Biochem. J. (1988) 255, 15–21 (Printed in Great Britain)

The primary structures of six human salivary acidic proline-rich proteins (PRP-1, PRP-2, PRP-3, PRP-4, PIF-s and PIF-f)

Donald I. HAY,*§ Anders BENNICK,† David H. SCHLESINGER,‡ Kiyoshi MINAGUCHI,† George MADAPALLIMATTAM† and Susan K. SCHLUCKEBIER*

*Department of Biochemistry, Forsyth Dental Center, 140 Fenway, Boston, MA 02115, U.S.A., †Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada, M5S 1A8, and ‡Department of Medicine and Cell Biology, New York University Medical Center, 550 First Avenue, New York, NY 10016, U.S.A.

Human glandular salivary secretions contain several acidic proline-rich phosphoproteins (PRPs). These proteins have important biological functions related to providing a protective environment for the teeth, and appear to possess other activities associated with modulation of adhesion of bacteria to oral surfaces. These functions and activities depend on the primary structures of the PRPs. Previously determined amino acid sequences of two 150-residue molecules, PRP-1 and PRP-2, and two related 106-residue proteins, PRP-3 and PRP-4, indicated that residue 4 was Asn in PRP-1 and PRP-3, and Asp in PRP-2 and PRP-4, and position 50 was Asn in all four proteins. Recent data from cDNA sequence studies and further structural studies, however, showed that the previously proposed sequences cannot be completely correct. The present work has shown that the protein previously designated as PRP-1 actually consisted of two positional isomers, PIF-s, which has Asn and Asp at positions 4 and 50 respectively, and authentic PRP-1, which has the reverse arrangement. The same isomerism is present in the smaller proteins, PIF-f and PRP-3. Since the isomeric pairs have identical compositions and charges, their presence was not previously detected. Also, by using a more highly purified preparation, it has been found that position 50 in PRP-2 and PRP-4 is Asp, rather than Asn previously reported. These new findings for the six PRPs define their complete primary structures, which are now consistent with those proposed for PRP-1 and PIF-s from cDNA data, and are also consistent with the chromatographic and electrophoretic behaviours of the six PRPs and their derived peptides. These corrected structures are important for understanding the biological functions and activities of these unusual proteins.

INTRODUCTION

Human parotid and submandibular salivas contain several closely related acidic proline-rich phosphoproteins (PRPs), which were identified by isolation [Proteins A and C (Bennick & Connell, 1971); PRP-1, -2, -3 and -4 (Oppenheim *et al.*, 1971)], or by studies of protein polymorphism [identified as PIF-s, PIF-f, Db-s, Db-f and Pa (Azen & Oppenheim, 1973; Friedman & Merritt, 1975; Azen & Denniston, 1974, 1981)]. The primary structures of four PRPs have been reported, namely Protein C (identical with PRP-1) (Wong & Bennick, 1980), PRP-2 (Schlesinger & Hay, 1986), Protein A (identical with PRP-3) (Wong *et al.*, 1979) and PRP-4 (Schlesinger & Hay, 1979). The designations PRP-1, -2, -3 and -4 will be used here. PRP-1 and -2 are 150-amino-acid-residue proteins which were reported to differ in that residue 4 was Asn in PRP-1 and Asp in PRP-2 (Schlesinger *et al.*, 1977). PRP-3 and -4 are composed of 106 residues, which are identical with the first 106 residues of PRP-1 and PRP-2 respectively. Because of this identity of structure, the co-inheritance of the protein pairs PRP-1/PRP-3 and PRP-2/PRP-4 (Azen & Oppenheim, 1973), and the presence in saliva of the 44-residue peptide corresponding to residues 107–150 of the larger proteins (Isemura *et al.*, 1980), the smaller proteins are considered to be derived from the larger molecules by post-translational cleavage (Karn *et al.*, 1979; Wong *et al.*, 1983).

PRPs isolated in two independent investigations (Bennick & Connell, 1971; Oppenheim *et al.*, 1971) appeared to be of high purity in that they each gave a single band on native and SDS/polyacrylamide-gel electrophoresis and conventional electrofocusing, a single peak on gel filtration, anion-exchange and C_{18} reverse-phase chromatography, and appeared homogeneous when sequenced. Nonetheless, it was found more recently (Schluckebier *et al.*, 1986) that some tryptic peptides (residues 1–30, 31–74 and 31–106) from preparations of PRP-1 and PRP-3 were heterogeneous when examined by methods with higher resolution than those used previously. The corresponding peptides from PRP-2 and PRP-4, however, appeared essentially homogeneous under the same conditions.

Coincident with these observations, Maeda *et al.* (1985) reported sequences of cDNA clones corresponding to acidic PRP mRNA from a subject in whom the expression of the PRPs was limited to the pairs PRP-1/PRP-3 and PIF-s/PIF-f. Two sequences of the same length as PRP-1 and PIF-s were deduced from the cDNA data, which differed in position 4 and 50. Whereas one sequence contained Asp in position 4 and Asn in

Abbreviations used: PRP, human acidic proline-rich phosphonrotein; FPLC, fast protein liquid chromatography.

§ To whom correspondence and reprint requests should be sent.

These sequence data have been submitted to the EMBL.

position 50 (PRP-1), the opposite arrangement was found in the other deduced sequence (PIF-s). These sequences agreed completely with the previously determined amino acid sequence of protein C, except that in this protein residues 4 and 50 were both identified as Asn, suggesting that it was a mixture of PRP-1 and PIF-s. An attempt was made by Kim & Maeda (1986) to assign these sequences to PRP-1 and PIF-s. They suggested that PIF-s had Asn at position 4 and Asp at position 50, but the assignment was ambiguous, because part of the cDNA was derived from a subject who was heterozygous for PRP-1, who synthesized PRP-2 as well.

