PROTEINS: Structure, Function, and Genetics 20:356–366 (1994)

A Model for the Structure of a Homodimeric Prohormone: The Precursor to the Locust Neuropeptide AKH I

T.J. Horne,[1] D.G. Doak,[1] R.C. Rayne,[2] G. Balacco,[1] M. O'Shea,[2] and I.D. Campbell[1]
[1]*Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, England; and* [2]*Sussex Centre for Neuroscience, School of Biological Sciences, University of Sussex, Brighton, BN1 9QG, England*

ABSTRACT **We have determined the structure in solution of a homodimeric protein that is a precursor to the locust neuropeptide adipokinetic hormone I using nuclear magnetic resonance spectroscopy. This precursor, called P1, is comprised of two 41 residue strands joined by a single inter-chain disulphide at Cys39. We have also determined the structure of an end product of P1 processing, called APRP1; this is a homodimer comprised of residues 14–41 of P1. Nuclear Overhauser Effect (nOe) data indicate that in both P1 and APRP1, residues 22–37 (numbered with respect to P1) form pairs of α-helices, with no evidence for any other secondary structure.**

Key words: NMR, structure determination, coiled coil

INTRODUCTION

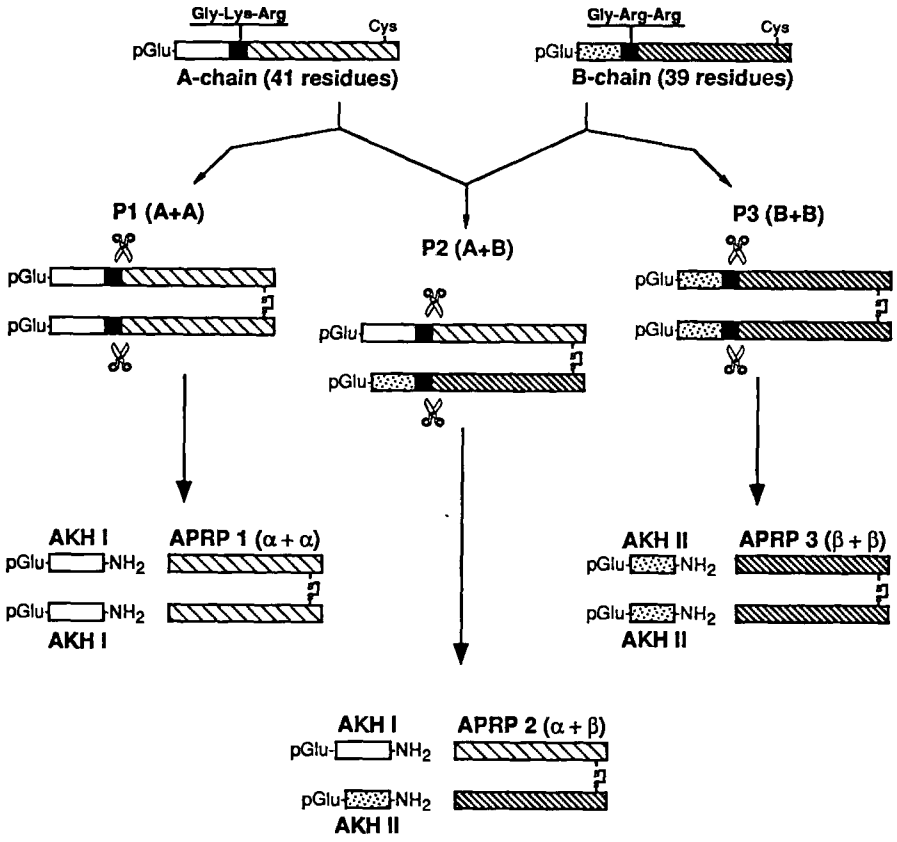
To enable locusts to use lipids as fuel during lengthy flights, the insect releases hormones that facilitate mobilization of diacylglycerols from the fat body into the circulation.[1,2] This mobilization arises from the action of two different but related neuropeptides, adipokinetic hormones I and II (AKH I and AKH II). In the desert locust, *Schistocerca gregaria,* these hormones are produced and stored in the glandular corpora cardiaca (CC), major neuroendocrine centers in the insect nervous system.[3,4]

The biosynthesis of AKH I and II involves the enzymatic processing of larger, precursor polypeptides[5]; the hormones are the N-terminal sequences of two independently synthesized prohormones, referred to as the A-chain and the B-chain, respectively (Fig. 1). These prohormones are not processed directly to yield AKH I and AKH II; rather, cysteine residues towards the C-termini are oxidized to form disulphide bonds, producing three dimeric precursors, P1 (A + A), P2 (A + B), and P3 (B + B). The hormones, like most bioactive peptides, are liberated from precursors by proteolysis C-terminal to pairs of basic amino acid residues[6–6]: enzymatic cleavage of the precursors, followed by further processing, yields the hormones, and leaves the AKH precursor-related peptides (APRP1, -2, or -3) (Fig. 1).

The endoproteolysis of precursors to yield active peptides is, in general, specific and limited, and the presence of a pair of basic residues within a precursor sequence is a necessary, but not sufficient, condition for cleavage. This has led to the proposal that structural motifs in the vicinity of potential cleavage sites act as recognition elements that determine whether cleavage will occur.[9,10] Efforts to determine the tertiary structure of prohormones as a test of this hypothesis are, however, hindered by the difficulty in obtaining quantities of prohormones large enough for biophysical studies; to date, proinsulin is the only prohormone to be studied by nuclear magnetic resonance (NMR).[11] The relatively small size of the prohormone of AKH I (41 residues) has allowed us to synthesize sufficient quantities of the dimeric prohormone precursor P1 to afford structural studies by NMR.

Both P1, and the end product of P1 processing, APRP1, are homodimers, and as such the calculation of their tertiary structure by standard NMR methods is potentially problematic. This is because of the difficulties caused by the inherent symmetry: a direct consequence of this symmetry is that, a priori, it is impossible to distinguish between intra- and inter-monomer Nuclear Overhauser Effects (nOes). However, several homodimers have been studied by NMR, among them the leucine zipper regions of GCN4[12,13] and Jun,[13] and the *Trp* repressor.[14] Various approaches have been used to overcome the ambiguities in nOe assignments: Saudek et al.[12] assumed that inter-monomer nOes would be weak, and thus in short mixing time NOESY spectra, only intra-monomer connectivites would be observed. O'Donoghue et al.[13] used models based on the crystal structure of GCN4[15] and coiled-coil

Received May 24, 1994; revision accepted August 26, 1994.
Address reprint requests to Dr. I.D. Campbell, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 1QU, U.K.
G. Balacco's present address is A. Menarini Industrie Farmaceutiche Riunite S.r.I., Via Setti Santi, 3-50131 Firenze, Italy.



