Isolation and structure of cyclosenegalins A and B, novel cyclopeptides from the seeds of *Annona senegalensis*

Alassane Wélé,[a] Yanjun Zhang,[a] Christelle Caux,[a] Jean-Paul Brouard,[a] Lionel Dubost,[a] Catherine Guette,[a] Jean-Louis Pousset,[a] Mamadou Badiane[b] and Bernard Bodo*[a]

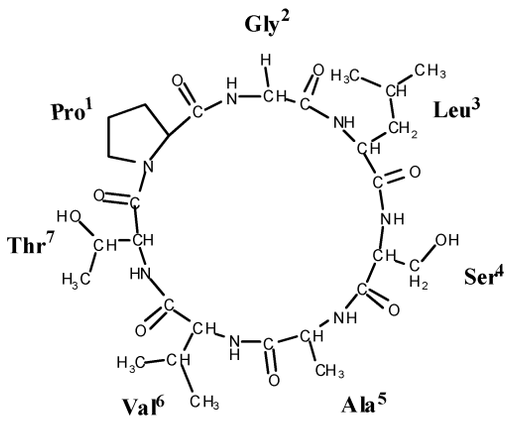
[a] *Laboratoire de Chimie des Substances Naturelles, ESA 8041 CNRS, Muséum National d'Histoire Naturelle, 63 rue Buffon, 75005 Paris, France*
[b] *Laboratoire de Chimie Organique et Thérapeutique, Faculté de Médecine et de Pharmacie, Université Cheikh Anta Diop Dakar-Fann, Sénégal*

Received (in Cambridge, UK) 23rd May 2002, Accepted 11th October 2002
First published as an Advance Article on the web 11th November 2002

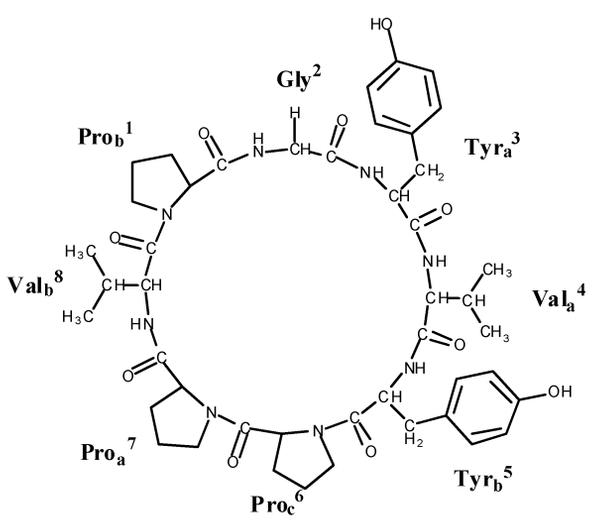
Two new cyclopeptides, cyclosenegalin A, cyclo(Pro[1]-Gly[2]-Leu[3]-Ser[4]-Ala[5]-Val[6]-Thr[7]-) (**1**) and cyclosenegalin B, cyclo(Pro[1]-Gly[2]-Tyr[3]-Val[4]-Tyr[5]-Pro[6]-Pro[7]-Val[8]-) (**2**), have been isolated from the methanol extract of the seeds of *Annona senegalensis* Pers., along with the known cyclic peptide, glabrin A. The structures were elucidated on the basis of the MS/MS fragmentation, using a Q-TOF mass spectrometer equipped with an ESI source, chemical degradation and extensive 2D-NMR.

Introduction

Annona senegalensis Pers. (Annonaceae) is a widespread small tree native in the woody savannah of Casamance, in the south of Senegal. There, traditional medicine uses this species, named "*digor*" by the natives, in the treatment of many diseases, especially as an antiseptic, healing substance and also against dermatosis and malaria fever.[1] In a previous study, we described the structural elucidation of cyclic peptides from the *Jatropha* species (Euphorbiaceae), some of which have antimalarial activity.[2,3] Continuing our investigation of cyclopeptides from plants, we report herein on the isolation from the seeds of *A. senegalensis* and the structural elucidation, based on tandem mass spectroscopy and 2D NMR, of two novel cyclic peptides, cyclosenegalins A (**1**) and B (**2**), along with the previously described cyclopeptide, glabrin A from *A. glabra*.[4,5] Previous phytochemical studies on *A. senegalensis* reported on the kauran diterpene derivatives obtained from the bark[6] and cytotoxic acetogenins obtained from the seeds.[7]



Cyclosenegalin A (1)



Cyclosenegalin B (2)

Results and discussion

1 Isolation of cyclopeptides

The seeds of *Annona senegalensis* were extracted with methanol and cyclosenegalins A (**1**), B (**2**) and glabrin A (**3**) were isolated from the ethyl acetate soluble fraction of this extract. They were purified successively by exclusion chromatography, silica gel column chromatography and C_{18} reversed-phase HPLC. A positive reaction with chlorine–*o*-tolidine reagent suggested they were peptides and the absence of coloration of the TLC spots with ninhydrin, that they were cyclic. Their total acidic hydrolysis and amino acid analysis of the hydrolysate after derivatization indicated the presence of Ala (1), Gly (1), Leu (1), Pro (1), Ser (1), Thr (1) and Val (1), for cyclosenegalin A (**1**), of Gly (1), Pro (3), Tyr (2) and Val (2) for cyclosenegalin B (**2**) and of Gly (1), Ile (1), Leu (1), Pro (1), Tyr (1) and Val (1) for glabrin A (**3**). The amino acids in the acidic hydrolysate were converted into the *n*-propyl esters of their *N*-trifluoroacetyl derivatives, analysed by gas chromatography on a chiral capillary column and their retention times compared with those of standards. All the chiral amino acids were L.

