Isolation and characterization of the two giant secretory proteins in salivary gland of *Chironomus tentans*

Lars RYDLANDER

Department of Histology, Karolinska Institutet, S-104 01 Stockholm, Sweden

(Received 28 November 1983/Accepted 20 February 1984)

The two giant secretory proteins, sp-Ia and sp-Ib, in salivary-gland cells of the larva of the fly *Chironomus tentans*, were isolated by preparative gel electrophoresis and characterized chemically. Their amino acid compositions are dominated by polar amino acids, with about 30% of basic amino acid residues. Crossed immunoelectrophoresis of sp-Ia and sp-Ib provided evidence that they share antigenic determinants. They also have major methionine-containing tryptic peptides in common. CNBr cleavage of sp-Ib gives a small number of low- M_r fragments, indicating that this protein has a repetitive structure.

Many insects spend most of their life cycle as larvae, like members of the genus *Chironomus* (Diptera). During their aquatic life the larvae construct dwelling-tubes and funnel-shaped catching nets from saliva (Grossbach, 1977). The secretion from *Chironomus tentans* salivary glands consists of five proteins (Grossbach, 1969) with M_r values ranging from 10^6 to 1.7×10^4 (Grossbach, 1977; Hertner *et al.*, 1980; Rydlander & Edström, 1980). The largest component, secretory protein I (sp-I), of molecular mass about 850 kDa, predominates, making up about 50% of the secretory protein mass. This 850 kDa protein is complex and normally consists of two subcomponents, designated 'sp-Ia' and 'sp-Ib'; usually sp-Ib is the dominating one. Another prominent component is secretory protein II (sp-II), of molecular mass about 235 kDa, amounting to 20–40% of the total secretory protein. The smaller components, III–V, account for the remaining 10–15% of the total secretory proteins (Rydlander & Edström, 1980).

Translation *in vitro* of 75S RNA, the giant transcripts from the Balbiani rings BR 1 and BR 2 (Daneholt, 1972), gives products with tryptic fingerprints similar to those of the 850 kDa proteins, suggesting that 75S RNA from BR 1 to BR 2 codes for sp-I (Rydlander *et al.*, 1980; Weber *et al.*, 1983).

Little is known about the structure of sp-Ia and sp-Ib, but considerably more is known about the BR genes. The different BR genes in the salivary gland are present in one or a few copies per genome

and contain internally repeated DNA (Lambert, 1972; Wieslander, 1979). Molecular cloning of BR 1 and BR 2 sequences of *C. tentans* revealed tandemly repeated sequences within BR DNA (Degelmann & Hollenberg, 1981; Sümegi *et al.*, 1982), which implies that a large part of protein I contains repetitive sequences. Furthermore, the analysis of cloned DNA sequences also shows that the BR 1 and BR 2 genes are related (Wieslander *et al.*, 1982).

To investigate whether the organization of the characterized BR genes is manifested in their gene products, i.e. in the structure of sp-Ia and sp-Ib, these two proteins were isolated and compared with respect to amino acid composition and sequence distributions. The results show that they are chemically closely related, with repeating amino acid sequences.

Materials and methods

Animals

Late-fourth-instar larvae of *Chironomus tentans* were obtained from a laboratory culture, originally from a stock kept by Professor W. Beermann. They were kept in aerated pans with 0.04% (w/v) NaCl in deionized water, and were given cellulose and nettle powder which have been previously fermented for 14 days. Salivary glands were excised and fixed in 70% (v/v) ethanol for 45–60 min at room temperature. The glands were then transferred to cold ethanol/glycerol (1:1, v/v), and kept at -20°C until dissection (1–24 h). The cells were removed from the secretion with insect needles

Abbreviations used: SDS, sodium dodecyl sulphate; bp, base-pair(s).

under a stereomicroscope and the fixed secretion (secretion clump) was used for further analyses.

Labelling of secretory proteins

The animals were placed for 24 h in culture medium containing $3 \mu\text{Ci}$ of [^{35}S]methionine (300–900 Ci/mmol)/1 ml. Usually about 20 animals were immersed in 30 ml of culture medium containing some cellulose powder.

Preparation of samples for SDS/agarose-gel electrophoresis

The fixed secretion clumps were transferred to small test tubes and dissolved with repeated shaking in ice-cold solubilizing buffer B [0.1 M-Tris/ H_3PO_4 (pH 7.2)/0.1 M-dithiothreitol/0.02 M-EDTA/3% (w/v) SDS]. A volume of 100 μl /ten secretion clumps was used. The samples were heated for 3 min in a boiling-water bath to complete the solubilization and to reduce the disulphide bridges in the proteins.