Because of the discrepancies between the primary structures of the PRPs determined by amino acid sequencing, and those proposed from cDNA data, and since PIF-s and PIF-f have not been isolated and characterized, the purpose of this study was to explain the above-noted discrepancies and to assign correct sequences to the primary gene products PRP-1, PIF-s and PRP-2, and to the derived proteins, PRP-3, PIF-f and PRP-4.

MATERIALS AND METHODS

Materials

Trypsin (EC 3.4.21.4), pancreatic trypsin inhibitor, cacodylic acid, Naphthol Blue Black (Color Index No. 20470) and pyroglutamate aminopeptidase (E.C. 3.4.19.3) from calf liver were obtained from Sigma Chemical Co., St. Louis, MO, U.S.A.; Stains-All was from Eastman-Kodak Co., Rochester, NY, U.S.A.; hydrolysed starch was from Fluka Chemical Corp., Ronkonkoma, NY, U.S.A.; Ampholines were from LKB Produkter, Stockholm, Sweden; Sepharose 4B, Sephadex G-50 and DEAE-Sephadex were from Pharmacia, Piscataway, NJ, U.S.A.; CM32 CM-cellulose was from Whatman, Maidstone, Kent, U.K. All other chemicals were reagent grade.

Collection of saliva

Individual parotid saliva samples were collected from two individuals (A. B. and D. H.) as described by Bennick & Connell (1971). For large-scale preparation of PRPs, stimulated parotid saliva collected from a group of 20 subjects was pooled, dialysed against water and freeze-dried to give a stock of parotid salivary proteins.

Isoelectric focusing

PRP phenotypes were determined and individual purified PRPs were identified by isoelectric focusing on a horizontal slab gel with a gradient from pH 3.5 to 5.2 as described by Azen & Denniston (1981). The protein bands were made visible by treating the gel with 20% (w/v) trichloroacetic acid for 20 min.

Column chromatography

A mixture of freeze-dried parotid salivary protein (5 g) from 20 subjects was chromatographed on DEAE-Sephadex as described by Hay (1975) to obtain the PRP-containing fraction. This was chromatographed with shallow chloride gradients (Oppenheim *et al.*, 1971; Hay & Oppenheim, 1974) to give partially purified preparations of the PRPs. The individual PRPs were obtained from these preparations by fast protein liquid chromatography (FPLC) on a 0.79 $cm^2 \times 10$ cm Mono-Q anion-exchange column (Pharmacia). Saliva samples from individual subjects were treated to remove kallikrein by chromatography on Sepharose 4B to which had been coupled pancreatic trypsin inhibitor as outlined by Wong *et al.* (1983). This was followed by separation on Sephadex G-50 and CM32 CM-cellulose as described by Bennick (1975, 1977*a*). Individual purified PRPs were then obtained by FPLC on a 0.2 $cm^2 \times 5$ cm Mono-Q anion-exchange column. Specific details for each fractionation are given in the Results section.

Tryptic digestion

Purified PRP-3, PIF-f and PRP-4 were digested with trypsin for 5 min as described by Wong *et al.* (1979), and the enzyme was inactivated by adding pancreatic trypsin inhibitor. Under these conditions, cleavage is essentially restricted to Arg_{30}–Gln_{31} (see the Results section; Fig. 6) to give the acidic, 30-residue, phosphorylated *N*-terminal peptide (TX peptide) and the 76-residue *C*-terminal segment, residues 31–106 (TY peptide). This selective cleavage stems from the much slower cleavage rate at Arg–Pro and Lys–Pro bonds, compared with other bonds (Allen, 1981). Purified PRP-1, PIF-s and PRP-2 were treated as described previously (Schlesinger & Hay, 1979, 1981, 1986) to give TX peptides and the peptide formed by residues 31–74.

Polyacrylamide-gel electrophoresis

Basic gels. For separation of TX peptides, a 30%-polyacrylamide slab gel (17 cm × 15 cm, 0.8 mm thick) was used. The gel buffer was 0.5 M-Tris sulphate, pH 9.0, and the reservoir buffer 0.065 M-Tris borate, pH 9.0 (Smith *et al.*, 1975). The samples contained approx. 1 μg of TX peptide. A potential of 130 V was applied for 22 h. The phosphopeptides were stained with Stains-All (Green *et al.*, 1973).

Acidic gels. To seperate TY peptides, electrophoresis was done in a 25%-polyacrylamide gel (8 cm × 6 cm, 0.75 cm thick; 'Minigel', from Hoefer Scientific Instruments, San Francisco, CA, U.S.A.). The gel and reservoir buffers were 0.25 M-sodium cacodylate, pH 5.5. Before use, persulphate was removed by electrophoresis at 150 V for 16 h. Approx. 1 μg of each of the TY peptides was applied. A current of 6 mA was maintained for 5–6 h at 5 °C. The gel was stained by a modification of the method described by Sung & Smithies (1969).

