Stereoregular Acyclic Polyalcohols and Polyacetates from Cellulose and Amylose

B. Casu,*[1a] A. Naggi,[1a] G. Torri,[1a] G. Allegra,[1b] S. V. Meille,[1b] A. Cosani,[1c] and M. Terbojevich[1c]

Istituto di Chimica e Biochimica G. Ronzoni, Milano, Italy, Politecnico di Milano, Milano, Italy, and Centro Studi Biopolimeri, CNR, Padova, Italy. *Received February 13, 1985*

ABSTRACT: Oxidative scission with periodate of all the C(2)–C(3) bonds of the D-glucopyranose rings of cellulose and amylose followed by reduction with borohydride of the resulting polydialdehydes afforded the acyclic polyalcohols RDC and RDA [RDC = poly[(2*r*,4*S*,5*R*)-2,4,5-tris(hydroxymethyl)-1,3-dioxopentamethylene]; RDA = poly[(2*s*,4*S*,5*R*)-2,4,5-tris(hydroxymethyl)-1,3-dioxopentamethylene]]. The corresponding triacetates (RDC-Ac and RDA-Ac) were obtained by conventional acetylation of the polyalcohols. The ^{13}C and ^{1}H NMR spectra of the products indicated a high degree of structural homogeneity, with retention of the different configuration at C(1) of the original polysaccharides. The concept of stereoregularity, and stereoisomerism of polymers obtained from cellulose and amylose, is reflected also by the high crystallinity easily achieved in solid samples, with X-ray powder patterns unambiguously different for pairs of cellulose and amylose derivatives. Light scattering measurements indicate a strong tendency to association for RDC in water, as well as for the corresponding peracetate RDC-Ac in acetone, acetonitrile, and trifluoroethanol. RDA forms microgels in water, while RDA-Ac associates on increasing concentration and/or time in acetone and acetonitrile.

Bonds between carbons bearing vicinal hydroxyl groups are easily split by periodate and tetraacetate ions, with conversion of the hydroxyl groups to aldehydes.[2] Splitting reactions are widely used in structural analysis of carbohydrates and their polymers. Under controlled conditions, splitting of unsubstituted diols in pyranose residues of polysaccharides occurs without cleavage of the glycosidic bonds.[3] However, glycol-split polysaccharides are usually degraded with acid, without isolation, for characterizing the "substitution pattern" of the original polymer through analysis of fragments.[3]

Although advantage of splitting reactions (most commonly, only at the level of a few residues) is currently taken for functionalizing polysaccharides by reacting the aldehydo groups formed at the sites of splitting,[4] these reactions have been seldom used for deliberately modifying the macromolecular properties of polysaccharides. A major modification of these properties is expected in 1 → 4 linked polysaccharides as a result of an increased flexibility associated with extra degrees of freedom instigated on each split residue, each "opened" pyranose ring acting as a flexible joint in an otherwise relatively rigid polymer backbone.[5,6]

In the framework of a systematic study on the influence of splitting reactions on conformational and rheological properties of polysaccharides,[6] these reactions were used for converting cellulose (poly-β-D-glucose) and amylose (poly-α-D-glucose) into acyclic, stereoregular polydicarboxylates,[7] which are good sequestering agents for calcium and magnesium[8] and other divalent ions.[7]

In the present work, acyclic polyalcohols were prepared by splitting with periodate the C(2)–C(3) bonds of all the glucopyranose residues of cellulose and amylose and reducing with borohydride the resulting polydialdehydes. The corresponding polytriacetates were prepared by conventional peracetylation of the polyalcohols.

Experimental Section

The polyalcohols RDC and RDA were prepared from commercial samples of cellulose and amylose, respectively. More specifically, RDC-I was prepared from microcrystalline cellulose (Merck, Darmstadt), MW ~ 20000 (viscometric determination); RDC-II, from cotton cellulose, assumed MW ~ 300000; RDA was prepared from amylose from potato starch (Serva), MW ~ 150000 (data from manufacturer), essentially unbranched as determined by ^{13}C NMR analysis. The periodate oxidation and reduction reactions were performed as described for Smith degradation of polysaccharides,[9] with slight modification and omission of the final step of acid hydrolysis. The polysaccharide (2 g) was suspended in 100 mL of a 0.1 M solution of sodium metaperiodate brought to pH 3.6–4.0 with concentrated HCl. The suspension was stirred in the dark at 4 °C for 24 h and 30 days for amylose and cellulose, respectively. In order to minimize possible depolymerization due to formation of free radicals, the supernatant of the reaction for cellulose was removed after 15 days and substituted with a fresh solution of metaperiodate. At the end of the reactions, ethylene glycol (5 mL) was added to react with the excess periodate. The reaction mixture was passed through a sintered-glass filter and the solid thoroughly washed with cold water (~5 °C) to remove iodates. The polydialdehydes were then suspended in water (20 mL) and the suspension added to sodium borohydride (1.2 g). The reduction was allowed to proceed at room temperature for 10 h and the excess borohydride was destroyed. The solution was brought to pH 4 by dropwise addition of glacial acetic acid. The solution was then concentrated under vacuum in the presence of methanol, dialyzed against water through a 14000 dalton membrane, and evaporated to dryness under vacuum. Yields varied from 60% to 90%. RDA-2,2′,3,3′,6,6′-d_6 was prepared by treating RDA with D_2O (99.7%) for 12 h in the presence of Raney nickel (W. R. Grace, no. 28), as described for deuteration of mono- and oligosaccharides.[10] The peracetates RDC-Ac and RDA-Ac were prepared by treatment of the corresponding polyalcohols with acetic anhydride in pyridine. One gram of RDC (or RDA) was dissolved in 40 mL of a 10:9 (v/v) pyridine/acetic anhydride mixture and kept 6 h at 70 °C, under stirring. After standing 72 h at room temperature, the mixture was added to H_2O to precipitate the product. The precipitate was redissolved in $CHCl_3$, and the solution was extracted with HCl (1%), Na_2CO_3 solution (1% w/v), and H_2O and finally evaporated to dryness under vacuum (yield = 0.80–0.85 g).

