Published in final edited form as:

Tetrahedron. 2013 December 23; 69(51): 10797–10804. doi:10.1016/j.tet.2013.10.095.

Euryjanicins E–G, poly-phenylalanine and poly-proline cyclic heptapeptides from the Caribbean sponge *Prosuberites laughlini*

Edward Avilés and Abimael D. Rodríguez*

Department of Chemistry, University of Puerto Rico, PO Box 23346, UPR Station, San Juan, PR 00931-3346, United States

Abstract

A new investigation of the active sponge extracts of *Prosuberites laughlini* collected off the West coast of Puerto Rico has yielded three new cyclic heptapeptides, namely euryjanicins E (1)–G (3), containing multiple phenylalanine and proline residues. In CDCl₃ solution, each euryjanicin F (2) and G (3) exists as an inseparable complex mixture of conformational isomers. The molecular structures of 1–3 were elucidated by a combination of chemical degradation, extensive ESI-MS/MSⁿ analyses, and 2D NMR methods. The elucidation of the absolute configuration was achieved by HPLC following analysis of the acid hydrolysates after derivatization with Marfey's reagent. When assayed against the National Cancer Institute 60 tumor cell line panel, the new cyclic peptides did not display significant *in vitro* cytotoxicity.

Keywords

Prosuberites laughlini; Cyclic peptides; Caribbean sponge; Cytotoxicity; Peptides

1. Introduction

Marine invertebrates, especially sponges and tunicates, represent an abundant resource of novel bioactive metabolites including terpenoids, alkaloids, macrolides, polyethers, and nucleoside derivatives.¹ Marine sponges in particular, are a rich source of non-ribosomal cyclic peptides with seven to ten amino acids that often exhibit a wide range of desirable pharmacological properties, such as antibiotic, anti-inflammatory, and cytotoxic activity.² Non-ribosomal cyclic peptides are often of considerable interest because they can also be used as models for studies of the structural features of proteins.^{3–6} A particularly important class of marine cyclic peptides is the so called proline-rich compounds. Among the 20 naturally occurring amino acids, proline is the only one in which the side chain atoms form a

*Corresponding author. Tel.: + 1 787 764 0000x4799; fax: + 1 787 756 8242; abimael.rodriguez1@upr.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Supplementary data

Full NMR, ESI-MS, and ESI-MS/MSⁿ amino acid sequence data for cyclic peptides 1–3, Table 4, and photos of the sponge and collection site. Supplementary data associated with this article can be found in the online version, at doi:

pyrrolidine ring with the backbone atoms. This ring structure leads to some conspicuous properties of proline, namely, it induces conformational constraints among the atoms in the pyrrolidine ring, and it is the reason for the slow isomerization between *cis/trans* conformations and for the secondary structure preferences of poly-proline sequences.⁷ Remarkably, the *cis/trans* preference of proline-AA peptide bonds (where AA is any amino acid) varies in different solvent environments.⁸ Cyclic peptides having between seven to ten amino acids of which two to four are proline residues are typically designated as “proline-rich”.⁹ Similarly, “phenylalanine-rich” cyclic peptides have been associated with strong anticancer and antimicrobial activity.¹⁰ Some examples of this class of natural products, expressly isolated from marine sponges, are the hymenamides B¹¹ and E,¹² phakellistatin 11,¹³ axinellin A,¹⁴ and stylissamides C¹⁵ and F.¹⁶

In 2006, the marine sponge *Prosuberites laughlini* (Diaz, Alvarez & van Soest, 1987)¹⁷ (phylum: Porifera; class: Demospongiae; order: Hadromerida; family: Suberitidae) was collected during an underwater expedition from a fringing reef off the northwestern coast of Puerto Rico, near the town of Aguadilla (18°27'46.90"N, 67°10'06.11"W). The sponge specimens were triturated and repeatedly extracted with MeOH, and the combined extracts concentrated and partitioned between H₂O and EtOAc followed by *n*-butanol. In-house biological screening of both extracts quickly revealed that the EtOAc fraction displayed *in vitro* cytotoxicity. Routine ¹H and ¹³C NMR analysis revealed that the EtOAc extract comprised a complex mixture of cyclic peptides, which were subsequently characterized as euryjanicins A–D^{18,19} and dominicin.²⁰ Oddly, while the EtOAc extract was active, none of the individual peptides could account for the primary level of cytotoxicity.¹⁹ This led us to hypothesize that this sponge might contain trace amounts of highly toxic peptides. To test this hypothesis more than twice as much sponge material was re-collected from the original location and extracted. Upon re-investigating the chemical composition of the new extracts, we now report on the isolation, structure elucidation, and biological evaluation of euryjanicins E (1)–G (3), three sparse cyclic peptides rich in phenylalanine and proline residues.

2. Results and Discussion

2.1. Extraction and isolation of cyclic peptides

Freeze-dried samples of *P. laughlini* (1.6 kg) were repeatedly extracted with MeOH and the extracts were combined and concentrated *in vacuo*. The crude MeOH-soluble material was subsequently purified by liquid/liquid partitioning between EtOAc, *n*-BuOH and H₂O. The EtOAc phase was concentrated to dryness to give an oily residue consisting of substances of a peptidic nature. TLC-guided fractionation of the EtOAc-soluble components via flash chromatography over Si gel using a gradient of increasing polarity with MeOH/NH₃ in

CHCl₃ (98:2–1:1) yielded 12 fractions. Fractions eluted with 5–10% of MeOH/NH₃ in CHCl₃ were further purified by reversed-phase HPLC on a C-18 Spheri-5 column eluting with MeOH/H₂O to yield pure euryjanicin E (**1**) (7.2 mg; 0.0005%), euryjanicin F (**2**) (10.8 mg; 0.0007%) and euryjanicin G (**3**) (5.2 mg; 0.0003%), along with the same previously isolated cyclic peptides.^{18–20}

