Basic Proline-Rich Proteins from Human Parotid Saliva: Complete Covalent Structure of Protein IB-9 and Partial Structure of Protein IB-6, Members of a Polymorphic Pair†

Dorothy Kauffman, Raymond Wong, Anders Bennick, and Patricia Keller*

ABSTRACT: The complete amino acid sequence of a basic proline-rich protein, IB-9, from human parotid saliva was determined by automated Edman degradation of peptides obtained by enzymatic cleavage of the intact protein with clostripain. The protein was digested with papain and elastase to obtain overlapping peptides. Automated Edman degradation of the intact protein was also performed. The protein consists of 61 amino acids, of which 26 are proline residues. The partial sequence of another human parotid basic proline-rich protein, IB-6, was also obtained. With one exception the first 54 residues of the two proteins are identical. An exceptional degree of internal reiteration occurs in both molecules, including several repeated sequences of 12–14 amino acids. The proteins show a high degree of homology with the C-terminal portion of the salivary acidic proline-rich protein C.

In recent years we have reported the purification and partial characterization of nine proline-rich proteins of human parotid saliva (Kauffman & Keller, 1979; Levine & Keller, 1977). The complete sequence of the most basic of these proteins, IB-9, has been published in abstract form (Kauffman et al., 1982). Work in our laboratory has identified protein IB-9 and another basic proline-rich protein, IB-6, as the proteins involved in the parotid middle-band polymorphism (Anderson et al., 1982a,b). Recently, Isemura et al. (1982) reported the sequence of a salivary basic peptide, P-E. The purpose of this paper is to document the sequence of IB-9 and its identity with P-E and to report the partial sequence of the polymorphic protein IB-6.

Materials and Methods

Clostripain and trypsin (TPCK[1] treated) were obtained from Worthington, papain suspension 2 times crystallized was from Sigma, and elastase suspension 2 times crystallized was from Mann Research Laboratories. Dansyl chloride, dansyl amino acids, PITC, and F_3CCOOH were obtained from Pierce and Cheng Chin polyamide sheets from Gallard-Schlesinger Chemical Corp.

Isolation of Salivary Protein IB-9. IB-9 was prepared from human parotid saliva of a single individual as previously described (Kauffman & Keller, 1979).

Clostripain Digestion of IB-9. Clostripain digestion of 10 mg of IB-9 was carried out at 37 °C for 2 h in 2 mL of 0.05 M ammonium hydrogen carbonate, pH 8, containing 0.27 M mercaptoethanol. The enzyme:substrate ratio was 1:50 on a weight basis.

Purification of Clostripain Peptides. The digest solution was spread along a 10-cm line on Whatman No. 1 paper and subjected to high-voltage electrophoresis at pH 6.5, 4 kV for 45 min. A guide strip was stained with cadmium ninhydrin to locate the peptide bands. The bands were cut out, and the peptides were eluted with 0.05 M acetic acid.

Papain Digestion of Clostripain Peptides 2 and 3. Clostripain peptides 2 and 3 were digested with papain for 2 h at 37 °C in 0.05 M pyridine–acetate, pH 6.5, containing 0.27 M mercaptoethanol. The enzyme:substrate ratio was 1:100 on a molar basis. The resulting peptides were purified by pH 6.5 high-voltage paper electrophoresis for 60 min at 3 kV. The area containing neutral peptides was stitched to a second Whatman No. 1 paper and subjected to high-voltage electrophoresis at pH 1.9, for 30 min at 3 kV. Peptides were eluted from the paper with 0.05 M acetic acid.

Elastase Digestion of IB-9. IB-9 (0.5 mg) was digested with elastase in 0.1 mL of *N*-ethylmorpholine, pH 8.5, at 37 °C for 16 h. The enzyme:substrate ratio was 1:100 on a weight basis. The resulting peptides were purified by high-voltage paper electrophoresis at pH 6.5, 30 min at 4 kV, and eluted from the paper with 0.05 M acetic acid.

Purification of IB-6. IB-6 was isolated by the procedure of Kauffman & Keller (1979) from the same individual from which IB-9 had been obtained and was further purified by gel filtration on Bio-Gel P-30 in 0.05 M ammonium hydrogen carbonate. The column size was 1.6 × 91 cm and the flow rate 15 mL/h.