2 Sequence determination by mass spectrometry

The molecular weight 625 for cyclosenegalin A (**1**) was deduced from the positive MALDI-QTOF spectrum, which displayed

 DOI: 10.1039/b205035h

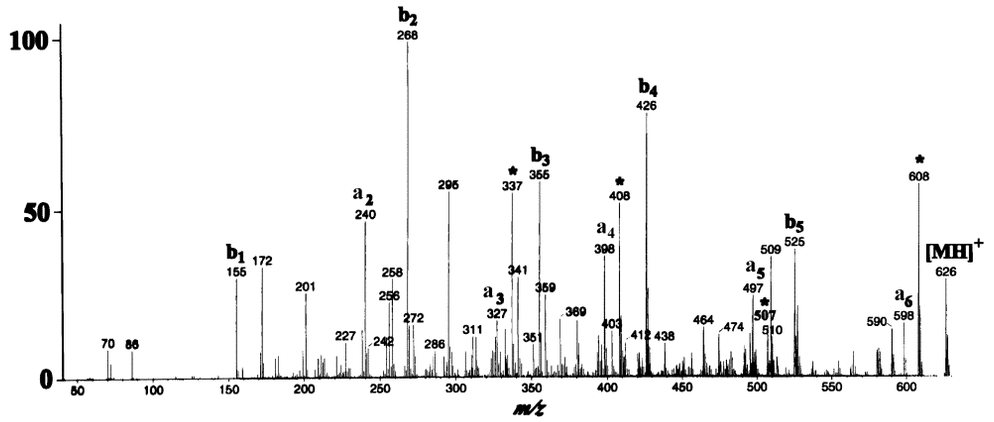


Fig. 1 CID mass spectrum of the $[M + H]^+$ ion (*m/z* = 626) of cyclosenegalin A (**1**) (* indicates the $b_n - H_2O$ fragment ions).

the $[M + Na]^+$ adduct ion at *m/z* 648 and the protonated molecular $[M + H]^+$ ion at *m/z* 626. According to the amino acid analysis, the molecular formula $C_{28}H_{47}N_7O_9$ was assigned to **1**. Cyclopeptides are not easily sequenced even by mass spectrometry. The reason is that multiple and indiscriminate ring-opening reactions occur during the CID of cyclic peptides, and in tandem mass spectra this results in the superimposition of random fragment ions, making the interpretation difficult.[8–10] However we have shown that when a proline is present in the sequence, a specific fragmentation occurs at the peptidyl-prolyl (Xaa-Pro) level, leading to a linear peptide C-ended by an acylium ion (b_n), which undergoes further fragmentation generating a series of acylium ions from which the sequence could be deduced.[2,11] This specific fragmentation is explained by the more basic nature of the proline nitrogen, relative to the other peptide bond nitrogen atoms. In this way, the more basic site at the proline level strongly directs the protonation making the fragmentation less complex.

The ESI-QTOF spectrum of **1** had the $[M + Na]^+$ adduct ion at *m/z* 648, and the protonated molecular $[M + H]^+$ ion at *m/z* 626. The protonated molecular ion $[M + H]^+$ of **1** at *m/z* 626 was subjected to CID experiments (Fig. 1). The ring opening began at the Thr-Pro amide bond level and a series of adjacent acylium ions (b_n) at *m/z* 525, 426, 355, 268 and 155 was generated from which the sequence could be deduced: amino acid residues were lost sequentially from the *C*-terminus to the *N*-terminus and for cyclopeptide **1** the successive loss of Thr, Val, Ala, Ser and Leu was observed, yielding the *N*-terminal dipeptide Pro-Gly (Fig. 2A). A second series of main peaks was observed at *m/z* 608, 507, 408 and 337, having 18 mass units less than the preceding b_n ions, corresponding to the loss of a water molecule. The absence of such ions associated with the b_n ions at 268 and 155 which included no Ser residue, suggested that this loss was due to the dehydration of the Ser residue. A third significant series of ions were observed at *m/z* 598, 497, 398, 327 and 240 which were assigned to adjacent a_n ions related to the above b_n ion series.

When analysing the fragmentation of the cationic adduct ion $[M + Na]^+$ at *m/z* 648, it was observed that the energy collision from fragmentation had to be higher and that the fragmentation differed from that of the protonated molecular ion. Then, three series of ions were observed: an abundant series of a_n ion fragments at *m/z* 519, 420, 349, 262 and 149, allowing the sequence to be deduced. Surprisingly the b_n series was not observed, but there was a $[b_n + Na - H]^+$ ion series at *m/z* 546, 447, 376 and 289 (Fig. 2B). Finally some abundant ions were observed at *m/z* 586, 489, 432, 319 and gave the tripeptide ion $[Ala\text{-}Val\text{-}Thr, Na]^+$ at *m/z* 232: these were assigned to y_n ion types leading to the partial sequence Pro-Gly-Leu-Ser. These observations for the sodiated peptide ion were corroborated by the analysis of the CID spectrum of $[M + Li]^+$ which had a similar spectrum to that of $[M + Na]^+$, with an ion series at *m/z*

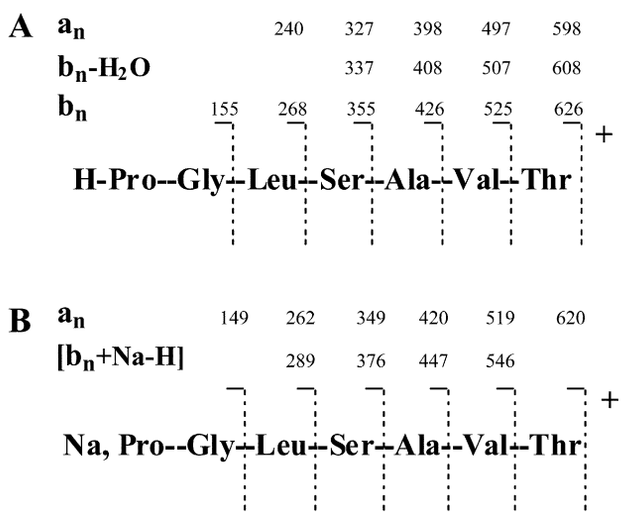


Fig. 2 MS/MS fragmentation of cyclopeptides with one proline. A) Protonated cyclosenegalin A (**1**) ion; B) sodiated cyclosenegalin A (**1**) ion.

133, 246, 333, 404, 503 and 604, a $[b_n + Li - H]^+$ series at *m/z* 273, 360, 431 and 530, and a short y_n series at *m/z* 303, 416, 473 and 570. These series are shifted by 16 mass units compared to the corresponding sodiated adduct ions, due to the mass difference between Na and Li.

The mass spectral results suggest the sequence $[H\text{-}Pro^1\text{-}Gly^2\text{-}Leu^3\text{-}Ser^4\text{-}Ala^5\text{-}Val^6\text{-}Thr^7]^+$ for the linearised peptide ion derived from cyclosenegalin A, and thus structure **1** for the natural cycloheptapeptide.

