Biochemistry

 Volume 30, Number 14 April 9, 1991

Articles

Basic Proline-Rich Proteins from Human Parotid Saliva: Relationships of the Covalent Structures of Ten Proteins from a Single Individual†

Dorothy L. Kauffman,‡ Anders Bennick,*,§ Max Blum,§ and Patricia J. Keller‡

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

Received August 10, 1990; Revised Manuscript Received December 27, 1990

ABSTRACT: Eleven basic proline-rich proteins were purified from the parotid saliva of a single individual. The complete amino acid sequences of six of these were determined by conventional protein sequence methodology, bringing to nine the number of known primary structures of nonglycosylated basic proline-rich proteins from the same individual. The partial sequence of one additional protein is also reported. All of the basic proline-rich proteins studied contain segments with identical or very similar sequences, but with two possible exceptions, none of the proteins is derived from another secreted proline-rich protein. The amino acid sequences of nine nonglycosylated basic proline-rich proteins were compared with primary structures deduced from published nucleotide sequences of DNA coding for human parotid proline-rich proteins. The sequences align well, in general, but differences also exist pointing to the complexity of the genetics of these proteins. Seven secretory basic proline-rich proteins appear to be formed from three larger precursors by selective posttranslational proteolyses of arginyl bonds. One of the basic proline-rich proteins appears to derive from human acidic proline-rich proteins. The remaining two proteins studied do not conform to any DNA structure as yet reported. Two of the basic proline-rich proteins studied are phosphoproteins and exhibit abilities to inhibit hydroxyapatite formation in vitro.

Human salivary proline-rich proteins comprise a family of proteins with unusual chemical structures and, in some cases, functional characteristics uniquely suited to the oral cavity. More than 20 such proteins have been isolated from human parotid saliva, which are currently categorized according to their charge characteristics (acidic or basic) and whether or not they contain covalently linked carbohydrate or phosphate groups.

Nine basic proline-rich proteins (BPRP's), containing up to 42 mol % proline and little or no carbohydrate, were isolated from human parotid saliva collected from a mixed population of donors (Levine & Keller, 1977) and from a single individual (Kauffman & Keller, 1979). These proteins constitute approximately 23% of the protein present in human parotid saliva (Kauffman & Keller, 1979). We reported previously the complete amino acid sequences of three parotid BPRP's from a single individual (Kauffman et al., 1982, 1986b). This paper expands our knowledge of the primary structures of salivary proline-rich proteins, all from the same individual, to include the remaining nonglycosylated BPRP's.

Bennick (1982) and Minaguchi and Bennick (1989) have reviewed the extensive chemical and genetic studies of the proline-rich proteins present in human parotid saliva, as well as their perceived functions. Activities of the acidic proline-rich proteins (APRP's) include the binding of ionic calcium (Bennick et al., 1981), strong adsorption to hydroxyapatite surfaces (Hay, 1973), and the stabilization of calcium phosphate salts in saliva (Moreno et al., 1979), all of which support the proposition that APRP's contribute to calcium–hydroxyapatite homeostatis in the mouth (Bennick, 1982). APRP's are also capable of mediating the binding of microorganisms to tooth surfaces (Gibbons et al., 1988), suggesting a possible role in the formation of dental plaque, the immediate cause of dental caries and periodontal disease. Basic proline-rich glycoproteins bind oral bacteria (Bergey, et al., 1986) and exhibit masticatory-lubricating properties (Hatton, et al., 1985).

† This work was supported by National Institutes of Health Grant DE05251-10 and Canada MRC Grant MT 4920.

‡ University of Washington.

§ University of Toronto.

0006-2960/91/0430-3351$02.50/0

Bailleul et al. (1977) have reported that human bronchial secretions contain basic proline-rich proteins with properties similar to those of the nonglycosylated proline-rich proteins of human parotid saliva. Using antibodies to salivary PRP's, Warner and Azen (1984) demonstrated immunoreactivity with serous cells of tracheobronchial glands and cultured tracheobronchial explants. Both Bailleul et al. (1977) and Warner and Azen (1984) hypothesize that PRP binding to mucins could modify the rheological characteristics of mucins.

Human salivary BPRP's are synthesized primarily by parotid glands (Robinson et al., 1989) and possibly by von Ebner's glands as well (Azen et al., 1990). Whereas both of these glands produce serous secretions, human salivary mucins are produced by submandibular-sublingual glands (Baig et al., 1973) and a variety of minor salivary glands (Milne & Dawes, 1973; Hensten-Pettersen, 1979). Interactions between the components of the serous and mucous secretions within the oral cavity seem probable but to date have not been studied systematically.