2.2. Structure elucidation of euryjanicin E

Compound **1**, the most polar compound, was isolated as an optically active white semi-solid with a molecular formula of C₄₄H₆₀N₈O₈ as deduced from HRESI-MS [M+H]⁺ (calcd 829.4607), requiring 19 sites of unsaturation. The presence of intense absorption peaks at 3317 and 1645 cm⁻¹ in the IR, corresponding to amino and carbonyl groups, respectively, along with some characteristic peaks in the ¹H- and ¹³C-NMR spectra (Table 1), strongly suggested a peptide like compound. In as much as eight amide carbonyls [δ 172.2, 172.1, 2 × 171.7, 171.6, 170.9, 170.8, and 170.4] but only seven peaks for amino acid α-carbons [δ 61.4, 61.2, 61.2, 56.2, 54.1, 52.8, 49.2] were detected in the ¹³C NMR spectrum, we surmised that **1** was a heptapeptide bearing either a glutamine or an asparagine residue. Additionally, eight resonances about the alkene region hinted at the presence of two monosubstituted phenyl rings [δ 135.8, 135.3, 129.6, 129.2, 128.9, 129.8, 127.5, 127.3], which suggested the occurrence in **1** of two phenylalanine residues. Analysis of the COSY, TOCSY and HSQC–TOCSY NMR data revealed only five peptide-bond secondary amide proton signals thus implying the presence of two proline units in the heptapeptide. Further analysis of these data allowed us to quickly assemble seven independent spin systems. TOCSY correlations between NH¹Ile¹, αH¹Ile¹, βH¹Ile¹, γH²Ile¹, δH³Ile¹, and εH³Ile¹ clearly established the presence of an isoleucine residue. Likewise, another spin system of the type –CH–CH(CH₃)–CH₂–CH₃ was observed between NH⁵Ile⁵, αH⁵Ile⁵, βH⁵Ile⁵, γH²Ile⁵, δH³Ile⁵, and εH³Ile⁵ that was attributed to another isoleucine unit. Moreover, the COSY spectrum allowed us to establish two independent spin systems of the type X–CH–CH₂–CH₂–CH₂–X' that were ascribable to a pair of proline residues. The first one, Pro², was profiled from a combination of COSY (αHPro², βH²Pro², γH²Pro², and δH²Pro²) and TOCSY correlations (αHPro² and γH²Pro²), while Pro⁶ was similarly outlined from COSY correlations between αHPro⁶, βH²Pro⁶, γH²Pro⁶, and δH²Pro⁶. Of the three spin systems remaining, two were assigned as phenylalanine residues from the strong COSY correlations between NH³Phe³, αHPhe³, and βH²Phe³, and those of NH⁷Phe⁷, αHPhe⁷, and βH²Phe⁷. As for the final spin system in **1**, we attributed it to an asparagine unit from TOCSY correlations between αHAsn⁴ and βHAsn⁴ and from subsequent analysis of the combined HMBC (Table 1) and electrospray tandem mass spectrometry (ESI-MS/MSⁿ) data (see Supplementary data).

The amino acid composition of euryjanicin E (**1**) was corroborated by HPLC analysis of the acid hydrolysate obtained from the reaction of the peptide with 6 N HCl followed by Marfey's derivatization with FDAA.²¹ One mol of aspartate (Asp) and two moles each of phenylalanine (Phe), proline (Pro), and isoleucine (Ile) residues were generated and determined to have the L-configuration at the α-carbon by comparison to amino acid standards. These seven amino acids, minus two hydrogens, accounted for all the atoms in the molecular formula of euryjanicin E (**1**) and all but one site of unsaturation. Hence, the degree of unsaturation leftover could be accounted for if **1** were cyclic, a contention

supported by its high solubility in organic solvents and the fact that it tested negative to a ninhydrin test.

The amino acid sequence in **1** was elucidated by interpretation of HMBC and ROESY data as illustrated in Fig. 1, and was later confirmed by ESI-MS/MSⁿ. The identity of all of the carbonyls in the cyclic heptapeptide was established from HMBC correlations to the corresponding protons in the α , β or NH positions within each residue. For instance, COPro² correlations to α HPro², β H₂Pro², and NHile¹, placed Pro² adjacent to Ile¹. The COAsn⁴ cross peaks to NHAsn⁴ and NHPhe³, together with correlations between COIle⁵ and NHAsn⁴ indicated that Asn⁴ was flanked by Phe³ and Ile⁵. This generated two amino acid fragments, namely, Pro²-Ile¹ and Phe³-Asn⁴-Ile⁵. In turn, COPro⁶ correlated to α HPro⁶, β H₂Pro⁶ and NHile⁵, conclusively elongating the latter amino acid chain to Phe³-Asn⁴-Ile⁵-Pro⁶. Key cross peaks in the ROESY spectrum between α Hile¹ and NHPhe⁷ and those of α HPro⁶ and α HPh⁷ allowed the merging of fragments Pro²-Ile¹-Phe⁷ and Phe³-Asn⁴-Ile⁵-Pro⁶-Phe⁷. Additional cross peaks between α HPro² and α HPh³ allowed us to bring about the entire amino acid sequence of euryjanicin E (**1**) as *cyclo*-(Phe-Pro-Ile-Asn-Phe-Pro-Ile-).

The amide bond conformations of Pro² and Pro⁶ within **1** were determined to be *cis* on the basis of the differential value of ¹³C chemical shifts of C _{β} and C _{γ} for each proline residue ($\delta_{\beta\gamma}$),^{22,23} namely, 9.6 and 8.8 ppm, respectively, and the strong ROESY cross peaks between α H of Pro² and Pro⁶ with the α H of the adjacent amino acids (i.e. α HPro²/ α HPh³ and α HPro⁶/ α HPh⁷). Hence, the absolute stereostructure for the solution conformation of euryjanicin E (**1**) in CDCl₃ at 25 °C was established as *cyclo*-(Phe-*cis*-Pro-Ile-Asn-Phe-*cis*-Pro-Ile-).

2.3. Structure elucidation of euryjanicin F

Compound **2** was obtained as a pure white amorphous powder. The molecular weight of euryjanicin F (**2**) (m/z 884.4756 [M+Na]⁺) was established from HES-MS and indicated the molecular formula C₄₉H₆₃N₇O₇Na, implying 22 degrees of unsaturation. The ¹H NMR spectrum in CDCl₃, like that of euryjanicin E (**1**), exhibited tell-telling resonances indicative of a compound of a peptidic nature. However, the ¹H- and ¹³C NMR spectra of **2** were visibly more complex than those of **1** as many, if not all, of the signals were doubled or broadened, suggesting that two interconverting solution conformations of comparable stability (1:1 ratio) existed at room temperature. Attempts to separate the two conformers by reversed-phased HPLC under a variety of experimental conditions (columns, solvent compositions, and flow rates) invariably led to a single sharp peak and a “clean” molecular ion in the HPLC-ESI-MS/MSⁿ. Neither changing the solvent (MeOH-*d*₄, DMSO-*d*₆, MeCN-*d*₃, acetone-*d*₆, and pyridine-*d*₅), heating the sample to ~60 °C or cooling it to -15 °C, nor adjusting the pH by addition of TFA or Et₃N, shifted the equilibrium in favor of a single conformer. Notwithstanding, when recorded in pyridine-*d*₅, the NMR signals were doubled in a 3:1 ratio making spectral interpretation more amenable to a structure elucidation study.²⁴

