Basic Proline-Rich Proteins from Human Parotid Saliva: Complete Covalent Structures of Proteins IB-1 and IB-6†

Dorothy Kauffman,‡ Theo Hofmann,§ Anders Bennick,§,‡ and Patricia Keller*,‡

Department of Oral Biology, University of Washington, Seattle, Washington 98195, and Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A8

Received July 22, 1985; Revised Manuscript Received December 10, 1985

ABSTRACT: The complete amino acid sequences of two basic proline-rich proteins, IB-1 and IB-6, from human parotid saliva have been determined. Fragments for sequence analysis were obtained by enzymatic digestions. The proteins have molecular weights of 9571 (IB-1) and 11 530 (IB-6) and contain 34 and 39 mol % proline, respectively. IB-1 and IB-6 contain an identical sequence of 54 residues except for an alanine in position 52 of IB-6, where IB-1 has proline. An unusually high number of repeated sequences occurs in both molecules. IB-1 has a blocked amino-terminal residue, pyroglutamic acid, and also contains one phosphoserine residue in position 8. The relationship of these proteins to the basic proline-rich protein IB-9 [Kauffman, D., Wong, R., Bennick, A., & Keller, P. (1982) *Biochemistry 21*, 6558–6562] and to other salivary proline-rich proteins is discussed.

Proline-rich proteins and glycoproteins are major secretory products of parotid and submandibular glands of humans (Bennick, 1982), subhuman primates (Arneberg et al., 1976; Belford et al., 1984), and other mammals (Keller et al., 1975; Muenzer et al., 1979a,b; Rajan & Bennick, 1983). More than 20 such components have been isolated from human parotid saliva during the past decade, and they are now recognized as constituting the greater proportion (up to two-thirds) of proteins secreted by this gland (Kauffman & Keller, 1979). They can be categorized according to their charge, acidic or basic, and according to whether or not they are glycosylated. Biochemical and genetic evidence exists that the acidic, basic, and glycosylated proline-rich proteins are closely related.

Four acidic proline-rich proteins have been extensively studied with respect to their primary and secondary structures (Schlesinger & Hay, 1979; Wong et al., 1979; Wong & Bennick, 1980) and their intracellular processing (Troxler et al., 1983). Nine nonglycosylated basic proline-rich proteins, designated IB-1 to IB-9, and several glycosylated basic proline-rich proteins have been isolated from human parotid saliva and characterized to varying degrees (Levine & Keller, 1977; Kauffman & Keller, 1979; Levine et al., 1969). We recently reported the complete covalent structure of one basic proline-rich protein (IB-9) and the partial sequence of another (IB-6) (Kauffman et al., 1982). The present paper reports the total covalent structure of IB-6 and that of another major basic proline-rich protein (IB-1).

Materials and Methods

Clostripain, trypsin (TPCK[1] treated), and carboxypeptidase B (PMSF treated) were obtained from Worthington, pyroglutamate aminopeptidase, carboxypeptidase B–DFP, carboxypeptidase P, phosphoserine, and *N*-ethylmorpholine were from Sigma, and elastase suspension, 2× crystallized, was from Mann Research Laboratories. Bio-Gels were obtained from Bio-Rad and Sephadexes from Pharmacia. Dansyl chloride, dansyl amino acids, PITC, F_3CCOOH, carboxypeptidase Y, and *Staphylococcus aureus* V8 protease were obtained from Pierce. Chen Ching polyamide sheets were obtained from Gallard-Schlesinger Chemical Corp., and dithiothreitol was obtained from Calbiochem-Behring Corp. Amino acid analyses were performed by AAA Laboratories, Mercer Island, WA.

Isolation of Salivary Proteins IB-1 and IB-6. IB-1 and IB-6 were prepared from human parotid saliva of a single individual as previously described (Kauffman & Keller, 1979).

Protease Digestions. Clostripain digestions were carried out in 0.05 M ammonium bicarbonate, pH 8, containing 2.5 mM dithiothreitol, at 37 °C for 2 (IB-1) and 16 h (IB-1A). The E:S ratio was 1:50 by weight. Elastase digestion was performed in 0.1 M *N*-ethylmorpholine, pH 8.5, for 16 h at 37 °C (molar E:S ratio, 1:50). TPCK–trypsin digestions of IB-6 were done in 0.1 M *N*-ethylmorpholine, pH 8.5, for 17 h at 37 °C (E:S ratio, 1:20 by weight). *S. aureus* V8 protease digestion was done in 0.1 M ammonium bicarbonate, pH 8, for 16 h at 37 °C (1:35 E:S molar ratio). PMSF–carboxypeptidase B digestions were carried out in 0.1 M *N*-ethylmorpholine, pH 8.5, for 2 and 16 h at 37 °C, at a molar E:S ratio of 1:100. Carboxypeptidase Y digestion was done in 0.1 M pyridine acetate, pH 6, at 37 °C for 4 and 8 h (E:S ratio, 1:20 by weight). After each time period the enzyme was inactivated by heating at 100 °C for 5 min.

