ANANTIN — A PEPTIDE ANTAGONIST OF THE ATRIAL NATRIURETIC FACTOR (ANF)

II. DETERMINATION OF THE PRIMARY SEQUENCE BY NMR ON THE BASIS OF PROTON ASSIGNMENTS

D. F. WYSS[a,b], H.-W. LAHM[b], M. MANNEBERG[b] and A. M. LABHARDT[b,*]

[a] Dept. of Organic Chemistry, University of Basel,
St. Johann Ring 19, CH-4056 Basel, Switzerland
[b] Central Research Units, F. Hoffman-La Roche Ltd.,
Grenzacherstrasse 124, CH-4002 Basel, Switzerland

(Received for publication August 18, 1990)

Anantin, a naturally occurring peptide from *Streptomyces coerulescens*, binds competitively to the receptor of atrial natriuretic factor (ANF) from bovine adrenal cortex ($K_d = 0.6\,\mu M$) and acts as ANF antagonist. Protein chemical data and FAB-MS have identified anantin to be a cyclic polypeptide consisting of 17 common L-amino acids. The molecule is highly stable and precludes the application of standard sequencing methods. The primary sequence of anantin was determined by 2D ^1H NMR spectroscopy and the application of advanced protein chemical methods to be Gly[1]-Phe[2]-Ile[3]-Gly[4]-Trp[5]-Gly[6]-Asn[7]-Asp[8]-Ile[9]-Phe[10]-Gly[11]-His[12]-Tyr[13]-Ser[14]-Gly[15]-Asp[16]-Phe[17]. The molecule is cyclized between the β-carboxyl group of Asp[8] and the amino group of Gly[1].

In our ongoing screening program for atrial natriuretic factor (ANF)-like activities anantin, a cyclic peptide produced by *Streptomyces coerulescens*, has been identified. The preceding paper[1] describes the fermentation, isolation, physico-chemical data and biological properties. During characterization a peculiarity associated with this peptide was observed: Although it consists of 17 common L-amino acids its enzymatic digestion under standard conditions proved to be very difficult. When treated with HCl the peptide is perfectly stable for 25 hours in 30 mM HCl at 40°C or for 135 minutes in 10 M HCl at 40°C. Given its amino acid composition, this is highly unexpected for such a small peptide, and we decided to investigate the primary structure and the peptide's conformation by 2D NMR methods.

We present here nearly complete ^1H sequential assignments of anantin. These data have enabled us to determine the primary sequence of the peptide solely on the basis of the amino acid composition. This was achieved by extending the AMX spin systems to full residues using 2D-relayed coherence transfer spectroscopy (relayed COSY) and rotating frame NOE spectroscopy (ROESY) spectra prior to sequential assignments. Partial sequences obtained from chemical analyses confirm the NMR results.

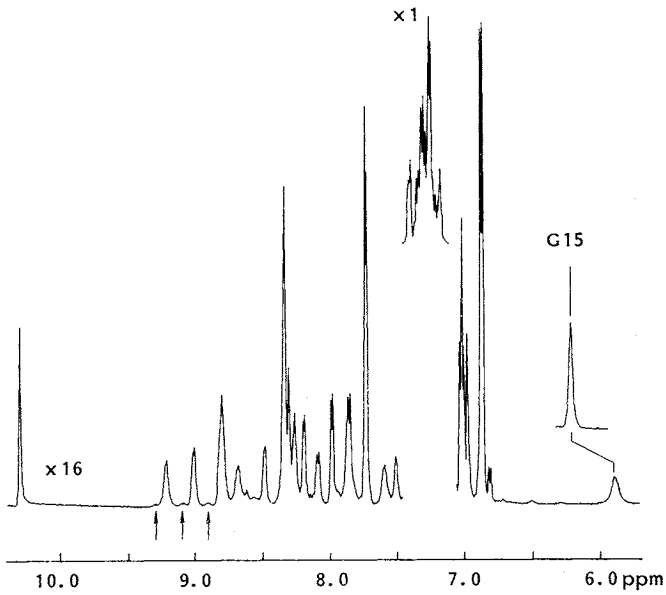
The 3D conformation of anantin determined by NMR in solution is presented elsewhere[2].

