Carbohydrate Research, 255 (1994) 41–59
Elsevier Science B.V., Amsterdam

Preliminary investigation into the purification, NMR analysis, and molecular modelling of chondroitin sulphate epitopes

Sinéad M.T. D'Arcy *, Stephen L. Carney and Trevor J. Howe

Lilly Research Centre Ltd., Eli Lilly and Co., Erl Wood Manor, Windlesham, Surrey GU20 6PH (United Kingdom)

(Received March 30th, 1993; accepted July 1st, 1993)

ABSTRACT

Electrophoretic methods are outlined for the rapid purification and analysis of chondroitin sulphate oligosaccharides on a milligram scale. Isomeric impurities however, exist within specific sized oligosaccharides. Detailed ^{1}H and ^{13}C NMR data for the chondroitin sulphate disaccharides Δ^4HexA(1-3)GalNAc6SO$_3^-$ and Δ^4HexA(1-3)GalNAc4SO$_3^-$ (prepared from shark chondroitin sulphate) are reported. Two-dimensional NMR methods have been employed in the assignment of spectra. Preliminary models of these disaccharides are proposed from molecular mechanics conformational searching and molecular dynamics procedures. This study provides ^{1}H and ^{13}C NMR reference data which will be useful in the investigation of larger chondroitin sulphate oligosaccharides, such as those prepared by the electrophoretic methods mentioned above.

INTRODUCTION †

Glycosaminoglycans (GAGs) are ubiquitous components of all connective tissue extracellular matrices where they serve a number of functions. The most studied aspect of GAG function relates to their space filling properties, particularly in cartilaginous tissues. Due to the high charge conferred on these molecules by virtue of the carboxylate and sulphate ester groups, GAGs (usually in the form of proteoglycans) can generate large Donnan osmotic pressures, which are essential for the function of tissues like cartilage. The most abundant class of GAGs is the

* Corresponding author.

† Abbreviations: Glycosaminoglycans, GAGs; ΔHexA-GalNAc6S/4S, derivatives of 2-acetamido-2-deoxy-3-*O*-(α-L-*threo*-hex-4-enopyranosyluronic acid)-D-galactose bearing a sulphate ester group at position 6 or position 4 of the hexosamine residue; COSY, correlated spectroscopy; HMQC, heteronuclear multiple quantum coherence; HMBC, heteronuclear multiple bond correlation; MBHSQC, multiple bond heteronuclear single quantum coherence; NOESY, nuclear Overhauser enhancement spectroscopy; ROESY, rotating-frame nuclear Overhauser enhancement spectroscopy; CE, capillary electrophoresis; ABNR, adopted-basis Newton Raphson.

SSDI 0008-6215(94)E0378-E

chondroitin sulphates[1]. The basic repeating unit of chondroitin is a disaccharide consisting of a glucuronic acid residue linked β(1-3) to an *N*-acetylgalactosamine residue. The polysaccharide is linear and unbranched. Diversity in the structure of the chondroitin arises as a result of the degree of sulphation and position of the sulphate esters.

Recent work[2] using monoclonal antibodies has demonstrated that chondroitin sulphate contains sequences within the chain which are almost certainly related to specific sulphation patterns. By analogy with the pentasaccharide anti-thrombin binding sequence of heparin[3] it has been proposed that such sequences in chondroitin sulphate may be responsible for binding molecules (for example, growth factors) within the extracellular matrix. Determination of the precise nature of these sequences is hampered by the lack of analytical techniques for the examination of chondroitin sulphate oligosaccharides. Disaccharides can be determined by methods such as HPLC[4] and more recently CE[5] but determination of oligosaccharides containing mixed disaccharide isomers is beyond the realms of either technique at present. In order to obtain definitive structural analysis of oligosaccharides, we must turn to alternative methodology, in this case ^{1}H and ^{13}C NMR techniques. This paper describes the parallel development of NMR, molecular modelling, and purification methods for chondroitin sulphate oligosaccharides as a prelude to investigation of structures, which contain the epitopes recognised by the monoclonal antibodies mentioned previously.

In particular, the ^{1}H and ^{13}C NMR spectra of *N*-acetylchondrosine [2-acetamido-2-deoxy-3-O-(β-D-glucopyranosyluronic acid)-D-galactose], ΔHexA-GalNAc6S and ΔHexA-GalNAc4S (see Fig. 1 for structures) have been fully assigned with recourse to two-dimensional NMR methods. ^{1}H NMR spectra were analysed using COSY and 1- and 2-step relayed-COSY experiments[6,7]. Alternatively, ^{13}C NMR spectra were assigned with the aid of the ^{1}H-detected (i.e., inverse) heteronuclear correlation or HMQC experiment[6,8,9]. The advantages of the HMQC experiment over conventional ^{13}C-detected heteronuclear correlation techniques are increased sensitivity (ca. eight-fold) and high resolution in the proton domain (i.e., ^{1}H–^{1}H coupling information is discernible). Furthermore the ^{1}H-detected long-range heteronuclear correlation (or HMBC[10]) experiment has facilitated the specific assignment of the anomeric protons of the ΔHexA residue

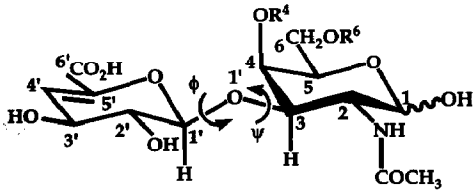

Fig. 1. Structure of the disaccharides ΔHexA-GalNAc6S ($R^4 = H$, $R^6 = SO_3^-$) and ΔHexA-GalNAc4S ($R^4 = SO_3^-$, $R^6 = H$).

in α and β anomers. These experiments have been successfully applied to ΔHexA-GalNAc4S/6S disaccharide samples available as mixed anomers in 1–3 mg quantities.

Molecular mechanics conformational searching and molecular dynamics experiments can give direct information on tertiary structure not available from any empirical source. Dynamic processes and solvent effects can be examined, indicating the possible role of charged groups and their function in determining tertiary structure. Of particular interest here is the function of the acetamido group and sulphate groups which are unique to heteropolymeric GAGs.

EXPERIMENTAL

Materials.—*N*-acetylchondrosine [2-acetamido-2-deoxy-3-O-(β-D-glucopyranosyluronic acid)-D-galactose] was prepared from chondrosine (Sigma Chemical Co.), 2-amino-2-deoxy-3-O-(β-D-glucopyranosyluronic acid)-D-galactose, as the barium salt. The unsaturated disaccharides of chondroitin 6-and 4-sulphate, ΔHexA-GalNAc4S/6S (1–3 mg mixed anomers), were prepared from shark chondroitin sulphate by digestion with chondroitinase ABC and purification by HPLC as reported previously[5]. Acrylamide, TEMED, and *N,N'*-methylenebisacrylamide (Electran grade) were purchased from BDH Ltd., Poole, UK. Gel filtration media (Sephadex) were obtained from Pharmacia, Milton Keynes, UK. Chondroitin lyases (ABC and ACII) were bought from ICN Flow, High Wycombe, UK. Water was HPLC grade and was obtained from FSA laboratory supplies, Loughborough, UK. All other reagents, unless otherwise stated, were of analytical grade.