The molecular weight 872 for cyclosenegalin B, **2**, was deduced from the positive MALDI-QTOF spectra, which displayed the $[M + K]^+$ adduct ion at *m/z* 911, the $[M + Na]^+$ adduct ion at *m/z* 895, and the protonated molecular $[M + H]^+$ ion at *m/z* 873. According to the amino acid analysis the molecular formula $C_{45}H_{60}N_8O_{10}$ was assigned to **2**. Such ions were also observed in the ESI-QTOF spectrum which had the $[M + Na]^+$ adduct ion at *m/z* 895 and the protonated molecular $[M + H]^+$ ion at *m/z* 873. The CID spectrum of the $[M + H]^+$ ion at *m/z* 873 showed a main series of adjacent b_n peaks at *m/z* 774, 677, 580, 417, 318, and 155, corresponding to the successive loss of Val, Pro, Pro, Tyr, Val and Tyr yielding the terminal dipeptide ion $[H\text{-}Pro\text{-}Gly]^+$ and suggesting the sequence $H\text{-}Pro^1\text{-}Gly^2\text{-}Tyr^3\text{-}Val^4\text{-}Tyr^5\text{-}Pro^6\text{-}Pro^7\text{-}Val^8$ for the linearised peptide (Fig. 3A). A second series of ions with peaks at *m/z* 710, 611, 448, 391, 294, and 195 was assigned to a second b_n ion series, showing the successive loss of Tyr, Val, Tyr, Gly, Pro, Val and yielding the terminal dipeptide ion $[H\text{-}Pro\text{-}Pro]^+$, which indicates the sequence H-Pro-Pro-Val-Pro-Gly-Tyr-Val-Tyr (Fig. 3B). This sequence was confirmed by observation of the

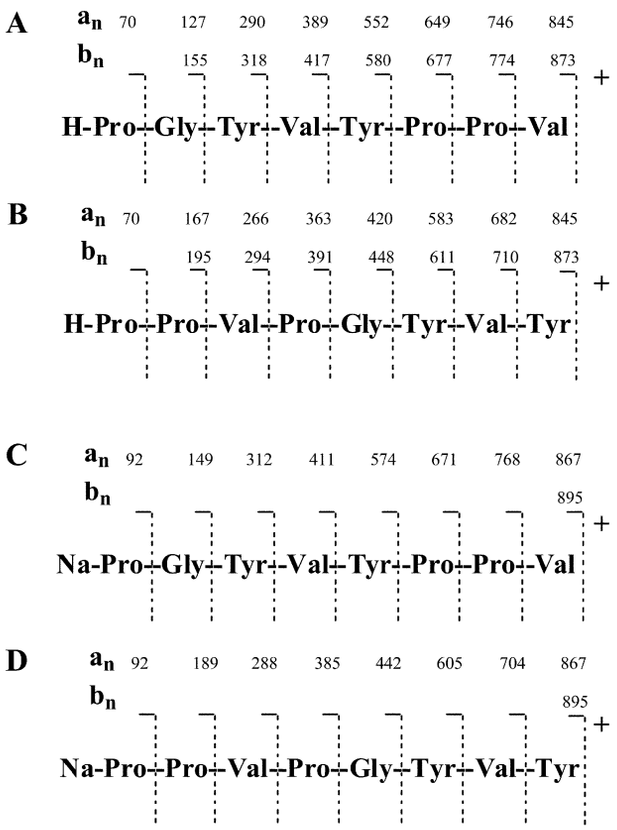


Fig. 3 MS/MS fragmentation $[M + H]^+$ and $[M + Na]^+$ of cyclosenegalin B (**2**) with three prolines. A) protonated ion, cleavage at Pro[1]; B) protonated ion, cleavage at Pro[6]; C) sodiated ion, cleavage at Pro[1]; D) sodiated ion, cleavage at Pro[6].

two corresponding a_n ion series. No significant ion fragmentation was observed due to the cleavage of the Pro-Pro amide bond.

When analysing the CID spectrum of the $[M + Na]^+$ ion at *m*/*z* 895, no related b_n ion series was observed, but two a_n ion series, corresponding to the sequences: [Na, Pro-Gly-Tyr-Val-Tyr-Pro-Pro-Val] and [Na, Pro-Pro-Val-Pro-Gly-Tyr-Val-Tyr] (Fig. 3C and 3D). All the data were in agreement with the linear sequence for cyclosenegalin B (**2**) indicated in Fig. 3. The cyclic peptide can be thus linearised by two different cleavages at two of the three possible sites before a proline. The amide bond between the vicinal prolines (-Pro[6]-Pro[7]-) appeared not to be cleaved significantly. Further fragmentation of the two linearised peptide ions allowed sequence determination of cyclosenegalin B as **2**.

The ESI-QTOF spectrum of cyclopeptide (**3**) revealed the $[M + Na]^+$ adduct ion at *m*/*z* 665 and the protonated molecular $[M + H]^+$ ion at *m*/*z* 643, and taking into account the amino acid composition, indicated $C_{33}H_{50}N_6O_7$ to be the molecular formula. The CID fragmentation of the $[M + H]^+$ ion showed the main fragments at *m*/*z* 480, 367, 268 and 155 corresponding to the successive loss of Tyr, Leu/Ile, Val, Leu/Ile to give the dipeptide fragment H-Pro-Gly. All the b_n ions were accompanied by the related a_n ions, 28 mass units lower. This hexapeptide cyclo(Pro-Gly-Leu-Val-Ile-Tyr-) (**3**) was identified as glabrin A, previously characterized in *A. glabra*.[4,5] The sequence and the relative position of Leu and Ile for **3** were determined by NMR.