Results

General

Fig. 1 shows the lowfield region of the spectrum of anantin in CD_3OH, 285 K and an apparent pH of 4.6. The following features from this spectrum will be relevant in the context of the data presented here: (1) In addition to the main component the existence of a minor variant of anantin is obvious. This variant is in slow exchange with the major form and at 320 K it shows positive exchange peaks in 2D

Fig. 1. Downfield region of the 1D spectrum of anantin in CD_3OH, 285 K and pH 4.6.



The inset on the right refers to 320 K. The arrows point to NH resonances of a minor variant of anantin (see text for details).

ROESY spectra (data not shown). This interconversion will be discussed in detail elsewhere. (2) The broad resonance labeled G15† will be shown to be the NH resonance of glycine 15. It is exchange broadened and gets considerably sharper at 320 K (compare inset). (3) The region from 7.19 to 7.37 ppm contains the phenyl ring resonances. Its integral suggests the occurrence of about 16 protons corresponding to the ring protons of 3 phenylalanines and one NH (compare F10 in Table 1). The two-quantum filtered COSY (2QF COSY) spectrum (*c.f.* Fig. 3A) reveals that no other aromatic protons resonate in this region.

The Basic Spin Systems

The amino acid composition of anantin consists of 5 glycines, 2 isoleucines, 3 phenylalanines, 2 aspartic acids and one of each asparagine, histidine, serine, tyrosine and tryptophan. We hence expect 5 glycine spin systems, 2 long side chain and 10 AMX spin systems. The fingerprint region of a 2QF COSY is shown in Fig. 2. All expected (NH, $C^{\alpha}H$) cross peaks are seen. The five glycines are identified by their dual cross peaks. Labeling refers to the final sequential assignments†† and is not yet obvious from the COSY spectrum except for the cross peak labeled S14 which was tentatively assigned to belong to the single serine on the basis of the typical low field β-proton chemical shifts (*c.f.* Table 1). For two (NH, $C^{\alpha}H$) peaks (labeled I3 and I9 in Fig. 2) *J*-connectivities are readily extended to the side chain protons and clearly identify two isoleucine residues (*c.f.* Fig. 3B). The aromatic part of the 2QF COSY (*c.f.* Fig. 3A) further documents the occurrence of the expected histidine, tryptophan and tyrosine ring systems in addition to the 3 phenylalanine ring systems mentioned above (*c.f.* Fig. 1).

† The one letter code for the amino acids residues is used.

†† For sake of clarity spin systems are consistently labeled with the sequence number and the residue type (part of) which they constitute.

Table 1. ^1H Chemical shifts of anantin and des-phe-anantin, determined in CD_3OH, at 285 K and the apparent pH of 4.6.

(A) Anantin:

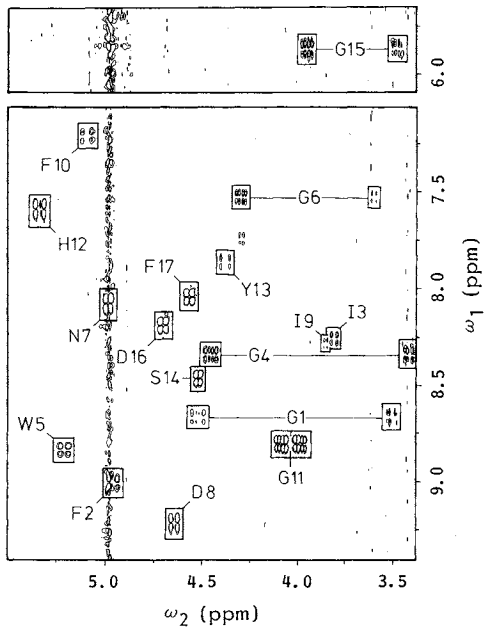
Sequence	Chemical shifts								
	NH	αH	αH′	βH	βH′	Others			
G1	8.68	4.515	3.52						
F2	9.005	4.96		3.44	3.15	QR1	7.31		
I3	8.26	3.825		1.745					
G4	8.34	4.435	3.43						
W5	8.81	5.20		3.32	3.12				
G6	7.505	4.285	3.595						
N7	8.085	4.98		3.00	2.68	δNH	7.725	δNH′	6.96
D8	9.215	4.645		3.58	2.63				
I9	8.305	3.86		1.575					
F10	7.20	5.09		3.745	2.62	QR3	7.23		
G11	8.80	4.09	3.98						
H12	7.59	5.33		4.215	2.49	2H	8.345	4H	7.17
Y13	7.86	4.375		2.99	2.83	2,6-H	7.39	3,5-H	6.87
S14	8.49	4.515		3.94	3.725				
G15	5.885	3.95	3.49						
D16	8.19	4.68		2.76	2.625				
F17	7.98	4.54		3.19	3.035	QR2	7.25		
I3	γH	γH′	γH_3	δH_3					
	1.56	1.10	0.86	0.875					
I9	γH	γH′	γH_3	δH_3					
	0.96	0.875	0.31	0.685					
W5	N1H	2H	4H	5H	6H	7H			
	10.30	7.255	7.73	7.015	7.165	7.395			