By combining ^{13}C and DEPT135 NMR spectra with data obtained from HSQC experiments, it was established that the major solution conformer of euryjanicin F had 49 well-resolved carbon resonances. Of these, seven resonances possessed chemical shifts appropriate for amide carbonyls [δ 173.4, 171.6, 171.5, 171.4, 171.1, 170.4, 169.9], twelve were for three monosubstituted phenyls [δ 137.9, 136.3, 136.1, 130.9, 129.9, 129.7, 129.1, 129.0, 128.7, 127.7, 127.6, 127.0], and seven had chemical shifts for amino acid α -carbons [δ 62.0, 61.7, 61.4, 56.8, 56.7, 55.5, 53.7] consistent with a heptapeptide with three phenylalanine residues (Table 2). Further inspection of the ^1H NMR spectrum revealed five amide NH signals indicative of a peptide containing two proline units. Careful interpretation of the HSQC, COSY, TOCSY, and HSQC-TOCSY NMR spectra led to the identification of the following seven amino acids: 2 x Pro, 2 x Ile, and 3 x Phe. As these amino acid residues accounted for only 21 of the 22 sites of unsaturation, and no further evidence was found for a terminal amino or carboxylic acid functionality, it stands to reason that heptapeptide **2** was also cyclic. The presence of only proline, isoleucine, and phenylalanine in this cyclic peptide was confirmed by acid hydrolysis of **2** followed by Marfey's HPLC analysis of the hydrolysate,²¹ which also established the absolute configuration of all the amino acids to be L.

Following identification of all of the amide protons in **2** from COSY correlations to their corresponding amino acid α -protons, the amino acid sequence for euryjanicin F (**2**) was easily established by combination of the ROESY and HMBC data. For instance, COPhe^2 gave a strong HMBC cross peak with NHPh^1 , which in turn correlated to COPhe^1 and $\beta\text{H}_2\text{Phe}^1$, establishing the first fragment as Phe^1 - Phe^2 . HMBC correlations $\text{COPro}^7/\alpha\text{HIle}^6$, $\text{COIle}^6/\text{NHPh}^5$, and $\text{COPro}^7/\text{NHle}^6$ provided the segment Phe^5 - Ile^6 - Pro^7 . Further analysis of the ROESY spectrum revealed a strong correlation between $\alpha\text{HPh}^1/\alpha\text{HPro}^7$ that allowed us to interconnect the latter two segments thus establishing the sequence $-\text{Phe}^5$ - Ile^6 - Pro^7 - Phe^1 - Phe^2 -. Additional HMBC and ROESY cross peaks between $\alpha\text{HPh}^5/\alpha\text{HPro}^4$ and $\text{NHPh}^2/\alpha\text{HIle}^3$ brought about the complete amino acid sequence of euryjanicin F (**2**) as *cyclo*-(- Phe-Phe-Ile-Pro-Phe-Ile-Pro-) (see Table 2 and Fig. 1).²⁴

Close inspection of $\delta_{\beta\gamma}$ in the ^{13}C NMR spectrum for the two conformational variants of euryjanicin F in pyridine- d_5 revealed the geometry for each proline peptide bond.²⁵ In the major solution conformer **2**, both proline amide bonds adopt the *cis* geometry ($\delta_{29.0-21.9} = 7.1$ ppm for Pro^4 ; $\delta_{30.9-22.5} = 8.4$ ppm for Pro^7). This geometry was further supported by the ROESY cross peaks $\alpha\text{HPh}^1/\alpha\text{HPro}^7$ and $\alpha\text{HPh}^5/\alpha\text{HPro}^4$ (Fig. 1). In the minor solution conformer (not shown), one of the proline amide bonds (Pro^4) adopts the *trans* geometry ($\delta_{29.9-25.8} = 4.1$ ppm) whereas the same bond of Pro^7 assumes the *cis* geometry ($\delta_{314-22.2} = 9.2$ ppm).²⁶ Additional support for the latter geometries stems from ROESY cross peaks $\text{NHle}^3/\text{H}\alpha\text{Pro}^4$ and $\text{H}\alpha\text{Phe}^1/\text{H}\alpha\text{Pro}^7$ plus the conspicuous absence of cross peaks between $\alpha\text{HPh}^5/\alpha\text{HPro}^4$. Interestingly, the sequences of euryjanicin E (**1**) and euryjanicin F (**2**) differ by only one amino acid, namely, the fragment Phe-Asn in **1** is replaced by Phe-Ile in **2**.

2.4. Structure elucidation of euryjanicin G

Peptide **3** was the least abundant peptide component isolated from the sponge, with an accurate mass measurement of m/z 868.4421 ($[M+Na]^+$ 4.7 mDa from that calculated for $C_{48}H_{59}N_7O_7$). Like euryjanicin F (**2**), euryjanicin G (**3**) generated a single molecular ion in the HRESI-MS and always gave a sharp well-resolved peak when analyzed by reversed-phase HPLC under a variety of solvent systems. Despite these indications of homogeneity, the 1H and ^{13}C NMR spectra of **3** in $CDCl_3$ were difficult to analyze due to the complexity caused by multiple conformations resulting in poor resolution. Subsequently, a single set of well-resolved resonances could be obtained at 25 °C when pyridine- d_5 was used as the solvent, revealing that **3** can exist as a single solution conormer.²⁷

The structure of euryjanicin G (**3**) was elucidated by interpretation of 1D and 2D NMR data in pyridine- d_5 (Table 3). The 1H NMR spectrum of **3** showed four amide protons at δ_H 10.42 (1H, br s, NHPhe⁵), 9.79 (1H, d, J 9.1 Hz, NHile¹), 9.73 (1H, d, J = 9.7 Hz, NHPhe⁷), and 7.35 (1H, d, J = 4.5 Hz, NHPhe³), while the ^{13}C NMR spectrum contained seven carbonyl signals at δ_C 173.0, 172.8, 172.0, 171.9, 171.4, 170.4, and 169.3 along with seven amino acid αC 's (δ_C 62.3, 61.3, 59.0, 56.5, 55.0, 54.6, and 53.2). The presence of 13 methylene carbons along with 23 sites of unsaturation suggested a substantial number of proline and phenylalanine moieties constituting the structural backbone. Further analysis of the 1H - 1H COSY, HSQC, HMBC, and HSQC-TOCSY spectra revealed spin systems corresponding to three phenylalanines, one isoleucine, and three proline residues (Table 3). The amino acid composition was confirmed by HPLC analyses of the acid hydrolysate, and since only 22 degrees of unsaturation could be accounted for by the functionality present in the seven individual amino acids, it was apparent that **3** too was a cyclic peptide. Marfey's analysis was used to establish the absolute configuration as L for all the constituent amino acids of euryjanicin G (**3**).²¹

The amino acid sequence in **3** was established by combination of HMBC, ROESY, and ESI-MS/MSⁿ data. The fragment -Pro²-Phe³- was assigned from the HMBC correlations of COPro²/NHPhe³ and COPhe³/NHPhe³. Furthermore, cross peaks for COPro⁴/NHPhe⁵ and COPhe⁵/ δH_2 Pro⁶ determined the partial sequence -Pro⁴-Phe⁵-Pro⁶- while the correlations between COPro⁶/NHPhe⁷ and COPhe⁷/NHile¹ lengthened the chain to Pro⁴-Phe⁵-Pro⁶-Phe⁷-Ile¹. Strong ROESY cross peaks between αH Pro²/ αH Ile¹ made it possible to link fragments -Pro²-Phe³- and Pro⁴-Phe⁵-Pro⁶-Phe⁷-Ile¹. Overall, this sequence was supported by the ROESY correlations between $NH_{(i)}$ and $\alpha H_{(i-1)}$ and, in case of proline residues, $\alpha H_{(i)}$ Pro and $\alpha H_{(i-1)}$. Pivotal NOE cross peaks between αH Phe³ and αH Pro⁴ allowed us to close the ring.

