Carbohydrate Research, 243 (1993) 99–114
Elsevier Science Publishers B.V., Amsterdam

Structure of *O*-(2-hydroxyethyl)celluloses with high degrees of molar substitution: a critical evaluation

Peter W. Arisz *, James A. Lomax [1] and Jaap J. Boon

Unit for Mass Spectrometry of Macromolecular Systems, FOM Institute for Atomic and Molecular Physics, Kruislaan 407, 1098 SJ Amsterdam (Netherlands)

(Received March 21st, 1992; accepted September 26th, 1992)

ABSTRACT

Methylation analysis complete with GLC–MS of three *O*-(2-hydroxyethyl)celluloses with degrees of molar substitution (ms) of 1.73, 2.44, and 3.56, determined by hydrolysis with hydriodic acid, revealed at least 35 new monomeric compounds. Except for S_{33333}, all of the isomers up to the monomers with ms 5 have been identified and quantified by GLC. The ms, degree of substitution (ds), positional ds, and the ms and ds distribution of the monomers have been derived. The resulting ms values were 1.89, 2.30, and 3.39, and the ds values were 1.21, 1.34, and 2.06, respectively. The relative reaction constants for HO-2,3,6 have been computed and that for HO-2 appears not to vary. These calculated constants have been used to determine the homogeneity of the substituent distribution.

INTRODUCTION

O-(2-Hydroxyethyl)celluloses (HECs) are commercial products which are used widely as adhesive and thickening agents. The synthesis of an HEC is a statistical process[1–3] in which HO-2,3,6 of the glucose residues react with ethylene oxide to form 2-hydroxyethyl substituents, the hydroxyl groups of which may also react with ethylene oxide to give poly(ethylene glycol)-like side chains. These reactions result in a large variety of monomeric units. The monomer composition and sequence determine the functional properties of the polymers[4], such as rheological behaviour and biodegradability[5,6]. The monomer composition becomes more complex when two-step processes are applied for the synthesis of HECs. Such HECs cannot be described adequately by the statistical models and independent parameters are necessary.

* Corresponding author.
[1] Present address: Rowett Research Institute, Bucksburn, Aberdeen AB2 9SB, UK.

HECs may be described in terms of the mole fraction of unsubstituted glucose residues (u%), the degree of substitution (ds), i.e., the average number of substituents per glucose residue, and the degree of molar substitution (ms), i.e., the average number of ethylene oxide units per glucose residue. The average chain length of the substituents is given by the ratio ms/ds. Although these parameters are correlated with some important properties of HECs, such as the biodegradability, they do not define the structures fully.

The study now reported is concerned with the detailed monomer composition of HECs with high ms values. The availability of such data allows the derivation of ms, ds, ds variously at HO-2,3,6 (pds), and the ms and ds distribution of the monomers, which will describe the structure of HECs accurately. Lindberg et al[7] identified the major monomers up to ms 4 in HECs with ms values up to 2.8 by methylation analysis. However, for an HEC with ms 3.6, there is a significant increase in the fraction of higher hydroxyethylated monomers. These monomers have been identified together with several isomers having ms 3 and 4 not reported previously.

The refinement of the methylation analysis involves a more complete interpretation of the mass spectra of the monomers, the development of an analysis technique based on regularities in their behaviour in GLC, the determination of the errors in the quantification, and a faster procedure for derivatisation which is suitable for HECs of high molecular weight. The method is illustrated by the monomer analysis of three HECs with ms values of 1.73, 2.44, and 3.56, determined by the hydriodic acid method.

RESULTS AND DISCUSSION

Identification.—Fig. 1 shows the gas chromatogram of the monomers obtained on methylation analysis of an HEC with MS 3.56. The numbers above the peaks refer to the monomers in Table I. The ms values of the monomers were determined by GLC–CIMS. Monomers with the same ms are separated into sets of isomers. The theoretical number of isomers in each set is $0.5(i^2 + 3i + 2)$, where i is the ms of the monomer. Fig. 1 shows that the fraction of monomers with ms > 4 is significant. These monomers include a large fraction with derivatised hydroxyethyl groups which therefore contribute substantially to the ms and ds.

The general structure of the monomers, shown in **1**, involves 1,4,5-tri-*O*-acetyl-D-glucitol-*1-d* with $(CH_2CH_2O)_xMe$ groups attached to positions 2, 3, or 6 and where x is k, l, or n, respectively. The monomer notation involves an S with a subscript for the positions and number of CH_2CH_2O units (cf. ref 7). Thus, subscript 2 connotes an $MeOCH_2CH_2$ substituent at position 2 in **1**, 22 connotes an $MeOCH_2CH_2OCH_2CH_2$ substituent at position 2, 222 connotes an $MeOCH_2CH_2OCH_2CH_2OCH_2CH_2$ substituent at position 2, etc. For example, S_{266} indicates an $MeOCH_2CH_2$ substituent at position 2 and an $MeOCH_2CH_2O$-CH_2CH_2 substituent at position 6.