Amino acid sequence analysis

Before sequence analysis PRP-1, PRP-2, PRP-3, PRP-4, PIF-s and PIF-f were treated with pyroglutamate aminopeptidase to remove *N*-terminal pyrrolidonecarboxylic acid as previously described (Schlesinger *et al.*, 1977). These preparations were sequenced as described below to identify the residue at position 4. To identify the residue at position 50 the same proteins were digested with trypsin, but not with pyroglutamate aminopeptidase, and subjected to the sequence analysis. Also, purified TY peptides derived from PRP-3, PRP-4 and PIF-f were subjected to sequence analysis. Automated Edman degradations were performed on a gas-phase protein sequenator (model 470A; Applied Biosystems Inc., Foster City, CA, U.S.A.), by using a single-coupling single-cleavage protein program. The resulting amino acid phenylthiohydantoins were identified on a reversed-phase small-bore on-line high-perfor-

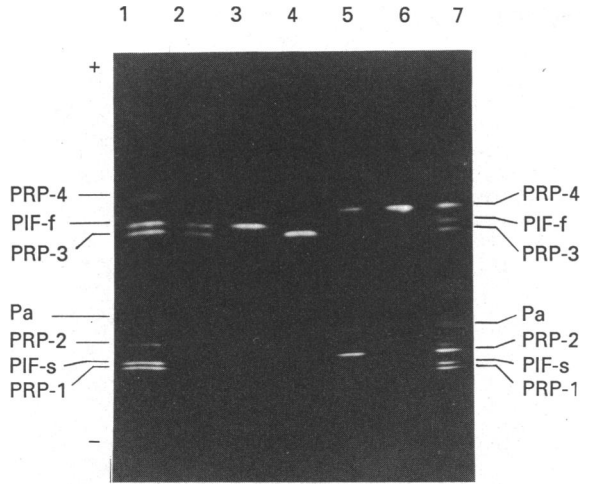

Fig. 1. Isoelectric focusing of salivary-protein samples and PRP preparations in a pH 3.5–5.2 gradient

The bands corresponding to PRP-1, PRP-2, PRP-3, PRP-4, PIF-s, PIF-f and Pa are labelled. Lanes: 1, PRP from parotid saliva from A.B. (contains PRP-1, PIF-s, PRP-3 and PIF-f); 2, Protein A (a mixture of authentic PRP-3 and PIF-f); 3, purified PIF-f; 4, purified PRP-3; 5, PRPs from parotid saliva from D.H. (contains PRP-2, PRP-4 and Pa); 6, purified PRP-4; 7, parotid saliva from an individual known to contain PRP-1, PIF-s, PRP-2, Pa, PRP-3, PIF-f and PRP-4.

mance liquid chromatograph (model 120A; Applied Biosystems Inc.).

RESULTS

Typing of saliva

Fig. 1 shows examples of the patterns obtained by isoelectric focusing of PRP from salivas of genetically different subjects. PIF-s, PIF-f, PRP-1 and PRP-3 are present in the parotid saliva from subject A.B. (lane 1) and PRP-2, PRP-4 and Pa in the saliva from subject D.H. (lane 5). An unexplained feature of these results is that even the highly purified proteins show unexpected and apparently artifactual minor bands. It is possible that this is due to deamidation, since incubation of pure PRPs under conditions known to cause deamidation (Lewis *et al.*, 1981) increases the intensity of the minor bands when proteins treated in this manner are examined by electrofocusing.

Purification of PRP-1, PIF-s and PRP-2

Previously reported procedures (Oppenheim *et al.*, 1971) gave the electrophoretically distinct and chromatographically separable proteins 'PRP-1' and 'PRP-2'. This separation was not easily achieved, and involved several repetitions of an anion-exchange step to give apparently homogeneous proteins. Analytical-scale (approx. 5 μg) re-examination of 'PRP-1' by FPLC on a Mono-Q anion-exchange column (0.2 cm^2 × 5 cm) with a chloride gradient (0.22–0.30 M-NaCl in 0.5 M-Tris/HCl, pH 8.0, flow rate 1 ml/min, over 25 min), with monitoring of the eluate at 220 nm, showed the previous preparation of 'PRP-1' to contain two major protein constituents (Fig. 2), now designated PIF-s and authentic PRP-1. The 'PRP-2' preparation was found to contain mainly PRP-2, but some PRP-1 was also present (Fig. 2).

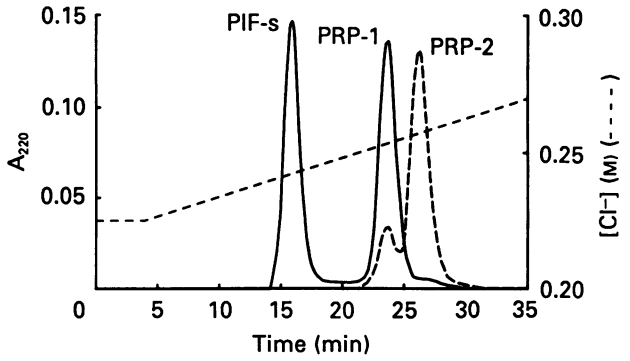

Fig. 2. Anion-exchange chromatography of PRP-1, PIF-s and PRP-2

A 0.2 cm^2 × 5 cm Mono-Q column was used, eluted with 0.05 M-Tris/HCl buffer, pH 8.0, containing a chloride gradient (----). (*a*) A 10 μg portion of 'PRP-1' (a mixture of authentic PRP-1 and PIF-s) was resolved into its constituent proteins (——). (*b*) Original PRP-2 (12 μg) was shown to be contaminated by PRP-1 (----). Comparable results were obtained from samples of 'PRP-3' and PRP-4.

Larger-scale purification (5 mg samples/run) of these three proteins on a preparative Mono-Q column (0.79 cm^2 × 10 cm) with a chloride gradient (0.220–0.276 M-NaCl in 0.05 M-Tris/HCl, pH 8.0, flow rate 4 ml/min, over 35 min), collecting the central portions of the individual peaks, gave good yields of highly purified PRP-1, PIF-s and PRP-2.