Depolymerisation of chondroitin sulphate.—Treatment of chondroitin sulphate with the hydrolase bovine testicular hyaluronidase (EC 3.2.1.35) cleaves the polysaccharide chains into oligosaccharides of the form [GlcA(β1-3)GalNAc(β1-4)]$_n$GlcA(β1-3)GalNAc. The position of the sulphates on the galactosamines is variable. Using the following reaction conditions it was possible to produce a range of oligosaccharides varying in size from hexasaccharide to relatively large oligosaccharides (docosasaccharide and larger), each varying in size by a single disaccharide unit. Chondroitin sulphate A from bovine trachea (Sigma Chemical Co.) was dissolved in 0.1 M sodium acetate–0.15 M NaCl, pH 5.8, to a final concentration of 10 mg/mL. The solution was depolymerised by the addition of bovine testicular hyaluronidase (272 U/mL) for 16 h at 37°C.

Gel electrophoresis of depolymerised chondroitin sulphate.—Polyacrylamide gel electrophoresis was carried out essentially by the method of Cowman et al.[11]. Polyacrylamide:bisacrylamide (3.1% crosslinking) gels of various concentrations (details will be outlined in the optimisation section) were prepared in the following buffer: 90 mM tris–90 mM boric acid–2.4 mM EDTA, pH 8.3. Initiation and catalysis of the polymerisation reaction was with ammonium persulphate (22.5 mg in 30 mL) and 0.1% (v/v) TEMED. Gels (15 × 15 × 0.2 cm) were cast and run in

the tris–boric acid–EDTA buffer outlined previously at 13.33 V/cm. Samples were dissolved in tris–boric acid–EDTA running buffer and contained Bromophenol Blue and Phenol Red (both 0.001% w/v) as tracker dyes. Sucrose was added to a final concentration of 5% (w/v) to facilitate loading. The gel was run until the Phenol Red tracking dye had migrated 12 cm. The gel was stained for 45 min with 0.5% (w/v) Alcian Blue 8GX in 2% (v/v) acetic acid, and then destained with 2% (v/v) acetic acid until clear.

Preparative gel electrophoresis.—1. Optimisation of gel concentration. Optimisation was performed according to the manufacturer's instructions for the BioRad model 491 Prep Cell. The optimal gel concentration was 16%, although 20% gels were run to allow separation of the smallest oligosaccharides.

2. Running preparative gels. Gels were cast using the BioRad model 491 Prep Cell, according to manufacturers recommendations. The gel composition, running buffers, and sample preparation were exactly as outlined in the previous section. Gels were typically 10–12 cm in length and were run at ca. 25 V/cm. A typical gel run would take ca. 5 h to complete. Fractions (2 min fractions, ca. 2 mL) from the preparative gel were collected when the Phenol Red tracker dye had migrated to within 3 cm of the end of the gel rod. Collected fractions were assayed using the Dimethyl Methylene Blue (DMB) GAG-dye binding assay of Farndale[12] modified for microtitre plate analysis.

Capillary electrophoresis of oligosaccharides.—Essentially, the method used for the electrophoresis of oligosaccharides of chondroitin was a modification of a previous method used for the determination of chondroitin sulphate disaccharides at low pH. The conditions were as follows: 2 min wash in 0.1 M NaOH; 5 min rinse with 200 mM orthophosphoric acid, pH 3. Samples, containing in the order of 0.1 mg/mL, were introduced by vacuum (typically 1–2 s, ca. 4–8 nL). Electrophoresis was performed with an Applied Biosystems 270A apparatus at 40°C and −15 kV in 200 mM orthophosphoric acid, pH 3.

Column chromatography.—Oligosaccharides were chromatographed on Sephadex G-50 eluted with 20 mM ammonium formate. BioGel P2 eluted with distilled water was used for the desalting of small oligosaccharide fractions.

Analytical methods.—Oligosaccharides were assayed by the DMB assay of Farndale[12], adapted for microtitre plates. Uronic acid determinations were by the method of Bitter and Muir[13], again adapted for use in microtitre plates.

NMR spectroscopy.—Samples were prepared for NMR analysis by repeatedly treating with ^2H_2O (99.8 atom%^{2}H, Aldrich Chemical Co.) with intermediate lyophilisation and finally dissolving in 0.5 mL ^2H_2O. Spectra were acquired on a Bruker AC-300 spectrometer equipped with a 5 mm dual ^{1}H and ^{13}C inverse probehead. COSY and both l- and 2-step relayed COSY spectra were acquired in the magnitude mode using 1K x 256 data matrices which were zero-filled to 1K × 512 prior to Fourier transformation (Ft). For resolution enhancement in both dimensions, the time domain data were multiplied by a nonshifted sine bell window function. The mixing time, D2, was set to 30 ms for the 1-step relayed

COSY experiments and the mixing times, D2 and D3, were both set to 39 ms for the 2-step relayed COSY experiments.

Two-dimensional ^{1}H-detected $^{1}H-^{13}C$ heteronuclear correlation experiments were obtained using the standard Bruker microprogrammes for HMQC[14] and HMBC[10]. The HMQC programme employs a BIRD (bilinear rotational decoupling) pulse to suppress signals from protons not directly bonded to ^{13}C, time-proportional phase incrementation (TPPI) to allow for phase-sensitive handling of the data and ^{13}C decoupling during acquisition using the GARP-1 scheme. The microprogramme was modified to include a presaturation pulse, which serves to greatly reduce the intensity of the residual $HO^{2}H$ peak in $^{2}H_2O$ solutions. The spectral width in the ^{1}H domain was 840 Hz and 4560 Hz in the ^{13}C domain. Data sets (512×128) were acquired, which were zero-filled to 512×512 prior to Ft. A sine bell window function shifted by $\pi/2$ was applied in both dimensions before Ft, followed by phase correction of the data.

HMBC experiments were acquired with $4K \times 512$ data sets, which were zero-filled to $4K \times 1K$, followed by application of a $\pi/2$ shifted sine bell window function, prior to Ft. The spectral width in the ^{1}H domain was 900 Hz and 9056 Hz in the ^{13}C domain. The time delay (D4), during which the long-range coupling evolves, was set to 50 ms.

In all experiments, chemical shifts were referenced to external sodium 3-trimethylsilyl[$^{2}H_4$] propionate

Molecular modelling.—All modelling experiments were performed on either a Silicon Graphics 4D/320 GTX or Indigo R3000. QUANTA 3.3 and CHARMm 22 (with standard parameters)[15] were used throughout. No missing parameters were reported.

