Comparative Biochemistry of Chondroitin Sulphate-Proteins of Cartilage and Notochord

By MARTIN B. MATHEWS

Departments of Pediatrics and Biochemistry, The La Rabida—University of Chicago Institute, and the Joseph P. Kennedy Jr. Mental Retardation Research Center, The University of Chicago, Chicago, Ill. 60637, U.S.A.

(Received 27 April 1971)

Fragments that consisted mainly of two polysaccharide chains joined by a short polypeptide bridge (doublets) were prepared from chondroitin sulphate-proteins of lamprey, sturgeon, elasmobranch and ox connective tissues after hydrolysis with trypsin and chymotrypsin. Consideration of molecular parameters, compositions and behaviour on gel electrophoresis and density-gradient fractionation leads to a proposed parent structure for chondroitin sulphate-proteins. A single polypeptide chain of about 2000 amino acid residues contains alternating short and long repeating sequences. A short sequence consists of less than 10 amino acid residues with one *N*-terminal and one *C*-terminal serine residue, each of which carries a polysaccharide chain linked glycosidically to its hydroxyl group. This structure constitutes the doublet subunit. Some variation is introduced when the doublet subunit carries only a single polysaccharide chain. The long sequence contains about 35 amino acid residues and is subject to cleavage by trypsin and chymotrypsin. The main polypeptide is probably homologous in the vertebrate sub-phylum with strong conservation of structure suggested for the short sequence. However, polymorphism of polypeptide structures cannot be excluded.

Mathews & Lozaityte (1958) proposed a comb-like model for the chondroitin sulphate-protein macromolecules of bovine hyaline cartilage. The basic unit consisted of an extended protein backbone to which side chains of acidic polysaccharide were covalently bound at intervals to form a macromolecule with a molecular weight of about 4×10^6 . It was subsequently shown that each linear polysaccharide chain was linked glycosidically, via a linkage region containing glucuronic acid, galactose and xylose, to the hydroxyl group of a serine residue (Rodén, 1968). Electron-microscopic studies support a branched structure for bovine chondroitin sulphate-protein (Serafini-Fracassini & Smith, 1966; Rosenberg, Hellmann & Kleinschmidt, 1970).

To account for the observation (Mathews, 1956) that chondroitin sulphate prepared by alkaline extraction of cartilage had a molecular weight about one-half of that of preparations obtained by trypsin digestion of chondroitin sulphate-protein, Anderson, Hoffman & Meyer (1965) suggested that two 'single' chains of chondroitin sulphate were joined by a short peptide sequence as 'doublets' containing an alkali-labile bond. Lusecombe & Phelps (1967) found that the weight-average molecular weight (43 300) of the product obtained by trypsin digestion was decreased to 21 000 by the action of papain or alkali, and thus supported this proposal.

The present paper deals with the composition and structure of doublets obtained by the action of both trypsin and chymotrypsin. Number-average molecular weights (M_n) were obtained so that amino acid contents could be related reliably to numbers of molecular chains under conditions of some polydispersity of polysaccharide chain weight. Since chondroitin sulphate-protein molecules in cartilage of many vertebrates are very similar in chemical composition, and in molecular size parameters (Mathews, 1967), it was decided to investigate the extent of homology at the doublet level of structure.

MATERIALS AND METHODS

Isolation of chondroitin sulphate-protein. The preparations of chondroitin sulphate-protein from ox nasal septal cartilage (SCSP4E) and from ray cartilage (W-2B) were described previously (Mathews & Lozaityte, 1958; Mathews, 1962). The physical constants for the preparations of chondroitin sulphate-protein from cartilage and notochord of lamprey and of sturgeon have been published (Mathews, 1967). However, the procedure for isolation from tissue of the latter preparations was sufficiently modified from the procedure employed earlier to merit a brief description.

About 10 g of tissue was homogenized with a Virtis blender in 200 ml of water, kept cold by an ice bath. Four

periods of blending at high speed of 5 min each were alternated with periods at low speed of 30 min each. The homogenate was diluted with 1.8 litres of 0.1 M-NaCl, heated within 10 min to 75–80°C and, after 1 min, was cooled to 45°C. The mixture was filtered with suction through a pad of Celite filter powder and aq. 4% (w/v) cetylpyridinium chloride was added to the filtrate to obtain complete precipitation. After 1 h at 30°C, the suspension was centrifuged at 2000g. The precipitate was stirred with 300 ml of 0.1 M-NaCl containing 0.1% cetylpyridinium chloride and recovered by centrifugation. To obtain the chondroitin sulphate-protein as the sodium salt, the precipitate was stirred at 40–50°C with a mixture of 160 ml of 2.5 M-NaCl and 80 ml of methanol for several hours and the resulting solution clarified by centrifugation at room temperature. The supernatant was diluted with 400 ml of 60% (v/v) ethanol and added with stirring to 1 litre of 90% ethanol. After setting overnight, the precipitate was recovered by centrifugation, washed successively with ethanol and ether and dried over P_2O_5 in a vacuum. A small content of hydroxyproline (less than 0.1%) was further decreased by the following steps. The precipitate, dissolved in 500 ml of 0.4 M-NaCl, was filtered through a Celite pad, diluted to 2 litres with water, re-precipitated with cetylpyridinium chloride and converted into the sodium salt as described above.