ELNFTPNWGTGKRDAADFGDPYSFLYRLIQAEARKMSGCSN
ELNFTPNWGTGKRDAADFGDPYSFLYRLIQAEARKMSGCSN P1

DAADFGDPYSFLYRLIQAEARKMSGCSN
DAADFGDPYSFLYRLIQAEARKMSGCSN APRP1

DAADFGDPYSFLYRLIQAEARKMSGCSN
YADPNADPMAFLTKLIQIEARKLSGCSN APRP2

Fig. 1. AKH I, AKH II, and APRP biosynthesis from three dimer precursors P1, P2, and P3. See reference 5 for additional details. The sequences of the three dimers studied in this work, P1, APRP1, and APRP2, are also shown.

models of the leucine zipper regions of GCN4 and Jun, to classify nOes as either intra- or inter-monomer. Arrowsmith et al.[14] studied several isotopomers of the *Trp* repressor, each with a different pattern of selective deuteriation, to delineate intra- and inter-monomer nOes. Prior to the structure calculations described in this paper, P1 and APRP1 were modeled using a modified version of a protocol developed by Nilges and Brünger.[16–18] These models were used to assess the likelihood of observing inter-monomer nOes; in addition, inter-monomer nOes were unambiguously assigned from the NOESY spectrum of a heterodimer closely related to APRP1, called APRP2. APRP2 consists of two chains, denoted α and β, disulphide bonded at Cys39, like APRP1. The α-chain of APRP2 is identical to those of APRP1, and the β-chain is 64% identical. The sequences of P1, APRP1, and APRP2 are shown in Figure 1.

MATERIALS AND METHODS

Peptide Synthesis and Purification

All monomers were produced by chemical synthesis (Peninsula Laboratories, St. Helens, U.K.), and the sequences confirmed using a Beckmann LF 3000 automated protein sequencer. Oxidation of monomers to dimers was accomplished using a glutathione-based oxido-shuffling system.[19] Dimers were purified by reverse phase high-performance liquid chromatography (HPLC), as described elsewhere.[20] Electro-spray mass spectroscopy was used

TABLE I. ^{1}H Chemical Shifts at 310 K and pH 3.2 of P1

	NH	αH	βH	γH	δH	Others
pGlu1	7.73	4.29	2.49, 2.01	2.42		
Leu2	8.16	4.29	1.55	1.45	0.84, 0.82	
Asn3	8.23		2.69, 2.65			7.41 6.75
Phe4	7.99		3.02, 2.94			7.23 (3,5)
						7.12 (2,6)
Thr5	7.88	4.49	4.04	1.12		
Pro6		4.17	1.92, 1.46	1.78	3.52	
Asn7	8.14		2.74, 2.68			7.45, 6.79
Trp8	7.86		3.28, 3.20			7.20 (H2)
						7.55 (H4)
						7.05 (H5)
						7.12 (H6)
						7.41 (H7)
						10.03 (NH)
Gly9	8.14	3.91, 3.86				
Thr10	7.91	4.31	4.26	1.25		
Gly11	8.33	4.01, 3.97				
Lys12	8.02	4.25	1.79	1.39, 1.33	1.60	3.00 (εCH_2)
Arg13	8.21	4.24	1.79, 1.71	1.58, 1.51	3.09	7.09, 6.60
Asp14	8.20	4.66	2.88, 2.81			
Ala15	8.09	4.17	1.35			
Ala16	8.00	4.19	1.28			
Asp17	7.95	4.64	2.77, 2.70			
Phe18	7.88	4.52	3.11, 2.97			7.23 (3,5)
						7.15 (2,6)
Gly19	8.07	3.87				
Asp20	7.99	4.92	2.98, 2.83			
Pro21		4.25	2.02, 1.77	1.62, 1.56	3.78, 3.70	
Tyr22	7.72	4.23	2.98			7.03 (2,6)
						6.75 (3,5)
Ser23	7.85	4.22	4.02, 3.97			
Phe24	7.93	4.22	3.19, 3.04			7.26 (3,5)
						7.09 (2,6)
Leu25	7.94	3.88	1.76, 1.70	1.65	0.91, 0.85	
Tyr26	8.08	4.05	3.06, 2.98			6.92 (2,6)
						6.69 (3,5)
Arg27	7.85	3.91	1.88	1.81, 1.66	3.15	7.18, 6.67
Leu28	7.93	3.97	1.57	1.48	0.72	
Ile29	8.20	3.48	1.87	1.85, 0.99	0.76	0.78 (γCH_3)
Gln30	8.02	3.83	2.02	2.17, 2.01		6.86, 6.68
Ala31	7.99	4.06	1.47			
Glu32	8.16	4.13	2.03	2.44, 2.36		
Ala33	8.48	3.98	1.42			
Arg34	7.88	4.06	1.89	1.80, 1.65	3.19	7.22, 6.68
Lys35	7.65	4.19	1.95, 1.88	1.52, 1.42	1.68	2.98 (εCH_2)
						6.68
Met36	7.87	4.31	2.13, 2.05	2.64, 2.54		2.08 (εCH_3)
Ser37	7.87	4.38	3.97			
Gly38	8.12	4.08, 3.95				
Cys39	8.17	4.74	3.31, 2.96			
Ser40	8.33	4.48	3.88			
Asn41	8.19	4.69	2.90, 2.83			7.48, 6.80

to confirm that the purified dimers had the correct molecular mass.

NMR and CD Spectroscopy

NMR spectra were recorded on samples of 1–2 mM solution of the homodimers P1 and APRP1, and the heterodimer APRP2, in either 90% D_2O/10% CD_3CN or 80% H_2O/10% D_2O/10% CD_3CN at a pH of 3.2 (uncorrected glass-electrode reading). Acetonitrile was required because P1 is not soluble at mM concentrations in H_2O/D_2O alone, and for consistency, spectra of the other peptides were also ob-