Six initial structures based on ΔHexA-GalNAc4S, ΔHexA-GalNAc6S and ΔHexA-GalNAc were built in pairs; a fully protonated form (SO_4 atom type) and an unprotonated form (sulphate and carboxy groups unprotonated, SO_4 atom type) carrying a unit negative charge of -2. These were minimised by adopted-basis Newton Raphson (ABNR) to an energy gradient tolerance of 0.001 and an energy value tolerance of 0.0. From these starting geometries grid conformational searches were executed about the O-5′–C-1′–O-1′–C-3 (ϕ) and C-1′–O-1′–C-3–C-2 (ψ) angles. Two types of grid conformational search were executed. Firstly a 'torsionally constrained' search at 10° intervals with the ϕ and ψ torsions fixed at each increment where side chain interactions were resolved by 1000 steps of ABNR at each angular change in starting conformation. Secondly, a fully relaxed search at 10° intervals without fixed ϕ and ψ torsions was executed. From each change in geometry all torsions were allowed to flex under 1000 steps of ABNR minimisation. In both cases the minimised starting geometry was used to generate the new conformation rather that the last previous conformation in the search. This avoids the production of many chair states which are often induced by side chain interactions in highly strained conformations; once one is produced this can be carried unrealistically into the remainder of the conformational search.

Verlet[16] dynamics simulations were executed on the lowest energy conformations of the protonated forms of the structures from the torsionally constrained grid searches. A solvent shell of TIP3P[17] water molecules and 15 Å diameter was built around each molecule. A positional harmonic constraint of 0.1 kcal/mol was applied to the water molecules to prevent boil off of the water in the dynamics runs, particularly during the heating stage. This was set at 8.0 Å for ΔHexA-GalNAc4S and ΔHexA-GalNAc. For ΔHexA-GalNAc6S which has a more extended conformation this value was set to 7.2 Å. The SHAKE[18] algorithm was used throughout all runs. The dynamics simulations were composed of 10 ps of heating from 0° to 300°C followed by equilibration for a further 10 ps and simulation for 20 ps. In the case of the unsulphated analogue, ΔHexA-GalNAc, the simulation phase was run for 30 ps. The coordinates of all atoms were saved every 0.1 ps for analysis.

RESULTS

Preparative gel electrophoresis.—Limited depolymerisation of chondroitin sulphate produces a number of oligosaccharides each differing by a single disaccharide from one another. Fig. 2 shows a typical gel exclusion chromatogram on Sephadex G-50 of such a digest. As can be clearly seen there are a number of peaks that are partially resolved by this technique, but preparation of single

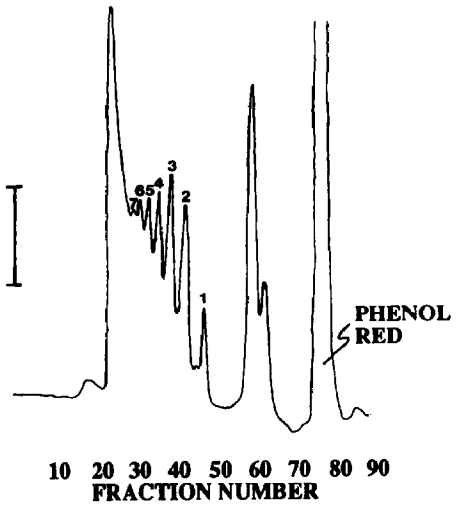

Fig. 2. Sephadex G-50 gel exclusion chromatography of a partially depolymerised chondroitin sulphate preparation. Details as outlined in text (column dimensions, 84.5 cm×4.5 cm^2; 5-mL fractions at a flow rate of 37.5 mL/h; bar represents 0.01 AU at 206 nm). Oligosaccharides are numbered in terms of increasing molecular mass each differing by a single disaccharide unit. The smallest oligosaccharide (numbered 1) represents a hexasaccharide. The numbering system used in Figs. 2, 4, and 5 is consistent for all oligosaccharides.

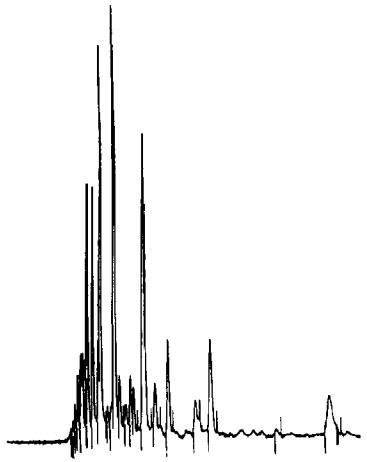

Fig. 3. Capillary electrophoretogram of the chondroitin sulphate preparation chromatographed in Fig. 2. Details of electrophoresis are as outlined in text.

oligosaccharides by this technique requires multiple chromatographic steps and it is difficult to obtain most of the larger oligosaccharides in a pure form. Capillary electrophoresis (Fig. 3) has the resolving potential to separate various oligosaccharides. For the requirements of a project such as this one, however, it seems impossible that one could ever obtain sufficient material (1–2 mg) to perform NMR analysis even using parallel capillary bundles, a technique that is receiving considerable attention at present. CE represents a rapid, convenient and highly sensitive technique for monitoring digestion and purity of prepared oligosaccharides. A new approach to preparation of oligosaccharides is to adopt polyacrylamide gel electrophoresis. We have adapted a method of analytical electrophoresis by Cowman et al.[11] for the preparation of oligosaccharides depolymerised by the action of hydrolases and eliminases. Fig. 4 shows the analytical separation of chondroitin sulphate oligosaccharides obtained from preparative electrophoresis. It is possible to see the degree of purity of oligosaccharides obtained from a single preparative electrophoresis run. Assessed densitometrically, the purity (with respect to size) appears to be better than 90% (results not shown). This is somewhat surprising in view of the degree of overlap in a typical preparative run as shown in Fig. 5, but may be principally due to cumulative overlap of consecutive fractions, half of which would not be represented in the purified sample. The concentration of oligosaccharide in each fraction was determined by the DMB dye binding assay. Fractions could be collected and reanalysed by capillary electrophoresis or by analytical gel electrophoresis as shown in Fig. 4.

NMR spectroscopy.—^{1}H and ^{13}C NMR data for the unsulphated reference compound, *N*-acetylchondrosine, are compiled in Table I. ^{1}H and ^{13}C NMR data for the unsaturated chondroitin sulphate disaccharides are compiled in Tables II

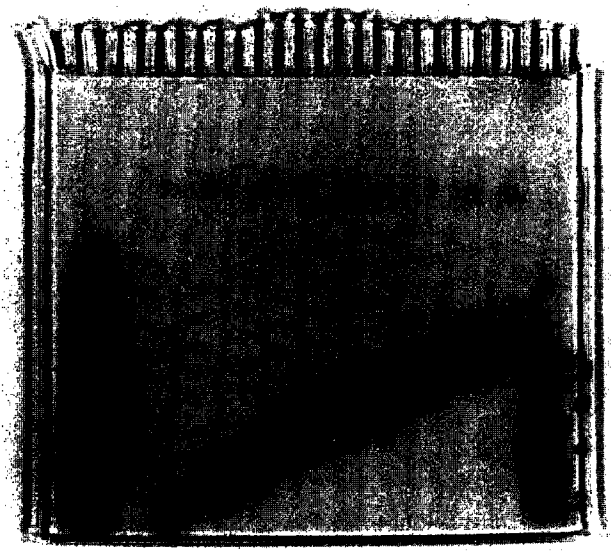

Fig. 4. Analytical gel electrophoretogram of the chondroitin sulphate preparation chromatographed in Fig. 2 (Track D). Tracks designated 1–9 represent the fractionated oligosaccharides obtained from preparative gel electrophoresis (see Fig. 5). The positions of the individual oligosaccharides in the mixture are numbered 1–9 in the right hand column.