Materials and Methods

Clostripain, trypsin (TPCK[1] treated), and carboxypeptidase B (PMSF treated) were obtained from Cooper-Worthington. Pyroglutamate aminopeptidase and carboxypeptidase B-DFP were from Sigma, and elastase suspension, 2× crystallized, was from Mann Research Laboratories. Bio-Gels were obtained from Bio-Rad and Sephadexes from Pharmacia. Carboxypeptidase Y and *Staphylococcus aureus* V8 protease were obtained from Pierce. Dithiothreitol was obtained from Calbiochem-Behring Corp. Dicalcium phosphodihydrate, $CaHPO_4 \cdot 2H_2O$ (DCPD), was a gift from Dr. D. I. Hay.

Source of Parotid Saliva. Parotid saliva was collected from a single individual (D.L.K.) as previously described (Kauffman & Keller, 1979). Genetic typing of the saliva was done by Dr. K. Minaguchi, Tokyo Dental College, Japan. The saliva was typed for basic proline-rich proteins as follows: PmF^+, PmS^+, Ps^0, Gl^1, $Con1^+$, $Con2^-$, Pe^+, Pmo 1^+, Pmo 2^-, Po^+.

Isolation of Salivary Proteins. The procedures leading to fractions IB-1–IB-9 are essentially the same as those described previously (Kauffman & Keller, 1979). Fractions originally designated IB-2 and IB-3, representing less than 1% of the total BPRP, were not studied further. To eliminate minor contamination, fractions IB-4, IB-5, IB-7, and IB-8 were purified further on Bio-Gel P-10 (200-400 mesh) gel filtration columns with 0.05 M ammonium bicarbonate as the eluant. As shown in the inset of Figure 1, Bio-Gel filtration of fraction IB-8 yielded two major fractions (Figure 1a, supplementary material). The second peak from the column was subsequently resolved into two fractions, 8b and 8c, using a SP-Sephadex column (1.5 × 85 cm), equilibrated with 0.05 M sodium borate, pH 9. After sample application, equilibrating buffer was run through the column for 6 h (14 mL/h), followed by a gradient of 200 mL of buffer and 200 mL of buffer containing 0.3 M NaCl (Figure 1b, supplementary material). Proteins II-1 and II-2 were separated on the Bio-Gel P-10 (200–400 mesh) column (Figure 2, supplementary material).

Protease Digestions. Clostripain digestions were carried out in 0.05 M ammonium bicarbonate, pH 8, containing 2.5 mM dithiothreitol, at 37 °C for 16 h. The E:S ratio was 1:50 by weight. Elastase digestion was performed in 0.1 M *N*-ethylmorpholine, pH 8.5, for 17 h at 37 °C (molar E:S ratio, 1:50). TPCK–trypsin digestions were done in 0.1 M *N*-ethylmorpholine, pH 8.5, for 17 h at 37 °C (molar 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). Papain digestions were performed in 0.05 M pyridineacetate, pH 6.5, containing 2.5 mM dithiothreitol for 2 h at 37 °C. The enzyme:substrate ratio was 1:100 on a molar basis. Protein II-2 was subjected to digestion with pyroglutamate aminopeptidase, as previously described (Kauffman et al., 1986b) using 0.05 M potassium phosphate, pH 7.4, containing 5 mM EDTA and 5 mM dithiothreitol, for 16 h at 37 °C. The enzyme:substrate ratio was 1:20 by weight. 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 pyridineacetate, pH 6, at 37 °C for 7 h (E:S ratio, 1:20 by weight). The enzyme was inactivated by heating at 100 °C for 5 min.

Identification of Phosphoserine. A limited acid hydrolysis of protein II-2 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.

Phosphorus Analysis. The quantification of phosphoserine was performed by analysis for P_i produced by alkaline hydrolysis (Martinsen, 1984). The protein (0.16 mg) was dephosphorylated in 0.3 mL of 1 N NaOH at 37 °C for 21.5 h. After dephosphorylation 0.3 mL of 1 N HCl was added to the hydrolysis. Samples of 0.18 and 0.36 mL were taken for analysis by the method of Kallner (1975). Peptides were analyzed in the same manner.

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

Sequence Analysis. Automated Edman degradation of the proteins and peptides was performed on a gas-phase protein sequenator (Model 470) from Applied Biosystems. Typically, 5–10-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. The HPLC method of Moser and Rickli (1979) was used for identification of the PTH-amino acids.

