Sequence and solution structure of cherimolacyclopeptides A and B, novel cyclooctapeptides from the seeds of *Annona cherimola*

Alassane Wélé,[a] Céline Landon,[b] Henri Labbé,[b,†] Françoise Vovelle,[b] Yanjun Zhang[a] and Bernard Bodo[a,*]

[a]*Laboratoire de Chimie des Substances Naturelles, UMR 5154 CNRS, Muséum National d'Histoire Naturelle, 63 rue Buffon, 75005 Paris, France*
[b]*Centre de Biophysique Moléculaire, UPR 4301 CNRS, affiliated to Orléans university, rue Charles Sadron, 45071 Orléans Cedex 02, France*

Abstract—Two new cyclooctapeptides, cherimolacyclopeptide A, *cyclo*(Pro^1-Gln^2-Thr^3-Gly^4-Met^5-Leu^6-Pro^7-Ile^8-) (**1**) and the related cherimolacyclopeptide B, *cyclo*(Pro^1-Gln^2-Thr^3-Gly^4-Mso^5-Leu^6-Pro^7-Ile^8-) (**2**), have been isolated from the methanol extract of the seeds of *Annona cherimola* Miller. The sequences 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-heteronuclear NMR. The three-dimensional solution structure of cherimolacyclopeptide A (**1**) determined by ^{1}H NMR data and molecular modelling is characterised by the presence of two β turns and a new type of β-bulge.
© 2003 Elsevier Ltd. All rights reserved.

1. Introduction

Annona cherimola Miller (Annonaceae) is a small tree native to Ecuador and Peru, now widespread in subtropical areas, America, Africa and Asia and even in the south of Europa where it is cultivated for its edible fruits. The plant is also used in folk medicine as a parasiticide and an insecticide. *A. cherimola* has been described to produce many natural compounds of biological interest, such as alkaloids,[1,2] mainly of the isoquinoline group for which no less than 70 compounds have been isolated, acetogenins,[3–6] and dimeric amides.[7,8]

In continuation of our programme on cyclopeptides from plants,[9–11] we have investigated *A. cherimola*, and isolated from the seeds two new cyclooctapeptides, cherimolacyclopeptides A (**1**) and B (**2**). Although several cyclic octapeptides have been isolated from various natural sources, such as agardhipeptin B[12] from cyanobacteria, hymenistatin I[13] and axinellin C[14] from marine sponges, pseudostellarin H,[15] cyclolinopeptides D and E,[16] cyclogossin B,[9] pohlianin C[10] and cyclosquamosins B–D[17] from plants, only few tri-dimensional studies are devoted to such peptides. The conformations of cyclic peptides are of interest by themselves, as they can be used as models for the study of recurring structural features of proteins. The very limited number of 3D studies on cyclooctapeptides concern mainly synthetic peptides involving L-amino acids,[18] and mostly structurally symmetrical compounds.[19,20]

In this paper we report on the isolation, the sequence determination based on tandem mass spectroscopy and 2D NMR of cherimolacyclopeptides A (**1**) and B (**2**), and on the solution structure of **1**. This conformation including two β-turns and a new type of β-bulge is compared to that of other cyclopeptides. Cherimolacyclopeptide A (**1**) was found to be cytotoxic against tumoral KB cells, with an IC$_{50}$ 0.6 μM, whereas cherimolacyclopeptide B (**2**) was less active with IC$_{50}$ 45 μM.

2. Results and discussion

2.1. Isolation of cherimolacyclopeptides A and B

The dried and ground seeds of *A. cherimola* were extracted with methanol and cherimolacyclopeptides A (**1**) and B (**2**) 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. Positive reaction with chlorine/ *o*-tolidine reagent suggested that they were peptides and the absence of coloration with ninhydrin of the spots on thin

* Corresponding author. Tel.: +33-1-40-79-3129; fax: +33-1-40-79-3135;
  e-mail: bodo@mnhn.fr
† Deceased on August 5, 2003.

layer chromatography (TLC), that these peptides were cyclic. The total acidic hydrolysis and amino acid analysis of the hydrolysate after derivatization indicated the presence of Glx (1), Gly (1), Ile (1), Leu (1), Met (1), Pro (2) and Thr (1), for cherimolacyclopeptide A (**1**) and Glx (1), Gly (1), Ile (1), Leu (1), Mso (1), Pro (2) and Thr (1), for cherimolacyclopeptide B (**2**). The amino acids in the acidic hydrolysate were converted into the *n*-propyl esters of their *N*-trifluoroacetyl derivatives. These esters were analysed by gas chromatography on a chiral capillary column and their retention times were compared with those of standards. All the chiral amino acids were L. The Glx in the hydrolysate was further identified as Gln from the absence of carboxylic acid group and from ^{1}H NMR data for both peptides, **1** and **2**.

2.2. Sequence determination by mass spectrometry

The molecular weight 837 for cherimolacyclopeptide A (**1**) was deduced from the positive ESI-qTOF mass spectrum, which displayed the protonated molecular $[M+H]^+$ ion at *m/z* 838, and the $[M+Na]^+$ and $[M+K]^+$ adduct ions at *m/z* 860 and 876, respectively. According to the amino acid analysis, the molecular formula $C_{38}H_{62}N_9O_{10}S$ was assigned to **1**, in agreement with the presence of eight residues in the cyclopeptide. Similarly, the molecular formula $C_{38}H_{63}N_9O_{11}S$ was assigned to **2**, taking into account the protonated molecular $[M+H]^+$ ion observed at *m/z* 854 in the ESI-qTOF spectrum, together with the $[M+K]^+$ adduct ion at *m/z* 992 and the $[M+Na]^+$ adduct ion at *m/z* 876 and its amino acid composition: **2** differs from **1** by the replacement of the Met residue by a methionine sulfoxyde (Mso).