Crystallization of the polyalcohols and the polyacetates was performed by addition of methanol to solutions in water and in chloroform, respectively. Melting points of the products were measured by differential scanning calorimetry on a Perkin-Elmer DCS-2B instrument. Appropriate annealing a few degrees below the melting points increased the crystallinity of all four polymers.

The NMR spectra were measured on a Varian CFT-20 spectrometer (20 MHz for ^{13}C) and a Bruker CXP-300 spectrometer (300 MHz for ^{1}H and 75 MHz for ^{13}C) from solution in D_2O (99.9%, following several exchanges with D_2O 99.7%) for the polyalcohols and in $CDCl_3$ for the acetates. Coupling constants for the anomeric centers were measured directly from splittings of the H-1 and C-1 signals (these latter in proton-coupled ^{13}C spectra).

Optical rotation measurements were made with a Perkin-Elmer Model 241 polarimeter at 20 °C. Viscosities were determined by

0024-9297/85/2218-2762$01.50/0

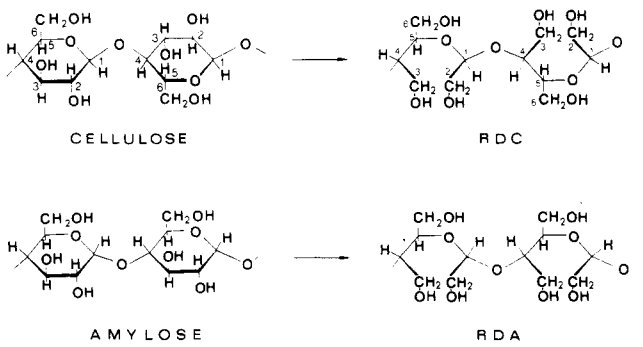


Figure 1. Conversion of cellulose and amylose to the corresponding "reduced dialdehyde" (RD) derivatives by periodate oxidation of the C(2)–C(3) bonds and borohydride reduction of the resulting dialdehyde groups.

using suspended level Ubbelohde viscometers at 25 ± 0.1 °C. The light scattering measurements were performed at 25 °C by a Sofica Model 42000 photometer with cylindrical cells immersed in toluene, using nonpolarized light at 436 and 546 nm (data obtained at the two wavelengths were substantially identical). Rayleigh ratios $(R_{90})_{436nm} = 48.5 \times 10^{-6}$ cm^{-1} and $(R_{90})_{546nm} = 16.3 \times 10^{-6}$ cm^{-1} were used for calibration of the instrument with benzene. Solvents (commercial, used without purification) and solutions were freed of dust by 2–3 h centrifugation at 25000 g. Concentrations were determined by weight. The samples of RDA and RDC were previously dialyzed against 0.5 M NaCl through a 10000 dalton membrane. Minor amounts of RDC-Ac insoluble in acetone were removed by centrifugation and filtration. Solubilization of RDC-Ac obtained from high molecular weight cellulose occasionally required heating at 60–80 °C. Cooling the solutions to 20 °C did not cause precipitation or turbidity.

Results

Synthesis and NMR Spectra. Cellulose and amylose were oxidized with sodium periodate at pH 3.5–4.0, and the resulting polydialdehydes were reduced with sodium borohydride at pH 8 to afford acyclic polyalcohols, as illustrated in Figure 1. Products are referred to as "reduced dialdehydes" (RDC and RDA). For simplicity, numbering of carbon atoms is maintained as in the original polysaccharides. More appropriate nomenclature and numbering are proposed further on.

Treatment of the polyalcohols with acetic anhydride in pyridine afforded the polyacetates RDC-Ac and RDA-Ac. $[\alpha]^{20}_D$ values for all the polyalcohols and polyacetates are ≥−1° for RDC and ≥+1° for RDA derivatives, i.e., smaller than 1 in absolute value.

The ^{13}C NMR spectra of typical preparations of RDC and RDA (Figure 2) consist of the expected four major signals (two of which account for two equivalent carbons each), indicating a high degree of overall conversion from the original polysaccharides. Minor signals (arising from residual unsplit glucopyranose rings and/or incompletely reduced aldehydes) account for less than 10% of the products.

Signal assignments were made according to established chemical shift correlation criteria for carbohydrate polymers.[11] Assignments for primary carbons were confirmed by the disappearance of the corresponding signals from the spectrum of RDA-2,2′,3,3′,6,6′-d_6 (stick diagram in Figure 2). Though quite similar to each other, the spectra of RDC and RDA display small but reproducible chemical shift differences. These differences were confirmed in the spectrum of a 1:1 mixture of the products where signals having different chemical shift in the spectrum of individual components were displayed as well-resolved doublets.