(B) Des-phe-anantin:

Sequence	Chemical shifts				
	NH	αH	αH′	βH	βH′
G1	8.54	4.44	3.51		
F2	8.98	4.89		3.38	3.15
I3	8.22	3.71		1.82	
G4	8.28	4.42	3.33		
W5	8.88	5.27		3.30	3.08
G6	7.56	4.31	3.62		
N7	7.99	4.96		3.00	2.64
D8	9.13	4.60		3.52	2.59
I9	8.13	3.84		1.57	
F10	blch	blch		3.72	2.66
G11	8.785	4.07	3.99		
H12	7.55	5.33		4.24	2.50
Y13	7.91	4.33		3.02	2.83
S14	8.36	4.51		3.92	3.74
G15	6.17	4.07	3.54		
D16	8.11	4.65		2.79	2.62
I3	γH	γH′	γH_3	δH_3	
	1.56	1.10	0.86	0.89	
I9	γH	γH′	γH_3	δH_3	
	1.00	0.88	0.37	0.70	

The one letter code for the amino acids is used; the sequence numbering corresponds to the sequential assignments (*cf.* Figs. 7 and 5). The abbreviations «QR1», «QR2» and «QR3» in column 7 stand for phenyl ring protons. They have not been assigned to individual ring positions. «blch» identifies resonances that were bleached due to presaturation of the OH line of methanol.

Fig. 2. Fingerprint region of a 2QF COSY spectrum of anantin in CD_3OH, pH 4.6, 285 K.



The (NH, $C^{\alpha}H$) J cross peaks of all 17 residues are seen and are labeled using the one letter codes for the amino acids. The numbering refers to the primary peptide sequence as deduced from the sequential assignments (*c.f.* Figs. 5 and 7).

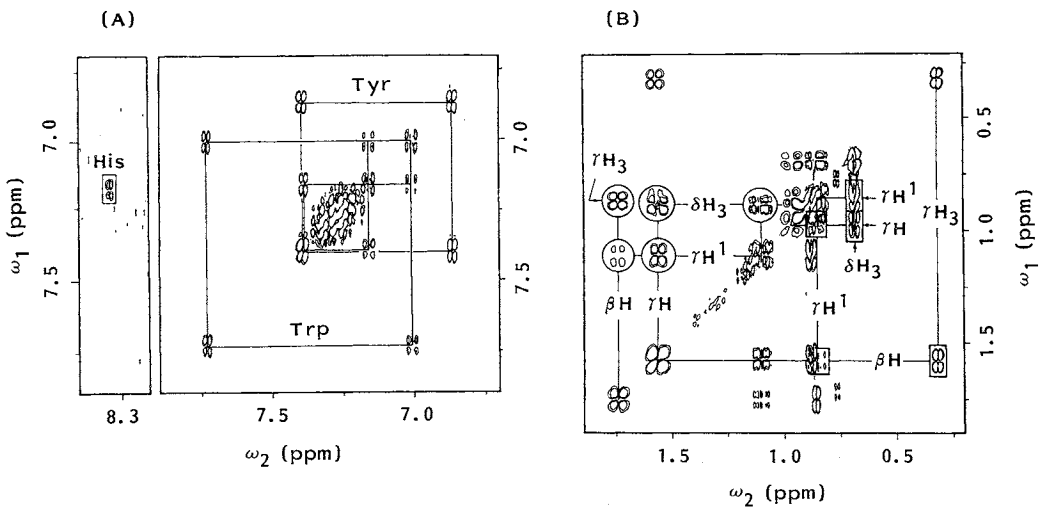
Identification of Amino Acid Residues

Relayed COSY and/or through space ROESY connectivities were observed between several of the AMX and their side chain spin systems. The spectra are not shown; the results are summarized in Fig. 4. Due to the existence of relayed COSY cross peaks (dashed lines in Fig. 4) the extension of the spin systems labeled F17 and W5 is unambiguous. The same holds true for the AMX system labeled N7 which is the only one that has a ROESY cross peak to the single NH_2. The AMX systems labeled Y13 and H12 are the most probable candidates for the tyrosine and histidine respectively as judged from the number and intensity of their cross peaks to the corresponding aromatics. In the range from 7.19 to 7.37 ppm the overlapping phenylalanine ring resonances were not assigned to individual rings and hence the two AMX systems labeled F2 and F10 were extended to phenylalanines assuming that the observed ROESY cross peaks would pertain to separate rings.