Purification of PIF-f and PRP-3

The preparation previously designated as Protein A (PRP-3), actually a mixture of authentic PRP-3 and PIF-f (Fig. 1, lane 2), was prepared from A.B.'s parotid saliva as described by Bennick (1975). Approx. 800 μg of this protein was applied to a Mono-Q column (0.2 cm^2 × 5 cm) and eluted with an NH_4HCO_3 gradient (0.25–1.0 M, pH 8.3), the eluate being monitored at 220 nm. Fractions forming the two peaks obtained (similar to those obtained for PRP-1 and PIF-s; Fig. 2) were pooled separately and chromatographed three times under the same conditions to give highly purified preparations of PRP-3 and PIF-f. Central cuts of the peaks were pooled, and the identity and purity of the proteins were confirmed by isoelectric focusing (Fig. 1, lanes 3 and 4).

Purification of PRP-4

Parotid saliva (400 ml) was collected from D.H. Kallikrein was removed from other salivary proteins by passing the sample through a Sepharose 4B column (15 ml) to which had been coupled pancreatic trypsin inhibitor. The column was equilibrated with 0.1 M-Tris/HCl, pH 8.6, and eluted with the same buffer until all unbound protein had been recovered. The product was dialysed against 10 mM-EDTA and then water, chromatographed on a Sephadex G-50 column, and the first peak obtained was chromatographed on a column of CM32 CM-cellulose as previously described (Bennick, 1975, 1977*a*). The first peak eluted from the column contained PRP-4, which was purified further by FPLC using a Mono-Q column as described for PRP-3 and PIF-f. The elution profile from the latter separation

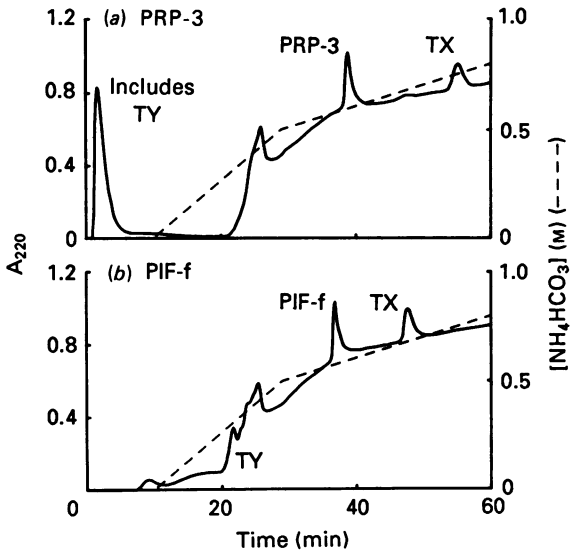

Fig. 3. Purification of tryptic peptides from PRP-3 and PIF-f

(*a*) A 5 min tryptic digest of PRP-3 (110 μg) was fractionated on a 0.2 cm^2 × 5 cm Mono-Q column at a flow rate of 0.4 ml/min with a gradient of 0.005–1 M-NH_4HCO_3 (----); 0.3 ml fractions were collected. The initial buffer concentration of 0.005 M was increased to 0.5 M over 15 min and then to 0.8 M over the next 30 min. The *C*-terminal TY peptide was eluted in the starting buffer, undigested PRP-3 at 0.58 M and the *N*-terminal TX peptide at 0.7 M-NH_4HCO_3. The peak at 0.42 M is a buffer artifact. (*b*) A 5 min tryptic digest of PIF-f (125 μg) was chromatographed in the same way. Peptide TY was eluted at 0.31 M, undigested PIF-f at 0.57 M and peptide TX at 0.65 M-NH_4HCO_3. The peak at 0.39 M is also a buffer artifact.

showed a single symmetrical peak. The product was shown by isoelectric focusing to contain only PRP-4 (Fig. 1, lane 6).

Purification and electrophoretic characterization of tryptic peptides

Chromatography of tryptic peptides from PRP-3 and from PIF-f on a Mono-Q column gave distinctly different elution profiles. For the PRP-3 digest (Fig. 3*a*), the *C*-terminal TY peptide was eluted in the starting buffer, the undigested PRP-3 at 0.58 M-NH_4HCO_3, and the *N*-terminal TX peptide at 0.7 M-NH_4HCO_3. In contrast, the TY peptide from the PIF-f digest was not eluted until after the gradient had been applied (Fig. 3*b*), at about 0.3 M-NH_4HCO_3, the undigested PIF-f was eluted slightly earlier than PRP-3, and the TX peptide from PIF-f was eluted significantly earlier than the corresponding peptide from PIF-s, at about 0.65 M-NH_4HCO_3.