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. This results in the superimposition of random fragment ions, making the interpretation difficult.[21−23] However, due to the presence of a proline in the sequence, a specific fragmentation occurs at the peptidyl-prolyl (Xaa-Pro) level, leading to one linear peptide C-ended by an acylium ion (b_n), which undergoes further fragmentation and generates a series of acylium ions from which the sequence could be deduced.[10]

The collisional induced decomposition (CID) experiment on the protonated molecular $[M+H]^+$ ion of **1** at *m/z* 838 allowed the sequence determination (Fig. 1). The ring opening occurred at the Ile-Pro amide bond level and a series of adjacent acylium ions (b_n) at *m/z* 725, 628, 515, 384, 327 and 226 was generated: amino acids residues were lost sequentially from the *C*- to the *N*-terminus of the linearised peptide derived from **1**. The successive loss of Ile/Leu, Pro, Leu/Ile, Met, Gly and Thr was observed, yielding to the *N*-terminal dipeptide Pro-Gln (Fig. 2A). A second series of main peaks was observed at *m/z* 820, 707, 610, 497, 366 and 198, having 18 mass units less than the preceding b_n ions, and corresponding to the loss of a water molecule at the Thr level. A third significant series of ions was observed at *m/z* 810, 697, 600, 487, 356, 299 and 198 which were assigned to adjacent a_n ions related to the above b_n ion series. The mass spectral results suggested the sequence [H-Pro^1-Gln^2-Thr^3-Gly^4-Met^5-Leu/Ile^6-Pro^7-Ile/Leu^8]$^+$ for the
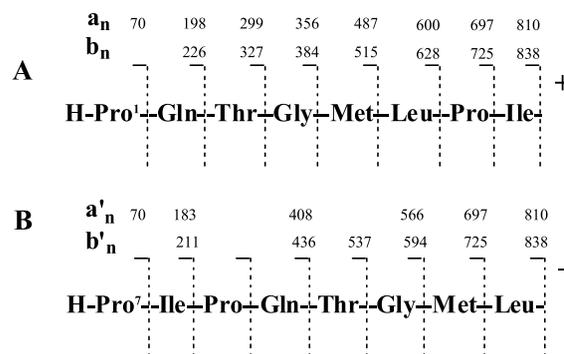
Figure 1. MS/MS fragmentation of the cherimolacyclopeptide A (**1**) $[M+H]^+$ ion (*m/z*=838): **A)** cleavage of the cyclopeptide at the Ile^8-Pro^1 amide bond level; **B)** cleavage of the cyclopeptide at the Leu^6-Pro^7 amide bond level.

linear peptide ion derived from cherimolacyclopeptide A, and thus the structure *cyclo*(Pro^1-Gln^2-Thr^3-Gly^4-Met^5-Leu^6-Pro^7-Ile^8) for the natural cyclooctapeptide **1**, with an ambiguity on the respective position of Leu and Ile.

A second linearised peptide was formed from the $[M+H]^+$ ion at *m/z* 838, due to the cleavage at the Leu^6-Pro^7 amide bond level and the resulting b'n and a'n ions were detected (Fig. 1B). The b'n ions series was characterized by ions at *m/z* 725, 594, 537, 436 and 211 corresponding to the successive loss of Leu/Ile, Met, Gly, Thr, (Pro-Gln) and yielding to the *N*-terminal dipeptide Pro-Ile/Leu, in agreement with the above proposed sequence.

Similarly, the CID spectrum of the $[M+H]^+$ ion at *m/z* 854 of cherimolacyclopeptide B (**2**) showed a main series of adjacent b_n peaks at *m/z* 741, 644, 531, 384, 327 and 226, corresponding to the successive loss of Ile/Leu, Pro, Leu/Ile, Mso, Gly and Thr, yielding the terminal dipeptide ion [H-Pro-Gln]$^+$ and suggesting the sequence H-Pro^1-Gln^2-Thr^3-Gly^4-Mso^5-Leu/Ile^6-Pro^7-Ile/Leu^8 for the linearised peptide. A second series of b_n type ions, at *m/z* 741, 594, 537, 436, 308 and 211, due to a cleavage at the Leu^6-Pro^7 amide bond level, indicated the successive loss of Leu, Mso, Gly, Thr, Gln, Pro and yielding the terminal dipeptide ion [H-Pro-Ile]$^+$, which confirmed the sequence *cyclo*(Pro^1-Gln^2-Thr^3-Gly^4-Mso^5-Leu^6-Pro^7-Ile^8) for **2**, with however, an ambiguity related to the respective location of Leu and Ile. Cherimolacyclopeptide **2** appeared to be an analogue of **1**, by Mso/Met substitution. An NMR study was undertaken to solve the above indecision, and in addition to assign the chemical shifts for the conformational study.

2.3. ^{1}H and ^{13}C NMR studies: sequence-specific assignment

The ^{1}H and ^{13}C NMR spectra of cherimolacyclopeptide A (**1**) were recorded in pyridine-d_5 (Table 1) and DMSO-d_6 solution (Table 2), and a major stable conformation (>90%) was observed in both solvents. However, the optimal conditions for **1** were in pyridine-d_5 at 263 K, where the minor component was negligeable. The six amide protons were clearly identified in the ^{1}H NMR spectrum, as well as the presence of eight carbonyl groups in the ^{13}C NMR spectrum, in agreement with an octapeptide structure including two prolines. The ^{1}H NMR spectrum of
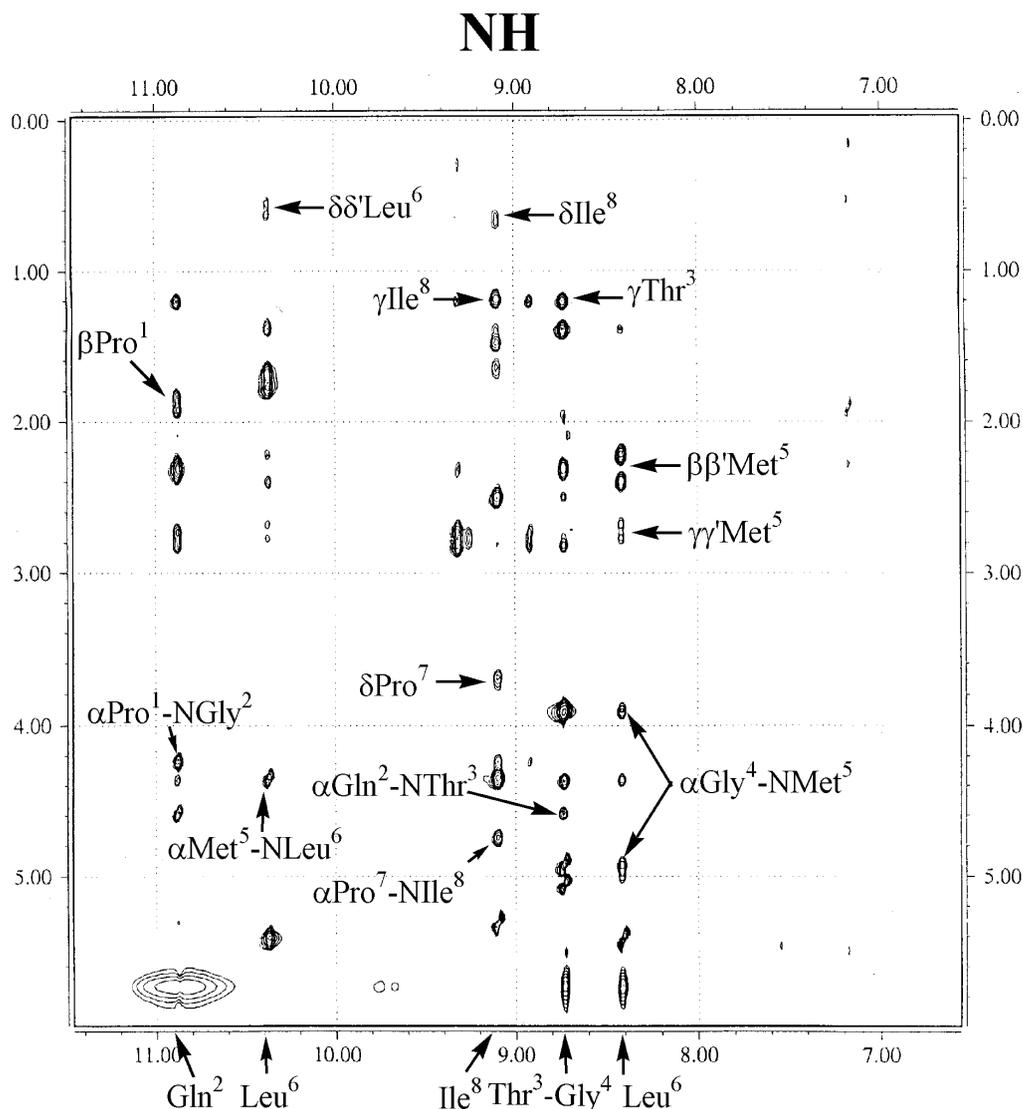
Figure 2. Part of the NOESY spectrum (120 ms) showing the NH to α and to side-chain proton correlations.