Sequential Assignments

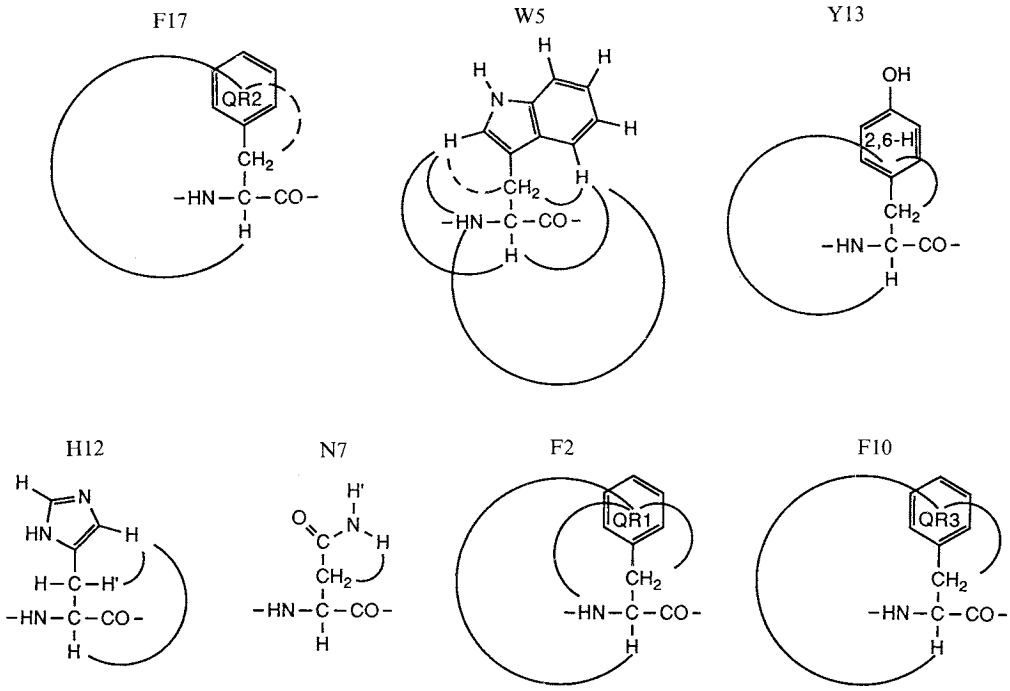
A set of sequential contacts is shown in Fig. 5.

Fig. 3. Partial side chain spin systems of anantin: Aromatic resonances for a single His, Tyr, Trp (Fig. 3A) and the side chain resonances for two Ile's (Fig. 3B) are apparent from 2QF COSY spectra.



Conditions are the same as in Fig. 2. The cross peaks referring to Ile 3 are marked by circles above the diagonal and those of Ile 9 by boxes below the diagonal (numbering refers to the sequential assignments, *c.f.* Figs. 5 and 7).

Fig. 4. Survey of the extension of 7 of the AMX spin systems to full amino acid residues on the basis of ^1H relayed COSY and ROESY (dashed lines) or ^1H ROESY (full lines) connectivities (spectral data not shown).



Lines pointing to CH_2 groups indicate that cross peaks to both methylen protons have been seen. The proton(s) responsible for the cross peaks to the phenylalanine rings (F17: at 7.25 ppm; F2 at 7.31 ppm; F10 at 7.23 ppm) were not individually assigned (see text for details). The residues are labeled with the one letter codes. Numbering refers to the peptide primary sequence as obtained from the sequential assignments (*c.f.* Figs. 5 and 7). The meaning of the abbreviations QR1, QR2, QR3 and 2,6-H are explained in Table 1.