As before, the amide bond configuration of Pro², Pro⁴, and Pro⁶ was determined to be *cis* on the basis of $\delta_{\beta\gamma}$ in all three proline residues, which was 9.0 ppm. Furthermore, pivotal ROESY cross peaks between αH 's of Pro², Pro⁴, and Pro⁶ and the αH 's of the adjacent amino acid, which are present in the *cis* conformation but absent in the *trans*-proline,²⁸ and the strong correlations between αH Pro²/ αH Ile¹, αH Pro⁴/ αH Phe³, and αH Pro⁶/ αH Phe⁵, confirmed that all three proline amide bonds adopted a *cis* geometry. Thus, the entire amino acid sequence for **3** was established as *cyclo*-(Ile¹-*cis*-Pro²-Phe³-*cis*-Pro⁴-Phe⁵-*cis*-Pro⁶-

Phe⁷). Remarkably, the sequences of euryjanicin G (3) and stylissamide C, a cyclic heptapeptide reported in 2007 from the Bahamian sponge *Stylissa caribica*, differ by only one amino acid.¹⁵ While all three proline amide bonds in these two compounds adopt the *cis* geometry, Phe⁵ in euryjanicin G is replaced by tyrosine in stylissamide C.

2.5. Biological activity

Euryjanicins 1–3 were tested at a single high dose (10 μ M) in the full NCI 60 cell panel, but none satisfied the NCI's predetermined threshold inhibition criteria for evaluation against the 60 cell panel at five concentration levels. Most of the cells displayed growths above 90%, with the exception of leukemia cell lines MOLT-4 and RPMI-8226, wherein euryjanicin F (2) elicited growth percentages of 49% and 43%, thus implying growth suppressions of 51% and 57%, respectively.

3. Conclusion

Like their predecessors, the new peptides could not account for the incipient cytotoxicity of the primary extracts.^{18, 19} This may be due to changes in the conformation of the cyclic peptides during the isolation process.²⁹ Compounds 1–3 are characterized by the co-occurrence of two to three proline and phenylalanine residues, one or two strongly hydrophobic isoleucine side chains and, in the case of 1, a polar uncharged asparagine residue. Remarkably, these metabolites are the first examples of cyclic peptides from a sponge to possess the same *cis*-Pro-Phe-Ile-*cis*-Pro-Phe segment. Each 2 and 3, exists as an inseparable complex mixture of conformational isomers about the proline residues, supporting the contention that proline-rich cyclic peptides often adopt well-defined conformations in solution, most likely from distinctive folds, hydrogen bonding patterns, and solvent accessible surfaces that contribute to the independent stability of the dominant conformer.³⁰ To no small extend, the fact that compounds 1–3 display a remarkable analogy to several other sponge-derived proline-rich antitumor cyclopeptides^{11–14} as to their amino acid content and sequential homology, might suggest that the construction of a suitable pharmacophore model for rationalizing their structure-activity relationship is possible. Thus, cyclopeptides such as 1–3 isolated from marine sponges could ultimately become indispensable tools for understanding the role of prolyl *cis*↔*trans* isomerization as a common means of switching between alternative functional states of folded proteins.²⁶

4. Experimental section

4.1. General procedures

Optical rotations were recorded with a Rudolph Autopol IV polarimeter. UV and IR spectra were recorded with Shimadzu UV-2401P and Nicolet Magna IR 750 spectrometers, respectively. ID and 2D NMR spectra were recorded on a Bruker DRX-500 or a Bruker AV-500 FT-NMR spectrometer. HPLC was performed on an Agilent 1260 infinity equipped with an Agilent 1260 photodiode array detector. Column chromatography was performed on Analtech Si gel (35–75 mesh) and monitored by TLC analysis carried out on Analtech glass precoated Si gel plates and visualized using UV light ($\lambda = 254$ nm) or I₂ vapors. The percentage yield of each compound is based on the weight of the dry sponge.

4.2. Animal material

The sponge *Prosuberites laughlini* (Diaz, Alvarez & van Soest, 1987)¹⁷ (phylum: Porifera; class: Demospongiae; order: Hadromerida; family: Suberitidae) was collected in April 2009 and identified by J. Vicente. A voucher specimen (No. PLAG09–02) is stored at the Chemistry Department of the University of Puerto Rico.¹⁹

4.3. Extraction and isolation of cyclic peptides

The freshly collected sponge was freeze-dried for 3 days, and the dried organism (1.6 kg) was repeatedly extracted with MeOH (20 L). The combined MeOH extracts were evaporated to dryness, and the resulting brown oil (196 g) was partitioned between EtOAc (5 × 1 L), *n*BuOH (5 × 1 L), and H₂O (1 × 2 L). The combined EtOAc extracts were concentrated *in vacuo* to give 18 g of a dark brown oil that was flash chromatographed over Si gel (600 g) using a gradient of increasing polarity with MeOH/NH₃ in CHCl₃ (98:2–1:1) as mobile phase and separated into 12 fractions on the basis of TLC and ¹H NMR analyses. Fraction 4 (2.8 g) was subsequently chromatographed over Si gel (100 g) using a gradient of increasing polarity with MeOH/NH₃ in CHCl₃ (98:2–4:1) as eluent resulting in 13 fractions. Subfractions 2–7 contained all of the cyclic peptides. Fraction 2 was purified by reversed-phase chromatography with a 5 mm × 250 mm Spheri-5 C18 column, 5 μm, with an isocratic solvent composition of (A) 42% H₂O and (B) 58% MeOH over 25 min (flow rate = 1 mL/min) with UV detection set at 220 nm to yield pure euryjanicin C (71 mg; 0.005%), euryjanicin D (80 mg; 0.005%), euryjanicin G (3) (5.2 mg; 0.0003%), and euryjanicin F (2) (10.8 mg; 0.0007%).¹⁹ Fraction 3 contained a mixture of euryjanicins C and D (143.0 mg, 0.009%). Fractions 4, 5 and 6 consisted of pure dominicin (40.0 mg, 0.003%), euryjanicin A (65.0 mg, 0.004%) and euryjanicin B (90.0 mg, 0.006%), respectively.^{18–20} Purification of fraction 7 via C18 reversed-phase HPLC using a 10 mm × 25 cm Ultrasphere ODS column, 5 μm, with 20% H₂O in MeOH yielded pure euryjanicin E (1) (7.2 mg; 0.0005%).