Enzyme treatments. A sample of 300 mg of chondroitin sulphate-protein was dissolved in 60 ml of 0.05 M-sodium phosphate buffer, pH 7.6, and placed in a Ubbelohde suspended-level viscometer in a water bath at 25.00°C. Three drops of toluene and 1.5 mg of crystalline trypsin (Calbiochem, Los Angeles, Calif., U.S.A.) were added. The viscosity of the solution decreased rapidly and reached a limiting value in 24 h. On addition of 1.5 mg of crystalline chymotrypsin (Calbiochem) a further fall in viscosity occurred for about 1 h. Subsequent changes in viscosity were small and became negligible after 48 h. After dialysis of the solution against 0.1 M-NaCl and precipitation with cetylpyridinium chloride, the product was recovered as a dry sodium salt. Further treatment as described above with 1.5 mg of chymotrypsin for 24 h produced changes in reduced viscosity of less than 3%. The recovered product in aqueous solution was then adsorbed on to a column of 250 ml of Dowex 1 resin (X2; 200–400 mesh; Cl^- form) and the resin eluted successively with 400–500 ml batches of water, 1.0 M-NaCl, 1.5 M-NaCl, 1.7 M-NaCl and 5.0 M-NaCl. The bulk of the product, recovered as dry sodium salt, was eluted by 1.5 M- and 1.7 M-NaCl. These two fractions represented a yield of 70–90% of the hexosamine of chondroitin sulphate-protein, and were substantially free of keratan sulphate (Rosenblum & Cifonelli, 1967). The 1.7 M-NaCl fractions from preparations of ox, sturgeon and ray cartilage (60–70% yield) were used further. The 1.5 M- and 1.7 M-NaCl fractions from preparations of lamprey tissues and of sturgeon notochord were combined in each case.

Hydrolyses with crystalline papain (Nutritional Biochemicals Corp., Cleveland, Ohio, U.S.A.) were carried out at 60°C for 15 h with concentrations of 1 mg of tryptic-chymotryptic hydrolysate product and 0.1 mg of enzyme/ml of 0.1 M-sodium phosphate buffer, pH 6.5, containing 0.01 M-cysteine and 0.01 M-EDTA. The hydrolysate was dialysed against water, clarified by filtration, precipitated

with cetylpyridinium chloride and the product finally recovered as the sodium salt by alcohol precipitation.

Hydrolyses with Pronase (B grade; Calbiochem) were carried out at 60°C for 2 days with concentrations of 2 mg of tryptic-chymotryptic hydrolysate product and 0.07 mg of enzyme/ml of 0.1 M-sodium phosphate buffer, pH 7.6. The principal product was recovered as the sodium salt (see above).

Alkali treatment. A solution of 50 mg of tryptic-chymotryptic hydrolysate product in 0.1 M-NaCl was made 0.05 M in NaOH by addition of 2.5 M-NaOH and kept at 25°C for 24 h. The solution was neutralized with 1 M-acetic acid and the product precipitated with cetylpyridinium chloride and converted into the sodium salt. During the period of exposure to alkali, the reduced viscosity decreased rapidly, at about an exponential rate, to reach a limiting value in about 6 h, with about 30 min required to produce one-half of the final decrease in viscosity. Although similar limiting values of reduced viscosity were found after 24 h exposure to 0.01 M-NaOH, the rate of fall of viscosity was about one-sixth of that in 0.05 M-NaOH. No appreciable changes of solution viscosity were found after 24 h in the presence of 0.02 M- Na_2CO_3 (pH approx. 10.5).

Analyses. The hexosamine content of preparations was determined after hydrolysis of samples in 4 M-HCl for 14 h at 100°C by a modification (Boas, 1953) of the method of Elson & Morgan (1933) and corrected for losses by comparison with a standard sample run concurrently. An 'anhydrous' content of polysaccharide was then calculated from the hexosamine content and the structural formula for the disaccharide repeating unit (Mathews & Dorfman, 1953). Amino acids, galactosamine and glucosamine were determined with the Technicon Amino Acid Analyzer after hydrolysis of samples for 24 h at 100°C in 6 M-HCl. Control experiments with known mixtures of amino acids and chondroitin sulphate indicated that losses of individual amino acids were less than 5%. Hence the amino acid results were not corrected for losses. However, since amino sugar losses during hydrolysis were about 30%, corrections were made on the basis of the previously determined values of hexosamine content. Values of intrinsic viscosity and M_n , obtained by previously described procedures (Mathews & Dorfman, 1953), were based also on the 'anhydrous' polysaccharide content of preparations. Tryptophan was not determined, but appears to be present only in very small amounts in bovine chondroitin sulphate-protein (Mathews & Lozaityte, 1958).

Cystine was determined as cysteic acid after performic acid oxidation of samples by the method of Hirs (1967). Crystalline bovine albumin was used as a reference standard in each analysis in the presence of 10 times its weight of chondroitin sulphate to approximate the proportion of protein and polysaccharide in chondroitin sulphate-protein. After the oxidation step, the samples were hydrolysed for 24 h at 100°C in 6 M-HCl and evaporated repeatedly to remove HCl. The amino acid residues were then spotted on Whatman 3MM paper together with standards of cysteic acid; and electrophoresis was performed at 10 V/cm in a buffer that was prepared by mixing 6 ml of pyridine, 20 ml of acetic acid and 1 litre of water and adjusting to pH 3.9. After the paper was stained with

ninhydrin, the densities of the cysteic acid spots were measured with a Photovolt model 525 densitometer. The albumin reference standard gave values for cystine within 10% of theoretical.