cherimolacyclopeptide A (**1**) was assigned via standard sequential assignment methods developed by Wüthrich.[24–26] The entire spin systems of individual amino acid residues were identified through COSY and TOCSY experiments and the ^{13}C NMR spectrum was assigned through HSQC and HMBC experiments (Table 1).

The final peptide sequence determination was based on the data of the HMBC experiment. All the amino acid spin systems were identified using scalar spin–spin couplings determined from the ^{1}H–^{1}H COSY and TOCSY experiments.[27] The ^{13}C NMR assignments of the protonated carbons were obtained from the proton detected heteronuclear HSQC spectrum. This experiment, with the HMBC experiment optimised for a long-range J-value of 7 Hz, for the non-protonated carbons allowed the carbonyl groups assignment. Carbonyl carbons of Pro^1, Gln^2, Thr^3, Met^5 and Pro^7 were easily identified from their intra-residue ^{3}J correlations with β protons, those of Gly^4 and Leu^6 from the strong intra-residue connectivities with the α protons (Table 1). Accordingly, the sequence determination was performed from the observation of the connectivities between the

carbonyl of residue i with the amide and/or α protons of residue $i+1$.

$^{3}J_{\mathrm{CH}}$ CO(i)/αH($i+1$) correlations between the leucine and Pro^7 and between the isoleucine and Pro^1 on the HMBC spectrum, indicates univocally that the Leu residue is at position 6 and Ile at position 8. Other connectivities are in full agreement with the structure deduced from the mass spectrometry study.

The NOESY spectrum recorded at 263 K in pyridine-d_5 clearly depicted strong or medium NOE $d_{NN(i,i+1)}$ interactions from Gln^2 to Met^5, a weak one between Met^5 and Leu^6 and a stretch of $d_{\alpha N(i,i+1)}$ sequential connectivities from Pro^1 to Leu^6 and between Pro^7 and Ile^8 (Figs. 2 and 3). The strong NOE between Gln^2 and Thr^3 suggested the presence of a β-turn with Gln^2 at the $i+2$ position. In addition, strong NOE correlations were observed between the α proton of Ile^8 and both δ and δ' protons of Pro^1, indicating that the Ile^8-Pro^1 amide bond is in *trans* configuration, while a strong correlation between the α protons of Leu^6 and Pro^7 indicates that the Leu^6-Pro^7 amide

Table 1. ^{13}C and ^{1}H NMR data and HMBC correlations of cherimolacyclopeptide A (**1**), in pyridine-d_5 (400 MHz, 308 K; s: singulet; d: doublet; dd: doublet of doublet; t: triplet; m: multiplet)

Residue	δ_C	δ_H	m (J Hz)	HMBC correlations with H
Pro$_b^1$ CO	174.7	—		αH, βH Pro[1]; NH Gln[2]
αCH	64.1	4.37	dd 8.7, 6.2	
βCH$_2$	30.1	2.16	m	
	—	2.03	m	
γCH$_2$	25.3	1.95	m	
	—	1.75	m	
δCH$_2$	48.7	4.22	m	
Gln[2] CO	172.9	—		αH, βH Gln[2]; NH Thr[3]
NH	—	10.50	d 3.9	
αCH	57.8	4.53	ddd 8.7, 4.2, 3.9	
βCH$_2$	26.2	2.35	m	
γCH$_2$	32.7	2.79	M	
	—	2.68	m	
δCO	177.7	—		βH, γH Gln[2]
εNH$_2$	—	8.83	s	
	—	8.37	s	
Thr[3] CO	172.3	—		αH, βH Thr[3]; NH Gly[4]
NH	—	8.59	d 9.2	
αCH	60.4	4.99	dd 9.2, 2.8	
βCH	69.1	4.95	dd 6.5, 2.8	
γCH$_3$	20.0	1.41	d 6.5	
Gly[4] CO	169.1	—		αH Gly[4]; NH, αH Met[5]
NH	—	8.55	dd 7.8, 4.8	
αCH$_2$	43.5	4.83	dd 17.0, 7.8	
	—	3.86	dd 17.0, 4.8	
Met[5] CO	173.3	—		αH, βH, Met[5]; NH Leu[6]
NH	—	8.21	d 9.2	
αCH	51.9	5.30	dt 9.2, 7.4	
βCH$_2$	33.9	2.39	m	
	—	2.28	m	
γCH$_2$	30.1	2.71	m	
	—	2.71	m	
εCH$_3$	14.9	1.85	s	
Leu[6] CO	172.0	—		αH Leu[6]; αH Pro[7]
NH	—	9.83	d 5.4	
αCH	52.2	4.40	m	
βCH$_2$	38.9	1.87	m	
	—	1.45	ddd 13.6, 10.0, 2.9	
γCH	25.3	1.76	m	
δCH$_3$	20.9	0.77	d 6.5	
δ'CH$_3$	23.2	0.70	d 6.5	
Pro$_a^7$ CO	171.1	—		αH, βH Pro[7]; NH, αH Ile[8]
αCH	61.5	4.72	d 7.6	
βCH$_2$	31.6	2.82	m	
	—	2.09	m	
γCH$_2$	22.5	1.80	m	
δCH$_2$	47.3	3.79	dd 11.0, 7.5	
	—	3.63	dd 11.0, 9.0	
Ile[8] CO	171.8	—		αH Ile[8]; αH Pro[1]
NH	—	8.99	d 9.8	
αCH	56.0	5.26	dd 9.8, 9.8	
βCH	35.1	2.50	ddd 9.8, 6.6, 3.3	
γCH$_2$	24.9	1.61	dqd 13.5, 7.4, 3.3	
	—	1.27	m	
γ'CH$_3$	16.8	1.24	d 6.6	
δCH$_3$	10.3	0.83	t 7.4	

