Carbohydrate Research, 168 (1987) 67–77
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

CASPER—A COMPUTERISED APPROACH TO STRUCTURE DETERMINATION OF POLYSACCHARIDES USING INFORMATION FROM N.M.R. SPECTROSCOPY AND SIMPLE CHEMICAL ANALYSES*

PER-ERIK JANSSON, LENNART KENNE, AND GÖRAN WIDMALM
Department of Organic Chemistry, Arrhenius Laboratory, University of Stockholm, S-106 91 Stockholm (Sweden)

(Received November 25th, 1986; accepted for publication, April 16th, 1987)

ABSTRACT

A computer program, CASPER, has been developed for the determination of the structure of polysaccharides composed of regular repeating-units. The program utilises the ^{13}C-n.m.r. spectrum of the polysaccharide and information from sugar and methylation analyses. Based on the identity of the monosaccharides present in the repeating unit and the positions of the linkages, all possible permutations are constructed. With the aid of a database containing ^{13}C-n.m.r. substituent shifts obtained from disaccharides, the spectra of the alternative structures are simulated and the best fit with the observed spectrum is selected. The program has been tested on two polysaccharides of known structures, for both of which the correct structure was selected.

INTRODUCTION

Until recently, the elucidation of polysaccharide structures involved specific degradations and analysis of the products. In recent years, n.m.r. spectroscopy has become increasingly important in such studies and the structures of several oligo- and poly-saccharides have been determined mainly or exclusively by n.m.r. techniques. The n.m.r. data were generally compared with data obtained from monosaccharide derivatives. Compilations of ^{13}C-n.m.r. spectra[1,2] and the rationalisation of some substituent shifts[3] have been reported.

The ^{1}H- and ^{13}C-n.m.r. chemical shifts of the resonances of a monosaccharide residue within a larger saccharide depend mainly on the structure of the monosaccharide and on the nature of the flanking sugar residues. In ^{1}H-n.m.r. spectroscopy, the resonances of some protons have chemical shifts outside the "bulk-region" (δ 3.5–4) and these resonances have been assigned to structural reporter

*Presented at the XIIIth International Carbohydrate Symposium, Cornell University, Ithaca, New York, U.S.A., August 10–15, 1986.

0008-6215/87/$ 03.50

groups[4] which include anomeric protons, H-2 of mannose, and H-5 of fucose. In ^{13}C-n.m.r. spectroscopy, a completely resolved spectrum can often be obtained and, when signals overlap, the integrated intensity will establish the number of carbons associated with each signal. Thus, all of the resonances can be observed.

The structures of unknown substances may be determined from their n.m.r. spectra when accurate data for reference substances are available. Computer programs for the analysis of n.m.r. spectra have been reported[5–7], but these deal mainly with low-molecular-weight compounds. Computerised approaches to the analysis of the structure of oligosaccharides by high-resolution n.m.r. spectroscopy, with special reference to *O*- and *N*-linked chains of glycoproteins, have been reported[8,9].

We now report on a computer program designed to suggest possible structures of oligosaccharide repeating-units in polysaccharides from n.m.r. data and information obtained from sugar and methylation analyses. In its current design, the program uses ^{13}C-n.m.r. data and may be applied to unbranched structures with at most six sugars in the repeating unit. Predictions for oligosaccharides can also be performed.

The acronym CASPER, used for the program, stands for Computer Assisted Spectrum Evaluation of Regular polysaccharides.

RESULTS AND DISCUSSION

General overview. — CASPER consists of three parts, namely, the database, the spectrum simulator, and the fitting procedure. The database contains the chemical and substituent shifts for the ^{13}C resonances of the monosaccharides obtained by comparison of the chemical shifts for the resonances of a disaccharide with those of the corresponding monomeric sugars. Input data are the identities of the constituent sugars in the repeating unit of the polysaccharide and the positions of the linkages. This information is obtained from sugar and methylation analyses and by determination of the absolute configuration. Using these data, all possible permutations of repeating units can be generated and their spectra simulated. The fit with the experimental spectrum is then calculated and the simulated spectra are ranked with respect to this fit. It is further possible to visualise the structure of the polysaccharide by the output of a datafile readable by the molecular modelling program CHEM-X*. The molecular model may be obtained either as a rough estimation of the conformation or as an energy-minimised structure.

Databases. — In order to obtain reliable data, the spectra should be recorded under uniform conditions. All the data used here were recorded for solutions in D_2O at 70°, using 1,4-dioxane (δ 67.40) as the internal reference. The chemical shifts of the ^{13}C resonances of the α- and β-pyranosidic forms of D-glucose, D-

*CHEM-X, developed and distributed by Chemical Design, Oxford, Great Britain.