Solutions of chondroitin sulphate-protein in 4.0M-guanidinium chloride-0.05M-acetate buffer, pH 5.8, and CsCl, of density 1.50 g/ml, were fractionated by the density-gradient procedure of Sajdera & Hascall (1969). A Spinco model L2-65 ultracentrifuge with a no. 65 head was operated for 48 h at 20°C and an average force of 75000g.

Electrophoresis. The disc-electrophoresis apparatus of the Buchler Instrument Co. was employed, with polyacrylamide gels made with different concentrations of monomer in sodium phosphate buffer, pH 7.5, I 0.13. The acidic polysaccharide bands were stained with Alcian Blue. The use of gels of different pore size permitted resolutions of homologous polymers based on molecular size (Mathews & Decker, 1971).

RESULTS

Homogeneity. Since preparations of chondroitin sulphate-protein may contain degradation products derived by the action of native proteolytic enzymes (Ali, 1964), they were examined by gel electrophoresis. Products with the molecular size of tryptic-chymotryptic hydrolysate products and lower-molecular-weight products were essentially absent. The tryptic-chymotryptic hydrolysate products from ray, ox and sturgeon cartilage

contained only trace amounts of chondroitin sulphate chains of the size of those obtained by papain hydrolysis of the respective chondroitin sulphate-protein preparations. However, the 1.5M-sodium chloride eluate fractions from the resin preparation of the tryptic-chymotryptic hydrolysate products from ox and sturgeon cartilage contained appreciable proportions (near 20% by visual estimate of stain in gel electrophoresis) of chondroitin sulphate of lower molecular size. Similar proportions of polysaccharide of low molecular weight were found in the tryptic-chymotryptic hydrolysate products from lamprey tissues and from sturgeon notochord.

Molecular parameters. Table 1 shows the molecular weights and the intrinsic viscosities of the products of hydrolysis of the chondroitin sulphate-protein with trypsin and chymotrypsin and of the products after further degradation by sodium hydroxide, papain and by Pronase. The M_n values for the products of sodium hydroxide and of papain treatments of the tryptic-chymotryptic hydrolysate products of ox, sturgeon and ray cartilage are close to one-half of the initial values. These results are consistent with additional scission of only one bond per molecule of the tryptic-chymotryptic hydrolysate products. The close agreement of M_n values of the tryptic-chymotryptic hydrolysate product and of the product from additional hydrolysis by

Table 1. *Molecular parameters of tryptic-chymotryptic hydrolysate of chondroitin sulphate-protein and products after various treatments*

M_v , viscosity-average molecular weight, was estimated from relationship $[\eta] = 3.1 \times 10^{-2} M_v^{0.74}$ for chondroitin sulphate A and $[\eta] = 5.8 \times 10^{-2} M_v^{0.74}$ for chondroitin sulphate C. Ox cartilage and the notochords of sturgeon and lamprey contain mainly chondroitin sulphate A; cartilage of the sturgeon contains primarily chondroitin sulphate C; cartilage of the ray contains an over-sulphated chondroitin sulphate C (Mathews, 1959, 1967).

Tissue	Molecular parameter	Treatment			
		—	NaOH	Papain	Pronase
Ox cartilage	$[\eta]$	102	60	60	96
	M_n	43000	21000	21800	—
	M_v	56000	27000	27000	51000
Sturgeon cartilage	$[\eta]$	215	135	130	200
	M_n	50000	25400	23000	—
	M_v	66000	35000	34000	61000
Ray cartilage	$[\eta]$	310	205	169	310
	M_n	80000	40100	40200	—
	M_v	110000	62000	52000	110000
Lamprey cartilage	$[\eta]$	38	29	28	27
	M_n	7000	5100	5200	—
	M_v	(6600)	(4600)	(4400)	(4300)
Sturgeon notochord	$[\eta]$	38	25	23	33
	M_n	10000	5900	7100	—
	M_v	15000	8100	7400	12000
Lamprey notochord	$[\eta]$	45	25	28	39
	M_n	11400	5900	6500	—
	M_v	18000	8300	9600	14600

Table 2. *Fractionation of chondroitin sulphate-protein in a caesium chloride density gradient*

The volumes and average densities at 20°C of zones were: A, 3 ml, 1.42; B, 5 ml, 1.47; C, 4 ml, 1.54. Protein in chondroitin sulphate-protein from each zone was calculated from amino acid analysis.

Source	Zone	Percentage of total		Protein in chondroitin sulphate-protein from each zone (%)
		Protein	Chondroitin sulphate	
Ox cartilage	A	10	0	—
	B	20	<5	—
	C	70	>95	10.4
Sturgeon cartilage	A	10	0	—
	B	10	<4	—
	C	80	>96	7.0
Ray cartilage	A	10	0	—
	B	10	<4	—
	C	80	>96	5.3
Lamprey cartilage	A	20	7	71
	B	30	30	46
	C	50	63	32
Sturgeon notochord	A	40	6	69
	B	40	30	30
	C	20	64	21
Lamprey notochord	A	30	17	58
	B	50	47	52
	C	20	36	40

papain for the ox preparations with the weight-average molecular weights of 43300 and 21000, respectively for comparable preparations obtained by Luscombe & Phelps (1967) suggests a high degree of molecular size homogeneity. Evidently, the cleavage of the tryptic-chymotryptic products divides them into two macromolecular fragments of approximately equal size. It is likely that similar statements apply to the homologous sturgeon and ray molecules. The values of the viscosity-average molecular weight (M_v) support these conclusions.