Effects on Hydroxyapatite Formation. The ability of the BPRP's IB-1 and II-2 to inhibit the seeded precipitation of hydroxyapatite on the surface of DCPD crystals was studied as described by Hay et al. (1979). The phosphate concentration in the supernatant of the suspension of DCPD crystals, determined according to Lowry and Lopez (1946), was used as a measure of inhibition. The initial phosphate concentration due to dissolution of DCPD to the point of saturation was 2.5 mM. It increased to 6.5 mM after 10–20 h because of hydrolysis of DCPD and formation of the more basic hydroxyapatite. If precipitation is inhibited, the concentration of phosphate in solution remains at 2.5 mM. Samples included buffer without protein and 0.001 μM polyaspartate which causes 100% inhibition (Hay et al., 1979). To compare the activities in different samples, the amount of protein necessary for 50% inhibition (HA_{50}) was calculated.

Results

Major stages in BPRP purification and the system of naming individual fractions are outlined in Figure 1. Fractions originally designated IB-2 and IB-3, representing less than 1% of the total BPRP, were not studied further. With this exclusion and the addition of three proteins purified from IB-8, the nine nonglycosylated BPRP's derived from peak IB are

[1] Abbreviations TPCK, L-1-(tosylamino)-2-phenylethyl chloromethyl ketone; PMSF, phenylmethanesulfonyl fluoride; HPLC, high-performance liquid chromatography; PTH, phenylthiohydantoin; EDTA, ethylenediaminetetraacetic acid; P_i, inorganic phosphate.

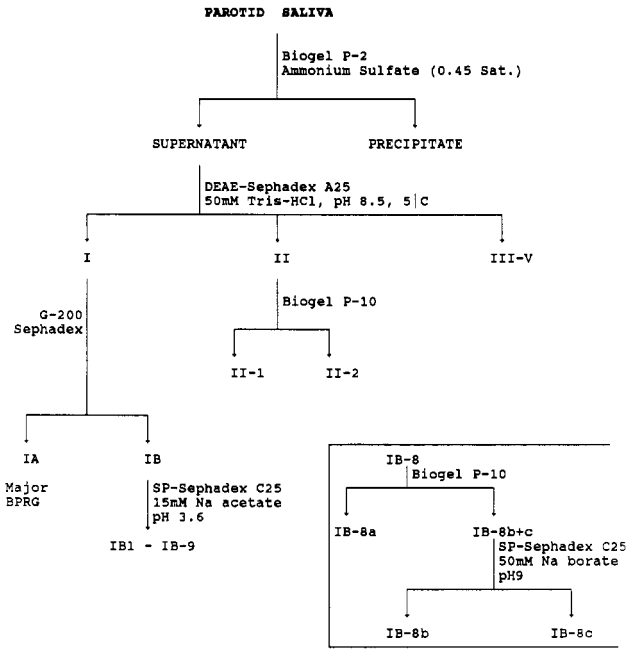

FIGURE 1: Purification scheme used for the isolation of the basic proline-rich proteins. The inset details the further purification of fraction IB-8.

1 16
S P P G K P Q G P P P Q G G N Q P Q G P P
22 37
P P P G K P Q G P P P Q G G N K P Q G P P
43 59
P P G K P Q G P P P Q G *D N K S R*

FIGURE 2: Complete amino acid sequence of salivary proline-rich protein IB-7. The residues in bold letters represent the repeat sequence described in the text.

now designated IB-1, IB-4–IB-7, IB-8a, -b, and -c, and IB-9. Two additional proteins, II-1 and II-2, represent proteins less basic than those derived from peak IB. This paper reports the complete sequences of IB-4, IB-5, IB-7, IB-8b, IB-8c, and II-2 and the partial sequence of IB-8a. The sequence of II-1, a glycosylated and phosphorylated BPRP, is currently being determined.

All human parotid BPRP's studied to date exhibit regions of structural commonality as well as regions of individual variation. To demonstrate this, we will describe in detail the amino acid sequence of a protein (IB-7) representative of the major shared features and will then relate the sequences of other BPRP's to this protein.

The total amino acid sequence of protein IB-7 is shown in Figure 2. The protein contains 59 amino acid residues, 25 of which are proline residues. The proline residues occur singly and in di-, tri-, and pentaprolyl units. The hexapeptide sequence PQGPPP occurs five times in IB-7, as residues 6–11, 17–22, 27–32, 38–43, and 47–52, respectively.

Protein IB-7 contains two contiguous 21-residue segments (1–21 and 22–42) with identical sequences, except in the case of residues 1 with 22 and 16 with 37. These segments are followed by a 12 amino acid segment (43–54) identical in sequence to residues 2–13 and 23–34, respectively, of the first and second segments. The contiguous repeat pattern exhibited in residues 1–54 of IB-7 recurs in other human parotid BPRP's and will be referred to in this paper as the repeat sequence. In IB-7 the repeat sequence is followed by the C-terminal pentapeptide, DNKSR.

