Journal of Peptide Science
J. Peptide Sci. **4**: 305–318 (1998)

A Conformational Study in Solution of Pro-Somatostatin Fragments by NMR and Computational Methods

LUCIA FALCIGNO[a], FRANCA FRATERNALI[a], DANIELA M. MANDUCA[a], GABRIELLA D'AURIA[a], MARIO SIMONETTI[b], CARLO DI BELLO[b] and LIVIO PAOLILLO[a,*]

[a] Department of Chemistry, University of Naples, 80134 Naples, Italy
[b] Department of Chemical Process Engineering, University of Padua, 35131 Padua, Italy

Received 24 June 1997
Accepted 8 October 1997

Abstract: The results of a conformational study by nuclear magnetic spectroscopy and computational methods on a series of point-mutated synthetic peptides, containing 14 amino acid residues and mimicking the region containing the Arg-Lys dibasic cleavage site of pro-somatostatin, have confirmed the possible role of a well defined secondary structure in the recognition phenomenon by processing enzymes.

The importance of the residues located near the Arg-Lys dibasic site in the C-terminal region of the pro-hormone for the cleavage of the precursor into somatostatin-14 has been confirmed. The present structural analysis indicates the occurrence of two β-turns in the 4–7 and 11–14 regions, flanking the cleavage site, for all the peptides recognized as substrates by the processing enzyme.

Interestingly, in the point-mutated analogue not processed by the enzyme and containing the replacement of proline by alanine in position 5 the first β-turn is displaced by one residue and involves the Ala^5-Arg^8 segment. This observation may explain the lack of recognition by the maturation enzyme. © 1998 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: pro-somatostatin; β-turn; NMR; computational methods; conformational analysis

INTRODUCTION

In recent years numerous studies have been made on the characterization of the structural parameters that are thought to play a key role in the enzyme–prohormone recognition process [1–5].

Pro-S is the common precursor of two biologically active peptides, S-14 and S-28. These hormones, differently localized within the cells, inhibit the release of growth hormone, insulin, and glucagon [6], and appear to be endogenous antiproliferative agents for a large variety of cells [7,8]. Besides the wide range of important biological activities, the pro-S system constitutes a very useful model in the study of the formation of multiple hormone products by differential processing of a single polyfunctional precursor [9]. In fact, in mammals a single gene encodes for the common precursor pro-S, which undergoes tissue specific processing to yield S-14 and/or S-28. This system appears quite sophisticated when compared to lower organisms like the teleostean fish [10], where two separate precursors, encoded by distinct genes, are processed in different cells to originate S-14 and S-28. The mammalian specific processing can be explained either by the presence of different processing enzymes in

Abbreviations: BFGS, Broyden–Fletcher–Goldfard–Shanno; CD, circular dichroism; DIANA, distance geometry algorithm for NMR application; DQFCOSY, double quantum filter correlated spectroscopy; E.COSY, exclusive correlation spectroscopy; FID, free induction decay; IR, infrared; mBHA, 4-methylbenzhydrylamine; MLEV, Malcom Levitt; MM, molecular mechanics; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; pro-S, prosomatostatin; rms, root mean square; ROE, rotating frame Overhauser enhancement; ROESY, rotating frame Overhauser spectroscopy; S-14, somatostatin-14; S-28, somatostatin-28; TOCSY, total correlation spectroscopy.

* Correspondence to: Department of Chemistry, University of Naples, via Mezzocannone 4, 80134 Naples, Italy.

CCC 1075–2617/98/050305-14$17.50

the cells or by the specific targeting of the precursor to different secretory pathways [11,12]. In both cases, however, the precursor undergoes a processing step where limited proteolysis occurs, mostly at basic pairs, but also at a single Arg residue. This step is thought to involve a specific recognition machinery discerning sequence and/or structure specific determinants encoded in the precursor [13,14].

In this paper, as part of a systematic study aimed at elucidating the structural elements essential to the enzyme–prohormone recognition, we present the results of a conformational study in solution by NMR and computational methods on synthetic fragments (Table 1) mutated in the region directly preceding the Arg-Lys dibasic cleavage site of pro-somatostatin. In particular, in the analogue Som II the proline in position 5 has been replaced with a Gly residue expected to promote the formation of β-turn structures in the segment preceding the Arg-Lys site [15]. In the analogue Som I the Pro^5 residue is instead replaced with an Ala residue expected to decrease the propensity toward β-turn formation, while in the analog Som III the entire sequence Pro-Arg-Glu-Arg is replaced by the Ser-Ser-Asn-Arg sequence shown to be part of a β-turn structure in 2SNS nuclease [16,17].

As linear peptides are intrinsically flexible molecules, the NMR data are often an average of many distinct conformational states. In these cases, the NMR data alone are not sufficient to describe physically reasonable conformations of the peptides, and other data (such as those from force-field theoretical calculations or from other experimental techniques) have to be used in order to produce plausible conformational models.

Structural information from CD, IR and sequence predictions suggest that secondary structures of the β-turn type can occur near the Arg-Lys cleavage site [18,19].

In the present study we have extended the structural work with the help of NMR and computational techniques in order to obtain a better insight in the local order around the dibasic site. The results demonstrate that the linear analogues examined in the present study do exhibit a rather high flexibility along the chains but in the proximity of the dibasic site. Therefore, since in this region the existence of β-turns seems confirmed, we have used theoretical calculations to assess their stability.

These turns appear stable and are present in all examined peptides, thus adding further experimental evidence to the hypothesis that they may be essential in directing the cleavage with the conformation appropriate for the processing enzyme.

Moreover, our approach permits to assign the peptide sequence in β-turn conformation near the dibasic cleavage site. It will also be shown that those peptides that are not recognized by the processing enzyme have a different β-turn architecture.

Table 1 Sequences of synthetic analogues of Pro-S[a]

Peptide	Sequence
Som WT_b	H-P^1-A-M-A-*P*^{5}-R-E-**R**-**K**-A-G-A-K-N^{14}-NH_2
Som I	H-P^1-A-M-A-*A*^{5}-R-E-**R**-**K**-A-G-A-K-N^{14}-NH_2
Som II	H-P^1-A-M-A-*G*^{5}-R-E-**R**-**K**-A-G-A-K-N^{14}-NH_2
Som III	H-P^1-A-M-A-*S*^{5}-*S*-*N*-**R**-**K**-A-G-A-K-N^{14}-NH_2

[a] The dibasic cleavage site is labeled with bold characters and the substituted residues are indicated in italics.

