allo*-Thr^3-CO), 159.5 ($^{\triangle 2,4}$Pro^5-CO)] (Table 1). On the basis of the above data, together with an extensive 2D NMR analysis using COSY, HMQC, HMBC, and NOESY, it was inferred the assignment of five amino acid residues as β-Phe, Ser, Thr, Abu, $^{\triangle 2,4}$Pro.

The sequence of amino acid residues in **1** was determined from HMBC correlations between the carbonyl group and amide of the adjacent residues (Fig. 2). HMBC cross-peaks of β-Phe^1-NH (7.65)/Ser^2-CO (170.7), Ser^2-NH (7.94)/Thr^3-CO (167.8), Thr^3-NH (7.98)/Abu^4-CO (171.7), and Abu^4-NH (8.22)/β-Phe^1-CO (169.3) demonstrated the connectivities of cyclo (β-Phe^1-Ser^2-Thr^3-Abu^4). The correlations was observed for Thr^3-H$_\beta$ (5.35) to $^{\triangle 2,4}$Pro^5-CO (159.5), confirming the presence of $^{\triangle 2,4}$Pro^5-CO-O-Thr^3-C$_\beta$.

Tataricin B (**2**)[13] showed a pseudomolecular ion peak at *m/z* 552 [M+Na]$^+$, and the molecular formula of $C_{25}H_{31}N_5O_8$ was determined by HREIMS [M$^+$ *m/z* 529.2179, calcd for 529.2173], 16 units larger than **1**. The ^{1}H and ^{13}C NMR data displayed similar structural features as those found in **1** (Table 1). A detailed analysis of 1D NMR spectra showed the presence of two methyl groups at 17.4 (Thr^3-C$_\gamma$)/1.28 (Thr^3-H$_\gamma$) and 19.5 (Thr^4-C$_\gamma$)/1.16 (Thr^4-H$_\gamma$), and one additional methine group at 65.4 (Thr^4-C$_\beta$)/3.99 (Thr^4-H$_\beta$) in comparison with **1**, which suggested that the difference in **2** was the replacement of Abu^4 in **1** with Thr^4. These observations were confirmed by further analysis of COSY, HMQC, HMBC, and NOESY spectra (see Supplementary data). Thus, the planar structure of **1** and **2** were established.

Since compounds **1** and **2** are the first two cyclopeptides that contained a unique cyclotetrapeptide backbone with a $^{\triangle 2,4}$Pro side chain isolated from the family of Compositae, the advanced Marfey's method and a total synthesis of **1** and **2** were performed to unambiguously establish their structures and absolute configurations. The advanced Marfey's method[14] was first utilized to confirm the absolute stereochemistry of the amino acid residues

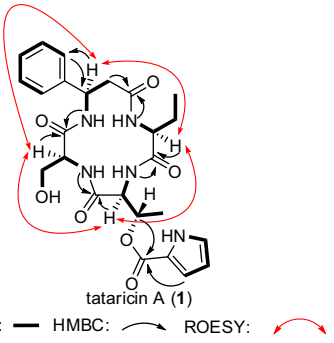
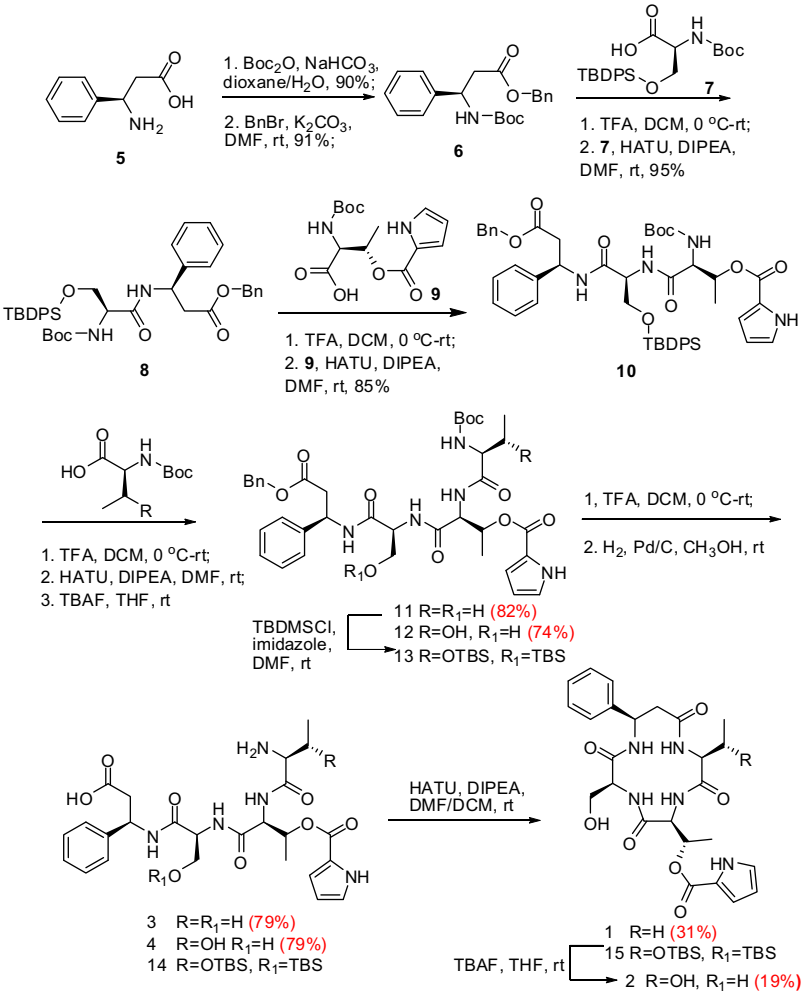