Purification of Peptides. Clostripain, elastase, and *S. aureus* V8 protease digests of IB-1 were purified by paper electrophoreses at pH 6.5 (3 kV, 45 min, or 4 kV, 20 min) and at pH 1.9 (4 kV, 25 min). Peptides were eluted from the paper with 0.05 M acetic acid or 0.05 M ammonium bicarbonate. Tryptic peptides of IB-6 were isolated by SP-Sephadex C25 chromatography (1.5 × 104 cm column) in 0.025 M sodium phosphate, pH 6.75, with a linear NaCl gradient to 0.3 M. The peptides were desalted on a Bio-Gel P-2 column equilibrated with 0.05 M ammonium hydrogen carbonate. Peaks 4 and 7 were further purified on columns of SP-Sephadex C25

† This work was supported by National Institutes of Health Grant DE-05251 and Canadian MRC Grant MT 4920.

‡ University of Washington.

§ University of Toronto.

[1] TPCK, L-1-(tosylamido)-2-phenylethyl chloromethyl ketone; PMSF, phenylmethanesulfonyl fluoride; PITC, phenyl isothiocyanate; DFP, diisopropyl fluorophosphate; HPLC, high-performance liquid chromatography; PTH, phenylthiohydantoin; dansyl, 5-(dimethylamino)-naphthalene-1-sulfonyl; EDTA, ethylenediaminetetraacetic acid; P_i, inorganic phosphate.

0006-2960/86/0425-2387$01.50/0

(0.9 × 84 cm) equilibrated with 0.05 M ammonium carbonate, pH 3.7, by using a 300-mL linear ammonium chloride gradient to 0.3 M. Peak 6 was further purified on paper high-voltage electrophoresis at pH 1.9.

Pyroglutamate Aminopeptidase Digestions. Intact IB-1 and the acidic peptide A were subjected to digestion with pyroglutamate aminopeptidase. Digestions were carried out in 0.05 M potassium phosphate, pH 7.4, containing 5 mM EDTA and 5 mM dithiothreitol, with an enzyme:substrate ratio of 1:20 by weight. The digests were incubated at 37 °C overnight and then applied to a Bio-Gel P-10 column (1 × 80 cm) equilibrated with 0.05 M ammonium hydrogen carbonate. The protein or peptide peak, after lyophilization, was subjected to amino acid analysis and NH_2-terminal analysis. The peak containing the pyrrolidonecarboxylic acid was lyophilized, dissolved in acetone, dried, and hydrolyzed with 6 N HCl.

Identification of Phosphoserine. Sequence analysis of the acidic peptide A showed close resemblance to the NH_2 terminal of the acidic proline-rich proteins A and C. The latter proteins contain phosphoserine; therefore, the acidic peptide was analyzed for the presence of phosphoserine. A limited acid hydrolysis was performed in 6 N HCl at 108 °C for 2 h, followed by high-voltage paper electrophoresis at pH 1.9. A marker of phosphoserine was placed next to the digest. The part of the digest that had the same mobility as phosphoserine was eluted and subjected to amino acid analysis.

Phosphorus Analysis. The quantification of phosphoserine was performed by analysis for P_i produced by alkaline hydrolysis (Martinsen, 1984). The protein (0.5 mg) was dephosphorylated in 0.3 mL of 1 N NaOH at 37 °C for 21.5 h. Samples of 25, 50, and 75 μL were taken for analysis by the method of Itaya and Ui (1966). Blanks and standards were prepared with the same amounts of 1 N NaOH that were present in the samples. Peptides (3–6 nmol) were analyzed in the same manner.

Amino Acid Analyses. Peptides and protein were hydrolyzed and analyzed by AAA Laboratories. Hydrolysis was performed in 6 N HCl at 110 °C for 24 h.

NH_2-Terminal Analysis. NH_2-terminal amino acids were determined by using the dansyl procedure for peptides (Gray, 1967).

Sequence Analysis: Manual. Manual dansylation and Edman degradation was performed by the micro method of Bruton and Hartley (1970). This method was used to sequence the five-residue peptide, peptide 6, of IB-6, the arginine-containing elastase peptide of IB-1 (elastase 2), and IB-1 peptides B2 and B3.

Sequence Analysis: Automated. Automated Edman degradation of the proteins and peptides, except for native IB-6, was performed on a gas-phase protein sequenator (Model 470) from Applied Biosystems. Typically, 20-nmol samples were sequenced. All parts of the protein were sequenced at least twice on the same or overlapping peptides. The standard program provided by the company was used for sequence analysis. In the case of native IB-6, the analysis was performed on a Beckman Model 890 sequenator by the method described by Wong et al. (1979). Polybrene was used as carrier, and the buffer was 0.1 M Quadrol. The amount analyzed was 5 mg. The HPLC method of Moser and Rickli (1979) was used for identification of the PTH-amino acids.

RESULTS

Strategy

To obtain the covalent structure of protein IB-1, we chose to use the enzyme clostripain, which previously had been employed to solve the structure of the basic proline-rich protein IB-9 (Kauffman et al., 1982). Two large fragments (A and B1), a dipeptide (B3), and a tripeptide (B2) were obtained from this digestion (Figure 1). The large fragments were sequenced completely by using the sequenator. In one clostripain digestion, the fragment B1 was cleaved between residues 72 and 73, yielding two peptides, B1a and B1b. These were also sequenced by the sequenator.

Since IB-1 has a blocked NH_2-terminal residue, both the intact protein and the NH_2-terminal fragment A were treated with pyroglutamate aminopeptidase before sequencing. The positions of the di- and tripeptides, B3 and B2, within the molecule were established by elastase and carboxypeptidase B digestions of the protein (Figure 1).

Trypsin was employed for the digestion of IB-6 because larger and more easily purified fragments were obtained than with clostripain. The large fragments (2, 4, 5, 7a, 7b, 7c) were sequenced with the sequenator as far as possible (Figure 3). The five amino acid fragment 6 was sequenced manually by using the dansyl Edman procedure (Bruton & Hartley, 1970). Intact IB-6 was sequenced far enough to establish the position of fragment 6 within the molecule.

The complete amino acid sequences of IB-1 and IB-6 are shown in Figures 1 and 3, respectively, and are compared in diagrammatic form with the sequences of other salivary proline-rich proteins in Figure 4.

Table I: Amino Acid Composition of Clostripain Peptides of IB-1[a]

	peptide				
	A	B1	B2	B3	IB-1
residues	1–35	36–91	94–96	92–93	
Asx	5.0 (5)	3.3 (3)			8.2 (8)
Ser	2.7 (3)	1.1 (1)	0.9 (1)	0.9	5.5 (6)
Glx	8.0 (8)	10.0 (9)			18.3 (17)
Pro	7.2 (7)	27.5 (25)	1.2 (1)		32.2 (33)
Gly	4.8 (5)	11.7 (13)			19.8 (18)
Ala	2.1 (2)				2.0 (2)
Val	1.0 (1)				0.6 (1)
Ile	0.9 (1)				1.5 (1)
Leu	1.9 (2)				2.1 (2)
Lys	1.0 (1)	5.0 (5)			6.0 (6)
Arg			1.0 (1)	1.0	2.3 (2)
yield (%)	77	82	96	100	

[a] Residues per molecule by amino acid analysis or (in parentheses) from the sequence (Figure 1).

Protein IB-1

Purification and Characterization of the Clostripain Peptides of IB-1. Paper high-voltage electrophoresis successfully separated the clostripain peptides so that no further purification was needed. One acidic (A) and three basic peptides (B1, B2, B3) were obtained in good yield. The amino acid compositions of the peptides are shown in Table I. In one digest, peptide B1 (residues 36–91) was cleaved by clostripain into two peptides, residues 36–72 (B1b) and 73–91 (B1a).

NH_2-Terminal Analysis of IB-1. No amino-terminal residue was found with dansylation, indicating that the NH_2-terminal residue might be blocked. Digestion of the protein with pyroglutamate aminopeptidase resulted in the removal of pyrrolidonecarboxylic acid and the appearance of an NH_2-terminal aspartic acid. Amino acid analysis of the digested protein showed one less glutamic acid residue, and in addition, the pyrrolidonecarboxylic acid isolated from the column was converted to glutamic acid with acid hydrolysis. The acidic clostripain peptide A, which also had a blocked NH_2 terminal, gave identical results when digested with pyroglutamate aminopeptidase.

Table II: Amino Acid Composition of *S. aureus* Protease Peptides of Fragment IB-1A[a]

	peptide					
	1			2	3	4
residues	6–11,[b] 1–11	6–11	1–11	6–35	11–35	12–35
Asx	1.9	(1)	(3)	3.6 (3)	2.1 (2)	2.1 (2)
Ser	1.0	(1)	(1)	2.6 (3)	1.9 (2)	1.6 (2)
Glx	3.6	(3)	(5)	6.1 (6)	3.8 (4)	3.1 (3)
Pro				6.7 (7)	6.9 (7)	6.7 (7)
Gly				4.7 (5)	4.7 (5)	4.6 (5)
Ala				2.0 (2)	2.0 (2)	2.0 (2)
Val	1.1	(1)	(1)	1.0 (1)		
Ile				1.0 (1)	1.0 (1)	1.0 (1)
Leu	0.5	(0)	(1)	1.2 (1)	1.1 (1)	1.1 (1)
Lys				1.0 (1)	1.0 (1)	1.0 (1)
NH_2 terminal	Asx			Asx	Glx	Ser

[a] Residues per molecule by amino acid analysis or (in parentheses) from the sequence (Figure 1). [b] See Results; phosphoserine in IB-1.

COOH Terminal of IB-1. PMSF–carboxypeptidase B cleaved only arginine from IB-1, as revealed by paper electrophoresis at pH 6.5 and pH 1.9.

Phosphoserine in IB-1. Quantitation of the amount of phosphate present in IB-1 gave a value of 0.85 mol of P_i/mol of protein. Mild acid hydrolysis of peptide A revealed that phosphoserine was located in this peptide. Because there are three serine residues in peptide A (Table I), it was necessary to determine which serine was phosphorylated. A *S. aureus* V8 protease digest resulted in four peptides, the amino acid compositions of which are shown in Table II. Peptide 1, which had the mobility of aspartic acid on pH 6.5 paper high-voltage electrophoresis, was a mixture of two peptides. It was not further purified since both peptides contained only the Ser-8. The other peptides contained the remaining serines. P_i determinations on peptide 1 showed 0.93 mol of P_i/mol of peptide, while peptide 3 containing the other serines had no P_i, thus establishing the first serine of peptide A as the site of phosphorylation.

COOH Terminal of Peptide A. No COOH-terminal amino acid was released from peptide A when digested with carboxypeptidases B, P, and Y. Amino acid analysis (Table I) demonstrated that the peptide contained no arginine and only one lysine, which was confirmed by integration of an NMR spectrum (Bennick, unpublished experiments). The lysine residue was shown to be located at position 30 of the 35-residue peptide by automatic sequence analysis of peptide A (Figure 1). It therefore appeared that an unusual split had occurred during digestion of IB-1 with clostripain, which usually is highly specific for the carboxyl-terminal peptide bond of arginine and, to a lesser extent, lysine. In sequencing peptide A, residues 34 and 35 could not be identified; however, the sequenator analysis of intact IB-1 showed proline residues in both positions (Figure 1).

To confirm the presence of proline in position 35 (the COOH-terminal residue of peptide A), an overnight clostripain digest of peptide A was performed. A peptide (A-c) with an amino acid composition of $GlyPro_2$ was obtained. Manual dansyl Edman degradation of this peptide gave the sequence Gly-Pro-Pro. Since there is no other Gly-Pro-Pro sequence in peptide A, it must be the COOH-terminal peptide. This confirms the intact IB-1 sequence analysis and corroborates the unusual clostripain cleavage.

Complete Sequence of IB-1. Automated sequence analyses of intact IB-1 and of the large clostripain fragments (A, B1, and B1a) are shown in Figure 1, and the yields at each step are shown in Tables IV–VII, located in the supplementary material (see paragraph at end of paper regarding supplementary material). In sequencing fragment B1a containing residues 73–91, both proline and glutamine were found in position 79, suggesting that the preparation contained two species of protein. The sequence of fragment B1b, containing residues 36–72, was identical with that of B1, residues 36–72, and is therefore not shown.

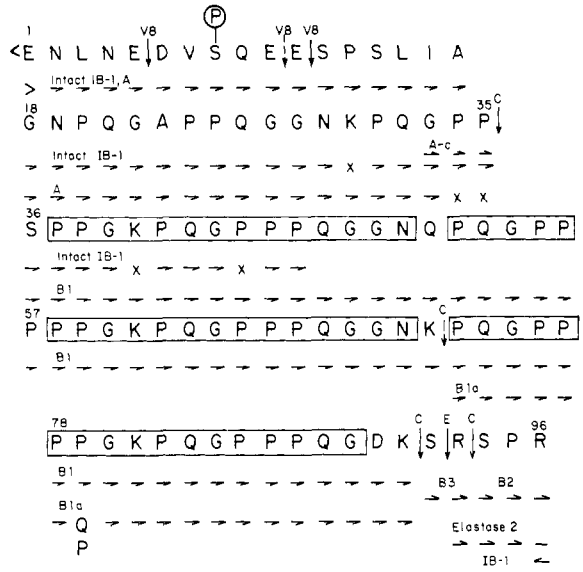

FIGURE 1: Summary proof and complete amino acid sequence of salivary proline-rich protein IB-1: (→) automated Edman degradation; (⟶) manual dansylation and Edman degradation; (⟶) carboxypeptidase B digestion; (<) pyroglutamate aminopeptidase digestion; (×) residue not identified. Sites of cleavage are indicated by vertical arrows; a broken vertical arrow symbolizes partial cleavage: V8, *S. aureus* V8 protease; C, clostripain; E, elastase. Boxed areas show repetitive identical sequences.

Residues 36–96 of the protein are identical with the total primary sequence of IB-9 (Kauffman et al., 1982), except for residue 72, where lysine is found instead of the arginine found in IB-9. The sequence of the first 10 residues of peptide A (and IB-1, see, Figure 1) is identical with that of the first 10 residues of acidic proline-rich protein C (Wong & Bennick, 1980) with the exception of residue 2. Ser-8 is phosphorylated in both proteins.

The di- and tripeptides, B3 and B2, as well as the arginine-containing elastase peptide 2 (residues 93–96), were sequenced manually by the dansyl Edman procedure (Bruton & Hartley, 1970). The elastase peptide (Arg-Ser-Pro-Arg) established the order of the B2 and B3 peptides as Ser-Arg-Ser-Pro-Arg. Since arginine is the COOH-terminal residue of IB-1, and B2 and B3 contain all of the arginine present in the molecule, the peptides must be located at the COOH terminal as shown in Figure 1.

Protein IB-6

Purification and Characterization of IB-6 Tryptic Peptides. SP-Sephadex C25 chromatography of the tryptic digest of IB-6 resulted in the elution profile shown in Figure 2. Peak 1 contains trypsin, which was determined by chromatographing trypsin on an identical column. Peak 3 is a minor peptide derived from the peak 4 peptide. Peak 7, when purified further, as described under Materials and Methods, resulted in the separation of three peptides, 7a, 7b, and 7c. The amino acid compositions and yields of the peptides are shown in Table III.

COOH-Terminal Analysis of IB-6. No enzymatic cleavage occurred when IB-6 was digested with carboxypeptidase B–DFP. With carboxypeptidase Y digestion, 1 mol of glutamine

Table III: Amino Acid Composition of Trypsin Peptides of IB-6[a]

	peptide							
	2	4	5	6	7a	7b	7c	IB-6
residues	63–112	63–118	1–~53	58–62	1–18	1–53	19–57	1–118
Asx	3.1 (3)	2.7 (3)	2.9 (2)		1.0 (1)	2.1 (2)	1.7 (1)	5.5 (5)
Ser	1.1 (1)	1.7 (2)	2.1 (1)	1.7 (2)	1.0 (1)	2.1 (2)	1.8 (2)	5.8 (6)
Glx	10.2 (11)	10.4 (12)	6.7 (9)	0.8 (1)	3.9 (4)	9.2 (9)	5.0 (5)	22.0 (22)
Pro	19.0 (19)	20.1 (22)	18.6 (24)		6.0 (7)	23.8 (24)	17.0 (17)	37.0 (46)
Gly	10.1 (10)	9.6 (10)	11.9 (12)		3.2 (4)	13.1 (14)	7.9 (10)	24.6 (24)
Ala	3.3 (3)	2.7 (3)	(1)	1.1 (1)		1.3 (1)	1.5 (1)	4.7 (5)
Lys	1.0 (1)	1.0 (1)	4.0 (4)		0.9 (1)	4.7 (5)	3.5 (4)	5.7 (6)
Arg	2.7 (2)	2.8 (3)		1.0 (1)				3.9 (4)
Yield (%)	25	29	19	16	10	25	9	

[a] Residues per molecule by amino acid analysis or (in parentheses) from the sequence (Figure 3).

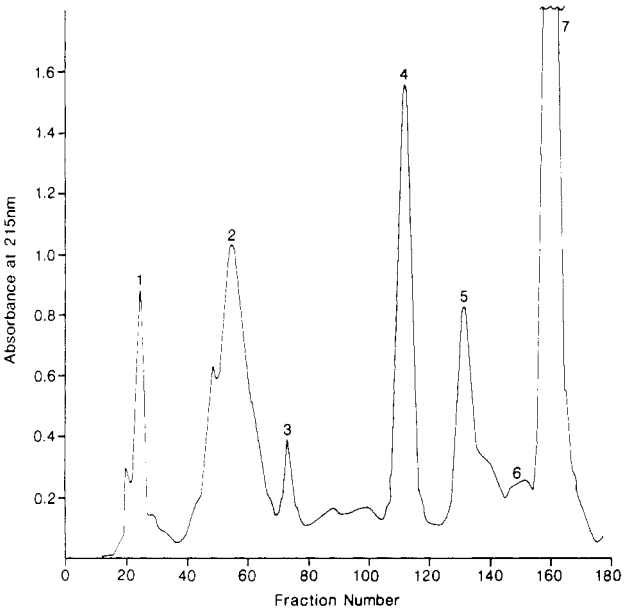

FIGURE 2: Elution profiles of IB-6 tryptic peptides from a SP-Sephadex C25 column (1.5 × 103 cm) equilibrated with 0.025 M sodium phosphate, pH 6.75. A gradient of 500 mL of buffer and 500 mL of buffer containing 0.3 M NaCl was initiated after sample application. The flow rate was 13.6 mL/h; the fraction volume was 3.4 mL.

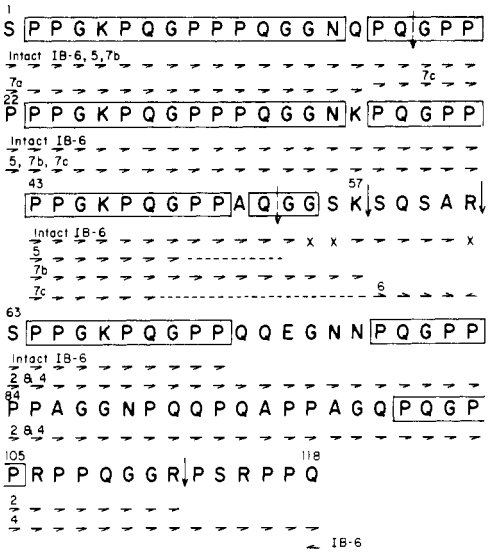

FIGURE 3: Summary proof and complete amino acid sequence of salivary proline-rich protein IB-6: (—) automated Edman degradation; (—) manual dansylation and Edman degradation; (—) carboxypeptidase Y or B digestion; (×) residue not identified; (---) region inferred from amino acid composition. Sites of tryptic cleavage are indicated by vertical arrows; a broken vertical arrow symbolizes partial cleavage. Boxed areas show repetitive identical sequences.

was removed from 1 mol of IB-6, thus establishing glutamine as the COOH-terminal residue of IB-6.

Complete Sequence of IB-6. The complete covalent sequence of IB-6, obtained from the peptide sequences and a partial sequence of the intact protein, is shown in Figure 3. Automated sequence analysis was performed on IB-6 and on peptides 2, 4, 5, 7a, 7b, and 7c. Peptide 2 is derived from peptide 4, and peptides 5, 7a, and 7c are derived from the larger peptide 7b. Peptide 6, a five-residue peptide, was sequenced manually and gave the sequence Ser-Gln-Ser-Ala-Arg. The yield of this peptide purified on SP-Sephadex was low (16%). However, when the tryptic digest was purified by paper electrophoresis, the yield was approximately 62%, thus excluding the possibility that this peptide was a contaminant. The positioning of the five-residue peptide, peptide 6, was established by the partial sequence analysis of intact IB-6. The yields from sequenator runs are given in Tables VIII–X in the supplementary material. We reported previously that residue 52 was proline (Kauffman et al., 1982). When alanine was found in this position with a new preparation of IB-6, the results of the previous analysis were reexamined. It was found that the HPLC chromatogram for residue 52 had been misread and that indeed it was alanine in that position. The sequence of the first 54 residues of Ib-6 is identical with that of residues 36–89 of IB-1, except for an alanine in position 52, where IB-1 has proline.

Discussion

Acidic and basic proline-rich proteins and glycoproteins are widespread in salivary glands and their secretions. In addition to the human proteins and glycoproteins, they have been isolated and characterized to varying degrees from *Macaca irus* (Arneberg, 1974, 1976), *Macaca fascicularis* (Oppenheim et al., 1982, 1985), rat (Keller et al., 1975; Muenzer et al., 1979a,b; Iversen et al., 1982), and rabbit (Rajan & Bennick, 1983). The proline-rich secretory products of these various species vary somewhat in number, size, and chemical composition but have in common their unusual covalent polypeptide structure in which two or three amino acids (glycine, proline, glutamic acid, or glutamine) predominate.

This discussion will attempt to relate the findings reported in this paper to other studies and will address the following topics: (1) the structural relationship of the basic proline-rich proteins IB-1, IB-6, and IB-9 to each other and to similar proteins studied by Sanada and associates, and (2) the structural relationships of the nonglycosylated basic proline-rich proteins of human parotid saliva to the acidic proline-rich proteins and the proline-rich glycoproteins.

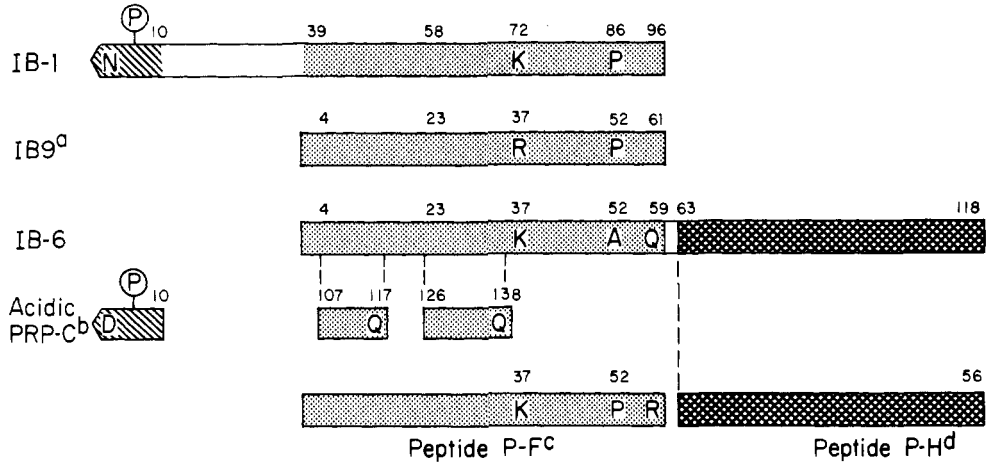

FIGURE 4: Relationships between salivary basic proline-rich proteins IB-1, -6, and -9 and between these and other salivary proline-rich proteins. Identical shadings signify identical sequences, except where indicated by amino acid symbols. References cited: (a) Kauffman et al. (1982); (b) Wong and Bennick (1980); (c) Saitoh et al. (1983a); (d) Saitoh et al. (1983b).

Proteins IB-1, IB-6, and IB-9 are three of the nine nonglycosylated basic proline-rich proteins isolated from the parotid saliva of a single person (Kauffman & Keller, 1979; Kauffman et al., 1982). They differ in their chemical composition, electrophoretic mobility, and molecular size, with IB-6 > IB-1 > IB-9. Nonetheless, these proteins bear close structural relationships to each other, as shown in Figure 4. With the exception of a single substitution (Lys-72 of IB-1 for Arg-37 of IB-9) the entire 61 amino acid sequence of IB-9 is reproduced in IB-1 as residues 36–96. With the same exception (Lys for Arg in position 37) the first 51 residues of IB-9 are reproduced also in IB-6.

Proteins IB-1 and IB-6 differ significantly from each other in the remainder of their covalent structures. In IB-1 the region homologous with IB-9 is preceded by an NH_2-terminal extension comprising 35 amino acid residues, 10 of which are homologous with acidic proline-rich protein C (Wong & Bennick, 1980). Included in this decapeptide is a phosphoserine residue in position 8 of both proteins. This represents the first example of a *basic* proline-rich phosphoprotein in human parotid saliva. Protein IB-6 also contains a molecular extension which, in contrast to IB-1, is located on the COOH-terminal side of the sequence that is homologous to IB-9. The data show that, despite extensive regions of identity, none of these three basic proteins could have derived from the other nor from a common precursor protein unless the precursor contains separate segments for each of these proteins.

Sanada and his coinvestigators have isolated seven basic proline-rich peptides from human parotid saliva (P-C through P-I) and have sequenced five of these. In comparing these peptide sequences to those of other proline-rich proteins, the following relationships were noted: peptide P-C (Isemura et al., 1980) is identical with residues 107–150 in the acidic proline-rich protein C (Wong & Bennick, 1980); peptide P-E (Isemura et al., 1982) is identical with the basic proline-rich protein IB-9. As shown in Figure 4, peptide P-F (Saitoh et al., 1983a) is homologous with residues 1–61 of basic proline-rich protein IB-6, and peptide P-H (Saitoh et al., 1983b) is identical with residues 63–118 of protein IB-6.

Human parotid and human submandibular saliva contain four major acidic proline-rich proteins, which have been extensively characterized with respect to their structure and function [see review by Bennick (1982)]. The four major acidic proline-rich proteins are closely related structurally, differing only in the nature of the residue in position 4 (Asp or Asn) and in the presence or absence of a 44-residue extension at the carboxyl terminal of the molecule. The covalent structures of the acidic proline-rich proteins reveal a structural polarity, with most of the negatively charged residues located near the amino terminal of the molecule and most of the proline, glycine, and glutamine located in the carboxyl terminal. As was noted in an earlier communication (Kauffman et al., 1982), the basic proline-rich protein IB-9 exhibits a high degree of homology with the carboxyl-terminal region of acidic proline-rich protein C, including sequences of 11 amino acids that occur repeatedly in both proteins.

Azen et al. (1984) and Maeda et al. (1985) have reported nucleotide sequences of cDNAs coding for human salivary proline-rich proteins. Both groups of investigators noted the presence of tandem, repetitive regions made up of identical or nearly identical nucleotide sequences. The finding of repetitive nucleotide sequences is consistent with amino acid sequence analyses reported in the present paper and other publications dealing with human salivary proline-rich proteins (Kauffman et al., 1982; Bennick, 1982).

Supplementary Material Available

Sequence analyses of IB-1 and IB-1 peptides A, B1, and B1a (Tables IV–VII) and of IB-6 and IB-6 fragments 7b and 4 (Tables VIII–X) (7 pages). Ordering information is given on any current masthead page.

Registry No. IB-1, 100993-19-9; IB-6, 100993-20-2.

References

Arneberg, P. (1974) Thesis, University of Oslo, Norway.

Arneberg, P., Helgeland, K., & Tjornhom, T. (1976) *Arch. Oral Biol. 21*, 379–383.

Azen, E., Lyons, K. M., McGonigal, T., Barrett, N. L., Clements, L. S., Maeda, N., Vanin, E. F., Carlson, D. M., & Smithies, O. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5561–5565.

Belford, H. S., Triffleman, E. G., Offner, G. D., Troxler, R. F., & Oppenheim, F. G. (1984) *J. Biol. Chem. 259*, 3977–3984.

Bennick, A. (1982) *Mol. Cell. Biochem. 45*, 83–99.

Bruton, C. J., & Hartley, B. S. (1970) *J. Mol. Biol. 52*, 165–178.

Gray, W. R. (1967) *Methods Enzymol. 11*, 139–151.

Isemura, S., Saitoh, E., & Sanada, K. (1980) *J. Biochem. (Tokyo) 87*, 1071–1077.

Isemura, S., Saitoh, E., & Sanada, K. (1982) *J. Biochem. (Tokyo) 91*, 2067–2075.