3 ^{1}H- and ^{13}C-NMR studies

The ^{1}H-NMR spectrum of cyclosenegalin A (**1**) in DMSO-d_6 solution (Table 1) showed a main stable conformational state (>90%) where six amide protons were clearly depicted. As well, the presence of seven carbonyl groups in the ^{13}C-NMR spectrum are in agreement with a heptapeptide structure including a proline. The peptide sequence determination was based on the data of the HMBC experiment. This heteronuclear methodology was preferred, when possible, to the homonuclear method described by Wüthrich and co-workers[12,13] and based on $d_{NN(i,i+1)}$ and $d_{\alpha N(i,i+1)}$ connectivities from the ROESY/NOESY spectra, because for small size cyclic peptides, the obtention of conformational information can interfere with the obtention of sequential information. All the amino acid spin systems were identified using scalar spin–spin couplings determined from the ^{1}H–^{1}H COSY and TOCSY experiments.[14] The ^{13}C-NMR assignments of the protonated carbons were obtained from the proton-detected heteronuclear HSQC spectrum and combined with the HMBC experiment optimized for a long-range *J* value of 7 Hz for the non-protonated carbons. This experiment especially allowed the carbonyl groups to be assigned (Fig. 4). In this way, the sequence determination was done from the observation of the connectivities between the carbonyl of residue i and the amide and/or α protons of residue i + 1. All the $^2J_{CH}$, CO (i) to NH (i + 1) correlations, shown in Fig. 5, were depicted in the HMBC spectrum, in accord with the structure deduced from the mass fragmentations.

The NOESY spectrum clearly showed the $d_{NN(i,i+1)}$ interactions from Gly[2] to Thr[7] (Fig. 6). The NOE between Ser[4] and

Table 1 ^{13}C- and ^{1}H-NMR data for cyclosenegalin A (**1**) (DMSO-d_6, 318 K). ^{1}H chemical shifts are given to three or two decimal places when obtained from 1D or 2D spectra, respectively

Residue	δ_C	δ_H	Multiplicity	*J*/Hz
Pro[1] CO	171.2	—		
α CH	60.6	4.133	dd	8.9, 7.1
β CH_2	28.8	2.123	m	
	—	1.791	m	
γ CH_2	24.4	2.02	m	
	—	1.93	m	
δ CH_2	47.9	3.862	ddd	10.3, 7.5, 2.8
	—	3.649	ddd	10.3, 9.5, 6.5
Gly[2] CO	168.4	—		
NH	—	8.772	dd	7.9, 4.8
α CH_2	42.5	4.030	dd	17.0, 7.9
	—	3.300	dd	17.0, 4.8
Leu[3] CO	170.7	—		
NH	—	7.812	d	10.4
α CH	52.3	4.524	dd	10.4, 5.9
β CH_2	42.3	1.42	m	
	—	1.37	m	
γ CH	24.0	1.507	m	
δ CH_3	21.7	0.847	d	6.3
δ′ CH_3	22.3	0.863	d	6.3
Ser[4] CO	168.9	—		
NH	—	8.472	d	6.9
α CH	53.7	4.486	ddd	6.9, 3.0, 2.1
β CH_2	62.7	4.205	dd	10.7, 3.0
	—	3.720	dd	10.7, 2.1
Ala[5] CO	172.1	—		
NH	—	8.382	d	4.5
α CH	50.8	3.966	dd	4.5, 7.4
β CH_3	16.7	2.045	d	7.4
Val[6] CO	170.1	—		
NH	—	7.406	d	10.2
α CH	58.5	4.105	dd	10.2, 6.5
β CH	30.0	2.040	m	
γ CH_3	17.9	0.813	d	6.8
γ′ CH_3	19.0	0.822	d	6.8
Thr[7] CO	168.7	—		
NH	—	7.102	d	9.2
α CH	56.2	4.626	dd	9.2, 9.0
β CH	67.1	3.756	dd	9.0, 6.3
γ CH_3	19.6	1.092	d	6.3

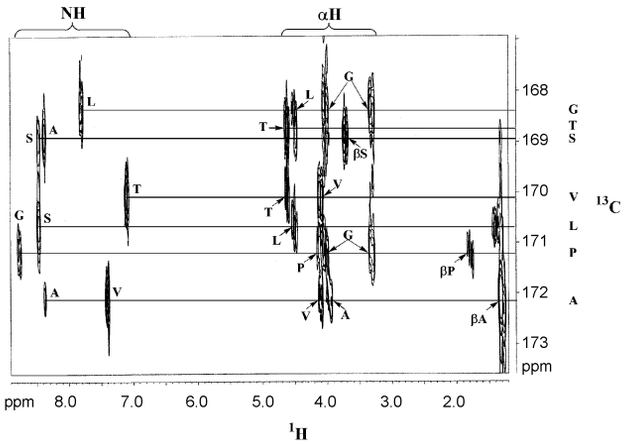


Fig. 4 Expansion of the HMBC spectrum of cyclosenegalin A (**1**) (DMSO d_6): regio for CO correlations (ω_1: 167.0 to 173.5 ppm, ω_2: 1.10 to 8.90 ppm).

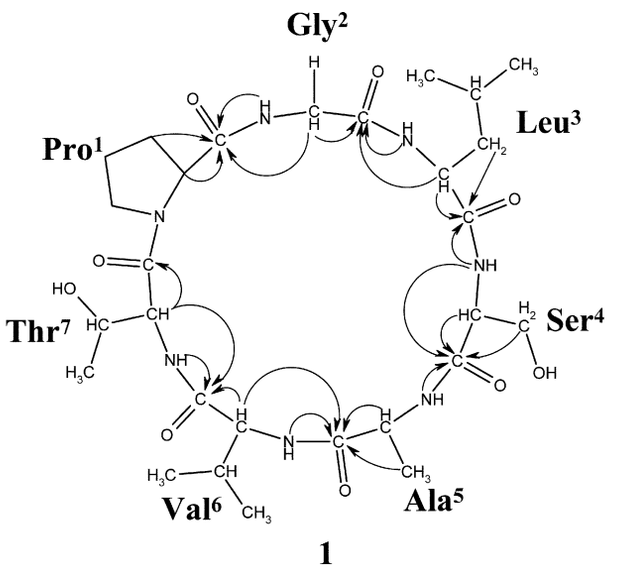


Fig. 5 Sequence of cyclosenegalin A (**1**): correlations from the HMBC spectrum.

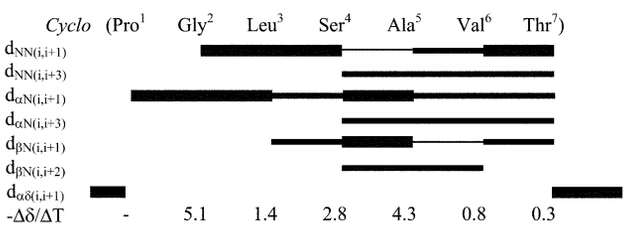


Fig. 6 Diagrammatic representation showing the pattern of strong, medium and weak connectivities involving the NH, α, β and δ protons, and the temperature coefficients of the amide protons (ppb K^{-1}) for cyclosenegalin A (**1**) in DMSO-d_6 solution.