Clostripain Digestion of IB-6. A 1-mg sample of IB-6 was digested with clostripain as described above. The peptides were purified by high-voltage electrophoresis at pH 6.5; the neutral area was further purified by high-voltage electrophoresis at pH 1.9. Peptides were eluted with 0.05 M acetic acid.

Amino Acid Analysis. Peptides were hydrolyzed in 6 N HCl at 108 °C in sealed, evacuated tubes for 16–24 h, and amino acid analyses of the hydrolysates were performed on a Beckman Model 121 analyzer.

End Group Analysis. N-Terminal amino acids of the purified peptides were determined by the dansyl procedure for peptides (Gray, 1967).

Sequence Analysis: Manual. Manual dansylation and Edman degradation was performed by the micromethod of Bruton & Hartley (1970).

Sequence Analysis: Automated. Automated Edman degradation of peptides was performed on a Beckman Model 890 sequenator by the method described by Wong et al. (1979). For sequencing of native proteins IB-6 and IB-9, Polybrene

† From the Department of Oral Biology, University of Washington, Seattle, Washington 98195, and the Department of Biochemistry, University of Toronto, Toronto, Canada M5S 1A8. *Received September 20, 1982.* This work was supported by National Institutes of Health Grants DE-02600 and DE-05251 and Canadian MRC Grant MT 4920.

[1] Abbreviations: TPCK, L-1-(tosylamido)-2-phenylethyl chloromethyl ketone; PITC, phenyl isothiocyanate; HPLC, high-performance liquid chromatography; PTH, phenylthiohydantoin; CD, circular dichroism.

0006-2960/82/0421-6558$01.25/0

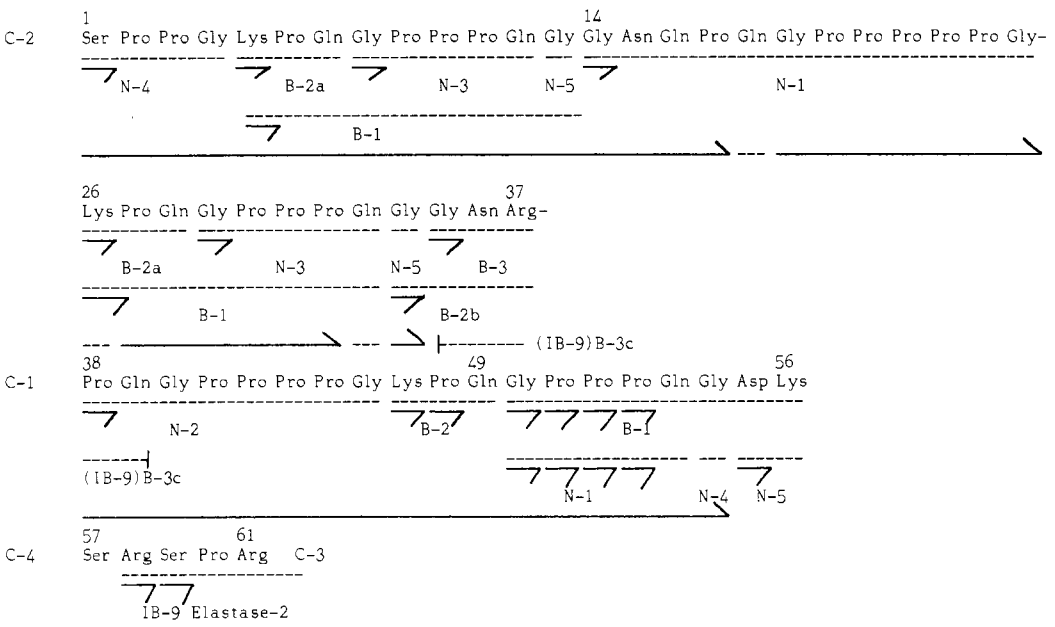

FIGURE 1: Summary of peptides used to establish complete amino acid sequence of salivary protein IB-9: (⇀) automated Edman degradation; (→) manual dansylation and Edman degradation; (- - -) regions inferred from amino acid composition.