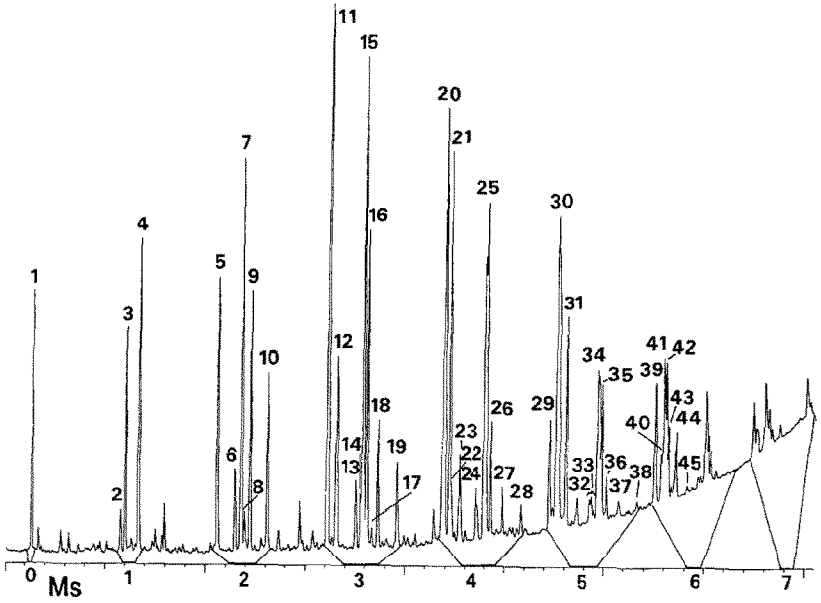
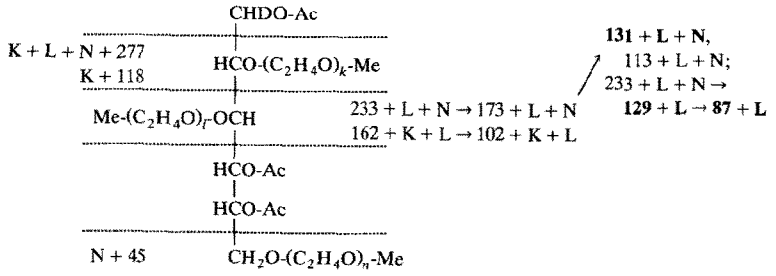


Fig. 1. The gas chromatogram of the monomers obtained by methylation analysis of an HEC with ms 3.56. The peak numbers refer to Table I.



1

The structure **1** has a distinctive fragmentation pattern in 70-eV EIMS[8]. The deuterium at position 1 facilitates the identification of fragment ions in which C-1 is incorporated. Many of the fragment ions shown in **1** were reported previously[7]. The newly identified fragment ions are given in bold type in **1** and in the text. The primary ions *m/z* 162 and 233 may lose a molecule of acetic acid via a McLafferty elimination[9] to give the secondary ions *m/z* 102 and 173, respectively. In addition, the secondary fragment ion *m/z* 173 may lose ketene or acetic acid to give the ions *m/z* **131** or 113, respectively. Analogous to the McLafferty elimination of acetic acid, the primary fragment ion *m/z* 233 may also lose methoxymethyl acetate, as illustrated in Scheme 1 (cf. ref 10).

The fragmentation of 1,4,5-tri-*O*-acetyl-2,3,6-tri-*O*-methyl-*i*-glucitol is influenced little by the insertion of CH_2CH_2O groups at positions 2, 3, and 6. Therefore, the masses of the fragment ions of the substituted monomers are given by the masses

TABLE I

Identified monomers and their corresponding mole fractions in the HECs

Peak number [a]	Monomer index (S)	Retention time (t_R) [b]	i [c]	j [d]	HEC 1.73 Mole fraction	HEC 2.44 Mole fraction	HEC 3.56 Mole fraction
1	0	1696	0	0	0.2289	0.1869	0.0377
2	3	2020	1	1	0.0152	0.0104	0.0057
3	2	2037	1	1	0.0934	0.0705	0.0329
4	6	2085	1	1	0.1288	0.1131	0.0686
5	23	2339	2	2	0.0233	0.0203	0.0356
6	36	2398	2	2	0.0125	0.0105	0.0109
7	26	2420	2	2	0.0582	0.0490	0.0536
8	33	2430	2	1	0.0040	0.0035	0.0037
9	22	2452	2	1	0.0418	0.0354	0.0330
10	66	2508	2	1	0.0734	0.0780	0.0278
11	236	2699	3	3	0.0188	0.0181	0.0722
12	233	2722	3	2	0.0181	0.0179	0.0278
	223	2728	3	2			
13	336	2784	3	2	0.0032	0.0025	0.0079
14	366	2804	3	2	0.0072	0.0059	0.0052
15	226	2814	3	2	0.0270	0.0247	0.0637
16	266	2823	3	2	0.0352	0.0472	0.0403
17	333	2837	3	1	0.0025	0.0028	0.0021
18	222	2857	3	1	0.0180	0.0193	0.0136
19	666	2916	3	1	0.0270	0.0399	0.0108
20	2336	3052	4	3	0.0154	0.0181	0.0636
	2236	3059	4	3			
21	2366	3075	4	3	0.0105	0.0150	0.0408
22	2233	3080	4	2	0.0033	0.0018	0.0043
23	2333	3100	4	2	0.0081	0.0096	0.0127
	2223	3106	4	2			
24	3366	3156	4	2	0.0022	0.0051	0.0088
n.i. [e]	3336	3161	4	2	0.0014	0.0009	0.0040
25	2266	3185	4	2	0.0279	0.0347	0.0627
	3666	3188	4	2			
	2226	3190	4	2			
26	2666	3204	4	2	0.0096	0.0141	0.0105
n.i. [e]	3333	3220	4	1	0.0001	0.0003	0.0001
27	2222	3240	4	1	0.0048	0.0068	0.0041
28	6666	3297	4	1	0.0071	0.0141	0.0031
29	22336	3376	5	3	0.0034	0.0043	0.0128
30	23366	3399	5	3	0.0156	0.0216	0.0607
	23336	3404	5	3			
	22366	3405	5	3			
	22236	3409	5	3			
31	22333	3429	5	2	0.0063	0.0091	0.0179
	22233	3429	5	2			
	23666	3431	5	3			
32	23333	3458	5	2	0.0023	0.0028	0.0025
	22223	3464	5	2			
33	33366	3503	5	2	0.0007	0.0033	0.0052
	33666	3508	5	2	0.0010	0.0006	0.0008
	33336	3517	5	2	0.0004	0.0006	0.0013