Following the ($C^{\alpha}H^{i}$, NH^{i+1}) cross peaks the entire chain can be traced from G1 to F17. Several of the sequential connectivities are also evident from ($C^{\beta}H^{i}$, NH^{i+1}) cross peaks. The contact S14→G15, which is missing in Fig. 5, has been observed in other spectra with shorter mixing time ($\tau_m = 60$ mseconds) or at elevated temperature when the exchange broadened NH of G15 sharpens (compare Fig. 1).

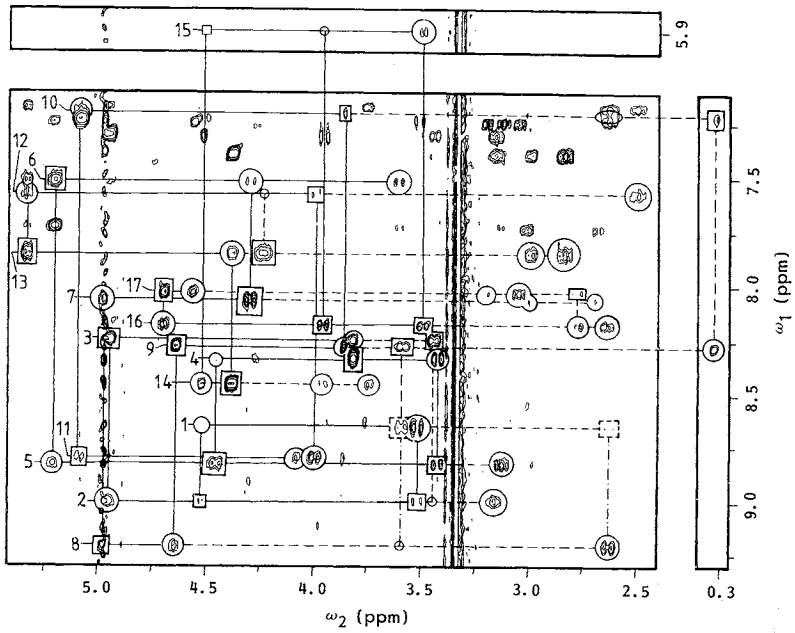
Particular attention should be drawn to the ROEs observed on the NH of G1 which are highlighted in Fig. 5 by dashed boxes. They indicate close proximity between G1 and the β-protons of the AMX spin system at position 8 (Fig. 6).

Table 1 lists the ^1H chemical shifts of all assigned protons in the order of the primary sequence of anantin (under (A)) and of des-phe-anantin (under (B), spectra of des-phe-anantin are not shown). The comparison of (B) with (A) reveals that it is the *C*-terminal F17 which is missing in des-phe-anantin.

Protein Chemistry

Initially *N*-terminal sequence analysis was performed by automated Edman degradation but it did not yield any amino acid sequence information. Due to the bacterial origin of anantin a modification of its terminal α-amino group was suspected. Indeed, amino acid analysis and FAB-MS established the cyclic nature of the peptide assembled entirely from common L-amino acids (*c.f.* Fig. 8). To obtain sequenceable

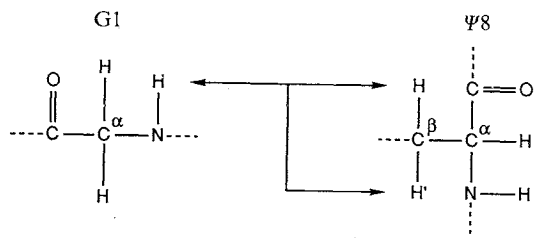
Fig. 5. Fingerprint region of a ROESY spectrum ($\tau_m = 120$ mseconds) of anantin.



Conditions are the same as in Fig. 2. The frequency positions of intraresidue COSY and relayed COSY cross peaks due to *J* coupling (compare Table 1) are marked by circles in the ROESY spectrum. Most but not all of these peaks are also present as through space ROESY connectivities. Cross peaks representing sequential connectivities are boxed. Sequential ($C^{\alpha}H^{i}$, NH^{i+1}) connectivities are marked by full lines, ($C^{\beta}H^{i}$, NH^{i+1}) and the ($C^{\gamma}H_3{}^{9}$, NH^{10}) connectivities by dashed lines. The entire sequence of anantin can be traced as a linear chain using the sequential ($C^{\alpha}H^{i}$, NH^{i+1}) connectivities in Fig. 5, with the exception of position i = 14 (this latter connectivity has been observed for $\tau_m = 60$ mseconds (data not shown)). No contact was ever observed between F17 and G1. The dashed boxes on the NH of G1 mark two "sequential" contacts to residue D8; anantin is cyclised between G1 and D8 (see Discussion and Fig. 8).