TABLE I

^{13}C-N.M.R. DATA[a] FOR GLYCOPYRANOSES

Sugar	*C-1*	*C-2*	*C-3*	*C-4*	*C-5*	*C-6*
α-Gal	93.18	69.35	70.13	70.28	71.30	62.04
β-Gal	97.37	72.96	73.78	69.69	75.93	61.84
α-Glc	92.99	72.47	73.78	70.71	72.37	61.84
β-Glc	96.84	75.20	76.76	70.71	76.76	61.84
α-Man	94.94	71.69	71.25	67.94	73.34	61.99
β-Man	94.55	72.13	74.03	67.69	77.00	61.99

[a]Obtained at 70° for solutions in deuterium oxide (internal 1,4-dioxane; δ 67.40).

TABLE II

GLYCOSYLATION SHIFTS (Δδ VALUES) USED IN THE CALCULATION OF SIMULATED SPECTRA

Disaccharide element[a]	*C-1*	*C-2*	*C-3*	*C-4*	*C-5*	*C-6*
α-D-Glc-(1→3)-α/β-D-Gal[b]	3.05	−0.27	−0.06	−0.13	0.37	−0.29
	0.01	−1.63	5.14	−3.59	−0.23	−0.03
β-D-Glc-(1→3)-α/β-D-Gal	7.59	−0.86	−0.14	−0.20	−0.04	−0.20
	−0.10	−0.96	9.86	−0.33	−0.30	−0.08
α-D-Glc-(1→3)-α/β-D-Man	8.53	0.35	0.24	−0.07	0.71	−0.15
	−0.16	0.02	7.43	−0.80	0.14	−0.08
β-D-Glc-(1→3)-α/β-D-Man	4.36	−1.28	−0.09	−0.16	0.26	−0.17
	−0.23	−2.40	7.66	−1.50	−0.28	−0.10
α-D-Gal-(1→3)-α-D-Glc	6.96	0.19	0.13	0.15	0.44	−0.36
	0.11	−1.40	7.41	0.15	−0.24	−0.12
α-D-Gal-(1→3)-β-D-Glc	6.95	0.13	0.13	−0.21	0.46	−0.19
	0.02	−1.28	7.15	0.18	−0.32	−0.37
β-D-Gal-(1→3)-α-D-Glc	6.83	−0.78	−0.16	−0.13	0.13	−0.20
	−0.15	−0.62	9.64	−1.55	−0.17	0.04
β-D-Gal-(1→3)-β-D-Glc	7.04	−0.80	−0.19	−0.14	0.16	−0.12
	−0.27	−0.48	9.17	−1.46	−0.34	0.02
α-D-Man-(1→3)-α-D-Glc	6.96	−0.31	0.13	0.15	0.44	−0.36
	0.11	−1.40	7.41	0.15	−0.24	−0.12
α-D-Man-(1→3)-β-D-Glc	6.96	−0.51	0.18	−0.17	0.52	−0.21
	0.12	−1.31	6.84	0.29	−0.15	−0.21
α-D-Man-(1→3)-α/β-D-Gal	3.05	−0.77	−0.06	−0.13	0.37	−0.29
	0.01	−1.63	5.14	−3.59	−0.23	−0.03
α-D-Man-(1→3)-α/β-D-Man	8.53	−0.35	0.24	−0.07	0.71	−0.15
	−0.16	0.02	7.43	−0.80	0.14	−0.08

[a]For the remaining disaccharide elements, approximations were performed as follows. The values are not listed since they are found above.

α-D-Gal-(1→3)-α/β-D-Gal	from	α-D-Glc-(1→3)-α/β-D-Gal
β-D-Gal-(1→3)-α/β-D-Gal	from	β-D-Glc-(1→3)-α/β-D-Gal
α-D-Gal-(1→3)-α/β-D-Man	from	α-D-Glc-(1→3)-α/β-D-Man
β-D-Gal-(1→3)-α/β-D-Man	from	β-D-Glc-(1→3)-α/β-D-Man
β-D-Man-(1→3)-α-D-Glc	from	β-D-Gal-(1→3)-α-D-Glc
β-D-Man-(1→3)-β-D-Glc	from	β-D-Gal-(1→3)-β-D-Glc
β-D-Man-(1→3)-α/β-D-Gal	from	β-D-Glc-(1→3)-α/β-D-Gal
β-D-Man-(1→3)-α/β-D-Man	from	β-D-Glc-(1→3)-α/β-D-Man

[b]The first rows refer to the glycosyl group and the second to the "aglycon".