Ala[5] was weak, and the NOE between the amide protons of Ser[4] and Thr[5] was not sequential, but explained by a conformational fold. Similarly a stretch of $d_{\alpha N(i,i+1)}$ sequential connectivities from Pro[1] to Thr[7] was depicted in agreement with the proposed sequence.

The α proton of Thr[7] gave strong NOE correlations with both δ and δ′ protons of Pro[1], indicating the Thr-Pro amide bond was *trans*. Chemical shifts of β and γ carbons of Pro were 28.8 and 24.4 ppm, respectively, giving further evidence of the *trans*-geometry of this amide bond.[15]

A β-turn of type II, with Pro[1] and Gly[2] at the two corners was suggested from the strong NOE correlation between the α-proton of Pro[1] and NH of Gly[2] which was in addition

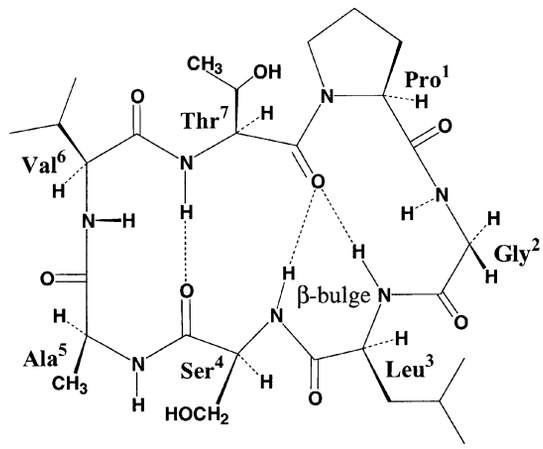


Fig. 7 Plausible solution conformation of cyclosenegalin A (**1**) proposed on the basis of NMR data (Hash lines showed transannular hydrogen bonds).

strongly correlated to the NH of Leu[3] (Fig. 7). A second β-turn involved Ala[5] and Val[6] at the two corners and is stabilised by a hydrogen bond between the NH of Thr[7] and the CO of Ser[4], in agreement with the strong NOE depicted between β-protons of Ser[4] and amide proton of Ala[5]. It is of type I, because NOEs were observed between the NH of Val[6] and both the NH of Thr[7], the β-methyl of Ala[5] and the β-proton of Val[6]. The significant NOEs observed between the NH of Thr[7] and Ser[4], Ser[4] and Leu[3] and finally Leu[3] and Gly[2], is explained by a β-bulge with a bifurcated hydrogen bond involving the CO of Thr[7] and the NH of Leu[3] and Ser[4] (Fig. 7). The pattern of the hydrogen bonding, as shown in Fig. 7, is in agreement with the thermal coefficients measured in DMSO-d_6 solution (Fig. 6) which indicates that the amide protons of Thr[7], Val[6] and Leu[3] are strongly involved in intramolecular hydrogen bonds, and that those of Gly[2] and Ala[5] are exposed to solvent.

All the data agreed with a cyclic structure for cyclosenegalin A (**1**), with a backbone conformation containing two β-turns, one of type I and the other of type II and incorporating a β-bulge. This structure seems to be a favorable motif for cyclic heptapeptides.[3,11]

The ^{1}H-NMR spectrum of cyclosenegalin B (**2**) in DMSO-d_6 solution (Table 2) showed a main stable conformational state (>90%) where the five amide protons were clearly depicted, one triplet (Gly) and four doublets of two Tyr (Tyr_a and Tyr_b) and two Val (Val_a and Val_b). The presence of two Tyr was evident, as characteristic signals for two *para*-disubstituted benzene rings were observed between 6.3 and 7.1 ppm. Assigment of protons and carbons to amino acid residues was achieved, as usual, from the COSY, TOCSY and HSQC data. Three prolines (Pro_a, Pro_b and Pro_c) were also depicted. Analysis of the long-range correlations of the eight carbonyls in the HMBC spectrum, allowed their assignment to definite residues, and also allowed complete sequence determination. The following correlations between the CO of residue (i) and the NH group of the residue (i − 1) were depicted: Pro_b[1] to Gly[2], Gly[2] to Tyr_a[3], Tyr_a[3] to Val_a[4], Val_a[4] to Tyr_b[5] and also Pro_a[7] to Val_b[8] (Fig. 8). Three peptide fragments with five, two and one residues, respectively, and interrupted by proline residues were thus defined: i) Pro_b[1]-Gly-Tyr_a-Val_a-Tyr_b[5], ii) Pro_a[7]-Val_b[8] and iii) Pro_c[6]. It is clear that they are connected to each other from the strong NOEs observed in the ROESY spectrum between the α proton of Tyr_b[5] and the α proton of Pro_c[6], and between the α proton of Val_b[8] and the δ protons of Pro_b[1] and between the α proton of Pro_c[6] and the δ protons of Pro_a[7] (Fig. 9). The complete cyclic sequence cyclo(Pro_b[1]-Gly[2]-Tyr_a[3]-Val_a[4]-Tyr_b[5]-Pro_c[6]-Pro_a[7]-Val_b[8]-) was thus defined. It was corroborated by two $d_{NN(i,i+1)}$ between Gly[2] and Tyr_a[3], and between Val_a[4] and Tyr_b[5] and by a stretch of $d_{\alpha N(i,i-1)}$ correlation from Pro_b[1] to Tyr_b[5] and between Val_b[8] and

Table 2 ^{13}C- and ^{1}H-NMR data for cyclosenegalin B (**2**) (DMSO-d_6, 298 K). ^{1}H chemical shifts are given to three or two decimal places when obtained from 1D or 2D spectra, respectively