Basic-gel electrophoresis of tryptic digests of PRP-3, PRP-4 and PIF-f, and TX peptides purified from the digests, gave the results shown in Fig. 4. Only the acidic phosphorylated *N*-terminal TX peptides enter the gel and appear as blue-coloured bands, typical of phosphorylated proteins and peptides (Azen, 1978). The original preparation of 'PRP-3' (now shown to be a mixture of authentic PRP-3 and PIF-f) gave two such bands, but the tryptic digests of PRP-3 and PRP-4 and the purified TX peptides prepared from PRP-3 gave only one band which corresponded to the faster-migrating

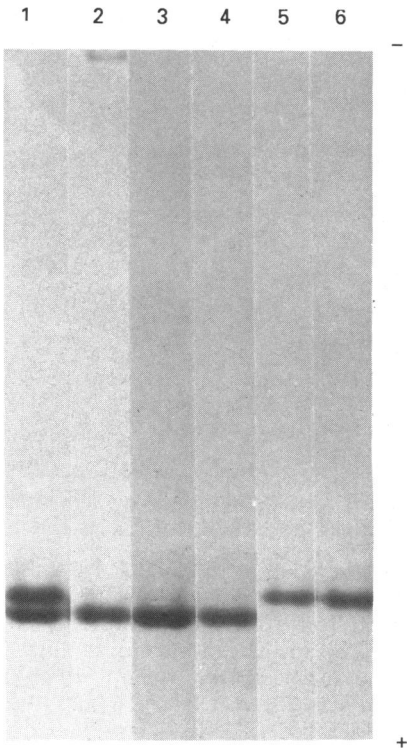

Fig. 4. Polyacrylamide-gel electrophoresis of *N*-terminal phosphorylated TX peptides

The origin is at the cathode (top). Approx. 1 μg of each of the TX peptides was applied and, after electrophoresis, the gel was stained with Stains-All. Lanes: 1, tryptic digest of 'PRP-3' (Protein A; a mixture of authentic PRP-3 and PIF-f); 2, tryptic digest of purified PRP-4; 3, tryptic digest of purified PRP-3; 4, TX peptide isolated from a digest of PRP-3; 5, tryptic digest of PIF-f; 6, peptide TX isolated from a digest of PIF-f.

band from 'PRP-3'. The tryptic digest of PIF-f and the TX peptide prepared from PIF-f gave a single band with a mobility identical with that of the slower-migrating band from 'PRP-3'. These results suggest that the TX peptides from PRP-3 and PRP-4 are identical and possess a higher negative charge than the corresponding peptide from PIF-f.

The results of acidic-gel electrophoresis of tryptic digests of PRP-3, PRP-4 and PIF-f, and *C*-terminal TY peptides prepared from these proteins, are shown in Fig. 5. Under the experimental conditions used, undigested protein and the *N*-terminal TX peptide do not enter the gel. The digest of 'PRP-3' (mixture of authentic PRP-3 and PIF-f) showed two heavily stained bands corresponding to the TY peptides, but digests of PRP-4 and PIF-f as well as TY purified from PIF-f showed only a single heavily stained band with mobility corresponding to the slower-migrating component of the 'PRP-3' digest. Electrophoresis of a tryptic digest of PRP-3 and the TY peptide prepared from the digest both showed a single band that migrated at the same rate as the faster band of the 'PRP-3' digest. Minor bands are seen because of additional but much slower tryptic cleavages at Arg–Pro and Lys–Pro bonds. These results are consistent with the TY peptide from both PIF-f and PRP-4 having an additional negative charge compared with the same peptide prepared from PRP-3.

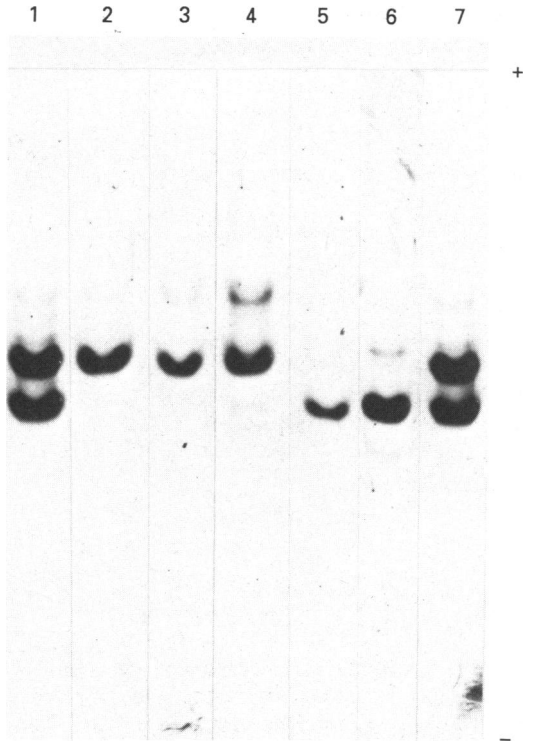

Fig. 5. Polyacrylamide-gel electrophoresis of *C*-terminal basic TY peptides

The origin is at the anode (top). About 1 μg of each of the TY peptides was applied. After electrophoresis the gel was stained with Naphthol Blue-Black. Lanes: 1 and 7, tryptic digest of 'PRP-3' (Protein A; mixture of authentic PRP-3 and PIF-f); 2, tryptic digest of PRP-4; 3, tryptic digest of PIF-f; 4, TY peptide isolated from a digest of PIF-f; 5, tryptic digest of purified PRP-3; 6, TY peptide prepared from purified PRP-3.

Sequencing of PRP-1, PIF-s, PRP-2, PRP-3, PIF-f and PRP-4

The results of amino acid sequence analysis of residue 4 of these six proteins showed unequivocally that this position is Asp in PRP-1, PRP-2, PRP-3 and PRP-4, and Asn in PIF-s and PIF-f. Also, the 'mixture' sequence analyses of the tryptic peptides from the six PRPs, and the purified TY peptides from PIF-f and PRP-3, provided unequivocal identification of the position 50 (cycle 20 of the automated Edman degradation) as Asn in PRP-1 and PRP-3, and Asp in PIF-s, PIF-f, PRP-2 and PRP-4. For example, the net yields from sequencing of the purified TY peptides are shown in Fig. 6.