Itaya, K., & Ui, M. (1966) *Clin. Chim. Acta 14*, 361–366.
Iversen, J. M., Johnson, D. A., Kauffman, D. L., Keller, P. J., & Robinovitch, M. R. (1982) *Arch. Oral Biol. 27*, 925–930.
Kauffman, D. L., & Keller, P. J. (1979) *Arch. Oral Biol. 24*, 249–256.
Kauffman, D. L., & Keller, P. J. (1983) *Arch. Oral Biol. 28*, 61–67.
Kauffman, D., Wong, R., Bennick, A., & Keller, P. (1982) *Biochemistry 21*, 6558–6562.
Keller, P. J., Robinovitch, M., Iversen, J., & Kauffman, D. L. (1975) *Biochim. Biophys. Acta 379*, 562–570.
Levine, M., & Keller, P. J. (1977) *Arch. Oral Biol. 22*, 37–41.
Levine, M. J., Weill, J. C., & Ellison, S. A. (1969) *Biochim. Biophys. Acta 188*, 165–167.
Maeda, N., Kim, H. S., Azen, E. A., & Smithies, O. (1985) *J. Biol. Chem. 260*, 11123–11130.
Martinsen, T. M. (1984) *Methods Enzymol. 107*, 3–22.
Moser, P. W., & Rickli, E. E. (1979) *J. Chromatogr. 176*, 451–455.
Muenzer, J., Bildstein, C., Gleason, M., & Carlson, D. M. (1979a) *J. Biol. Chem. 254*, 5623–5628.
Muenzer, J., Bildstein, C., Gleason, M., & Carlson, D. M. (1979b) *J. Biol. Chem. 254*, 5629–5634.
Oppenheim, F. G., Offner, G. D., & Troxler, R. F. (1982) *J. Biol. Chem. 257*, 9271–9282.
Oppenheim, F. G., Offner, G. D., & Troxler, R. F. (1985) *J. Biol. Chem. 260*, 10671–10679.
Rajan, A. I., & Bennick, A. (1983) *Arch. Oral Biol. 28*, 431–439.
Saitoh, E., Isemura, S., & Sanada, K. (1983a) *J. Biochem. (Tokyo) 93*, 883–888.
Saitoh, E., Isemura, S., & Sanada, K. (1983b) *J. Biochem. (Tokyo) 94*, 1991–1999.
Schlesinger, D. H., & Hay, D. I. (1979) *Pept., Proc. Am. Pept. Symp. 6th, 1979*, 133.
Troxler, R. F., Belford, H. S., & Oppenheim, F. G. (1983) *Biochem. Biophys. Res. Commun. 111*, 239–246.
Wong, R. S. C., & Bennick, A. (1980) *J. Biol. Chem. 255*, 5943–5948.
Wong, R. S. C., Hofmann, T., & Bennick, A. (1979) *J. Biol. Chem. 254*, 4800–4808.

Histidinoalanine, a Naturally Occurring Cross-Link Derived from Phosphoserine and Histidine Residues in Mineral-Binding Phosphoproteins†

Mary E. Marsh
The University of Texas Health Science Center, The Dental Science Institute, Houston, Texas 77225
Received October 22, 1985; Revised Manuscript Received December 30, 1985

ABSTRACT: Native mineral-containing phosphoprotein particles were isolated from the Heterodont bivalve *Macrocallista nimbosa*. The native particles are discrete structures about 40 nm in diameter which migrate as a single band during electrophoresis in agarose gels. Removal of the mineral component with ethylenediaminetetraacetic acid dissociates the native protein into nonidentical subunits. The lower molecular weight subunits, representing 8% of the total protein, were obtained by differential centrifugation. The native protein is characterized by a high content of aspartic acid, phosphoserine, phosphothreonine, histidine, and the bifunctional cross-linking residue histidinoalanine. The low molecular weight subunits have the same amino acid composition except for a reduction in histidinoalanine and a corresponding increase in phosphoserine and histidine residues, demonstrating that the alanine portion of the cross-link is derived from phosphoserine residues. Ion-exchange chromatography and molecular sieve chromatography show that the low molecular weight subunits have a similar charge density but differ in molecular weight, and the relative mobilities of the subunits on agarose gels indicate that they are polymers of a single phosphoprotein molecule. The minimum molecular weight of the monomer is about 140000 on the basis of the amino acid composition. The high molecular weight subunits are rich in histidinoalanine and too large to be resolved by either molecular sieve chromatography or gel electrophoresis. On the basis of the ultrastructural, electrophoretic, chromatographic, and compositional evidence, native phosphoprotein particles are composed of subunits ionically cross-linked via divalent cations. These subunits are variable molecular weight aggregates of a single phosphoprotein molecule covalently cross-linked via histidinoalanine residues. Evidence for a nonenzymatic cross-linking mechanism is discussed.

Histidinoalanine or N^{τ}-(2-amino-2-carboxyethyl)histidine is a bifunctional cross-linking amino acid. It was first identified in dentin and bone (Fujimoto et al., 1982) and then in other vertebrate connective tissues including tendon, cartilage, and aorta (Fujimoto, 1982a; Fujimoto & Yu, 1984). In dentin, histidinoalanine occurs exclusively in free phosphophoryn and an insoluble phosphophoryn–protein conjugate (Kuboki et al., 1984). Phosphophoryns are high-capacity calcium-binding proteins in which about 80% of the amino acid residues are phosphoserine and aspartic acid (Lee et al., 1977). The conjugated form of phosphophoryn is probably complexed to collagen (Dimuzio & Veis, 1978).

Histidinoalanine also occurs in another group of high-capacity calcium-binding proteins called phosphoprotein particles (Sass & Marsh, 1983; Marsh & Sass, 1985). The phosphoprotein particles sequester calcium, magnesium, and inorganic phosphate ions and exist in a stable colloidal suspension in the blood and extrapallial fluid of Heterodont bivalves (Marsh & Sass, 1983, 1984, 1985). The particles are rich in phosphoserine, aspartic acid, and histidine residues and have a histi-

† Supported by National Institutes of Health Grant AM-36239.

0006-2960/86/0425-2392$01.50/0