4.3.1. Euryjanicin E (1)—White semi- solid; [α]_D²⁰ = −78.0 (*c* 1.0, CHCl₃); IR (film) $\nu_{\max}$ 3317 (br), 3063, 3030, 2963, 2932, 2878, 1645 (br), 1526, 1437, 750, 702 cm^{−1}; UV (MeOH) $\lambda_{\max}$ (log ϵ) 192 (ϵ 4.8) nm; ¹H NMR (500 MHz, CDCl₃) and ¹³C NMR (125 MHz, CDCl₃) (see Table 1); HR(+)-ESI-MS *m/z* [M+H]⁺ 829.4606 (calcd for C₄₄H₆₁N₈O₈, 829.4607).

4.3.2. Euryjanicin F (2)—White powder; [α]_D²⁰ = −102.0 (*c* 1.0, CHCl₃); IR (film) $\nu_{\max}$ 3312 (br), 3063, 3030, 2963, 2932, 2878, 1643 (br), 1520, 1447, 750, 702 cm^{−1}; UV (MeOH) $\lambda_{\max}$ (log ϵ) 203 (ϵ 4.7) nm; ¹H NMR (500 MHz, C₅D₅N) and ¹³C NMR (125 MHz, C₅D₅N) (see Table 2); HR(+)-ESI-MS *m/z* [M+Na]⁺ 884.4756 (calcd for C₄₉H₆₃N₇O₇Na, 884.4681).

4.3.3. Euryjanicin G (3)—White semi-solid; [α]_D²⁰ = −62.0 (*c* 1.0, CHCl₃); IR (film) $\nu_{\max}$ 3282 (br), 3062, 3030, 2956, 2852, 1714, 1658, 1633 (br), 1552, 1452, 1346, 1247, 1188, 748, 702 cm^{−1}; UV (MeOH) $\lambda_{\max}$ (log ϵ) 203.5 (ϵ 4.5) nm; ¹H NMR (500 MHz, C₅D₅N) and ¹³C NMR (125 MHz, C₅D₅N) (see Table 3); HR(+)-ESI-MS *m/z* [M+Na]⁺ 868.4421 (calcd for C₄₈H₅₉N₇O₇Na, 868.4374).

4.4. Acid hydrolysis of euryjanicins E (1)–G (3)

Pure euryjanicins E–G (0.5 mg) were hydrolyzed in 0.5 mL of 6 N HCl at 110 °C for 12 h in a 1.0 mL reaction vial. The cooled reaction mixture was evaporated to dryness and traces of HCl were removed from the residual hydrolysate by repeated evaporation from H₂O (3 × 0.5 mL) using N₂ gas.

4.5. Absolute configuration of amino acids²¹

To a 4 mL vial containing 1 μmol of pure amino acid standards in 200 μL of H₂O was added 1 μmol of *N*-R-(2,4-dinitro-5-fluorophenyl)-L-alanine amide (L-FDAA) in 400 μL of acetone followed by 100 μL of 1 N NaHCO₃. The mixture was heated for 1 h at 40 °C. After cooling to rt, 100 μL of 2 N HCl was added and the resulting solution was filtered through a small 4.5 mm filter and stored in the freezer until HPLC analysis. Half of each peptide hydrolysate mixture was dissolved in 200 μL of H₂O, and to this was added 1.5 μmol of L-FDAA in 400 μL of acetone followed by 100 μL of 1 N NaHCO₃. The derivatization reaction was carried out and worked up as described above. An 8 μL aliquot of the resulting mixture of L-FDAA derivatives was analyzed by reversed-phase HPLC. A 5 mm 250 mm Spheri-5 C18 column, 5 μm, with a linear gradient of (A) 9:1 triethylammonium phosphate (50 mM, pH 3.0)/CH₃CN and (B) CH₃CN with 0% B at start to 40% B over 55 min (flow rate = 1 mL/min) was used to separate the L-FDAA derivatives with UV detection set at 340 nm. Each chromatographic peak was identified by comparing its retention time with the L-FDAA derivative of the pure L-amino acid standard and by co-injection. In all cases a peak at 33.9 min was observed, which was attributed to excess L-FDAA. Retention times (min) are given in parentheses: L-Pro (30.37), D-Pro (34.14), L-Ile (43.52), D-Ile (51.25), L-Phe (43.54), D-Phe (49.35), L-Asp (20.98), D-Asp (22.13).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank B. Vera and J. Ascencio for collecting the sponge, K. Nieves, M. Muñoz, and C. Jiménez for logistic support, Dr. I. E. Vega for ESI-MS/MSⁿ data, and the NCI for cytotoxicity assays. NTH Grant 1SC1GM086271-01A1 and the UPR-RISE/Fellowship Program provided financial support.

References and notes

1. Mayer AMS, Rodriguez AD, Tagliatela-Scafati O, Fusetani N. *Marine Drugs*. 2013; 11:2510–2573. [PubMed: 23880931]
2. Matsunaga S, Fusetani N. *Curr. Org. Chem.* 2003; 7:945–966.
3. Auvin-Guette C, Baraguey C, Blond A, Xavier HS, Pousset JL, Bodo B. *Tetrahedron*. 1999; 55:11495–11501.
4. Laupacis A, Keown PA, Ulan RA, McKenzie N, Stiller CR. *Can. Med. Assoc. J.* 1982; 126:1041–1043. [PubMed: 7074504]
5. Kirkpatrick P, Raja A, LaBonte J, Lebbos J. *Nat. Rev. Drug Discov.* 2003; 2:943–948. [PubMed: 14756140]
6. Sandhu P, Xu X, Bondiskey PJ, Balani SK, Morris ML, Tang YS, Miller AR, Pearson PG. *Antimicrob. Agents Chemother.* 2004; 48:1272–1277. [PubMed: 15047529]