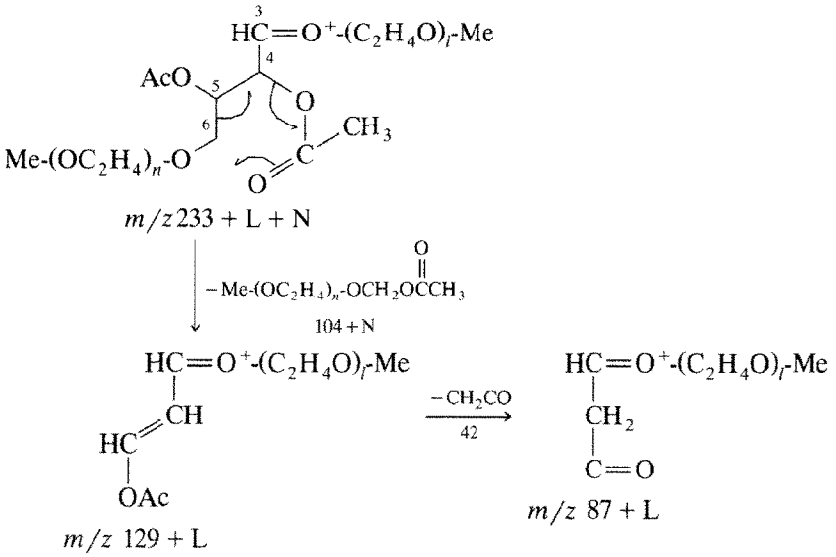
TABLE I (continued)

Peak number [a]	Monomer index (S)	Retention time (t_R) [b]	i [c]	j [d]	HEC 1.73 Mole fraction	HEC 2.44 Mole fraction	HEC 3.56 Mole fraction
34	22266	3531	5	2	0.0103	0.0154	0.0264
	22666	3535	5	2			
35	22226	3542	5	2	0.0042	0.0063	0.0098
	36666	3543	5	2			
36	26666	3558	5	2	0.0026	0.0049	0.0033
37	22222	3595	5	1	0.0011	0.0019	0.0011
38	66666	3651	5	1	0.0015	0.0035	0.0009
39	223366	3696	6	3	0.0037	0.0049	0.0118
	223336	3698	6	3			
	222336	3699	6	3			
40	233366	3719	6	3	0.0010	0.0023	0.0049
41	222366	3727	6	3	0.0039	0.0073	0.0154
42	233666	3734	6	3	0.0024	0.0037	0.0099
	223666	3734	6	3			
43	222236	3740	6	3	0.0015	0.0017	0.0036
44	236666	3765	6	3	0.0016	0.0033	0.0051
45	222223	3800	6	2	0.0002	0.0006	0.0003
n.i. [e]	333666	3820	6	2	0.0001	0.0000	0.0002
	R6		6	2	0.0053	0.0146	0.0171
	R7		7	3	0.0042	0.0112	0.0211

[a] Peak numbers refer to Fig. 1. [b] In scan numbers (0.9 s/scan). [c] Number of attached CH_2CH_2O units. [d] Number of substituents. [e] Not indicated in Fig. 1.

of the fragment ions of the parent glucitol derivative plus the K, L and N values (i.e., for 44 times k, l, and n, respectively) as shown in **1**. The secondary fragment ions provide the same information as the primary fragment ions, except for m/z **129** + L and **87** + L which are formed from m/z 233 + L + N, as shown in Scheme 1. The ms values for positions 3 and 6 can be established unambiguously by these fragment ions.

Information about the length of the longest substituent in a monomer is obtained by the ion series Cx with the masses m/z 59 + X, where X = K, L, or N. Complementary information is given by the fragment ion m/z **M** − **Cx**, where M is the molecular ion. Each monomer has a unique set of fragment ions, which is illustrated by the mass spectra of the isomers S_{236} and S_{222} (Fig. 2). The K, L, and N values are 132, 0, 0, respectively, for S_{222}, and 44 for S_{236}. In general, the C1 fragment, m/z 59, is the base peak in the mass spectra. The Cx series m/z 59, 103, and 147 in the spectrum of S_{222} reveals that the longest substituent contains three CH_2CH_2O units. This inference is confirmed by the fragment ion m/z **320**, corresponding to S_{222}(M) − OC3. Most of the other ions in the spectra can be assigned to one of the fragments derived from the glucitol skeleton (**1**). The most abundant fragment ions of **1** are always m/z 118 + K and 233 + L + N, which can be used routinely for the identification of trace compounds by mass chromatography.



Scheme 1

Many monomeric compounds were identified by mass chromatography on specific ions[7]. However, several isomers could not be identified because (*a*) mass chromatography cannot be used when more than two compounds coelute or when the difference in the retention times of two compounds is less than the chromatographic time resolution of the GLC–MS, i.e. 0.9 s, which was chosen as the cycle time of the mass spectrometer; and (*b*) the signal-to-noise ratio is insufficient for trace compounds, such as S_{3333}. These problems can be solved by an analysis of the

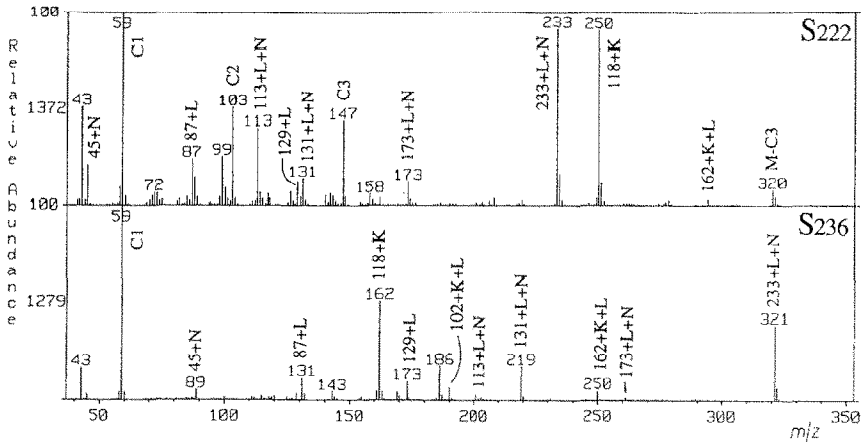


Fig. 2. The 70 eV EI-mass spectra of the monomers S_{222} and S_{236}. Most of the mass peaks correspond to the fragment ions predicted by I or to fragment ions of the side chains (Cx-series). The values of K, L, and N for S_{222} are 132, 0, and 0, respectively, and they are all 44 for S_{236}.