Table I: Amino Acid Compositions and N-Terminal Analysis of Clostripain Peptides of IB-9

	residues per mol				
amino acid	C-1	C-2	C-3	C-4	IB-9
Asx	1.2	2.0			3.1
Ser		0.8	1.0	1.0	2.6
Glx	2.8	6.0			9.4
Pro	8.5	16.6	1.0		25.9
Gly	3.8	9.1			13.0
Lys	1.7	2.3			4.0
Arg		0.8	1.0	1.0	2.9
yield (%)	100	84	100	50	
N terminal	Pro	Ser	Ser	Ser	Ser

was used as a carrier, and the buffer was 0.1 M Quadrol. Before application of the samples to the sequenator cup, three preliminary cycles with Polybrene and glycylglycine were performed. Instead of gas–liquid chromatography, the HPLC method of Moser & Rickli (1979) was used for identification of the PTH-amino acids.

Immunological Reactivity. The reactivity of IB-9 was tested by means of immunoradiometric assay for native human acidic proline-rich proteins (Bennick & Madapallimattan, 1982). The antibody used was specific antibody to the acidic proline-rich protein C. Protein C was used as the standard.

CD Spectroscopy. CD spectra were recorded as described by Wong & Bennick (1980) of clostripain peptide C-1 and C-2 (0.1 mg/mL) obtained from a digest of protein IB-9.

Results

Purification and Characterization of Clostripain Peptides from IB-9. Four peptides were obtained from clostripain digestion of IB-9. These were readily purified by paper high-voltage electrophoresis at pH 6.5. The peptides were characterized by amino acid analysis and amino-terminal analyses, as shown in Table I. The sum of the amino acids of clostripain peptides C-1 through C-4 agrees well with the amino acid composition of intact IB-9. In a previous publication (Kauffman & Keller, 1979) we reported on the isolation of five peptides after clostripain digestion of IB-9. However, one of the five, a neutral clostripain peptide always obtained in low yield, was subsequently identified as a noncovalently bound impurity. Sequence analyses were performed on IB-9 that had been freed from this component.

Papain Digestions of Peptides C-1 and C-2 and of IB-9. A summary of the compositions of the peptides obtained by papain digestion of peptides C-1 and C-2 is shown in Table II. The sum of these peptides is in close agreement with the amino acid compositions of C-1 and C-2. Yields of greater than 100% for peptides N3 and B2a from clostripain fragment C-2 indicate that these peptides occur twice within this fragment (Figure 1). Also shown in Table II is the composition of the overlapping peptide between C-1 and C-2, which was obtained by papain digestion of intact IB-9. Figure 1 shows the placement of the peptides as determined by the automated

Table II: Amino Acid Compositions of Papain Peptides[a]

	C-1						C-2								B3c of
	N1	N2	N4	N5	B1	B2	N1	N3	N4	N5	B1	B2a	B2b	B3	IB-9
Asx				1.0	1.0		1.0						1.0	0.9	0.9
Ser									0.9						
Glx	0.9	0.9			1.2	1.0	2.1	0.8			2.0	1.0			1.9
Pro	3.1	5.0			3.0	0.8	5.4	3.0	2.2		4.0	1.2			0.9
Gly	1.4	2.1	1.0		2.1		3.2	1.0	1.0	1.0	1.5		1.9	1.0	1.3
Lys				0.9	1.0	1.0					1.2	1.0			
Arg													1.0	0.9	0.9
yield (%)	68	48	7.5	9	64	80	46	160	89	23	37	200	84	15	25

[a] N indicates a neutral peptide and B a basic peptide at pH 6.5.

