The cyclic structure of the enterococcal peptide antibiotic AS-48

Bart Samyn[a], Manuel Martinez-Bueno[b], Bart Devreese[a], Mercedes Maqueda[c], Antonio Gálvez[c], Eva Valdivia[c], Jacques Coyette[b], Jozef Van Beeumen[a],*

[a]*Department of Biochemistry, Physiology and Microbiology, University Gent, K.L. Ledeganckstraat, 35, B-9000 Gent, Belgium*
[b]*Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie B6, Sart Tilman, B-4000 Liège, Belgium*
[c]*Departamento de Microbiologia, Facultad de Ciencias, Universidad de Granada, E-18071 Granada, Spain*

Abstract The complete primary structure of the peptide antibiotic AS-48 produced by *Enterococcus faecalis* has been determined by chemical degradation analysis. The cyclic nature of this 70 residues containing peptide was demonstrated by plasma desorption mass analysis of the generated peptides and electrospray ionisation mass analysis of the native polypeptide. As far as we know, this is the first example of an antibiotic protein cyclised by a tail–head peptide bond formation and not by branching of the polypeptide side chains.

Key words: Primary structure; Cyclic peptide; Antibiotic; Mass analysis

1. Introduction

The AS-48 antibiotic protein from *Enterococcus faecalis* subsp. *liquefaciens* S-48 has a broad antimicrobial activity against Gram-positive and Gram-negative bacteria [1,2]. This activity is due to insertion of the polypeptide into the cytoplasmic membrane of the target cell which thereby becomes permeable to small molecules and starts loosing cytoplasmic material [3]. A similar phenomenon of cell lysis is known to occur also as a result of the action of bacteriocins and peptide antibiotics produced by lactic acid bacteria [4]. In contrast to the lantibiotics such as nisin [5] and lactacin 491 [6] which contain unusual amino acids, the antibiotic AS-48 appears to contain no other than the natural amino acids as could be deduced from its amino acid compositional analysis [2]. Here we report the complete primary structure of the antibiotic protein, a goal which could only be achieved by Edman degradation analysis in combination with accurate mass analysis methods.

2. Materials and methods

Peptide AS-48 was purified as in [2]. About 30 nmol of the AS-48 protein were digested with Lys-C endoproteinase (Wako, Osaka, Japan). The digest was carried out in 100 mM Tris-HCl buffer, pH 8.0, using an E(nzyme)/S(ubstrate) ratio of 1/30, during 12 h at 37°C. A second digest, on the same amount, was performed using the endoproteinase Glu-C isolated from *Staphylococcus aureus* strain V8 (Miles, Slough, UK); E/S was 1/30 in 20 mM NaH_2PO_4, pH 7, and the incubation 5 h at 37°C. A final digest, with chymotrypsine (Worthington, Freehold, NY), was then performed (30 nmol AS-48, E/S = 1/30) in 100 mM NH_4HCO_3, pH 8.0, for 4 h at 37°C.

The peptides resulting from these digests were subjected to reversed-phase high-performance liquid chromatography on a PTC C18 column (2.1 × 220 mm, 5 μm) (Brownlee, Applied Biosystems, Foster City, USA). The gradient was formed with a 140A solvent delivery system (Applied Biosystems). The composition of solvent A was 0.1% trifluoroacetic acid (TFA) in Milli Q water (Millipore, Bedford, USA); solvent B was 0.1% TFA in 70% acetonitrile/water. Detection of the peptides was done with a diode array detector (Model 1000S, Applied Biosystems) and the absorbance at 220 nm registered on a chart recorder (speed: 0.5 cm/min). Fractions were collected manually in polypropylene tubes and dried in a Speed Vac concentrator (Savant, Hicksville, USA). Prior to sequence analysis peptides were redissolved in 0.1% TFA/water; the remaining peptide solutions were stored at −18°C.