and III. The bulk of the proton resonances in the spectrum of ΔHexA-GalNAc6S occur in the region 3.8–4.4 ppm where there are sixteen overlapping proton multiplets. The only proton signals occurring outside this region are the olefinic (H-4) proton of the HexA moiety (at 5.88 ppm) and the anomeric protons of each

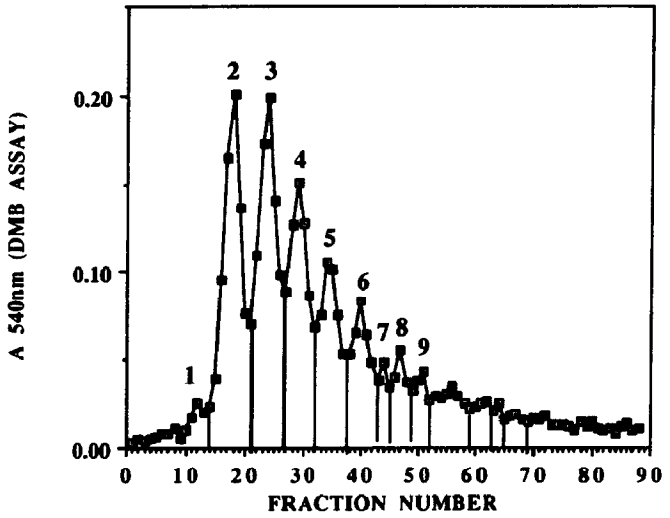

Fig. 5. A preparative electrophoretogram of the digests used in Figs. 2–4. Detection of oligosaccharides was by the method of Farndale[12].

TABLE I

^1H and ^{13}C NMR chemical shifts of *N*-acetylchondrosine

Residue	Proton	^1H Chemical shift		Carbon	^{13}C Chemical shift	
		α	β		α	β
GlcA	H-1′	4.56	4.50	C-1′	106.9	106.9
	H-2′	3.33	3.33	C-2′	75.6	75.6
	H-3′	3.45–3.55	3.45–3.55	C-3′	78.2	77.8
	H-4′	3.45–3.55	3.45–3.55	C-4′	74.6	74.6
	H-5′	3.6–3.8	3.6–3.8	C-5′	79.0	79.0
				C-6′	177 –178	177 –178
GalNAc	H-1	5.21	4.68	C-1	94.0	98.0
	H-2	4.29	3.98	C-2	51.8	51.8
	H-3	4.00	3.82	C-3	80.2	83.2
	H-4	4.24	4.18	C-4	71.3	70.6
	H-5	4.12	4.12	C-5	73.2	73.2
	H-6a	3.6–3.8	3.6–3.8	C-6	64.1	63.9
	H-6b	3.6–3.8	3.6–3.8	NAc(CH_3)	25.1 [a]	24.9 [a]
	NAc(CH_3)	2.025	2.025	NAc(CO)	n.d. [b]	n.d.

[a] α/β Assignments are interchangeable. [b] n.d., Not determined.

residue. Spectral assignment was achieved with the aid of a COSY experiment and 1- and 2-step relayed COSY experiments. Once the anomeric proton (H-1) of a residue has been assigned, H-2 can be readily assigned from the COSY spectrum, H-3 from the 1-step relayed COSY spectrum, and H-4 from the 2-step relayed COSY spectrum. This is particularly well demonstrated for the four spin system of

TABLE II

^1H NMR chemical shifts of ΔHexA-GalNAc6/4S

Residue	Proton	^1H Chemical shift (ppm)			
		ΔHexA-GalNAc6S		ΔHexA-GalNAc4S	
		α	β	α	β
ΔHexA	H-1′	5.23	5.19	5.29	5.25
	H-2′	3.79	3.79	3.85	3.85
	H-3′	4.12	4.12	3.93	3.93
	H-4′	5.88	5.88	5.96	5.96
GalNAc	H-1	5.22	4.72	5.20	4.76
	H-2	4.29	3.98	4.30	4.14
	H-3	[a]	3.93	4.30	4.14
	H-4	[a]	[a]	4.67	4.60
	H-5	[a]	3.94.0	4.2–4.3	4.2–4.3
	H-6a	4.37	4.37	3.65–3.8	3.65–3.8
	H-6b	[a]	[a]	3.65–3.8	3.65–3.8
	NAc(CH_3)	2.046	2.046	2.072	2.072

[a] Denotes the region 4.10–4.25 ppm.

TABLE III

^{13}C NMR chemical shifts of ΔHexA-GalNAc6/4S

Residue	Carbon	^{13}C Chemical shift (ppm)			
		ΔHexA-GalNAc6S		ΔHexA-GalNAc4S	
		α	β	α	β
ΔHexA	C-1′	104.1	104.1	103.1	103.1
	C-2′	72.4	72.4	71.7	71.7
	C-3′	68.7	68.7	67.7	67.7
	C-4′	110.1	110.1	109.8	109.8
	C-5′	147.3	147.3	n.d. [a]	n.d.
	C-6′	177.4	177.4	n.d.	n.d.
GalNAc	C-1	94.1	97.9	94.4	98.0
	C-2	51.7	55.2	52.9	56.4
	C-3	79.7	82.5	76.1	78.7
	C-4	70.9 [b]	70.3 [c]	80.4	79.3
	C-5	71.3	75.6	73.6	77.7
	C-6	71.0 [b]	70.6 [c]	64.2 [d]	64.3 [d]
	NAc(CH_3)	25.0 [d]	24.8 [d]	25.2 [d]	25.4 [d]
	NAc(CO)	173.9 [d]	172.2 [d]	n.d.	n.d.

[a] n.d., Not determined. [b] Assignments are interchangeable. [c] Assignments are interchangeable. [d] α/β Assignments are interchangeable.

the HexA residue, in the 2-step relayed COSY of ΔHexA-GalNAc6S (Fig. 6). The chemical shifts of H-6a and H-6b of the GalNAc ring have been assigned at 4.37 ppm and in the region 4.10–4.25 ppm, respectively. These have shifted significantly downfield relative to those of H-6a and H-6b in *N*-acetylchondrosine (3.6–3.8 ppm). This is consistent with sulphation at C-6 as reported previously for chondroitin sulphate[19,20] and keratan sulphate[21] saccharides. Also the *N*-acetyl methyl signal has shifted downfield by +0.021 ppm.