Table III: Sequence Analysis of IB-9[a]

protein amount																					repetitive yield (%)	no. of runs
6 mg of IB-9	PTH-aa	Ser	Pro	Pro	Gly	Lys	Pro	Gln	Gly	Pro	Pro	Gln	Gly	Gly	Asn	Gln	Pro	Gln	Gly	Pro	Pro	96	3[c]
	yield (nmol)	(249)[b]	516	548	403	266	493	503	332	381	362	477	399	333	385	343	350	314	306	210	275	347	
													Pro 113										
	PTH-aa	Pro	Pro	Pro	Gly	Lys	Pro	Gln	Gly	Pro	Pro	Pro	Gln	Gly	Gly	Asn	Arg[d]	Pro	Gln	Gly	Pro	Pro	
	yield (nmol)	344	345	355	189	192	260	219	173	213	322	310	171	128	178	122		124	132	103	142	178	
					Pro 112								Pro 128				Asn 56						
	PTh-aa	Pro	Pro	Gly	Lys	Pro	Gln	Gly	Pro	Pro	Pro	Gln[e]	Gly	Asp	Lys	Ser	Arg[d]	Ser	Pro	(Arg)			
	yield (nmol)	176	177	93	93	121	67	74	83	155	163	66	61	63	49	25		28	28				
				Pro 84								Pro 85											

[a] The first line of peptide yield indicates the amount recovered of the amino acid (aa) identified in that cycle. The second line gives the yield and the amino acid identified in the previous cycle. [b] Yield was not corrected for destruction of this residue. [c] Only the results of the last run were presented. PTH-amino acids were identified by HPLC. [d] Residue identified by back-hydrolysis. [e] The identification of glutamine is based on the absence of glutamine in the previous cycle and the decrease in proline from the preceding cycle.

Table IV: Amino Acid Composition of IB-6

	residues per 100		residues per 100
Asx	5.4	Gly	19.2
Ser	5.2	Ala	4.9
Glx	19.4	Lys	5.6
Pro	37.0	Arg	3.3

sequence analyses of C-1 and C-2. These peptides clarified areas of sequence that were in doubt in the automated sequencing of C-1 and C-2.

Elastase Digestion of IB-9. Only one peptide bond was hydrolyzed during the elastase digestion of IB-9, giving a large core (E-1) and a small peptide (E-2) consisting of two arginine residues, one serine residue, and one proline residue. The sequence of this peptide, obtained by manual dansylation Edman degradation is Arg-Ser-Pro-Arg (Figure 1).

Automated Sequence Analysis. The automated sequence analyses performed on peptides C-1 and C-2 are shown in Figure 1. The sequence analysis performed on the intact protein is shown in Table III. The protein has 61 amino acids, with a molecular weight of 5988. When the residues are aligned as in Figure 2A, it can be seen that a sequence of 12 amino acids is repeated 3 times in the molecule (residues 2–13, 23–34, 43–54). This is extended to a 14 amino acid identity in the case of two of the three repeating units. In addition, after a single residue "break" the identity continues for another five residues. The carboxyl-terminal region of the molecule is rich in basic residues; of the six final amino acids, three are positively charged.

Purification and Partial Sequence of IB-6. IB-6 eluted from the gel filtration column ahead of the contaminating protein, IB-7, and gave a single band on cationic gel electrophoresis. The amino acid composition is shown in Table IV. Clostripain digestion resulted in a more complex pattern of peptides than was seen in clostripain digests of IB-9. One large fragment obtained from IB-6 was identical in composition with IB-9 C-2, except that a lysine had replaced the arginine residue. The ability of clostripain to cleave this lysine bond in IB-6 (and also lysine-56 in IB-9) indicates that the specificity of clostripain is not restricted to arginine bonds. This secondary specificity has been noted previously (Mitchell & Harrington, 1970).

Automated Edman degradation of residues 1–54 in native IB-6 (Table V) confirms that the C-2 peptide is identical with C-2 of IB-9, except that in position 37 a lysine has replaced the arginine. For confirmation of this difference, the aqueous phases of cycles 36, 37, and 38 from IB-6 were subjected to back-hydrolysis and amino acid analysis. Similar analysis was done of the organic phase. In no instance could arginine be detected. The sequence continues its identity to IB-9 through residue 54.

Immunological Reactivity. Immunoradiometric assays of SP-9 demonstrated reactivity within the range of the assay. The intensity of reaction of IB-9 was weak and varied between 5×10^{-6} and 37×10^{-4} of that obtained with the same weight amount of acidic proline-rich protein C.