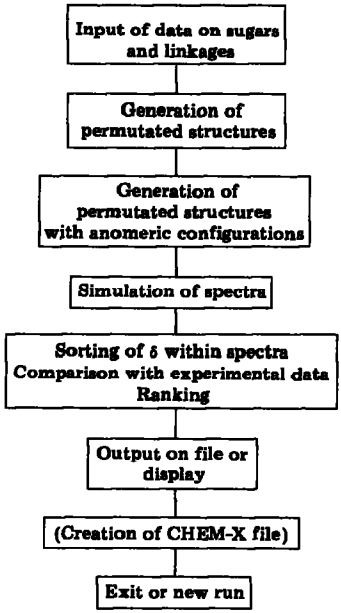

Fig. 1. Flow diagram of the data handling routines in CASPER.

galactose, and D-mannose are included in the database and are given in Table I. In addition, the substituent shifts for all ^{13}C resonances from all combinations of these monosaccharides in a disaccharide are compiled in a $\Delta\delta$-file. Ail data used in these calculations are presented in Table II and a full account of them will be published elsewhere. Useful data were obtained from refs. 2, 10, and 11. CASPER is at present also capable of handling L-fucose, L-rhamnose, and 2-acetamido-2-deoxy-D-glucose residues. ^{13}C-N.m.r. data for all disaccharides were not available and approximations of some substituent shifts had to be made. A check number is therefore added to each set of values in order to indicate the quality of the data. This is used later in the output for evaluation of the suggested structures. Thus, $\Delta\delta$ values coming from real ^{13}C-n.m.r. data have a check number of 0.01. If a disaccharide is only slightly different from a compound with existing data, its $\Delta\delta$ values are assigned a check number of 0.1, for example, when a glucopyranosyl residue is substituted for a galactopyranosyl residue.

When coarser approximations are made, the check number assigned is 1. The $\Delta\delta$ values for disaccharides for which no proper approximation can be performed have a check number of 10, and the substituent shifts of the resonances of the α- and the β-carbons are given values of 7 and -1 p.p.m., respectively.

Data handling routines. — The flow diagram for CASPER, shown in Fig. 1, contains all the essential steps. The requested inputs are, firstly, the identities of the glycopyranosyl residues and their absolute configurations. These residues are then permutated by the program to all possible structures of the repeating unit,

TABLE III

CALCULATION OF THE CHEMICAL SHIFTS[a] OF THE ^{13}C RESONANCES FOR ONE OF THE SIMULATED *Klebsiella* K8S STRUCTURES

	→3)-β-D-Glcp-(1→						→3)-β-D-Galp-(1→						→3)-α-D-Galp-(1→					
	C-1	*C-2*	*C-3*	*C-4*	*C-5*	*C-6*	*C-1*	*C-2*	*C-3*	*C-4*	*C-5*	*C-6*	*C-1*	*C-2*	*C-3*	*C-4*	*C-5*	*C-6*
δ, for monomer	96.8	75.2	76.8	70.7	76.8	61.9	97.4	73.0	73.8	69.7	75.9	61.8	93.2	69.4	70.1	70.3	71.3	62.0
δΔ for 1st disaccharide	7.6	−0.9	−0.1	−0.2	0.0	−0.2	−0.1	−1.0	9.9	−0.3	−0.3	−0.1						
Modified δ	104.4	74.3	76.7	70.5	76.8	61.7	97.3	72.0	83.7	69.4	75.6	61.7						
δΔ for 2nd disaccharide							7.6	−0.9	−0.1	−0.2	0.0	−0.2	−0.1	−1.0	9.9	−0.3	−0.3	−0.1
Modified δ							104.9	71.1	83.6	69.2	75.6	61.5	93.1	68.4	80.0	70.0	71.0	61.9
δΔ for 3rd disaccharide	0.0	−1.3	7.2	0.2	−0.3	−0.4							7.0	0.1	0.1	−0.2	0.5	−0.2
Simulation	104.4	73.0	83.9	70.7	76.5	61.3	104.9	71.1	83.6	69.2	75.6	61.5	100.1	68.5	80.1	69.8	71.5	61.7

[a]Rounded up or down to one place of decimals for simplicity.

and, for each structure, all possible combinations of anomeric configurations are also generated.

The simulated ^{13}C-n.m.r. spectrum for each combination, *e.g.*, A–B–C, is then calculated according to the following procedure, exemplified for the Smith-degraded *Klebsiella* K8 capsular polysaccharide[12] (Table III). The disaccharide (A–B) formed by the first two monosaccharide residues is identified. The substituent shifts for all of the ^{13}C resonances of A–B are then located in the $\Delta\delta$-file and added to the corresponding chemical shifts of A and B. The procedure is then repeated for B–C and C–A, and the simulated spectrum has then been calculated. A signal-by-signal comparison between the experimental spectrum and the sorted simulated spectrum then yields a number of chemical shift differences which may be positive or negative. The sum of the absolute values are calculated and this δ sum (delta-sum in the sample runs) represents the fit between the calculated and the experimental spectrum.