7. Balbach, J.; Schmid, FX. Mechanisms of Protein Folding. Pain, RH., editor. New York: Oxford University Press; 2000. p. 212-249.
8. Buskirk AR, Green R. Science. 2013; 339:38–39. [PubMed: 23288527]
9. Kay BK, Williamson MP, Sudol M. FASEB J. 2000; 14:231–241. [PubMed: 10657980]
10. Dahiya R. Acta Pol. Pharm. Drug Res. 2007; 64:509–516.
11. Kobayashi J, Tsuda M, Nakamura T. Tetrahedron. 1993; 49:2391–2402.
12. Tsuda M, Shigemori H, Mikamia Y, Kobayashi J. Tetrahedron. 1993; 49:6785–6796.
13. Pettit G, Tan R, Ichihara Y, Williams MD, Doubek DL, Tackett LP, Schmidt JM. J. Nat. Prod. 1995; 58:961–965. [PubMed: 7673945]
14. Randazzo A, Dal Piaz F, Orrù S, Debitus C, Roussakis C, Pucci P, Gomez-Paloma L. Eur. J. Org. Chem. 1998:2659–2665.
15. Schmidt G, Grube A, Köck M. Eur. J. Org. Chem. 2007:4103–1110.
16. Cychon C, Köck M. J. Nat. Prod. 2010; 73:738–742. [PubMed: 20349940]
17. (a) Díaz MC. Caribb. J. Sci. 2005; 41:465–475. (b) Van Soest, RWM. Systema Porifera: A guide to the Classification of Sponges. Hooper, JNA.; Van Soest, RWM., editors. Vol. Vol. 1. New York: Kluwer Academic/Plenum Publishers; 2002. p. 227-244.
18. Vicente J, Vera B, Rodríguez AD, Rodríguez-Escudero I, Raptis RG. Tetrahedron Lett. 2009; 50:5493–5496.
19. Vera B, Vicente J, Rodríguez AD. J. Nat. Prod. 2009; 72:1555–1562. [PubMed: 19743810]
20. Williams DE, Patrick BO, Behrisch HW, Van Soest R, Roberge M, Andersen RJ. J. Nat. Prod. 2005; 68:327–330. [PubMed: 15787430]
21. Marfey P. Carlsberg Res. Commun. 1984; 49:591–596.
22. Dormann DE, Borvey FA. J. Org. Chem. 1973; 38:1719–1722.
23. Dormann DE, Borvey FA. J. Org. Chem. 1973; 38:2379–2383.
24. Although the co-occurrence of a minor solution conformer made the structural elucidation of 2 more difficult, it also provided an opportunity to confirm the proposed structure for euryjanicin F. Indeed, analysis of the 2D NMR data ascribable to the minor solution conformer of 2 led us to the same amino acid sequence (see Table 4 in the Supplementary data).
25. The coexistence of stable conformers in solution is a common feature of proline-rich cyclic peptides. Cyclic peptides having proline residues switching from the *cis* to *trans* configuration in solution have been reported to exist as inseparable mixtures of conformers; see Siemon IZ, Wieland T, Pook KH. Angew. Chem. Int. Engl. 1975; 14:702–703. Angew. Chem. 1975, 88, 712–714. Tabudravu JN, Jaspars M, Morris LA, Van den Bosch JJK, Smith N. J. Org. Chem. 2002; 67:8593–8601. and refs. 22 and 23. [PubMed: 12444643]
26. Cyclic peptides differing by only the *cis-trans* isomerization of the proline peptide bonds can manifest different folds, hydrogen-bonding patterns, and solvent accessible surfaces. These characteristics can also render large differences in biological activity as the particularly low activation barrier for the prolyl amide bond to switch between *cis* and *trans* conformations would impact the three-dimensional structure of the peptide; see: Montaser R, Abboud KA, Paul VJ, Luesch H. J. Nat. Prod. 2011; 74:109–112. [PubMed: 21138309] Williams DE, Yu K, Behrisch HW, Van Soest R, Andersen RJ. J. Nat. Prod. 2009; 72:1253–1257. and refs. 9 and 25b. [PubMed: 19583251]
27. It has been found that the rate for this interconversion is solvent dependent, being higher in nonpolar solvents, where the charge separation and resonance stabilization of the ground state of all the conformers are lost, lowering the *cis-to-trans* torsion barrier; see: John JS, Kang YK. J. Phys. Chem. A. 1999; 103:5436–5439. Eberhardt ES, Loh SN, Hinck AP, Raines RT. J. Am. Chem. Soc. 1992; 114:5437–5439. and ref. 23. [PubMed: 21451730]
28. Tonelli AE. J. Am. Chem. Soc. 1973; 95:5946–5948. [PubMed: 4733830]
29. The antitumor activity of some cyclic peptides have been shown to be highly dependent on the solvent (typically DMSO or MeOH) and the time the sample is left in the solvent before the assay is conducted; see ref. 25b
30. Two stable conformers for phakellistatin 2 (*cyclo*-[Phe-*cis*-Pro-Ile-Ile-*cis*-Pro-Tyr-*cis*-Pro and *cyclo*-[Phe-*trans*-Pro-Ile-Ile-*cis*-Pro-Tyr-*cis*-Pro]) have been separated by HPLC, each lasting for a

period of weeks in MeOH solution (ref 25b). Similarly, rolloamide B (*cyclo*-[Ser-Ile-*trans*-Pro-Leu-Phe-*cis*-Pro-Ile]) is a stable conformational variant of euryjanicin C (*cyclo*-[Ser-Ile-*cis*-Pro-Leu-Phe-*cis*-Pro-Ile]) (refs. 19 and 26b)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1.
Key HMBC (C→H) and ROESY correlations (dashed arrows) for euryjanicin E (**1**) and euryjanicin F (**2**).

Table 1
 ^1H NMR (500 MHz), ^{13}C NMR (125 MHz), TOCSY, HMBC, and ROESY correlations for euryjanicin E (1)^a