MATERIALS AND METHODS

Peptide Synthesis and Purification

All peptides (Table 1) were synthesized by standard solid-phase methods on an Applied Biosystems 431A apparatus using Boc chemistry and a mBHA resin purchased from Novabiochem (Läufelfingen, Switzerland). The Met residue was introduced as Met sulfoxide. Peptide removal from the resin and deprotection of all functions were performed by reaction at -5–0°C for 1 h with a mixture of HF/anisole/dimethylsulphide (100/10/10 v/v) containing 10 mg/ml of 2-mercaptopyridine. Analyses of the crude deprotected peptides by HPLC, using a CH_3CN mixture containing 0.05% TFA as eluent (Waters C_{18} Delta-Pak 100 Å, 3.9 × 150 mm column), showed two major peaks. For each peptide, separation and purification of these two products with the higher retention times gave the expected results, while the masses of the other components corresponded to that of the correct sequences, but containing Met sulfoxide. Amino acid composition, after acid hydrolysis and amino-terminal sequencing gave the expected results.

NMR Measurements

Samples used in all ^{1}H-NMR experiments were prepared by dissolving 6–7 mg of purified solid peptides in 700 μl DMSO-d_6 (99.96% isotopic purity, CIL) or DMSO-d_6/H_2O 80/20 (v/v).

The chemical shifts were referred to internal tetramethylsylane (TMS) in the case of DMSO and to internal 3-(trimethylsilyl)-propionic acid-2,2′,3,3′-d_4 (sodium salt) in the case of aqueous solutions.

All experiments were carried out on a Varian UNITY 400 NMR spectrometer.

Two-dimensional (2D) spectra were measured at 298 K with 2048 or 4096 data points along t_2 and 256 t_1 increments. All 2D experiments were always run in the phase-sensitive mode by using the hypercomplex technique known as the States–Haberkorn–Ruben method [20]. The data points were zero-filled to 1k in F1 before transformation. In all cases $\pi/2$-shifted, squared sine-bell weighting functions were used.

The total correlation spectroscopy (TOCSY) [21] spectra were acquired by using the basic sequence proposed by Bax [22]. Sixteen scans were accumulated per experiment with a 2 ms trim pulse preceding the 70 ms Malcom Levitt (MLEV-17) spin lock period.

Due to resonance overlaps in the one-dimensional (1D) spectra, the temperature dependence of the amide protons was calculated by measuring the chemical shifts in the TOCSY maps, acquired with 4096 data points in the temperature range of 298–310 K.

The nuclear Overhauser enhancement spectroscopy (NOESY) experiments [23,24] were carried out by using 200, 250 and 300 ms mixing times and 32 scans per t_1 increment. The rotating frame nuclear Overhauser enhancement spectroscopy (ROESY) [25,26] experiments were recorded with a continuous spin lock. Two 90° hard pulses immediately preceding and following the spin lock period were added to compensate for the non-negligible offset dependence of the cross-peaks (compensated ROESY) [27]. ROESY spectra were performed with mixing times of 100–120 ms and 32 scans per FID. Intensity of NOE effects was evaluated by volume integration of cross-peaks in the NOESY and ROESY spectra by means of the VARIAN software. These contacts were classified as strong (s), medium (m) and weak (w), to define upper and lower limits for internuclear distances in the building-up of molecular models (Table 2).

Table 2 Principal NOE backbone contacts of the peptides Som WT_b, Som I, Som II and Som III in DMSO at 298 K. NOE intensities are classified as strong (s), medium (m) and weak (w)

Contacts	Som WT_b	Som I	Som II	Som III
NHP^1-αP^1		m	w	m
NHA^2-αP^1	s	s	s	s
NHA^2-αA^2	w	w	w	m
NHA^2-NHM^3	w	w	w	w
NHM^3-αM^3	s[a]	s		s[a]
NHM^3-αA^2	s		s	s[a]
NHA^4-αA^4	m	s[a]		s[a]
NHA^4-αM^3	s[a]			s[a]
αA^4-$\delta\delta' P^5$	s			
$NH\mathbf{X}^5$-$\alpha\mathbf{X}^5$		s[a]	s[a]	s[a]
$NH\mathbf{X}^5$-αA^4		s[a]		s[a]
$NH\mathbf{X}^6$-$\alpha\mathbf{X}^5$		s[a]	s[a]	s
NHR^6-NHG^5			s[a]	
$NH\mathbf{X}^6$-$NH\mathbf{X}^7$	s	w		s[a]
NHE^7-$\alpha\alpha' G^5$			w	
$NH\mathbf{X}^7$-$\alpha\mathbf{X}^6$	s	s[a]		s
$NH\mathbf{X}^7$-$\alpha\mathbf{X}^7$	m	s[a]	s[a]	s
$NH\mathbf{X}^7$-NHR^8		m		
NHR^8-αN^7				s
NHK^9-αR^8				s[a]
NHK^9-αK^9		s[a]	s[a]	s[a]
NHA^{10}-αA^{10}		s[a]		s
NHG^{11}-αA^{10}	s	s[a]	s	s
NHG^{11}-$\alpha\alpha' G^{11}$	s	s	s[a]	s
NHA^{12}-NHK^{13}	s[a]	s[a]	s[a]	s[a]
NHA^{12}-$\alpha\alpha' G^{11}$	s	s	s[a]	s
NHA^{12}-αA^{12}	s			s[a]
NHK^{13}-αA^{12}	m	s[a]		
NHK^{13}-αK^{13}	m	m	m	s
NHK^{13}-NHN^{14}	s[a]	s[a]	s[a]	s[a]
NHN^{14}-αK^{13}	s	s	s	
NHN^{14}-αN^{14}	m	m	m	s

[a] Overlapped contact.
$\mathbf{X}^5$ indicates P^5 for Som WT_b, A^5 for Som I, G^5 for Som II and S^5 for Som III.
$\mathbf{X}^6$ indicates R^6 for Som WT_b, for Som I and for Som II, S^6 for Som III.
$\mathbf{X}^7$ indicates E^7 for Som WT_b, for Som I and for Som II, N^7 for Som III.

Vicinal $^3J_{NH\alpha CH}$ coupling constants were estimated by examining the correlation patterns in the double quantum filtered correlated spectroscopy (DQFCOSY) [28] and exclusive correlation spectroscopy (E.COSY) [29,30] spectra.

