Biochem. J. (1975) **145**, 557–567
Printed in Great Britain

Chemical and Physical Characteristics of a Phosphoprotein from Human Parotid Saliva

By ANDERS BENNICK
Department of Biochemistry and Faculty of Dentistry, University of Toronto, Toronto M5S 1A8, Canada

(*Received* 9 *September* 1974)

The isolation of a highly purified phosphoprotein, previously named protein A, from human parotid saliva is described. This protein has an unusually high amount of glycine, proline and dicarboxylic amino acids. Together these amino acids account for 80% of all residues. The protein contains 1.9mol of P/mol of protein, probably as phosphate in an ester linkage to serine, and about 0.5% carbohydrate, but no hexosamine. The *N*-terminal is blocked and the following *C*-terminal sequence is proposed: -Ala-Asp-Ser-Gln-Gly-Arg-Arg. The isoelectric point is 4.43. The molecular weight of the protein determined by ultracentrifugation is 9900 and from chemical analyses 11000. Circular-dichroism and nuclear-magnetic-resonance spectra indicate the absence of polyproline and triple-helical-collagen-like structure for the protein. There is little restriction on the orientation of the single phenylalanine residue in the protein, but there is also an indication of conformational restraint in the protein.

Studies on human salivary secretions have demonstrated the presence of several proteins which all have an unusually high content of dicarboxylic amino acids, proline and glycine. Together these amino acids account for 70–85% of all amino acid residues in the proteins (Bennick & Connell, 1971; Oppenheim *et al.*, 1971; Armstrong, 1971; Levine *et al.*, 1973). The biological significance of these salivary proteins is not known. A number of biological activities that can be demonstrated in salivary secretions cannot be attributed to these proteins (Bennick & Connell, 1971) and their composition is distinct from that of structural proteins such as collagen (Fleischmajer & Fishman, 1965), elastin (Anwar, 1965) and enamel proteins (Eggert *et al.*, 1973). It has been demonstrated that some of these salivary proteins adsorb to hydroxyapapite (Hay, 1973) and this may be an important factor in the biological interactions which occur at the tooth surface. Another of these proteins has been demonstrated to be the protein core precursor of a glycoprotein, decreasing in concentration in saliva secreted at high flow rates (Levine *et al.*, 1973).

A genetic relationship between some of these proteins has been established (Azen & Oppenheim, 1973) and it is interesting that proteins have been demonstrated in rat salivary secretions which have a composition similar to the human proteins (Muenzer & Carlson, 1974).

The amount of these proteins in the parotid saliva is difficult to ascertain because of lack of specific assays. Some of these proteins correspond to heavily stained bands which can be seen after staining of polyacrylamide gels on which saliva has been electrophoresed, and at least some of these proteins appear to be present in significant amounts in saliva.

It appears from these investigations that saliva contains a number of unique proteins which may be chemically related to each other and which may have biological functions in the oral cavity or elsewhere, although it is possible that the primary function of these proteins is associated with the secretory mechanism of the gland. It was therefore decided to investigate further the function and physical and chemical properties of four proteins named A, B, C and D, which have previously been isolated from human parotid saliva (Bennick & Connell, 1971). The present paper describes chemical and physical properties of protein A, which has been found to be a calcium-binding protein (A. Bennick, unpublished work).

Experimental

Materials

Sephadex chromatographic resins were obtained from Pharmacia, Stockholm, Sweden. CM-cellulose and DEAE-cellulose (Whatman 32) were obtained from W. and R. Balston, Maidstone, Kent, U.K. Ultrafiltration membranes (UM2) were purchased from Amicon, Lexington, Mass., U.S.A. Bovine serum albumin was from Armour Pharmaceutical Co., Kankakee, Ill., U.S.A., and acrylamide, *NN*-methylenebisacrylamide and *NNN'N'*-tetramethylenedi-

amine were from Eastman–Kodak, Rochester, N.Y., U.S.A. Coomassie Brilliant Blue (Michrome no. 137) was obtained from E. Gurr, London S.W.14, U.K., and Bromophenol Blue was from BDH Chemicals, Poole, Dorset, U.K. Ampholine (pH3–6) was purchased from LKB Produkter, Stockholm, Sweden, and polyamide layer plates were from the Cheng Chin Trading Co., Taipei, Taiwan. Penicillo-carboxy-peptidase-S1 was a gift from Dr. T. Hofmann, Toronto, and ^2H_2O was purchased from Bio-Rad, Richmond, Calif., U.S.A. All other chemicals were analytical-grade reagents.

Collection and concentration of saliva and chromatographic procedures

Saliva was collected as described previously from one donor in 50ml volumes (Bennick & Connell, 1971). The method of saliva collection was approved by the committee on human experimentation at the University of Toronto and was performed with the consent of the donor. A 0.2M solution of EDTA, pH7.0, was added to the saliva to give a final concentration of 0.01M-EDTA. The saliva samples were quick-frozen in liquid N_2 and stored at −10°C.

The chromatographic resins were treated according to the manufacturer's instructions. Saliva was desalted in 100ml volumes on a Sephadex G-15 column (4.5cm×35cm). The desalted saliva was freeze-dried and stored at −10°C. When needed, the freeze-dried solids obtained from a total of 500ml of parotid saliva was redissolved in 0.1M-Tris–HCl buffer, pH8.0, in about one-tenth of the original volume.

Gel filtration on Sephadex G-50 was performed as described previously on a column (5cm×80cm) (Bennick & Connell, 1971). The column was developed with 0.1M-Tris–HCl buffer, pH8.0. For desalting, a column (2.9cm×49cm) of Sephadex G-25 was used. The column was developed with water. For cation-exchange chromatography, a column (2.5cm×65cm) of CM-cellulose CM-32 was used. The resin was equilibrated in sodium acetate buffer, pH4.4, $I = 0.01$, and the sample was dialysed against the same buffer before it was applied to the column. The column was developed with a linear gradient composed of 900ml of sodium acetate buffer, pH4.4, $I = 0.01$, and an equal volume of sodium acetate buffer, pH4.4, $I = 0.20$.