TABLE II. ^{1}H Chemical Shifts at 310 K and pH 3.2 of APRP1

	NH	αH	βH	γH	δH	Others
Asp14			2.92			
Ala15	8.60	4.26	1.33			
Ala16	8.10	4.14	1.22			
Asp17	8.01	4.57	2.73, 2.68			
Phe18	7.93	4.48	3.05, 2.93			7.20 (3,5) 7.10 (2,6)
Gly19	8.08	3.83				
Asp20	8.00	4.88	2.92, 2.73			
Pro21		4.19	1.99, 1.74	1.58, 1.51	3.71, 3.67	
Tyr22	7.72	4.20	2.93			6.98 (2,6) 6.68 (3,5)
Ser23	7.82	4.16	3.98, 3.92			
Phe24	7.90	4.21	3.16, 2.99			7.13 (3,5) 7.06 (2,6)
Leu25	7.87	3.82	1.68	1.60	0.85, 0.79	
Tyr26	8.06	3.98	2.99, 2.91			6.87 (2,6) 6.64 (3,5)
Arg27	7.82	3.86	1.88, 1.79	1.61	3.15	
Leu28	7.83	3.94	1.53	1.45	0.68	
Ile29	8.13	3.43	1.82	1.69, 0.96	0.68	0.73 (γCH_3)
Gln30	8.00	3.79	1.88	2.11, 1.95		6.83, 6.64
Ala31	7.92	4.02	1.43			
Glu32	8.08	4.10	1.98	2.37, 2.28		
Ala33	8.39	3.94	1.37			
Arg34	7.82	4.03	1.85, 1.76	1.61	3.15	
Lys35	7.63	4.16	1.89, 1.84	1.47, 1.39	1.63	2.92 (εCH_2)
Met36	7.96	4.33	2.08, 2.01	2.58, 2.49		1.92 (εCH_3)
Ser37	7.88	4.36	3.93			
Gly38	8.11	4.03, 3.94				
Cys39	8.14	4.71	3.16, 2.94			
Ser40	8.34	4.45	3.83			
Asn41	8.18	4.63	2.87, 2.77			7.43, 6.77

tained in 10% CD_3CN. The chemical shifts of P1, APRP1, and APRP2 are shown in Tables I–III. Spectra were referenced to water, which was in turn referenced to sodium trimethylsilyl-propionate (TSP) at 0.000 ± 0.004 ppm and dioxan at 3.743 ± 0.003 ppm.

Far UV (185–260 nm) circular dichroism (CD) spectra of P1 were recorded using a Jasco J600 spectropolarimeter; high-precision, single-scan spectra were obtained using a spectral bandwidth of 2 nm and a 4 s time constant (10 nm min^{-1}). Sample concentrations were in the range of 0.30–0.50 mg ml^{-1}, as determined both gravimetrically and by UV absorption. CD spectra were obtained from 0.05 cm pathlength cells at 295 K. The results (Fig. 2) are reported as the differential molar CD extinction coefficients, Δε, vs. wavelength, where Δε is $\varepsilon_L - \varepsilon_R$, and are based on a mean molecular mass per residue of 112.3. It is clear that the CD spectra of P1 in 100% H_2O and 85% H_2O/15% AcCN are very similar, which indicates that the addition of acetonitrile does not perturb the structure significantly; in addition, both spectra are characteristic of a high α-helical content, with a positive Δε maximum at 190 nm and negative Δε maxima at 208 nm and 220 nm.

Solution state ^{1}H NMR spectra were recorded on either a Bruker AM 600 or AM 500 spectrometer, at 300 K, 310 K, and 320 K. All spectra recorded were phase-sensitive, with time-proportional phase incrementation (TPPI) used for quadrature detection in F1. Nuclear Overhauser effect (NOESY) spectra[21] and homonuclear Hartmann-Hahn (HOHAHA) spectra[22,23] were recorded using 2 K real data points in t2 and 750–1,024 t1 increments. The receiver phase and pre-acquisition delay were adjusted to minimize baseline distortions.[24] Solvent saturation was achieved by presaturation in the case of spectra recorded in D_2O/CD_3CN, and by the use of a jump-return sequence[25] for spectra recorded in $H_2O/D_2O/CD_3CN$. The free induction decays obtained from samples in $H_2O/D_2O/CD_3CN$ were deconvoluted in the time domain with a cosine function prior to Fourier transformation in F2.[26] Correlated (COSY) spectra and double-quantum-filtered (DQF) COSY spectra[27] were acquired in 90% H_2O/10% D_2O/10% CD_3CN, and primitive exclusive (PE) COSY[28] spectra recorded in 90% D_2O/10% CD_3CN, all with 4 K real data points and 800–1,024 t1 increments.

Spectra of each peptide collected at 310 K were assigned in detail; spectra at other temperatures

TABLE III. ^{1}H Chemical Shifts at 310 K and pH 3.2 of APRP2

	NH	αH	βH	γH	δH	Others
α-Chain						
Asp14			2.96			
Ala15	8.60	4.27	1.34			
Ala16	8.10	4.16	1.23			
Asp17	8.02	4.58	2.78, 2.70			
Phe18	7.94	4.50	3.06, 2.95			7.22 (3,5) 7.11 (2,6)
Gly19	8.09	3.80, 3.71				
Asp20	8.03	4.92	2.96, 2.82			
Pro21		4.22	2.02, 1.78	1.59, 1.54	3.74, 3.71	
Tyr22	7.69	4.19	2.99, 2.95			7.00 (2,6) 6.71 (3,5)
Ser23	7.78	4.19	3.98, 3.94			
Phe24	7.91	4.23	3.18, 3.01			7.17 (3,5) 7.09 (2,6)
Leu25	7.86	3.84	1.66	1.61	0.86, 0.82	
Tyr26	8.11	4.03	3.03, 2.98			6.92 (2,6) 6.68 (3,5)
Arg27	7.82	3.90	1.88, 1.78	1.63	3.18	7.18, 6.62
Leu28	7.81	3.97	1.58, 1.55	1.50	0.71	
Ile29	8.14	3.43	1.84	1.71, 0.97	0.72	0.74 (γCH_3)
Gln30	8.06	3.78	1.97, 1.89	2.13		6.92, 6.67
Ala31	7.92	4.04	1.44			
Glu32	8.15	4.11	1.99	2.34		
Ala33	8.38	3.97	1.39			
Arg34	7.79	4.04	1.88, 1.78	1.63	3.18	7.18, 6.62
Lys35	7.60	4.17	1.89, 1.84	1.50, 1.40	1.63	2.95 (εCH_2) 7.48
Met36	7.89	4.36	2.11, 2.03	2.60, 2.51		1.99 (εCH_3)
Ser37	7.86	4.36	3.92			
Gly38	8.17	4.02, 3.92				
Cys39	8.15	4.70	3.27, 2.98			
Ser40	8.33	4.46	3.84			
Asn41	8.22	4.69	2.80			7.46, 6.78
β-Chain*						
Tyr14			3.11, 3.05			7.10 (2,6) 6.83 (3,5)
Ala15†	8.32	4.29	1.28			
	8.34	4.28	1.31			
Asp16†	8.32	4.85	2.81, 2.72			
	8.34	4.93	2.95, 2.79			
Pro17		4.37	2.10, 1.13	1.90	3.40, 3.35	
Asn18	7.96	4.55	2.83, 2.78			7.48, 6.80
Ala19	7.76	4.24	1.34			
Asp20	8.04	4.91	2.97, 2.79			
Pro21		4.22	2.02, 1.78	1.59, 1.54	3.74, 3.71	
Met22	8.08	4.34	2.04, 1.98	2.54, 2.48		1.92 (εCH_3)
Ala23	7.69	4.09	1.43			
Phe24	7.81	4.24	3.16, 3.10			7.18 (3,5) 7.09 (2,6)
Leu25	8.07	3.85	1.69	1.67	0.89, 0.94	
Thr26	7.99	3.72	4.18	1.13		
Lys27	7.46	4.02	1.81	1.48, 1.36	1.62	2.89 (εCH_2)
Leu28	7.81	3.91	1.60	1.48	0.70	
Ile29	8.14	3.47	1.92	1.71, 0.97	0.72	0.79 (γCH_3)
Gln30	7.79	3.92	2.24, 2.08	2.51		7.14, 6.79
Ile31	8.14	3.68	1.90	1.72, 1.10	0.76	0.86 (γCH_3)
Glu32	8.15	4.12	1.99	2.40, 2.30		
Ala33	8.57	3.92	1.42			
Arg34	7.79	4.04	1.88, 1.78	1.63	3.18	
Lys35	7.77	4.14	1.89	1.51, 1.40	1.62	2.95 (εCH_2)
Leu36	7.86	4.24	1.72, 1.70	1.54	0.78	
Ser37	7.81	4.32	3.94			
Gly38	8.00	4.02, 3.95				
Cys39	8.10	4.71	3.27, 2.94			
Ser40	8.35	4.40	3.84			
Asn41	8.19	4.66	2.80			7.50, 6.82