The revised sequence of PRP-1 is shown in Fig. 7, and the variations in the structures of the six PRPs investigated are given in Fig. 8. Differences from the previously published sequences are at residues 4 and 50. PRP-1 and PIF-s are now shown to be positional isomers, in which the aspartate and asparagine residues at positions 4 and 50, respectively, in PRP-1 are reversed in PIF-s. In PRP-2, both positions are aspartate. These sequences are now consistent with those proposed for PRP-1 and PRP-3, and PIF-s and PIF-f, from cDNA sequence determinations (Maeda *et al.*, 1985), and are consistent with the observed differences in chromatographic and electrophoretic behaviours of the proteins and derived peptides discussed below.

DISCUSSION

Knowledge of the structures of the PRPs is important for several reasons. Thus these proteins form 10–40% of the protein present in human glandular salivas (Kousvelari *et al.*, 1980; Baum *et al.*, 1982; Mandel & Bennick, 1983). Functionally, they act as inhibitors of surface-induced calcium phosphate precipitation on to teeth from saliva (Hay & Gron, 1976; Hay *et al.*, 1979; Moreno *et al.*, 1979), which is highly supersaturated with respect to enamel mineral (Hay *et al.*, 1982), and they also act as Ca^{2+}-binding proteins (Bennick, 1977*b*). These properties consort to provide a protective and reparative environment for dental enamel, which is important for the integrity of the teeth.

For the PRPs to act as inhibitors of surface-induced calcium phosphate precipitation, they must adsorb on mineral surfaces (Hay, 1973), such as tooth surfaces, where they form part of the adsorbed protein film, or pellicle, normally present on teeth (Kousvelari *et al.*, 1980; Bennick *et al.*, 1983). In this location the PRP may also act as receptors for commensal and pathogenic organisms which colonize teeth, and recently the PRP have been identified as selective mediators of bacterial adhesion to hydroxyapatite (Gibbons & Hay, 1988*a*,*b*), which is the prototype mineral for dental enamel. This observation seems highly significant in that differences in the phenotypic expression of the PRPs have been related to variations in oral disease (Friedman *et al.*, 1980; Yu *et al.*, 1986). All the foregoing properties depend on structural features of the PRPs (Bennick *et al.*, 1979, 1981; Moreno *et al.*, 1982, 1984; Aoba *et al.*, 1984; Hay *et al.*, 1987). For example, the single-residue substitutions of Asp for Asn, identified in the present studies, significantly change the adsorption affinities of the PRP for hydroxyapatite surfaces (Hay & Moreno, 1979; Moreno *et al.*, 1982, 1984).

	1								10									20				
TY, PIF-f	Q	G	P	P	L	G	G	Q	Q	S	Q	P	S	A	G	D	G	N	Q	D	D	G P Q Q
Net yield																				35	17	
TY, PRP-3	Q	G	P	P	L	G	G	Q	Q	S	Q	P	S	A	G	D	G	N	Q	N	D	G P Q Q
Net yield																				154	225	

Fig. 6. Amino acid sequencing of TY peptides from PIF-f and PRP-3

The sequences above were determined in a gas-phase sequencer by using 1.75 nmol of TY from PIF-f, and 3 nmol of TY from PRP-3, purified from their respective tryptic digests. The net yields are shown (below the sequence) for cycles 20–21. In these cycles there was no net yield of other amino acids except for TY from PIF-f, where 3 pmol of P was found in cycle 21. The repetitive yield based on the recovery of P in cycles 3 and 23 was 91%.

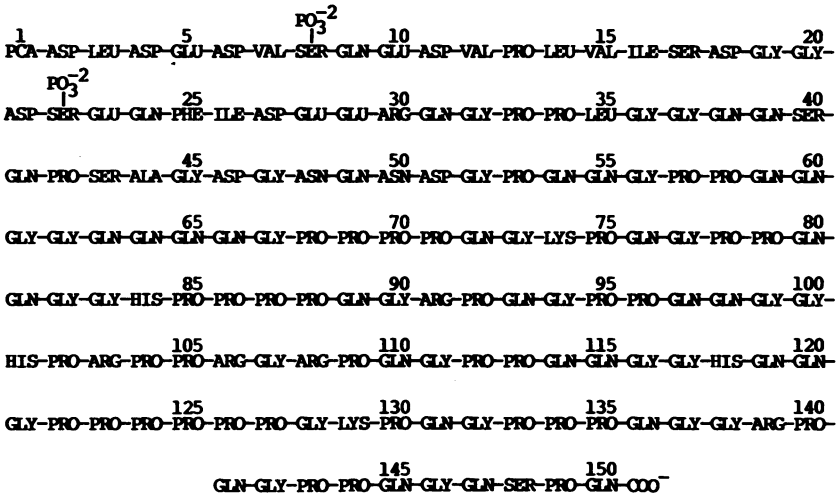

Fig. 7. Revised sequence of PRP-1

Abbreviation: PCA, pyrrolidonecarboxylic acid.

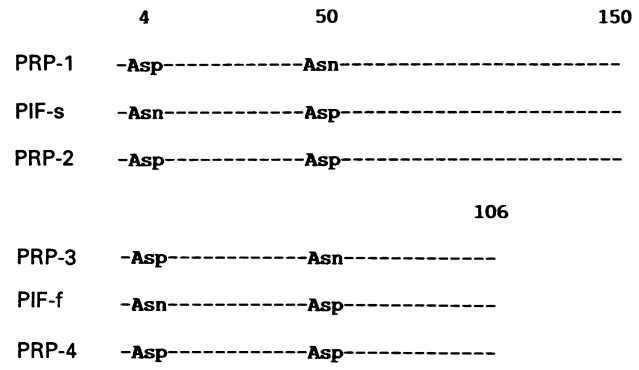

Fig. 8. Summary of the structural variations determined by sequencing for six PRPs

Data for PIF-s and PIF-f, and PRP-1 and PRP-3, are in agreement with proposed structures based on cDNA sequences.