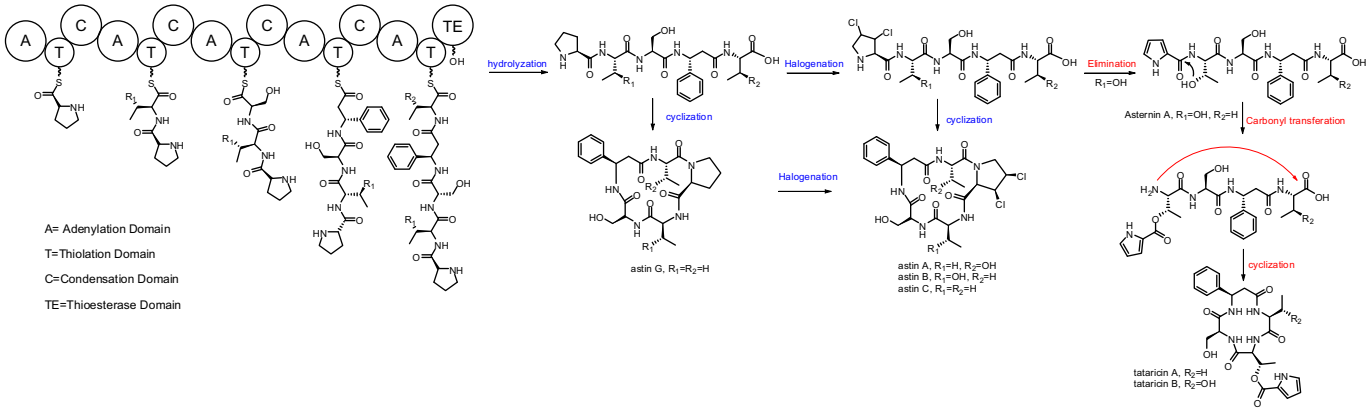
Figure 2. Selected ^{1}H–^{1}H COSY, HMBC, and ROESY correlations of **1**.



Scheme 1. Synthesis of **1** and **2**.

in **1** and **2**. The LC/MS analyses indicated that all amino acid residues and Thr were of L configuration, that is, L-β-Phe, L-Ser, and L-Aub (see Supplementary data). The L-*allo*-Thr and L-Thr standards were carried out and were subjected to the advanced Marfey's analysis using HPLC for confirming the L-configuration of the Thr in compounds. The identical retention times of the L-Marfey derivatives L-*allo*-Thr and Thr in the hydrolyzates of **1** and **2** suggested an L-configuration of the *allo*-Thr moiety. Then, the linkages between the amino acid residues and the $^{\triangle 2,4}$Pro moiety in **1** and **2** were further confirmed by a total synthesis.

Ultimate closure of the tetrapeptide was chosen to be a peptide linkage between the β-Phe and Abu or *allo*-Thr units.[15] This is sufficiently unhindered to allow a smooth reaction, and should be possible to achieve without any racemization. Following the strategy, (*R*)-3-amino-3-phenylpropionic acid (**5**) as the C-terminal amino acid, was transformed to **6** in two steps (Scheme 1).



Scheme 2. Plausible biogenetic pathway of **1** and **2**.

Removal of the Boc group of **6** using TFA was coupled with **7** using HATU and DIPEA as a coupling reagent,[16] leading to the dipeptide **8** (95% yield). Compound **8** was N-deprotected and coupled with **9** to afford tripeptide **10** (85% yield). The corresponding building block **9** was prepared from L-*allo*-threonine by protection with Boc_2O and benzyl bromide followed by the reaction with 1*H*-pyrrole-2-carbonyl chloride. Sequential addition of *N*-Boc-(*S*)-Abu or *N*-Boc-L-*allo*-Thr under standard procedures (HATU/DIPEA as coupling reagent) and cleavage of silyl ether of the hydroxyl group with TBAF delivered tetrapeptide **11** or **12**. The cyclization is a key step in the synthesis of cyclic peptides. After the Boc group and benzyl in tetrapeptide **11** were respectively removed, the free amine/free acid linear tetrapeptide **3** was dissolved in DMF slowly dropwise to HATU/DIPEA in CH_2Cl_2. The crude product was purified by preparative HPLC and provided the pure cyclic peptide **1** in 31% yield.