*Substitutions in the β-chain sequence are denoted by an underlining of the residue's name.
†Two resonances are seen for Ala15 and Asp16, probably to cis-trans isomerisation of the Asp16–Pro17 bond. The shifts listed first are due to the major form, and the ratio major:minor is approximately 2:1.

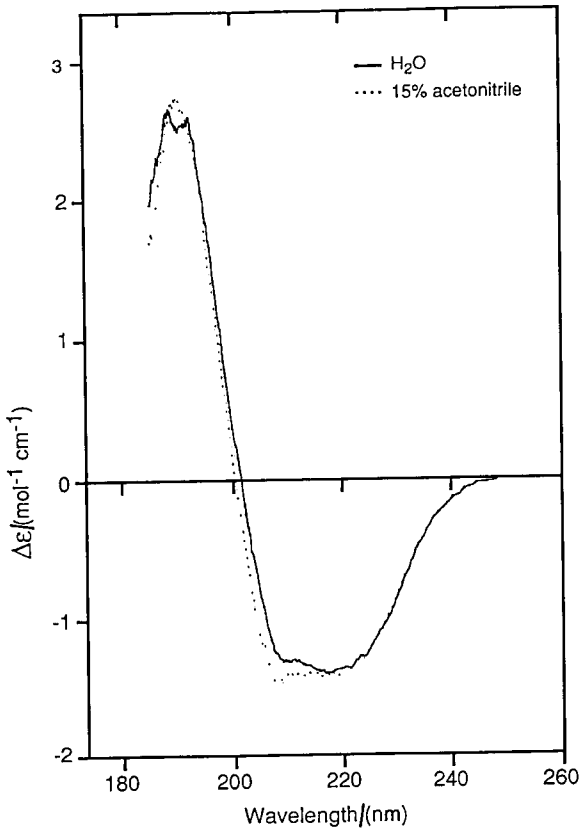


Fig. 2. CD spectra of P1 in 100% H_2O and 85% H_2O/15% AcCN.

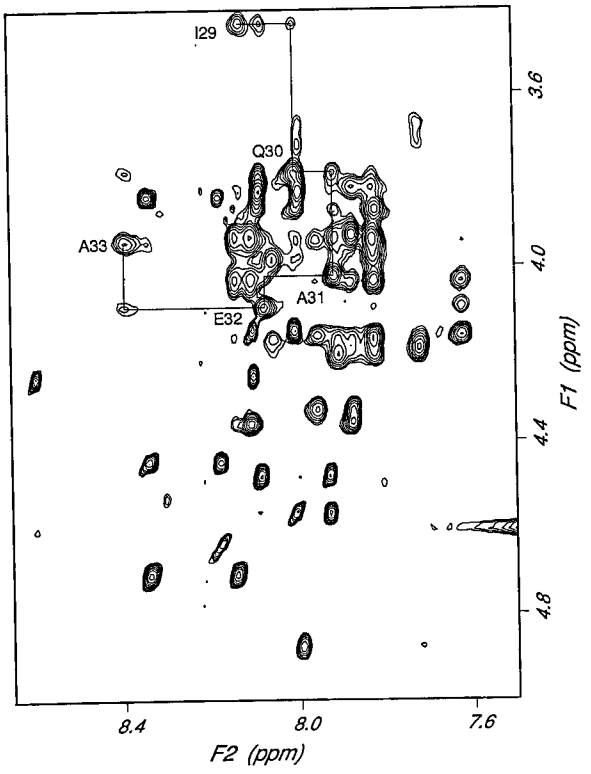


Fig. 3. The sequential assignment of part of the backbone of APRP1, from Ile29 to Ala33.

were used only to assist in overcoming ambiguities due to degeneracy of resonances. Sequential assignment was carried out by the methods described by Wüthrich[29]; spin systems were assigned from DQF-COSY, HOHAHA, and NOESY spectra, and correlated with particular residues after consideration of the primary sequences of each peptide. Figure 3 shows the sequential assignment of part of an APRP1 NOESY spectrum. Structures for APRP1 were calculated from nOe data derived from 80 ms mixing time NOESY spectra. Structures for P1 were calculated from nOe data derived from both the 80 ms NOESY of APRP1, together with a 200 ms NOESY of P1 itself. Inter-chain nOes used in the calculation procedure were obtained from a 100 ms NOESY of APRP2.

nOe cross-peaks were classified as strong, medium, or weak, which were interpreted in the structure calculations as distance restraints of < 2.5 Å, $<$ 3.5 Å, and < 5.0 Å, respectively. Appropriate corrections were made for degenerate methyl and methylene protons, and aromatic H^{δ} and H^{ϵ} protons. In addition, if only one methylene proton was observed to participate in an nOe interaction, center averaging and appropriate restraint correction were used. For non-stereospecifically assigned methylene protons, distance restraints were set at the value obtained from the weaker nOe cross-peak.

The $J_{NH\text{-}\alpha H}$ coupling constant of residues (except proline and glycine) were estimated by simulating the corresponding NH-αH cross peak in the fingerprint region of a COSY spectrum.[30] $J_{\alpha\text{-}\beta}$ coupling constants were measured from PE-COSY spectra, and, together with examination of the strength of α-β nOes, used to determine the conformation about the residues' C_α-C_β bonds.