In all NMR experiments performed in DMSO/H_2O, the water resonance from the solvent was suppressed during the relaxation delay by presaturation. The decoupler offset was set equal to the

STEP 1: Substitution of the β-turn residues with Ala residues

P^1-A^2-M^3-A^4-**G**^{5}-R^6-E^7-R^8-K^9-A^{10}-G^{11}-A^{12}-K^{13}-N^{14}

| |

A^2-A^3-A^4-**G**^{5}-A^6-A^7 => MODEL PEPTIDE *SOM IIALA*

↓

STEP 2: Imposition of **backbone restraints** + imposition of a **type I β-turn** in the sequence A^2-A^3-A^4-**G**^{5}-A^6-A^7

Energy Minimization

↓

STEP 3: Substitution of Ala residues with the **original residues** for SWT, SII and SIII + **imposition** of the second **type I β-turn** in the region G^{11}-A^{12}-K^{13}-N^{14}

Energy Minimization

↓

STEP 4: Imposition of **side-chain restraints**

Energy Minimization

↓

STEP 5: Comparison between the **structure minimized in (3)** with the **structure minimized in (4)**

↓

STEP 6: Systematic search of side-chain conformations on the structure obtained in **(3)** imposing **all NOE restraints**

Series of **Energy Minimization**

Scheme 1 Procedure used for the construction of starting structures

transmitter to prevent phase distortions around the water peak due to imperfect cancellation of the dispersive component of residual water magnetization [31].

Computational Methods

All studied systems present long side-chains in the β-turn region so that the imposition of a folded conformation results geometrically quite complex. For this reason, a model peptide (peptide model *SOMIIALA*, Scheme 1) has been generated with all Ala residues in the turn region and by imposing a β-turn conformation as suggested from the NMR data. This starting model, after being minimized, has been substituted by the original residues (STEPS 4, 5, 6 of Scheme 1).

Energy minimizations with NOE distance restraints (Table 2) have been performed for the native sequence Som WT_b and for peptides Som II and Som III. The Som I peptide, not recognized by the endoprotease, has not been examined in this section.

All the calculations have been performed with the package SYBYL (Version 5.3); the parametrization chosen is that of TRIPOS [32]. The dielectric function has been chosen as:

(i) Distance dependent (simulations labelled as SWT(r), SII(r), SIII(r) for Som WT_b, Som II, Som III, respectively);
(ii) Constant with a high value (40), in order to cut-off the electrostatic contribution (simulations labelled as SWT(40), SII(40), SIII(40), for Som WT_b, Som II, Som III, respectively).

On the obtained minima the restraints have been relaxed and the structures SWTU, SIIU and SIIIU have been generated.

Table 3 Chemical shifts of peptides Som WT_b (Pro^5-Arg^6-Glu^7), Som I (Ala^5-Arg^6-Glu^7), Som II (Gly^5-Arg^6-Glu^7) and Som III (Ser^5-Ser^6-Asn^7) in DMSO at 298 K

Amino acid	Proton	Som WT_b	Som I	Som II	Som III
Pro^1	NH	9.34; 8.50	9.32; 8.50	9.34; 8.50	9.29; 8.50
	αCH	4.19	4.19	4.19	4.18
	βCH	2.27; 1.86	2.27; 1.86	2.27	2.27; 1.86
	γCH	1.86	1.86	1.85; 1.74	1.86
	δCH	3.21	3.21	3.21	3.20
Ala^2	NH	8.68	8.69	8.69	8.68
	αCH	4.36	4.36	4.36	4.36
	βCH	1.24	1.25	1.24	1.24
Met^3	NH	8.12	8.14	8.12	8.14
	αCH	4.33	4.31	4.33	4.32
	βCH	1.86; 1.74	1.89;1.78	1.92;1.76	1.90; 1.77
	γCH	2.42	2.43	2.45	2.44
	S-CH3	2.03	2.03	2.03	2.03
Ala^4	NH	8.13	8.06	8.05	8.05
	αCH	4.50	4.25	4.48	4.33
	βCH	1.18	1.22	1.24	1.22
a.a.^{5}	NH	–	8.03	8.15	8.02
	αCH	4.33	4.24	3.68; 3.82	4.33
	βCH	2.04	1.19		3.64; 3.55
	γCH	1.82			5.07 (γOH)
	δCH	3.61; 3.53			
a.a.^{6}	NH	8.09	8.06	8.00	7.93
	αCH	4.19	4.25	4.30	4.30
	βCH	1.70	1.66;1.51	1.67	3.64; 3.55
	γCH	1.49	1.48	1.49	5.09 (γOH)
	δCH	3.09	3.09	3.09	
	εNH	7.64	7.64	7.65	
a.a.^{7}	NH	7.80	7.87	8.04	8.16
	αCH	4.28	4.27	4.24	4.55
	βCH	1.86;1.72	1.86;1.72	1.86; 1.74	2.57; 2.45
	γCH	2.23	2.23	2.25	7.46; 6.98(γNH)
Arg^8	NH	8.04	8.00	8.00	7.97
	αCH	4.24	4.30	4.30	4.19
	βCH	1.66	1.66; 1.51	1.67	1.51
	γCH	1.49	1.48	1.49	1.48
	δCH	3.09	3.09	3.09	3.07
	εNH	7.64	7.64	7.65	7.54
Lys^9	NH	8.04	8.03	8.03	7.98
	αCH	4.24	4.25	4.24	4.19
	βCH	1.68	1.66;1.51	1.68	1.69;1.53
	γCH	1.32	1.32	1.32	1.31
	δCH	1.52	1.52	1.50	1.51
	εCH	2.76	2.76	2.76	2.76
	εNH	7.72	7.73	7.74	7.70
Ala^{10}	NH	8.04	8.06	8.12	7.92
	αCH	4.24	4.25	4.33	4.21
	βCH	1.22	1.22	1.23	1.24

Table 3 (*Continued*)

Amino acid	Proton	Som WT_b	Som I	Som II	Som III
Gly^{11}	NH	8.16	8.15	8.15	8.11
	αCH	3.79;3.66	3.79;3.66	3.68;3.82	3.77;3.66
Ala^{12}	NH	8.00	8.00	8.00	7.98
	αCH	4.29	4.24	4.30	4.19
	βCH	1.24	1.24	1.22	1.22
Lys^{13}	NH	8.17	8.17	8.17	8.15
	αCH	4.13	4.12	4.12	4.12
	βCH	1.68	1.66;1.51	1.67	1.67;1.54
	γCH	1.32	1.32	1.32	1.31
	δCH	1.51	1.52	1.54	1.51
	εCH	2.76	2.76	2.76	2.76
	εNH	7.75	7.73	7.74	7.70
Asn^{14}	NH	7.98	7.97	7.97	7.96
	αCH	4.42	4.42	4.42	4.41
	βCH	2.48	2.48	2.48	2.48
	γNH	7.38;6.92	7.39;6.92	7.38;6.92	7.38;6.92
	CONHt	7.08;7.02	7.09;7.03	7.09;7.03	7.09;7.02

Atomic charges have been calculated with the Pullman method [33]. Distance restraints were imposed by means of a harmonic function with a constant value of 25 kcal/Å^2 per mol and distance ranges of 1.5–2.5 (s), 2.5–3.5 (m), 3.5–5.0 (w) Å.

Total root mean square (rms) deviations have been calculated from the imposed ranges. Violations from the data have been considered equal to zero if the observed distance falls within the imposed ranges.

All the energy minimizations were performed with the BFGS algorithm [34] with a convergence criterion of 10^{-3} on the gradient norm.

As for Som II peptide, NMR restraints have been utilized for DIANA calculations [35]. All structures superimpose well in the turn region (Ala^4-Glu^7).

RESULTS

NMR Analysis

Identification of all amino acid spin systems is obtained from the TOCSY 2D spectra of the four peptides. Then, NOESY and/or ROESY spectra allow a complete assignment [36]. Proton resonances for all peptides are summarized in Tables 3 and 4, in DMSO and DMSO/H_2O, respectively.

In the latter solvent mixture, most of the NH resonances of the N-terminal residue are absent, due to fast exchange with H_2O, as the peptides are unprotected.

Chemical shifts in all peptides appear quite similar, with the obvious exception of those residues that have been varied and, to a lesser extent, of the neighbouring residues.

ROESY and NOESY experiments were carried out for all peptides and they provided similar results. In each peptide a set of NOE contacts involving consecutive backbone and side-chain residues was identified in the NOESY or in the ROESY spectra, although some ambiguities in the attribution of NOE or ROE contacts are present due to large resonance overlaps.

In the peptide Som WT_b the strong NOE contact between the $Ala^4\alpha$CH and the $\delta\delta'$ protons of Pro^5 is diagnostic of a *trans* peptide bond between the Ala^4-Pro^5 residues [37]. In the amide region, both NOESY and ROESY spectra reveal NOE contacts for the pairs NH Arg^6-NH Glu^7, NH Ala^{12}-NH Lys^{13} and NH Lys^{13}-NH Asn^{14}. These NOEs can be interpreted in terms of a β-turn like structure [36], as indicated by theoretical predictions [13,19]. In particular, they suggest a β-turn between Ala^4-Glu^7 and Gly^{11}-Asn^{14} residues.

For peptide Som II the wide overlap of NH signals does not allow an easy attribution of the NOE contact at 8.00/8.15 ppm. This very strong contact could be attributed to one or more of the NH pairs: Gly^5-Arg^6 or Gly^{11}-Ala^{12} or Ala^{12}-Lys^{13} or Lys^{13}-

Table 4 Chemical shifts of peptides Som WT_b (Pro^5-Arg^6-Glu^7), Som II (Gly^5-Arg^6-Glu^7) and Som III (Ser^5-Ser^6-Asn^7) in DMSO/H_2O at 298 K

Amino acid	Proton	Som WT_b	Som II	Som III
Pro^1	NH	–	–	8.49; 7.74
	αCH	–	4.24	4.18
	βCH	2.34	2.33	2.32
	γCH	1.91	1.91	1.90
	δCH	3.28; 3.19	3.27;3.22	3.27;3.20
Ala^2	NH	8.70	8.72	8.70
	αCH	4.30	4.38	4.31
	βCH	1.28	1.29	1.27
Met^3	NH	8.17	8.23	8.19
	αCH	4.33	4.32	4.33
	βCH	1.90,1.78	1.95;1.82	1.94;1.82
	γCH	2.46	2.47;2.44	2.47
	S-CH3	2.03	2.03	2.04
Ala^4	NH	8.20	8.19	8.18
	αCH	4.48	4.22	4.31
	βCH	1.23	1.27	1.27
a.a.^{5}	NH	–	8.24	8.01
	αCH	4.29	3.78	4.32
	βCH	2.13; 1.81		3.62; 3.69
	γCH	1.92		
	δCH	3.66; 3.56		
a.a.^{6}	NH	8.20	8.05	8.05
	αCH	4.19	4.26	4.32
	βCH	1.72	1.73	3.66;3.60
	γCH	1.54	1.56	
	δCH	3.10	3.10	
	εNH	7.47	7.48	
a.a.^{7}	NH	7.90	8.08	8.19
	αCH	4.25	4.23	4.56
	βCH	1.93;1.80	1.95;1.82	2.62; 2.53
	γCH	2.28	2.32	7.54; 7.00 (γNH)
Arg^8	NH	8.12	8.13	7.99
	αCH	4.25	4.22	4.20
	βCH	1.72	1.70	1.75
	γCH	1.55	1.55	1.52
	δCH	3.10	3.10	3.09
	εNH	7.47	7.48	7.45
Lys^9	NH	8.10	8.12	8.04
	αCH	4.20	4.52	4.20
	βCH	1.71	1.55	1.72
	γCH	1.36	1.34	1.34
	δCH	1.56	1.70	1.55
	εCH	2.80	2.81	2.80
	εNH	7.75	7.75	7.74
Ala^{10}	NH	8.19	8.04	8.09
	αCH	4.48	4.47	4.22
	βCH	1.23	1.27	1.27

Table 4 (*Continued*)

Amino acid	Proton	Som WT_b	Som II	Som III
Gly^{11}	NH	8.24	8.24	8.19
	αCH	3.77	3.78	3.72
Ala^{12}	NH	8.04	8.05	8.03
	αCH	4.24	4.26	4.22
	βCH	1.29	1.27	1.27
Lys^{13}	NH	8.24	8.23	8.21
	αCH	–	4.20	4.14
	βCH	1.72	1.60	1.72
	γCH	1.36	1.34; 1.28	1.34
	δCH	1.56	1.72	1.55
	εCH	2.80	2.81	2.80
	εNH	7.75	7.75	7.74
Asn^{14}	NH	8.03	8.04	8.00
	αCH	4.47	4.47	4.46
	βCH	2.56	2.57	2.54
	γNH	7.54; 6.95	7.54; 6.96	7.50; 6.95
	CONHt	7.21; 7.15	7.21; 7.15	7.15; 7.14

Asn^{14} (Figure 1a). First, it is very likely due to the Gly^5-Arg^6 pair by considering the existence of NOE contacts between the Glu^7 NH and the Gly^5 $\alpha\alpha'$ protons (Figure 1b). Furthermore, it may be also assigned to the pairs Ala^{12}-Lys^{13} and Lys^{13}-Asn^{14} analogously to the other peptides since the C-teminus is unchanged.

Again, these results support the possibility of β-turns in the segments Ala^4-Glu^7 and Gly^{11}-Asn^{14}, as in the peptide previously discussed.

Similar conclusions can be drawn for Som III, where NOE contacts among the NH resonances of the pairs Ser^6-Asn^7, Ala^{12}-Lys^{13} and Lys^{13}-Asn^{14} have been observed.

In the case of Som I a slightly different NOE pattern is observed. Cross-peaks between Glu^7-Arg^8 NHs and between Arg^6-Glu^7 NHs are observed. From these data, the presence of a β-turn, involving the segment Ala^5-Arg^6-Glu^7-Arg^8 can be postulated. This result is interpreted in terms of a β-turn shifted by one amino acid towards the C-terminus of the molecule, thus involving the Arg^8 residue which is part of the dibasic site.

The same NOE patterns have been identified both in DMSO and DMSO/H_2O.

Computational Results

As already pointed out in the Materials and Methods section, in the course of energy minimizations the dielectric constant was considered both distance dependent [ε(r)] and constant with a value of 40. Calculations with a distance dependent dielectric constant gave rise to unphysical results due to the larger number of highly polar residues close in the sequence. Calculations carried out with a dielectric constant of 40 represented, instead, better solutions and yielded physically reliable models. Such a high dielectric constant has already been applied to other analogous systems [38] and has been shown to affect very little the backbone folding when compared with *in vacuo* calculations. Sidechains, on the contrary, are affected by different dielectric constants, especially charged amino acid residues, where electrostatic interactions become predominant at lower dielectric values. For these reasons, in the following discussion we just refer to the simulations with a dielectric constant of 40.

As NMR data do not help in discriminating between type I and II β-turns because of a large spectral overlap in the region of interest, both types have been included in the 4–7 segment for the calculations. Computational results were quite similar independently of the starting structure. For this reason only the data for calculations with a β-turn I have been shown.

The energy contributions for the minimized structures are reported in Table 5. It is worth noting that for peptide Som II two possible turns (β-turn type I and I′) in the 4–7 segment have been considered,

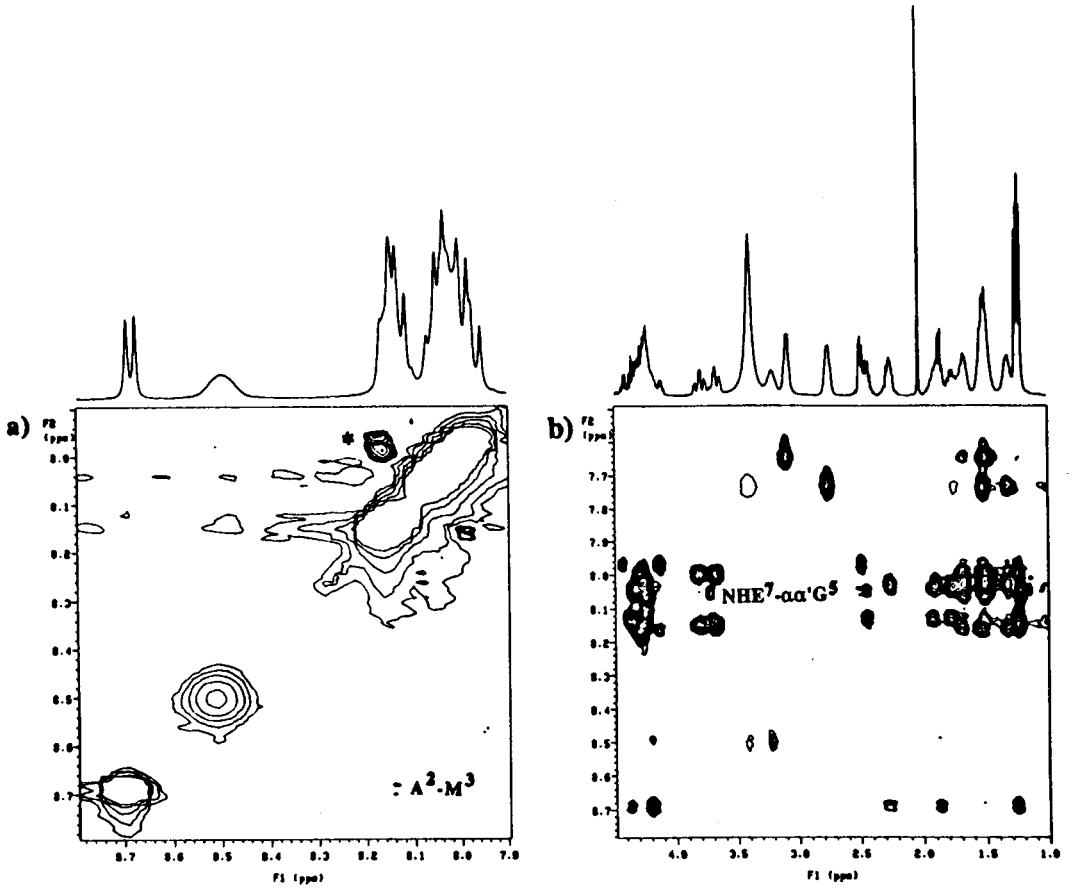

Figure 1 400 MHz ROESY spectrum acquired for peptide Som II in DMSO at 298 K: (a) NH–NH correlation region; (b) NH-high field region.

because of the Gly residue in position i + 1 [39]. The unrestrained minimizations are always lower in energy, relaxing the strain from the imposed distances essentially by lessening bending, torsional and van der Waals energy contributions.

The restraint energy for SII(40) and SIII(40) is significantly higher in value than that for SWT(40). Analysis of Table 6 indicates that the total number of imposed restraints for SII(40) and SIII(40) is considerably higher than for SWT(40). The main difference resides in the total side-chain number of restraints. This simply means that higher values for the restraint contributions are expected and they essentially depend on side-chain deviations from imposed distances. Since this study is primarily addressed to backbone folding, any irrelevant side-chain deviation should not alter the main final conclusions.

In Figure 2 superimpositions of restrained and unrestrained structures for SWT(40)–SWTU(40), SII(40)–SIIU(40) and SIII(40)–SIIIIU(40) are reported. By relaxing the restraints, no structure undergoes any significant backbone conformational change suggesting that the reached minima are local stable minima.

The total rms deviations from the imposed distances of the restrained and unrestrained structures are reported in Table 7. As expected, upon relaxing the restraints, the rms deviation slightly increases. As far as the native sequence Som WT_b is concerned, most of the observed NOE contacts fall within the imposed ranges and the total rms is the lowest in the series. In general, a high value of the total rms, together with a lower value for the backbone rms, indicates that the higher violation is to be attributed to side-chain contacts.

In the same table, contacts diagnostic of types I and II β-turns have been reported, together with the violations from the expected intensity ranges. The simultaneous presence of low violations for the

Table 5 Energy contributions of the calculated structures[a,b]

Minimized Structure	SWT	SWTU	SII (turn I)	SIIU (turn I)	SII (turn I′)	SIIU (turn I′)	SIII	SIIIU
$E_{stretching}$	3.6	4.9	9.0	5.3	9.0	5.1	8.0	5.0
$E_{bending}$	34.7	25.3	59.9	24.2	56.7	21.6	49.8	25.9
$E_{torsional}$	33.7	30.2	52.0	28.9	62.0	28.8	45.0	26.6
$E_{bending\ out\ of\ plane}$	1.1	0.7	3.1	0.4	4.8	0.4	2.1	0.5
$E_{1-4\ van\ der\ Waals}$	14.6	15.9	39.2	16.6	40.8	16.4	36.7	15.4
$E_{van\ der\ Waals}$	−27.2	−36.6	−23.2	−37.7	−22.2	−41.2	−15.2	−41.2
$E_{1-4\ electrostatic}$	4.8	9.5	3.1	3.1	3.2	3.2	9.4	9.3
$E_{electrostatic}$	−7.6	−7.8	−2.4	−2.5	−2.8	−3.2	−5.3	−5.8
$E_{restraint}$	35.9	0	124.1	0	128.1	0	142.1	0
E_{TOT}	57.7	42.0	140.8	38.2	151.4	30.2	130.5	35.8

[a] Energy contributions refer to simulations carried out with a high value dielectric constant (40).
[b] In the restrained minimizations the $E_{restraint}$ contribution has been subtracted from E_{TOT}.

NH_{i+2}–NH_{i+3} and αCH_{i+1}–NH_{i+2} contacts is diagnostic of a type II β-turn, in contrast with a type I β-turn, in which a simultaneous low violation for NH_{i+1}–NH_{i+2} and NH_{i+2}–NH_{i+3} is expected [36].

In our case the observed violations are slightly in favour of a type II β-turn, although the presence of a type I β-turn can not be excluded. It is then fair to postulate that an equilibrium between the two structures in solution would more likely represent the system in solution.

Table 6 Summary of NOE restraints imposed in the minimizations

NOE restraint	SWT[a]	SII[a]	SIII[a]
Strong	19	32	48
Medium	23	21	17
Weak	6	24	20
Total	48	77	85
$Strong_{backbone}$	10	12	25
$Medium_{backbone}$	6	5	2
$Weak_{backbone}$	2	4	1
$Total_{backbone}$	18	21	28
$Strong_{sidechains}$	9	20	23
$Medium_{sidechains}$	17	16	15
$Weak_{sidechains}$	4	20	19
$Total_{sidechains}$	30	56	57

[a] Minimized structure.

DISCUSSION

Pro-somatostatin contains in the C-terminal region a monobasic as well as a dibasic site from which S-28 and S-14 are, respectively, released.

In analogy with work already done on ocytocin precursors [40,41], in the present study NMR and MM techniques have been used to define the architecture of the peptide region around the dibasic cleavage site of pro-S. This work has been carried out on several peptides where substitutions in positions 5 (Som I, Som II) and 5, 6, 7 (Som III) have been made. This strategy originates from the rationale that any conclusion derived from the conformational analysis of a single peptide, for example the native Som WT_b, can only be of a qualitatively nature. If, on the contrary, a series of highly homologous compounds is investigated by the combined use of different spectroscopic techniques and computational methods, then similar results can be thought of as more convincing from a structural point of view. Thus, conclusions can be extended to correlations with the biological activity.

A previous paper [18] has already reported biological and spectroscopic (CD, FT-IR absorption) data on the peptides examined here (Table 1). By taking those data into account and comparing them with our results, the following conclusions are reached:

(i) All examined peptides show an equilibrium in solution between ordered and randomly coiled structures;
(ii) NMR data suggest the presence of conformers characterized by β-turn structures;

Table 7 Characteristic distances and violation of the calculated structures[a,b]

	SWT(40)[c]	SWTU(40)[c]	SII(40)[c] (turn I)	SIIU(40)[c] (turn I)	SII(40)c (turn I′)	SIIU(40)[c](turn I′)	SIII(40)[c]	SIIIU(40)[c]
rms_{tot}	0.25	0.49	0.35	0.97	0.35	1.1	0.37	0.70
$rms_{backbone}$	0.30	0.47	0.33	0.44	0.32	0.4	0.31	0.44
Dist CO_i–NH^d_{i+3}	2.6	2.5	2.6	2.2	2.9	4.8	2.7	2.1
Violation (1.5–2.5)[e]	0.1	0.0	0.1	0.0	0.4	2.3	0.2	0 0
Dist $NH_{i+1}/\delta\delta'_{i+1-}$ NH_{i+2}	4.0	4.0	3.9	4.4	4.1	4.5	4.4	4.7
Violation (2.5–3.5)[e]	0.5	0.5	0.4	0.9	0.6	1.0	0.9	1.2
Dist NH_{i+2}–NH_{i+3}	2.6	3.2	2.5	2.7	2.3	3.0	2.8	3.4
Violation (1.5–2.5)[e]	0.1	0.7	0.0	0.2	0.0	0.5	0.3	0.9
Dist NH_{i+2}–$\alpha\alpha'_{i+1}$	3.6	3.6	3.1; 2.1	3.5; 2.4	3.0; 2.1	2.5; 2.8	2.5	2.5
Violation (1.5–2.5)[e]	1.1	1.1	0.6; 0.0	1.0; 0.0	O.5; 0.0	0.0; 0.3	0.0	0.0

[a] Total rms deviation (Å) from the NOE restraints have been calculated from the violations of upper and lower imposed bounds.
[b] The subscript i is referred to the residue in position 4.
[c] Minimized structure.
[d] Typical hydrogen bonding distance for a β-turn structure.
[e] Upper and lower bounds in Å of the expected distance range.

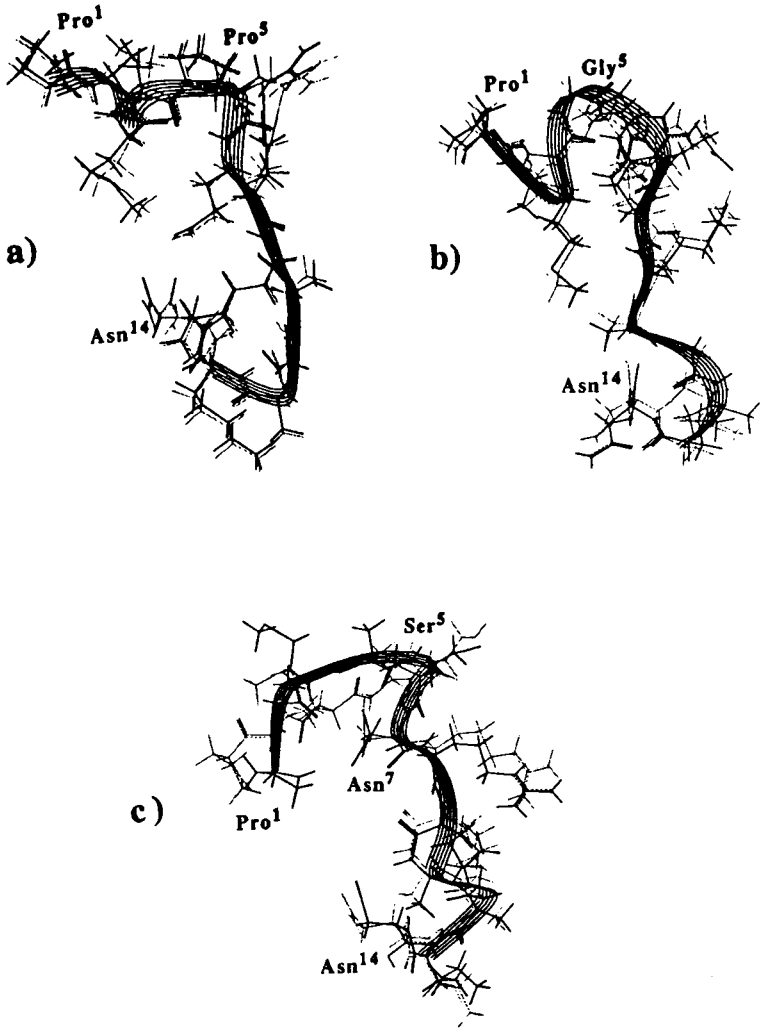

Figure 2 Superimposition of structures from restrained and unrestrained simulations for: (a) SWT–SWTU; (b) SII–SIIU; (c) SIII–SIIIU.

(iii) Computational methods indicate that the postulated structures are energetically stable and physically reliable;
(iv) Peptides showing common structural elements exhibit the same behaviour as substrates in the presence of the characteristic processing enzyme.

In particular, for peptides Som WT_b, Som II and Som III, all recognized as substrates by the enzyme [18], the NOE data indicate the occurrence of β-turns in the 4–7 and 11–14 regions of the polypeptide chain, on the left and right sides of the dibasic site (Arg^8-Lys^9), respectively. Therefore, substitution of the Pro in position 5 (Som WT_b) with Gly or Ser does not seem to affect the stability of the β-turn. On the contrary, in Som I, where Pro^5 is substituted by Ala, the NOE contacts suggest a β-turn between Ala^5-Arg^8 and Gly^{11}-Asn^{14}. In this case, the first β-turn is shifted by one amino acid and includes the first residue (Arg^8) of the cleavage site.

The structural NMR information have been confirmed by a series of energy minimizations. Experimental NMR data are generally consistent with the minimizations carried out. In the final structures the turns in the regions 4–7 and 11–14 are found to be quite stable, even though modest deviations from the canonical β-turns are observed.

Interestingly, these findings fit very well with the CD and biological data [18]. Som WT_b, Som II and Som III are all recognized by the enzyme, while Som I is not.

We believe that our structural data can explain this behaviour in that a β-turn including one of the amino acids (Som I) of the dibasic pair modifies the structure that is no longer recognized by the enzyme.

The β-turn motif seems, then, a quite general motif for those prohormone molecules where cleavage occurs in the proximity of a dibasic Arg-Lys pair. This has already been suggested in the case of pro-ocytocin fragments [4,5] and is confirmed again in the present study.

Acknowledgements

The program DIANA was kindly provided by Professor K. Wutrich, Institute für Molekularbiologie und Biophysik ETH, CH-8093 Zürich, Switzerland.

REFERENCES

1. N. Brakch, H. Boussetta, M. Rholam and P. Cohen (1989). Processing endoprotease recognizes a structural features at the cleavage site of peptide prohormones. *J. Biol. Chem. 264*, 15912–15916.
2. P. Cohen, S. Gomez, P. Gluschankof, C. Clamagirand, I. Plevrakis, N. Brakch, C. Créminon, A. Lepage, A. Morel, M. Rholam and M. Camier. 'Proteolytic processing of polypeptide hormone precursors' in: *The Second Forum Peptides*, A. Aubry, M. Marraud and B. Vitoux, Eds., Vol. 174, p. 3–9, Colloque INSERM, Libbey Eurotext, London 1989.
3. M. Rholam, P. Cohen, N. Brakch, A. Scatturin, L. Paolillo and C. Di Bello (1990). Evidence for β-turn structures in model peptides reproducing pro-ocytocin/neurophysin proteolytic processing site. *Biochem. Biophys. Res. Commun. 168*, 1066–1073.
4. L. Paolillo, M. Simonetti, N. Brakch, G. D'Auria, M. Saviano, M. Dettin, M. Rholam, A. Scatturin, C. Di Bello and P. Cohen (1992). Evidence for the presence of a secondary structure at the dibasic site of prohormone: the pro-ocytocin model. *EMBO J. 7*, 2399–2405.
5. C. Di Bello, M. Simonetti, M. Dettin, L. Paolillo, G. D'Auria, L. Falcigno, M. Saviano, A. Scatturin, G. Vertuani and P. Cohen (1995). Conformational studies on synthetic peptides reproducing the dibasic processing site of pro-ocytocin–neurophysin. *J. Pept. Sci. 1*, 251–265.
6. M. Brown, J. Rivier and W. Vale (1983). Somatostatin-28: selective action on the pancreatic β-cell and brain. *Endocrinology 108*, 2391–2393.
7. J.P. Moreau and F.V. De Feudis (1987). Pharmacological studies of somatostatin and somatostatin-analogs: therapeutic advances and perspectives. *Life Sci. 40*, 419–437.
8. A.V. Schally (1988). Oncological application of somatostatin analogues. *Cancer Res. 48*, 6977–6985.
9. N. Brakch, M. Rholam, C. Nault, G. Boileau and P. Cohen (1991). Differential processing of hormone precursor: independent production of somatostatins 14 and 28 in transfected neuroblastoma 2A cells. *FEBS Lett. 282*, 363–367.
10. A. Morel, P. Kuks, J. Bourdais and P. Cohen (1988). Pro-somatostatin processing in anglerfish brain, gut and pancreas. *Biochem. Biophys. Res. Commun. 151*, 347–354.
11. P. Gluschankof, S. Gomez, A. Morel and P. Cohen (1987). Enzymes that process somatostatin precursors. *J. Biol. Chem. 262*, 9615–9620.
12. J. Bourdais, A.R. Pierotti, H. Boussetta, N. Barre, G. Devilliers and P. Cohen (1991). Isolation and functional properties of an arginine-selective endoprotease from rat intestinal mucosa. *J. Biol. Chem. 266*, 23386–23391.
13. P. Argos, W.L. Taylor, C.D. Minth and J.E. Dixon (1983). Nucleotide and amino acid sequence comparisons of prepro-somatostatins. *J. Biol. Chem. 258*, 8788–8793.
14. A. Lepage-Lezin, P. Joseph-Bravo, G. Devilliers, L. Benedetti, J.-M. Launay, S. Gomez and P. Cohen (1991). Pro-somatostatin is processed in the Golgi apparatus of rat neural cells. *J. Biol. Chem. 266*, 1679–1688.
15. P.Y. Chou and G.D. Fasman (1978). Empirical prediction of protein conformation. *Annu. Rev. Biochem. 47*, 251–276.
16. S. Gomez, G. Boileau, L. Zollinger, C. Nault, M. Rholam and P. Cohen (1989). Site-specific mutagenesis identifies amino acid residues critical in prohormone processing. *EMBO J. 8*, 2911–2916.
17. F.A. Cotton, E.E. Hazen and M.J. Legg (1979). Staphylococcal nuclease: proposed mechanism of action based on structure of enzyme–thymidine 3′,5′-biphosphate–calcium ion complex at 1.5-Å resolution. *Proc. Natl. Acad. Sci. USA 76*, 2551–2555.
18. N. Brakch, G. Boileau, M. Simonetti, C. Nault, P. Joseph-Bravo, M. Rholam and P. Cohen (1993). Pro-somatostatin processing in Neuro2A cells. Role of a β-turn structure in the vicinity of the Arg–Lys cleavage site. *Eur. J. Biochem. 216*, 39–47.
19. M. Rholam, P. Nicolas and P. Cohen (1986). Precursors for peptide hormones share common secondary structures forming features at the proteolytic processing site. *FEBS Lett. 207*, 1–6.
20. D.J. States, R.A. Haberkorn and D.J. Ruben (1982). A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson. 48*, 286–292.
21. L. Braunschweiler and R.R. Ernst (1983). Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *J. Magn. Reson. 53*, 521–528.

22. A. Bax and D.G. Davis (1985). MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson. 65*, 355–360.
23. J. Jeener, B.H. Meier, P. Bachmann and R.R. Ernst (1979). Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys. 71*, 4546–4553.
24. A. Kumar, R.R. Ernst and K. Wütrich (1980). A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun. 95*, 1–6.
25. A.A. Bothner-By, R.L. Stephens, J. Lee, C.D. Warren and R.W. Jeanloz (1984). Structure determination of a tetrasaccharide: transient nuclear Overhauser effects in the rotating frame. *J. Am. Chem. Soc. 106*, 811–813.
26. A. Bax and D.G. Davis (1985). Pratical aspects of two-dimensional transverse NOE spectroscopy. *J Magn. Reson. 65*, 355–360.
27. C. Griesinger and R.R. Ernst (1987). Frequency offset effects and their elimination in NMR rotating-frame cross-relaxation spectroscopy. *J. Magn. Reson. 75*, 261–271.
28. U. Piantini, O.W. Sørensen and R.R. Ernst (1982). Multiple quantum filters for elucidating NMR coupling networks. *J. Am. Chem. Soc. 104*, 6800–6801.
29. C. Griesinger, O.W. Sørensen and R.R. Ernst (1985). Two-dimensional correlation of connected NMR transitions. *J. Am. Chem. Soc. 107*, 6394–6396.
30. C. Griesinger, O.W. Sørensen and R.R. Ernst (1987). Practical aspects of the E. COSY technique. Measurement of scalar spin–spin coupling constants in peptides. *J. Magn. Reson. 75*, 474–492.
31. D.I. Hoult (1976). Solvent peak saturation with single phase and quadrature Fourier transformation. *J. Magn. Reson. 21*, 337–347.
32. M. Clark, R.D. Cramer III and N. Van Opdenbosch (1989). Validation of the general purpose Tripos 5.2 force field. *J. Comput. Chem 10*, 982–1012.
33. H. Berthod and A. Pullmann (1965). Calculation of the σ structure of conjugated molecules. *J. Chim. Phys. 62*, 942–946.
34. W.A. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling in: *Numerical Recipes, the Art of Scientific Computing*, p. 307–312, Cambridge University Press, Cambridge 1986.
35. P. Güntert, W. Braun and K. Wütrich (1991). Efficient computation of three dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol. 217*, 517–530.
36. K. Wütrich in: *NMR of Proteins and Nucleic Acids*, Wiley, New York 1986.
37. O. Jardetzky and G.C.K. Roberts in: *NMR in Molecular Biology*, Academic Press, New York 1981.
38. L. Wesson and D.E. Eisenberg (1992). Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci. 1*, 27–35.
39. G.D. Rose, L.M. Gierasch and J.A. Smith (1985). Turns in peptides and proteins. *Adv. Protein Chem. 37*, 1–109.
40. L. Paolillo, G. D'Auria, L. Falcigno, F. Fraternali, M. Saviano, M. Simonetti and C. Di Bello 'Conformational studies on peptide fragments related to ocytocin and somatostatin precursors', in: *Current Topics in Peptide and Protein Research. 1*, p. 193–222, Research Trends Publ., Trivandrum, India 1994.
41. L. Falcigno, L. Paolillo, G. D'Auria, M. Saviano, M. Simonetti and C. Di Bello (1996). NMR conformational studies on a synthetic peptide reproducing the [1–20] processing domain of the pro-ocytocin–neurophysin precursor. *Biopolymers 39*, 837–848.