Figure 3 presents the sequences of seven BPRP's isolated from human parotid saliva of a single individual (D.L.K.). They are arranged with intent to call attention to the recurring repeat sequence, as well as the range of dissimilarities observed in human parotid BPRP's. For completeness, the comparison includes three BPRP sequences previously reported by us (IB-1, IB-6, and IB-9) along with four not previously reported (IB-7, IB-8c, IB-8a, and II-2).

Figure 3a shows the sequences of IB-7, IB-8c, and IB-9, which are the most similar BPRP's in size and sequence. Protein IB-8c is a 61 amino acid protein containing the entire

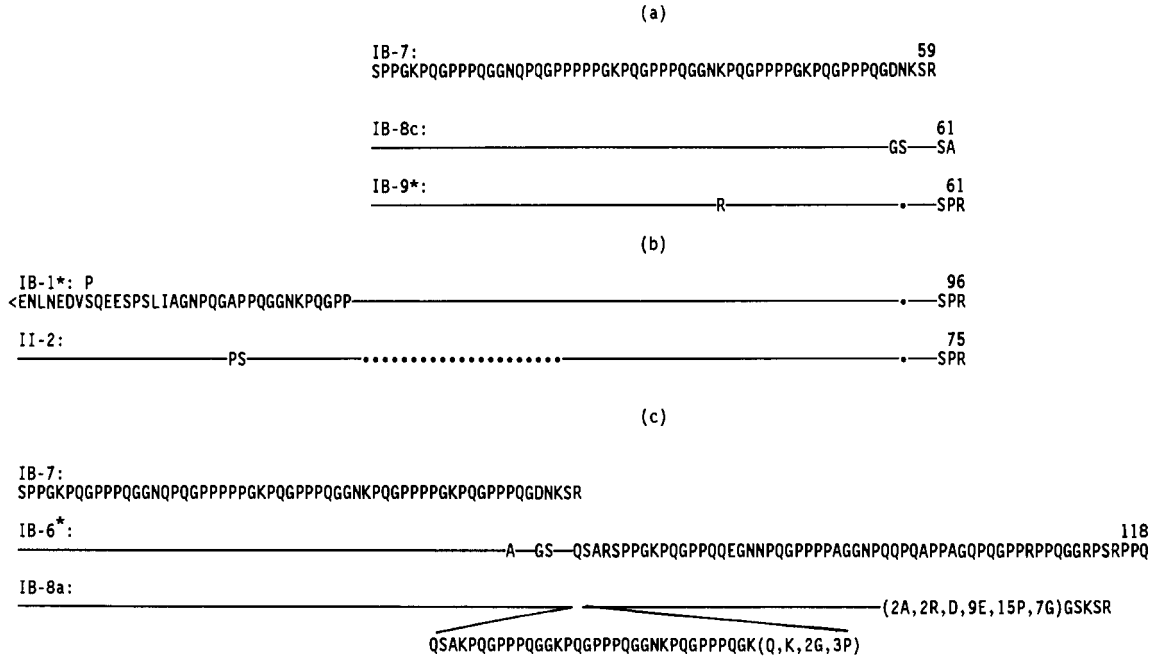

FIGURE 3: Comparison of amino acid sequences of the basic proline-rich proteins IB-7, IB-8c, IB-9, IB-1, II-2, IB-6, and IB-8a, isolated from the parotid saliva of a single individual. (a) Complete sequence of IB-7 and two BPRP's exhibiting the closest sequence homology and size. (b) Complete sequence of two BPRP's with major amino-terminal extensions to the repeat sequence of IB-7. (c) Sequences of two BPRP's with major carboxyl-terminal extensions to the repeat sequence of IB-7. Brackets indicate unknown sequences. An internal sequence in IB-8a was offset to emphasize the homology with IB-6. Protein sequences previously published are signified by an asterisk; lines (—) indicate sequence identity; dots (···) indicate amino acid deletions; the symbol (<) indicates a blocked amino-terminal residue.

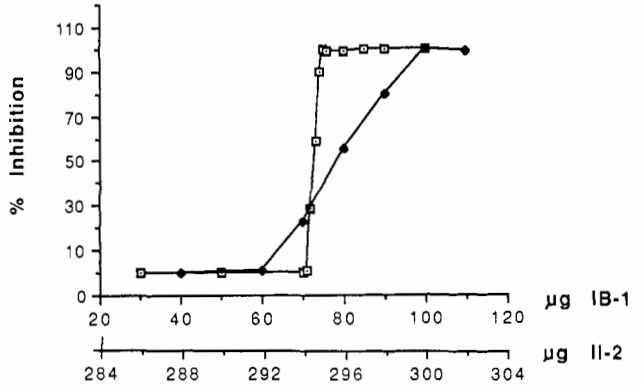

FIGURE 4: Inhibition of hydroxyapatite formation by the basic proline-rich phosphoproteins IB-1 (□) and II-2 (◆). Experimental details are presented in the text.

repeat sequence of IB-7 followed by a C-terminal septapeptide, GSKSRSA. Protein IB-9 contains 61 amino acids, comprising the entire 54 amino acid repeat sequence, with the substitution of arginine for lysine in position 37. The repeat sequence is followed by a C-terminal septapeptide, DKSRSPR.