Anion-exchange chromatography was carried out on a column (24cm×11.5cm) of DEAE-Sephadex A-25. The resin was equilibrated in 0.05M-Tris–HCl buffer, pH8.0, containing 0.14M-NaCl, and the sample was dialysed against the same buffer before it was applied to the column. A linear gradient of 200ml of 0.14M-NaCl and 200ml of 0.25M-NaCl in 0.05M-Tris–HCl buffer, pH8.0, was used to develop the column.

All chromatographic procedures were performed at 3–5°C. The protein fractions were concentrated either by ultrafiltration through UM2 membranes or by freeze-drying solutions of protein in water. Concentrated protein fractions were quick-frozen in liquid N_2 and all protein fractions were stored at −10°C. Protein concentrations were determined by the method of Lowry *et al.* (1951), indicated by (L) in the text, with bovine serum albumin as a standard, or from the amino acid composition, indicated by (A) in the text.

Analytical polyacrylamide-gel electrophoresis

This was performed in 0.1M-Tris adjusted to pH9.2 with glycine, separating proteins migrating toward the anode, or in 0.05M-sodium formate buffer, pH2.8, separating components migrating to the cathode as described previously (Bennick & Connell, 1971). For detection of minor components in the gels the Coomassie Blue staining method described by Chrambach *et al.* (1967) was preferred. Because of the varying penetration of the dye into the different protein bands, these gels had inferior photographic qualities, and Amido Black was the preferred stain for photographic purposes.

Analytical isoelectric focusing

Samples of 40μg of protein (L) were incorporated into 12cm-long polyacrylamide gels. The pH of the Ampholines was 3–6 and the gels were polymerized with riboflavin by exposure to u.v. light. The electrolyte solutions were 1% phosphoric acid for the anode, 2% (w/v) ethylenediamine for the cathode. A potential of 100V was applied to the electrodes for 15h at room temperature (20°C). After isoelectric focusing the gels were soaked in 12.5% (w/v) trichloroacetic acid until the proteins appeared as bands of precipitated material in the gels. A blank gel was included in the experiment. After isoelectric focusing this gel was sliced into 2mm-thick slices. The slices were placed in individual test tubes containing 0.5ml of water in order to elute the Ampholines. The pH of these Ampholine solutions was determined on an extended-scale pH-meter. The isoelectric point of protein A was determined from the pH of the blank gel slice corresponding to the position of protein A in the experimental gel.

Amino acid analysis

From three different preparations of protein A were taken samples of approx. 0.1mg of protein. To each sample was added 60nmol of norleucine and 60nmol of α-aminoguanidopropionic acid. The samples were hydrolysed in constant-boiling HCl for 24h in evacuated sealed tubes at 110°C. For determination of tryptophan a volume of 20μl of thioglycollic acid was added to a fourth sample of protein before hydrolysis for 24h (Matsubara & Sasaki, 1969). The hydrolysates were analysed for amino acid

content on a Beckman model 120 C amino acid analyser. The mean content and standard deviation of each amino acid was calculated. No corrections were made for loss of threonine and serine during hydrolysis and the values for valine and isoleucine were not corrected for incomplete recoveries.

Carbohydrate analysis

The total carbohydrate content was determined by the anthrone method (Seifter *et al.*, 1950) with a glucose standard. The hexosamine concentration was determined on a sample that had been hydrolysed in the same manner as the samples used for amino acid analysis except that the time of hydrolysis was 8.75h. The hydrolysate was analysed on a Beckman model 120 C amino acid analyser.

Phosphorus analysis

This was done by the method described by Bartlett (1959). The phosphorus concentration was determined on samples that had been hydrolysed in boiling 72% (w/v) $HClO_4$ and on samples to which cold $HClO_4$ had been added in order to determine the concentration of total as well as inorganic phosphorus. In an attempt to discover if the phosphorus was bound in an ester linkage to the polypeptide, a sample of protein was subjected to mild alkaline hydrolysis in 1M-NaOH at 37°C for 4h before analysis for inorganic and total phosphorus. A further sample of protein was extracted with chloroform–methanol (2:1, v/v) and the extract and residue were analysed for total phosphorus in order to determine the presence of phospholipids.

N-Terminal determination

Samples of 0.1mg of protein (L) were dansylated, hydrolysed and subjected to t.l.c. on polyamide plates by the method of Percy & Buchwald (1972) in three different solvent systems consecutively. Standards of dansyl-amino acids were run along with the hydrolysate.

C-Terminal amino acid sequence

From a sample of protein, portions were taken for amino acid analysis and for digestion with penicillocarboxypeptidase-S1 (Jones & Hofmann, 1972). Amounts of protein varying from 24 to 47nmol were incubated with carboxypeptidase in 0.6M-pyridine adjusted with acetic acid to pH3.1 and diluted to 0.06M. An enzyme/protein molar ratio between 1:50 and 1:200 was used. Hydrolysis was allowed to proceed at 37°C at incubation periods varying from 20min to 10h. A further sample of protein was incubated at 37°C for 10h, but without added enzyme. After incubation the samples were heated on a boiling-water bath for 15min and subjected to amino acid analysis on a Beckman Spinco model 120C amino acid analyser by using a single column and a lithium salt system designed for amino acid analysis of physiological fluids (Benson *et al.*, 1967).

Determination of molecular weight

Samples of protein were subjected to sedimentation–equilibrium centrifugation by the low-speed method of Richards & Schachman (1959) on a Beckman model E analytical ultracentrifuge equipped with Rayleigh interference optics. The initial concentration of the protein solution was determined from a synthetic-boundary run. Ultracentrifugation was performed at 20°C on a sample with a concentration of 2.6mg of protein (A)/ml in a buffer containing 0.008M-NaH_2PO_4, 0.004M-Na_2HPO_4 and 0.18M-NaCl, pH7.4, at speeds varying from 21000 to 26000 rev./min. The partial specific volume was calculated from the amino acid composition.