The results for the remaining three tryptic-chymotryptic hydrolysate products shown in Table 1 are not amenable to a simple interpretation. Although the products from notochord of sturgeon and of lamprey and from cartilage of lamprey are similarly degraded by alkali and by papain, $\bar{M}_n$ values are decreased by less than 50%. Also, Pronase shows an appreciable effect on $[\eta]$ of all three of these tryptic-chymotryptic hydrolysate products, in contrast to a lack of effect on the first three preparations in Table 1. The values of M_v are only approximations, since they have been calculated from intrinsic-viscosity values outside the range of the values used to establish the empirical relationships.

Composition of chondroitin sulphate-protein. For preparations of chondroitin sulphate-protein from cartilage of ox, sturgeon and ray, density-gradient fractionation (Table 2) achieved separation of extraneous protein with virtually complete concentration of the polysaccharide content in the

densest fraction. Analyses are given for these fractions (zone C) in Table 3, computed for assumed doublets (two average polysaccharide chains). These computations were based on values of M_n of the products derived from alkali treatments of tryptic-chymotryptic hydrolysate products (Table 1). The gram-molecular weight for the parent disaccharide, which is defined (Mathews, 1967) as consisting of one residue of glucuronic acid, one residue of acetylgalactosamine and one residue of ester sulphate, was 503 g; for the over-sulphated ray polysaccharide (Mathews, 1962) the corresponding value was 515 g. Relatively small proportions of the total protein in each tube were contained in zones of low density. These results differed from those of Hascall & Sajdera (1969) and may be attributed to a comparatively low content of contaminating protein in the preparations.

Considerable variation of protein content with zone density is shown by the results in Table 2 for chondroitin sulphate-protein fractions of the remaining preparations. Zone B as well as zone C was analysed but analytical results for zone B fractions only are given in Table 3. The corresponding zone C fractions have lower contents of total amino acid residues per doublet, as indicated by results on protein contents of the chondroitin sulphate-protein of each zone. Although the protein of a zone C fraction was similar in amino acid composition to the protein of the corresponding zone B fraction, the differences in content for some amino acids, especially cystine, clearly indicated

Table 3. *Molar composition per doublet of chondroitin sulphate-protein fractions from density-gradient centrifugation*

Residue	Ox cartilage zone C	Sturgeon cartilage zone C	Ray cartilage zone C	Lamprey cartilage zone B	Sturgeon notochord zone B	Lamprey notochord zone B
Galactosamine	86	100	155	20	23	23
Glucosamine	6	5	5	0.2	1.0	0.0
Serine	5.3	4.1	4.0	10.8	4.2	8.2
Glycine	5.4	4.6	6.2	8.0	4.8	6.2
Glutamic acid	5.3	4.4	6.2	15.3	8.1	24.8
Aspartic acid	3.2	2.7	3.9	7.0	4.2	9.3
Proline	3.3	1.7	5.2	5.4	4.5	10.6
Threonine	2.3	1.8	2.4	4.7	3.5	12.0
Alanine	3.2	2.3	3.2	5.2	3.5	7.6
Valine	2.2	2.1	2.8	5.2	2.5	3.7
Isoleucine	1.5	3.6	3.8	3.3	1.6	8.1
Leucine	3.4	1.4	2.4	3.5	3.3	8.2
Tyrosine	0.7	2.5	0.8	0.0	1.3	3.2
Phenylalanine	1.5	0.8	1.2	1.2	1.3	2.3
Lysine	1.5	1.0	1.8	2.2	2.2	2.5
Histidine	0.7	0.6	0.5	4.9	1.4	4.6
Arginine	1.8	1.0	1.9	3.2	1.9	3.6
Half-cystine	1.0	0.6	0.9	0.5	0.9	1.0
Sum of amino acid residues	42	35	46	80	49	116

Table 4. *Molar composition of tryptic-chymotryptic hydrolysate of chondroitin sulphate-protein as chondroitin sulphate doublets*

Residue	Ox cartilage	Sturgeon cartilage	Ray cartilage	Lamprey cartilage	Sturgeon notochord	Lamprey notochord
Galactosamine	85	100	155	20	23	23
Glucosamine	0.7	0.8	1.9	0.2	0.2	
Serine	1.9	2.0	2.1	2.2	2.1	2.0
Glycine	1.9	2.3	2.4	2.2	2.3	2.0
Glutamic acid	1.3	1.0	1.9	1.9	2.6	3.5
Aspartic acid	0.5	0.8	1.1	0.6	0.8	1.1
Proline	0.9		2.8	0.5	1.7	1.1
Threonine	0.4	0.6	0.3	0.4	1.1	1.3
Alanine	0.6	0.6		0.7	1.5	0.8
Valine	0.6			0.9	1.3	0.4
Isoleucine	0.4	0.5	1.9	0.4	0.7	
Leucine	0.8	0.8	1.8	0.3	0.7	
Total amino acids	9.3	8.6	14.3	10.1	14.8	12.2

heterogeneity. The half-cystine contents of zone C fractions of chondroitin sulphate-protein from lamprey cartilage, sturgeon notochord and lamprey notochord were respectively 0.1, 0.2 and 0.4 mol/doublet.