Figure 3b shows two basic proline-rich phosphoproteins (IB-1 and II-2) that contain a common 35 amino acid N-terminal sequence. In IB-1, the amino-terminal segment is followed by the complete repeat sequence and C-terminal peptide of IB-9. Protein II-2 is identical with IB-1 except for the differences shown in Figure 3b and the fact that the first 21-residue segment of the repeat sequence has been deleted.[2] Thus IB-1 contains a total of 96 amino acids and II-2 contains 75. The amino-terminal segments of these proteins have blocked N-terminals, are less proline-rich than the repeat sequence (20 mol % vs 40 mol %), and are less basic than other BPRP's as a consequence of phosphorylation of the serine residues in position 8.

Both IB-1 and II-2 have the ability to inhibit hydroxyapatite formation in vitro. From Figure 4 it can be calculated that the amount needed for 50% inhibition (HA_{50}) is 73 µg for IB-1 and 296 µg for II-2. Acidic proline-rich proteins (APRP's) have been shown to have an HA_{50} of 19 µg (Madapallimattam & Bennick, 1990), indicating that the inhibitory role of BPRP is probably minor relative to that of APRP.

Figure 3c presents two BPRP's, IB-6 and IB-8a, which contain the repeat sequence of IB-7 as amino-terminal segments, with extensive additions on their carboxyl-terminal sides. The carboxyl-terminal segments of these proteins contain regions of homology with each other but show limited homology with the repeat sequence of IB-7 and other BPRP's. The major differences between IB-6 and IB-8a are in the sequences between the amino and carboxyl segments. In IB-6 the intervening segment is the septapeptide GSKSQSA while in IB-8a the intervening segment contains more than 42 residues in which the octapeptide PQGPPPQG occurs three or more times.

The primary structures of the three remaining nonglycosylated BPRP's isolated from human parotid saliva (D.L.K.) are shown in Figure 5. These proteins contain numerous examples of reiterated segments but do not exhibit the extended repeat sequence observed in the proteins discussed heretofore in this report. Protein IB-8b is a 44-residue polypeptide that corresponds in sequence to the C-terminal (TZ) segment of the major human parotid *acidic* PRP-C (Wong & Bennick, 1980). IB-4 is a 56 amino acid polypeptide

```
                                                                    44
IB-8b GRPQGPPPQQGGHQQGPPPPPPGKPQGPPPQGGRPQGPPQGQSPQ

                                                                    56
IB-4  SPPGKPQGPPQQEGNNPQGPPPPAGGNPQQPQAPPAGQPQGPPRPPQGGRPSRPPQ

                                                                    70
IB-5  ————————————K————GKPQGPPPPGGNPQQPQAPPAGKPQGPPPPPQGGRPPRPAQGQQPPQ
```

FIGURE 5: Amino acid sequences of basic proline-rich proteins IB-8b, IB-4, and IB-5. The line (—) indicates identity of sequence with the protein above it.

corresponding exactly to the C-terminal extension of IB-6 (residues 63–118). These are the only examples observed by us of secretory BPRP that could derive directly from other secretory BPRP. IB-5 is a 70-residue polypeptide exhibiting close homology with IB-4 for approximately 25 residues.[3]

Discussion

The primary structures of nine human parotid nonglycosylated BPRP's and the partial structure of another (IB-8a), all from a single individual, have been determined by conventional protein sequence methodology as reported in this paper and previous papers (Kauffman et al., 1982, 1986b). The nonglycosylated proteins all contain, to some degree, repetitive segments with identical or very similar sequences. The most similar in size and sequence are the BPRP's designated IB-7, IB-8c, and IB-9, which differ only in the substitution of arginine for lysine in position 37 of the contiguous repeat sequence and in the composition and sequence of the terminal 5–7 residues at the carboxyl end of the molecules. Modifications observed in other human parotid BPRP include major extensions at the amino terminus, as in BPRP's II-2 and IB-1 or at the carboxyl terminus, as in BPRP's IB-6 and IB-8a. Notwithstanding many similarities in their primary structures, it is apparent that, with the possible exception of IB-4 and IB-8b, none of the BPRP's isolated from human parotid saliva (D.L.K.) is derived from another secretory parotid BPRP. Nor is any secreted nonglycosylated BPRP the precursor of the glycosylated BPRP (Kauffman & Keller, 1983).