Isolation of secretory proteins by preparative gel electrophoresis

A 0.3 cm $\times$ 14 cm $\times$ 10 cm slab gel of 2% (w/v) agarose (SeaKem ME, Marine Colloids Division, Rockland, ME, U.S.A.) in 0.1 M-Tris/ H_3PO_4 (pH 7.2)/0.1% SDS was cast in a Bio-Rad model 220 electrophoresis cell. After the gel had set, a 10 cm-long sample well was cut.

Secretion from 60 glands, which corresponds to about 0.7 mg of protein, was dissolved in buffer B, reduced and applied to the gel after addition of 5% (v/v) glycerol. Electrophoresis was run with a constant current of 80 mA until the Bromophenol Blue marker had migrated 9.5 cm. The gel was stained for 30 min with 0.1% Coomassie Brilliant Blue R250 in ethanol/acetic acid/water (5:2:13, by vol.) and subsequently destained with ethanol/acetic acid/water (5:2:13, by vol.). The gel was then washed with water until all acetic acid was removed. The stained protein bands were excised, cut into small pieces (about 1 mm³) and transferred to test tubes. The proteins were eluted from the gel slices with 10 vol. of 0.05 M-Tris/HCl (pH 7.9)/0.05 M-NaCl/1 mM-phenylmethanesulphonyl fluoride/2% (w/v) SDS/1% (w/v) 2-mercaptoethanol by slowly rotating the test tubes at room temperature. After 24 h the elution buffer was collected and the extraction procedure was repeated once.

An alternative procedure was used to isolate secretory proteins for tryptic peptide mapping and CNBr cleavage. Instead of Coomassie Brilliant Blue staining the different protein bands in the preparative gel were located with KCl as described by Hager & Burgess (1980). After the protein bands had been located, they were excised, cut into small pieces, rinsed with distilled water and equili-

brated with electrophoresis buffer. The gel pieces were then mixed with a hot solution of 0.5% agarose in electrophoresis buffer and poured into a polythene tube (14 mm $\times$ 50 mm) with a sealed lower end. A piece of dialysis tubing (18 mm $\times$ 50 mm), previously boiled in 0.01 M-EDTA and washed with distilled water, was tied off tightly at one end, filled with electrophoresis buffer and then mounted with the open end over the lower end of the polythene tube. The elution tubes, thus prepared, were inserted into a vertical electrophoresis apparatus built in the laboratory. The upper-chamber unit was lowered until half the length of the dialysis bags was below the surface of the lower buffer. The protein in the agarose was electrophoretically eluted into the dialysis sac at 40–60 V and 10–15 mA per tube in the cold overnight. Before the dialysis sac was emptied with a hypodermic needle, the electrophoresis direction was reversed for a few minutes.

Quality control of the purified secretory proteins

Labelled sp-Ia and sp-Ib were purified after 'KCl staining'. The different components were precipitated overnight at 4°C from one-fifth of the electroeluate after addition of acetone to 70% (v/v). The precipitate was collected by centrifugation [Sorvall SS34 centrifuge; 10000 rev./min (r_{av} 8.4 cm) for 10 min], dissolved in buffer B, heated in a boiling-water bath for 3 min and then applied to a 2% (w/v) agarose gel. Unfractionated secretory proteins diluted to approximately the same radioactivity were applied to a parallel lane of the same gel. Electrophoresis was stopped when the Bromophenol Blue marker had migrated 10 cm. After fixation with 10% (w/v) trichloroacetic acid, the gel was prepared for fluorography with en³Hance (New England Nuclear Corp.) according to the manufacturer's instructions.

Amino acid analysis

Eluates of the different secretory proteins from the Coomassie Brilliant Blue-stained preparative gel containing between 30 and 150 μg of protein were divided into three parts. These were precipitated with 20% (w/v) trichloroacetic acid, spun down [Sorvall SS34 centrifuge; 10000 rev./min (r_{av} 8.4 cm) for 30 min], washed twice with 10% trichloroacetic acid, further washed twice with acetone at -20°C and then dried. One part was oxidized with performic acid at 2°C (Hirs, 1967); the other two parts were hydrolysed without oxidation. Hydrolysis was performed under vacuum at 110°C for 22 h in 6 M-HCl containing 0.5% phenol. Amino acids were determined on a Beckman 121 M analyser. No corrections were made for hydrolysis losses of serine and threonine or incomplete release of valine.