The ^{13}C NMR spectra of the acetates RDC-Ac and RDA-Ac are similar to those of the corresponding polyalcohols, except for substituent-induced shifts and the additional signals due to the acetyl substituents. The one-bond C(1)–H(1) coupling of these products was measured from the proton-coupled ^{13}C spectra.

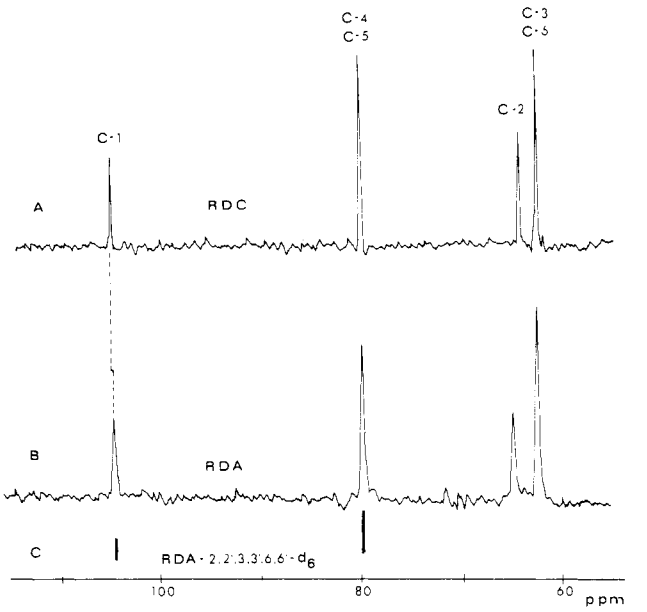


Figure 2. Proton-decoupled ^{13}C NMR spectra (20 MHz, D_2O) of RDC (A) and RDA (B), with stick diagram of RDA-2,2′,3,3′,6,6′-d_6 (C).

Table I
^{13}C and ^{1}H Chemical Shifts (δ)[a,b]

	RDC	RDA	RDC-Ac	RDA-Ac
C-1	104.6	104.1	102.7	100.4
($J_{C(1)H(1)}$)	(163.0)	(160.9)	(160.8)	(164.8)
C-2	63.6	64.0	64.7	63.4
C-3	61.8	61.8	63.6	62.4
C-4	79.3	79.3	74.7	75.0
C-5	79.3	79.3	74.7	75.0
C-6	61.8	61.8	63.6	62.8
CH_3			21.3	19.4
C=O			171.0	169.1
H-1	4.94	5.08	4.83	5.12
($J_{H(1)H(2)}$)	(4.5)	(4.6)	(4.75)	(5.1)
H-2,2′	3.67	3.67	4.04	4.04
H-3,6	3.76	3.81	4.39	4.39
H-4	3.97	4.09	3.92	4.13
H-5	3.97	4.09	3.92	4.13
H-3′,6′	3.76	3.81	4.17	4.13
CH_3-2			2.05	2.05
CH_3-3,6			2.10	2.08

[a] In D_2O for RDC and RDA, in $CDCl_3$ for RDC-Ac and RDA-Ac. Coupling constants (Hz) are given only for the anomeric centers. ^{13}C chemical shifts in D_2O are referred to internal methanol = 50.04 ppm from external Me_4Si, and, for solutions in $CDCl_3$, to internal $CHCl_3$ = 79.6 ppm from internal Me_4Si. ^{1}H chemical shifts are referred to internal $Me_3Si(CM_2)_3SO_2OH$ and Me_3Si for solutions in D_2O and $CDCl_3$, respectively. [b] Carbon (and hydrogen) numbering as in Figure 1.

The ^{1}H NMR spectra of the pair of polyalcohols and polyacetates displayed the expected major signals, assigned by their multiplicity and by spin decoupling. The ^{1}H spectra of typical preparations of RDC-Ac and RDA-Ac are shown in Figure 3. Table I reports ^{13}C and ^{1}H chemical shifts for RDC, RDA, and the corresponding acetates, together with $^1J_{C,H}$ and $^3J_{H,H}$ values for the anomeric centers.

Crystallization and X-ray Diffraction Data. After removal (by dialysis) of minor low molecular weight fractions, RDC and RDA were precipitated with methanol from aqueous solutions, affording highly crystalline pow-

Table II
Melting Points and Debye Spectra[a]

RDC,[b,c] mp 235 °C	RDA,[b,c] mp 120 °C	RDC-Ac,[c] mp 125 °C	RDC-Ac,[b] mp 118 °C	RDA-Ac,[b] mp 108 °C
7.05 vw	7.81 w		11.87 vs	12.67 s
6.35 vs	6.15 s	8.71 s	8.53 s	10.49 s
6.13 s	4.82 vs	8.24 vs	7.00 m	7.48 m
5.64 s	4.71 vw	6.37 ms	5.94 m	6.04 m
4.26 vs	4.21 vw	5.42 w	5.40 ms	5.61 w
3.85 vw	4.12 s	5.05 m	4.98 mw	5.29 w
3.74 m	3.77 s	4.93 m	4.54 ms	4.72 w
3.55 m	3.40 w	4.62 w	4.22 w	4.40 w
			4.02 vs	
			3.72 vw	
			3.51 vw	
3.39 m	3.08 w	4.35 s	3.35 ms	4.26 m
3.23 m	2.94 m	4.11 s		3.76 vs
3.03 w	2.60 m	3.87 s	3.11 w	3.52 m
2.59 m	2.23 m	3.69 vw	2.78 w	3.32 w
2.38 w	2.01 w	3.41 w	2.44 w	
2.27 w		3.31 w		
2.07		3.13 w		