Residue	δ_C	δ_H	Multiplicity	J/Hz
Pro^1 (b) CO	171.3	—		
α CH	58.2	3.119	d	8.1
β CH_2	28.5	1.964	dd	11.7, 8.1
	—	1.616	m	
γ CH_2	25.2	1.996	m	
	—	1.871	m	
δ CH_2	46.6	3.719	m	
	—	3.437	m	
Gly^2 CO	168.1	—		
NH	—	8.643	dd	9.4, 3.3
α CH_2	42.2	3.982	dd	17.0, 9.4
	—	3.052	dd	17.0, 3.3
Tyr^3 (a) CO	170.3	—		
NH	—	8.432	d	9.9
α CH	53.7	4.483	ddd	9.9, 3.8, 3.8
β CH_2	34.7	2.874	d	3.8
	—	2.874	d	3.8
1′	129.8	—		
2′,6′	130.4	7.021	m	8.5
3′,5′	114.2	6.520	m	8.5
4′	155.1	—		
OH	—	8.932	s	
Val^4 (a) CO	170.6	—		
NH	—	8.041	d	9.4
α CH	60.7	3.828	dd	9.8, 9.4
β CH	29.5	1.874	m	
γ CH_3	19.5	0.834	d	6.6
γ′ CH_3	19.3	0.720	d	6.7
Tyr^5 (b) CO	169.9	—		
NH	—	7.225	d	8.1
α CH	54.1	4.292	dd	8.1, 7.6
β CH_2	34.9	2.707	d	7.6
	—	2.707	d	7.6
1′	127.0	—		
2′,6′	130.0	6.960	m	8.5
3′,5′	115.0	6.650	m	8.5
4′	156.0	—		
OH	—	9.210	s	
Pro^6 (c) CO	168.9	—		
α CH	61.9	4.053	m	
β CH_2	30.1	1.90	m	
	—	1.64	m	
γ CH_2	21.6	1.536	m	
	—	1.536	m	
δ CH_2	46.0	3.237	m	
	—	3.103	m	
Pro^7 (a) CO	170.3	—		
α CH	59.7	4.43	d	8.1
β CH_2	25.0	2.00	dd	11.7, 8.1
	—	1.85	m	
γ CH_2	25.2	1.184	m	
	—	1.723	m	
δ CH_2	46.4	3.36	m	
	—	3.36	m	
Val^8 (b) CO	170.7	—		
NH	—	7.834	d	9.6
α CH	55.1	3.43	m	
β CH	28.9	2.018	m	
γ CH_3	19.3	0.724	d	6.6
γ′ CH_3	16.7	0.557	d	6.8

Pro_a^7. The NOE interaction between the α protons of Tyr_b^5 and Pro_c^6 indicated the amide Tyr_b^5-Pro_c^6 bond to be *cis*, whereas the NOEs between the α proton of Val_b^8 and the δ protons of Pro_b^1, as well as the NOEs between the α proton of Pro_c^6 and the δ protons of Pro_a^7 indicated that the two Val_b^8-Pro_b^1 and

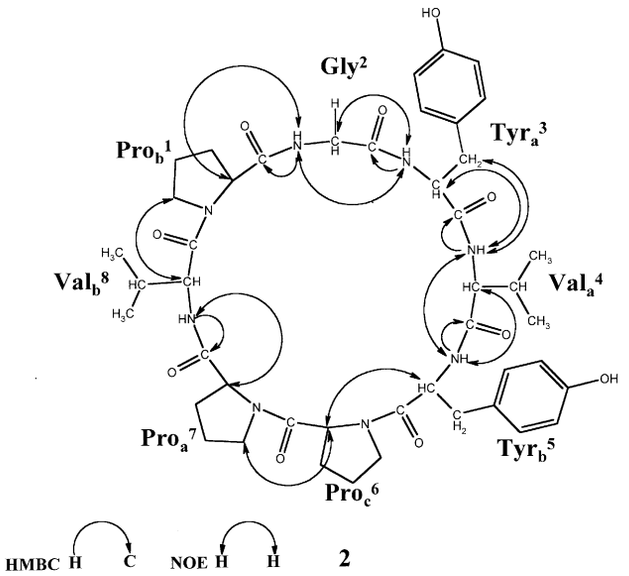


Fig. 8 HMBC and ROESY ($d_{NN(i,i+1)}$ and $d_{\alpha N(i,i+1)}$) correlations for cyclosenegalin B (**2**).

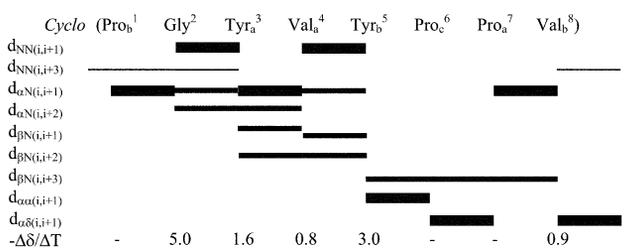


Fig. 9 Diagrammatic representation showing the pattern of strong, medium and weak connectivities involving the NH, α, β and δ protons, and the temperature coefficients of the amide protons (ppb K^{-1}) for cyclosenegalin B (**2**) in DMSO-d_6 solution.

Pro_c^6-Pro_a^7 amide bonds were *trans*. The chemical shift of the proline γ carbon is expected at 21–22 ppm for *cis* prolines and 24–26 ppm for *trans* ones.[15] The observed ^{13}C shift values for Pro_b^1, Pro_c^6 and Pro_a^7, at 25.2, 21.6 and 25.2, respectively are in concordance with the proposed amide bond stereochemistry (Table 2). The sequence of cyclosenegalin B (**2**) was determined to be cyclo(Pro^1-Gly^2-Tyr^3-Val^4-Tyr^5-Pro^6-Pro^7-Val^8-). The determination of the amide proton chemical shift dependence indicated that the NH protons of Val^4, Val^8 and Tyr^3 were engaged in strong intramolecular hydrogen bonding, and that the NH proton of Gly^2 was exposed to solvent (Fig. 9).

It is remarkable that the peptides isolated from the Annonaceae family, hexa- to nonapeptides include at least one proline residue.[16–19] As usual, this residue induces conformational constraints in cyclic peptides of such a small size. In addition, its presence is very useful as it induces one linearised peptide (or a few linearised peptides if there are several prolines) from which the mass fragmentation allows the sequence to be unambiguously determined by mass spectrometry. Mass spectrometry was also used to locate cyclopeptides either in the slightly oily endosperm, or in the ruminated integument of the seeds. In the MALDI-TOF MS of the two MeOH extracts the characteristic ions for the $[M + H]^+$ and $[M + Na]^+$ of **1**, **2** and **3** were observed only in the endosperm part of the seeds and not in the integument. This location, together with the absence of antibiotic–antifungal and cytotoxic activity, suggest that these peptides do not protect against predators or parasites.