Sequence analysis was performed on the models 477A or 476A pulsed liquid phase sequenators with on-line PTH-analysis on a 120A analyser (Applied Biosystems). Sequencing reagents and solvents were obtained from the same firm.
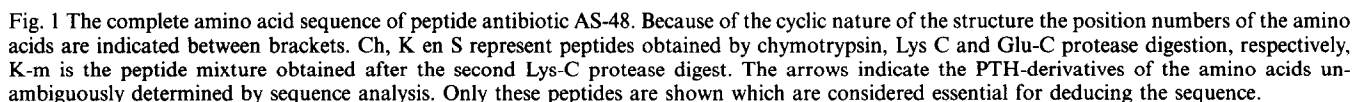
Mass determination of the peptides was performed by plasma desorption mass spectrometry (PDMS) on a BIO-ION 20K time of flight instrument (BioIon, Uppsala, Sweden). Samples were prepared by adsorption on a nitrocellulose layer in a 50% $MeOH/H_2O$ solution for 15 min. After adsorption, the targets where dried with a N_2 stream. The applied acceleration voltage was 15 kV for 1 million counts (start detector). Mass determination of the complete protein was determined on a BIO-Q electrospray mass spectrometer (Fisons Instruments, Altrincham, UK). In this case the protein was dissolved in 50% MeOH/H_2O, 1% acetic acid and introduced by a 10 μl sample loop injector in the ionisation chamber at a flow rate (same solvent) of 4 μl/min. The capillary voltage was set at 4 kV, the sample cone voltage was 50 V.

3. Results and discussion

3.1. Sequence analysis

The primary structure determination of the AS-48 protein encountered the initial difficulty to be not susceptible to Edman degradation, suggesting that the N-terminal residue was blocked. Also treatment with carboxypeptidases A and B did not cleave off any C-terminal residues (results not shown). A first digest, with endoproteinase Lys-C, revealed only four peptides of relatively short length, two of which (K3 and K4) had the same sequence containing a tryptophan residue at Edman-cycle 6. The position of this residue is shown as the last residue in Fig. 1. Since some large peptides were apparently lost on the C18 column during separation, a second Lys-C digest was performed and the resulting peptide mixture was subjected to sequence analysis without separation. Apart from the sequences already known, this mixture revealed the sequence of a region covering 49 residues (Glu^4-Lys^{52}) which could be identified as a single polypeptide from Edmancycle 10 onwards. The hydrophobic nature of this region explains why the peptide remained adsorbed to the C18 reversed phase column using the elution conditions described in section 2.

(1)   (10)   (20)
— Met Ala Lys Glu Phe Gly Ile Pro Ala Ala Val Ala Gly Thr Val Leu Asn Val Val Glu Ala Gly Gly Trp

K-m

S6 a   S6 c
S7 d
S6 b
S7 a   b   c

Ch3   Ch1 5   Ch9 a
Ch6 a

(30)   (40)
Val Thr Thr Ile Val Ser Ile Leu Thr Ala Val Gly Ser Gly Gly Leu Ser Leu Leu Ala Ala Ala Gly Arg

K-m

Ch13   Ch5 a
Ch6 b

(50)   (60)   (70)
Glu Ser Ile Lys Ala Tyr Leu Lys Lys Glu Ile Lys Lys Lys Gly Lys Arg Ala Val Ile Ala Trp ---

K-m   K-m   K-m
K2   K1   K3/4

S4   S5/6 a
S2   S7 d

Ch4 a   Ch2   Ch9 b
Ch4 b   Ch11

Fig. 1 The complete amino acid sequence of peptide antibiotic AS-48. Because of the cyclic nature of the structure the position numbers of the amino acids are indicated between brackets. Ch, K en S represent peptides obtained by chymotrypsin, Lys C and Glu-C protease digestion, respectively, K-m is the peptide mixture obtained after the second Lys-C protease digest. The arrows indicate the PTH-derivatives of the amino acids unambiguously determined by sequence analysis. Only these peptides are shown which are considered essential for deducing the sequence.

A second digest, envisaged by the presence of a Glu-residue at position 20, was performed using endoproteinase Glu-C (Fig. 2A). Some peptides, S2, S4, S6a and S7d, were found to cover the assumed C-terminal region of the protein deduced from the gene sequence [7]. The sequence of the peptides S6a and S7d, however, did not end at the assumed C-terminal residue Trp[70] but continued at the assumed N-terminal sequence as found in the sequence analyses of peptides K3 and K4. Other peptides, S6b, S6c and S7a,b,c covered the N-terminal region. Several aspecific cleavages were found: peptide S2 ends with a Tyr residue, peptide S6b starts after an Ala residue. This residue is also one of the three C-terminal residues of the C-terminal ragged peptide S7. Again, also in this digest, no peptides covering the hydrophobic sequence in the middle of the protein could be recovered from the RPLC-column.