[a] *d* in angstroms. [b] Solution-crystallized samples. [c] Melt-crystallized samples.

ders. Also, the corresponding peracetylated derivatives RDC-Ac and RDA-Ac were precipitated (with methanol from chloroform solutions) in crystalline form. Appropriate annealing of the four polymers a few degrees below their melting points for 2 h significantly improved the crystallinity and the quality of the X-ray diffraction spectra of the samples.

The Debye X-ray diffraction data of the annealed polymers (*d*-spacings and visual estimation of relative intensities) are given in Table II. Oriented crystalline samples were obtained for the cellulose derivatives RDC and RDC-Ac; their fiber diffraction patterns will be discussed elsewhere.

The existence of more than one crystalline form in the case of RDC-Ac is indicated by the X-ray diffraction spectra of solution-crystallized ("native") samples and of melt-crystallized samples. (See Table II.) For the other three polymers (RDC, RDA, RDA-Ac) there are no indications of the occurrence of different polymorphs upon crystallization from the melt.

Striking differences in crystallization behavior were observed between the two polyalcohols and the corresponding peracetylated derivatives. RDC and RDA crystallize very rapidly from the melt, and are significantly crystalline even when quenched from the melt to room temperature. By contrast, in spite of their lower melting points, the peracetylated derivatives do not crystallize at room temperature (20 °C), and their crystallization is slow even at higher temperatures.

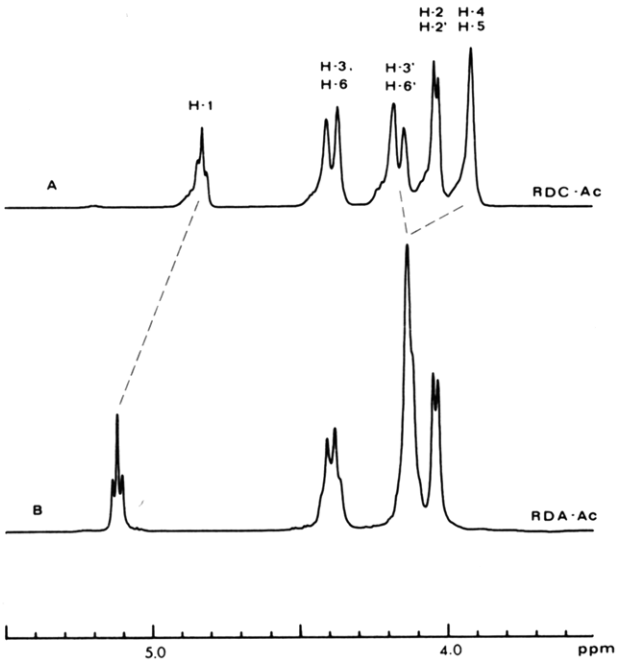


Figure 3. Partial ^{1}H-NMR spectra (300 MHz) of RDC-Ac (A) and RDA-Ac (B) in $CDCl_3$.

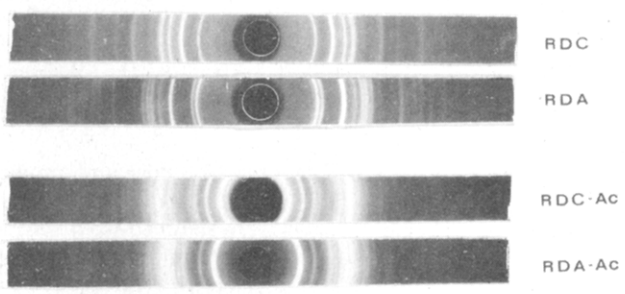


Figure 4. X-ray diffraction powder spectra of polyalcohols RDC and RDA and the corresponding acetates (annealed samples).

Solubility, Viscosity, and Light Scattering. Polyalcohols obtained from low molecular weight (microcrystalline) cellulose and from amylose (RDC-I and RDA) are soluble in water at 20 °C. The RDC preparations obtained from cotton cellulose (RDC-II), which are sparingly soluble in water at room temperature, could be solubilized at 80 °C. Reduced viscosities of the solutions are reported in Table III.