TABLE II

Retention behaviour of the homologous series [a] with the basic elements S_0 and S_2

	S_0						S_2			
	Substitution position						Substitution position			
	2		3		6		3		6	
i [b]	t_R [c]	Δt_R [d]	t_R	Δt_R	t_R	Δt_R	t_R	Δt_R	t_R	Δt_R
0	0		0		0					
		341		324		389				
1	341		324		389		341	302	341	383
		415		410		423				
2	756		734		812		643	383	724	
		405		**407**		**408**				**403**
3	1161		1141		1220		1026		1127	
		383		**383**		**381**		**376**		**381**
4	1544		1523		1601		1404		1508	
		355				**354**		**358**		**354**
5	1899				1955		1762		1862	

[a] S_0 is the non-substituted glucitol skeleton (**1**), S_2 is the *O*-(2-hydroxyethyl)glucitol skeleton. [b] Number of CH_2CH_2O units. [c] Retention time expressed in scans relative to that of S_0. [d] Difference in retention time between two subsequent members of the series.

retention behaviour of the monomers, which allows the retention times of missing compounds to be predicted.

The mixtures of monomers contain various homologous series, each of which involves an increasing number of CH_2CH_2O groups inserted at a particular position (2, 3, or 6). As examples, the homologous series with the basic elements S_0 and S_2 are given in Table II. A comparison of the Δt_R values for these series reveals that they become independent of the position of the substituents when more than three CH_2CH_2O units are attached to the glucitol skeleton. This regularity in the Δt_R values allows the retention time of the next member in an homologous series to be predicted. Thus, for the trace compound S_{3333}, which could not be identified by the mass spectrometric data because of the low signal-to-noise ratio, the data in Table II show that the Δt_R value between S_{333} (t_R 1141) and S_{3333} has to be 382 scans. The tentative identification of this minor peak in the chromatogram with the predicted retention time was confirmed by mass chromatography of the most abundant fragment ions m/z 118 + K and 233 + L + N. The actual measured Δt_R value was found to be 383 scans. Another example is S_{233666}, which has a retention time exactly the same as that of the major compound S_{223666}. The deviation in the predicted retention time of the compounds identified was never > 4 scans, corresponding to an accuracy of 3.6 s.

The reliability of the identification of the monomeric compounds was enhanced by fitting their retention times into the empirical homologous series. Any incorrect interpretation of a monomer was recognisable by an irregularity in the Δt_R values. The sequence of elution of monomers was the same as that observed by Lindberg et al.[7], with the exception of S_{333} which may have been confused with S_{366} since their mass spectra contain similar major fragment ions. However, these isomers can be distinguished unambiguously by their secondary ions m/z **129** + **L** and **87** + **L**. The fragment ions m/z 162 + K + N and 102 + K + N, proposed[7] for the

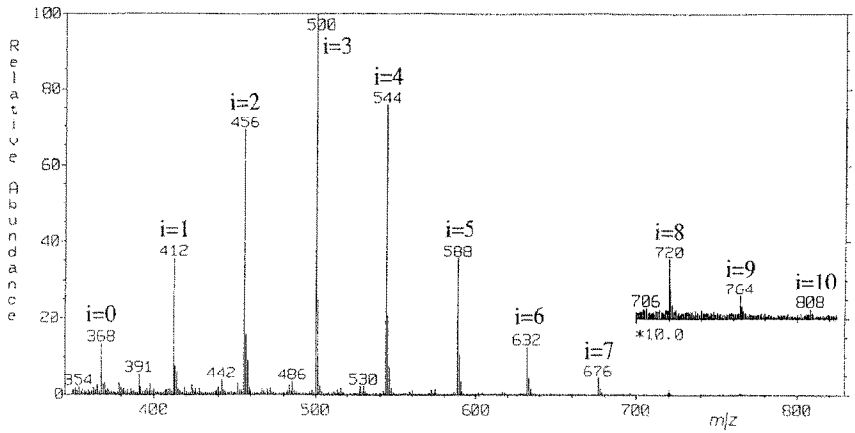


Fig. 3. The DCI(ammonia)-mass spectrum of the derivatised monomers of HEC 3.56.

determination of the positions of substitution, were not present in the spectrum of S_{333}. Also, the retention time of S_{333} points to the identity of S_{366}, an isomer which was not reported[7].

The monomers identified and their retention times, expressed in scan numbers, are listed in Table I. All the isomers up to monomers with ms 5 have been identified, except for S_{33333}. The identification of monomers with ms > 6 was difficult due to their relatively low abundances and limited transmission through the GLC–MS system. Possible transmission losses by the GLC capillary column were checked by DCIMS (see Fig. 3) of the monomers obtained from HEC 3.56. The ion series corresponds to the ammonium-ion adducts of the monomers, which were reduced in this sample with sodium borohydride instead of borodeuteride (see Experimental). The ms of the monomers (i) is indicated above the peaks. This spectrum reveals that the abundance of monomers with ms > 7 is low. The intensity of the mass peaks for ms 8 is only 1.57% of that of the base peak m/z 500, which corresponds to ms 3.