Structure Calculation

Structures of P1 and APRP1 modeled from purely theoretical considerations (i.e., not using any NMR data)[16–18] predicted that intra-monomer proton-proton distances would always be shorter than the corresponding inter-monomer distances; in most cases, the difference in distance was considerable, greater than 5 Å. Thus all nOes observed were assumed to arise from intra-monomer interactions. Structures were calculated using the program X-PLOR, version 3.0.[31] Initial structures were generated with atomic coordinates randomized over 1 Å.[3] Simulated annealing was carried out using a robust variable length protocol described by Brünger.[31] Hydrogen bonds inferred from the diagnostic pattern of medium-range nOes for residues 22–37 were represented by 12 distance restraints, from O_i to N_{i+4}. In

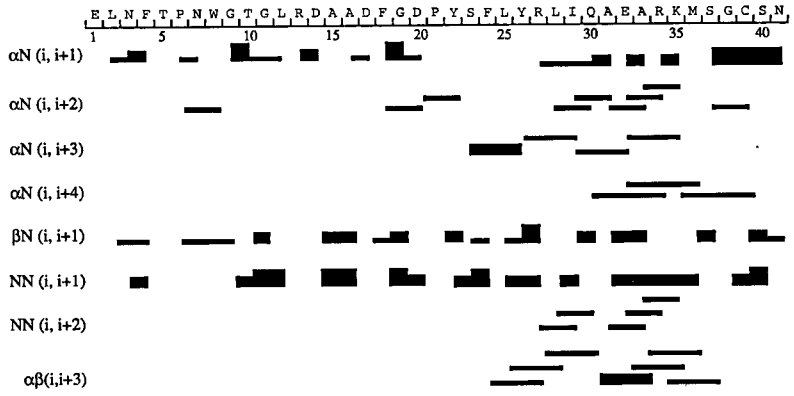


Fig. 4. A summary of the short and medium range nOe data for P1. The strength of an nOe is proportional to the thickness of the line representing it.

addition, ten inter-monomer restraints were applied over the α-helical portion of the molecule (as defined by the nOe data) between the geometric centers of each set of seven consecutive C_α atoms, similar to those used in modeling studies of coiled-coils.[16,17] Dyadic/dimeric symmetry was maintained by non-crystallographic symmetry restraints. Structures were accepted on the basis of low intra- and inter-chain nOe violations.

RESULTS

Spectra of P1, and also of APRP1, showed complete degeneracy of resonances from equivalent protons on different monomers, indicating that both of these homodimeric peptides have symmetrical structures. Proton chemical shifts from the heterodimer APRP2 were very similar to those of the equivalent proton in APRP1, except for small differences around the substituted residues. In addition, shifts for corresponding protons in P1 and APRP1 were very similar except for residues close to the N-terminus of APRP1. The chemical shifts are shown in Tables I–III.

NOESY spectra of P1, APRP1, and APRP2 showed a pattern of nOes indicating that residues 22–37 form an α-helix (the numbering is with respect to P1, so the N-terminal residues of APRP1 and APRP2 are labeled 14), with no evidence for any other secondary structure. The chemical shift and nOe data, taken together, indicate that the structures of APRP1 and APRP2 and the C-terminal 28 residues of P1 are almost identical. Figures 4, 5, and 6 show a summary of the nOe data for the three peptides.

It was possible to measure $J_{NH\text{-}\alpha H}$ constants for residues outside the α-helical regions of the three peptides only; within an α-helix, a NH-αH coupling constant is small, typically 4-6Hz[29] and this, together with the antiphase nature of the COSY cross-peaks, meant that the signal/noise ratio for couplings arising from residues within the helices was

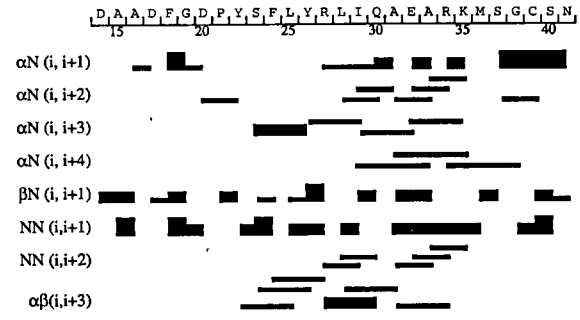


Fig. 5. A summary of the short and medium range nOe data for APRP1. The strength of an nOe is proportional to the thickness of the line representing it.

too poor for accurate measurement. For residues outside the α-helices, as defined by the nOes, the fingerprint cross-peaks had much higher signal/noise ratio; this is illustrated clearly in Figure 7, which is the fingerprint region of a DQF-COSY spectrum of P1 in which the contour level is such that the only peaks visible are those of residues outside the helix, with the exception of K35 and S37, both of which are close to or at the C-terminus of the helix. The greater signal/noise ratio is indicative of either larger coupling constants, or smaller T_2, both of which would be expected in residues with greater conformational freedom, i.e., not in the helix, which is in agreement with the nOe evidence.

Proton-proton $J_{\alpha\beta}$ coupling constants and α-β nOes were measured for all AMX spin systems having non-degenerate β-protons. In all cases, it was found that there was conformational freedom about the residue's α-β bond; thus it was not possible to make stereospecific assignments for these residues.

Three hundred one intra-monomer distance restraints were derived from nOes in the APRP1 and P1 spectra. Unambiguous inter-monomer nOe restraints observed in the NOESY spectrum of the heterodimer APRP2 (α29γ3-β33NH, α33α-β33β,

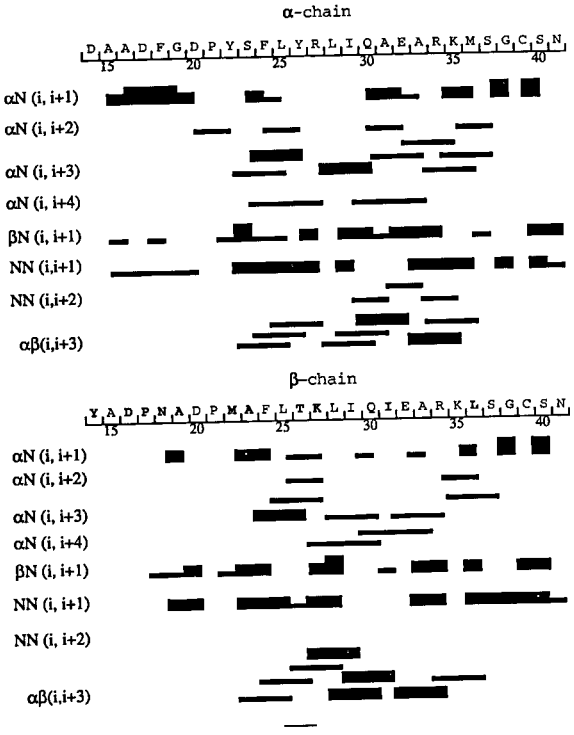


Fig. 6. A summary of the short and medium range nOe data for APRP2. The strength of an nOe is proportional to the thickness of the line representing it.