From the Zimm plot for RDC-II in water (Figure 4a), a rather high DP_w value ($\sim 6 \times 10^3$) is evaluated, corresponding to M_w values consistently higher than the ones assumed for the starting material. Similar data were ob-

Table III
Viscosity and Light Scattering

	solvent	η_{sp}/c (c, %)	M_w	DP_w[a]	B, mol mL g^{-2}	R_g, Å
RDC-I[b]	H_2O	0.11 (0.6)	$0.6–1 \times 10^5$	400–600	5×10^{-4}	900 ± 300
	0.5 M NaCl		$5–8 \times 10^4$	300–500	5×10^{-4}	900 ± 300
RDC-II	H_2O	0.13 (0.75)	1×10^6	6200	3×10^{-4}	1600
	0.5 M NaCl	*c*	1×10^5	620	1.5×10^{-4}	500
RDA	H_2O	0.21 (0.7)	7×10^4	430	8×10^{-4}	*d*
RDC-Ac[b]	$(CH_3)_2CO$	0.24 (0.3)	5×10^5	1700	3.5×10^{-5}	1050
	CH_3CN	*c*	1.3×10^6	4500	5×10^{-5}	2500
	CF_3CH_2OH	*c*	4×10^6	14000	0	2600
RDA-Ac	$(CH_3)_2CO$	*c*	7×10^4	250	1.5×10^{-4}	*d*
	CH_3CN	*c*	7×10^4	250	4×10^{-4}	*d*
	$(CH_2)_2Cl_2$	0.15 (1.2)	*c*	*c*	*c*	*c*

[a] Monomeric unit weight taken equals 164 for the polyalcohols and 290 for the polyacetates. [b] From low molecular weight (microcrystalline) cellulose, dp 125. [c] Not determined. [d] Not measurable.

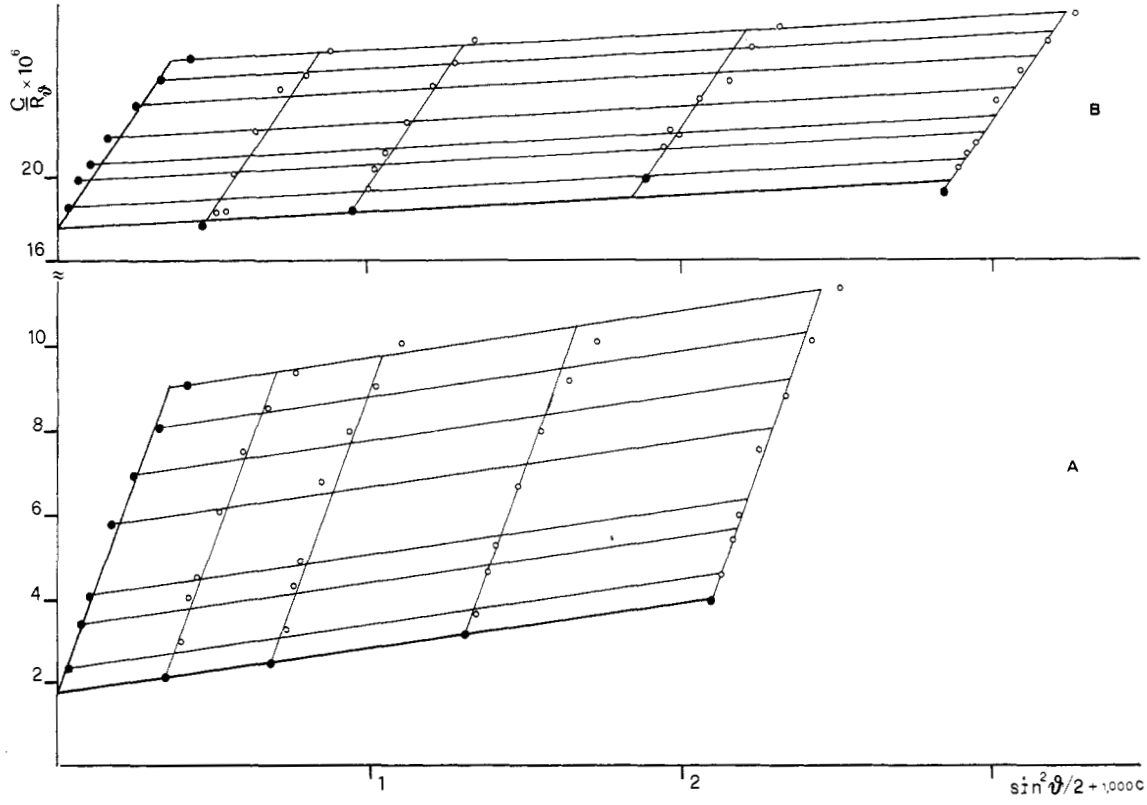


Figure 5. Zimm plots of RDC obtained from cotton cellulose (RDC-II) in H_2O (A) and 0.5 M NaCl (B).

tained after extensive heating (up to 8 h) and for measurements at 60–70 °C. In contrast, a different behavior at room temperature was observed when the polymer was dissolved in 0.5 M NaCl at 80 °C (Figure 4b). Addition of NaCl causes a dramatic drop in the estimated M_w (and R_G) values (see Table III). These values are lower than for the original polysaccharide. For the polyalcohols from amylose, measured M_w values were lower than for the starting material even in the absence of salt (measurements were not made in the presence of NaCl).

The angular dependence of scattered light of solutions in H_2O of polyalcohols from amylose was consistently more marked at low angles (plots not shown), a behavior substantially unmodified by several cycles of heating–cooling or by the treatment with $CHCl_3$/isoamyl alcohol[12] for removing dust. This effect, which steadily increased with time of standing of the solutions with no apparent level off even after 180 h, was substantially decreased by centrifugation at 35 000 g for 3–8 h of fresh solutions and completely removed after standing for 8 days if solutions were centrifuged in the same conditions. M_w values measured on the latter solutions correspond to the values obtained for samples centrifuged for short periods of time when only data at moderately high angles are extrapolated to $\theta = 0$.