Experimental

Optical rotations were measured with a Perkin-Elmer model 341 Polarimeter and the $[\alpha]_D^{22}$ values are given in deg cm^2 g^{-1}.

Melting points were determined on a Büchi melting point B-545 apparatus. ^{1}H- and ^{13}C-NMR spectra were recorded either (1D ^{13}C) with a Bruker AC 300 spectrometer, equipped with an Aspect 3000 computer using DISNMR software or (2D spectra) with a Bruker Avance 400 spectrometer operating at 400.13 MHz. The coupling constant used to establish the necessary delay for the selection of the proton coupled to the carbon in the HSQC spectrum was 135 Hz, corresponding to a delay of 3.7 ms; the delay for the HMBC spectra was 70 ms corresponding to a long-range coupling constant of 7 Hz. The phase sensitive ROESY experiments were obtained with a mixing time of 150 ms. Mass spectra were recorded on a MALDI-TOF or an API Q-STAR PULSAR i of Applied Biosystem. For the CID spectra, the collision energy (CE) was 35 to 90 eV and the collision gas was nitrogen.

Plant material

Seeds of *Annona senegalensis* Pers. (Annonaceae) were collected in Casamance (Senegal) in November 2000. Samples were immediately washed with distilled water and were dried at room temperature. A voucher has been deposited at the National Museum of Natural History (Paris).

Extraction and isolation

The dried and powdered seeds of *Annona senegalensis* (4.2 kg) were macerated three times with cyclohexane (3 l), and the combined extracts yielded an oil which was discarded. The seeds were then extracted three times with MeOH (3 l) at room temperature to give after evaporation of the solvent under reduced pressure the MeOH extract (296 g) which was partitioned between EtOAc and water. The organic phase was concentrated to dryness and the residue (92 g) was dissolved in MeOH and chromatographed on a Sephadex LH-20 column with MeOH. The peptide fraction (12.15 g) was then repeatedly subjected to silica gel column chromatography (Kieselgel 60 H Merck) and eluted with CH_2Cl_2 containing increasing amounts of MeOH from 5% to 20%. Peptide purification was monitored by TLC (silica gel 60 F_{254} Merck) with CH_2Cl_2–MeOH (9 : 1) as the eluent system and the peptides were detected with Cl_2–*o*-toluidine reagent: the peptides exhibited two blue spots with R_f 0.26 (glabrin A) and R_f 0.30 (compounds **1** and **2**) which were separated by SiO_2 column chromatography yielding two peptide mixtures I (R_f 0.26) and II (R_f 0.30). The peptide mixtures were finally purified by isocratic reversed phase HPLC (Kromasil C_{18}, 250 × 7.8 mm, 5 µm, AIT France; flow rate 2 ml min^{-1}, detection 220 nm). Mixture I, using MeOH–H_2O (65 : 35) with 1% TFA, yielded glabrin A (**3**, t_R 9.9 min, 151.5 mg); mixture II, using MeOH–H_2O (55 : 45) with 1% TFA, yielded cyclosenegalin A (**1**, t_R 9.3 min, 26.6 mg) and cyclosenegalin B (**2**, t_R 13,1 min, 55.8 mg).

Absolute configuration of amino acids

Solutions of **1** and **2** (each containing 1 mg of peptide) in 6 M HCl (1 ml) were heated at 110 °C for 24 hours in sealed tubes. After cooling, each solution was concentrated to dryness. The hydrolysates were dissolved in anhydrous solution of 3 M HCl in propan-2-ol and heated at 110 °C for 30 min. The reagents were evaporated under reduced pressure. The residues were dissolved in CH_2Cl_2 (0.5 ml) and 0.5 ml trifluoracetic acid was added. The mixtures were kept in a screw-capped tube at 110 °C for 20 min. The reagents were evaporated and the mixtures analysed on a Chirasil-*L*-Val (*N*-propionyl-L-valine-*tert*-butylamide polysiloxane) quartz capillary column with helium (1.1 bar) as carrier gas and a temperature program of 50–130 °C at 3 °C min^{-1} and 130–190 °C at 10 °C min^{-1}, with a HEWLETT PACKARD series 5890 apparatus. Comparison of t_R values with those of standards amino acids was used: L-Ala (11.6), L-Val (13.9), Gly (14.6), L-Thr (15.2), L-Pro (18.2), L-Ser (18.8), L-Leu (19.2) and L-Tyr (31.9).

Cyclosenegalin A (1). Colourless amorphous solid, $[\alpha]_D^{22}$ −2.6 (*c* 0.7, MeOH). ^{1}H- and ^{13}C-NMR, see Table 1. ESI-QTOF *m/z*: 664 $[M + K]^+$, 648 $[M + Na]^+$, 626 $[M + H]^+$. ESI-QTOF MS/MS on *m/z* 626 $[M + H]^+$ (CE 39 eV) *m/z* (%): 626 (50), 608 (56), 598 (14), 581 (12), 525 (37), 509 (35), 507 (16), 497 (26), 474 (14), 464 (16), 438 (11), 426 (83), 408 (52), 398 (38), 380 (18), 369 (17), 359 (25), 355 (58), 341 (34), 337 (54), 327 (17), 295 (51), 268 (100), 258 (34), 256 (25), 240 (55), 227 (11), 201 (27), 172 (33), 155 (40), 86 (5), 70 (4). ESI-QTOF MS/MS on $[M + Na]^+$ (CE 70 eV) *m/z* (%): 648 (12), 630 (14), 620 (15), 604 (66), 586 (100), 576 (68), 574 (44), 558 (63), 546 (48), 530 (25), 519 (28), 515 (27), 505 (34), 501 (19), 489 (53), 487 (33), 477 (27), 463 (25), 459 (27), 447 (47), 445 (35), 433 (30), 432 (26), 420 (59), 418 (51), 406 (39), 402 (33), 390 (30), 376 (59), 360 (21), 349 (63), 347 (15), 335 (14), 319 (70), 305 (31), 289 (7), 266 (14), 262 (49), 252 (18), 220 (14), 206 (10), 177 (9), 155 (1) 149 (1), 70 (3). ESI-QTOF MS/MS on $[M + Li]^+$ (CE 61 eV) *m/z* (%): 632 (37), 614 (28), 604 (13), 588 (87), 570 (100), 560 (41), 543 (18), 542 (49), 530 (21), 503 (15), 489 (16), 473 (23), 457 (15), 431 (14), 429 (30), 416 (13), 404 (24), 386 (21), 360 (26), 345 (19), 333 (29), 303 (28), 289 (24), 273 (5), 246 (32), 236 (11), 177 (12), 176 (2), 161 (2), 133 (6).