The ^1H NMR spectrum of the ΔHexA-GalNAc4S disaccharide has been assigned in a similar manner. In this case, the most significant downfield shift (relative to *N*-acetylchondrosine) is observed for the chemical shift of H-4 (+0.4 ppm) of the GalNAc ring, indicating sulphation at C-4 of this sugar residue. For this sample, a downfield shift of +0.047 ppm is observed for the *N*-acetyl methyl signal.

Assignment of ^{13}C NMR data (and further assignment of ^1H NMR data) for both disaccharides was achieved with the aid of two-dimensional ^1H-detected (i.e., inverse) heteronuclear correlation experiments. In the case of the ΔHexA-GalNAc6S disaccharide, a conventional ^{13}C spectrum was acquired with a total of 98 304 scans which amounted to 36 h accumulation time. In contrast, the HMQC experiment (Fig. 7) containing the additional information, as described in the introduction, was acquired overnight. The sulphation position of the GalNAc ring is also evident from the ^{13}C data[22], a significant downfield shift being observed for the chemical shift of the sulphated carbon. In the case of the ΔHexA-GalNAc6S,

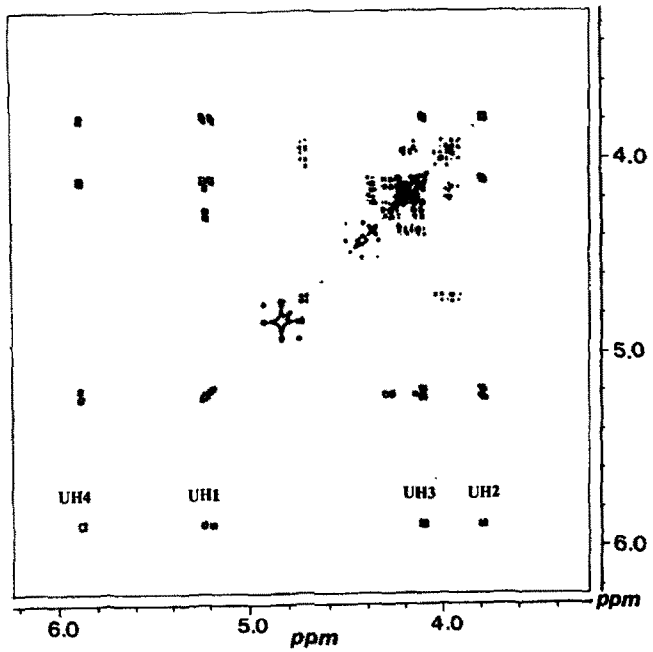

Fig. 6. 300-MHz 2-step relayed-COSY spectrum of ΔHexA-GalNAc6S. Signals labelled "U" refer to the HexA residue.

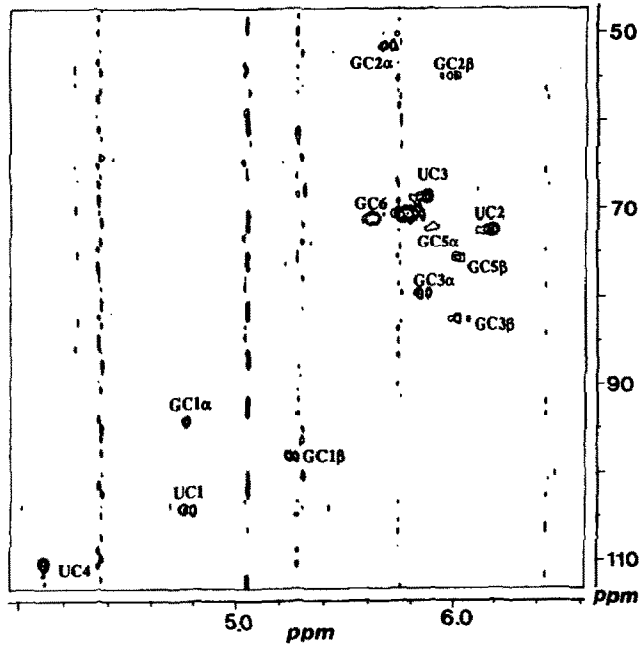

Fig. 7. ^{1}H-detected ^{13}C–^{1}H heteronuclear correlation (HMQC) spectrum of ΔHexA-GalNAc6S. Signals labelled "U" refer to the HexA residue, while those of the GalNAc residue are designated "G".

TABLE IV

ϕ, ψ Angles of the lowest energy conformations of ΔHexA-GalNAc4S, ΔHexA-GalNAc6S and ΔHexA-GalNAc using protonated acidic groups from type 1 (fixed torsions) and type 2 (relaxed torsions) conformational searches

Structure	ϕ Angles		ψ Angles		Energy (kcal mol^{-1})
	O-5′–C-1′ –O-1′–C-3	H-1′–C-1′ –O-1′–C-3	C-1′–O-1′ –C-3–C-2	C-1′–O-1′ –C-3–H-3	
ΔHexA-GalNAc6S type 1	−150	−33	180	−65	−260.5
ΔHexA-GalNAc6S type 2	−168	−53	178	−63	−261.3
ΔHexA−GalNAc4S type 1	−80	35	50	169	−185.1
ΔHexA-GalNAc4S type 2	−84.8	35.0	54.5	174	−189.5
ΔHexA-GalNAc type 1	30	150	−170	−52	−32.4
ΔHexA-GalNAc type 2	63	57	−155	−34	−48.8

downfield shifts of +6.9 and +6.7 ppm are observed for GalNAc C-6 in the α- and β-anomers, respectively. The chemical shift of GalNAc C-5 is also perturbed, an upfield shift of −1.9 ppm being observed for the α anomer and a downfield shift of +2.4 ppm for the β-anomer. On the other hand, the ΔHexA-GalNAc4S disaccharide exhibits downfield shifts of +9.1 and +8.7 ppm for GalNAc C-4 in α- and β-anomers, respectively. In this case, the chemical shift of GalNAc C-3 shows an upfield shift of −4.1 ppm for the α-anomer and −4.5 ppm for the β anomer. This observation is consistent with previous data obtained for native chondroitin sulphate[23]. Interestingly, the chemical shift of GalNAc C-5 of the β anomer shifts substantially downfield (+4.5 ppm), while that of the α anomer shifts only slightly (+0.4 ppm).

The HMBC experiment, which defines the long-range (i.e., within a two- or three-bond distance) ^{13}C–^{1}H correlations has allowed the specific assignment of the anomeric protons of the HexA residue to α and β anomers. The observed long-range correlations from the α and β anomeric (H-1) protons of the HexA residue to the α and β-aglyconic (C-3) carbons of the GalNAc residue (i.e., across the glycosidic linkage) have distinguished the chemical shifts of these protons.

Molecular modelling.—The minimum energy structures of both the type 1 and type 2 conformational searches were broadly similar as shown by Tables IV and V. Analysis of these data is focused on the protonated searches (Table IV) since as the monosaccharide units close upon themselves, the electrostatic interactions in the charged portions of the molecule can begin to dominate in the force field. In a highly charged unprotonated molecule this is an unrealistic simulation of the physical system in aqueous media.