bond is in *cis* configuration. These stereochemistries are further confirmed by the γ carbons ^{13}C chemical shifts of Pro[1] and Pro[7] at 25.3 and 22.5 ppm, respectively, in agreement with the presence of *trans*-Pro[1] and *cis*-Pro[7] amide bonds.[28] All the data indicate the cyclic structure **1** for cherimolacyclopeptide A, including only one *cis*-amide bond.

The ^{1}H and ^{13}C NMR spectra of cherimolacyclopeptide B (**2**) were recorded in DMSO-d_6, assigned via 2-D homo- and heteronuclear experiments and compared to those of **1** (Table 2). The CO of Pro[7] gave a strong correlation with the NH of Ile[8] and the CO of Mso[5] with the NH of Leu[6] leading to a non-ambiguous location of the position of the Leu and Ile residues in the sequence of **2**. A strong Hα(8)-Hδ(1)

Table 2. ^{13}C and ^{1}H NMR data for cherimolacyclopeptides A (**1**) and B (**2**) in DMSO-d_6 (298 K; s: singulet; d: doublet; dd: doublet of doublet; t: triplet; m: multiplet; Mxx: Met for **1** and Mso for **2**)

Residue	1			2		
	δ_C	δ_H	m (J Hz)	δ_C	δ_H	m (J Hz)
Pro$_b^1$ CO	173.3	—		173.5		
αCH	50.9	3.85	m	62.6	3.82	m
βCH$_2$	29.0	2.10	m	29.0	2.12	ddd 12.2, 8.4, 6.1
	—	1.83	m	—	1.85	m
γCH$_2$	24.4	1.93	m	24.5	1.95	m
	—	1.74	m	—	1.75	m
δCH$_2$	47.7	3.78	m	47.7	3.78	m
	—	3.49	m	—	3.69	m
Gln^2 CO	171.7	—		171.8	—	
NH	—	9.41	d 4.0	—	9.42	d 3.8
αCH	55.9	3.91	m	56.0	3.90	m
βCH$_2$	25.4	1.89	m	25.4	1.90	m
γCH$_2$	31.3	2.26	t 6.5	31.4	2.28	t 6.5
δCO	175.2	—		175.2	—	
εNH$_2$	—	7.57	s	—	7.58	s
	—	7.13	s	—	7.13	s
Thr^3 CO	170.9	—		171.0	—	
NH	—	7.99	d 8.9	—	7.97	d 9.0
αCH	59.1	4.17	dd 8.9, 3.1	59.1	4.18	dd 9.0, 3.0
βCH	67.7	4.23	dd 6.7, 3.1	67.5	4.26	dq 6.6, 3.0
γCH$_3$	19.2	1.12	d 6.7	19.3	1.12	d 6.6
OH	—	4.41	d 10.7	—		
Gly^4 CO	167.8	—		168.0	—	
NH	—	7.84	dd 7.9, 4.8	—	7.84	dd 7.5, 5.0
αCH$_2$	42.1	4.10	dd 17.3, 7.9	42.1	4.12	dd 17.2, 7.5
	—	3.33	dd 17.3, 4.8	—	3.33	dd 17.2, 5.0
Mxx^5 CO	171.2	—		171.0	—	
NH	—	7.34	d 9.2	—	7.35	d 9.1
αCH	50.4	4.53	ddd 9.2, 7.0, 7.0	49.8	4.63	ddd 9.1, 7.5, 7.5
βCH2	33.3	1.72	m	26.3	1.81	m
	—	1.64	m	—	1.81	m
γCH$_2$	28.8	2.42	m	48.3	2.81	dt 13.7, 7.6
	—	2.30	m	—	2.66	dt 13.7, 7.6
δCH$_3$	14.5	2.00	s	37.7	2.49	s
Leu^6 CO	170.6	—		170.7	—	
NH	—	8.61	d 4.7	—	8.73	d 5.0
αCH	62.6	3.84	m	51.0	3.84	m
βCH$_2$	37.5	1.40	m	37.6	1.50	m
	—	1.31	m	—	1.36	m
γCH	24.1	1.69	m	24.2	1.70	m
δCH$_3$	20.3	0.78	d 6.6	20.3	0.79	d 6.5
δCH$_3$	23.1	0.89	d 6.6	23.1	0.89	d 6.6
Pro$_a^7$ CO	170.1	—		170.1	—	
αCH	60.2	4.38	m	60.2	4.39	d 7.8
βCH$_2$	30.7	2.39	m	30.7	2.39	ddd 11.6, 7.8, 6.3
	—	2.01	m	—	2.02	m
γCH$_2$	21.7	1.85	m	21.7	1.87	m
	—	1.48	m	—	1.50	m
δCH$_2$	46.4	3.40	m	46.4	3.40	m
	—	3.27	m	—	3.30	m
Ile^8 CO	170.2	—		170.1	—	
NH	—	8.22	d 9.6	—	8.20	d 9.7
αCH	54.8	4.62	dd 10.0, 9.6	54.7	4.62	dd 10.2, 9.7
βCH	34.1	2.00	m	34.0	2.02	m
γCH$_2$	23.7	1.41	m	23.7	1.42	m
	—	0.98	m	—	1.00	m
γ'CH$_3$	16.2	0.89	d 6.6	16.2	0.89	d 6.6
δCH$_3$	10.2	0.81	t 7.5	10.2	0.81	t 7.4

NOE connectivity together with a Cγ chemical shifts at 24.5 ppm for Pro[1] allowed to identify a *trans* configuration for the Leu[8]-Pro[1] amide bond. The Leu[6]-Pro[7] peptide bond was easily identified as a *cis*-isomer by its Cγ chemical shifts at 21.7 ppm and the presence of a strong correlation between α protons of Leu[6] and Pro[7] in the NOESY spectrum. Peptide **2** appears thus to differ from **1** only by the Met/Mso substitution.