Quantification.—The quantitative data were derived from the peak areas measured by GLC fitted with an FID. In order to correct for the effective carbon numbers (ECN) of the monomers, the peak areas measured were multiplied by a molar response factor (MRF), which consists of the ECNs of the glucitol skeleton (7.25) and additional CH_2CH_2O units (1.03) and is given by eq *1* where i is the number of CH_2CH_2O units.

$$\mathrm{MRF} = \frac{7.25}{7.25 + 1.03i} \qquad (1)$$

The ECN of the 1,4,5-tri-*O*-acetyl-2,3,6-tri-*O*-methyl-*i*-glucitol appears not to have been reported, but an approximate value can be calculated, using the empirical rules[11], as 7.25 by taking the contribution of the acetyl esters as the sum of the alcohol and the acetyl group. However, Sweet et al.[12] have reported a value

of 7.40. The experimental value[13] of the ECN of 1.57 for methyl acetate is in good agreement with the value assumed for AcO-1 in the glucitol skeleton, but also a value of 1.04 for methyl acetate has been reported[14]. The average ECN of inserted CH_2CH_2O units was derived from the literature[14,15] to be 1.03. There are indications that the ECN of the CH_2CH_2O units is dependent on the length of the chain in which they are incorporated. The increment of the ECN by an additional CH_2CH_2O is 4–7%, an effect which has been neglected in this paper.

The mole fractions of the monomers of HECs with ms 1.73, 2.44, and 3.56, listed in Table I, were calculated by dividing the corrected peak areas (measured peak area × MRF) by the total corrected peak area. The peak areas of monomeric compounds that were eluted together are combined. The peak areas of the unidentified monomers in the ms 6 and 7 regions are indicated as R6 and R7. The reliability and reproducibility of the monomer data were checked by the determination of the errors in the calculated ms values (Table III) introduced at the various stages of the procedure.

The reproducibility of the synthesis and purification processes of the methylation analysis was examined by duplicate analyses of the HECs 2.44 and 3.56; the deviations in the calculated ms values were 8 and 3%, respectively. This error could reflect the inhomogeneity of the samples and/or the losses of particular monomers during synthesis and purification. The inhomogeneity of the samples was reduced by ball mailing, and losses were minimised by using low temperatures (40–50°C) for the evaporation of solvents and by extracting the samples several times.

The reproducibility of the GLC measurements and the integration of the chromatograms was determined by duplicate measurements of 9 samples with different ms values. The deviation in the calculated ms values was 0.47% with a standard deviation of 0.21.

Fig. 4 shows the ms distribution of the monomers calculated from the data in Table I. The calculated distribution for HEC 3.56 is similar to that in the DCI-mass spectrum (Fig. 3). The monomers with ms > 7 could not be measured by GLC-FID and, in order to estimate their mole fractions, a curve was fitted through the data points of Fig. 4 and the mole fraction of the monomers with ms 12 was adjusted to zero. The mole fractions with ms 8 and 9 for HEC 3.56 were found to be 0.048 and 0.008, respectively. If these mole fractions are taken into account, then the calculated ms of HEC 3.56 becomes 3.39 and that for HEC 2.44 becomes 2.30. The absence of the measurements of highly substituted monomers (ms > 7) introduces systematic errors of −1.8% and −0.9% in the calculated ms value of HEC 3.56 and 2.44, respectively, which can be corrected.

There is disagreement in the literature concerning the ECNs of the glucitol skeleton and, to a lesser extent, the additional CH_2CH_2O units. The difference between our calculated values of ms and those reported elsewhere could be due to the use of incorrect ECNs. Therefore, the sensitivity of the ms calculated for HEC 2.44 to deviations in the ECNs of the glucitol skeleton (7.25) and the CH_2CH_2O

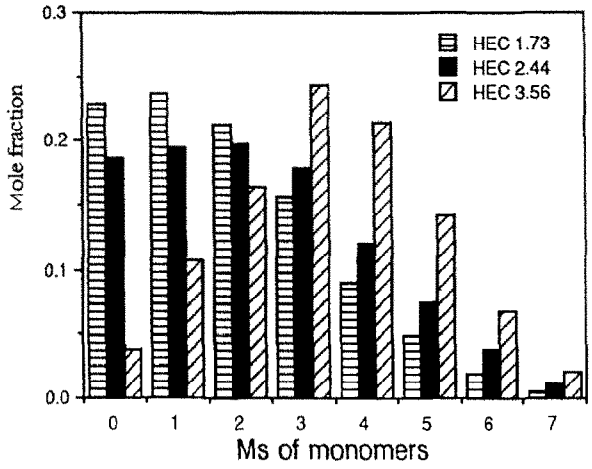


Fig. 4. The ms distribution of the monomers calculated from the GLC data in Table I.

units (1.03) was studied and is shown in Fig. 5. A deviation of 10% leads to an error of only 1% in the ms. Lindberg et al.[7] used ECNs of 7.40 for the glucitol skeleton and 1.00 for the CH_2CH_2O units, and use of their data results in a ms of 2.29 for HEC 2.44, which is a difference of $< 0.5\%$. Hence, small deviations in the ECNs result in insignificant changes in the calculated ms values.

The ms value calculated for HEC 1.73 is higher than expected which may reflect relatively greater losses in the synthesis and purification processes. The ms values

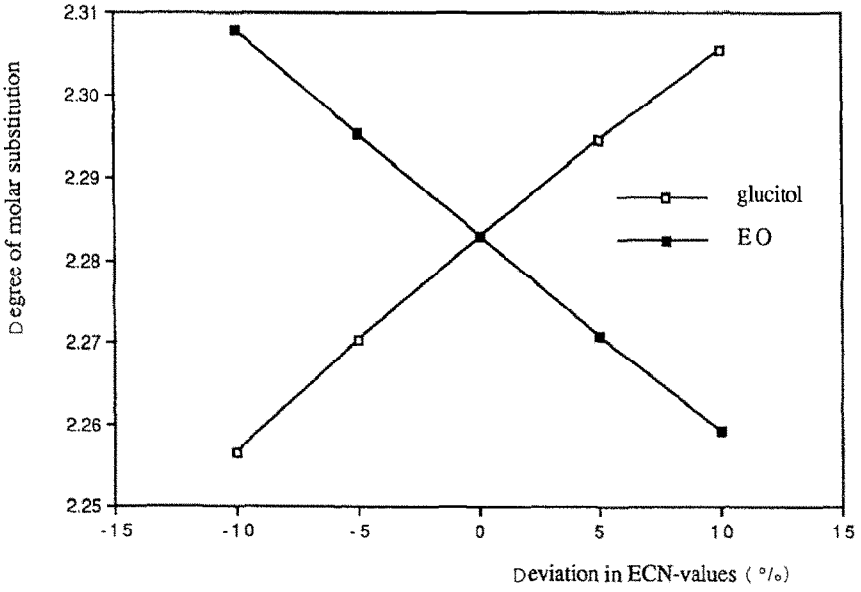


Fig. 5. The sensitivity of the calculated ms of HEC 2.44 to deviations in the ECNs of the glucitol skeleton (7.25) and the inserted CH_2CH_2O (EO) units (1.03).

calculated for HEC 2.44 and 3.56 are slightly below those obtained using the hydriodic acid procedure, which could be due to underestimation of high molecular weight compounds by the FID[16].