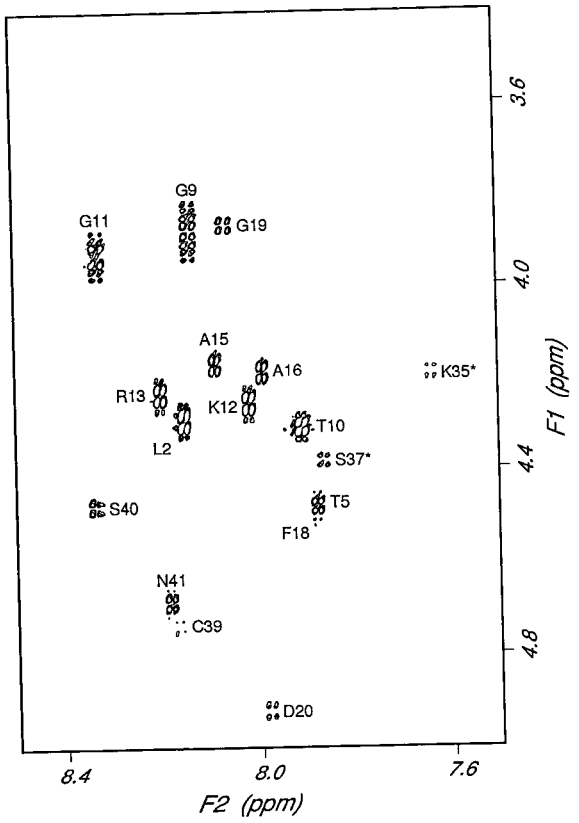


Fig. 7. Fingerprint region of a DQF-COSY of P1. The contour level is set such that only peaks arising from residues outside the helix are visible, except for K35 and S37, which are starred.

and α33α-β33NH, together with the corresponding 3 β-chain to α-chain nOes) were incorporated into the calculation as a set of six inter-monomer distance restraints. Table IV shows the distance restraints and violations.

A total of 15 structures was retained out of 100; this convergence of structures may seem poor, but may be due to the limits on the conformational freedom of the system imposed in the calculation by the symmetry restraints; this question has been addressed by Nilges.[32] These 15 structures have an average crossing angle for the helices, Ω, of Ω = −19.9° ± 3.9°, and an inter-helical separation of 10.8 Å ± 1.8 Å. The root-mean-square differences (RMSDs) for the converged structures are given in Table V. Figure 8 shows a stereo view of the helical region of P1 and APRP1, and figure 9 shows one chain of the dimer P1.

Figure 10 shows plots of the chemical shift indices[33] of P1, APRP1, and APRP2. The chemical shift index (CSI) of a residue is defined as −1 if the α-proton chemical shift is more than 0.1 ppm upfield of its random coil value (as defined in ref. 33), +1 if the shift is more than 0.1 ppm downfield of its random coil value, and zero otherwise. A run of four or more consecutive residues having a CSI of −1, not interrupted by a +1, is indicative of an α-helix; three or more +1 not interrupted by a −1 indicates a β-sheet. The secondary structure that would be predicted on the basis of these indices is largely in agreement with the nOe data; however, the index predicts that for P1, APRP1, and the α-chain of

TABLE IV. Distance Restraints Used in the Structure Calculations

Restraint	Number	Violations (Å)
Intra-monomer	301	0.16 ± 0.02
Inter-monomer	3 (×2)	0.08 ± 0.03
Coiled coil	10	0.03 ± 0.02
Hydrogen bond	12	0.16 ± 0.02

TABLE V. RMSDs for the 15 Converged Structures

Atoms	RMSD (Å)
Backbone (C, C_α, N, O, helix only) dimer	0.56 ± 0.20
C_α (helix only) dimer	0.55 ± 0.21
C_α (helix only) monomer	0.41 ± 0.17
C_α (helix only) to other monomer*	1.02 ± 0.62

*This is the RMSD for one monomer when the structures have been superimposed over the C_α atoms in the helix of the other monomer. It indicates that the relative orientations of the monomers are fairly constant.

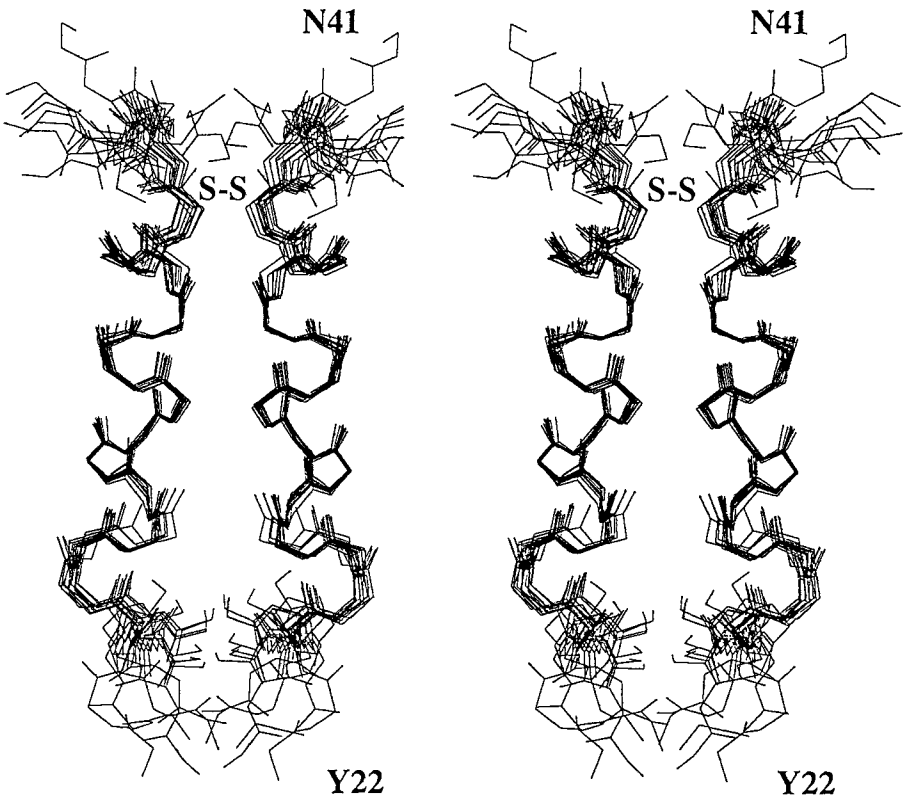


Fig. 8. Stereoview of the helical region of P1 and APRP1.