TABLE V

ϕ, ψ Angles of the lowest energy conformations of ΔHexA-GalNAc4S, ΔHexA-GalNAc6S and ΔHexA-GalNAc using unprotonated acidic groups from type 1 (fixed torsions) and type 2 (relaxed torsions) conformational searches

Structure	ϕ Angles		ψ angles		Energy (kcal mol^{-1})
	O-5′–C-1′ –O-1′–C-3	H-1′–C-1′ –O-1′–C-3	C-1′–O-1′ –C-3–C-2	C-1′–O-1′ –C-3–H-3	
ΔHexA-GalNAc6S type 1	−40	82	−60	63	−18
ΔHexA-vGalNAc6S type 2	−34	87	−63	59	−16
ΔHexA-GalNAc4S type 1	−100	20	50	168	−57
ΔHexA-GalNAc4S type 2	−89	30	40	160	−59
ΔHexA-GalNAc type 1	−80	−42	180	−61	−24
ΔHexA-GalNAc type 2	45	164	−109	14	−28

Initial inspection of the ϕ, ψ plots of these conformational searches shows them to have several potential minima. A representative example from the type 1 search on ΔHexA-GalNAc6S is shown in Fig. 8. The range of energy shown here is up to 100 kcal from the lowest energy structure. If conformations within only 10 kcal of the lowest energy structure are considered then small discrete clusters of structures emerge. The low energy conformations derived from the ϕ, ψ plots clearly do not describe all conformational space, it is not within the scope of a preliminary study such as this to cover all possible side-chain geometries. In determining

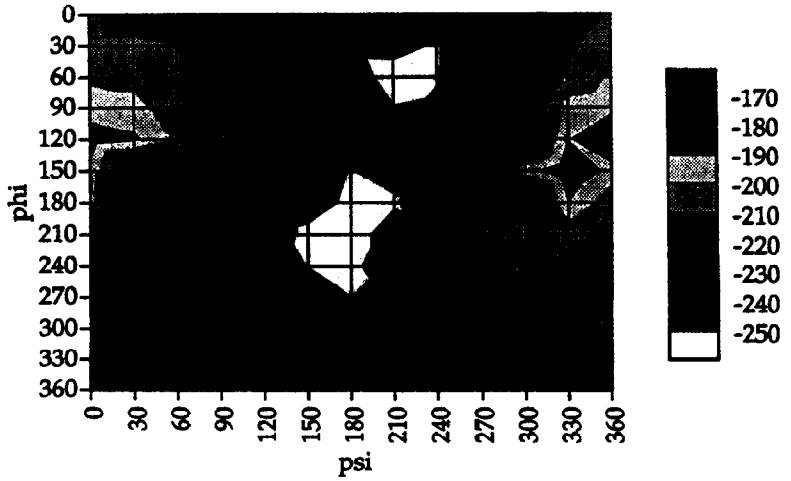

Fig. 8. 2D Plot of ϕ and ψ angles (0 to 360°) vs. energy (−250 to −170 kcal mol^{-1}) for ΔHexA-GalNAc6S.

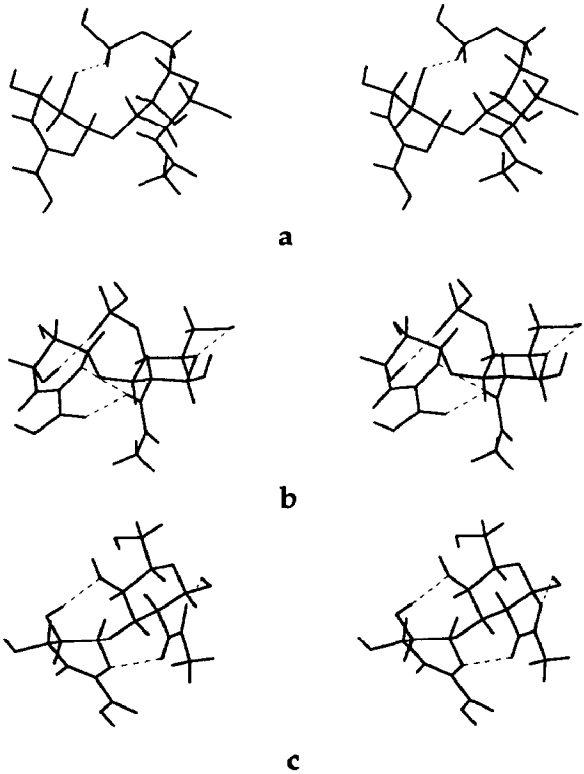

Fig. 9. Stereoviews of the minimum energy conformations of: (a) ΔHexA-GalNAc6S, (b) ΔHexA-GalNAc4S, (c) ΔHexA-GalNAc. Dashed lines indicate hydrogen bonds.

oligosaccharide conformation it is first necessary to determine likely ϕ, ψ angles before embarking on a thorough investigation of side-chain geometry. However, it is recognised that ϕ, ψ angles will be influenced by pendant groups so energy minimisations were carried out at each incremental change in the torsional angle to resolve side-chain interactions during the conformational searches.

ΔHexA-GalNAc6S. When a 5 kcal cut-off is applied to the type 1 search the ϕ, ψ plot shows only a small area of available conformational space centred around ϕ −159.7°, ψ 170.6°. The key electrostatic interaction stabilising these structures appears to be between the S=O oxygen and the 2′ hydroxyl as shown in Fig. 9a. At its shortest, the distance between these two groups is ~1.7 Å. This interaction is also observed in the dynamics experiment, the 2′ hydroxyl and S=O oxygen maintaining an intermolecular hydrogen bond throughout the 30 ps simulation phase. When a 10 kcal cut-off is applied a further minima centred around a second structure (ϕ 60.0°, ψ −140.0°) is seen. The structures in this potential well occur as the *N*-acetylgalactosamine ring wraps further under the glucuronic acid ring and a stabilising electrostatic interaction can take place between the S=O oxygen and the 3′ hydroxyl proton. This is a slightly longer interaction than that seen with the

2′ hydroxyl, being ~2.0 Å at its shortest. The type 2 search shows similar results to the type 1 search with the lowest energy minima centred at ϕ −154.9°, ψ −174.6°. At the 10 kcal cut-off level a further area of low energy structures is observed centred at ϕ 57.0 ± 5.0° and ψ 138.0 ± 2.0°. There is an additional hydrogen bond interaction not observed in the conformational searches which is seen in the dynamics simulations. The C-3–C-2–N–H torsion rotates to allow the proton to form a stable 5-membered hydrogen bonded ring with the oxygen of the glycosidic linkage. The total energy trace is linear throughout the simulation phase of the dynamics run.

ΔHexA-GalNAc4S.—Both the type 1 and type 2 searches show the 4-sulphated analogue to be much more conformationally restricted with only one real area of low energy structures being generated. The type 1 search at the 5 and 10 kcal cut-off levels shows this to be centred in the region ϕ −90.0 ± 18°, ψ 40.0 ± 17°. The 10 kcal cut-off shows a further very restricted potential well of ϕ 170.0°, ψ 140.0 ± 10.0°.