Cyclo	(Pro$_b^1$)	Gln^2	Thr^3	Gly^4	Met^5	Leu^6	Pro$_a^7$	Ile^8
d$_{NN(i,i+1)}$								
d$_{NN(i,i+2)}$								
d$_{\alpha N(i,i+1)}$								
d$_{\alpha N(i,i+2)}$								
d$_{\beta N(i,i+1)}$								
d$_{\beta N(i,i+2)}$								
d$_{\beta N(i,i+3)}$								
d$_{\alpha\alpha(i,i+1)}$								
d$_{\alpha\delta(i,i+1)}$								
-Δδ/ΔT	-	8.0 ε: 10.7,11.7	3.2	4.0	4.3	12.0	-	4.6
H/D exchange	-	m	m	m	m	f	-	s

Figure 3. NMR data for peptide **1**: sequential connectivities (strong, medium and weak), temperature coefficients of amide protons (ppb k^{-1}) and H/D exchange rate for cherimolacyclopeptide A in pyridine-d_5 [s (slow): exchanged in more than 5 h; m (medium): exchanged in 1 to 5 h; f (fast): exchanged in less than 1 h, at 293 K].

2.4. Solution structure of 1

The 3-D structure of cherimolacyclopeptide A (**1**) was determined using a restraint file including 30 intra-residual, 40 sequential restraints, six ($i,i+2$) restraints between Gln^2 and Ile^8, Thr^3 and Met^5 and between Pro^7 and Pro^1 and finally, ten non-sequential restraints (Table 3) among which three connect Thr^3 to Ile^8 and three involve Pro^1. Five dihedral angle restraints were deduced from the $^3J_{NH,H\alpha}$ coupling constants.

Among the 100 structures generated with DYANA, 87 display a target function smaller than 1 Å^2. The RMS deviation between the Cα atoms of 40 best structures (Target function ≤0.21 Å^2) was only 0.06±0.04 Å. In order to enhance the conformational space sampling, these 40 structures were subjected to simulated annealing and energy minimizations in the cartesian coordinates space as described in the Section 3. Twenty structures showing the smallest number of residual violations were selected for further analysis. The pairwise RMSD between the Cα backbone atoms of the 20 selected structures (0.18±0.08 Å) remains very small, showing that the structures are well defined by the NMR data. Most side-chains adopt a well-defined conformation due to the presence of numerous NOEs, the only side-chain showing a large conformational variability is Gln^2.

The 20 selected structures are in very good agreement with all experimental data and the standard covalent geometry. There are no distance violation larger than 0.2 Å, no angle violations larger than 5° and the root-mean-square deviations (RMSD) with respect to the standard covalent geometry are low. Both negative van der Waals and

Table 3. Non-sequential NOEs

NH Gln^2	αH Ile^8
NH Gln^2	γCH$_3$ Ile^8
NH Thr^3	βCH Ile^8
NH Thr^3	γCH$_3$ Ile^8
γCH$_3$ Thr^3	NH Met^5
γCH$_3$ Thr^3	NH Ile^8
αH Pro^7	δCH$_2$ Pro^1
εH or ε'H Gln^2	γCH$_3$ Ile^8
δCH$_3$ or δ'CH$_3$ Leu^6	βH or β'H Pro^1
δCH$_3$ or δ'CH$_3$ Leu^6	δH or δ'H Pro^1

Table 4. Structural statistics of the 20 models of cherimolacyclopeptide A (**1**)[a]

|	
Restraint violations	
Distance restraints >0.2 Å	0
Deviation from standard geometry	
Bond lengths>0.05 Å	0
Bond angles >10°	0
Final energies (kcal mol^{-1})[a]	
$E_{electrostatic}$	−61±3
E_{vdw}	−21±4
RMSD	
	Pairwise (Å)
Backbone	0.18±0.08
All atoms	1.04±0.18

[a] The energy terms were calculated using the CHARMM force field.

electrostatic energy terms are indicative of favorable non-bonded interactions (Table 4). Moreover, 100% of the (ϕ,ψ) angles of all structures are in the most favored regions of the Ramachandran plot and additional allowed regions according to the PROCHECK software nomenclature.[38]

The overall fold of cherimolacyclopeptide A (**1**) (Figs. 4–6) includes two β turns. The first one, formed with Met^5 at position $i+1$ and Ile^8 at position $i+3$, can be classified as a VIa turn. It is characterized by the presence of a *cis*-proline at position $i+2$ and (ϕ,ψ) angles for residues $i+1$ and $i+2$ (Table 5) closed from the standard values: (−60, +120) and (−90, 0), respectively.[29] This turn is stabilized by the
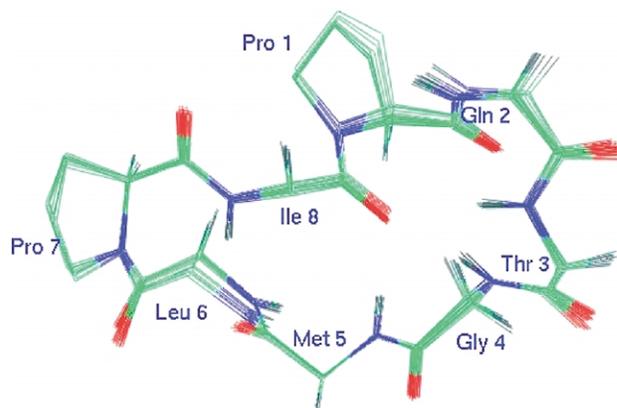
Figure 4. Superposition of the 20 NMR derived structures of cherimola-cyclopeptide A (**1**) in pyridine-d_5 (263 K).
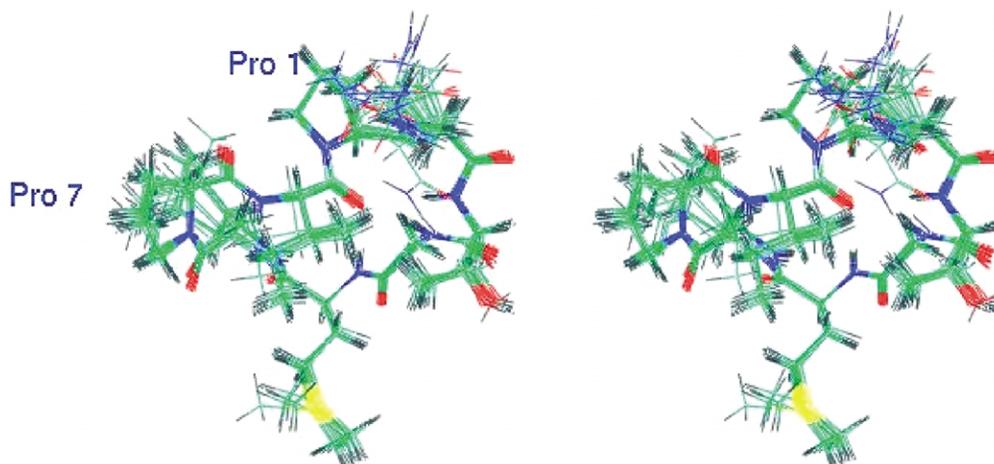
Figure 5. Stereoview of a superposition of the 20 NMR structures of cherimolacyclopeptide A (**1**) in pyridine-d_5 (263 K) showing the side-chain and the polypeptide backbone.