Extensive studies of inheritance patterns of human parotid BPRP over the past several decades (Azen & Maeda, 1988) led to a postulate of 13 genes with null alleles. This number was reduced by Maeda (1985) to six, comprising two subfamilies: an *Hae*III type (PRH 1 and 2) encoding APRP and a *Bst*N1 type (PRB 1–4) encoding BPRP. Maeda et al. (1985) subsequently constructed a cDNA library from the total poly(A+) RNA of a single human parotid gland, from which cDNAs coding for PRP were isolated and sequenced. They reported that clones cP_3 to cP_7 encoded precursors for basic and glycosylated PRP and that the precursors contained similar signal (S) and N-terminal segments (N), variable repetitive regions in which *Bst*N1 sites occur repeatedly ($B_1B_2B_3$), and C-terminal regions (CB). Although the amino acid sequences of the BPRP's shown in Figure 3 agree, in general, with the $B_1B_2B_3$ repeat pattern of cDNA sequences described by Maeda et al. (1985), a notable exception is IB-1, for which the protein sequence agrees better with a $B_1B_1B_2B_3$ repeat pattern. The latter repeat pattern is as yet unreported in published DNA nucleotide sequences (Maeda et al.; 1985; Lyons et al., 1988).

Comparison of the amino acid sequences reported in this paper and previous papers (Kauffman et al., 1982, 1986b) with sequences predicted from existing DNA data on human salivary BPRP (Maeda et al., 1985; Lyons et al., 1988) reveals the following relationships. BPRP's II-2, IB-9, and IB-6

[2] Contrary to an earlier abstract (Kauffman et al., 1986a), BPRP II-2 does not contain carbohydrate.

[3] Saitoh et al. (1983b,c) and Isemura et al. (1982) have reported the sequences of four BPRP's obtained from pooled parotid saliva, called P-D, P-E, P-F, and P-H. We have determined that these correspond respectively to IB-5, IB-9, IB-8c, and IB-4.

correspond exactly and in the order stated to the precursor molecule predicted by the PRB1^s loci reported by Lyons et al. (1988) for PmS$^+$,PmF$^+$ individuals, except for one amino acid difference in IB-6. The truncated clone (cP_7), reported by Maeda et al. (1985), contains nucleotide sequences corresponding to BPRP IB-7 (nucleotides 42–218), except for the C-terminal amino acid, as well as an unidentified protein (nucleotides (219–413) and BPRP's IB-8c (nucleotides 414–596) and IB-4 (600–767). The agreement of the two latter proteins with the cDNA sequences is exact. Clone cP_6 (Maeda et al., 1985) contains the nucleotide sequences (565–774) corresponding to BPRP IB-5, except for residue 32, which is proline in the protein and alanine in the cDNA. Thus, five BPRP's correspond exactly to the DNA sequences cited and two BPRP's differ in one amino acid only. It should be noted that the basic and glycosylated PRP are encoded by four genes (Maeda et al., 1985), all of which have several alleles (Lyons et al., 1988). It is therefore not surprising that there are differences between the results obtained by conventional amino acid sequencing of the proteins and those obtained by deduction of sequences from cDNA of unrelated individuals. Determination of the amino acid sequences of all the secreted PRP's as well as the nucleotide sequences encoding these proteins, all from the same individual, is in progress and will provide a unique opportunity to assign BPRP genes unequivocably.

No DNA sequences encoding IB-8a, IB-1, or IB-8b were found on any of the clones for BPRP sequenced by Maeda et al. (1985) or Lyons et al. (1988). However, in a recent paper (Robinson et al., 1989) we reported that a 44 amino acid polypeptide identical with IB-8b is the major nonglycosylated BPRP in human submandibular-sublingual saliva and that mRNA studies indicate that this protein arises by proteolytic cleavage of APRP's. It seems reasonable to conclude that human parotid BPRP IB-8b also derives from APRP's or their precursors. IB-8a and IB-1 remain unaccounted for in the published data of Maeda et al. (1985) and Lyons et al. (1988).

With the exception of those BPRP's ascribed to C-terminal regions of the cDNA clones, all except IB-8c have arginine in the C-terminal position, consistent with the postulate that the multiplicity of secretory BPRP in human parotid saliva results, in part at least, from posttranslational proteolytic cleavage of a limited number of precursor molecules. The finding of alanine as the C-terminal in IB-8c is puzzling. Whereas alanine has been observed consistently by us and is reported by Saitoh et al. (1983a), the nucleotide sequences of Maeda et al. (1985) predict an Ala-Arg sequence in this position of the precursor. The existence of IB-8c and IB-4 as separate secretory products is by itself presumptive evidence of proteolytic cleavage at this site of the precursor. The mechanism by which the putative Arg residue is removed from IB-8c remains a mystery although tryptic-like cleavage, followed by carboxypeptidase B like removal of the basic residue, has been observed in other processing systems (Swartz, 1986). The sequences surrounding the cleaved arginines are as follows: SP$\overline{R}$SPP in three sites, SA$\overline{R}$SPP in two sites, and KS$\overline{R}$SPP in one site. The consistent presence of SPP immediately following the cleaved arginine on the carboxyl side is noteworthy, as is the presence of proline immediately preceding the arginine in three instances. The latter qualify as examples of proline-directed arginine cleavage described by Schwartz (1986). It is interesting to note that BPRP IB-6 contains an uncleaved SA$\overline{R}$SPP sequence (residues 60–65).