Composition of tryptic-chymotryptic hydrolysate products. The amino acid compositions of the tryptic-chymotryptic hydrolysate products computed on the basis of chondroitin sulphate doublets are shown in Table 4. Several compositional features are common to all preparations, i.e.

approx. 2 mol each of serine and glycine, a minimum of 1 mol of glutamic acid or glutamine and a total amino acid content less than one-third of the number of amino acids per doublet in chondroitin sulphate-protein indicated in Table 3. Apparently, each serine residue is bound via a hydroxyl group to a polysaccharide chain.

Effects of enzyme and alkali treatments. The treatment of the trypsin-chymotrypsin hydrolysate product with alkali or with enzymes results in products containing polysaccharide with altered

Table 5. *Molar composition per doublet of products from treatment of tryptic-chymotryptic hydrolysate of chondroitin sulphate-proteins of ox, sturgeon and ray cartilages*

Residue Treatment ...	Ox cartilage			Sturgeon cartilage			Ray cartilage		
	NaOH	Pronase	Papain	NaOH	Pronase	Papain	NaOH	Pronase	Papain
Galactosamine	83	86	86	100	100	92	155	155	155
Glucosamine	0.7	0.7	0.4	1.0	1.0	1.0	1.9	1.9	0.8
Serine	0.7	2.0	2.2	0.3	2.0	1.8	0.8	2.2	1.9
Glycine	0.9	1.8	1.7	0.9	1.5	1.8	1.3	2.2	1.0
Glutamic acid	0.7	1.5	0.7	0.5	1.1	0.6	1.3	1.8	1.0
Aspartic acid	0.4	0.5	0.5	0.3	1.0	0.6	0.7	1.0	0.8
Proline	0.7	1.1	0.5				1.5	2.6	0.5
Threonine	0.2	0.3	0.3	0.3	0.7	0.2			0.3
Alanine	0.3	0.6	0.6	0.2	0.7				
Valine	0.3	0.6	0.5						
Isoleucine							2.1		
Leucine			0.5	0.4			1.2		

Table 6. *Molar composition per doublet of products from treatment of tryptic-chymotryptic hydrolysate of other chondroitin sulphate-proteins*

Residue Treatment ...	Lamprey cartilage			Sturgeon notochord			Lamprey notochord		
	NaOH	Pronase	Papain	NaOH	Pronase	Papain	NaOH	Pronase	Papain
Galactosamine	20	20	20	23	23	28	23	23	26
Glucosamine	0.2	0.2	0.2	0.2	0.2	0.3	0.0	0.0	0.0
Serine	1.8	2.2	2.0	0.6	1.9	1.7	0.4	1.8	1.7
Glycine	1.8	2.2	1.5	0.7	1.6	1.0	0.4	1.8	1.6
Glutamic acid	1.8	1.6	0.8	0.7	2.0	0.8	1.2	1.4	1.0
Aspartic acid	0.6	0.7	0.4	0.3	0.6	0.4	0.2	0.9	0.9
Proline	0.4	0.6	0.3	0.7	1.0	0.7	0.2	0.4	0.3
Threonine	0.4	0.4	0.3	0.7	0.9	0.7	0.6	0.5	0.3
Alanine	0.8	0.6	0.5	0.4	0.8	0.5	0.4	0.5	0.1
Valine	0.8	0.6	0.4	0.5	0.8	0.6	0.1	0.1	

amino acid compositions, as well as molecular parameters. Tables 5 and 6 show the compositions of these products. The structural interpretation of the compositional results, as was the case for the results on molecular parameters, is relatively simple for the products listed in Table 5. The products listed in Table 6 cannot be interpreted in the same way and are discussed below.

To account for the observation that papain bisects the 'doublet fraction', presumably by peptide-bond cleavage, one may assume the presence in this portion of the polypeptide chain of a minimum peptide sequence with one serine residue as the *N*-terminal group and a second serine residue as the *C*-terminal group. For purposes of brevity, this hypothetical peptide sequence between two serine residues, including the two terminal serine residues as well, will be referred to as sequence X. A polysaccharide chain is linked to the hydroxyl group of each serine residue. In the tryptic-chymotryptic hydrolysate products some amino acid residues may be present that are linked

to the serine residue and external to sequence X. Also, all of the amino acid residues in Table 4 may be components of sequence X, including those present in fractional molar amounts, since the extent of polymorphism of sequence X is unknown.

Papain digestion (Table 5) not only bisects the tryptic-chymotryptic hydrolysate product but also partially removes amino acid residues, except for serine. These results do not permit the assignment of any residue to sequence X with certainty. However, the retention of glycine suggests that at least one glycy residue is vicinal to the serine. Pronase digestion, which does not bisect the tryptic-chymotryptic hydrolysate product, removes primarily leucine and isoleucine, which are apparently external to sequence X. Consequently, a probable maximum of nine amino acid residues comprise sequence X.