CD Spectroscopy. The spectra obtained of clostripain peptides C-1 and C-2 from protein IB-9 were very similar to the spectrum of the C-terminal CTZ peptide obtained from the salivary acidic proline-rich protein C (Wong & Bennick, 1980), further attesting to the structural similarity of these proteins. In the spectral region from 190 to 250 nm there was no positive ellipticity but a single intense minimum at 202 nm (peptide C-1) or 201 nm (peptide C-2). The mean residue

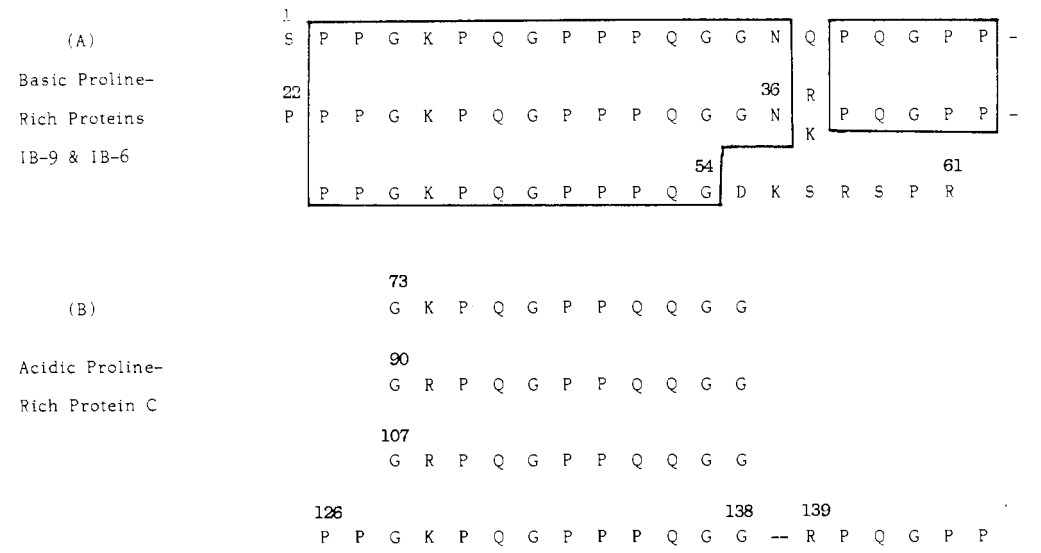

FIGURE 2: (A) Complete amino sequence of salivary basic proline-rich protein IB-9 and partial sequence of IB-6 through residue 54 by automated Edman degradation. At residue 37 the raised amino acid is in IB-9; the lower one in IB-6. They are otherwise identical through residue 54. The sequence has been arranged to emphasize the degree of reiteration present, and the boxed areas indicate the areas of internal repetition in both IB-9 and IB-6. (B) sequences present in salivary protein C (Wong & Bennick, 1980) that show partial identity with those in IB-9 and IB-6.

Table V: Sequence Analysis of IB-6

protein amount													repetitive yield (%)	no. of runs
5 mg of IB-6	PTH-aa	Ser	Pro	Pro	Gly	Lys	Pro	Gln	Gly	Pro	Pro	Pro	96.7	1
	yield (nmol)	nd[a]	204	208	169	196	221	130	151	189	171	171		
	PTH-aa	Gln	Gly	Gly	Asn	Gln	Pro	Gln	Gly	Pro	Pro	Pro		
	yield (nmol)	85	97	118	35	151	109	79	64	111	128	130		
		Pro 62												
	PTH-aa	Pro	Pro	Gly	Lys	Pro	Gln	Gly	Pro	Pro	Pro	Gln		
	yield (nmol)	131	143	66	56	91	56	54	63	81	85	50		
				Pro 65								Pro 62		
	PTH-aa	Gly	Gly	Asn	Lys	Pro	Gln	Gly	Pro	Pro	Pro	Pro		
	yield (nmol)	44	52	28	35	49	35	32	73	80	101	100		
	PTH-aa	Gly	Lys	Pro	Gln	Gly	Pro	Pro	Pro	Gln	Gly	4 Asx		
	yield (nmol)	39	22	78	31	32	71	84	63	21	32			
		Pro 74												
		4 Ser,	9 Glx,	9 Pro,	4 Gly,	4 Ala,	1 Lys,	3 Arg[b]						

[a] Recovery of residue not determined. [b] The number of residues not identified is based on the assumption that the protein contains three arginines.

ellipticities at the minima were -17×10^3 deg·cm^2·dmol^{-1} (peptide C-1) and -29×10^3 deg·cm^2·dmol^{-1} (peptide C-2).