Amino acid	Position	δC (mult) ^b	δH , mult (J in Hz)	TOCSY	HMBC ^c	ROESY ^d
L-isoleucine ¹	CO	171.7, C			α	
	α	56.2, CH	4.13, t (9.2)	β, γ	β, γ	NHPhe ⁷ , β
	β	31.5, CH	2.02, m	α, γ, δ	α, γ, δ	$\alpha, \gamma, \varepsilon$
	γ	24.3, CH ₂	1.41, m 1.05, m	$\alpha, \beta, \delta, \varepsilon$	$\alpha, \beta, \delta, \varepsilon$	β
	δ	9.8, CH ₃	0.78, t (7.5)	$\beta, \gamma, \varepsilon$	$\beta, \gamma, \varepsilon$	γ
	ε	16.1, CH ₃	0.87, d (6.6)	γ, δ	δ	β
L-proline ² cis	NH		7.94, d (8.5)	α		$\alpha, \alpha\text{Phe}^3$
	CO	172.2, C			$\alpha, \beta, \text{NHlle}^1$	
	α	61.4, CH	3.71, d (7.9)	β, γ	β, γ	$\alpha\text{Phe}^3, \beta$
	β	31.3, CH ₂	2.02, m 1.04, m	α, γ, δ	α, γ, δ	α, γ
	γ	21.7, CH ₂	1.68, m 1.52, m	$\alpha, \beta, \gamma, \delta$	α, β, δ	β, δ
	δ	45.9, CH ₂	3.50, m 3.32, m	β, γ	β, γ	γ
L-phenylalanine ³	CO	171.6, C			α, β	
	α	52.8, CH	4.67, dd (7.8, 15.9)	β	β, γ	$\alpha\text{Pro}^2, \text{NHlle}^1, \beta$
	β	37.9, CH ₂	3.00, m	α	α, γ	α, γ
	γ	135.8, C			β, ortho	
	ortho	129.6, CH	7.21, m		γ, meta	meta
	meta	128.9, CH	7.32, m		ortho, para	ortho, para
L-asparagine ⁴	para	127.5, CH	7.28, m		meta, ortho	meta
	NH		7.46, d (6.0)	α		$\alpha, \alpha\text{Asn}^4$
	CO	172.1, C			$\beta, \text{NH}, \text{NHPhe}^3$	
	α	49.8, CH	4.38, m	β	β	NHPhe ³
	β	37.0, CH ₂	2.99, dd (5.8, 15.8) 2.84, dd (5.8, 15.8)	α	α	α
	γ	171.7, C				
γ NH			6.48, s			γ' NH
γ' NH			5.34, s			γ NH

Amino acid	Position	δC (mult) ^b	δH , mult (<i>J</i> in Hz)	TOCSY	HMBC ^c	ROESY ^d
L-isoleucine ⁵	NH		7.69, d (5.2)	α		α , NHlle ⁵
	CO	170.4, C			NH, NHAsn ⁴	
	α	61.2, CH	3.93, t (8.5)	β , γ	β , γ	β
	β	35.8, CH	1.96, m	α , γ , δ	α , γ , δ	α , γ , ϵ
	γ	25.6, CH ₂	1.48, m 1.12, m	α , β , δ , ϵ	α , β , δ , ϵ	β
	δ	10.8, CH ₃	0.81, t (7.4)	β , γ , ϵ	β , γ , ϵ	γ
L-proline ⁶ cis	ϵ	15.8, CH	0.90, d (6.8)	γ , δ	δ	β
	NH		8.18, d (8.2)	α		α , $\alpha\beta he^7$, NHAsn ⁴
	CO	170.9, C			α , β , NHlle	
	α	61.2, CH	3.76, t (7.6)	β , γ	β , γ	$\alpha Phe^7 \beta$
	β	30.4, CH ₂	2.13, m 0.85, m	α , γ , δ	α , γ , δ	α , γ
	γ	21.6, CH ₂	1.65, m 1.43, m	α , β , γ , δ	α , β , γ , δ	β , δ
L-phenylalanine ⁷	δ	46.2, CH ₂	3.51, m 3.36, m	β , γ	β , γ	γ
	CO	170.8, C			β	
	α	54.1, CH	4.53, m		β , γ	αPro^6 , β
	β	38.4, CH ₂	3.03, m		α , γ	α , γ
	γ	135.3, C			β , ortho	
	ortho	129.2, CH	7.23, m		γ , meta	meta
	meta	128.8, CH	7.32, m		ortho, para	ortho, para
	para	127.3, CH	7.28, m		meta, ortho	meta
	NH		7.50, d (3.1)	α		α , αlle^1

^a Spectra were recorded in CDCl₃ at 25 °C. Chemical shift values are in ppm relative to the residual CHCl₃ (7.26 ppm) or CDCl₃ (77.0 ppm) signals.

^b ¹³C NMR multiplicities were obtained from DEPTQ experiments.

^c HMBC mixing time = 50 ms.

^d ROESY mixing time = 200 ms.

Table 2
¹H NMR (500 MHz), ¹³C NMR (125 MHz), TOCSY, HMBC, and ROESY correlations for euryjanicin F (2)^a

Amino acid	Position	δC (mult) ^b	δH, mult (J in Hz)	TOCSY	HMBC ^c	ROESY ^d
L-phenylalanine ¹	CO	169.9, C			β, NH	
	α	55.5, CH	4.58, m		β, γ	APro ⁷ , β
	β	37.9, CH ₂	3.06, t (12.2) 2.79, dd (4.2, 12.1)		α, γ	α, γ
	γ	136.1, C			β, Ortho	
	ortho	129.9, CH	7.19, m	meta	γ, meta	meta
	meta	129.1, CH	7.20, m	ortho	ortho, para	ortho, para
L-phenylalanine ²	para	127.6, CH	7.22, m		meta, ortho	meta
	NH		10.16, s			αPhe ² , α, β
	CO	171.6, C			NHPhe ¹	
	α	56.7, CH	4.83, m		β, γ	β
	β	38.8, CH ₂	3.56, m 3.44, m		α, γ	α, γ
	γ	137.9, C			β, ortho	
L-isoleucine ³	ortho	130.9, CH	7.71, d (7.5)	meta	γ, meta	meta
	meta	129.0, CH	7.27, m	ortho	ortho, para	ortho, para
	Para	127.0, CH	7.22, m		meta, ortho	meta
	NH		8.42, s			αIle ³
	CO	171.5, C			α	
	α	61.4, CH	4.76, t (9.0)	β, γ	β, γ	β
L-proline ⁴ cis	β	37.7, CH	3.20, dd (7.8, 14.8)	α, γ, δ	α, γ, δ	α, γ, ε
	γ	25.6, CH ₂	1.67, m 1.27, m	α, β, δ, ε	α, β, δ, ε	β
	δ	10.9, CH ₃	0.90, t (7.3)	β, γ, ε	β, γ, ε	γ
	ε	16.0, CH ₃	1.11, d (6.7)	γ, δ	δ	β
	NH		8.43, d (2.6)			αPhe ⁵
	CO	170.4, C				
L-proline ⁴ cis	α	61.7, CH	3.98, d (7.0)	β, γ	β, γ	αPhe ⁵ , β
	β	29.0, CH ₂	2.30, dd (11.7, 6.5) 0.49, m	α, γ, δ	α, γ, δ	α, γ