Table 5. Main chain torsional angles

Residue	ϕ (°)	ψ (°)
Pro[1]	−42.6±3.4	−44.8±4.5
Gln[2]	−31.4±3.5	−40.1±6.4
Thr[3]	−112.5±2.1	−34.9±11.6
Gly[4]	−117.1±4.9	36.6±4.5
Met[5]	−132.8±2.5	65.1±3.8
Leu[6]	−36±1.8	117.0±5.1
Pro[7]	−99.3±2.1	31.0±2.4
Ile[8]	−115.1±0.7	62.0±4.5

canonical hydrogen bond between the C=O group of Met[5] and the NH of Ile[8] with an average O−H_N hydrogen bond distance of 1.86 Å. This is in agreement with the low value of the chemical shift dependence of the NH of Ile[8], as well as its low exchange rate with D_2O. Another β-turn is formed between Ile[8] and Thr[3]. This turn is a type I β-turn with *trans*-Pro[1] at position $i+1$, it is very distorted because of the proline. A rather similar conformation was previously found for the sequence Phe-Pro-Ala-Arg of the cyclic peptide *cyclo*(RGDFPA)[30] where the proline was also in *trans*-configuration. However, as in cherimolacyclopeptide A (**1**)

structure the turn is stabilized by a trifurcated hydrogen bond between the C=O group of Ile[8] and the NH groups of Thr[3], Gly[4] and Met[5], with average O−H_N distances of 2.14, 1.97 and 1.97 Å, respectively, the distortion could be due to this particular hydrogen bonding. The temperature coefficients ($\Delta\delta/\Delta T$) for the amide resonances of these three residues, in pyridine as well as in DMSO solution, are in accordance with the hydrogen bond network found on the 20 structures (Table 3, Fig. 2). It should be noted that the values measured in pyridine are about two times larger than those measured in DMSO. The low values, in the range −3 to −5 ppb K^{-1} for Thr[3], Gly[4], Met[5] and Ile[8] in pyridine (Fig. 3) and which are in the range −1 to −2 ppb K^{-1} in DMSO (Table 6) indicated that they are involved in intramolecular hydrogen bonds, whereas the large temperature coefficient values found in pyridine for Leu[6] (−12 ppb K^{-1}), for Gln[2] (−8 ppb K^{-1}) and also for the εNH₂ protons of Gln[2] side-chain (−10.7 and −12 ppb K^{-1}) indicated they are solvent exposed (Fig. 3). Large values are also observed in DMSO-d_6 (Table 6).

The cyclooctapeptide **1** appeared to be structured with two β-turns of type VIa and I, the latter being stabilized by a
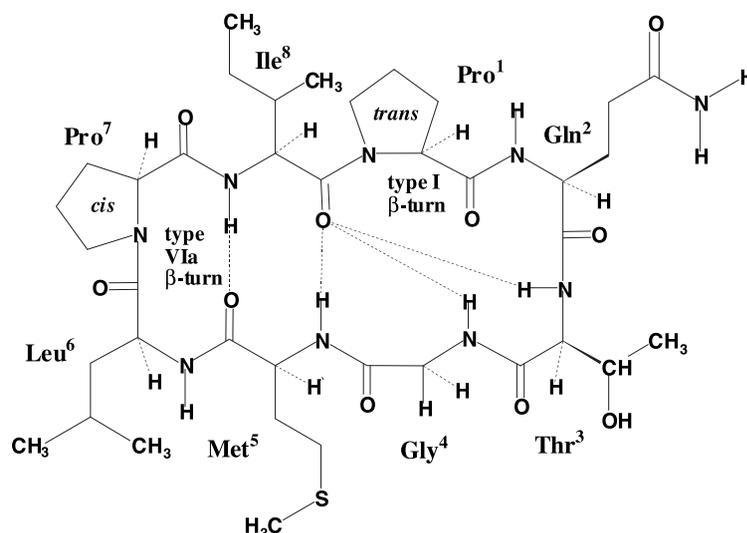
Figure 6. Schematic diagram of the turn types and the hydrogen bonds (broken lines) for cherimolacyclopeptide A (**1**).

Table 6. Temperature coefficients ($-\Delta\delta/\Delta T$, ppb K^{-1}) for amide protons of cherimolacyclopeptides A (**1**) and B (**2**) in DMSO-d_6

	Pro$_b^1$	Gln^2	Thr^3	Gly^4	Mxx^5	Leu^6	Pro$_a^7$	Ile^8	ε Gln^2
(**1**)	—	4.2	1.6	2.0	1.9 (Met)	5.0	—	1.2	4.8; 4.0
(**2**)	—	4.0	1.6	1.7	1.8 (Mso)	4.9	—	1.0	3.6; 3.6

trifurcated hydrogen bond. It is interesting to note that cyclohexapeptides are usually organized with two β-turns stabilized by two hydrogen bonds, and that cycloheptapeptides are usually organized with two β-turns, one stabilized by a normal hydrogen bond and the second by a bifurcated hydrogen bond forming a β-bulge.[10] A new type of β-bulge is observed in the structure of **1**, involving a trifurcated hydrogen bonding instead of a bifurcated one. All the NMR data (Tables 2 and 6) indicates that cyclopeptides **1** and **2** adopt a similar conformation in solution.

Due to the presence of side-chain side-chain and side-chain backbone connectivities (Table 3), most side-chains adopt a well-defined orientation. This is particularly striking for Thr^3 and Ile^8, the only residues showing variability are Gln^2 and to a less extend Leu^6 (Fig. 5). However, these residues adopt always the same orientation with respect to the backbone and the variability is limited to the δCH$_3$ groups of Leu^6 and to the amide group of Gln^2.

3. 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. Mass spectra were recorded on an API Q-STAR PULSAR *i* of Applied Biosystem. For the CID spectra, the collision energy was 40 eV and the collision gas was nitrogen.

^{1}H NMR spectra were recorded either on an INOVA 600 Varian spectrometer operating at 600 MHz or a Bruker Avance 400 spectrometer operating at 400.13 MHz equipped with X-WIN NMR (version 2.6). 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.

3.1. Plant material

Fruits of *Annona cherimola* Miller (Annonaceae) were collected in the south of Spain in December 2000. The seeds were collected and were immediately washed with distilled water and were dried at room temperature. Samples were deposited in the Herbarium of the National Museum of Natural History (Paris).