Preparation of antisera

Antiserum was prepared by injection of total salivary buffer-B-soluble material intradermally into the thicker part of the skin above the scapula in rabbits weighing 3–4 kg as described by Rydlander & Edström (1980). Serum was harvested after centrifugation [Sorvall SS34 centrifuge; 2000 rev./min (r_{av} , 8.4 cm) for 10 min] and was stored at -20°C until used.

Crossed immunoelectrophoresis

Reduced secretory proteins were separated in the first dimension by electrophoresis in SDS/1% agarose/2.2% polyacrylamide composite gel. One lane from the first-dimension gel was cut out and run in the second dimension as described by Rydlander & Edström (1980). After electrophoresis the gel was washed overnight with 0.9% NaCl containing 0.1% Nonidet P40, then with 0.9% NaCl only, and finally with distilled water. The plate was air-dried and stained with Coomassie Brilliant Blue (Weeke, 1973).

Preparation of tryptic peptides

The purified secretory proteins from one preparative gel, or secretory proteins Ia/Ib prepared by gel filtration on Sepharose 6B (Rydlander *et al.*, 1980), were precipitated with 20% (w/v) trichloroacetic acid and treated as described above. The dry protein (50 to 300 μg) was dissolved in 0.3 ml of 98% formic acid, and 2 vol. of performic acid were added. Oxidation was carried out for 4 h at 2°C , after which the sample was diluted 10-fold with water and freeze-dried. The dried material was dissolved in 1 ml of 1% NH_4HCO_3 containing 2–10 μg of trypsin (EC 3.4.21.4; Sigma type III), which gives an enzyme/substrate ratio of 1:30. Digestion was carried out for 4 h, after which a new portion of trypsin solution was added and the digestion continued for 18 h. The sample was finally freeze-dried, dissolved in water and freeze-dried again; this was repeated three to four times.

Two-dimensional tryptic-peptide analysis

The tryptic digest was dissolved in 5 μl of 20% (v/v) pyridine, and 2 μl of the material was applied to a thin-layer cellulose sheet (Polygram Cel 300; Machery, Nagel and Co., Düren, Germany). The sheet was subjected to electrophoresis at pH 6.5 in pyridine/acetic acid/water (25:1:225, by vol.) at 400 V until a Methyl Green dye marker had migrated 5 cm, and dried in air overnight. Ascending chromatography at right angles to the direction of electrophoresis was carried out with freshly prepared butanol/acetic acid/water/pyridine (15:3:12:10, by vol.). The sheets were finally dried for autoradiography (Kodak XAR-film).

Determination of tryptic-peptide molecular masses

Bio-Gel P-30 (Bio-Rad Laboratories) was packed to a height of 55 cm in a Pharmacia K9/60 column. The column was equilibrated with de-aerated 8 M-urea/0.1 M-Tris/ H_3PO_4 buffer, pH 7.2 (running buffer). [^{35}S]Methionine-labelled tryptic peptides obtained by digestion of different secretory proteins were dissolved in buffer B containing 8 M-urea, heated for 10 min at 60°C , and stored overnight at room temperature. The different tryptic peptides, separated by electrophoresis and chromatography on thin-layer plates, were located by autoradiography. The spot containing the peptide was scraped off from the plate and the peptide was eluted with 0.1 M-pyridine/acetic acid, pH 6.5. A portion of the eluate was used for radioactivity determination, the rest was freeze-dried and the peptide dissolved as described above. Before application of the samples to the column, Blue Dextran 2000 (Pharmacia) and $^3\text{H}_2\text{O}$ (Packard) were added. The column was eluted with running buffer at a flow rate of 1.0 ml/h, 1 ml fractions being collected. Then 6 ml of Bio-Fluor (New England Nuclear Corp.) was added to each fraction and radioactivity counting was done in a Packard liquid-scintillation counter. The void volume (V_0) and the inner volume (V_i) were determined from the elution position of Blue Dextran and $^3\text{H}_2\text{O}$. The column was calibrated with cytochrome *c* (Sigma type VI; M_r 12400) and the following peptides: aprotinin (M_r 6500), bacitracin (M_r 1450) (Sigma), insulin B-chain (M_r 3495), Suc-Ala-Ala-Pro-Phe *p*-nitroanilide (M_r 625) and Phe-Ala (Serva; M_r 236). Each protein/peptide was run three times and its elution position recorded at 280 nm. The average distribution coefficient (V_e/V_0) was calculated for each protein/peptide.