In addition to the posttranslational modifications such as phosphorylation, glycosylation, and proteolysis which occur during intracellular processing of salivary PRP's, other mechanisms by which a disproportionate number of proteins might arise from a relatively small number of genetic loci include (1) differential RNA splicing (Maeda et al., 1985) and (2) allelic gene-length variants (Lyons et al., 1988). The findings reported in this paper include an example of allelic gene-length variation but do not provide examples of differential RNA splicing.

The putative precursor molecules have not yet been observed in human parotid saliva despite the fact that their predicted sizes (ca. 30 000) are within the range suitable for chromatographic or electrophoretic resolution by the procedures used. Furthermore, no incomplete proteolysis products have been reported. The sequences of the secreted proteins reported in this paper together with determination of the corresponding DNA sequences from the donor of the sequenced proteins will make it possible to locate exactly the cleavage sites in the precursor proteins and thereby facilitate studies on postribosomal cleavages of the proteins.

The reason for the large number of isoforms of PRP is not entirely clear, although the possibility exists that there is no need to maintain strong genetic control over the structure of the proline-rich proteins. For example, proline-rich proteins show a high affinity for tannins, leading to the suggestion that they protect the digestive system from the deleterious effects of dietary tannins (Hagerman & Butler, 1981; Manshansho et al., 1987). The ability to bind tannin is shared with other proteins and polymers with high proline content such as collagen (Hagerman & Butler, 1981). Since salivary PRP and collagen have quite different primary structures, it would appear that there is no strong amino acid sequence dependence on tannin binding to protein.

Supplementary Material Available

Summary of the proof of protein sequences; three tables, giving the amino acid compositions of BPRP's IB-4, IB-5, IB-7, IB-8a, IB-8b, IB-8c, and II-2 (Table I), sequence analyses of these proteins and peptides therefrom (Table II), and quantitative CP-B digestion of IB-8a (Table III); and two figures, showing the chromatographic purification of proteins from IB-8 (Figure 1) and from fraction II (Figure 2) (11 pages). Ordering information is given on any current masthead page.

References

Azen, E. A., & Maeda, N. (1988) *Adv. Hum. Genet. 17*, 141–199.

Azen, E. A., Hellekant, L. M., Sabatini, L. N., & Warner, T. F. (1990) *J. Dent. Res. 69*, 1724–1730.

Baig, M. M., Winzler, R. D., & Renner, O. M. (1973) *J. Immunol. 111*, 1826–1833.

Bailleul, V., Ricket, C., Hayen, A., & Degand, P. (1977) *Clin. Chim. Acta 74*, 115–123.

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

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

Bergey, E. J., Levine, M. J., Reddy, M. S., & Bradway, S. D. (1986) *Biochem. J. 234*, 43–48.

Gibbons, R. J., Hay, D. I., Cisar, J. O., & Clark, W. B. (1988) *Infect. Immun. 56*, 2990–2993.

Hagerman, A. E., & Butler, L. G. (1981) *J. Biol. Chem. 256*, 4494–4497.

Hatton, M. N., Loomis, R. E., Levine, M. J., & Tabak, L. A. (1985) *Biochem. J. 230*, 817–820.

Hay, D. I. (1973) *Arch. Oral Biol. 18*, 1517–1529.