3.2. Extraction and isolation

The dried and powdered seeds of *A. cherimola* (3.0 kg) were macerated three times with cyclohexane (3 L), the combined extracts yielded an oil (531 g) 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 (126 g) which was partitioned between EtOAc and water. The organic phase was concentrated to dryness and the residue (61.5 g) was dissolved in MeOH and chromatographed on Sephadex LH-20 column with MeOH. The head fraction (33.6 g) containing peptides and acetogenins was subjected to repeated silica gel column chromatography (Kieselgel 60 H Merck) eluted with CH$_2$Cl$_2$ containing increasing amount of MeOH from 5 to 20% yielding to three peptide fractions (I-III), characterised by TLC on silica gel 60 F$_{254}$ Merck, with CH$_2$Cl$_2$/MeOH 9:1 as eluent system. The peptides were detected with Cl$_2$/*o*-tolidine reagent as blue spots with *Rf* 0.24 (I, 260 mg), 0.43 (II, 690 mg) and 0.45 (III, 410 mg). The two last peptide fractions II and III were purified by isochratic reversed phase HPLC (Kromasil C$_{18}$, 250×7.8 mm, 5 µm, AIT France; flow rate 2 mL/min, detection 220 nm). Fraction II using MeOH/H$_2$O: 55:45 with 0.1% TFA, yielded cherimolacyclopeptide A (**1**, retention time (t_R) 23.9 min, 185 mg); Fraction III using MeOH/H$_2$O: 55:45 with 0.1% TFA, yielded cherimolacyclopeptide B (**2**, t_R 10.6 min, 199 mg).

3.3. Absolute configuration of amino acids

Solutions of **1** and **2** (each containing 1 mg of peptide) in 6 N HCl (1 mL) were heated at 110 °C for 24 h in sealed tubes. After cooling, the solutions were concentrated to dryness. The hydrolysates were dissolved in anhydrous solution of 3 N HCl in 2-propanol and heated at 110 °C for 30 min. The reagent were evaporated under reduce pressure. The residues were dissolved in CH$_2$Cl$_2$ (0.5 mL) and 0.5 mL trifluoracetic anhydride was added. The mixtures were kept in a screw-capped tubes 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 temperature program of 50–130 °C at 3 °C/min and 130–190 °C at 10 °C/min, with a HEWLETT PACKARD series 5890 apparatus. Comparison of Rt values with those of standard amino acids was used: L-Glu (29.3), Gly (14.6), L-Ile (16.9), L-Leu (19.2), L-Met (27.9), L-Mso (27.8), L-Pro (18.2) and L-Thr (15.2).

3.3.1. Cherimolacyclopeptide A. C$_{38}$H$_{63}$N$_9$O$_{10}$S: colourless solid, mp 192–193 °C (MeOH); $[\alpha]_D^{22}$ −8.5 (*c* 0.9, MeOH). ESI-qTOF, *m/z*: 876 [M+K]$^+$, 860 [M+Na]$^+$, 838 [M+H]$^+$. ESI-qTOF MS/MS on [M+H]$^+$(ce 40 eV) *m/z* (%): 838 (11), 820 (3), 810 (9), 725 (23), 707 (13), 697 (17), 649 (7), 628 (37), 611 (20), 610 (24), 600 (41), 594 (7), 583 (14), 566 (5), 552 (14), 537 (5), 515 (57), 497 (24), 487 (13), 439 (50), 436 (11), 408 (3), 404 (10), 403 (21), 401 (14), 384 (44), 366 (11), 356 (2), 327 (70), 309 (32), 299 (5), 290 (37), 283 (8), 243 (9), 226 (100), 211 (11), 198 (57), 183 (10), 115 (10), 170 (10).

3.3.2. Cherimolacyclopeptide B. $C_{38}H_{63}N_9O_{11}S$: colourless solid, mp 228–229 °C (MeOH); $[\alpha]_D^{22}$ −8.3 (c 2, MeOH). ESI-qTOF, m/z: 892 [M+K]$^+$, 876 [M+Na]$^+$, 854 [M+H]$^+$.

ESI-qTOF MS/MS on [M+H]$^+$(ce 40 eV) m/z (%): 854 (23), 836 (5), 826 (22), 741 (27), 723 (7), 713 (22), 677 (26), 649 (22), 644 (70), 626 (27), 616 (85), 599 (17), 594 (5), 580 (53), 566 (6), 562 (19), 552 (53), 537 (12), 531 (100), 516 (25), 513 (26), 509 (2), 503 (3), 496 (16), 467 (85), 449 (18), 439 (39), 436 (21), 419 (48), 418 (7), 408 (7), 391 (23), 384 (43), 366 (15), 356 (6), 327 (82), 318 (14), 309 (35), 306 (65), 299 (7), 226 (59), 211 (8), 198 (4), 183 (7), 170 (5).

3.4. Nuclear magnetic resonance spectroscopy

The NMR spectra were recorded either in DMSO-d_6 or in pyridine-d_5 (0.7 mL of 25 mM solution). At 298 K the ^{1}H NMR spectrum showed two sets of signals, indicating the occurrence of two slowly interconverting conformations (I and II), presumably due to a *cis/trans* isomerism of one proline. NMR spectra were recorded at 258, 263, 273, 278, 383, 298 and 318 K, showing a temperature dependence of the I:II ratio. At 263 K, the second set of signals become negligible, therefore, NMR spectra subsequently used for NOE quantification were recorded at 263 K.

The temperature coefficients of amide protons were obtained by 5 K temperature increments, in the 278–318 K range: the variation was linear indicating no variation of the conformation with the temperature and the NH coupling constants were unchanged.

A conventional set of one (1-D) and two dimensional (2-D) ^{1}H NMR TOCSY and NOESY spectra, was acquired at a temperature of 263 K on a VARIAN INOVA NMR spectrometer equipped with a z-axis field-gradient unit and operating at a proton frequency of 600 MHz. The TOCSY spectrum was collected with a spin lock time of 80 ms using the MLEV-17 mixing scheme[31] and NOESY spectra were recorded with mixing times of 120 and 300 ms. 2D NMR spectra were processed on a SGI O$_2$ workstation using the NMR Pipe/Draw software.[32] The NMR data sets were analysed with XEASY.[33]

NOESY cross peaks recorded with a mixing time of 120 ms were converted into upper distance limit restraints using the CALIBA program included in the DYANA package. In order to assess possible contributions from spin diffusion effects, some NOEs only observable on the 300 ms mixing time NOESY map were taken into account with a 6 Å upper bound. The minimum distance constraint between two protons was limited by their van der Waals radii (2.0 Å). All these restraints were brought together in a distance restraint file used as input to initial steps of molecular modelling. In addition, in DYANA, the cyclization was ensured by a set of distance constraints on the backbone atoms of the first and last residue, the corresponding van der Waals constraints were removed. Backbone ϕ dihedral angle restraints were deduced from the 1D COSY spectrum. These ϕ angles were restrained to −60±30° for $^3J_{NH-H\alpha}$<6.5 Hz and −120±40° for $^3J_{NH-H\alpha}$≥8 Hz.