Circular dichroism

C.d. (circular-dichroism) spectra were recorded on a Jasco ORD/CD-15 spectropolarimeter equipped with an SS-20 CD modification. Spectra of aqueous solutions of protein at pH6.5 at room temperature were recorded between 320 and 180nm. For spectra recorded between 250 and 180nm a cell with a 1mm optical path-length was used and the concentration of the protein solution was 0.1mg/ml (A). For spectra recorded between 320 and 250nm a cell with a 1cm optical path-length was used and the concentration of the protein solution was 4.55mg/ml. Spectra were also recorded of solutions of the same preparation of protein A containing 4.60mg of protein (A)/ml in 6M-guanidine hydrochloride at wavelengths from 320 to 250nm in a cell with a 1cm optical path. The absorption of the solution containing 4.45mg of protein/ml at 280nm was approx. 0.01. This may be due to light-scattering which would introduce an artifact in the c.d. spectrum. Even so it would still be valid to compare the spectra recorded in the presence and absence of guanidine hydrochloride. The results are given as mean amino acid residue ellipticity measured in degrees·cm^2·$dmol^{-1}$. A mean residue weight of 104 was calculated from the amino acid composition. The calculations of ellipticity were done as described by Dorrington & Smith (1972).

Proton n.m.r. (nuclear-magnetic-resonance) spectroscopy

N.m.r. spectra were recorded by continuous-wave spectroscopy on a Varian HR-220MHz spectrometer at the Canadian 220 MHz n.m.r. centre. A tube with 5mm diameter was used and spectroscopic scans were accumulated in a computer of average transients. About 120 scans were accumulated for each spectrum recorded.

Samples were prepared for n.m.r. spectroscopy by freeze-drying the protein and redissolving it in 99.8% ^2H_2O, repeating this procedure, then freeze-

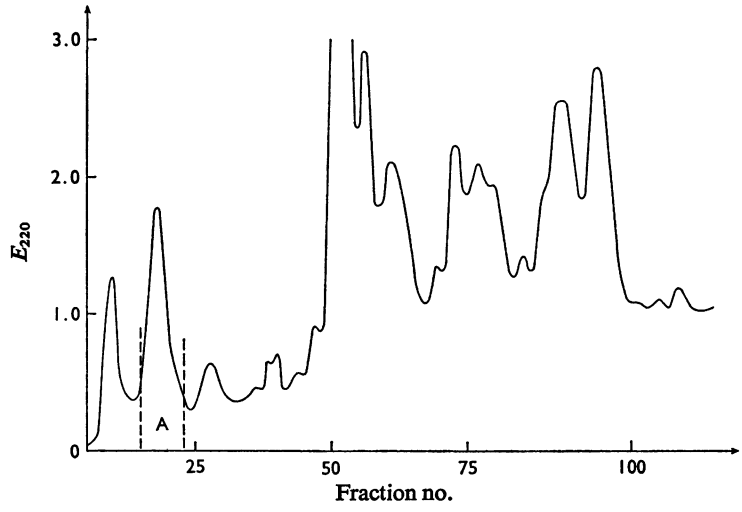

Fig. 1. *CM-cellulose chromatography of fraction I*

Fraction I [650mg of protein (L)] obtained from gel filtration of parotid saliva on Sephadex G-50 was fractionated on a CM-cellulose column (2.5cm×65cm) equilibrated in sodium acetate buffer, pH4.4, $I = 0.01$. The column was developed with a linear gradient composed of 900ml of sodium acetate buffer, pH4.4, $I = 0.01$, and an equal volume of sodium acetate buffer, pH4.4, $I = 0.2$. The flow rate was 40ml/h. After sample application, 57ml of buffer was eluted from the column before fraction collection was started. The volume of the fractions was 15ml and E_{220} was read in a cell with a 1mm optical path-length. A indicates the location of protein A in the elution diagram.

drying and redissolving the protein in 100% ^2H_2O at a concentration of 10mg of protein (A)/ml. Spectra were recorded at a probe temperature of 18°C. Spectra with a sweep width of 2500Hz with and without 3-trimethylsilyltetradeuteriopropionic acid as internal standard were recorded of solutions at pH5.6.

A spectrum of the denatured protein was computed from the amino acid composition by the method of McDonald & Philips (1969). In this calculation it was assumed that there was no glutamine and asparagine in the protein. The computed spectrum was compared with the spectrum of the native protein. All measurements on the spectra were done in Hz relative to the position of a 2,2-dimethyl-2-silapentane-5-sulphonate reference signal. The position of this reference signal relative to the 3-trimethylsilyltetradeuteriopropionic acid signal was determined experimentally.

Results

After being desalted and freeze-dried, the redissolved concentrated saliva was fractionated by gel filtration on Sephadex G-50. This fractionation gave rise to four chromatographic fractions, I, II, III and IV, as previously described (Bennick & Connell, 1971). Protein A, which was located in fraction I, was further purified by ion-exchange chromatography on CM-cellulose. The elution pattern from this column is illustrated in Fig. 1. By means of polyacrylamide-gel electrophoresis at pH9.2, protein A was located in the fraction indicated in Fig. 1. Protein A was further purified by ion-exchange chromatography on DEAE-Sephadex A-25. The elution diagram is shown in Fig. 2. Protein A was located in the fraction indicated in the diagram. Finally protein A was desalted on Sephadex G-25 and concentrated either by ultrafiltration or freeze-drying and stored at −90°C. From 100ml of unconcentrated parotid saliva 6mg of protein A was recovered as determined by the method of Lowry *et al.* (1951).

Criteria of purity

Samples of pure protein A were subjected to analytical electrophoresis at pH9.2 and 2.8 on polyacrylamide gels varying in concentration from 3% to 15% acrylamide; 3% of the total acrylamide was bisacrylamide. Gels were stained with Coomassie Blue or Amido Black. In all instances only one heavily stained band could be seen and no minor components were visible. An example of a gel containing pure protein A is illustrated in Fig. 3.