Cyclosenegalin B (2). Colourless solid, mp 246–247 °C (MeOH), $[\alpha]_D^{22}$ −5.7° (*c* 0.4, MeOH). ^{1}H- and ^{13}C-NMR, see Table 2. ESI-QTOF, *m/z*: 910 $[M + K]^+$, 895 $[M + Na]^+$, 873 $[M + H]^+$. ESI-QTOF MS/MS on $[M + H]^+$ (CE 50 eV) *m/z* (%): 873 (46), 845 (23), 774 (7), 746 (2), 710 (18), 682 (9), 677 (10), 649 (4), 611 (18), 583 (10), 580 (57), 552 (43), 535 (8), 515 (7), 481 (9), 457 (26), 448 (7), 429 (14), 420 (1), 417 (58), 391 (4), 389 (39), 363 (1), 358 (14), 318 (68), 294 (100), 266 (24), 233 (11), 212 (7), 195 (6), 167 (1), 155 (7), 127 (3), 72 (22), 70 (6). ESI-QTOF MS/MS on $[M + Na]^+$ (CE 90 eV) *m/z* (%): 895 (24), 867 (51), 796 (1), 768 (14), 746 (8), 704 (37), 699 (1), 690 (3), 661 (16), 605 (100), 574 (48), 550 (37), 508 (20), 451 (46), 442 (35), 439 (5), 411 (70), 409 (5), 385 (16), 345 (24), 312 (50), 288 (69), 189 (4), 149 (1), 92 (1), 70 (4).

Glabrin A (3). Colourless microcrystals, mp 225–226 °C (MeOH), $[\alpha]_D^{22}$ −17.0 (*c* 0.9, MeOH). ESI-QTOF, *m/z*: 681 $[M + K]^+$, 665 $[M + Na]^+$, 643 $[M + H]^+$. ESI-QTOF MS/MS on $[M + H]^+$ (CE 40 eV) *m/z* (%): 643 (41), 615 (36), 598 (12), 544 (10), 530 (33), 502 (22), 485(17), 480 (45), 452 (17), 431 (18), 403 (17), 386 (12), 367 (100), 339 (56), 318 (14), 268 (52), 240 (10), 261 (6), 213 (4), 172 (5), 155 (33), 136 (5), 127 (3).

Acknowledgements

The French "Ministère de la Coopération" (EGIDE) is gratefully acknowledged for granting a fellowship to one of us (AW), and the "Région Ile-de-France" for its generous contribution to the funding of the 400 MHz NMR and the ESI-TOF mass spectrometers.

References

1 J. Kerharo and J. G. Adam, in *La pharmacopée sénégalaise traditionnelle. Plantes médicinales et toxiques*, ed. Vigot Frères, Paris, 1974, p. 147.
2 C. Baraguey, C. Auvin-Guette, A. Blond, F. Cavelier, F. Lezenven, J.-L. Pousset and B. Bodo, *J. Chem. Soc., Perkin Trans. 1*, 1998, 3033.
3 C. Baraguey, A. Blond, F. Cavelier, J.-L. Pousset, B. Bodo and C. Auvin-Guette, *J. Chem. Soc., Perkin Trans. 1*, 2001, 2098.
4 C.-M. Li, N.-H. Tan, Q. Mu, H.-L. Zheng, X.-J. Hao, H.-L. Liang and J. Zhou, *Phytochemistry*, 1998, **47**, 1293.
5 C.-M. Li, N.-H. Tan, H.-L. Zheng, Q. Mu, X.-J. Hao, Y.N. He and J. Zhou, *Phytochemistry*, 1999, **50**, 1047.

6 I. T. U Eshiet, A. Akisanya and D. A. H. Taylor, *Phytochemistry*, 1971, **10**, 3294.
7 S. Sevser, C. Gonzalez, R. Hocquemiller, M. C. Zafra-Polo and D. Cortes, *Phytochemistry*, 1996, **42**, 103.
8 S. Lin, G. Liehr, B. S. Cooperman and R. J. Cotter, *J. Mass Spectrom.*, 2001, **36**, 658.
9 K. B. Tomer, F. W. Crow, M. L. Gross and K. D. Kopple, *Anal. Chem.*, 1984, **56**, 880.
10 M. Yuan, M. Namikoshi, A. Otsuki, K. L. Rinehart, K. Sivonen and M. F. Watanabe, *J. Mass Spectrom.*, 1999, **34**, 33.
11 C. Auvin-Guette, C. Baraguey, A. Blond, S. X. Xavier, J.-L. Pousset and B. Bodo, *Tetrahedron*, 1999, **55**, 11495.
12 K. Wüthrich, G. Wider, G. Wagner and W. Braun, *J. Mol. Biol.*, 1982, **155**, 311.
13 K. Wüthrich, M. Billeter and W. Braun, *J. Mol. Biol.*, 1984, **180**, 715.
14 G. Wagner, A. Kumar and K. Wüthrich, *Eur. J. Biochem.*, 1981, **114**, 375.
15 D. E. Douglas and F. A. Bovey, *J. Org. Chem.*, 1973, **38**, 2379.
16 C.-M. Li, N.-H. Tan, Y.-P. Lu, H.-L. Liang, Q. Mu, H.-L. Zheng, X.-J. Hao, Y. Wu and J. Zhou, *Acta Bot. Yunnanica*, 1995, **17**, 459.
17 C.-M. Li, N.-H. Tan, H.-L. Zheng, Q. Mu, X.-J. Hao, Y. N. He and J. Zhou, *Phytochemistry*, 1998, **48**, 555.
18 H. Morita, Y. Sato and J. Kobayashi, *Tetrahedron*, 1999, **55**, 7509.
19 C.-M. Li, N.-H. Tan, Q. Mu, H.-L. Zheng, X.-J. Hao, Y. Wu and J. Zhou, *Phytochemistry*, 1997, **45**, 521.