The errors in the calculated ms values are ±8% for the duplicate methylation analyses, and ±0.47% for the GLC-measurements and the integration of the chromatograms. A systematic maximal error of −1.8% (HEC 3.56) was introduced by the lack of data for the monomers with ms > 7. This error can be corrected and will be smaller for HECs with a lower ms. The error that is introduced by incorrect ECNs is estimated to be in the range ±1%. Thus, the overall error in the method will be < 10%. The ms values of 1.89, 2.30, and 3.39 calculated for HECs 1.73, 2.44, and 3.56, respectively, are in good agreement with reference values determined by the hydriodic acid method.

Structural parameters.—A series of parameters which characterise the HECs can be derived from the monomer data in Table I. The classical parameters, which are normally used, are *u*%, ms, and ds. The value of *u*% is given directly by S_0, and the ms and the ds values of the samples are derived according to eqs *2* and *3* where *n* is the number of measured monomers, *i*(*n*) is the number of CH_2CH_2O units in monomer *n*, *j*(*n*) is the number of substituents in monomer *n*, and mf(*n*) is the mole fraction of monomer *n*. The results are listed in Table III. The ms is usually determined by the hydriodic acid method[17] and the values derived from the GLC-FID data are in good agreement with the reference values. The assumption made for the calculation of the ds values is that the unidentified monomers R6 and R7 have ds 2 and 3, respectively. NMR spectroscopy can also be used to determine the ds, although assumptions have to be made because of overlapping signals in the spectra. The ds value of 1.34 for HEC 2.44 is significantly higher than that (1.13) determined[3] by NMR spectroscopy for an HEC with a comparable ms of 2.57. It is not evident whether this variation is due to different response factors of the analysis techniques or to the compositions of the samples. There are no other reliable wet chemical methods for the determination of ds[6].

$$\mathrm{ms} = \sum_{n=1}^{n} i(n) \cdot \mathrm{mf}(n) \qquad (2)$$

$$\mathrm{ds} = \sum_{n=1}^{n} j(n) \cdot \mathrm{mf}(n) \qquad (3)$$

The structures of HECs can be described more accurately by such parameters as the pds values, which are given by the sum of the monomers in which position x_i is substituted, the pms values, and the ms and ds distributions of the monomers. These parameters are especially suitable for the description of samples which are manufactured by an unavailable process, such as the Aqualon HECs.

Table IV shows the pds values and related parameters derived from our samples and those of Lindberg et al.[7]. Our pds values were derived from the data in Table I, but, because of the lack of resolution of some compounds, the

TABLE III

Classical parameters of the HECs calculated from the data in Table I

	u%	Ms	Ds	Ms/ds	Ms determined by HI-procedure [a]
HEC 1.73	22.9	1.89±0.01 [b]	1.21	1.56	1.73
HEC 2.44	18.7	2.28±0.18 [c]	1.34±0.08 [c]	1.70	2.44
HEC 3.56	3.6	3.33±0.11 [c]	2.06±0.05 [c]	1.62	3.56

[a] Determined by Aqualon. [b] Average value from duplicate measurements. [c] Average values from duplicate methylation analyses.

monomers S_{3666}, S_{22333}, S_{22233}, S_{36666}, R6, and R7 were ignored for the calculation. The difference between the ds values calculated from the total and the selected list (sum pds) of monomers is $< 5\%$. The pds values can also be derived from NMR data[18], but, as for the ds measurements, there are problems due to overlapping signals.

The pds values have been used frequently in the literature to calculate relative reaction constants. However, the meaning of such constants is doubtful for HECs of unknown origin. For samples synthesised by a two-step process, the calculated reaction constants are composite values. Consequently, the distribution of substituents in an unknown sample would not be described adequately by the statistical models. On the assumption that the samples studied here were prepared by a one-step process, the relative reaction constants in Table IV could be calculated according to model I described by Reuben[19]:

$$k_i = -\ln(1 - \mathrm{pds}i) \tag{4}$$

$$k_x = \frac{\mathrm{ms} - \mathrm{ds}}{3 - \mathrm{pds2}/k_2 - \mathrm{pds3}/k_3 - \mathrm{pds6}/k_6} \tag{5}$$

The value of k_x is calculated from the complete data set, which is more representative than a computation from the equations of only one of the reaction pathways used by Lindberg et al.[7]. This situation is illustrated by the re-calculated k_x values of the samples used by Lindberg et al.[7], presented in Table IV, which are significantly lower (1.8, 2.1, 2.5, and 0.42) than those (2.5, 2.7, 2.9, and 0.42)

TABLE IV

pds-Values and derived parameters of *A* our samples and *B* those of Lindberg et al. (ref 7)

Sample	pds2	pds3	pds6	sum ds	%pds2	%pds3	%pds6	k_2	k_3	k_6	k_x
A HEC 1.73	0.473	0.190	0.525	1.188	39.8	16.0	44.2	1.000	0.328	1.162	1.630
HEC 2.44	0.492	0.209	0.583	1.284	38.3	16.2	45.4	1.000	0.345	1.292	1.946
HEC 3.56	0.764	0.463	0.745	1.973	38.8	23.5	37.8	1.000	0.431	0.944	0.747
B Lindberg 1.3	0.166	0.041	0.225	0.432	38.4	9.5	52.1	1.000	0.231	1.404	1.800
Lindberg 2.3	0.258	0.082	0.360	0.700	36.9	11.7	51.4	1.000	0.287	1.496	2.145
Lindberg 3.3	0.325	0.151	0.498	0.974	33.4	15.5	51.1	1.000	0.416	1.753	2.496
Lindberg 2.8	0.785	0.415	0.733	1.933	40.6	21.5	37.9	1.000	0.349	0.859	0.421

reported[7]. Reuben and Casti[3] have shown that HECs obey model II rather than model I. However, a computation of the relative reaction constants according to model II could not be accomplished, because the formulas for the calculation of k_3' have not been reported.