Samples of the fractions containing protein A collected from the column used in the final purification were subjected to isoelectric focusing. In all gels a single pronounced precipitated band with no traces of minor components was seen. The band had the same location in all gels, corresponding to an isoelectric pH of 4.43.

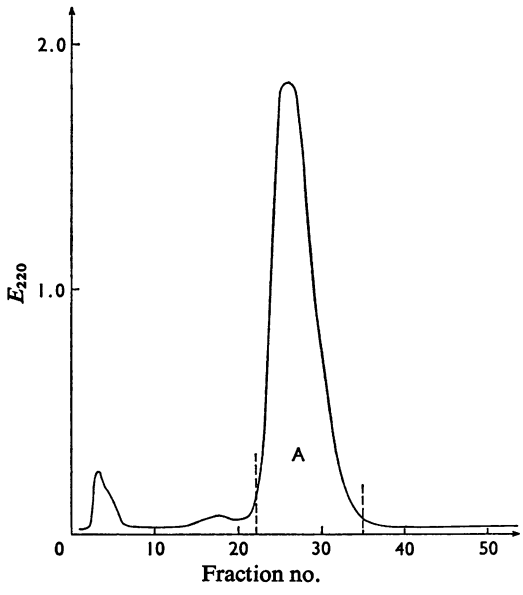

Fig. 2. *Purification of protein A by chromatography on DEAE Sephadex*

Protein A [27mg (L)] obtained by CM-cellulose chromatography of fraction I was chromatographed on a Sephadex A-25 column (2.4cm×11.5cm) equilibrated in 0.05M-Tris–HCl buffer, pH8.0, containing 0.14M-NaCl. The column was developed with a linear gradient of 200ml of the equilibrating buffer and 200ml of 0.05M-Tris–HCl, pH8.0, containing 0.25M-NaCl. The flow rate was 20ml/h. Fractions (7ml) were collected and E_{220} was read in a cell with a 1mm optical path-length. A indicates the location of protein A in the elution diagram.

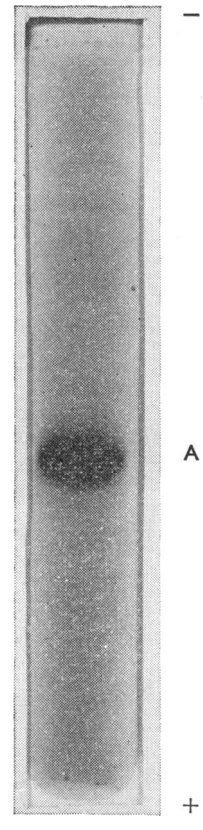

Fig. 3. *Analytical polyacrylamide-gel electrophoresis of pure protein A*

A sample of 20μg of protein A (L) was subjected to electrophoresis on a 4.5% polyacrylamide gel (0.5cm×6cm) in a continuous buffer system of 0.1M-Tris–glycine, pH9.2. The cathode was at the top end of the gel and the anode at the bottom end. A potential of 125V was applied to the electrodes for 39min. The gel was stained with Amido Black.

Amino acid composition

The result of the amino acid analysis is given in Table 1. There are no sulphur-containing amino acids in the protein and no tryptophan, tyrosine and threonine. The four most prominent amino acids are proline, glutamic acid, glycine and aspartic acid. Together these four amino acids account for 80% of all the residues in the protein. From the amino acid composition a minimum molecular weight of 10800 was calculated.

Carbohydrate analysis

The carbohydrate content of protein A as determined with a glucose standard was 0.47μg/100μg of protein (A). No hexosamines were found in protein A.

Phosphorus analyses

The results of the phosphorus assays are given in Table 2. The total phosphorus concentration was 0.55μg/100μg of protein (A) or 1.9μg-atoms of P/μmol of protein, determined as the minimum molecular weight from amino acid composition. In the same manner it can be calculated that the protein contained 1.3μg-atoms of organic P/μmol of protein. The results indicate that at least 90% of the phosphorus in the protein preparation was organic phosphorus. It is possible that some release of organically linked phosphorus occurred in samples that had not been boiled in $HClO_4$, since the assay method also includes a short heating in the presence of acid. The content of organic phosphorus may therefore be higher than these assays indicate. After NaOH hydrolysis 87% of the phosphorus could be recovered as inorganic phosphorus. It is possible that hydrolysis for a longer period than was used in these determinations might have released all organically

bound phosphorus as inorganic phosphorus. After extraction with chloroform–methanol all the phosphorus could be recovered in the residue, indicating the absence of lipid-bound phosphorus.

N-Terminal analysis

After chromatography of samples of protein A on thin-layer polyamide plates, no dansyl-amino acids could be detected. Because it is difficult to distinguish dansyl-alanine from dansyl-amine in this chromatographic system and since co-chromatography of standard dansyl-alanine gave ambiguous results, a 5mg sample of protein A was subjected to automatic sequence analysis by the Edman technique (Hermodson *et al.*, 1972) on a Beckman model 890 C Sequencer. After 10 cycles of degradation the protein was removed, hydrolysed and subjected to amino acid analysis as described in the Experimental section. Apart from a lower concentration of lysine due to only partial recovery of lysine during hydrolysis of phenolthiocarbamoyl-lysine this protein preparation had the same composition as the untreated protein. These results indicate that the *N*-terminal amino acid is blocked.

C-Terminal amino acid sequence

The experiments with penicillo-carboxypeptidase-S1 are summarized in Table 3.

These experiments clearly establish that arginine is the *C*-terminal amino acid. The next most abundant

Table 1. *Amino acid composition of protein A*

The results are mean values±S.D. of 24h hydrolysis of three different preparations of protein A except for tryptophan (determined from a sample hydrolysed in the presence of thioglycollic acid). The value for serine has not been corrected for loss during hydrolysis and the recoveries of valine and isoleucine have not been corrected.