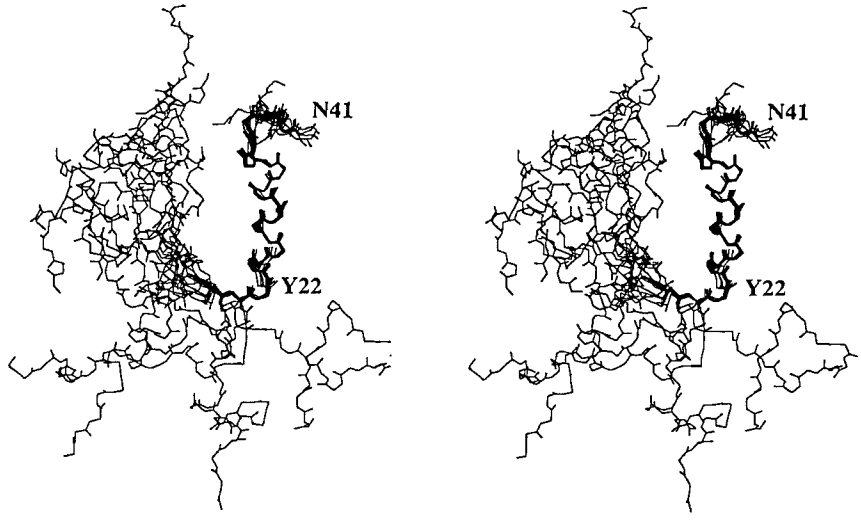


Fig. 9. Stereoview of one chain of P1.

APRP2, there should be a short helix from Ala15 to Phe18, which is not supported by the nOe data, or by $J_{NH\text{-}\alpha H}$ coupling constants. It is also of interest to note that all residues immediately preceding a proline (i.e., residue X in a X-Pro dipeptide sequence) have a CSI of +1, which is seen in other peptides.[34]

DISCUSSION

P1 and APRP1 are symmetrical homodimers, with a C-terminal region, residues 22–37 (numbered with respect to the N-termini of P1), consisting of two intimately associated parallel α-helices, and unstructured N-termini. The helices cross at an angle of approximately −20°, and the average distance be-

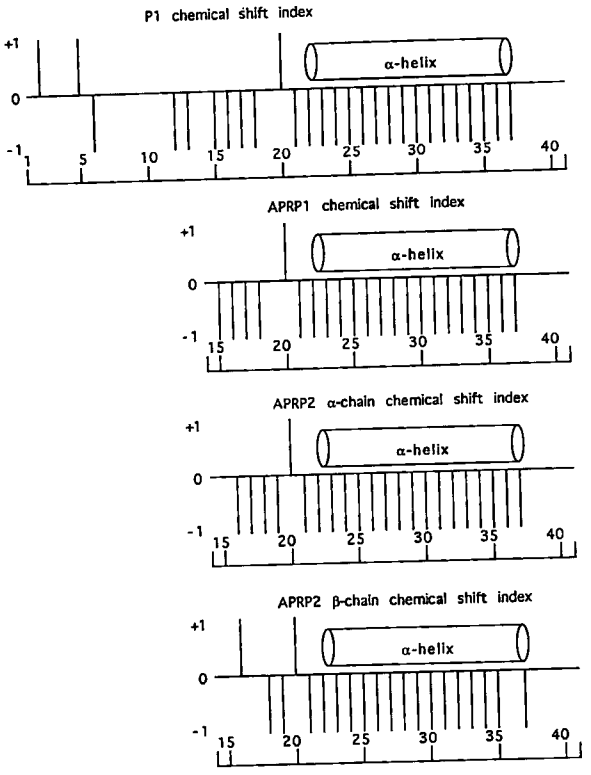


Fig. 10. Chemical shift indices of P1, APRP1, and APRP2.

tween them is 10.8 Å; the negative crossing angle corresponds to a right-handed supercoiling. The tertiary structure of the heterodimer APRP2 is probably almost identical to the homodimeric APRP1 over the entire molecule, except for minor effects attributable to the substitutions on the non-identical chain.

The only other complete peptide hormone precursor to be studied by NMR is proinsulin.[11] Other workers have studied incomplete prohormone sequences by biophysical methods, including NMR, with the aim of identifying structural motifs associated with enzymatic cleavage sites. For example, Rholam, Cohen, and co-workers have studied peptide sequences from the N-terminal portion of prooxytocin/neurophysin, a region that includes the site of enzymatic cleavage.[35,36] This work has provided evidence for a β-turn adjacent to the basic amino acid pair that precedes the appropriate endopeptidase cleavage site. It is believed that this motif may serve as a recognition element for the prohormone processing endopeptidase.[37] In the present study, there is no evidence of secondary structure in the N-terminal region that contains the endopeptidase cleavage site of P1, an AKH I precursor. Thus this work is unable to support the proposal that an Ω-loop may serve as a recognition element for P1 cleavage.[38] Our results do, however, show that another basic amino acid pair in the sequence of P1 (Arg34-Lys35), which does not mark a processing site, lies within an α-helix. Brakch et al.[39] have shown that an Arg-Lys sequence in prosomatostatin that is a cleavage site normally lies adjacent to a β-turn, and that disruption of this motif abolishes processing. The inability of processing endopeptidases to cleave the bond adjacent to Arg34-Lys35 in P1 thus supports further the idea that sites of cleavage in prohormones generally reside near flexible structures and do not lie in helical regions.[10,40]

ACKNOWLEDGMENTS

This is a contribution from the Oxford Centre for Molecular Science, which is supported by SERC and MRC, and from the Sussex Centre for Neuroscience, which is supported by SERC. T.J.H. thanks the MRC for the award of a Training Fellowship, D.G.D. thanks the Department of Education for Northern Ireland for a studentship, and G.B. thanks MURST for a training fellowship (IMI grant 53704. We thank Dr. Paul Barlow and Dr. David Norman for their advice and assistance with this work, Dr. Jennifer Potts for very interesting discussions about X-Pro sequences, Dr. Colin Wheeler and Dr. Chris Kowalczyk for peptide sequencing, Dr. Alex Drake of the National Chiroptical Spectroscopy Laboratory, Birkbeck College, University of London, for his expertise in interpreting the CD spectra, and Beulah Banfield for dedicated technical assistance.