Amino acid	Position	δC (mult) ^b	δH , mult (J in Hz)	TOCSY	HMBC ^c	ROESY ^d
L-phenylalanine ⁵	γ	21.9, CH ₂	1.56, m	$\alpha, \beta, \gamma, \delta$	α, β, δ	β, δ
	δ	45.9, CH ₂	3.77, m 3.41, m	β, γ	β, γ	γ
	CO	171.4, C				
	α	53.7, CH	4.94, m		β, γ, NH	$\alpha Pro^4, \beta$
	γ	38.4, CH ₂ 136.3, C	3.16, t (12.2) 2.94, dd (5.4, 12.7)		α, γ $\beta, ortho$	α, γ
	ortho	129.7, CH	6.86, d (6.4)	meta	$\gamma, meta$	meta
	meta	128.7, CH	7.27, m	ortho	ortho, para	ortho, para
	Para	127.7, CH	7.22, m		meta, ortho	meta
	NH		10.82, d (6.3)			$\alpha Ile^6, \alpha, \beta$
	CO	173.4, C			$\alpha, NHPh^5$	
L-isoleucine ⁶	α	56.8, CH	4.88, t (10.4)	β, γ	β, γ	β
	β	34.3, CH	2.54, m	α, γ, δ	α, γ, δ	α, γ, ϵ
	γ	25.0, CH ₂	1.50, m	$\alpha, \beta, \delta, \epsilon$	$\alpha, \beta, \delta, \epsilon$	β
	δ	9.3, CH ₃	0.68, t (7.4)	β, γ, ϵ	β, γ, ϵ	γ
	ϵ	15.8, CH ₃	1.03, d (6.7)	γ, δ	δ	$\alpha Phe^1, \alpha, \beta$
L-proline ⁷ cis	NH		9.78, d (9.3)			
	CO	171.1, C			$\alpha Ile^6, NH Ile^6$	
	α	62.0, CH	3.79, d (7.5)	β, γ	β, γ	$\alpha Phe^1 \beta$
	β	30.9, CH ₂	2.44, dd (5.9, 12.1) 0.97, m	α, γ, δ	α, γ, δ	α, γ
	γ	22.5, CH ₂	1.51, m	$\alpha, \beta, \gamma, \delta$	α, β, δ	β, δ
	δ	46.7, CH ₂	3.67, m	β, γ	β, γ	γ

^aSpectra were recorded in pyridine-*d*₅ at 25 °C. Chemical shift values are in ppm relative to the residual pyridine (8.71 ppm) or pyridine-*d*₅ (149.9 ppm) signals.

^b¹³C NMR multiplicities were obtained from DEPTQ experiments.

^cHMBC mixing time = 50 ms.

^dNOESY mixing time = 200 ms.

Table 3
¹H NMR (500 MHz), ¹³C NMR (125 MHz), TOCSY, HMBC, and ROESY correlations for euryjanicin G (3)^a

Amino acid	Position	CO	δC (mult) ^b	δH , mult (J in Hz)	TOCSY	HMBC ^c	ROESY ^d
L-isoleucine ¹			173.0, C				
	α		55.0, CH	4.96, t (9.3)		NH	α Pro ²
	β		38.3, CH	1.89, m	β , γ	β , γ	β
	γ		24.0, CH ₂	1.55, m 1.16, m	α , γ , δ	α , γ , δ	α , γ , ϵ
	δ		11.1, CH ₃	0.69, t (7.3)	α , β , δ , ϵ	α , β , δ , ϵ	β
	ϵ		15.3, CH ₃	0.83, d (6.8)	β , γ , ϵ	β , γ , ϵ	γ
L-proline ² cis	NH			9.79, d (9.1)	γ , δ	δ	α Phe ⁷ , β
	CO		171.4, C			α , NHPhe ³	
	α		62.3, CH	4.83, d (7.9)	β , γ	β , γ	α lle ¹
	β		32.3, CH ₂	2.23, m	α , γ , δ	α , γ , δ	α , γ
	γ		22.5, CH	2.00, m 1.69, m	α , β , γ , δ	α , β , δ	β , δ
	δ		47.0, CH ₂	3.79, m 3.53, m	β , γ	β , γ	γ
L-phenylalanine ³	CO		169.3, C			α , β , NH	
	α		53.2, CH	5.30, m			α Pro ⁴
	β		37.9, CH ₂	3.47, m 3.22, m		α , γ	α , γ
	γ		136.5, C			β , Ortho	
	ortho		129.7, CH	7.15, m	meta	γ , meta	meta
	meta		128.5, CH	7.17, m	ortho	ortho, para	ortho, para
L-proline ⁴ cis	para		127.2, CH	7.19, m		meta, ortho	meta
	NH			7.35, d (4.5)		β , NHPhe ⁵	
	CO		172.8, C				
	α		59.0, CH	4.69, d (8.0)	β , γ	β , γ	α Phe ³

Amino acid	Position	CO	δC (mult) ^b	δH , mult (J in Hz)	TOCSY	HMBC ^c	ROESY ^d
L-phenylalanine ⁵	β		31.2, CH ₂	2.11, m 1.88, m	α, γ, δ	α, γ, δ	α, γ
	γ		22.2, CH ₂	2.28, m 1.69, m	$\alpha, \beta, \gamma, \delta$	α, β, δ	β, δ
	δ		47.6, CH ₂	4.03, m	β, γ	β, γ	γ
				3.62, m			
	CO		170.4, C			α, β, NH	
	α		54.6, CH	5.07, m			αPro^6
	β		37.2, CH ₂	3.15, m		α, γ	α, γ
				3.03, m			
	γ		136.2, C			$\beta, ortho$	
	ortho		129.6, CH	7.22, m	meta	$\gamma, meta$	meta
	meta		128.4, CH	7.17, m	ortho	ortho, para	ortho, para
	para NH		126.4, CH	7.19, m 10.42, s		meta, ortho	meta $\alpha Pro^4, \alpha, \beta$
L-proline ⁶ cis	CO		172.0, C			$\alpha, NHPro^7$	
	α		61.3, CH	3.58, d (8.5)	β, γ	β, γ	αPhe^5
	β		31.1, CH ₂	2.38, m 1.13, m	α, γ, δ	α, γ, δ	α, γ
	γ		22.0, CH ₂	1.69, m 1.29, m	$\alpha, \beta, \gamma, \delta$	α, β, δ	β, δ
	δ		46.8, CH ₂	3.53, m 3.49, m	β, γ	β, γ	γ
	CO		171.9, C			α, β, NH	
L-phenylalanine ⁷	α		56.5, CH	4.99, m		NH	
	β		37.2, CH ₂	3.45, m		α, γ	α, γ
				3.17, m			
	γ		139.9, C			$\beta, ortho$	
	ortho		130.4, CH	7.07, m	meta	$\gamma, meta$	meta
	meta		129.1, CH	7.29, m	ortho	ortho, para	ortho, para
	para a		127.3, CH	7.19, m		meta, ortho	meta
	NH			9.73, d (8.3)			$\alpha Phe^5, \beta$

^aSpectra were recorded in pyridine-*d*₅ at 25 °C. Chemical shift values are in ppm relative to the residual pyridine (8.71 ppm) or pyridine-*d*₅ (149.9 ppm) signals.

^b¹³C NMR multiplicities were obtained from DEPTQ experiments.

$^p\text{NOESY}$ mixing time = 200 ms.
 $^c\text{HMBC}$ mixing time = 50 ms.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript