Journal of Peptide Science
J. Peptide Sci. **9**: 125–131 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.438

JOURNAL OF PEPTIDE SCIENCE

Synthesis and Circular Dichroism Study of the Human Salivary Proline-Rich Protein IB7

CÉCILE SIMON,[a] ISABELLE PIANET[a,b] and ERICK J. DUFOURC[a]*

[a] Institut Européen de Chimie et Biologie, FRE CNRS 2247, Pessac, France
[b] Laboratoire de Chimie Organique et Organométallique, UMR CNRS 5802, Talence, France

Received 18 July 2002
Accepted 2 August 2002

Abstract: The solid phase synthesis of a 59 amino acid human salivary protein IB7 has been accomplished using Fmoc strategy. Because the protein contains 25 proline, 13 glycine and 9 glutamine residues the coupling time, piperidine delivery and acetic anhydride reaction time were increased. Yield after HPLC purification was 35%. Circular dichroism studies revealed that about one third of IB7 residues adopted a type II helix secondary structure, as found in collagen helices. The rest of the sequence adopts a random coil secondary structure. Copyright © 2003 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: proline-rich protein; salivary protein; solid phase synthesis; circular dichroism; type II helix

INTRODUCTION

Proline-rich proteins (PRP) are ubiquitous and constitutively expressed in the human body. They are involved in numerous enzymatic reactions or in biological recognition [1]. They are predominant in human saliva and represent up to 70% of the proteins secreted by the parotid gland [2]. About 20 have been identified from a single human donor and sorted into three groups: acidic, alkaline and glycosylated proteins [3]. Acidic PRPs are involved in calcium-binding and inhibit crystal growth [4]. Glycosylated proteins could have lubricating properties especially in the mouth [5]. Basic PRPs do not bear any known biological functions but bind polyphenols very efficiently [4,6]. Nine of them have been sequenced and contain about 40% proline, 21% glycine and 17% glutamine residues. The basic PRP IB7 serves as a building unit for other members of the family because its sequence is often contained in their amino-terminal part [3]. Interestingly, IB7 itself is made of three repeated sequences and of five PQGPP patterns as sketched in Figure 1.

On the other hand, many plant species contain complex polyphenols also named tannins. They act as chemical defences [6,7] and are divided in condensed tannins, i.e. polymers of procyanidin and hydrolysable tannins (mainly galloic tannins). Grape seeds contain a large amount of condensed tannins that are partially responsible for the organoleptic properties of grapes and wines (e.g. astringency and bitterness). During ingestion, tannins interact with many kinds of proteins present in saliva, stomach, intestines and blood [8,9]. The interaction between tannins and salivary proteins is thought to be the primary source of astringency [10,11]. The high affinity of PRPs for polyphenols has been attributed to their open and extended structure and to their high proline content. Murray and coworkers [12] propose by analogy with PRPs studied so far that IB7 would also possess a random coil structure. Other works put forward the existence of a poly(L)proline type II conformation (collagen helix) [13] owing to the high number of proline residues. The favourable interaction between tannins and IB7 has been

* Correspondence to: Dr Erick J. Dufourc, IECB-Polytechnique, 16 Avenue Pey Berland, 33607 Pessac, France;
e-mail: erick.dufourc@iecb-polytechnique.u-bordeaux.fr
Contract/grant sponsor: Conseil Interprofessionnel du Vin de Bordeaux (CIVB).

SPPGKPQGPPPQGGNQPQGPP

PPPGKPQGPPPQGGNKPQGPP

PPGKPQGPPPQGDNKSR[59]

Figure 1 Amino acid sequence of the human salivary proline-rich protein IB7. Underlined fragments are present five times in the total sequence.

proposed to be due to the short size of the protein [6] and to the presence of the repeated (5 times) PQGPP pattern in its sequence.

All the above-mentioned studies of PRP structure and interaction with polyphenols were carried out on peptides of about 20 residues long. Currently the interaction between polyphenols and salivary proteins is being studied and as a representative element of the salivary PRP family, the protein IB7 was chosen. It is a 59 residue protein and because its sequence is contained in 5 of the 9 human basic salivary proteins it is believed to be a good model protein to follow the interaction. The present study reports the challenging solid-phase chemical synthesis of the proline-rich protein IB7 as well as its secondary structure characterization by circular dichroism.

MATERIALS AND METHODS

Peptide Synthesis

The chemical synthesis of IB7 peptide was performed on an Applied Biosystems Peptide Synthesiser 433A (PE Biosystem, Courtaboeuf, France) using Fmoc strategy [14]. Fluorenylmethylocarbonyl (Fmoc) -L-Arg(Pmc)-Wang-resin, (Fmoc) -L-Arg(Pmc)-NovaSyn-resin, *N*-α-Fmoc-protected amino acids, 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HOBt) and *N*-hydroxybenzotriazole (Hbtu) were purchased from Novabiochem (Läufelfingen, Switzerland); *N*-methylpyrrolidone (NMP), piperidine, dichloromethane (DCM), dimethylformamide (DMF), diisopropylethylamine (DIEA), trifluoroacetic acid (TFA) and acetic acid were from SDS (Peypin, France). The *N*-α-Fmoc-amino acids were protected on the side chain as follows: *t*-butyl for serine and aspartic acid, *t*-butoxycarbonyl (Boc) for lysine, trityl (trt) for glutamine and asparagines, and 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) for arginine.

Cleavage of the Protein from the Resin

The final peptide mixture was cleaved from the resin and removal of the protecting groups accomplished by TFA/thioanisol/Milli Q-water (93/3/2/2 by volume). Total deprotection was achieved by allowing a reaction time of 150 min at 0 °C. TFA was evaporated under reduced pressure. The crude peptide mixture was then precipitated in cold diethyl ether and the cloudy aqueous phase was freeze dried.

Purification and Analysis

The powder was dissolved in distilled water and analysed by HPLC (Waters 2487 dual wavelength absorbance detector) using a Nova Pak C_{18} column (5 µm, 100 Å, 3.9 × 150 mm) reverse-phase column (Saint-Quentin-en-Yvelynes, France). The elution was carried out with 0.1% aqueous TFA (eluent A) and acetonitrile (0.08% TFA, eluent B) at a flow rate of 1 ml/min. Absorption was monitored at 214 and 254 nm. Maximum resolution for the synthetic peptide mixture was achieved with a 30 min linear gradient, where eluent B varied from 18% to 45%. Semi-preparative HPLC was performed with a Waters Delta Pak C_{18} column (15 µm, 100 Å, 300 × 7.8 mm) at a flow rate of 2.5 ml/min under the same elution conditions.

Mass Spectroscopy (MS)

Matrix assisted laser desorption and ionization time of flight (MALDI TOF) spectra were obtained on a Reflex III mass spectrometer (Bruker Daltonik, Bremen, Germany). The crude product was loaded on the target using the dried droplet method [15]. Sinapic acid was chosen as the matrix. Spectra were externally calibrated using the $[M+H]^+$ ions of two peptide standards. The amount of peptide deposited on each spot was typically 3–5 pmol.

Circular Dichroism (CD)

Far-ultraviolet (190–270 nm) circular dichroism spectra were recorded on a Mark VI Jobin-Yvon dichrograph (Longjumeau, France), calibrated using iso-androsterone (Roussel-Uclaf, France) in dioxane and camphorsulfonic acid (Sigma-Chimie, France). The protein was dissolved in different solvents: water or water/ethanol (88/12 v/v) to mimic the composition of wine. Each one was done in the absence or presence of $CaCl_2$. All experiments were performed at room temperature, pH 3.5, and by using cells of

1 and 0.1 mm path length. Poly(L)proline (PLP, average amino acid number of 5500, Sigma, St Louis, MO, USA) was dissolved under the same conditions, and home made poly(L)lysine (20 amino acids long, PLL_{20}) was prepared at pH 7.

RESULTS AND DISCUSSION

Chemical Synthesis

The feasibility of a peptide solid phase synthesis depends on the nature and frequency of amino acid residues in the protein. Hydrophobic peptides are known to be difficult to synthesize because of their poor water solubility [16]. Other amino acids may also bring some problems. For example, proline promotes a turn formation that may prevent the attachment of the following residue. On the contrary, glycine, which is the least bulky amino acid, has less availability for peptide elongation due to its capacity for entering into the small cavities of the resin beads. The IB7 composition (Figure 1) is made of 59 amino acids among which are 25 prolines, 13 glycines and 9 glutamines. Despite this challenging composition, solid phase synthesis was carried out using the FastMoc strategy [14] in double coupling mode, with 2.5 mmol of resin and a four-fold excess of amino acids. In order to optimize the coupling yield several steps needed to be modified. The coupling time between amino acids and the peptide in elongation was extended up to 30 min per residue. Likewise,

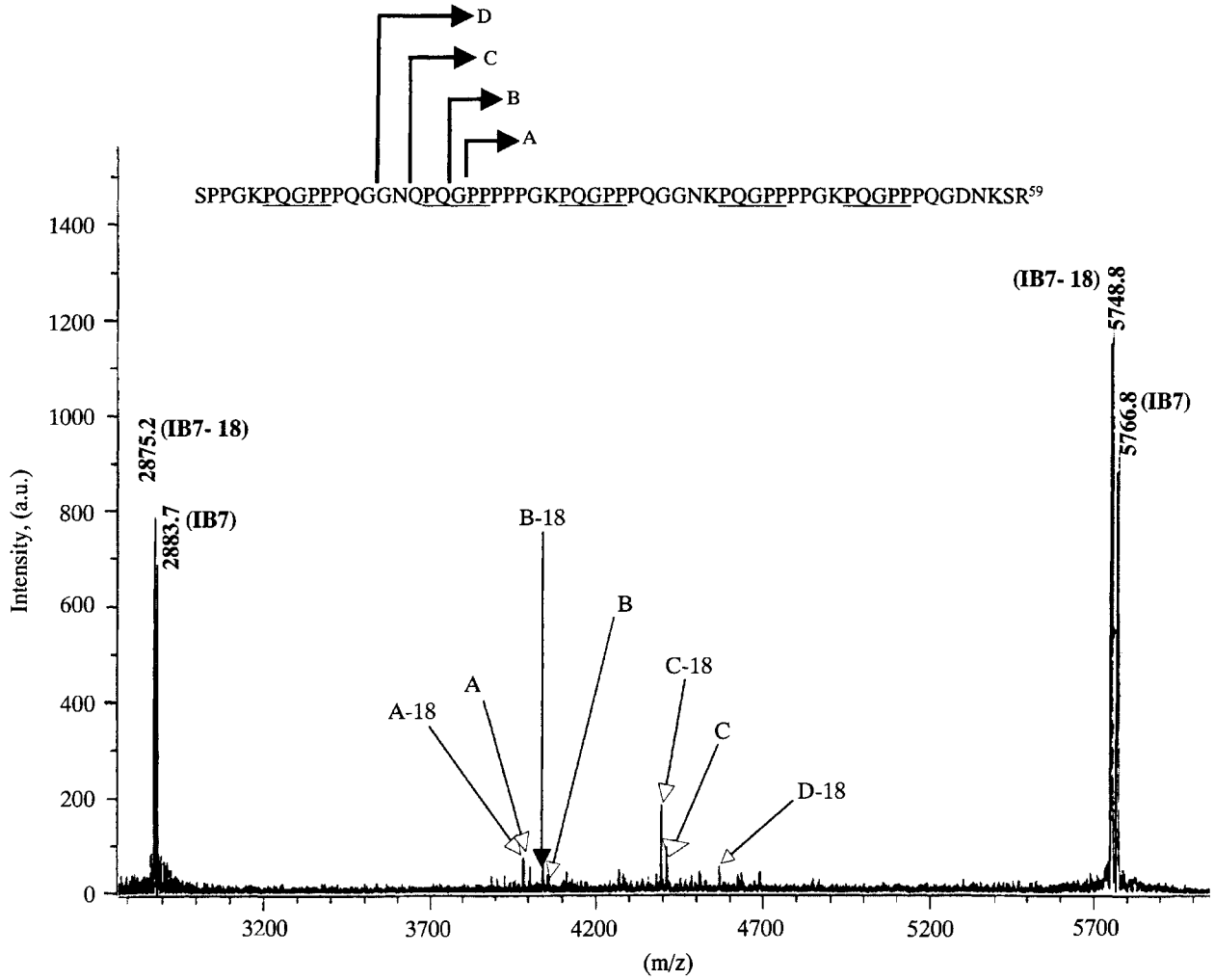

Figure 2 MALDI-TOF MS spectrum of the crude reaction mixture after dissolution in water. Major peaks correspond to the simply (1^+) and doubly (2^+) charged species of IB7 and IB7-18. Secondary peaks represent truncated peptides with or without a loss of 18. Some of the truncated peptides are schematized at the top of the figure.

increasing the delivery time of piperidine optimized the yield of the α-Fmoc-NH_2 groups cleavage. This implementation can be done by continuously monitoring the conductivity change that follows the removal of the Fmoc groups. Because the access to the chain was expected to be difficult, the reaction time of acetic anhydride as a capper was increased (5%). This avoided the formation of hybrid peptides with missing residues. In order to increase the yield it appeared judicious to use a Nova resin rather than the Wang resin classically used for long peptides (Novabiochem, France).

Purification

The peptide was cleaved from the resin by TFA treatment. In the same cleavage mixture ethanedithiol was added in order to remove the side chain protecting groups. MS (MALDI-TOF) was used to check the content of the crude synthesis products (Figure 2). Two major peaks were identified, each one being split into mono-protonated (1^+) and di-protonated (2^+) forms. IB7 has an exact mass of 5765.9 g.mol^{-1} that corresponds to the peak of 5766.8 m/z for the (1^+) form and of 2883.7 m/z for the (2^+). Peaks at m/z of 5748.8 and 2875.2 correspond to the mass of the protein minus 18 for (1^+) and (2^+) forms, respectively. This deserves a comment: a central problem in the chemical synthesis of aspartic acid-containing protein is aspartimide formation [17]. Some factors promote this reaction such as acidic solvent, and the presence of Asp-Gly or Asp-Asn sequences that might involve imide formation. Because residue 55 in IB7 is aspartate, followed by asparagine the conditions favour aspartimide formation. Its presence could be avoided in decreasing the cleavage time with TFA. Unfortunately, this would render the removal of some protecting groups very difficult. Because aspartate is in position 55 in the IB7 sequence, i.e. almost at the protein end, aspartimide formation is expected to have little importance for further structural studies.

In the MS spectrum other minority peaks are present and correspond to the truncated peptides of the bulk reaction mixture. Figure 3 shows some of them, e.g. acetylated fragments, from Arg^{59}-Pro^{20} to Arg^{59}-Gln^{14} with m/z ranging from 3984.5 to 4584.3. The presence of these peptides could be explained by the occurrence of five consecutive proline residues from position 22–18 (see Figure 3), which could prevent efficient elongation due to possible turns formation.

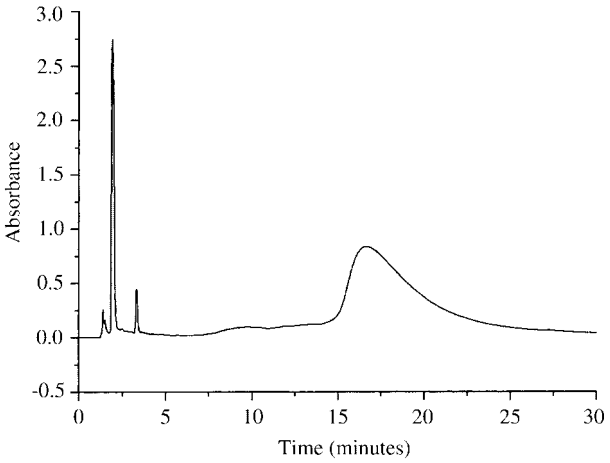

Figure 3 HPLC profile of the crude product using reverse phase chromatography on an analytic C18 column. Elution took place within 30 min using a linear gradient where eluent B (see text) varied from 18% to 45%. The first peaks in the chromatogram originated from the protecting groups and the major fraction eluted between 15 and 20 min contained a mixture of IB7 and truncated peptides.

A very broad and major peak mainly constitutes the resulting HPLC profile (Figure 4). Indeed, all of the peptides (including the truncated ones) that were obtained after cleavage from the resin were eluted at similar elution times because of their composition was the same. The purification was carried out by a fine fractionation of the major peak, and the purity in each fraction was checked by MALDI TOF mass spectrometry. Figure 4 shows a MS spectrum of a fraction that contained only IB7 and IB7-18. Both peptides were non-dissociable by the HPLC. The total yield after purification was 35%.

Secondary Structure Analysis by Circular Dichroism

Circular dichroism was used to assess the secondary structure of the synthetic protein. The IB7 spectrum in water is presented in Figure 5 with a minimum ellipticity at 200 nm of $[\theta] = -20000$ deg.dmol^{-1}.cm^2. No change was noticed when increasing the peptide concentration from 50 µM to 2 mM. Because several studies [12,18] predicted a random coil conformation for PRP proteins IB7 was compared with poly(L)lysine (PLL_{20}) that adopts a random coil under specific conditions such as neutral pH at room temperature. As IB7 contains 40% proline, it was also compared with the left-handed type II helical structure of the poly(L)proline, PLP

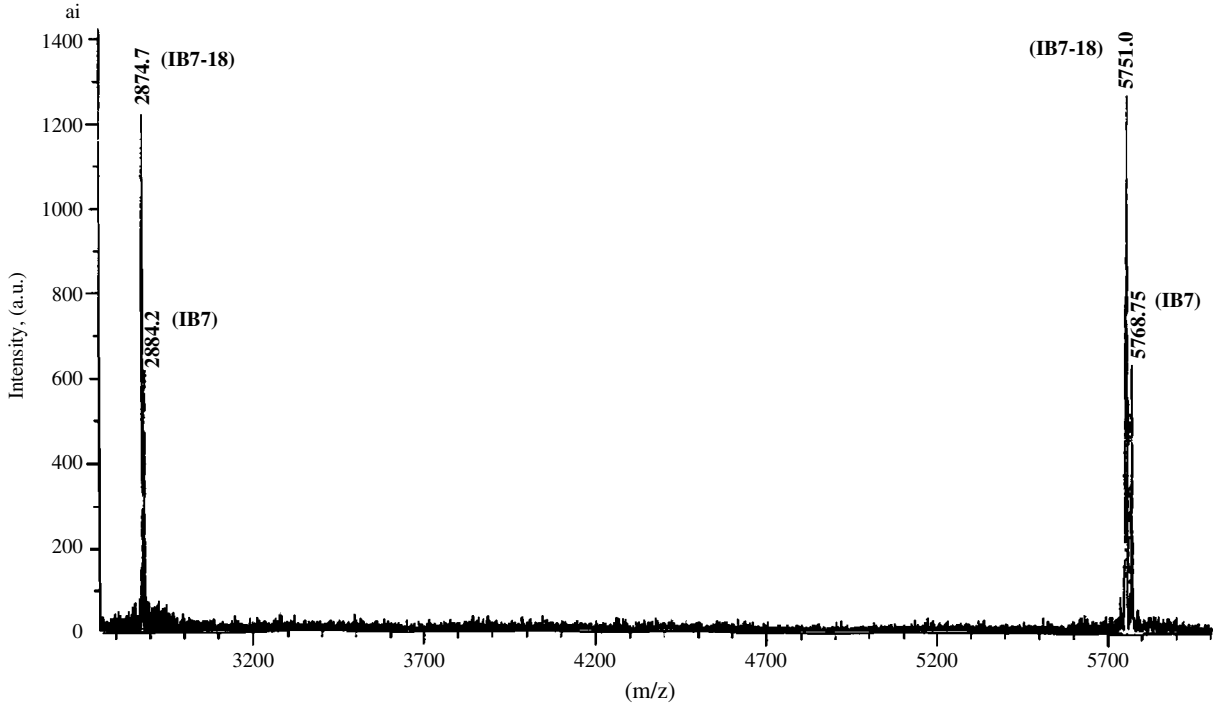

Figure 4 MALDI TOF spectrum of the purified peptide after fractionation by HPLC. The two major species IB7 and IB7-18 are well characterized by their (1^+) and (2^+) ions.

(see Figure 5). PLL_{20} exhibited a minimum ellipticity at 196 nm of $[\theta] = -12000$ deg.dMol^{-1}.cm^2, whereas PLP showed a minimum at 206 nm of $[\theta] = -45000$ deg.dMol^{-1}.cm^2. The IB7 spectrum clearly stands in between. The IB7 curve was submitted to a spectral deconvolution that accomplishes linear combinations of random, alpha helix, beta sheet and helix II CD standard curves (FORTRAN routine, EJ Dufourc, unpublished). Use of 64% random coil, 0% alpha helix, 0% beta sheet and 36% helix II standard curves could best simulate the IB7 experimental CD data. Interestingly the proline content is 42% and the amount of proline residues engaged in sequences of three or more proline blocks is 31%. This suggests a strong relationship between the number of proline blocks and a helix of type II secondary structure.

IB7 was also dissolved in water/ethanol (88/12) in order to mimic the composition of wine. Similar results were obtained in water and water/ethanol (data not shown). Because calcium ions are known to promote secondary structure randomization [19] CD spectra were recorded in the presence of 4 mM $CaCl_2$. As seen in Figure 5 calcium addition did not change the observed CD profile. As a control, the CD spectrum of PLP were recorded also in the presence of the same calcium concentration (Figure 5). No change was observed, suggesting that 4 mM $CaCl_2$ is not able to denature helices of type II.

CONCLUSIONS

This is the first time, to our knowledge, that an entire proline-rich human protein has been chemically synthesized with the correct yield in spite of a large number of proline and glycine residues in the sequence. Aspartimide formation at the end of the sequence could not be avoided, but is expected not to be crucial for structural studies. A major finding is that the secondary structure of IB7 is not entirely in a random coil, as Murray and co-workers would have predicted. About one third of its amino acids are in a helix of type II conformation (collagen helix). It is noteworthy that a similar secondary structure prediction was already reported for IB9, a PRP very similar to IB7, by using CD, proton NMR spectroscopy [20] and by hydrophobicity profile analysis [13]. In order to further detail our findings for IB7, NMR studies combined with molecular modelling are under way in our laboratory.

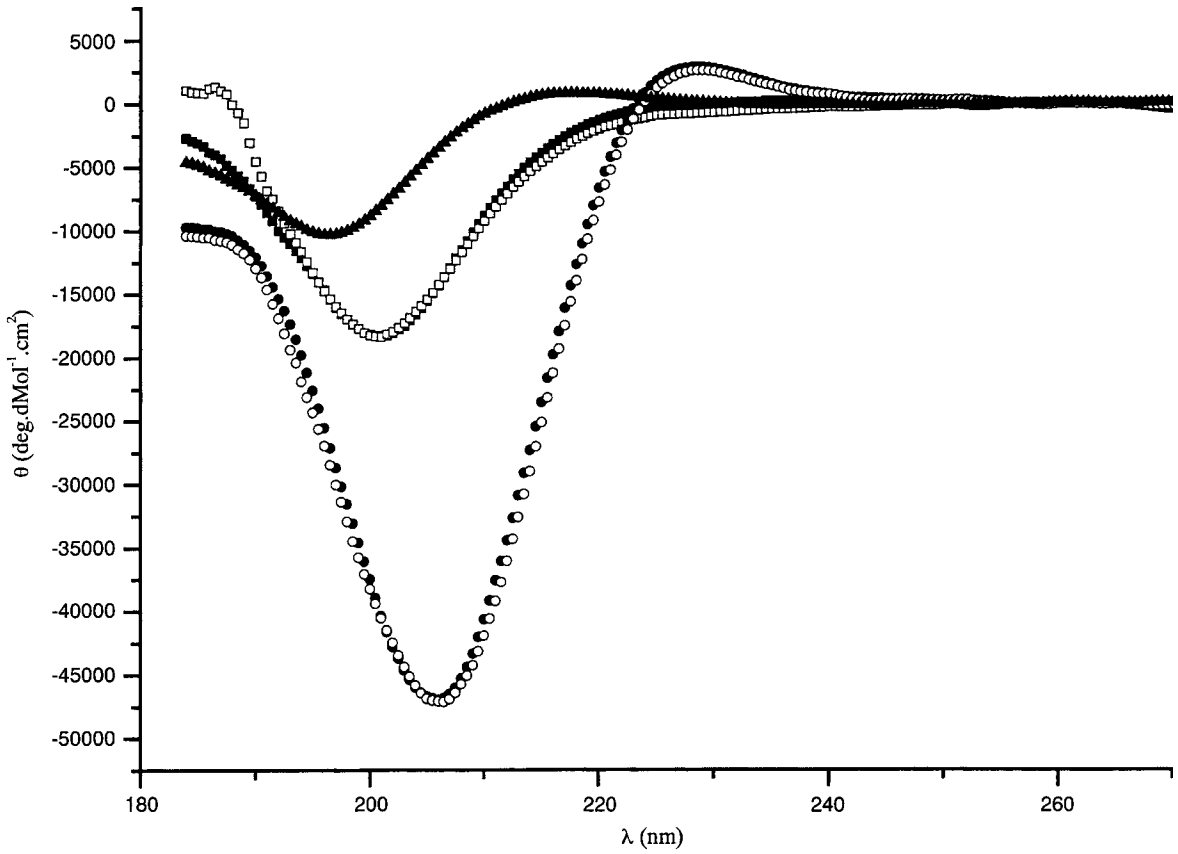

Figure 5 Far-ultraviolet CD spectrum for IB7 (200 μM) and standards (PLL_{20}, 20 μM, PLP, 16 μM) in water. Top curve (▲) shows the random coil conformation of a poly(L)lysine of 20 residues at pH 7 and room temperature. Curve at the bottom (●) is characteristic of poly(L)proline (PLP) type II helix. In the middle is shown the spectrum of IB7 (■). Plots with empty symbols (○,□) present the effect of a final concentration of 4 mM $CaCl_2$.

Acknowledgements

The authors are indebted to Professor Jean-Marie Schmitter and Katel Bathany for mass spectrometric analyses.

REFERENCES

1. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* 2000; **14**: 231–241.
2. Kauffman DL, Keller PJ. The basic proline-rich proteins in human parotid saliva from a single subject. *Arch. Oral. Biol.* 1979; **24**: 249–256.
3. Kauffman DL, Bennick A, Blum M, Keller PJ. Basic proline-rich proteins from parotid saliva: relationships of the covalent structures of ten proteins from a single individual. *Biochemistry* 1991; **30**: 3351–3356.
4. Bennick A, McLaughlin AC, Grey AA, Madapallimattam G. The location and nature of calcium-binding sites in salivary acidic proline-rich phosphoproteins. *J. Biol. Chem.* 1981; **256**: 4741–4746.
5. Hatton MN, Loomis RE, Levine MJ, Tabak LA. Masticatory lubrication. The role of carbohydrate in the lubricating property of a salivary-albumin complex. *Biochem. J.* 1985; **230**: 817–820.
6. Lu Y, Bennick A. Interaction of tannin with human salivary proline-rich proteins. *Arch. Oral Biochem.* 1998; **43**: 717–728.
7. Beart JE, Lilley TH, Haslam E. Plant polyphenols — secondary metabolism and chemical defense: some observations. *J. Chem. Soc. Perkin Trans. II* 1985; **24**: 33–38.
8. Haslam E, Lilley TH, Cai Y, Martin R, Magnolato D. Traditional herbal medicines — the role of polyphenols. *Plant. Med.* 1989; **55**: 1–8.
9. Foley WJ, MacArthur C. *The Effects and Costs of Allelochemicals for Mammaliam Herbivores: An Ecological Perspective.* Cambridge University Press: Cambridge, 1994; 370–391.

10. Bate-Smith EC. Astringency in foods. *Foods* 1954; **23**: 124.
11. Prinz JF, Lucas PW. Saliva tannin interactions. *J. Oral Rehab.* 2000; **27**: 991–994.
12. Murray NJ, Williamson MP, Lilley TH, Haslam E. Study of the interaction between salivary proline-rich proteins and a polyphenol by 1H-NMR spectroscopy. *Eur. J. Biochem.* 1994; **219**: 923–935.
13. Cid H, Vargas V, Bunster M, Bustos S. Secondary structure prediction of human salivary proline-rich proteins. *FEBS Lett.* 1986; **198**: 140–144.
14. Fields CG, Lloyd DH, MacDonald RL, Otteson KM, Noble RL. HBTU activation for automated Fmoc solid-phase peptide synthesis. *Pept. Res.* 1991; **4**: 95–101.
15. Kussmann M, Nordhoff E, Rahbek-Nielsen H, Haebel S, Rossel-Larsen M, Jakobsen L, Gobom J, Mirgorodskaya E, Kroll-Kristensen A, Palm L, Roepstorff P. Matrix-assisted laser desorption/ionization mass spectroscopy sample preparation techniques designed for various peptide and protein analytes. *J. Mass Spectr.* 1997; **32**: 593–601.
16. Goetz M, Schmitter JM, Geoffre S, Dufourc EJ. Chemical synthesis of yeast mitochondrial ATP synthase membranous subunit 8. *J. Pept. Sci.* 1999; **5**: 245–250.
17. Tam JP, Riemen MW, Merrifield RB. Mechanisms of aspartimide formation: the effects of protecting groups, acid, base, temperature and time. *Pept. Res.* 1988; **1**: 6–18.
18. Murray NJ, Williamson MP. Conformational study of a salivary proline-rich protein repeat sequence. *Eur. J. Biochem.* 1994; **219**: 915–921.
19. Tiffany ML, Frimm S. Circular dichroism of poly-L-proline in an unordered conformation. *Biopolymers* 1968; **6**: 1767–1770.
20. Shibata S, Asakura J, Isemura T, Isemura S, Saitoh E, Sanada K. Conformational study of the basic proline-rich polypeptides from human parotid saliva. *Int. J. Pept. Prot. Res.* 1984; **23**: 158–165.