3.5. Structure calculations

Structure calculations were performed using an hybrid method including simulated annealing in the torsion angle space with DYANA[34,35] followed by simulated annealing in cartesian coordinates with XPLOR 3.1.[36]

A set of 100 structures was generated from random-built initial models using the annealing procedure of the variable target function program DYANA. In order to increase the sampling of the conformational space, a subset of 40 structures, with a target function ≤0.21 Å^2 were submitted to simulated annealing using the standard protocol of XPLOR 3.1. Finally these structures were energy minimised using Powell's algorithm and CHARMM force field parameters[37] implemented in X-PLOR 3.1 software. Twenty structures showing the smallest number of residual violations were selected as representative of the cherimolacyclopeptide A (**1**) structure and analysed using PROCHECK[38] and PROMOTIF.[39]

References and Notes

1. Leboeuf, M.; Cavé, A.; Bhaumik, P. K.; Mukherjee, B.; Mukherjee, R. *Phytochemistry* **1982**, *21*, 2783–2813.
2. Chen, C.-Y.; Chang, F.-R.; Pan, W.-B.; Wu, Y.-C. *Phytochemistry* **2001**, *56*, 753–757.
3. Rupprecht, J. K.; Hui, Y.-H.; McLaughlin, J. L. *J. Nat. Prod.* **1990**, *53*, 237–278.
4. Sevser, S.; Gonzalez, C.; Hocquemiller, R.; Zafra-Polo, M. C.; Cortes, D. *Phytochemistry* **1996**, *42*, 103–107.
5. Chen, C.-Y.; Chang, F.-R.; Pan, W.-B.; Wu, M.-J.; Wu, Y.-C. *Phytochemistry* **1999**, *51*, 429–433.
6. Woo, M.-H.; Kim, D.-H.; Fotopoulos, S. S.; McLaughlin, J. L. *J. Nat. Prod.* **1999**, *62*, 1250–1255.
7. Chen, C.-Y.; Chang, F.-R.; Pan, W.-B.; Wu, Y.-C. *Tetrahedron Lett.* **1998**, *39*, 407–410.
8. Chen, C.-Y.; Chang, F.-R.; Yen, H. F.; Pan, W.-B.; Wu, Y.-C. *Phytochemistry* **1998**, *49*, 1443–1447.
9. Auvin-Guette, C.; Baraguey, C.; Blond, A.; Pousset, J. L.; Bodo, B. *J. Nat. Prod.* **1997**, *60*, 1155–1157.
10. Auvin-Guette, C.; Baraguey, C.; Blond, A.; Xavier, S. X.; Pousset, J.-L.; Bodo, B. *Tetrahedron* **1999**, *55*, 11495–11510.
11. Baraguey, C.; Blond, A.; Cavelier, F.; Pousset, J.-L.; Bodo, B.; Auvin-Guette, C. *J. Chem. Soc., Perkin Trans. 1* **2001**, 2098–2103.

12. Murakami, M.; Yamaguchi, K. *Tetrahedron* **1996**, *52*, 13129–13136.
13. Cebrat, M.; Wieczorek, Z.; Siemion, I. Z. *Peptides* **1996**, *51*, 191–196.
14. Tabudravu, J. N.; Morris, L. A.; Kettenes-van den Bosch, J. J.; Jaspars, M. *Tetrahedron* **2002**, *58*, 7863–7868.
15. Morita, H.; Kayashita, T.; Takeya, K.; Itokawa, H. *J. Nat. Prod.* **1995**, *58*, 943.
16. Morita, H.; Shishido, A.; Matsumoto, T.; Itokawa, H.; Takeya, K. *Tetrahedron* **1999**, *55*, 967.
17. Morita, H.; Sato, Y.; Kobayashi, J. *Tetrahedron* **1999**, *55*, 7509–7518.
18. Saviano, M.; Isernia, C.; Rossi, F.; Di Blasio, B.; Iacovino, R.; Mazzeo, M.; Pedone, C.; Benedetti, E. *Biopolymers* **2000**, *189-199*, 19.
19. Kopple, K. D.; Parameswaran, K. N.; Yonan, J. P. *J. Am. Chem. Soc.* **1984**, *106*, 7212–7217.
20. Kopple, K. D.; Kartha, G.; Bhandary, K. K.; Romanowska, K. *J. Am. Chem. Soc.* **1985**, *107*, 4893.
21. Lin, S.; Liehr, G.; Cooperman, B. S.; Cotter, R. J. *J. Mass Spectrom.* **2001**, *36*, 658–663.
22. Tomer, K. B.; Crow, F. W.; Gross, M. L.; Kopple, K. D. *Anal. Chem.* **1984**, *56*, 880–886.
23. Yuan, M.; Namikoshi, M.; Otsuki, A.; Rinehart, K. L.; Sivonen, K.; Watanabe, M. F. *J. Mass Spectrom.* **1999**, *34*, 33–43.
24. Wüthrich, K. *NMR of proteins and nucleic acids*; Wiley: New York, 1986.
25. Wüthrich, K.; Wider, G.; Wagner, G.; Braun, W. *J. Mol. Biol.* **1982**, *155*, 311–319.
26. Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715–740.
27. Wagner, G.; Kumar, A.; Wüthrich, K. *Eur. J. Biochem.* **1981**, *114*, 375–384.
28. Douglas, D. E.; Bovey, F. A. *J. Org. Chem.* **1973**, *38*, 2379–2383.
29. Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1–109.
30. Müller, G.; Gurrath, M.; Kurz, M.; Kessler, H. *Proteins* **1993**, *15*, 235–251.
31. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355–360.
32. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.
33. Bartels, C.; Xia, T. E.; Billeter, M.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1995**, *5*, 1–10.
34. Güntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283–298.
35. Güntert, P.; Mumenthaler, C.; Herrmann, T. DYANA version 1.5—User's manual, ETH. Honggerberg: Zürich, 1998.
36. Brünger, A. T. The X-PLOR software manual Version 3.1. Yale University Press: New Haven, 1992.
37. Brooks, B.; Brucoli, R.; Olafson, B. O.; States, D.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187–217.
38. Laskowski, R. A.; Rullmann, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8*, 477–486.
39. Hutchinson, E. G.; Thornton, J. M. PROMOTIF a program to identify and analyse structural motifs in proteins. *Protein Sci.* **1996**, *5*, 212–220.