Discussion

Salivary glands of humans (Oppenheim et al., 1971; Bennick & Connell, 1971; Levine & Keller, 1977; Levine et al., 1969), subhuman primates (Arneberg, 1974; Oppenheim et al., 1979), rats (Keller et al., 1975; Muenzer et al., 1979), and rabbits (Castle & Palade, 1978) have been reported to synthesize and secrete unique groups of proteins in which the imino acid, proline, constitutes between 24 and 42 mol % of the total residues. More than twenty such proteins have been isolated from human parotid saliva during the past decade that can be categorized according to their charge characteristics and whether or not they are glycosylated. The covalent structures of several human acidic proline-rich proteins have been reported (Wong et al., 1979; Wong & Bennick, 1980; Schlesinger & Hay, 1979, 1981). This paper provides experimental documentation to our previous brief report on the covalent structure of one of the basic proline-rich proteins, IB-9, and establishes its identity with peptide P-E (Isemura et al., 1982). The partial sequence of protein IB-6 is also reported.

IB-9 and IB-6, each comprising approximately 10% of the basic proline-rich proteins (unpublished data), correspond respectively to two polymorphic proteins, Pm and PmS, found in human parotid saliva (Anderson et al., 1982a). On the basis of family studies, Azen & Denniston (1980) suggested that Pm and PmS were products of a single gene locus having allelic frequencies of 0.15 for Pm (IB-9) and 0.12 for PmS (IB-6). In agreement with the estimated gene frequencies, comparison of parotid salivas has demonstrated that approximately 17% of the population secretes both IB-9 and IB-6, while 7% secretes only IB-9, and the rest have neither protein in their saliva (Anderson et al., 1982b). Comparison of the partial sequence of IB-6 with that of IB-9 (Figure 2A) shows that only one amino acid substitution occurs within the first 54 residues; namely, an arginine at position 37 in IB-9 is replaced by a lysine in IB-6. This 54-residue domain of identity constitutes approximately 90% of protein IB-9 but only 59% of protein IB-6. Thus, a major difference between these proteins is that

IB-6 contains an extension of approximately 30 amino acid residues located at the carboxyl end of the molecule. However, the observed substitution of Lys for Arg at position 37 establishes protein IB-9 as a unique protein, different from IB-6, as opposed to only a posttranslational derivative of IB-6. A similar extension at the carboxyl terminal was demonstrated in acidic proline-rich proteins A and C from human parotid saliva (Wong & Bennick, 1980).

Proteins IB-9 and IB-6 exhibit an extraordinary degree of internal repetition as shown in Figure 2A (boxed area). The sequence Pro-Pro-Gly-Lys-Pro-Gln-Gly-Pro-Pro-Pro-Gln-Gly-Gly-Asn, with 14 amino acids, appears twice, in positions 2–15 and 23–36. Twelve of the fourteen residues recur in positions 43–54. In addition, the sequence Pro-Gln-Gly-Pro-Pro, present in the three peptides just described, occurs twice more in IB-9 and IB-6. These repetitions indicate extensive gene duplication.

Comparison of IB-9 with acidic proline-rich protein C (Wong & Bennick, 1980) from human parotid saliva reveals a high degree of homology with the carboxyl region of protein C (Figure 2B). The amino acid sequence in IB-9 from residues 2–14 and 23–35 is found also in protein C, residues 126–138. The five amino acid sequence Pro-Gln-Gly-Pro-Pro occurs 5 times in protein C as it does in IB-9. Additional homologies exist, all of which are located in the carboxyl-terminal portion of the acidic protein C. The similarity between IB-9 and the acidic proline-rich proteins is also evident from the immunological cross-reactivity and the CD spectra. As found in the acidic proline-rich proteins, the spectral characteristics do not indicate the presence of helical polyproline structure in protein IB-9.