fragments several hydrolysis techniques were used. (1) Limited alkaline hydrolysis yielded several fragments which could be isolated by RP-HPLC and sequenced (Fig. 7). (2) Oxidative cleavage at the *C*-terminus of tryptophan with *N*-chlorosuccinimide resulted in generating a further fragment amenable to sequencing (Fig. 7). The latter experiment was repeated after the free carboxyl groups had been modified with 2-diethylaminoethylamine. During sequencing PTH-aspartic acid-β-(2-diethylaminoethylamide) corresponding to D16 could be identified. The amino acid at sequence position 8 remained unaffected, indicating that D16, but not D8, has a free carboxyl group.

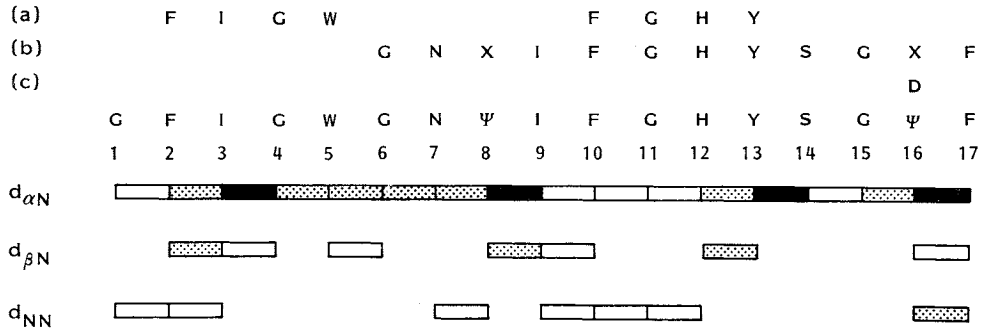
Fig. 6. Schematic presentation of the ROEs observed between the NH of G1 and the two β-protons of the AMX spin system 8 (abbreviated Ψ8).



Discussion

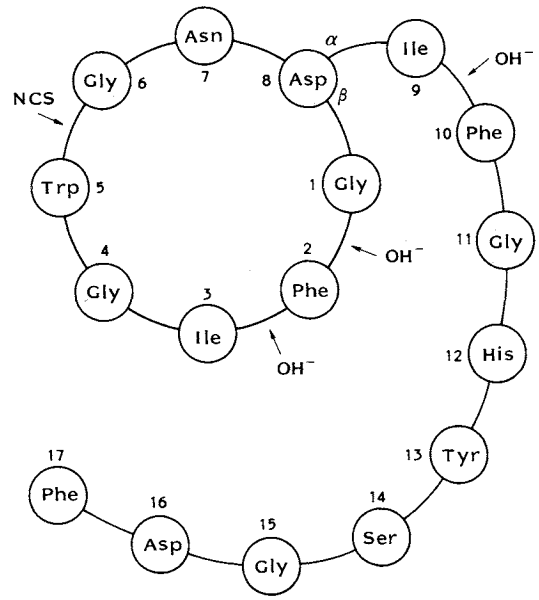
In summary, except for two AMX systems, all spin systems were assigned to full residues. The

Fig. 7. Survey of the sequential connectivities of anantin as determined by ^{1}H NMR and comparison to the fragment sequences obtained by chemical analyses.



The fragment sequences in row (a) were analyzed after hydrolysis of anantin, the one in row (b) after *N*-chlorosuccinimide oxidation. The one letter code for the amino acid residues is used. X stands for residues that could not be assigned during amino acid sequencing; using 2-diethylaminoethylamine to label free carboxyl groups the X in position 16 was shown to be an aspartic acid residue with a free β-carboxyl group [row (c)]. Ψ stands for AMX spin systems. The rows labeled $d_{\alpha N}$, $d_{\beta N}$ and d_{NN} summarize the sequential ($C^{\alpha}H^{i}$, NH^{i+1}), ($C^{\beta}H^{i}$, NH^{i+1}) and (NH^{i}, NH^{i+1}) cross peaks. Filled boxes refer to strong, shaded boxes to intermediate and open boxes to weak ROEs.

Fig. 8. Primary structure of anantin.