Amino acid	Amino acid (% of total residues)	Nearest integral per 10800 molecular weight
Lysine	0.96±0.04	1
Histidine	1.9±0.12	2
Arginine	3.8±0.05	4
Aspartic acid	10.3±0.27	11
Threonine	0	0
Serine	4.3±0.08	4–5
Glutamic acid	27.2±1.11	29
Proline	23.6±1.54	25
Glycine	18.9±0.57	20
Alanine	1.07±0.07	1
Half-cysteine	0	0
Valine	2.5±0.05	3
Methionine	0	0
Isoleucine	1.6±0.11	2
Leucine	3.0±0.13	3
Tyrosine	0	0
Phenylalanine	0.95±0.02	1
Tryptophan	0	0

Table 2. *Phosphorus analysis of protein A*

Inorganic and total phosphorus was determined by the method of Bartlett (1959). For details of analyses see the text.

Treatment of protein before analysis	Type of phosphorus determination	μg of P/ 100 μg of protein (A)
None	Inorganic	0.08
	Total	0.56
Hydrolysis in 1M-NaOH at 37°C for 4h	Inorganic	0,48
	Total	0.55
Chloroform–methanol extraction		
Supernatant	Total	0
Residue	Total	0.55

Table 3. *Release of amino acids from protein A by penicillo-carboxypeptidase-S1*

Samples of protein A were incubated with penicillo-carboxypeptidase-S1 at pH4.7 as described in the Experimental section. The amount of protein A digested is based on a minimum molecular weight of 10800 determined from the amino acid analysis. Glutamic acid is probably due to deamidation of glutamine during heat inactivation of the enzyme. 'Trace' indicates an amount too small to be measured but the amino acid is unambiguously present. n.d. indicates not determined, owing to an irregularity in the recording of the chromatogram.

	Protein A digested	Enzyme/ substrate ratio	Time of incubation	Amino acid released (nmol)						
Experiment	(nmol)	(w/w)	(min)	Ala	Asp	Ser	Glu	Gln	Gly	Arg
1	47	1:200	60	0	1.7	2.1	0.7	1.3	8.2	36
2	44	1:200	180	Trace	1.7	1.9	0	2.8	9.1	47
	44	1:200	480	Trace	2.7	2.4	1.2	2.7	8.4	50
3	24	1:50	20	1.3	1.5	3.2	1.3	2.6	5.2	10
	24	1:50	40	n.d.	1.6	3.8	1.8	3.2	5.9	12
	24	1:50	60	1.4	1.6	3.8	1.6	3.3	5.8	18
	24	1:50	600	2.1	3.6	4.3	2.7	5.1	7.0	47

amino acid present in all the hydrolysates is glycine, followed by glutamine, if it is assumed that the small amount of glutamic acid is due to deamidation of glutamine during heat inactivation of the enzyme. Except for one experiment with a hydrolysis period of 1 h, in which the same amount of glutamine and serine was present, serine was consistently found to be present in slightly smaller amounts than glutamine. Serine is followed in abundance by aspartic acid in all hydrolysates of protein A incubated for less than 8 h. In addition to these amino acids small amounts of alanine could be detected in most hydrolysates.

Because of the amount of arginine released there must be at least two arginine residues present in the *C*-terminal sequence. Since the protein contains four arginine residues the possible *C*-terminal sequences can be written: Ala-$(Arg)_p$-Asp-$(Arg)_q$-Ser-$(Arg)_r$-Gln-$(Arg)_s$-Gly-$(Arg)_t$, where p, q, r and s may vary from 0 to 3 and t from 1 to 4. In order to fit possible sequences to the experimental data it is not likely that any of p, q, r and s are larger than 0 and that t is larger than 2. It is therefore proposed that the protein has the following *C*-terminal amino acid sequence: Ala-Asp-Ser-Gln-Gly-Arg-Arg.

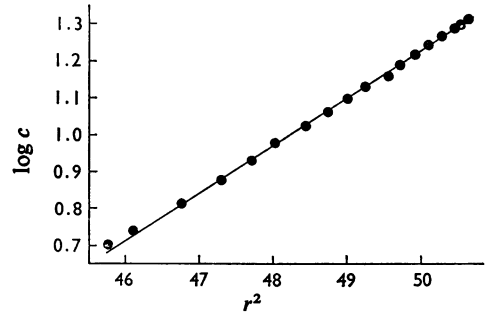

Fig. 4. *Ultracentrifugation of protein A by the low-speed sedimentation–equilibrium method*

The concentration of the solution was 2.6 mg of protein (A)/ml. The sample was dissolved in 0.012 M-sodium phosphate buffer containing 0.18 M-NaCl, pH 6.4. The speed of the ultracentrifuge was 21 053 rev./min, and the temperature of the rotor was 20°C. The log of the concentration measured as fringe displacement is plotted against the square of the distance from the centre of the rotor (r^2) measured in cm^2. The data were fitted to a straight line by the least-squares method. From the data obtained in this experiment and a partial specific volume of 0.695 the molecular weight was found to be 9900±80 (S.D.).