The polyacetates RDC-Ac and RDA-Ac are soluble in a number of organic solvents. Solubilization of the cellulose derivative in acetone or acetonitrile required some heating. As for the polyalcohols, light scattering data for the cellulose derivatives indicated high angular dependence of scattering, especially in trifluoroethanol. As shown by the data in Table III, DP_w values were substantially higher when compared to the unsubstituted polyalcohol RDC, a trend opposite to that observed for the corresponding amylose derivative using low polymer concentrations. Moreover, the scattering envelope of RDA-Ac in acetone is substantially modified on increasing the concentration of the polymer (Figure 5a). In acetonitrile, and only at

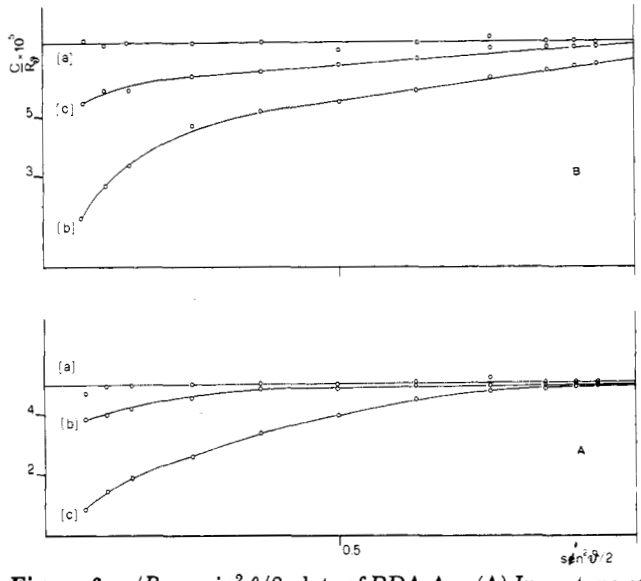


Figure 6. c/R_θ vs. $\sin^2 \theta/2$ plots of RDA-Ac: (A) In acetone at different concentrations (a, 5×10^{-3}; b, 8×10^{-3}; c, 14×10^{-3} g/L); (B) in acetonitrile at different times after dissolution (a, 3 h; b, 24 h; c, 24 h standing, followed by 15 min of heating at 45 °C).

relatively high polymer concentrations, the angular dependence of scattering was found to vary as a function of standing time (Figure 5b, curve b). Heating 5 min at 45 °C reversed the latter phenomenom to some extent (Figure 5b, curve c).

Discussion

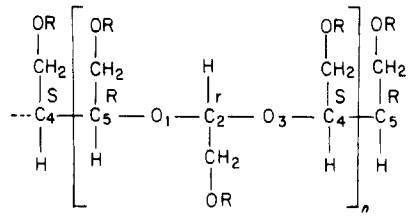
Opening the D-glucopyranose rings of cellulose and amylose by splitting all the C(2)–C(3) bonds with periodate and reduction (with borohydride) of the resulting polydialdehydes afforded the "reduced dialdehydes" RDC and RDA. In spite of being common intermediates in fragmentation analysis of polysaccharides,[3,9] these acyclic

polyalcohols were not previously isolated and characterized and can be considered as polymers of essentially new type.[13] Also, the corresponding peracetylated derivatives (RDC-Ac and RDA-Ac) are markedly different from the triacetates of cellulose and amylose.

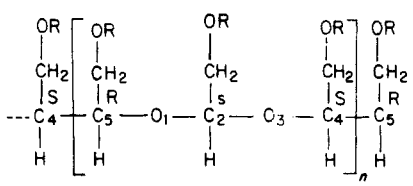
The NMR spectra of the polyalcohols and polyacetates (Figures 2 and 3) indicate a high degree of structural homogeneity, with minor signals accounting for less than 10% of the products. Also, significant chemical shift differences between the derivatives of cellulose and amylose (Table I) indicate that polymers obtained from the two polysaccharides are different from each other. The largest difference (confirmed in a 1:1 mixture of RDC and RDA) is for C-1 and H-1, suggesting that the "anomeric" carbon has retained the configuration of the original polysaccharide. This is supported also by differences in coupling constants associated with the "anomeric center" ($J_{C(1),H(1)}$ and $J_{H(1),H(2)}$, Table I). Spectral differences are especially striking for the acetates (Figure 3) where H-1 and H-4/H-5 of the RDA derivative are remarkably shifted downfield with respect to the RDC derivative. Since these shifts are associated also with significant differences in coupling constants $^3J_{H_{3,3'}}$ and $^3J_{H_{3,3'}}$ (data not reported), the different configuration at C(1) seems to involve also major differences in the conformation of the polymer chains.

The high crystallinity easily achieved for the polyalcohols and the corresponding acetates is likely to reflect the high degree of structural regularity of the products. A further proof of this stereoregularity was provided by the X-ray fiber pattern of good quality obtained from RDC-Ac.[14] Unambiguous differences between the X-ray powder patterns of cellulose and amylose derivatives support the conclusion that RDC and RDA are indeed stereoisomers.