	A	B
Trp	1.02	1.2
His	1.01	1.0
NH_3	1.49	—
Asp	3.00	3.0
Ser	0.93	1.1
Gly	5.00	4.8
Ile	1.95	2.0
Tyr	1.03	0.8
Phe	3.00	2.5

Monoisotopic protonated molecular ion: obsd, 1,870.89; calcd, 1,870.81[c].

The amino acid chain is numbered from *N*-terminus to *C*-terminus with the amino group of Gly^1 connected to the β-carboxyl group of Asp^8. Major cleavage points are indicated by arrows; NCS: *N*-chlorosuccinimide; OH^-: sodium hydroxide. The amino acid composition is given as determined by (A) postcolumn ninhydrin reaction, and (B) precolumn (+)-1-(9-fluorenyl)ethylchloroformate reaction (L-amino acids only). The calculation of the molecular ion from FAB-MS analysis (c) is based on a cyclic peptide.

comparison with the amino acid composition shows that both of these unidentified AMX spin systems (referred to as Ψ) belong to aspartic acids residues. When combining the ROESY connectivities observed for different τ_m's and at different temperatures, all 17 residues can be connected as a linear chain. The cumulated evidence including (NH, NH) contacts (data not shown) are summarized in Fig. 7. The figure

moreover compares the NMR results to the partial sequences obtained from chemical analyses. The tentative NMR side chain assignments of S14, F2 and F10 are fully confirmed and Ψ16 is reconfirmed to be an aspartic acid residue.

No ROESY cross peaks have ever been observed between F17 and G1. This is the reason for choosing to start the numbering at "G1". On the other hand, G1 cannot be a free *N*-terminal residue since its NH is seen by NMR (the free *N*-terminal NH_3^+ would exchange rapidly with the solvent and be bleached and hence disappear from the spectrum) and close proximity of this NH to the β-protons of D8 is observed (*c.f.* Fig. 6). In fact, mass spectrometry (FAB-MS) had shown[1] that the MW of anantin corresponds to its amino acid composition minus the weight of one water molecule. Together with the negative Edman degradation data this implies that the molecule must be cyclized. Three potential sites for cyclization with G1 exist: The β-carboxyls of D8 or D16 or the α-COOH of F17. The latter can be ruled out: Des-phe-anantin is cyclic[1]; Table 1 demonstrates that it lacks residue F17. Similarly D16 can be ruled out since its β-carboxyl group is free to be labeled [*c.f.* Fig. 7, row (c)]. This leaves D8 as sole possibility, in agreement with the expectation from the ROESY data which prove close spatial proximity between the NH of G1 and the β-CH_2 group of D8 (compare Figs. 5 and 6). The resulting complete structure of anantin is depicted in Fig. 8.

A total of 112 ROEs have been assigned and quantified for anantin. The full solution conformation/structure as determined from these ROE measurements and by distance geometry calculations will be presented elsewhere[2].

Experimental

NMR

All NMR spectra were measured on a Bruker AM-500 spectrometer. Anantin was dissolved in deuterated methanol (CD_3-OH). The remaining methanol OH resonance was eliminated from the spectrum by presaturation. The NH exchange kinetics between the backbone of the peptide and the OH of the solvent at apparent pH ≤ 6 were sufficiently slow to enable observation of the amide protons. Spectra were measured between pH 6 and 3.2 and temperatures between 280 K and 320 K. It was observed that parts of the molecule, depending on the apparent pH and the temperature, would go into intermediate exchange leading to loss of some of the resonance lines from the spectrum.

2QF COSY[3,4], relayed COSY[5] and ROESY[6~8] spectra were measured using the TPPI method. Routinely 2 K real time data points and 650 parametrically incremented experiments were collected per 2D spectrum. In the ROESY spectra 1.3 K of the real time data points, in all other spectra 2K data points were used for the 2D transformations after zero filling to 4 K·2 K. Apodization involved multiplication with shifted sine- and squared sine-bells. ROESY mixing times of 60, 120, 200 and 400 mseconds were used. The delay in the relayed COSY spectra was tuned for 8.3 Hz ^{1}H-^{1}H coupling constants.

Protein Chemistry

Amino acid analyses were performed applying two independent methods. Conventional peptide hydrolysis using methane sulfonic acid/0.2% 3-(2-aminoethyl)indole (Pierce; 4 M)[9] at 110°C for 24 hours *in vacuo* or microwave hydrolysis for 3 minutes[10]. Hydrolysates were analyzed on a Liquimat III amino acid analyzer (Kontron) employing ion exchange chromatography and postcolumn reaction with ninhydrin. Precolumn derivatization with (+)-1-(9-fluorenyl)ethylchloroformate followed by RP-HPLC was applied to distinguish between L- and D-amino acids[11].