To determine the complete sequence, a third digest, with the less specific protease chymotrypsine was performed. Separation of the peptides yielded a large number of smaller fragments (Fig. 2B). Sequence analysis of these fragments confirmed the sequence deduced from the gene [7] and also covered the hydrophobic region of the AS-48 protein. The necessary overlaps between the Lys-C and Glu-C peptides were detected so that the complete sequence of the AS-48 protein was thereby unambiguously determined.

3.2. Mass analysis

The obtained sequence was verified by determination of the molecular weight of the separated peptides by plasma desorption mass spectrometry. This method, in which the molecular ions are analysed by a time of flight technique, is known to have

an accuracy of 0.1% for peptides with a molecular weight below 5,000 Da [8]. The results of the mass analyses, given in Table 1, confirmed the results obtained by the sequence determination. Peptides K3, S5 and Ch3 all yielded an experimental mass of approximately 16 atomic mass units (amu) higher than the value calculated from the sequence. Since these peptides only have three residues in common, (M1-K3), we propose that this difference is due to a partial oxidation of the Met[1] residue. The non-oxidised form of these peptides was also found in respectively K4, S6a and Ch6a. The masses of the peptides K3, K4, S5, S6a and S7d, which link the C-terminal residue, Trp[70], to the N-terminal residue Met[1] all correspond to the calculated mass based on the sequence. Since the mass calculations of these peptides suggested the formation of a normal peptide bond and because the sequencer took the formed link without any problem, we conclude that this tail–head junction is indeed formed by a normal peptide bond.

The mass of the complete peptide antibiotic AS-48 was determined by electrospray mass spectrometry, a technique known to have an accuracy of 0.01% [9]. The calculated mass of the AS-48 protein, based on the average residual mass of each amino acid, is 7,167. 58 Da. The electrospray mass spectrum, however, indicates three masses of 7,150.17 Da, 7,166.5 Da and 7,182.66 Da (indicated respectively as A, B and C in Fig. 3, top).

Table 1
PDMS mass results for peptides obtained by digesting the AS-48 protein with Lys-C protease (K), Glu-C protease (S) and chymotrypsin (Ch)

Peptide	Sequence	Mass (DA)	
		Calculated	Experimental
K1	E(58)-K(60)	388.5	388.6
K2	A(53)-K(56)	493.6	493.7
K3	R(65)-K(3)	1045.3	1061.7
K4	R(65)-K(3)	1045.3	1045.8
S1	I(59)-A(69)	1211.5	1211.9
S2	S(50)-Y(54)	580.7	581.3
S4	S(50)-E(58)	1079.3	1080.0
S5	I(59)-E(4)	1857.3	1874.0
S6 (a)	I(59)-E(4)	1857.3	1860.0
(b)	G(13)-E(20)	830.0	n.f.
S7 (a)	F(5)-A(12)	744.9	745.0
(b)	F(5)-G(13)	801.9	801.9
(c)	F(5)-T(14)	903.0	903.1
(d)	V(67)-E(4)	947.2	947.9
S10	F(5)-N(17)	1229.4	1230.4
Ch1	L(55)-K(62)	1014.0	1014.2
Ch2	L(55)-K(60)	757.8	759.2
Ch3	M(1)-F(5)	624.8	641.9
Ch4 (a)	A(44)-Y(54)	1136.3	1137.0
Ch5 (a)	T(33)-L(40)	660.7	662.1
(b)	A(44)-Y(54)	1136.3	1138.5
Ch6 (a)	M(1)-F(5)	624.8	626.0
(b)	T(33)-L(40)	660.7	n.f.
Ch8	K(61)-W(70)	1156.4	1157.5
Ch9 (a)	N(17)-W(24)	830.9	831.3
(b)	K(61)-W(70)	1156.4	1158.1
Ch10	G(63)-W(70)	900.0	900.9
Ch11	R(65)-W(70)	714.9	715.6
Ch12	G(6)-L(16)	968.2	968.7
Ch13	V(25)-L(32)	845.0	845.0
Ch14	N-terminal sequence chymotrypsin		
Ch15	G(6)-L(16)	968.2	n.f.