Peptide mapping of sp-Ib after cleavage with CNBr

Sp-Ib, located by KCl and eluted from a preparative gel was precipitated by adding trichloroacetic acid to 10%. The precipitate was spun down, washed with 10% trichloroacetic acid, then acetone, and dried. The dry protein (0.2–0.3 mg) was dissolved in formic acid containing 1% (w/v) mercaptoethanol, diluted to 70% (v/v) formic acid and dialysed against 70% formic acid with 0.05% mercaptoethanol at 4°C overnight. CNBr (0.5 g) was added, and the samples were left at room temperature overnight. The reagent and the solvent were removed by repeated evaporation. The peptides were dissolved in buffer B, heated for 3 min at 100°C , and analysed by electrophoresis on SDS/15%-polyacrylamide gel (Laemmli, 1970). After electrophoresis the gel was fixed with 12%

trichloroacetic acid, stained with Coomassie Brilliant Blue, and destained.

Results

Isolation of *sp-Ia* and *sp-Ib*

The diffusion method for extraction is easy to handle and gives a recovery in each elution step of 60–70%, resulting in a total recovery of 85–90%. Gel electrophoresis of the eluted protein material showed, however, that *sp-Ia* and *sp-Ib* are degraded to fragments with M_r values between 200 000 and 400 000 (results not shown). The diffusion method was used for amino acid analysis because it gives low blank values.

The electroelution technique, which is more laborious, gives a total recovery in the same range. Subsequent electrophoresis shows that *sp-Ia* and *sp-Ib* are undegraded, and that there is little if any cross-contamination (Fig. 1).

Amino acid analyses of *sp-Ia* and *sp-Ib*

The results of the analyses are presented in Table 1. Both secretory proteins have 62–70%

Fig. 1. Purity of *sp-Ia* and *sp-Ib* prepared by agarose electrophoresis

A photometric scan of a Coomassie Brilliant Blue-stained preparative gel is shown. The hatched areas were used for elution of *sp-Ia* and *sp-Ib*. Roman numerals indicate different secretory proteins. The inset shows a fluorograph of a 2%-agarose gel from another experiment with [^{35}S]methionine-labelled proteins prepared from 30 glands. (A) purified *sp-Ia*; (B) purified *sp-Ib*; and (C) unfractionated secretory proteins. The exposure time was 40 h.

polar residues and about 30% of these are basic. Lysine, proline and serine constitute 50% of the total amino acids in *sp-Ia*. In *sp-Ib*, lysine, arginine and serine amount to about 42%. There are significant differences between *sp-Ia* and *sp-Ib* for three amino acids. Proline and lysine are more common in *sp-Ia*; glycine is more common in *sp-Ib*. Nevertheless, the amino acid compositions of *sp-Ia* and *sp-Ib* are very similar. As a test for similarities in amino acid sequences, an immunological comparison was performed.

Immunological comparison of *sp-Ia* and *sp-Ib*

A typical precipitin pattern of the major secretory proteins run in crossed immunoelectrophoresis is shown in Fig. 2. Two precipitin arcs are found: one complex system formed by *sp-Ia*/*sp-Ib*, and a weak single arc corresponding to a protein slightly smaller than *sp-Ib*. The precipitin peak from *sp-Ia*/*sp-Ib* is asymmetric, with two sharp edges above components Ia and Ib and with a slope towards the low- M_r side. The arc from *sp-Ia* is fused with the arc from *sp-Ib*, which implies that some of the antigenic determinants are identical. The weak single arc below the complex from proteins Ia/Ib probably represents the inducible secretory protein Ic, which is sometimes seen in small amounts in normal animals (Edström *et al.*, 1980). These results show that *sp-Ia* and *sp-Ib* have some cross-reacting amino acid sequences, and that *sp-Ia* also has unique ones. This is in essential agreement with previous studies on the related species,

Table 1. Amino acid compositions of secretory proteins *sp-Ia* and *sp-Ib*

Tryptophan was disregarded; cysteine was determined as cysteic acid. Values are averages from three different preparative gels (three duplicate analyses for each secretory protein).

Amino acid	Composition (mol/100 mol) ($\pm$ S.D.)	
	<i>Sp-Ia</i>	<i>Sp-Ib</i>
Cys	4.1 ± 0.2	5.0 ± 0.2
Asx	3.5 ± 0.3	2.4 ± 0.3
Thr	2.1 ± 0.4	2.3 ± 0.3
Ser	14.1 ± 1.9	15.1 ± 2.2
Glx	8.2 ± 0.7	8.5 ± 0.6
Pro	14.9 ± 1.5	8.6 ± 1.1
Gly	4.4 ± 1.9	11.1 ± 2.3
Ala	4.6 ± 1.1	7.6 ± 0.7
Val	1.4 ± 0.6	1.9 ± 0.9
Met	1.2 ± 0.7	0.9 ± 0.6
Ile	1.1 ± 0.6	2.3 ± 0.4
Tyr	0.7 ± 0.6	1.5 ± 1.1
Phe	1.3 ± 0.7	1.9 ± 0.5
Lys	21.1 ± 3.2	13.8 ± 2.2
His	4.8 ± 1.0	1.5 ± 0.5
Arg	11.2 ± 0.7	13.5 ± 0.9

C. pallidivittatus (Edström *et al.*, 1980). To obtain information about these sequence similarities, methionine containing tryptic peptides were prepared.