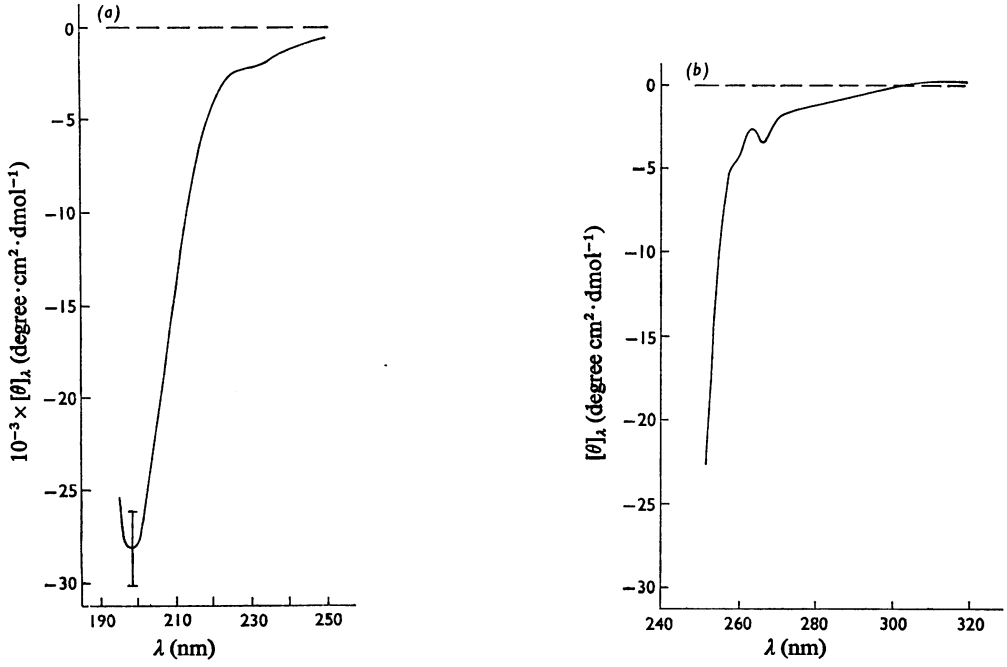

Fig. 5. *C.d. spectra of protein A*

In the near-u.v. region (250–320 nm; *b*) the spectrum was recorded in a cell with an optical path of 1 cm of a solution containing 4.55 mg of protein (A)/ml. In the far-u.v. region (190–250 nm; *a*) a cell with an optical path of 1 mm was used and the concentration of the solution was 0.1 mg/ml. In the far-u.v. region the error was estimated from repeated scans of the same solution. The magnitude of the error at 198 nm is indicated by a vertical bar. Under the experimental conditions the error in the near-u.v. region is estimated to be 0.2 degrees·cm^2·$dmol^{-1}$. E_{280} of the solution was 0.01. Spectrum of protein solution, ——; blank recording, – – – –.

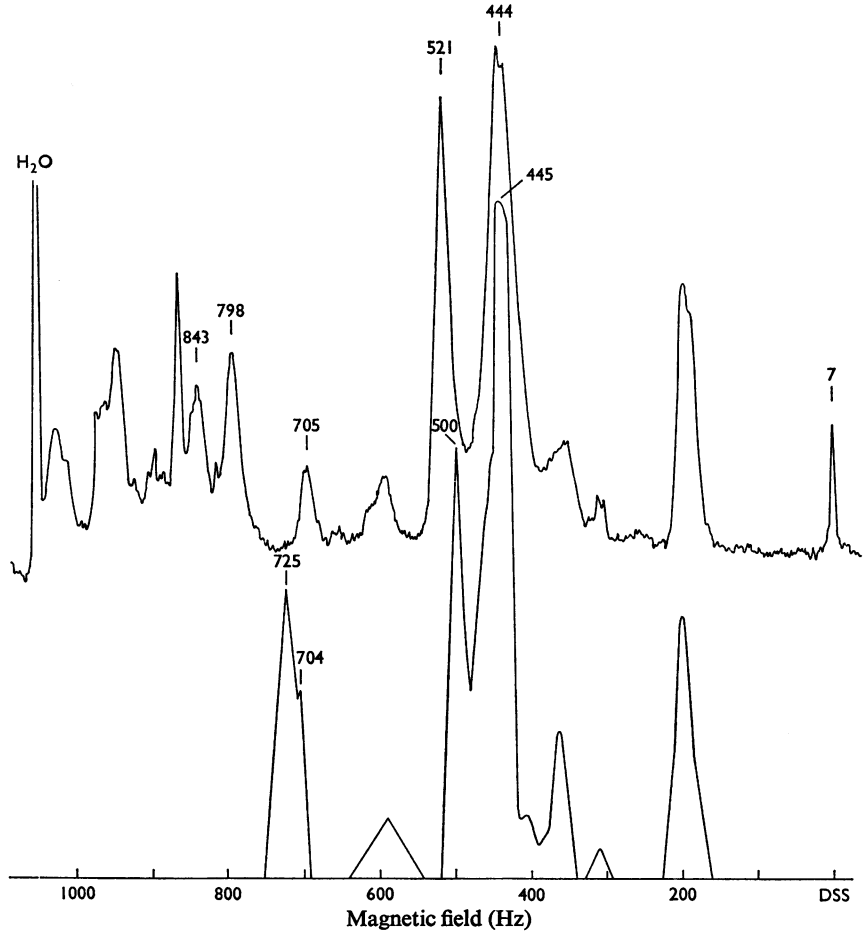

Fig. 6. *N.m.r. spectrum of protein A*

Only the part of the spectrum upfield to the H_2O resonance is illustrated. All measurements are in Hz downfield to the 2,2-dimethylsilyltetradeuteriopropionic acid (DSS) signal. The upper spectrum is an illustration of an n.m.r. spectrum of a solution containing 10mg of protein (A)/ml in ^2H_2O. The lower spectrum illustrates a spectrum computed as described in the text. The resonances of the α protons in the amino acids have not been included in the computed spectrum. The resonance position of a number of signals in both spectra have been indicated, above the recording of protein A and below the computed spectrum respectively. For experimental details see the text.

Molecular-weight determination

The molecular weight determined by equilibrium sedimentation did not show any dependence on the speed of the ultracentrifuge. An experimental plot of r^2 (distance from the rotor)^2 versus log c (concentration of protein expressed as fringe displacements) is illustrated in Fig. 4. The partial specific volume was 0.695 and the average molecular weight of three determinations was 9900 ± 340 (S.D.). Attempts were made to determine the molecular weight of protein A by sodium dodecyl sulphate–polyacrylamide-gel electrophoresis by the method of Weber & Osborn (1969). Proteins with known molecular weight were applied to separate gels. Under the electrophoretic conditions protein A migrated only 2mm into the gel, and myoglobin with a molecular weight of 17200 migrated 31mm. Sodium dodecyl sulphate–polyacrylamide-gel electrophoresis is therefore not a suitable method for determining the molecular weight of protein A.