In essence, glycol-split cellulose and amylose are polymers based on 1,3-dioxopentamethylene. This is better shown by formulas 1 and 2 (with stereochemical notations)



1, RDC/RDC-Ac



2, RDA/RDA-Ac
R = H or Ac

for RDC [poly(2*r*,4*S*,5*R*)-2,4,5-tris(hydroxymethyl)-1,3-dioxopentamethylene] and RDA [poly(2*s*,4*S*,5*R*)-2,4,5-tris(hydroxymethyl)-1,3-dioxopentamethylene] and the corresponding acetates. (Note that C-1 of the original polysaccharide (Figure 1) becomes C-2 in the dioxopentamethylene numbering, while O-5 and O-1 become O-1 and O-3, respectively.) Formulas 1 and 2 also make evident the internal compensation leading to very low optical rotation values for both polymers.

Light scattering measurements indicate a strong tendency of the present cellulose derivatives to associate in solution (H_2O for the alcohol, and acetone, acetonitrile, and trifluoroethanol for the acetate). For RDC, this association is disrupted by 0.5 M NaCl. For RDA in H_2O, tendency to aggregation is reflected by a strong light scattering at small angles, suggesting formation of microgels. These microgels, which account for only a few percent of the material, are stabilized on standing and can be eliminated by high-speed centrifugation.

The highest M_w values obtained from Zimm plots are 4×10^6 (RDC-Ac in trifluoroethanol), corresponding to association of about seven polymer chains if the assumption is made that no depolymerization has occurred during the splitting and derivatization reactions. (Some depolymerization is indicated at least in the case of RDA, whose M_w values in H_2O are somewhat lower (less than one-half) the accepted value for the original polysaccharide.)

The marked tendency to chain–chain association is most likely associated with the stereoregularity of the present acyclic polyalcohols and polyacetates. Because of differences in molecular weight, the observed differences in solubility and solution behavior between RDC and RDA (as well as those between the corresponding acetates) cannot be interpreted only in terms of their different stereochemistry. In view of the importance of chain–chain association in solution for promoting formation of ordered structures and for modulating the conformational and rheological properties of polysaccharides and related polymers,[15] further studies with products of similar molecular weight are in progress. The possibility that a few unmodified glucopyranose residues (especially if arranged in blocks) contribute to the aggregation properties of the polymers should also be considered. However, since these unmodified residues (if any) are present in amounts lower than detectable by NMR spectra and optical rotation, it seems more likely that aggregability is associated with the acyclic stereoregular chains.

Acknowledgment. We gratefully thank Prof. Terence Painter for stimulating discussions, Prof. Mario Farina for suggestions on polymer nomenclature, and Romano Poletti for technical assistance in obtaining the NMR spectra. This work was supported by the Italian Research Council (CNR), Target-Oriented Project on Fine Chemistry.

References and Notes

(1) (a) Istituto di Chimica e Biochimica G. Ronzoni, 20133 Milano, Italy. (b) Dipartimento di Chimica, Politecnico di Milano, 20133 Milano, Italy. (c) Centro Studi Biopolimeri, C.N.R., Istituto di Chimica Organica dell'Università, 35100 Padova, Italy.
(2) Dryhurst, G. "Periodate Oxidation of Diol and Other Functional Groups"; Pergamon: Oxford, 1970.
(3) Perlin, A. S. "Glycol-Cleavage Oxidation", in "The Carbohydrates"; Pigman, W.; Horton, D.; Wander, J. D., Eds; Academic Press: New York, 1980; p 1167.
(4) Kennedy, J. F. *Adv. Carbohydr. Chem. Biochem.* **1974**, *29*, 306.
(5) Smidsrød, O.; Painter, T. *Carbohydr. Res.* **1973**, *26*, 125.
(6) Casu, B.; Meille, S. V.; Naggi, A.; Su, P.; Torri, G.; Zoppetti, G.; Allegra, G. *Carbohydr. Polym.* **1982**, *2*, 283.
(7) Casu, B.; Gennaro, U.; Meille, S. V.; Morrone, M.; Naggi, A.; Occhipinti, M. S.; Torri, G. *Int. J. Biol. Macromol.* **1984**, *6*, 89.
(8) Crescenzi, V.; Dentini, M.; Meoli, C.; Casu, B.; Naggi, A.; Torri, G. *Int. J. Biol. Macromol.* **1984**, *6*, 142.
(9) Goldstein, I. J.; Hay, G. W.; Lewis, B. A.; Smith, F. In "Methods Carbohydr. Chem." Whistler, R. L.; BeMiller, J. N. Eds.; Academic Press: New York, 1965; Vol. 5, p 361.
(10) Koch, H. J.; Stuart, R. S. *Carbohydr. Res.* **1977**, *59*, C1. Balza, F.; Cyr, N.; Hamer, G. K.; Perlin, A. S.; Koch, H. J.; Stuart, R. S. *ibid.* C7.
(11) Perlin, A. S.; Casu, B. "Spectroscopic Methods" in "The Polysaccharides"; Aspinall, G. O., Ed.; Academic Press: New York, 1982; Vol. 1, p 133.
(12) Bernardi, G. *Makromol. Chem.* **1964**, *72*, 205.
(13) Reduced-dialdehyde derivatives of cellulose with different degrees of glycol splitting and conversion to polyalcohols (in-

cluding a RDC of >95% conversion, as in the present work) were independently prepared by Prof. T. Painter (private communication).
(14) Meille, S. V.; Allegra, G.; Naggi, A.; Torri, G.; Casu, B. presented at the XIIth International Carbohydrate Symposium, Utrecht, 1984; Abstract p 525.
(15) Rees, D. A.; Morris, E. R.; Thom, D.; Madden, J. K. "Shapes and Interactions of Carbohydrate Chains" in "The Polysaccharides" Aspinall, G. O., Ed.; Academic Press: 1982; Vol. 1, p 196.