Treatment of the tryptic-chymotryptic hydrolysate product with alkali may be expected to result in β -elimination of the polysaccharide chains from serine residues, accompanied by conversion

of the serine residues into dehydroalanine residues provided that both the carboxyl group (Anderson *et al.* 1965) and the amino group (Bella & Danishefsky, 1968) of each serine residue are substituted. Alkali treatment of the tryptic-chymotryptic hydrolysate products shown in Tables 5 and 6, except for the preparation from lamprey cartilage, produces a loss of slightly more than one serine residue and about one glycine residue. It would therefore appear that at least one serine residue of sequence X in each molecule of the tryptic-chymotryptic hydrolysate products is substituted at both amino and carboxyl positions. However, the recovery of polysaccharide containing some serine shows that an appreciable amount of serine is also in a terminal position of peptide sequences. In addition, the supernatant solution obtained after precipitation of alkali-treated tryptic-chymotryptic hydrolysate products with cetylpyridinium chloride is free of amino sugar, but contains amino acids including 0.1–0.2 mol of serine per doublet. The latter result indicates the presence in the tryptic-chymotryptic hydrolysate products of some serine that is not glycosylated.

DISCUSSION

Structure of chondroitin sulphate-protein. Fig. 1 summarizes the observations on the tryptic-

chymotryptic hydrolysate products and suggests a structure for chondroitin sulphate-protein. This elementary schematic formulation assumes the presence of a single polypeptide chain to which side chains of chondroitin sulphate of uniform length are attached at regular intervals; it will be utilized here as a parent model structure.

Structure of the tryptic-chymotryptic hydrolysate product. The hypothesis that the product of tryptic and chymotryptic hydrolysis of chondroitin sulphate-protein contains a polypeptide sequence X, comprising less than ten amino acid residues, with polysaccharide chains bound glycosidically to the hydroxyl groups of two serine residues, is in accord with the results for preparations from cartilage of ox, sturgeon and ray. The persistence of a glutamate residue in association with polysaccharide chains after various treatments of the tryptic-chymotryptic hydrolysate products (Tables 5 and 6) suggests that at least one such residue is close to a serine residue and is possibly included in sequence X. Indeed, a peptide of the structure Glu-Gly-dehydroAla-Gly has been isolated by Katsura & Davidson (1966) from a mixture of peptides resulting from alkali treatment of a papain digest of porcine chondroitin sulphate-protein. However, the results of alkali treatment of the tryptic-chymotryptic hydrolysate products are incompatible with the assignment of this sequence about both serine residues of the doublet. This is so since the observed β -elimination

Fig. 1. A segment of the proposed structure for a chondroitin sulphate-protein molecule which contains a single polypeptide chain of 2000–3000 amino acid residues to which side chains of chondroitin sulphate are attached. In the simplest, parent, structure the polypeptide chain consists of alternating X and Y sequences and every X sequence bears two polysaccharide side chains.

reaction would require minimally either the sequence -Gly-Ser-Gly-An-Gly-Ser- or the sequence -Ser-Gly-An-Gly-Ser-Gly-, where An represents other amino acid residues of sequence X. Thus if each serine residue had been flanked by two glycine residues in chondroitin sulphate-protein, then the tryptic-chymotryptic hydrolysate product should have contained minimally 3 mol of glycine instead of the assayed value of 2 mol per doublet.

Polypeptide structure. Considerations of the reduction of protein contents from starting material observed in zone C after density sedimentation requires a revision of the molecular weight for the basic molecular units of chondroitin sulphate-protein from bovine cartilage (Mathews & Lozaityte, 1958) and from chondrichthye (Mathews, 1962) to about 3.5×10^6 . Hascall & Sajdera (1970) obtained a mean molecular weight of 2.5×10^6 for the bovine chondroitin sulphate-protein. Therefore the polypeptide chains contain 2000–3000 amino acid residues per molecule of chondroitin sulphate-protein. Of the many possible different ways of visualizing the main polypeptide chain in terms of subunit sequences only the simplest ones will be considered. Although Serafini-Fracassini, Peters & Floreani (1967) had suggested that peptide subunits may be assembled in a random manner, there is no evidence for covalent linkages other than peptide bonds. Accordingly, the polypeptide chain of a molecule of chondroitin sulphate-protein is assumed to be a single biosynthetic entity whose structure is determined by a single structural gene. Division of the chain into subunits is solely for descriptive purposes.

The structure shown in Fig. 1 may be composed of only two types of subunits in regular order so that the polypeptide chain is represented as a repeating polymer (-X-Y-)_n. Alternatively, the order of X and Y sequences may not be regular, or more than two types of sequences may be present. The available results do not warrant further detailed speculation at this moment. The situation is complicated further by the presence in chondroitin sulphate-protein preparations of variable proportions of keratan sulphate-peptides, in part covalently linked to the chondroitin sulphate-protein molecule (Rosenblum & Cifonelli, 1967; Tsiganos & Muir, 1967).

It is evident from the variation in composition of chondroitin sulphate-proteins with buoyant density that the molecules are heterogeneous. Tsiganos, Hardingham & Muir (1971) showed that chondroitin sulphate chains were of uniform size in comparable fractions from pig laryngeal cartilage and concluded that the heterogeneity was due to the presence of different polypeptide chains differing in length and also in the type, distribution and number of carbohydrate chains attached. However, Hascall, Riolo, Hayward & Reynolds (1971) showed

that the polypeptide chains of chondroitin sulphate-proteins of bovine nasal cartilage were of closely similar size and composition differing primarily in the number of keratan sulphate chains attached. These findings are compatible with the presence generally of more than one type of polypeptide chain and of a variable extent of heterogeneity of chondroitin sulphate-proteins dependent on tissue, metabolic state etc.