Circular dichroism of protein A

In an aqueous solution of protein A resolved bands can be seen in the near-u.v. region at 267nm (–) and 264nm (+) (Fig. 5). These bands correspond closely to those for *N*-acetylphenylalanine amide (Horwitz

et al., 1969). Bands can also be seen at 260nm (–) and 257nm (+) in *N*-acetylphenylalanine amide. These bands could not clearly be observed in protein A, probably because of the low contribution of phenylalanine to the total optical activity of the protein at these wavelengths.

No other bands could be seen in the near-u.v. region, in agreement with the absence of aromatic residues other than phenylalanine from protein A. When protein A was dissolved in 6M-guanidine hydrochloride a very similar c.d. spectrum was obtained in the near-u.v. region. There was little if any change in the intensities of the bands observed at 267 and 264nm.

In the far-u.v. region there was a minimum at 198nm. The mean residue ellipticity at this wavelength is 28×10^3 degrees·cm^2·dmol^{-1}.

N.m.r. spectroscopy

A spectrum of protein A recorded at pH5.6 at 18°C is illustrated in Fig. 6. For comparison a spectrum computed as described in the Experimental section is also illustrated. The two spectra have been aligned according to a 2,2-dimethyl-2-silapentane-5-sulphonate reference signal (the 3-trimethylsilyltetradeuteriopropionic acid signal is 3.4Hz upfield to the 2,2-dimethyl-2-silapentane-5-sulphonate signal).

In general there is good agreement of the position and relative intensities of the signals in the recorded and the computed spectrum, although there are some obvious differences.

The resonance at 444Hz in the recorded spectrum corresponds well to the resonance at 445Hz in the computed spectrum, but the intensity of the signal at 444Hz relative to the other resonances in the recorded spectrum is somewhat lower than expected from the computed spectrum. Since the region from 415 to 500Hz contains three overlapping signals from the protons on the γ and β carbon atoms in proline and the protons on the β carbon atom in glutamic acid, it is possible that slight shifts in resonance position and widths in the recorded spectrum could account for the discrepancies in the two spectra. It is also possible that a better fit of the recorded spectrum with the computed spectrum could be obtained if the number of glutamine and asparagine residues in the protein was known.

The recorded spectrum of protein A shows a resonance at 7Hz. Such upfield resonances are common in native proteins (McDonald & Phillips, 1969). This resonance is not seen in the computed spectrum, in agreement with the absence of such resonances from denatured proteins. The recorded spectrum shows a resonance at 705Hz, which agrees well with a resonance of 704Hz in the computed spectrum owing to the protons attached to the δ carbon atom in arginine. This resonance overlaps partly with an intense resonance in the computed spectrum at 725Hz from the δ portons in proline. No corresponding signal can be seen at 725Hz in the recorded spectrum.

Discussion

The amino acid composition of protein A shows some variation from the previously reported results (Bennick & Connell, 1971), reflecting a higher degree of purity of the present preparation. In comparison with other salivary proteins, the composition of protein A is very similar to that of protein III (Oppenheim *et al.*, 1971). On the other hand the two proteins show some variation in the concentration of some amino acids, notably threonine, glutamic acid and proline. Differences are also noted in the isoelectric points. For protein A a value of 4.43 was obtained, but for protein III it was 4.14. On the basis of the available data it is not possible to determine if protein A is identical with protein III, or if the two proteins are related but distinct from each other. The phosphate analysis suggests that protein A contains 2 atoms of P/mol of protein. The alkali-lability of the organic phosphate suggests that the phosphorus is linked in an ester linkage to serine in the protein (Anderson & Kelly, 1959). A phosphorus-containing protein has been identified in human submaxillary saliva (Boat *et al.*, 1974). The composition of this protein indicates that it is not identical with protein A. The presence of phosphorus in protein A is of interest because of the high affinity of phosphoprotein for hydroxyapatite. It has also been demonstrated by Hay (1973) that proteins with a similar composition (proteins I, II, III and IV) show a higher tendency for adsorption to hydroxyapatite than do most other proteins from parotid saliva. These proteins may therefore be important in formation of dental deposits. The small amount of neutral carbohydrate and the absence of hexosamine make it unlikely that protein A is a glycoprotein. The carbohydrate could be a contaminant from the dextran resins, although it is possible that protein A is a glycoprotein of unusual composition. Of the four arginine residues in the protein two are present as the *C*-terminal and penultimate *C*-terminal residues. There is a total of 107 amino acid residues in the protein, and on average proline should occupy every fourth position in the polypeptide chain. It is interesting that of the seven amino acids identified in the *C*-terminal none are proline. It would be valuable to compare the *C*-terminal sequence of other salivary proteins with the sequence in protein A in order to determine possible relationships between these proteins. The molecular weight of 9900 determined from ultracentrifugation is somewhat lower than the minimum molecular weight of 11000 obtained from chemical analysis of the protein. This discrepancy in molecular weights may

in part be due to inaccuracies in the amino acid analysis.

That part of the protein is in an unordered conformation is suggested from the c.d. spectra in the near-u.v. region, since there is no change in this part of the spectrum when the protein is dissolved in 6M-guanidine hydrochloride, indicating that there may be little restriction on the orientation of phenylalanine in native protein A.

In the far-u.v. region the c.d. spectrum of protein A is similar to spectra of proteins in unordered conformation, which show a single minimum between 200 and 205nm, but the mean residue elipticity of protein A corresponding to the minimum is somewhat larger than the values obtained for proteins with unordered conformations (Fasman *et al.*, 1970) and closer to values which have been obtained for collagen, collagen-like polypeptides (Brown *et al.*, 1972) and a protein containing polyproline-like structure (Kanaya & Fujimoto, 1973). These proteins and polypeptides all show a single minimum in the far-u.v. region, between 198 and 205nm, but protein A lacks the positive band at about 222nm seen in molecules with the triple-helical structure (Brown *et al.*, 1972).

From the c.d. spectra no secondary structure can be recognized in protein A. More information about the conformation of protein A can be obtained from the n.m.r. spectra.