The biological functions of the proline-rich proteins in parotid saliva are a topic of considerable interest. Acidic proline-rich proteins have been shown to bind calcium, and the binding constants are consistent with a role in maintaining the concentration of ionic calcium in saliva. They have, in addition, the ability to inhibit hydroxyapatite formation, and their role thus might be to prevent growth of hydroxyapatite crystals onto surfaces. Information about the function of the glycosylated and nonglycosylated basic proline-rich proteins is scanty and for the most part speculative. Neither protein IB-9, protein IB-6, nor any other of the basic proline-rich proteins isolated in our laboratory appears to be precursors to the basic proline-rich glycoprotein of human parotid saliva (Kauffman & Keller, 1982). They are nonetheless of interest because of their widespread, abundant, and apparently unique occurence in salivary secretion, as well as their specific responsiveness to several physiological and pharmacological conditions affecting gland function (Muenzer et al., 1979; Johnson, 1980; Anderson & Johnson, 1981).

It is evident that basic biochemical approaches to the isolation and structural characterization of salivary macromolecules are providing a new insight into salivary composition. It can be expected that such studies will expand our knowledge of the structural relationships of the acidic, basic, and glycosylated proteins and might contribute to understanding of their biological raison d'etre as well.

References

Anderson, L. C., & Johnson, D. A. (1981) *Comp. Biochem. Physiol. B 70B*, 725–730.

Anderson, L. C., Kauffman, D. L., & Keller, P. J. (1982a) *Biochem. Genet. 20*, 1301–1307.

Anderson, L. C., Lamberts, B. L., & Bruton, W. F. (1982b) *J. Dent. Res. 61*, 393–396.

Arneberg, P. (1974) *Arch. Oral Biol. 19*, 921–928.

Azen, E. A., & Denniston, C. (1980) *Biochem. Genet. 18*, 483–501.

Bennick, A., & Connell, G. E. (1971) *Biochem. J. 123*, 455–464.

Bennick, A., & Madapallimattam, G. (1982) *Arch. Oral Biol.* (in press).

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

Castle, J. D., & Palade, G. E. (1978) *J. Cell Biol. 76*, 323–340.

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

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

Johnson, D. A. (1980) *J. Dent. Res. 59A*, 541, Abstr. 1089.

Kauffman, D. L., & Keller, P. J. (1979) *Arch. Oral Biol. 24*, 249–256.

Kauffman, D. L., & Keller, P. J. (1982) *J. Dent. Res. 61*, 294, Abstr. 1043.

Kauffman, D. L., Wong, R. S. C., Bennick, A., & Keller, P. J. (1982) *J. Dent. Res. 61*, 294, Abstr. 1043.

Keller, P. J., Robinovitch, M. R., Iveresen, 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.

Mitchell, W. M., & Harrington, W. F. (1970) *Methods Enzymol. 19*, 635–642.

Moser, P. W., & Rickli, E. E. (1979) *J. Chromatogr. 176*, 451–455.

Muenzer, J., Bildstein, C., Gleason, M., & Carlson, D. M. (1979) *J. Biol. Chem. 254*, 5623–5628.

Oppenheim, F. C., Hay, D. I., & Franzblau, C. (1971) *Biochemistry 10*, 4233–4238.

Oppenheim, F. C., Kousvelari, E., & Troxler, R. F. (1979) *Arch. Oral Biol. 24*, 595–599.

Schlesinger, D. H., & Hay, D. I. (1979) in *Peptides, Proceedings of the American Peptide Symposium, 6th* (Gross, E., & Meienhofer, J., Eds.) pp 133, Pierce Chemical Co., Rockford, IL.

Schlesinger, D. H., & Hay, D. I. (1981) *Int. J. Pept. Protein Res. 17*, 34–41.

Wong, R. S. C., & Bennick, A. (1980) *J. Biol. Chem. 255*, 5943–5848.

Wong, R. S. C., Hofmann, T., & Bennick, A. (1979) *J. Biol. Chem. 254*, 4800–4808.