The %pds values of HEC 1.73 and 2.44 are similar. There are strong indications that the reaction conditions used for the preparation of these samples were the same and only the amount of added ethylene oxide or the reaction time was varied[20]. This finding implies that, under heterogeneous conditions of reaction, the relative rates of reaction rather than the relative reaction constants of the hydroxyl groups of the cellulose remained constant. Also, the increase of k_x with increasing ms of the similarly produced samples (1.3, 2.3, and 3.3 in ref 7) supports this idea. The data cast doubt on the validity of all statistical models[1–4,7] which assume that the relative reaction constants remain constant.

Both the pds ratios and the relative reaction constants of HEC 3.56 are different, which implies that this HEC was produced under different conditions. The relative reactivities of positions 6 and x in HEC 3.57 are decreased, whereas that of position 3 is increased, which would imply that HEC 3.57 was produced under alkali concentrations lower than those used for the other samples[21].

Striking in all of the Aqualon samples is the constant %pds2 with values of 38–40%, including eight other samples with different ms values which are not reported here. This phenomenon is also reflected to some extent in the data of Lindberg et al.[7] (see Table IV). The deviations in their data, especially for the higher molar-substituted samples, could be due to the incomplete analysis of the monomers. Because of this constant relative reactivity of position 2, it would be better to express the relative reactivities of the other hydroxyl groups in the cellulose to that of HO-2, rather than to the more-often used HO-3.

The pms values should provide specific information on the reactivity of the hydroxyl groups of the substituents, but the chromatographic resolution is insufficient at present for reliable calculation.

Information on the homogeneity of the distribution of substituents over the cellulose backbone is not given by the structural parameters independent of the statistical theories. According to the statistical models, this distribution is random which should be reflected in the ds and ms distribution of the monomers. The deviation between the theoretical and the experimental distributions will be a measure of the homogeneity.

Fig. 6 shows the differences in ds distributions of the experimental minus the theoretical values, calculated according to model I[2]. It appears that the fraction of non-substituted and tri-substituted monomers are underestimated by model I, whereas those of the mono- and di-substituted monomers are overestimated. These results correspond to a heterogeneous reaction in which the more exposed exterior of the cellulose fibrils is substituted better than predicted by the model, whereas the core of the fibril is poorly derivatised, resulting in larger fractions of tri- and non-substituted monomers, respectively. Although model II[3,4] may fit the

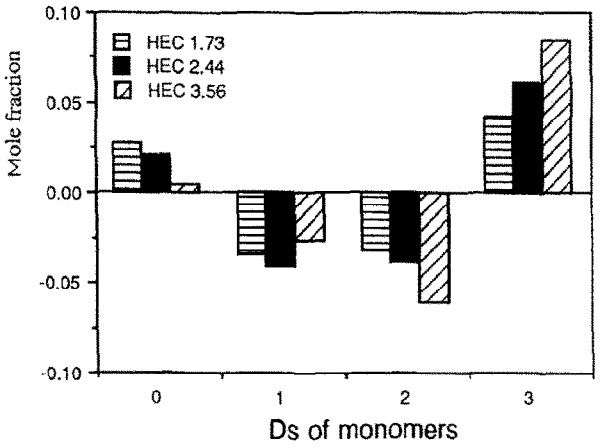


Fig. 6. Differences between the experimental and theoretical (model I) ds distributions.

experimental data more accurately, the result of such a computation will be uncertain, because of shortcomings in the statistical theories, and the real distribution can be provided only by an oligomer analysis.

EXPERIMENTAL

Samples and materials.—The *O*-(2-hydroxyethyl)celluloses were synthesised by the A&D Laboratory of Aqualon BV (Zwijndrecht, Netherlands). The ms values, determined by the hydriodic acid method, were 1.73, 2.44, and 3.56. Dry methyl sulphoxide was obtained from Pierce.

Methylation analysis[22,23].—Each HEC (~ 3 mg) was dried over P_2O_5 in vacuo at 50°C overnight in a screw-capped tube. All reactions were carried out under N_2. Each sample was suspended in dry Me_2SO (1.5 mL) and 2 M potassium methylsulphinylmethanide (dimsyl) in Me_2SO (1.0 mL) was added. Each mixture was shaken for 2 h and then cooled with ice, MeI (63 μL) was added, and mixing was continued for 2 h at room temperature. At this stage, the sample had not dissolved totally. The whole procedure was repeated, using 1.5 mL of 2 M dimsyl, an incubation time of 1 h, 2 mL of MeI, and finally mixing for 64 h. Each red-brown sample (iodine) was then poured into satd aq KCl (3 mL), the aqueous layer was acidified with HCl (pH < 1.5), and the methylated HECs were extracted with CH_2Cl_2 (4 × 2 mL). The combined extracts were washed in sequence with 0.05 M HCl (3 mL), aq 1% $NaBH_4$, and water (4 × 3 mL), and concentrated.

The residue was treated with aq 90% formic acid (1 mL for 5 h at 100°C), the solvent was evaporated under reduced pressure, and the residue was treated with 0.25 M H_2SO_4 (1 mL for 16 h at 100°C). The acid was neutralised with Amberlite IRA-400 (HCO_3^-) resin, the resin was filtered off and washed with water, and the combined filtrate and washings were concentrated.