Tetrapeptide **12** has an identical sequence to **11**, but the major difference came from the last residue (**11** R = H, **12** R = OH). Therefore, protection of two hydroxyl groups by TBDMSCI in **12** afforded the corresponding compound **13**, followed by the deprotection of Boc protecting group with TFA and then reduced by hydrogenation with Pd/C to give the linear tetrapeptide **14**. The resulting peptide was cyclized in the presence of HATU/DIPEA to give **15**. Treatment of cyclic peptide **15** with TBAF to cleave silyl ether of the hydroxyl group and purification by preparative HPLC provided the pure cyclic peptide **2**. The yield of 4 steps was 19% (from **13** to **2**). Synthetic **1** and **2** were identical to the natural **1** and **2** in all respects ($[\alpha]_D$, MS, ^1H, ^{13}C NMR, COSY, HMQC, HMBC, NOESY spectra, and HPLC analysis, see Supplementary data).

To the best of our knowledge, tataricins A (**1**) and B (**2**) represent a new type of cyclotetrapeptide backbone with a $^{\triangle\ 2,4}$Pro side chain in nature. Therefore, the possible biosynthetic pathway of **1** and **2** should be quite different from other common Compositae-type cyclopeptides. In general, linear peptidyl backbone of Compositae-type cyclopeptides could be generated by Nonribosomal Peptide Synthetase (NRPS).[17] After the hydrolyzation, common Compositae-type cyclopeptides could be reached by immediate cyclization and chlorination. The hydrolyzed linear peptide might be further modified by elimination, carbonyl transferation, and cyclization, which could lead to compounds **1** and **2** (Scheme 2).

Biological activity of **1** and **2** was tested for cytotoxicity against BGC-823 and Hela cells as well as immunosuppressive[9] activity. Unfortunately, **1** and **2** are inactive.

In summary, we have isolated two rare cyclotetrapeptides, that is, tataricins A (**1**) and B (**2**), from the traditional Chinese medicine *A. tataricus*. Their structures and absolute configurations were unambiguously determined using a combination of spectroscopic data, the advanced Marfey's method, and a total synthesis. Compounds **1** and **2** possess a unique cyclotetrapeptide backbone and a $^{\triangle\ 2,4}$Pro side chain, which add a new skeleton to the diverse group of Compositae-type cyclopeptides. Additionally, the efficient synthesis of **1** and **2** should allow for the preparation of various analogues of this kind of cyclic peptide.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (U1032602, 91013002, 21102152, and 30725048), the National Basic Research Program of China (2009CB522300 and 2013CB127505), the National New Drug Innovation Major Project of China (2011ZX09307-002-02), the Foundation of Chinese Academy of Sciences (Hundred Talents Program), and the Natural Science Foundation of Yunnan Province (2012GA003 and 2011FZ206).

Supplementary data

Supplementary data (detailed isolation procedures, Marfey's reaction, 1D and 2D NMR spectra, MS, IR, UV, CD, $[\alpha]_D$ of compounds **1** and **2**, synthetic intermediates are supplied) associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.tetlet.2012.12.111. These data include MOL files and InChiKeys of the most important compounds described in this article.