Chain Closure with Bond Angle Variations

Robert E. Bruccoleri† and Martin Karplus*

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138.
Received February 13, 1985

ABSTRACT: The chain-closure algorithm of Gō and Scheraga (Gō, N.; Scheraga, H. A. *Macromolecules* **1970**, *3*, 178) has been modified to allow bond angle variations. This modification greatly increases the domain of applicability of the algorithm. Of particular interest is its use for bridging deletions or introducing additions in the homology modeling of proteins. Examples from an α-helix, a β-sheet, a cyclic polypeptide, and several proteins are presented.

1. Introduction

In the modeling of a homologous or mutant protein from a known X-ray structure, a deletion or insertion in the sequence requires the construction of a peptide backbone across two fixed endpoints. Generally, this problem is solved by energy minimization,[1] by plastic or metal models (which solve part of the minimization problem mechanically),[2] or by human manipulation of a model using molecular graphics.[3] A related problem arises in the conformational search of the possible structures for a polypeptide loop or cyclic polypeptide. In such a search, some of the backbone torsions in the loop are varied freely and the remainder must be assigned to values such that the loop closes back onto the attached end.

The ring closure and local conformational deformation procedure of Gō and Scheraga[4] solves this problem algorithmically. It determines the set of dihedral angles of a polymer chain required to bridge the chain across two fixed, oriented endpoints. Their procedure assumes that the bond lengths and bond angles of the polymer chain are fixed. We have found that this restriction is very severe in that it greatly reduces the range of endpoints for which a closure can be found. It is known that bond angles do vary by a few degrees in biopolymers[5,6] and that the entropy of proteins has significant contributions from bond angle variations.[7,8] We have developed a simple modification to the Gō and Scheraga procedure that allows small bond angle variations, and we illustrate its applicability by choosing examples from known protein structures.

Section 2 describes the method, and Section 3 presents the results and discussion.

2. Method

In this section, we develop the chain-closure algorithm including both dihedral angle and bond angle variation. To make clear the nature of our modifications to the original approach of Gō and Scheraga, we must briefly review their procedure. In what follows, we refer to equation n in their paper by the notation eq GSn; details are omitted, and the interested reader should consult the original paper.[4] For the free torsion angles of a polymer chain, we define coordinate systems as shown in Figure 1. For atom i, the x axis is along the bond vector between atom $i-1$ and the succeeding atom, the y axis is the component of the succeeding bond vector perpendicular to the x axis, and the z axis is given by the cross product of the other two axes. The bond vector between atom $i-1$ and the succeeding atom is the axis of rotation for the free torsion angle, ω_i.[9] In the case of a protein, the peptide torsion angle (i.e., the torsion angle about the C–N bond) is assumed to be planar and fixed (either *cis* or *trans*) so one atom in the backbone does not need to have a coordinate set defined for it. Gō and Scheraga omit the carbonyl carbon, and thus, the α-carbon and the peptide nitrogen are used for defining the coordinate systems.

The essence of the Gō and Scheraga procedure is the transormation of the coordinate systems given by eq GS1–3

$$\mathbf{r}_{i-1} = \mathbf{T}_{i-1}\mathbf{R}_i\mathbf{r}_i + \mathbf{p}_{i-1} \qquad \text{(GS1)}$$

$$\mathbf{T}_{i-1} = \begin{bmatrix} \cos\theta_{i-1} & -\sin\theta_{i-1} & 0 \\ \sin\theta_{i-1} & \cos\theta_{i-1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{(GS2)}$$

$$\mathbf{R}_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega_i & -\sin\omega_i \\ 0 & \sin\omega_i & \cos\omega_i \end{bmatrix} \qquad \text{(GS3)}$$

$\mathbf{p}_{i-1}$ is a vector that specifies the translation between two coordinate systems. In the case of peptides, the z component of $\mathbf{p}_{i-1}$ is zero, but the y component is nonzero for translations between two coordinate systems that skip the carbonyl carbon (as shown for $\mathbf{p}_{i-1}$ on the left in Figure 1). $\mathbf{T}_{i-1}$ specifies the rotation associated with the angle, θ_{i-1}, and operates in the xy plane. $\mathbf{R}_i$ specifies the rotation associated with a torsion angle, ω_i, and operates in the yz plane.[10] In these equations, the degrees of freedom are the torsion angles, ω_i. For a peptide, Gō and Scheraga define the atom numbering such that the odd-numbered torsions are the φ torsion angles and the even-numbered torsions are the ψ torsion angles.

If we know the position of two amino acid residues that we wish to connect with a peptide chain containing a known number of residues, the problem of chain closure is equivalent to finding a series of coordinate transformations that transform one endpoint coordinate system into the other. To transform one arbitrary coordinate system into another requires six degrees of freedom, three

† Present address: Cellular and Molecular Research Laboratory, Massachusetts General Hospital, Boston, MA 02114.

0024-9297/85/2218-2767$01.50/0