The different polypeptide chains, in accord with the results of the present paper, may be similar to each other in amino acid composition and to some degree in amino acid sequences. Their similarities in structure and in structural relationships to polysaccharide chains suggests that these polypeptides may be polymorphically related. The presence of polymorphic polypeptides could account not only for the small variations in amino acid composition of the protein of fractions of different buoyant densities but also for the presence of non-integral amino acid residues in the doublet fractions shown in Table 4. However, such non-integral values may also arise as a consequence of incomplete proteolysis, as indicated by the action of Pronase on doublets.

The degree of difference in amino acid sequence between polymorphic or homologous proteins is usually variable along the polypeptide chains with the smallest number of residue differences occurring generally in regions associated with strict functional requirements. Since X sequences are involved in linkage to carbohydrate we might expect little variation of structure in this region of the polypeptide chains of molecules from the same tissue. Thus in chondroitin sulphate-protein (J. Baker & L. Rodén, unpublished work), as in the structurally related dermatan sulphate-protein molecule (Bella & Danishefsky, 1968; Lindahl & Rodén, 1971), a glycine residue is linked to the carboxyl group of serine residues that bear a chondroitin sulphate chain. As a consequence, if the amino terminus of the peptide chain is chosen to lie to the left of the sequence shown in Fig. 1, then both B and B' are glycine residues. The probable presence of at least one glutamate residue in sequence X has already been noted.

Polysaccharide chains. Chondroitin sulphate-protein preparations (Table 3) contain a minimum of two serine residues per doublet that are not involved in linkage to chondroitin sulphate chains. Thus it is possible that glycosylation of serine residues may be limited to specific regions of the polypeptide chain by requirements for enzyme specificity or by steric factors. These conditions set an upper limit to the number of chondroitin sulphate chains that may be linked to a given polypeptide. Also, gel electrophoresis showed the presence of single chondroitin sulphate chains as

well as doublets in tryptic-chymotryptic hydrolysate products, particularly those from lamprey tissues and from sturgeon notochord. In addition, the relative proportions of chondroitin sulphate and protein vary in zone B and zone C fractions.

To explain these observations we may consider the hypothetical examples of chondroitin sulphate-protein structure of Fig. 2. The structures shown for zone B reflect a chondroitin sulphate/polypeptide weight ratio lower than that of the structure for zone C, in accordance with the lower sedimentation density. Since the intrinsic viscosities of chondroitin sulphate isolated by papain digestion of material in zones A, B and C of preparations from sturgeon notochord and from lamprey tissues are virtually identical, the chain sizes of chondroitin sulphate must be the same. Hence structure 1 of zone B cannot be of major significance. It is probable therefore that structures similar to those of types 2 and 3 are present in regions of the chondroitin sulphate-protein molecule, particularly in chondroitin sulphate-protein from lamprey tissues. Fig. 1 shows that the two serine residues of the doublet are distinguishable and that one of these residues may be preferentially glycosylated. The particular examples in Fig. 2 were selected to account for the presence of single chondroitin

sulphate chains of uniform size in chondroitin sulphate-protein and should not imply (see above) that the polypeptide chain necessarily consists of X and Y sequences in regular alternating order.

A similar type of variation in numbers of chains may account for the observed variation in peptide content, mean sedimentation density and molecular weight of fractions of chondroitin sulphate-protein from heart valves (Lowther, Preston & Meyer, 1970). Hascall & Sajdera (1970) observed similar variations in these characteristics of fractions of chondroitin sulphate-protein from bovine cartilage. With the exception of half-cystine, amino acids were present in essentially the same relative proportions in all fractions. Hascall & Sajdera (1970) suggested that polydispersity in molecular weight of bovine chondroitin sulphate-protein was due primarily to differences in the number of chondroitin sulphate chains attached to one type of protein. However, Serafini-Fracassini (1968) obtained results suggesting that bovine chondroitin sulphate-protein contains three different polypeptide chains. Also, Brandt & Muir (1971) found appreciable polydispersity in size of chondroitin sulphate chains from chondroitin sulphate-proteins of pig cartilage.

Phylogeny and ontogeny. The hypothesis that the chondroitin sulphate-protein molecules from different vertebrate species form an homologous class is supported indirectly by structural data showing the occurrence of very similar, if not identical, polysaccharide chains bound to polypeptide in an identical fashion (Mathews, 1967). Strong evidence for this hypothesis, however, requires the demonstration that the polypeptide structures, which are referable directly via the genetic code to structural genes, are also very similar. My present results on the short peptide sequence of the doublet are indicative of homology and also suggest conservation over an extended period of evolution. Immunological results, which show the localization of a common antigenic determinant in the region of the short sequence, are consistent with this conclusion (Sandson, Damon & Mathews, 1970). The general similarity in number and type of amino acid residues composing the long polypeptide sequences of chondroitin sulphate-protein molecules of ox, sturgeon, ray and embryonic chick (Z. Nevo, unpublished work) cartilage suggests homology for this section of the chemical structure as well.