For the sequence-derived molecular mass calculation, the residual masses (monoisotopic values) of the individual amino acids were used. (n.f.: not found in mass spectrum).
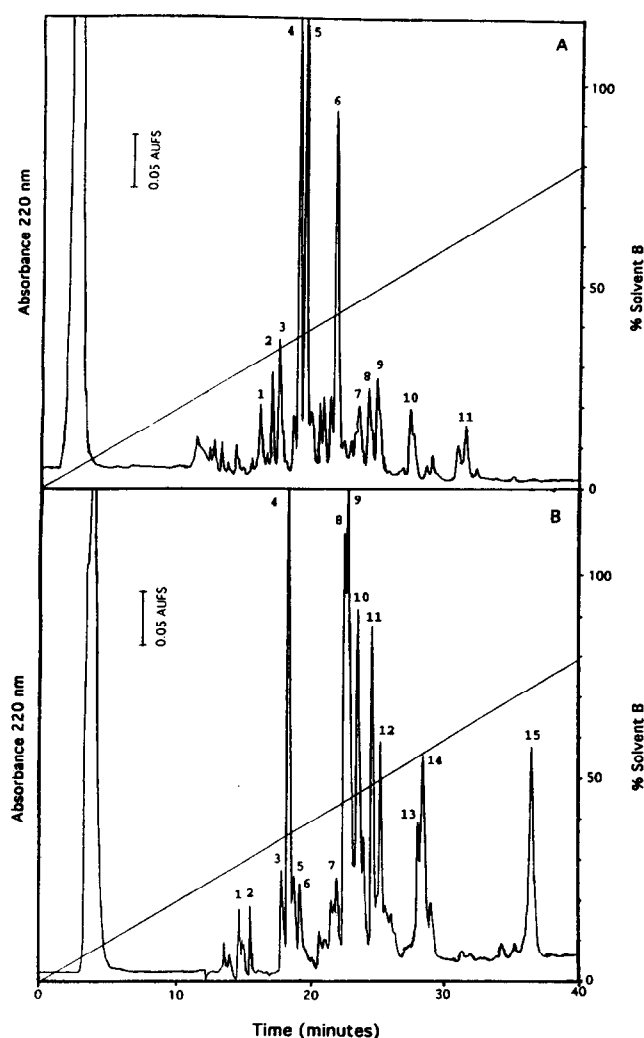

Fig. 2 Separation by reversed phase HPLC of the peptide mixtures obtained after digest of AS-48 with Glu-C endoproteinase (A) and chymotrypsin (B) The gradient profile of solvent B is indicated by the diagonal line. Numbers designate the resulting peptides.

The lowest experimental mass value, 7,150.17 Da, corresponds exactly to the calculated mass value if the dehydratation of one water molecule, necessary for the tail–head peptide-junction formation, is taken into account (calculated value: 7,149.57 Da). The second mass, 7,166.50 Da, with a mass difference of 16.33 amu to that of peak A, represents the oxidised AS-48 protein. A third peak, again with a mass difference of approximately 16 amu, indicates a second oxidation of the protein, most likely of tryptophan 70. Evidence herefore was obtained from the electrospray spectrum of peptide S5 which, apart from the mass of 1,873.26 corresponding to the oxidized peptide (which was also detected by PDMS) displayed an additional mass corresponding to an additional oxygen atom (result not shown).