References and notes

1. National Pharmacopoeia Committee of China, Ch. P, Chinese Medical Science Press, Beijing, 2010, p 322.
2. Lai, G. F.; Chen, J. J.; Luo, S. D. *Nat. Prod. Res. Dev.* **2002**, *14*, 65–70.
3. Tan, N. H.; Zhou, J. *Chem. Rev.* **2006**, *106*, 840–895.
4. (a) Morita, H.; Nagashima, S.; Takeya, K.; Itokawa, H.; Iitaka, Y. *Tetrahedron* **1995**, *51*, 1121–1132; (b) Morita, H.; Nagashima, S.; Shirota, O.; Takeya, K.; Itokawa, H. *Chem. Lett.* **1993**, *22*, 1877–1880; (c) Morita, H.; Nagashima, S.; Takeya, K.; Itokawa, H. *Heterocycles* **1994**, *38*, 2247–2252; (d) Morita, H.; Nagashima, S.; Takeya, K.; Itokawa, H. *Chem. Lett.* **1994**, *23*, 2009–2010.
5. Morita, H.; Nagashima, S.; Takeya, K.; Itokawa, H. *Chem. Pharm. Bull.* **1995**, *43*, 271–273.
6. (a) Cheng, D. L.; Shao, Y.; Hartman, R.; Roder, E.; Zhao, K. *Phytochemistry* **1994**, *36*, 945–948; (b) Cheng, D. L.; Shao, Y.; Hartmann, R.; Roeder, E.; Zhao, K. *Phytochemistry* **1996**, *41*, 225–227.
7. Wang, Z. T.; Lu, Y. H.; Ye, W. C.; Xu, L. S.; Xu, G. J.; Shu, Y. Z. *J. Chin. Pharm. Sci.* **1999**, *8*, 171–172.
8. Morita, H.; Nagashima, S.; Takeya, K.; Itokawa, H. *Chem. Pharm. Bull.* **1993**, *41*, 992–993.
9. Shen, Y.; Luo, Q.; Xu, H. M.; Gong, G. Y.; Zhou, X. B.; Sun, Y.; Wu, X. F.; Liu, W.; Zeng, G. Z.; Tan, N. H.; Xu, Q. *Biochem. Pharmacol.* **2011**, *82*, 260–268.
10. (a) Saviano, G.; Benedetti, E.; Cozzolino, R.; Capua, A. D.; Laccetti, P.; Palladino, P.; Zanotti, G.; Amodeo, P.; Tancredi, T.; Rossi, F. *Biopolymers* **2004**, *76*, 477–484; (b) Cozzolino, R.; Palladino, P.; Rossi, F.; Cali, G.; Benedetti, E.; Laccetti, P. *Carcinogenesis* **2005**, *26*, 733–739.
11. (a) Fan, J. T.; Chen, Y. S.; Xu, W. Y.; Du, L. C.; Zeng, G. Z.; Zhang, Y. M.; Su, J.; Li, Y.; Tan, N. H. *Tetrahedron Lett.* **2010**, *51*, 6810–6813; (b) Fan, J. T.; Su, J.; Peng, Y. M.; Li, Y.; Li, J.; Zhou, Y. B.; Zeng, G. Z.; Yan, H.; Tan, N. H. *Bioorg. Med. Chem.* **2010**, *18*, 8226–8234; (c) Han, J.; Ji, C. J.; He, W. J.; Shen, Y.; Leng, Y.; Xu, W. Y.; Fan, J. T.; Zeng, G. Z.; Kong, L. D.; Tan, N. H. *J. Nat. Prod.* **2011**, *74*, 2571–2575.
12. *Tataricin A (**1**):* white powder; $[\alpha]_D^{15.0}$ −26.1 (*c* 0.13, C_5H_5N); UV (MeOH) λ_{max} (logε) 204 (4.26), 268 (4.20); CD (*c* 0.098, MeOH) λ ($^{\triangle}\varepsilon$) 227 (−7.48), 271 (0.89); IR (KBr) ν_{max} 3424, 1660, 1533, 1410, 1305, 1149, 1075, 752, 578 cm^{-1}; ^1H NMR (600 MHz, DMSO-d_6) and ^{13}C NMR (150 MHz, DMSO-d_6), see Table 1; HRESIMS *m*/*z* 514.2295 $[M+H]^+$ (calcd for $C_{25}H_{32}N_5O_7$, 514.2301).
13. *Tataricin B (**2**):* white powder; $[\alpha]_D^{13.8}$ −5.3 (*c* 0.15, C_5H_5N); UV (MeOH) λ_{max} (logε) 207 (4.15), 267 (4.17); CD (*c* 0.098, MeOH) λ ($^{\triangle}\varepsilon$) 227 (−6.92), 271 (0.58); IR (KBr) ν_{max} 3426, 2930, 1725, 1679, 1548, 1289, 1127, 1076, 749 cm^{-1}; ^1H NMR (600 MHz, DMSO-d_6) and ^{13}C NMR (150 MHz, DMSO-d_6), see Table 1; HREIMS *m*/*z* 529.2179 $[M]^+$ (calcd for $C_{25}H_{31}N_5O_8$, 529.2173).
14. (a) Fujii, K.; Ikai, Y.; Mayumi, T.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 3346–3352; (b) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146–5151.
15. (a) Brady, S. F.; Varga, S. L.; Freidinger, R. M.; Schwenk, D. A.; Mendlowski, M.; Holly, F. W.; Veber, D. F. *J. Org. Chem.* **1979**, *44*, 3101–3105; (b) Tang, Y. C.; Xie, H. B.; Tian, G. L.; Ye, Y. H. *J. Pept. Res.* **2002**, *60*, 95–103.
16. Davies, J. S. *J. Pept. Sci.* **2003**, *9*, 471–501.
17. Xu, W. Y.; Li, L. L.; Du, L. C.; Tan, N. H. *Acta Biochim. Biophys. Sin.* **2011**, *43*, 757–762.