McDonald & Phillips (1969) suggested an assignment of 725Hz for the δ-H proline resonance in randomly coiled proteins based on observations of n.m.r. spectra of free proline and proteins in a random-coil conformation. In a number of low-molecular-weight derivatives it has been shown that the resonance position is different from that observed in free proline. For these derivatives values between 760 and 785Hz have been obtained for the δ protons in proline (Madison & Schellman, 1970). The lack of a signal at 725Hz in protein A does not therefore necessarily mean that proline is participating in maintaining a native conformation of protein A. In polyproline type I, which exists as a compact right-handed helix, as well as in polyproline type II, which forms a left-handed extended helix, it has been observed that the δ protons in proline have resonance positions at 705Hz and about 741Hz (Deber *et al.*, 1970). If we assume that the resonance at 705Hz in the spectrum of protein A is due to δ protons in arginine, as predicted from the computed spectrum, it may be concluded that there is no evidence of polyproline in protein A. It is also interesting to compare the n.m.r. spectrum of protein A with the spectra obtained by Kobayashi & Kyogoku (1973) on $(Pro\text{-}Pro\text{-}Gly)_{10}$. At room temperature these molecules aggregated into a triple-helical structure, which on n.m.r. spectroscopy show resonances at 700 and 780Hz which have been assigned to δ protons in proline and a resonance at 850Hz which has been assigned to α protons in glycine. The spectra of protein A show similar resonances at 705, 798 and 843Hz. The formation of the triple helix in $(Pro\text{-}Pro\text{-}Gly)_{10}$ is, on the other hand, associated with a number of spectral changes in the region from 410 to 510Hz. The spectrum of protein A in this region cannot be easily interpreted because of the predicted presence of resonances from protons in glutamic acid as well as proline. The c.d. spectrum of protein A shows no presence of triple-helical structure in the protein, and there is therefore no indication of any known secondary structure in protein A. This does not necessarily mean that the protein exists as a random coil. Indeed the upfield resonance at 7Hz is indicative of conformational restraint in the protein.

I thank Mr. R. Dove and Mrs. M. Cannon for invaluable technical assistance. I also thank Mr. C. Yu for performing the amino acid analyses, Dr. D. Kells for undertaking the ultracentrifugations and Dr. K. J. Dorrington for assistance with the c.d. spectroscopy. The n.m.r. spectra were recorded by Dr. A. A. Grey and Mrs. V. Robinson at the Canadian 220MHz n.m.r. centre. This work was supported by grant no. MA 4920 from the Canadian Medical Research Council.

References

Anderson, L. & Kelly, J. J. (1959) *J. Amer. Chem. Soc.* **81**, 2275–2276
Anwar, R. A. (1965) *Can. J. Biochem.* **44**, 725–734
Armstrong, W. G. (1971) *Caries Res.* **5**, 215–227
Azen, E. A. & Oppenheim, F. G. (1973) *Science* **180**, 1067–1069
Bartlett, G. R. (1959) *J. Biol. Chem.* **234**, 466–468
Bennick, A. & Connell, G. E. (1971) *Biochem. J.* **123**, 455–469
Benson, J. V., Gordon, M. J. & Patterson, J. A. (1967) *Anal. Biochem.* **18**, 228–240
Boat, T. F., Wiseman, U. N. & Pallavicini, J. C. (1974) *Pediat. Res.* **8**, 531–539
Brown, F. R., di Corato, A., Lorenzi, G. P. & Blout, E. R. (1972) *J. Mol. Biol.* **63**, 85–99
Chrambach, A., Reisfeld, R. A., Wyckoff, M. & Zaccari, J. (1967) *Anal. Biochem.* **20**, 150–154
Deber, C. M., Bovey, F. A., Carver, J. P. & Blout, E. R. (1970) *J. Amer. Chem. Soc.* **92**, 6191–6198
Dorrington, K. J. & Smith, B. R. (1972) *Biochim. Biophys. Acta* **263**, 70–81
Eggert, M., Allen, G. A. & Burgess, R. C. (1973) *Biochem. J.* **131**, 471–484
Fasman, G. D., Hoving, H. & Timasheff, S. N. (1970) *Biochemistry* **9**, 3316–3324
Fleischmajer, R. & Fishman, L. (1965) *Nature (London)* **205**, 264–266
Hay, D. I. (1973) *Arch. Oral Biol.* **18**, 1517–1530
Hermodson, M. A., Ericson, L. H., Titani, K. & Neurath, H. (1972) *Biochemistry* **11**, 4493–4502
Horwitz, J., Strickland, E. H. & Billups, C. (1969) *J. Amer. Chem. Soc.* **91**, 184–190

Jones, S. R. & Hofmann, T. (1972) *Can. J. Biochem.* **50**, 1297–1310
Kanaya, S. & Fujimoto, D. (1973) *J. Mol. Biol.* **81**, 415–418
Kobayashi, Y. & Kyogoku, Y. (1973) *J. Mol. Biol.* **81**, 337–347
Levine, M. J., Ellison, S. A. & Bahl, O. P. (1973) *Arch. Oral Biol.* **18**, 827–837
Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275
Madison, V. & Schellman, J. (1970) *Biopolymers* **9**, 511–588
Matsubara, H. & Sasaki, R. M. (1969) *Biochem. Biophys. Res. Commun.* **35**, 175–181
McDonald, C. C. & Phillips, W. D. (1969) *J. Amer. Chem. Soc.* **91**, 1513–1521
Muenzer, J. & Carlson, D. M. (1974) *Fed. Proc. Fed. Amer. Soc. Exp. Biol.* **33**, 1497
Oppenheim, F. G., Hay, D. I. & Franzblau, C. (1971) *Biochemistry* **10**, 4233–4238
Percy, M. E. & Buchwald, B. M. (1972) *Anal. Biochem.* **45**, 60–67
Richards, E. G. & Schachman, H. K. (1959) *J. Phys. Chem.* **63**, 1578–1591
Seifter, S., Dayton, S. & Muntwyler, E. (1950) *Arch. Biochem.* **25**, 191–200
Weber, K. & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406–4412