The findings reported here now define the complete primary structures of the three 150-residue primary gene products, PRP-1, PIF-s and PRP-2, and the three 106-residue derived proteins, PRP-3, PIF-f and PRP-4, considered to be formed by post-translational cleavage of the larger PRP at Arg_{106} (Fig. 7). The experimentally determined structures of PRP-1 and PIF-s are now consistent with those proposed from cDNA sequences (Maeda *et al.*, 1985) and correct the previous assignments of Asn at position 4 in PRP-1 (Wong & Bennick, 1980), now corrected to Asp, and Asn at position 50 in PRP-2 (Schlesinger & Hay, 1986), also corrected to Asp. In addition, PIF-s, previously identified by electrofocusing (Azen & Denniston, 1974), has now been isolated in highly purified form and its structure established. Corresponding corrections are reported for the 106-residue derived proteins, PRP-3 and PRP-4, and a third derived protein, PIF-f, has now been isolated and characterized. It is noteworthy that in the present studies no net yield of Asn was detected at the positions now shown to be Asp, and similarly, at the positions now designated as Asn there was no net yield of Asp.

Further confirmatory evidence of the proposed sequences is obtained from the electrophoretic and chromatographic properties of the PRPs and their tryptic peptides. These properties are completely consistent with the structures and charge distribution now proposed. To examine the TX peptides (residues 1–30), a 30%-polyacrylamide gel was used to prevent this highly negatively charged small peptide from travelling at the dye front. Because they are phosphorylated, the TX peptides stain readily with Stains-All, although they stained poorly with conventional silver stains. The electrophoretic patterns (Fig. 4) confirm the purity of the preparations, and the mobilities of the TX peptides are consistent with the presence of Asp in position 4 in PRP-3 and PRP-4, and with Asn at this location in PIF-f.

To maximize differences in electrophoretic mobilities of the TY peptides, a buffer system at pH 5.5 was used. At this pH all the acidic side chains are charged (Bennick *et al.*, 1981), so increasing the possibility of determining a difference of one negative charge, but the imidazole rings in histidine would be positively charged, thereby increasing the cationic mobility of the peptides. These peptides stained easily by a modified version of the Naphthol Blue/cobalt nitrate method of Sung & Smithies (1969), although they could not be stained by conventional silver staining. The relative mobilities of the TY peptides (Fig. 5) are consistent with the TY peptide from PRP-3 containing Asn, and the TY peptides from PIF-f and PRP-4 containing Asp at residue 50. Similar results were obtained for PRP-1, PIF-s and PRP-2.

The chromatographic behaviour of tryptic peptides from PIF-f and PRP-3 on a Mono-Q anion-exchange column also fits with the revised structures. Fig. 3(*a*) shows that peptide TY from PRP-3 (Asn at 50) is not retarded on this column, whereas the corresponding peptide from PIF-f (Asp at 50) is significantly retarded (Fig. 3*b*). Similarly, the TX peptide from PRP-3 (Asp at 4) is considerably more retarded than the corresponding TX peptide from PIF-f (Asn at 4). The later elution of PRP-3 compared with PIF-f is also seen.

The electrofocusing results (Fig. 1) are also consistent with the revised structures. Thus PRP-2, which has an additional negative charge compared with both PIF-s

and PRP-1, has a significantly lower isoelectric point than the last two proteins. Since PIF-s and PRP-1 have identical charges and compositions, it might be expected that they would not be resolved. In fact, PIF-s has a slightly lower isoelectric point than PRP-1. Presumably the local environments at positions 4 and 50 are sufficiently different to affect the p*K* of the Asp at these two locations. The fact that PIF-s was eluted earlier than PRP-1 from the Mono-Q column, but has a lower isoelectric point, is not contradictory, since all acidic residues will be fully ionized at the pH at which this chromatography was performed, a point which emphasizes the significance of the role of structural features in chromatographic behaviour (Regnier, 1987). The foregoing considerations also apply to the behaviours of PRP-3, PIF-f and PRP-4.

The biological significance of these variations in the primary structures of the PRPs, which are widely distributed throughout the human population, and the significance of the occurrence of the 150-residue and 106-residue proteins, are not yet known. Similar variations may be expected for the incompletely characterized protein pair Db-s and Db-f, although these proteins appear to be larger than the other PRPs (Azen & Denniston, 1974), and Pa differs from the other PRPs in that it contains cysteine (Friedman & Merritt, 1975). Further work is needed to clarify these problems.

Note added in proof (received 18 July 1988)

Since the preparation of this manuscript, a paper by Azen *et al.* (1987), reporting the DNA sequences coding for the proteins Db-s, Db-f and Pa, came to our attention.

These studies were supported by U.S. Public Health Service grants DE 03915, DE 07009 and DE 06159 from the National Institute for Dental Research, grant MT 4920 from the Canadian Medical Research Council, and a grant to D.H.S. from the Kaplan Cancer Center.