The chemistry of chondroitin sulphate-protein of vertebrates varies with stage of embryonic development and maturation in extent and site of sulphation of the amino sugar residues of carbohydrate chains (Mathews, 1967). Concomitant variation in the extent of glycosylation of serine residues may also occur. In addition, the possible presence of variable proportions of two or more

Fig. 2. Some hypothetical and schematic structures of representative segments of chondroitin sulphate-protein molecules illustrating variation in number and arrangement of polysaccharide chains attached to a given polypeptide sequence.

types of polypeptide, differing mainly in half-cystine content, may be of particular significance, since this residue is involved in aggregation of chondroitin sulphate-protein molecules (Sajdera, Hascall, Gregory & Dziewiatkowsky, 1970). All such variations in molecular structure influence interactions of chondroitin sulphate-protein with collagen and hence affect the molecular organization and properties of connective tissues (Mathews, 1970; Öbrink & Wasteson, 1971).

I am indebted to Miss LaVerne Decker for able laboratory assistance and to Dr Lennart Rodén of the University of Chicago for constructive criticism of the manuscript. This work was supported by Grants AM-05996 and HD-04583 of the U.S. Public Health Service and by a grant from the Chicago Heart Association.

REFERENCES

- Ali, S. Y. (1964). *Biochem. J.* **93**, 611.
- Anderson, B., Hoffman, P. & Meyer, K. (1965). *J. biol. Chem.* **240**, 156.
- Bella, A., jun. & Danishefsky, I. (1968). *J. biol. Chem.* **243**, 2660.
- Boas, N. F. (1953). *J. biol. Chem.* **204**, 553.
- Brandt, K. D. & Muir, H. (1971). *Biochem. J.* **121**, 261.
- Elson, L. A. & Morgan, W. T. J. (1933). *Biochem. J.* **27**, 1824.
- Hascall, V., Riolo, R., Hayward, J. & Reynolds, C. (1971). *Fedn Proc. Fedn Am. Socs exp. Biol.* **30**, 1278.
- Hascall, V. C. & Sajdera, S. W. (1969). *J. biol. Chem.* **244**, 2384.
- Hascall, V. C. & Sajdera, S. W. (1970). *J. biol. Chem.* **245**, 4920.
- Hirs, C. H. W. (1967). In *Methods in Enzymology*, vol. 11, p. 197. Ed. by Hirs, C. H. W. New York: Academic Press.
- Katsura, N. & Davidson, E. A. (1966). *Biochim. biophys. Acta*, **121**, 120.
- Lindahl, U. & Rodén, L. (1971). In *Glycoproteins*, 2nd ed. Ed. by Gottschalk, A. Amsterdam: Elsevier (in the Press).
- Lowther, D. A., Preston, B. N. & Meyer, F. A. (1970). *Biochem. J.* **118**, 595.
- Luscombe, M. & Phelps, C. F. (1967). *Biochem. J.* **103**, 103.
- Mathews, M. B. (1956). *Archs Biochem. Biophys.* **61**, 367.
- Mathews, M. B. (1959). *Biochim. biophys. Acta*, **35**, 9.
- Mathews, M. B. (1962). *Biochim. biophys. Acta*, **58**, 92.
- Mathews, M. B. (1967). *Biol. Rev.* **42**, 499.
- Mathews, M. B. (1970). In *Chemistry and Molecular Biology of the Intercellular Matrix*, vol. 2, p. 1155. Ed. by Balazs, E. A. London: Academic Press.
- Mathews, M. B. & Decker, L. (1971). *Biochim. biophys. Acta*, **244**, 30.
- Mathews, M. B. & Dorfman, A. (1953). *Archs Biochem. Biophys.* **42**, 41.
- Mathews, M. B. & Lozaityte, I. (1958). *Archs Biochem. Biophys.* **74**, 158.
- Öbrink, B. & Wasteson, Å. (1971). *Biochem. J.* **121**, 227.
- Rodén, L. (1968). In *Biochemistry of Glycoproteins and Related Substances*, p. 185. Ed. by Rossi, E. & Stoll, E. New York: S. Karger.
- Rosenberg, L., Hellmann, W. & Kleinschmidt, A. K. (1970). *J. biol. Chem.* **245**, 4123.
- Rosenblum, E. L. & Cifonelli, J. A. (1967). *Fedn Proc. Fedn Am. Socs exp. Biol.* **26**, 282.
- Sajdera, S. W. & Hascall, V. C. (1969). *J. biol. Chem.* **244**, 77.
- Sajdera, S. W., Hascall, V. C., Gregory, J. D. & Dziewiatkowsky, D. D. (1970). In *Chemistry and Molecular Biology of the Intercellular Matrix*, vol. 2, p. 851. Ed. by Balazs, E. A. London: Academic Press.
- Sandson, J., Damon, H. & Mathews, M. B. (1970). In *Chemistry and Molecular Biology of the Intercellular Matrix*, vol. 3, p. 1563. Ed. by Balazs, E. A. London: Academic Press.
- Serafini-Fracassini, A. (1968). *Biochim. biophys. Acta*, **170**, 289.
- Serafini-Fracassini, A., Peters, T. J. & Floreani, L. (1967). *Biochem. J.* **105**, 569.
- Serafini-Fracassini, A. & Smith, J. W. (1966). *Proc. R. Soc. B*, **165**, 440.
- Tsiganos, C. P., Hardingham, T. E. & Muir, H. (1971). *Biochim. biophys. Acta*, **229**, 529.
- Tsiganos, C. P. & Muir, H. (1967). *Biochem. J.* **104**, 26c.