REFERENCES

1. Beenakkers, A.M.T. The influence of corpus allatum and corpus cardiacum on lipid metabolism in *Locusta migratoria*. Gen. Comp. Endocrinol. 13:492, Abstract 12, 1969.
2. Mayer, R.J., Candy, D.J. Control of haemolymph lipid concentration during locust flight: An adipokinetic hormone from the corpora cardiaca. J. Insect Physiol. 15:611–620, 1969.
3. Stone, J.V., Mordue, W., Batley, K.E., Morris, H.R. Structure of locust adipokinetic hormone, a neurohormone that regulates lipid utilization during flight. Nature 263:207–211, 1976.
4. Siegert, K., Morgan, P., Mordue, W. Primary structures of locust adipokinetic hormones. II. Biol. Chem. Hoppe-Seyler 366:723–727, 1985.
5. O'Shea, M., Rayne, R.C. Adipokinetic hormones: Cell and molecular biology. Experientia 48:431–438, 1992.
6. Gluschankof, P., Cohen, P. Proteolytic enzymes in the post-translational processing of polypeptide hormone precursors. Neurochem. Res. 12:951–958, 1987.
7. Fisher, J.M., Scheller, R.H. Prohormone processing and the secretory pathway. J. Biol. Chem. 263:16515–16518, 1988.
8. Darby, N.J., Smyth, D.G. Endopeptidases and prohormone processing. Biosci. Rep. 10:1–13, 1990.
9. Geisow, M.J., Smyth, D.G. Proteolysis in prohormones and pro-proteins. In: "The Enzymology of Post-Translational Modification of Proteins." Vol 1. Freedman, R.B., Hawkins, H.C., eds. London: Academic Press, 1980:259–287.
10. Rholam, M., Nicolas, P., Cohen, P. Precursors for peptide hormones share common secondary structures forming features at the proteolytic processing sites. FEBS Lett. 207:1–6, 1986.
11. Weiss, M.A., Frank, B.H., Khait, I., Pekar, A., Heiney, R., Shoelson, S.E., Neuringer, L.J. NMR and photo-CIDNP studies of human proinsulin and prohormone processing

intermediates with application to endopeptidase recognition. Biochemistry 29:8389–8401, 1990.
12. Saudek, V., Pastore, A., Castiglione Morelli, M.A., Frank, R., Gausepohl, H., Gibson, T., Weih, F., Roesch, P. Solution structure of the DNA-binding domain of the yeast transcriptional activator GCN4. Protein Eng. 4:3–10, 1990.
13. O'Donoghue, S.I., Junius, F.K., King, G.F. Determination of the structure of symmetric coiled-coil proteins from NMR data: Application of the leucine zipper proteins Jun and GCN4. Protein Eng. 6:557–564, 1993.
14. Arrowsmith, C.H., Pachter, R., Altman, R.B., Iyer, S.B., Jardetzky, O. Sequence-specific ^{1}H NMR assignments and secondary structure in solution of *Escherichia coli trp* repressor. Biochemistry 29:6332–6341, 1990.
15. O'Shea, E.K., Klemm, J.D., Kim, P.S., Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:539–543, 1991.
16. Nilges M., Brünger, A.T. Automated modeling of coiled coils: Application to the GCN4 dimerisation region. Protein Eng. 4:649–659, 1991.
17. Nilges, M., Brünger, A.T. Successful prediction of the coiled coil geometry of the GCN4 leucine zipper domain by simulated annealing: Comparison to the X-ray structure. Proteins 15:133–146, 1993.
18. Doak, D.G. D. Peptide models of transmembrane proteins. Phil. Thesis, University of Oxford, 1994.
19. Jaenicke, R., Rudolf, R. Folding proteins. In: "Protein Structure, a Practical Approach." Creighton, T.E., ed. Oxford: IRL Press, 1989:191–223.
20. Rayne, R.C., O'Shea, M. Reconstitution of adipokinetic hormone biosynthesis in vitro indicates steps in prohormone processing. Eur. J. Biochem. 219:781–789, 1994.
21. Kumar, A., Ernst, R.R., Wüthrich, K. A Two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological molecules. Biochem. Biophys. Res. Commun. 95:1–6, 1981.
22. Braunschweiler, L., Ernst, R.R. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy, J. Magn. Reson. 53:521–528, 1983.
23. Davis, D.G., Bax, A. Assignment of complex ^{1}H NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy. J. Am. Chem. Soc. 107:2820–2821, 1985.
24. Marion, D., Bax, A. Baseline distortion in real-Fourier-transform NMR spectra. J. Magn. Reson. 79:352–356, 1988.
25. Plateau, P., Gueron, M. Exchangeable proton NMR without baseline distortion, using new strong-pulse sequences. J. Am. Chem. Soc. 104:7310–7311, 1982.
26. Marion, D., Ikura, M., Bax, A. Improved solvent suppression in one- and two-dimensional NMR spectra by convolution of time-domain data. J. Magn. Reson. 84:425–430, 1989.
27. Rance, M., Sorensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R., Wüthrich, K. Improved spectral resolution in COSY ^{1}H NMR spectra of proteins via double quantum filtering. Biochem. Biophys. Res. Commun. 117:479–485, 1983.
28. Müller, L. PE COSY: A simple alternative to E COSY. J. Magn. Reson. 72:191–196, 1987.
29. Wüthrich, K. "NMR of Proteins and Nucleic Acids." New York: Wiley, 1986.
30. Smith, L.J., Sutcliffe, M.J., Redfield, C., Dobson, C.M. Analysis of ϕ and χ_1 torsion angles for hen lysozyme in solution from ^{1}H NMR spin-spin coupling constants. Biochemistry 30:986–996, 1991.
31. Brünger, A.T. "X-PLOR. A System for X-Ray Crystallography and NMR." New Haven: Yale University Press, 1992.
32. Nilges, M. A calculation strategy for the structure determination of symmetrical dimers by ^{1}H NMR. Proteins 17: 297–309, 1993.
33. Wishart, D.S., Sykes, B.D., Richards, F.M. The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31:1647–1651, 1992.
34. Murray, N.J., Williamson, M.P. Conformational study of a salivary proline-rich protein repeat sequence. Eur. J. Biochem. 219:915–921, 1994.
35. Brakch, N., Boussetta, H., Rholam, M., Cohen, P. Processing endopeptidase recognizes a structural feature at the cleavage site of peptide prohormones. J. Biol. Chem. 264: 15912–15916, 1989.
36. Paolillo, L., Simonetti, M., Brakch, N., D'Auria, G., Saviano, M., Dettin, M., Rholam, M., Scatturin, A., DiBello, C., Cohen, P. Evidence for the presence of secondary structure at the dibasic processing site of a prohormone: The prooxytocin model. EMBO J. 11:2399–2405, 1992.
37. Brakch, N., Rholam, M., Boussetta, H., Cohen, P. Role of β-turn in proteolytic processing of peptide hormone precursors at dibasic sites. Biochemistry 32:4925–4930, 1993.
38. Rayne, R.C., O'Shea, M. Structural requirements for processing of proadipokinetic hormone I. Eur. J. Biochem. 217:905–911, 1993.
39. Brakch, N., Boileau, G., Simonetti, M., Nault, C., Joseph-Bravo, P., Rholam, M., Cohen, P. Prosomatostatin processing in Neuro2A cells: Role of β-turn structure in the vicinity of the Arg-Lys cleavage site. Eur. J. Biochem. 216: 39–47, 1993.
40. Bek, E., Berry, R. Prohormonal cleavage sites are associated with Ω-loops. Biochemistry 29:178–183, 1990.