REFERENCES

Allen, G. (1981) Sequencing of Proteins and Peptides, p. 53, North-Holland Publishing Co., Amsterdam

Aoba, T., Moreno, E. C. & Hay, D. I. (1984) Calcif. Tissue Int. **36**, 651–658

Azen, E. A. (1978) Arch. Oral Biol. **23**, 1173–1176

Azen, E. A. & Denniston, C. L. (1974) Biochem. Genet. **12**, 109–120

Azen, E. A. & Denniston, C. (1981) Biochem. Genet. **19**, 475–485

Azen, E. A. & Oppenheim, F. G. (1973) Sceince **180**, 1067–1069

Azen, E. A., Kim, H. S., Goodman, P., Slynn, S. & Maeda, N. (1987) Am. J. Hum. Genet. **41**, 1035–1047

Baum, B. J., Kousvelari, E. E. & Oppenheim, F. G. (1982) J. Gerontol. **37**, 392–395

Bennick, A. (1975) Biochem. J. **145**, 557–567

Bennick, A. (1977*a*) Biochem. J. **163**, 229–239

Bennick, A. (1977*b*) Biochem. J. **163**, 241–245

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

Bennick, A., Cannon, M. & Madapallimattam, G. (1979) Biochem. J. **183**, 115–126

Bennick, A., McLaughlin, A. C., Grey, A. A. & Madapallimattam, G. (1981) J. Biol. Chem. **256**, 4741–4746

Bennick, A., Chau, G., Goodlin, R., Abrams, S., Tustian, D. & Madapallimattam, G. (1983) Arch. Oral Biol. **28**, 19–27

Friedman, R. D. & Merritt, A. D. (1975) Am. J. Hum. Genet. **27**, 304–314

Friedman, R. D., Azen, E. A., Yu, P. L., Green, P. A., Karn, R. C. & Merritt, A. D. (1980) Hum. Hered. **30**, 372–375

Gibbons, R. J. & Hay, D. I. (1988*a*) Infect. Immun. **56**, 439–445

Gibbons, R. J. & Hay, D. I. (1988*b*) in Molecular Mechanisms of Microbial Adhesion: Symposium (Beachey, E. H., Hook, M., Switalski, L. & Wells, R., eds.), Springer-Verlag, New York, in the press

Green, M. R., Pastewka, J. V. & Peacock, A. C. (1973) Anal. Biochem. **56**, 43–51

Hay, D. I. (1973) Arch. Oral Biol. **20**, 1517–1529

Hay, D. I. (1975) Arch. Oral Biol. **20**, 553–558

Hay, D. I. & Gron, P. (1976) Microb. Aspects Dental Caries: Spec. Suppl. Microbiol. Abstr. **1**, 143–150

Hay, D. I. & Oppenheim, F. G. (1974) Arch. Oral Biol. **19**, 627–632

Hay, D. I. & Moreno, E. C. (1979) J. Dent. Res. **58**, 930–942

Hay, D. I., Moreno, E. C. & Schlesinger, D. H. (1979) Inorg. Perspect. Biol. Med. **2**, 271–285

Hay, D. I., Schluckebier, S. K. & Moreno, E. C. (1982) Calcif. Tissue Int. **34**, 531–538

Hay, D. I., Carlson, E. R., Schluckebier, S. K., Moreno, E. C. & Schlesinger, D. H. (1987) Calcif. Tissue Int. **40**, 126–132

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

Karn, R. C., Freidman, R. D. & Merritt, A. D. (1979) Biochem. Genet. **17**, 1061–1077

Kim, H. S. & Maeda, N. (1986) J. Biol. Chem. **261**, 6712–6718

Kousvelari, E. E., Baratz, R. S., Burke, B. & Oppenheim, F. G. (1980) J. Dent. Res. **59**, 1430–1438

Lewis, U. J., Singh, R. N. P., Bonewald, L. F. & Seavey, B. K. (1981) J. Biol. Chem. **256**, 11645–11650

Maeda, N., Kim, H. S., Azen, E. A. & Smithies, O. (1985) J. Biol. Chem. **260**, 11123–11130

Mandel, I. D. & Bennick, A. (1983) J. Dent. Res. **62**, 943–945

Moreno, E. C., Varughese, K. & Hay, D. I. (1979) Calcif. Tissue Int. **28**, 7–16

Moreno, E. C., Kresak, M. & Hay, D. I. (1982) J. Biol. Chem. **257**, 2981–2989

Moreno, E. C., Kresak, M. & Hay, D. I. (1984) Calcif. Tissue Int. **36**, 48–59

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

Regnier, F. E. (1987) Science **238**, 319–323

Schlesinger, D. H. & Hay, D. I. (1979) Pept. Struct. Biol. Funct.: Proc. Am. Pept. Symp. 6th 133–135

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

Schlesinger, D. H. & Hay, D. I. (1986) Int. J. Pept. Protein Res. **27**, 373–379

Schlesinger, D. H., Jacobs, R. & Hay, D. I. (1977) Pept. Proc. Am. Pept. Symp. 5th 56–58

Schluckebier, S. K., Hay, D. I., Schlesinger, D. H., Kresak, M. & Moreno, E. C. (1986) J. Dent. Res. **65**, 356 (abstr. 1682)

Smith, Q. T., Shapiro, B. L. & Hamilton, M. J. (1975) Arch. Oral Biol. **20**, 369–373

Sung, M. & Smithies, O. (1969) Biopolymers **7**, 39–58

Wong, R. S. & Bennick, A. (1980) J. Biol. Chem. **255**, 5943–5948

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

Wong, R. S., Madapallimattam, G. & Bennick, A. (1983) Biochem. J. **211**, 35–44

Yu, P. L., Bixler, D., Goodman, P. A., Azen, E. A. & Karn, R. C. (1986) Genet. Epidemiol. **3**, 147–152

Received 26 February 1988/27 April 1988; accepted 4 May 1988