Fig. 2. Crossed immunoelectrophoresis of *sp-Ia* and *sp-Ib*. Secretion from 12 fixed glands was dissolved in buffer B, reduced, and separated on a 1%-agarose/2.2%-acrylamide composite gel. One separation was fixed and stained (lower part of the Figure), the other was used for crossed immunoelectrophoresis (upper part). The dark spots at the lower edge of the immunoelectrophoretic separation are staining artefacts. Roman numerals indicate different secretory proteins. The starting point of electrophoresis is indicated by an arrow.

Fig. 3. Gel filtration in 8M-urea on Bio-Gel P-30 of [^{35}S]-methionine-labelled tryptic peptides from *sp-Ia/sp-Ib*. Labelled *sp-Ia/sp-Ib* from 20 glands was isolated by gel filtration on Sepharose 6B. Tryptic peptides were prepared as described in the Materials and methods section, dissolved in buffer B containing 8M-urea and applied to the column.

Preparation of tryptic peptides from *sp-Ia* and *sp-Ib*

For complete digestion of proteins Ia and Ib, which contain large amounts of lysine and arginine, it was necessary to use a high enzyme/substrate ratio and to extend the digestion time. Previous results (Rydlander *et al.*, 1980) showed ten major and nine minor methionine-containing tryptic peptides after 4h of digestion. An increased trypsin concentration and prolonged digestion time decreased the number of peptides. Since further digestion did not change the peptide pattern, 4h+18h was used as standard digestion time.

Gel filtration of [^{35}S]methionine-containing tryptic peptides on Bio-Gel P30 in 8M-urea (Fig. 3) shows that 10% of the labelled material was eluted close to the void volume. This may represent large peptide(s) or undigested material, whereas most of

Fig. 4. Calibration of the Bio-Gel P-30 column: a semilogarithmic plot of M_r versus V_e/V_0 for different standard peptides and proteins

The different substances were dissolved in buffer B containing 8mM-urea as described in the Materials and methods section and applied to the column after the addition of Blue Dextran 2000 and $^3\text{H}_2\text{O}$. Standard proteins/peptides used were: cytochrome *c* (M_r 12400), aprotinin (6500), insulin B-chain (3495), bacitracin (1450) and Suc-Ala-Ala-Pro-Phe *p*-nitroanilide (625) and Phe-Ala (236). ●, Standard protein/peptide; □, major [^{35}S]methionine-labelled tryptic peptides from secretory protein Ia/Ib.

Table 2. *Partial characterization of methionine-containing tryptic peptides from sp-Ia/sp-Ib*

Before the tryptic digest was applied to the thin-layer plate, a portion was used to determine the total radioactivity (= 100%). After electrophoresis and chromatography the different peptides were localized by autoradiography and eluted from the thin-layer plate. Peptides occurring in both sp-Ia and sp-Ib are designated 'common' and those occurring in only one of the secretory proteins 'unique'. One-tenth of the eluate was used for radioactivity determination (expressed as a percentage of the total $\pm$ S.D.), the rest was used for M_r determination by gel filtration on Bio-Gel P-30 in 8 M urea ($M_r \pm$ S.D.).

Tryptic peptide from secretory protein		Chemical nature	M_r	(n)	Label (%)	(n)
Ia	Common	Basic	550 ± 80	(5)	25 ± 5	(4)
Ib					45 ± 10	(5)
Ia	Common	Neutral	1400 ± 50	(5)	8 ± 5	(4)
Ib					22 ± 5	(5)
Ia	Unique	Neutral	1100 ± 50	(4)	25 ± 7	(4)
Ia	Common	Acid	650 ± 50	(5)	20 ± 5	(4)
Ib					19 ± 5	(5)
Ia	Unique	Acid	1800 ± 100	(3)	12 ± 5	(4)

the labelled peptides were eluted in three broad peaks. The M_r values of the peptides in the two dominating peaks are about 1400 and 550 (Fig. 4), which correspond to about thirteen and five amino acids respectively (average M_r of amino acids 110). These three peaks contained about 80% of the label.