The pH of a solution of the partially methylated sugars in water (1 mL) was adjusted to >7 with ammonia, and $NaBD_4$ or $NaBH_4$ (2–3 mg) was added. The mixture was stored at room temperature for 3 h, the excess of borodeuteride or borohydride was destroyed with glacial acetic acid, the solvent was evaporated under reduced pressure at 40°C, and borate was removed from the residue by co-distillation with 1% acetic acid in MeOH (4 × 2 mL) followed by MeOH (4 × 2 mL). The residue was dried overnight over P_2O_5 in vacuo at room temperature, then treated with Ac_2O (1 mL for 16 h at 100°C). The Ac_2O was removed by co-distillation with toluene (reduced pressure, 40°C). A solution of the residue in CH_2Cl_2 (3 mL) was washed with water (4 × 1 mL), then concentrated, and the residue was dried over P_2O_5 at 300 mmHg. The sample was redissolved in CH_2Cl_2 (0.1–0.3 mL) and an aliquot (0.3 μL) was used for GLC-FID or GLCMS.

Gas chromatography with flame ionisation detection (GLC-FID).—Capillary GLC was performed on a Carlo Erba SFC 3000 gas chromatograph equipped with an on-column injector. During the injection, the carrier gas was held at constant pressure. The system was switched to a constant flow of 2 mL of He/min 3 s after the injection and the secondary cooling of the injector was turned off after 10 s. A 25-m CP Sil 5 fused-silica capillary column (i.d., 0.32 mm; film thickness, 0.20 μm) was used. The temperature programme was 5 min at 50°C, then to 180°C at 20°C/min, and from 180°C to 325°C at 3°C/min. The FID-detector was set at 350°C (air flow, 340 mL/min; H_2 flow, 30 mL/min; and make-up gas, 30 mL of He/min). The data were recorded using a 'Note Book' data acquisition package on an Olivetti M280 PC.

Gas chromatography–mass spectrometry (GLC–MS).—A Hewlett–Packard 5890 Series II gas chromatograph equipped with a temperature programmable on-column injector was used connected to a Finnigan INCOS 50 quadrupole mass spectrometer equipped with an EI or a CI (ammonia) ion source. The conditions for chromatography were as described above. For GLC–MS, the mass range m/z 40–500 was scanned with a cycle time of 0.9 s.

Desorption chemical ionisation-mass spectrometry (DCIMS).—A Jeol DX-303 instrument was used with ammonia as the reactant gas (20 Pa in the ion source). The ion-source temperature was 160°C, the accelerating voltage was 3 kV, the scan slope was 0.5 s/decade, and the cycle time was 1 s. The mass range m/z 60–1000 was scanned with a resolution of 1500 and the post accelerating voltage was set at −5 kV. The filament of the desorption probe was heated at 13.5°C/s (1 A/min). The data were processed with a Jeol DA-5000 data system.

ACKNOWLEDGMENTS

This work is part of the research program of the Foundation for Fundamental Research on Matter (FOM). Support from the Dutch National Programme Committee on Carbohydrates (IOP-k) and from Aqualon B.V. are gratefully acknowledged. We thank Jan Commandeur and Jos Pureveen for technical assistance,

Ivana Pastronova for translating the Czechoslovak article, and Tina Weeding for help with the text.

REFERENCES

1 H.M. Spurlin, *J. Am. Chem. Soc.*, 61 (1939) 2222–2227.
2 J. Reuben, *Macromolecules*, 17 (1984) 156–161.
3 J. Reuben and T.E. Casti, *Carbohydr. Res.*, 163 (1987) 91–98.
4 H.M. Spurlin, in E. Ott, H.M. Spurlin, and M.W. Grafflin (Eds.), *Cellulose and Cellulose Derivatives*, Part II, Interscience, New York, 1954, pp 673–712.
5 M.G. Wirick, *J. Polym. Sci., Part A*, 6 (1968) 1705–1718.
6 D.P. Klug, D.P. Winquist, and C.A. Lewis, *Polym. Sci. Technol.*, 2 (1973) 401–416.
7 B. Lindberg, U. Lindquist, and O. Stenberg, *Carbohydr. Res.*, 170 (1987) 207–214.
8 H. Björndal, B. Lindberg, and S. Svensson, *Carbohydr. Res.*, 5 (1967) 433–440.
9 N.K. Kochetkov and O.S. Chizhov, *Adv. Carbohydr. Chem.*, 21 (1966) 39–93.
10 K. Axberg, H. Åbjorndal, Å. Pilotti, and S. Svensson, *Acta Chem. Scand.*, 26 (1972) 1319–1325.
11 R. Kaiser (Ed.), *Chromatographie in der Gasphase III*, Teil 2, Bibliographisches Institut, Mannheim/Zürich, 1969, pp 252–282.
12 D.P. Sweet, R.H. Shapiro, and P. Albersheim, *Carbohydr. Res.*, 40 (1975) 217–225.
13 R.G. Ackman, *J. Gas Chromatogr.*, 2 (1964) 173–179.
14 J.T. Scanlon and D.E. Willis, *J. Chromatogr. Sci.*, 23 (1985) 333–340.
15 M. Singliar, *Petrochemia*, 21 (1981) 113–120.
16 R.G. Ackman, *J. Gas Chromatogr.*, 6 (1968) 497–501.
17 P.W. Morgan, *Ind. Eng. Chem., Anal. Ed.*, 18 (1946) 500–504.
18 Y. Tezuka, K. Imai, M. Oshima, and T. Chiba, *Polymer*, 30 (1989) 2288–2291.
19 J. Reuben, *Carbohydr. Res.*, 157 (1986) 201–213.
20 P.W. Arisz and J.J. Boon, unpublished results.
21 R. Dönges, *Br. Polym. J.*, 23 (1990) 315–325.
22 S. Hakomori, *J. Biochem. (Tokyo)*, 55 (1964) 205–208.
23 S.C. Fry, in M. Wilkins (Ed.), *The Growing Plant Cell Wall: Chemical and Metabolic Analysis*, Wiley, New York, 1988, p 141.