These minima have reasonably close contacts (see Fig. 9b) between the *N*-acetylgalactosamine amide proton and the acid carbonyl on the glucuronic acid moiety so providing a stabilising hydrogen bond. The amide proton is also only 2.4 Å from the 5′ oxygen and so, with very little movement in the ϕ, ψ angles, a further hydrogen bond type stabilisation may occur. The lowest energy structures in the 5 kcal search allow an electrostatic interaction between the S=O oxygen and the 3′ hydroxyl group. This is also observed in the dynamics simulations through a water molecule hydrogen bonding to the oxygen of the glycosidic linkage to complete a twist-boat type 6-membered ring. The amide proton–acid carbonyl interaction seen in the conformational searches is also seen in the dynamics simulations and, in addition, the amide proton is also seen to hydrogen bond with the 5′ oxygen. No radical changes in the ϕ, ψ angles are observed throughout the simulation and the total energy during the simulation phase is virtually linear.

ΔHexA-GalNAc.—This is easily the most flexible structure, the type 1 search indicating several large potential wells at both cut-off levels. Indeed, at the 10 kcal level so much conformational space is covered the only structures which appear to be precluded are those which encounter severe steric repulsion through the hexuronate and *N*-acetylgalactosamine rings resting on top of one another. It serves then, only to examine the minima at the 5 kcal type 1 search. The largest potential well is centred around ϕ −35.0 ± 25.0° with ψ at either −145.0 ± 35.0° or −65.0 ± 25.0°. These minima have a facile transition of ~2 kcal to a further area of low energy conformations centred at ϕ 25.0 ± 25.0°, ψ −110.0 ± 30.0°. Another area of low energy conformers ~2 kcal higher than the lowest energy structure is centred at ϕ −110.0 ± 30.0°, ψ 5.0 ± 36.0°. The barrier to transition between these two structures from the minimum energy conformation, as measured in the type 2 search is only ~5 kcal. Since this structure is unsulphated the only charged residues on this structure which may provide hydrogen bonded and/or electrostatic stabilisation are the amide and acid groups. The lowest

energy minimum is stabilised by a hydrogen bond between the amide hydrogen and the 5′ oxygen as shown in Fig. 9c.

DISCUSSION

Although the structure of chondroitin sulphate[24] has been known on the basis of its component disaccharides since 1955, it is clear from recent work that sequence information may be contained within these chains as a result of the arrangement of the various sulphated isomers. In related glycosaminoglycans such as heparin it has been demonstrated that a specific pentasaccharide is responsible for binding to antithrombin[3]. More recently using monoclonal antibodies, the presence of unusual yet presumably specific chondroitin sulphate sequences (which are synthesised in response to a variety of biological stimuli[2]) has been demonstrated. First, it was shown that there was a specific spatial and temporal arrangement of such structures in the developing chick bursa. More significantly perhaps, it was demonstrated that cartilage produced these unusual chondroitin sulphate structures in response to the development of experimental osteoarthritis[25]. Their function may be related to binding of biologically active molecules such as growth factors and cytokines. As a result, there has been considerable interest in determining the structure and shape of these epitopes in order to gain a greater understanding of the mechanisms which may underpin the development of osteoarthritis.

Although there has been a great deal of interest in the elucidation of chondroitin sulphate sequence structure, the field has progressed relatively slowly because of two major factors. First, there is a lack of useful cleavage and derivatisation protocols for such linear heteropolysaccharides. Furthermore, purification of defined oligosaccharide structures has not been possible. In this paper we have shown alternative and rapid, high capacity techniques for the purification of oligosaccharides of defined length. This does not however solve the problem of microheterogeneity. Studies are in progress to address this problem, using ion exchange HPLC in conjunction with specific sulphatases and affinity chromatography steps. The success of these techniques is currently being assessed by CE. Such techniques have shown initial promise and we are at present preparing milligram amounts of oligosaccharides with defined sulphate sequence. We will soon be able to apply the experience obtained in the NMR analysis and molecular modelling of disaccharides to the definitive structural assignment and three-dimensional shape of these oligosaccharides. This will be of great significance in the future when it may be necessary to design new antigens for the production of monoclonal antibodies which occur in pathological states. Previous NMR investigations of chondroitin sulphate have been limited to ^{1}H and ^{13}C NMR of native polysaccharide preparations[19,22,23] (which very often consist of mixtures of isomers and give broad signals in ^{1}H NMR spectra) and one study of ^{1}H NMR of disaccharides[20]. ^{13}C NMR spectroscopy is advantageous in terms of both the greater chemical shift

range and the singlet nature of the signals. Conventional two-dimensional heteronuclear correlation experiments have been employed previously[23] in assigning ^{13}C spectra, of the polymeric species, the results of which have necessitated some reassignments of signals.

In this study, we report both ^{1}H- and ^{13}C NMR data for purified chondroitin sulphate disaccharides. The spectra have been completely assigned by a combination of two-dimensional methods. Although sample quantities of 1–3 mg were available here, ^{13}C enrichment of oligosaccharides would further increase the potential of these experiments. Moreover, we expect to be able to examine chondroitin sulphate oligosaccharides of greater length. With increasing chain length however, there is an exponential increase in complexity. For example, a tetrasaccharide monosulphated on the galactosamine rings has four potential structures, a hexasaccharide, eight; an octasaccharide, sixteen; and so on. However, the glucuronate may also be sulphated on C-2 (some reports[26] have shown that on rare occasions the C-3 may also be sulphated) and the galactosamine could also be unsulphated or indeed oversulphated at C-4 and C-6, forming the disulphate. Obviously the degree of complexity conferred by these post-translational modifications is greatly increased, producing a myriad of potential structures for tetrasaccharides and larger oligosaccharides. As has been suggested in the past[27] the potential for structural (and functional) diversity in the carbohydrates (in this case the glycosaminoglycans) is far greater than that for proteins and may be of great significance in the differentiation, development and degeneration of tissues.

Characterisation of chondroitin sulphate epitopes (in terms of composition, sequence and sulphation pattern) is of key importance in determining their biological function. The data presented here for these model disaccharides will be of use in investigating these larger structures. Also, a strategy for full assignment of ^{1}H NMR data is necessary for conformational studies involving generation of distance constraints (from NOESY and ROESY experiments) for molecular modelling routines. Furthermore, assignment of both ^{1}H and ^{13}C NMR data is a prerequisite for measurement of long-range coupling constants across the glycosidic linkage. The use of NMR in determining distance/angle constraints is essential for preparing accurate molecular models since at present crystallographic data for these molecules is not available.