Analysis of a sample of AS-48 to which myoglobin (16,951.50 Da) has been added as an internal standard, followed by calculation of the masses using the MaxEnt program [10] confirmed the mass differences between the species A and B, respectively, B and C, to be very close to 16.0 (effectively 15.95 and 16.86; data not shown). These data exclude the pos-
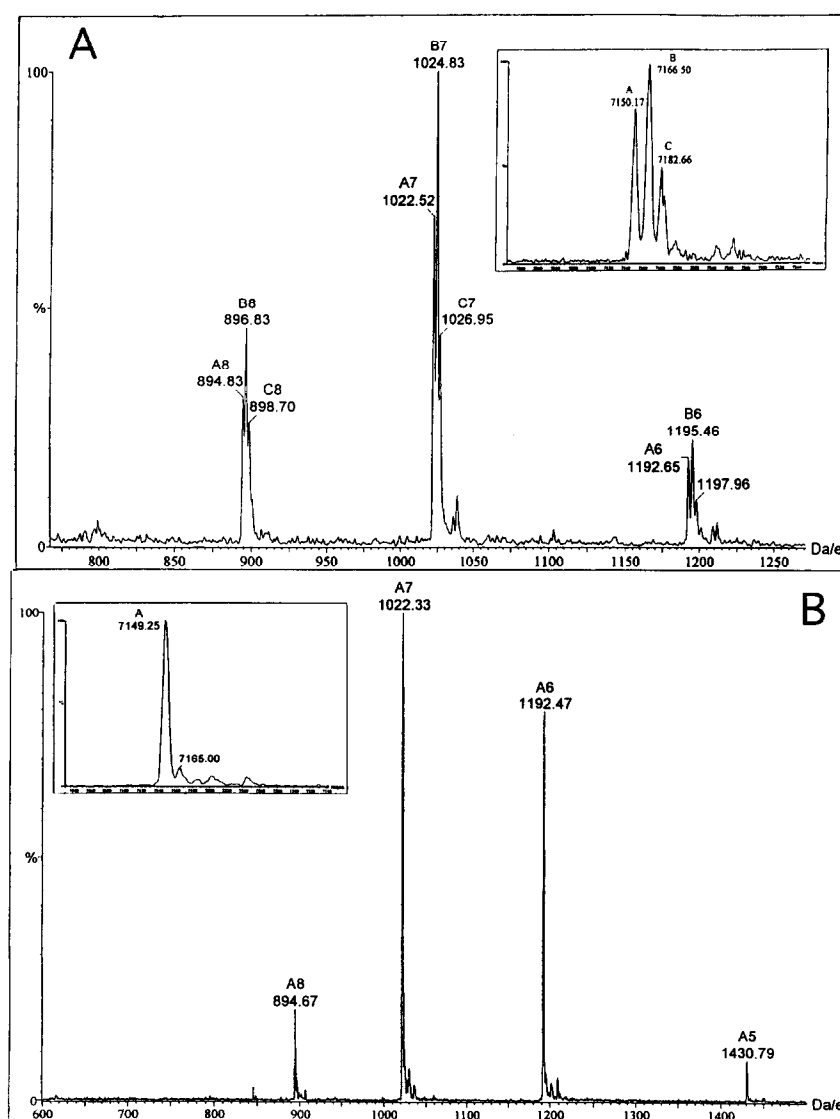
Fig. 3 Electrospray ionisation mass spectra of native antibiotic AS-48. The top spectrum is from the sample used for sequence determination. The bottom spectrum is from a freshly prepared sample. The insets are the converted spectra obtained with the Transform program.

sibility that species B would have arisen from ring opening of the cyclic structure during the molecular ion formation in the electrospray, a phenomenon which would cause a mass increase of 18 instead of 16 Da.

We should also add that we very recently, after the complete sequence determination, measured a new sample of AS-48 which was freshly prepared from new cells and found a mass of 7149.25 Da only (Fig. 3 bottom). This sample thus represents the cyclic form. It underwent only minor oxidation as indicated by the small peaks in the spectrum.

As far as we know, the cyclic structure of the AS48 antibiotic is the first example of a post-translational modification in which a cyclic structure arises from a 'tail-to-head' linkage of the gene product.

References

[1] Gálvez, A., Maqueda, M., Martinez-Bueno, M. and Valdivia, E. (1989) Res. Microbiol. 140, 57–68.

[2] Gálvez, A., Giménez-Gallego, G., Maqueda, M. and Valdivia, E. (1989) Antimicrob. Agents Chemother. 33, 437–441.

[3] Gálvez, A., Maqueda, M., Martinez-Bueno, M. and Valdivia, E. (1991) J. Bacteriol. 173, 886–892.

[4] Venema, K., Abee, T., Haandrikman, A.J., Leenhouts, K.J., Kok, J., Konings, W.N. and Venema, G. (1993) Appl. Environm. Microbiol. 59, 1041–1048.

[5] Gao, F.H., Abee, T. and Konings, W.N. (1991) Appl. Environm. Microbiol. 57, 2164–2170.

[6] Piard, J.-C., Kuypers, O.P., Rollema, H.S., Desmazeaud, M.J. and De Vos, W.M. (1993) J. Biol. Chem. 268, 16361–16368.

[7] Martinez-Bueno, M., Maqueda, M., Gálvez, A., Samyn, B., Van Beeumen, J., Coyette, J. and Valdivia E. (1994) J. Bacteriol., in press.

[8] Cotter, R.J. (1988) Anal. Chem. 60, 781A.

[9] Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M. (1989) Science 246, 64–71.

[10] Ferrige, A.G., Seddon, M.J., Green, B.N., Jarvis, S.A. and Skilling, J. (1992) Rapid Commun. Mass Spectr. 6, 707–711