Tryptic-peptide maps of [35 S]methionine-labelled sp-Ia and sp-Ib

Complete digestion with trypsin of [35 S]-methionine-labelled proteins Ia/Ib and subsequent separation of peptides by electrophoresis in one dimension, and chromatography in the second dimension, gives a simple pattern with four to five major spots (Figs. 5a and 5b). Sp-Ia and sp-Ib give rise to three methionine-containing peptides with the same electrophoretic and chromatographic mobilities: one basic, one neutral and one acid peptide. One difference between the two fingerprints is that sp-Ia has an extra neutral peptide with a slightly lower hydrophobicity than the common neutral peptide. Furthermore, component Ia has an acid peptide by a slower migration in the second dimension. Another difference between the two fingerprints, concerning the minor peptides, is that sp-Ia has an extra acid peptide with a high electrophoretic mobility, and that sp-Ib has an extra large neutral peptide (Fig. 5c).

Sizes and amounts of the major tryptic peptides from sp-Ia and sp-Ib are given in Table 2. Two of the common peptides have different abundances in the secretory proteins. The basic peptide and the common neutral peptide are more frequent in protein Ib than in Ia. This difference also applies to the common neutral peptide. With one methionine residue per peptide the results suggest that the

Fig. 5. *Autoradiographs of peptide maps of [35 S]methionine-labelled tryptic peptides*

Labelled sp-Ia and sp-Ib were purified from one preparative gel by the electroelution technique after 'KCl staining' (for further details, see the Materials and methods section). Protein Ia/Ib from another batch of labelled glands was purified as described in the legend to Fig. 3. Tryptic peptides were prepared as described in the Materials and methods section, applied to the cellulose thin-layer plates and separated by electrophoresis followed by ascending chromatography. (A), sp-Ia; (B), sp-Ib; and (C), sp-Ia/sp-Ib. The chromatography run in (C) was extended to 1.5 times the normal time to achieve better separation of the neutral and acid peptides. The radioactivity applied was 5400 c.p.m. (A), 7300 c.p.m. (B) and 25800 c.p.m. (C). Exposure times were 14 days (A and B) and 4 days (C). The origin of each fingerprint is marked 'X', the cathode is to the left and the anode to the right. ▷, peptides common to both sp-Ia and sp-Ib; ▲, peptides unique to sp-Ia.

homology between peptides from proteins Ia and Ib is about 70%.

CNBr peptide mapping of sp-Ib

The simple pattern of methionine-containing tryptic peptides from proteins Ia/Ib and the rather high methionine content suggest that both components contain repetitive sequences. To test this possibility, sp-Ib was isolated, cleaved with CNBr and the resulting peptides separated by electrophoresis on an SDS/15%-polyacrylamide gel. The electrophoretic profile (Fig. 6) shows that CNBr cleavage of sp-Ib gives only six to eight differently sized fragments with the following approximate M_r values: 15000, 18000, 20000, 25000, 32000 and 41000. The same M_r values were obtained when peptides were run on a 10–17%-polyacrylamide-gradient gel (result not shown). Calculation of peak areas from densitometer scans shows that 70% of the CNBr fragments have M_r values of 20000 or lower.

Fig. 6. Photometric scan of gel-electrophoretic pattern of CNBr fragments from sp-Ib

Sp-Ib was purified from one preparative gel after 'KCl staining'. The protein was cleaved with CNBr. The resulting peptides were dissolved in buffer B, separated on a SDS/15%-polyacrylamide gel, and stained with Coomassie Brilliant Blue. In an adjoining slot, low- M_r marker proteins of M_r ($\times 10^{-3}$) 94, 67, 43, 30, 20.1 and 14.4 (Pharmacia) were run. Their migrations in the gel are indicated by arrowheads ($\blacktriangledown$).

Discussion

Since the nucleotide sequences of the genes coding for sp-Ia and sp-Ib are known in principle, it is interesting to compare the present findings with conclusions made from the analysis of the base sequences of the corresponding gene. The BR 2 gene is claimed to be mapped to about 85% and is built from at least two types of major repeat units. Each repeat contains a repetitive and a non-repetitive part, and two different repeat lengths have been recorded, one 215bp and the other 180bp long. These lengths correspond to 72 and 60 amino acids respectively. It has also been proposed that the 72-residue repeat covers about 60% and the 60-residue repeat about 25% of the gene (Wieslander & Lendahl, 1983).