Amino acid sequence analyses were carried out on a gas phase sequenator model 470A (Applied Biosystems). The resulting PTH-amino acids were separated isochratically on a Supersphere RP-8 column (250 × 1.6 mm; 4 μm; MZ-Analysentechnik, Mainz)[12].

Fragments of anantin were generated by partial hydrolysis with 1 M sodium hydroxide at 100°C for 30 minutes. Chemical cleavage was performed with 5 nmol anantin dissolved in 0.5 M acetic acid-DMF and treatment with a 15-fold molar excess of *N*-chlorosuccinimide for 3 hours[13]. The resulting reaction products were separated by RP-HPLC and analyzed by amino acid analysis and Edman degradation. Free carboxylic groups were amidated by activation with 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU)[14] and reaction with 2-diethylaminoethylamine.

Acknowledgments

We thank URS RÖTHLISBERGER for excellent technical assistance and AGOSTINO FEDE for help in preparing some of the figures.

References

1) WEBER, W.; W. FISCHLI, E. HOCHULI, E. KUPFER & E. K. WEIBEL: Anantin—A peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity. J. Antibiotics 44: 164~171, 1991
2) WYSS, D. F.; D. GILLESSEN, E. KUPFER, A. TRZECIAK & A. M. LABHARDT: A novel motive of peptide chain folding: The solution conformation of anantin as determined by NMR. NATURE, to submitted
3) RANCE, M.; O. W. SØRENSEN, G. BODENHAUSEN, G. WAGNER, R. R. ERNST & K. WÜTHRICH: Improved spectral resolution in COSY ^{1}H NMR spectra of proteins via double quantum filtering. Biochem. Biophys. Res. Commun. 117: 479~485, 1983
4) NEUHAUS, D.; G. WAGNER, M. VÁSÀK, J. H. R. KÄGI & K. WÜTHRICH: Systematic application of high resolution, phase sensitive ^{1}H NMR techniques for the identification of the amino acid proton spin systems in proteins: Rabbit metallothionein-2. Eur. J. Biochem. 151: 257~273, 1985
5) WAGNER, G.: Two-dimensional relayed coherence spectroscopy of a protein. J. Magn. Reson. 55: 151~156, 1983
6) BRAUNSCHWEILER, L. & R. R. ERNST: Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. J. Magn. Reson. 53: 521~528, 1983
7) BAX, A. & D. G. DAVIS: Practical aspects of two-dimensional transverse NOE spectroscopy. J. Magn. Reson. 63: 207~213, 1985
8) RANCE, M.: Improved techniques for homonuclear rotating-frame and isotropic mixing experiments: Overview and improvements for TOCSY, CAMELSPIN and ROESY. J. Magn. Reson. 74: 557~564, 1987
9) MOORE, S.: The precision and sensitivity of amino acid analysis. *In* Chemistry and Biology of Peptides. *Ed.*, J. MEIENHOFER, pp. 629~653, Ann Arbor Publishers, 1972
10) LAHM, H.-W.; W. LERGIER, M. MANNEBERG & R. KNORR: An improved method for fast amino acid analysis. J. Protein Chem. 7: 258, 1988
11) EINARSSON, S.; B. JOSEFSSON, P. MÖLLER & D. SANCHEZ: Separation of amino acid enantiomers and chiral amines using precolumn derivatisation with (+)-1-(9-fluorenyl)ethyl chloroformate and reversed-phase liquid chromatography. Anal. Chem. 59: 1191~1195, 1987
12) LOTTSPEICH, F. J.: Microscale isochratic separation of phenylthiohydantoin amino acid derivatives. J. Chromatogr. 326: 321~327, 1985
13) SHECHTER, Y.; A. PATCHORNIK & Y. BURSTEIN: Selective chemical cleavage of tryptophanyl peptide bonds by oxidative chlorination with *N*-chlorosuccinimide. Biochemistry 15: 5071~5075, 1976
14) KNORR, R.; A. TRZECIAK, W. BANNWARTH & D. GILLESSEN: New coupling reagents in peptide chemistry. Tetrahedron Lett. 30: 1927~1930, 1989