Carbohydrate conformational analysis is often the subject of debate since there is still little consensus as to the exact nature of the exoanomeric effect[28] in aqueous media[29,30]. Although parameters are available in force fields to account for the subtleties of the exoanomeric effect[31,32], the force field[15] used here, CHARMm 22, is in its native form. For ΔHexA-GalNAc4S and ΔHexA-GalNAc6S the exoanomeric effect predicts that the expected preference of the *N*-acetylgalactosamine ring is to be in near synclinal orientation to the plane of the glucuronic acid ring oxygen and C-1′ carbon. In a force field unparameterised for the exoanomeric effect it cannot be expected that low energy conformers of this type will be preferred. The lowest energy conformers found in this study are stabilised

by hydrogen bonding and/or electrostatic interactions which tend to be at variance with a positive exoanomeric effect. The observed interaction between the acetamido group and the glucuronic acid moiety in ΔHexA-GalNAc4S provide a particular rigidity to this disaccharide. Although this type of interaction is possible in the ΔHexA-GalNAc6S analogue the major electrostatic stabilising interactions are seen through the S=O oxygens–glucuronic acid hydroxyl group, thus inducing a different conformational preference in ϕ, ψ space (Tables IV and V). The differences observed in these disaccharide conformations will clearly have profound effects on oligosaccharide conformation.

Under the regime of these conformational searches the optimised orientations of the charged groups which lead to the stabilisation of minimum energy conformers is by no means complete. For example, it is equally possible in the case of ΔHexA-GalNAc4S that the opposite hydrogen bond donor/acceptor orientation of amide and acid moieties from that currently described may occur. Until the exoanomeric effect is more clearly understood and force fields adjusted to mimic it, then the key interactions which contribute to oligosaccharide conformation cannot be fully delineated. It will be interesting to compare these models with structures derived from purely empirical techniques. A Karplus-type curve, relating the long range $^{1}H–^{13}C$ coupling constants ($^{3}J_{CH}$) across the glycosidic linkage to the torsional angles, ϕ and ψ (see Fig. 1), has previously been characterised[33,34]. These $^{3}J_{CH}$ values may be determined using another ^{1}H-detected heteronuclear experiment, multiple bond heteronuclear single quantum coherence[35] (MBHSQC). Further studies employing this particular experiment are being carried out and the predicted torsional angles related to the results obtained from the molecular modelling studies reported here. These experimentally determined values will then be incorporated into dynamics procedures on larger chondroitin sulphate oligosaccharides.

ACKNOWLEDGMENTS

We would like to thank David Dobson for the preparation of *N*-acetylchondrosine, and David Osborne and Sandra Woodhouse for preparation and purification of the chondroitin sulphate disaccharides.

REFERENCES

1 S.L. Carney and H. Muir, *Phys. Rev. A*, 68 (1988) 858–910.
2 B. Caterson, F. Mahmoodian, J.M. Sorrell, T.E. Hardingham, M.T. Bayliss, S.L. Carney, A. Ratcliffe, and H. Muir, *J. Cell Sci.*, 97 (1990) 411–417.
3 U. Lindahl, D.S. Feingold, and L. Roden, *Trends Biochem. Sci.*, 11 (1986) 221–225.
4 K. Sugahara, Y. Okumara, and I. Yamashina, *Biochem. Biophys. Res. Commun.*, 162 (1989) 189–197.
5 S.L. Carney and D.J. Osborne, *Anal. Biochem.*, 195 (1991) 132–140.
6 H. Kessler, M. Gehrke, and C. Griesinger, *Angew. Chem. Int. Ed. Engl.*, 27 (1988) 490–536.
7 M. Ikura and Hikichi, *Carbohydr. Res.*, 163 (1987) 1–8.

8 M.A. Skelton, R. Cherniak, L. Poppe, and H. Van Halbeek, *Magn. Reson. Chem.*, 29 (1991) 786–793.
9 G.E. Martin and A.S. Zektzer, *Two-dimensional NMR methods for establishing molecular connectivity*, Vol. 18, VCH, UK., 1988, pp. 213–221.
10 A. Bax and M.F. Summers, *J. Am. Chem. Soc.*, 108 (1986) 2093–2094.
11 M.K. Cowman, M.F. Slahetka, D.M. Hittner, J. Kim, M. Forino, and G. Gadelrab, *Biochem. J.*, 221 (1984) 707–716.
12 R.W. Farndale, C.A. Sayers and, A.J. Barrett, *Connect. Tissue Res.*, 9 (1982) 247–248.
13 T. Bitter and H. Muir, *Anal. Biochem.*, 4 (1962) 330–334.
14 A. Bax and S. Subramanian, *J. Magn. Reson.*, 67 (1986) 565–569.
15 QUANTA3.3 and CHARMm 22, Molecular simulations Inc., Burlington, MA 01803-5279, USA.
16 L. Verlet, *Phys. Rev.*, 159 (1967) 98.
17 W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, *J. Chem. Phys.*, 79 (1983) 926.
18 J.P. Ryckaert, G. Ciccotto and H.J.C. Bedersen, *J. Comput. Phys.*, 23 (1977) 327.
19 D. Welti, D.A. Rees, and E.J. Welsh, *Eur. J. Biochem.*, 94 (1979) 505–514.
20 S. Yamada, K. Yoshida, M. Sugiura and K. Sugahara, *J. Biochem.*, 112 (1992) 440–447.
21 E.F. Hounsell, J. Feeney, P. Scudder, P.W. Tang, and T. Feizi, *Eur. J. Biochem.*, 157 (1986) 375–384.
22 S.M. Bociek, A.H. Darke, D. Welti, and D.A. Rees, *Eur. J. Biochem.*, 109 (1980) 447–456.
23 K.R. Holme and A.S. Perlin, *Carbohydr. Res.*, 186 (1989) 301–312.
24 E.A. Davidson and K. Mayer, *J. Am. Chem. Soc.*, 77 (1955) 4796–4798.
25 S.L. Carney, M.E.J. Billingham, B. Caterson, A. Ratcliffe, M.T. Bayliss, T.E. Hardingham, and H. Muir, *Matrix*, 12 (1992) 137–147.
26 R.P. Vieira, B. Mulloy, and P.A.S. Mourão, *J. Biol. Chem.*, 266 (1991) 13530–13536.
27 D.A. Rees, *Polysaccharide Shapes*, Chapman and Hall, London, 1977.
28 W.A. Szarek and D. Horton (Eds.), *Anomeric Effect, Origin and Consequences*, Vol. 87, ACS Symp. Ser., Washington, 1979.
29 R. Montaganani and J. Tomasi, *Int. J. Quantum Chem.*, 39 (1991) 851–870.
30 O. Kysel and P. Mach, *J. Mol. Struct.:* THEOCHEM, 227 (1991) 285–293.
31 J.W. Brady, M. Field, A. Giammona, and S.N. Ha, *Carbohydr. Res.*, 180 (1988) 207–221.
32 S.W. Homans, *Biochemistry*, 29 (1990) 9110–9118.
33 B. Mulloy, T.A. Frenkiel, and D.B. Davies, *Carbohydr. Res.*, 184 (1988) 39–46.
34 1. Tvaroska, M. Hricovini, and E. Petrakova, *Carbohydr. Res.*, 189 (1989) 359–362.
35 T.J. Norwood, J. Boyd, J.E. Heritage, N. Soffe, and I.D. Campbell, *J. Magn. Reson.*, 87 (1990) 488–501.