The coded amino acid compositions of the two repeats give lysine, serine and proline as dominating amino acids for the long repeat (Sümegei *et al.*, 1982), and arginine, serine and glycine for the short repeat (Wieslander & Lendahl, 1983). Although the direct analyses of protein Ib agree in showing that four of these amino acids are the most abundant ones, two other amino acids show little agreement, namely glutamic acid and histidine. The coded amino acid composition of a 505bp-long complementary-DNA clone representative for at least 25% of the BR 1 gene (Wieslander *et al.*, 1982) also shows good agreement with the amount of the abundant amino acids in sp-Ia, i.e. lysine, serine, proline and arginine.

On the basis of the available amino acid sequence data, obtained from cloning of complementary DNA derived from 75S RNA, four different methionine-containing tryptic peptides could be expected. Two of these are Cys-Gly-Ser-Ala-Met-Arg, from BR 1 (Degelmann & Hollenberg, 1981), and Cys-Gly-Ser-Ala-Met-Lys, from BR 2 (Sümegei *et al.*, 1982), both containing cysteine, which during oxidation is converted into cysteic acid. Thus these peptides will be negatively charged at pH 6.5 and migrate towards the anode. Similar peptides are probably represented by the common acid peptide from proteins Ia and Ib. The unique acid tryptic peptide from sp-Ia might be the peptide Gly-Ser-Lys-Pro-Arg-Pro-Glu-Gly-Cys-Gly-Ser-Ala-Met-Arg, as deduced from a nucleotide sequence from BR 1 (Wieslander *et al.*, 1982). The common basic peptide could be the one that is found in the 60-amino-acid-long repeat from BR 2 and has the sequence Met-Arg(-Arg). The neutral methionine-containing tryptic peptides, which amount to about 30% of proteins Ia and Ib, have not yet been found in any complementary-DNA or genomic clone from BR 1 or BR 2.

The amino acid composition, and the sizes and chemical properties of methionine-containing

tryptic peptides from sp-Ib, are compatible with the proposed model of the BR 2 gene (Wieslander & Lendahl, 1983). Nevertheless, some points are still unclear concerning the sequences of neutral methionine-containing tryptic peptides and the differences in the amounts of two amino acids. It is possible that BR 2 contains more than one gene (Wieslander, 1979) for sp-Ib, or that it codes for sp-Ia, and that these genes have different repeats. This could explain why some sequences have not been detected by the molecular-cloning technique, as every clone is an amplification of a small part of the gene, whereas the protein data reflect the average activity of several copies of one gene or of similar genes.

The limited number of CNBr peptides derived from secretory protein Ib, the most abundant component in the secretion, confirms its repetitive structure. The actual sizes and numbers of the CNBr peptides differ considerably from those estimated from cloning data (Sümegei *et al.*, 1982; Wieslander & Lendahl, 1983). As there is only one methionine residue per repeat, and as the methionine residues are tandemly arranged, one should expect CNBr peptides with lengths of 72 and 60 amino acid residues, corresponding to molecular masses of 7.8 and 6.5 kDa respectively. However, these values must be corrected, as it was recently shown that sp-Ib is both glycosylated and phosphorylated. The amount of carbohydrate and phosphate is 15.3 (Hertner *et al.*, 1983) and 8 molecules/100 amino acid residues (Galler *et al.*, 1984) respectively. Assuming an M_r of 169 as the mean value for the carbohydrate moiety, and 79 for the phosphate, increases the M_r values of the expected CNBr fragments to 10200 and 8500 respectively. Still there is a large difference between the expected and the recorded sizes. Possible reasons for the discrepancies between the estimations might be that the high degree of post-translatory modifications causes the fragments to migrate anomalously during electrophoresis in SDS/polyacrylamide gels. Experimental evidence in support of this suggestion has been obtained from CNBr cleavage of secretory proteins from the salivary glands of *C. thummi* (Serfling *et al.*, 1983). When the CNBr peptides were electrophoresed in a borate buffer system, which partly compensates for the lower SDS-binding capacity of glycoproteins (Poduslo, 1981), instead of the ordinary Laemmli system, the apparent M_r of the major peptide decreased from 22000 to 17000. Another possible explanation of the discrepancies is an incomplete cleavage due to partial occurrence of methionine sulphoxide in the starting material, or its formation during cleavage (Joppich-Kuhn *et al.*, 1982), since the sulphoxide is resistant to attack by CNBr (Gross, 1967). This seems, however, un-

likely, since there is no trace of methionine sulphoxide in the starting material and the cleavage is performed under acidic and reducing conditions.

Instead of the predicted number of two different sized CNBr peptides from sp-Ib, five to six were found. This points towards the possibility that there are three additional repeat lengths within the molecule(s), or that the post-translatory modifications are unevenly distributed along the protein chains.