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

Hensten-Pettersen, A. (1979) in *Saliva and Dental Caries, Proceedings of a Workshop* (Kleinberg, I., Ellison, S. A., & Mandel, I. D., Eds.) Special Suppl. to Microbiology Abstracts, pp 81-87, IRL, Oxford.
Isemura, S., Saitoh, E., & Sanada, K. (1982) *J. Biochem. (Tokyo) 91*, 2067-2075.
Kallner, A. (1975) *Clin. Chim. Acta. 59*, 35-39.
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 2*, 6558-6562.
Kauffman, D. L., Bennick, A., & Keller, P. J. (1986a) *J. Dent. Res. 65*, Special Issue, Abstract 86.
Kauffman, D., Hoffman, T., Bennick, A., & Keller, P. (1986b) *Biochemistry 25*, 2387-2392.
Levine, M., & Keller, P. J. (1977) *Arch. Oral Biol. 22*, 37-41.
Lowry, O. H., & Lopez, J. A. (1946) *J. Biol. Chem. 162*, 421-428.
Lyons, K. M., Stein, J. H., & Smithies, O. (1988) *Genetics 120*, 267-278.
Madapallimattan, G., & Bennick, A. (1990) *Biochem. J. 270*, 297-304.
Maeda, N. (1985) *Biochem. Genet. 23*, 455-464.
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.
Menshansho, H., Butler, L. G., & Carlson, D. M. (1987) *Annu. Rev. Nutr. 7*, 423-440.
Milne, E. J. & Dawes, C. (1973) *Vox Sang. 25*, 298-307.
Minaguchi, K., & Bennick, A. (1989) *J. Dent. Res. 68*, 2-15.
Moreno, E. L., Varaghese, K., & Hay, D. I. (1979) *Calcif. Tissue Int. 28*, 7-16.
Moser, P. W., & Rickli, E. E. (1979) *J. Chromatogr. 176*, 451-455.
Robinson, R., Kauffman, D. L., Waye, M. M. Y., Blum, M., Bennick, A., & Keller, P. J. (1989) *Biochem. J. 263*, 497-503.
Saitoh, E., Isemura, S., & Sanada, K. (1983a) *J. Biochem. (Tokyo) 93*, 883-888.
Saitoh, E., Isemura, S., & Sanada, K. (1983b) *J. Biochem. (Tokyo) 93*, 495-502.
Saitoh, E., Isemura, S., & Sanada, K. (1983c) *J. Biochem. (Tokyo) 94*, 1991-1999.
Schwartz, T. W. (1986) *FEBS Lett. 200*, 1-10.
Warner, T. F., & Azen, E. A. (1984) *Am. Rev. Respir. Dis. 130*, 115-118.
Wong, R. S. C., & Bennick, A. (1980) *J. Biol. Chem. 255*, 5943-5948.

^{1}H Assignments and Secondary Structure Determination of the Soybean Trypsin/Chymotrypsin Bowman-Birk Inhibitor†

Milton H. Werner‡ and David E. Wemmer*,§

Chemical Biodynamics Division, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, Berkeley, California 94720

Received September 28, 1990; Revised Manuscript Received January 2, 1991

ABSTRACT: The ^{1}H resonance assignments and secondary structure of the trypsin/chymotrypsin Bowman-Birk inhibitor from soybeans were determined by nuclear magnetic resonance spectroscopy (NMR) at 600 MHz in an 18% acetonitrile-d_3/aqueous cosolvent. Resonances from 69 of 71 amino acids were assigned sequence specifically. Residues Q11-T15 form an antiparallel β-sheet with residues Q21-S25 in the tryptic inhibitory domain and an analogous region of antiparallel sheet forms between residues S38-A42 and Q48-V52 in the chymotryptic inhibitory domain. The inhibitory sites of each fragment (K16-S17 for trypsin, L43-S44 for chymotrypsin) are each part of a type VI like turn at one end of their respective region of the antiparallel β-sheet. These structural elements are compared to those found in other Bowman-Birk inhibitors.

Bowman-Birk inhibitor (BBI) is a small serine protease inhibitor containing seven disulfide bonds (Birk, 1985). Isoforms of the Bowman-Birk type are found in a wide variety of leguminous plants with up to 87% sequence homology among different species (Morhy & Ventura, 1987), thus forming a distinct family of serine protease inhibitors in plants (Steiner & Frattali, 1969). BBI is an intriguing protein among the serine protease inhibitors since it is the only inhibitor capable of binding 2 equiv of serine protease/equiv of inhibitor with kinetically independent binding sites (Harry & Steiner, 1970; Odani & Ikenaka, 1972).

Classical BBI from soybeans simultaneously inhibits trypsin and chymotrypsin, although several isoinhibitors have been isolated which are capable of inhibiting either two trypsin molecules or trypsin at one site and either chymotrypsin or elastase at a second site (Odani & Ikenaka, 1977). Each is characterized by 7 disulfide bonds, an isoelectric point near 4.2, and between 60 and 76 amino acids. The trypsin/chymotrypsin inhibitor (hereafter referred to as BBI-I) may be

† This work was supported by the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the U.S. Department of Energy, under contract DE-AC03-76 SF 00098 and through instrumentation grants from the U.S. Department of Energy (DE FG05-86 ER 75281) and the National Science Foundation (DMB 86-09035 and BBS 87-20134).

* To whom correspondence should be addressed.
‡ Chemical Biodynamics Division.
§ Department of Chemistry.

0006-2960/91/0430-3356$02.50/0