Both silk proteins and secretory proteins from salivary gland have a limited number of their amino acids in a highly redundant fashion, as in other fibrous proteins such as collagen (Kefalides, 1966) and elastin (Sandberg *et al.*, 1981). For example, in fibroin, 87% is glycine, alanine and serine; in sericin, 70% is serine, glycine and aspartic acid (Garel, 1982); in secretory protein I, 56% is lysine, serine, proline and arginine. Moreover, many extracellular fibrous proteins are also repetitive, for example fibroin (Gage & Manning, 1980), sericin (Okamoto *et al.*, 1982), collagen (Wozney *et al.*, 1981) and elastin (Sandberg *et al.*, 1981). The reason why collagen has a repetitive substructure is known, i.e. it must have a glycine residue in every third position to be able to form a stable superhelix, but why the two largest secretory proteins from salivary gland of *Chironomus* are repetitive is unknown. Explanations may be found in their evolutionary history and/or functional demands and these questions might be answered when additional sequence analyses of similar fibrous proteins and their corresponding genes are available.

I thank Professor Hans Jörnvall for advice regarding the isolation of the proteins for amino acid analysis and Mrs. Carina Palmberg who performed the analyses. Professor J.-E. Edström and Dr. Ulf Friberg provided criticism of the manuscript. This investigation was supported from the Funds of the Karolinska Institutet and by the Swedish Medical Research Council (13X-3532).

References

- Daneholt, B. (1972) *Nature (London) New Biol.* **240**, 229–232
- Degelmann, A. & Hollenberg, C. P. (1981) *Chromosoma* **83**, 295–313
- Edström, J.-E., Rydlander, L. & Francke, C. (1980) *Chromosoma* **81**, 115–124
- Gage, L. P. & Manning, R. F. (1980) *J. Biol. Chem.* **255**, 9444–9450
- Galler, R., Rydlander, L., Riedel, N., Kluding, H. & Edström, J.-E. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1448–1462
- Garel, J.-P. (1982) *Trends Biochem. Sci.* **7**, 105–108
- Gross, E. (1967) *Methods Enzymol.* **11**, 238–255
- Grossbach, U. (1969) *Chromosoma* **28**, 136–187

- Grossbach, U. (1977) in *Results and Problems in Cell Differentiation*, vol. 8 (Beermann, W., ed.), pp. 147–196, Springer-Verlag, Berlin
- Hager, D. A. & Burgess, R. R. (1980) *Anal. Biochem.* **109**, 76–86
- Hertner, T., Meyer, B., Eppenberger, H. M. & Mähre, R. (1980) *Wilhelm Roux's Arch. Dev. Biol.* **189**, 69–72
- Hertner, T., Eppenberger, H. M. & Lezzi, M. (1983) *Chromosoma* **88**, 194–200
- Hirs, C. H. W. (1967) *Methods Enzymol.* **11**, 198–199
- Joppich-Kuhn, R., Corkill, J. A. & Giese, R. W. (1982) *Anal. Biochem.* **119**, 73–77
- Kefalides, N. A. (1966) *Biochem. Biophys. Res. Commun.* **22**, 26–32
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Lambert, B. (1972) *J. Mol. Biol.* **72**, 65–75
- Okamoto, H., Ishikawa, E. & Suzuki, Y. (1982) *J. Biol. Chem.* **257**, 15192–15199
- Poduslo, J. F. (1981) *Anal. Biochem.* **114**, 131–139
- Rydlander, L. & Edström, J.-E. (1980) *Chromosoma* **81**, 85–99
- Rydlander, L., Pigon, A. & Edström, J.-E. (1980) *Chromosoma* **81**, 101–113
- Sandberg, L. B., Soskel, N. T. & Leslie, J. G. (1981) *N. Engl. J. Med.* **304**, 566–579
- Serfling, E., Meyer, L., Rudolph, A. & Steiner, K. (1983) *Chromosoma* **88**, 16–23
- Sümegei, J., Wieslander, L. & Daneholt, B. (1982) *Cell* **30**, 579–587
- Weber, F., Mähr, R., Meyer, B., Eppenberger, H. M. & Lezzi, M. (1983) *Wilhelm Roux's Arch. Dev. Biol.* **192**, 200–203
- Weeke, B. (1973) *Scand. J. Immunol.* **2**, Suppl. 1, 26–33
- Wieslander, L. (1979) *J. Mol. Biol.* **134**, 347–367
- Wieslander, L. & Lendahl, U. (1983) *EMBO J.* **2**, 1169–1175
- Wieslander, L., Sümegei, J. & Daneholt, B. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6956–6960
- Wozney, J., Hanahan, D., Tata, V., Boedtker, H. & Doty, P. (1981) *Nature (London)* **294**, 129–135