Calculation of all simulated spectra and comparison with experimental data yields, for each spectrum, the $\Delta\delta$ sum, which is low if the fit is good and high if it is poor. For structures with two or more identical sugar residues, identical spectra are produced and the duplicates may be excluded in the output. A further way of monitoring the fit is by summation of the squares of the chemical shift differences and then taking the square root of that value (SSS). This procedure gives a low ranking for spectra having some large chemical shift differences compared to spectra having evenly distributed chemical shift differences.

For each calculated spectrum, the check number sum is also obtained. A value of 0.30 for the above calculation is good since it indicates that $\Delta\delta$ values from disaccharides with similar stereochemistry have been used. A value such as 11.1 indicates that less accurate values have been used and that uncertainty is associated with the suggested structure.

Determination of the structures of polysaccharides by computer evaluation. — The complete calculation of the structure of two polysaccharides will be discussed. The backbones of the capsular polysaccharides from *Klebsiella* K35 (ref. 13) and K8 (ref. 12) were obtained after removal of pyruvic acid (K35S) and removal of terminal D-glucuronic acid side-chains by Smith degradation (K35S, K8S), and they have the following structures:

K35S →3)-α-D-Gal*p*-(1→3)-α-D-Man*p*-(1→3)-α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→

K8S →3)-β-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→3)-α-D-Gal*p*-(1→

Calculation of the K35S structure. — The test run for K35S is given in Scheme 1. After the input of the relevant data, permutations of the structure are calculated and the spectra are simulated. The number of ranked simulated structures displayed may be chosen and, generally, it is sufficient to inspect the best 5–10 structures. For K35S, the highest ranked structure is the correct one. The calculated and

```
Write spectrumfile                  :KL-K35S.DAT
Number of sugars                    :4
Poly- or Oligo-saccharide? P/O      :P
Sugar residue 1                     :DGALP
Linkage position.
Enter number                        :3
Sugar residue 2                     :DMANP
Linkage position.
Enter number                        :3
Sugar residue 3                     :DMANP
Linkage position.
Enter number                        :3
Sugar residue 4                     :DGLCP
Linkage position.
Enter number                        :3
You have entered as

sugar residue 1: 3 DGALP
sugar residue 2: 3 DMANP
sugar residue 3: 3 DMANP
sugar residue 4: 3 DGLCP

Change sugar or linkage position?   :N
Write on file or display? F/D       :D
Delete identical structures? Y/N    :Y
First simulated structure.
Enter number                        :1
Last simulated structure.
Enter number                        :8

        KL-K35S
No.     Polysaccharide.
2       3ADGALP     -3ADMANP    -3ADMANP    -3BDGLCP
4       3BDGALP     -3ADMANP    -3ADMANP    -3BDGLCP
6       3BDGALP     -3ADGLCP    -3ADMANP    -3ADMANP
8       3ADGALP     -3ADGLCP    -3ADMANP    -3BDMANP

Information about simulated structures.
First simulated structure.
Enter number                        :1
Last simulated structure.
Enter number                        :8

No.     Delta-sum       SSS         Check#
2        5.3            1.4         0.31
4       12.2            4.4         0.40
6       13.5            4.5         0.40
8       13.7            3.8         0.40

First simulated spectrum.
Enter number                        :2
Last simulated spectrum.
Enter number                        :2

Experimental spectrum.
104.4   102.8   101.5   101.3   83.0    79.9    79.5    79.1    76.5    74.3
 73.9    73.1    72.0    70.8   70.8    70.7    70.0    68.7    67.1    66.9
 62.1    62.0    61.7    61.5
```

Scheme 1. Sample run of K35S.

```
Spectrum number 2.
104.6   103.3   101.7   101.6   83.5   80.2   78.9   78.9   76.6   74.2
 74.0    73.0    71.7    71.4   71.2   70.8   69.9   68.7   67.1   67.0
 61.8    61.8    61.7    61.4

First non-sorted spectrum.
Enter number                    :2
Last non-sorted spectrum.
Enter number                    :2

Experimental spectrum.
104.4   102.8   101.5   101.3   83.0   79.9   79.5   79.1   76.5   74.3
 73.9    73.1    72.0    70.8   70.8   70.7   70.0   68.7   67.1   66.9
 62.1    62.0    61.7    61.5

Non-sorted spectrum number 2.
101.6   68.7   80.2   69.9   71.7   61.8
103.3   71.4   78.9   67.1   74.2   61.8
101.7   71.2   78.9   67.0   74.0   61.7
104.6   73.0   83.5   70.8   76.6   61.4

Create CHEM-X file? Y/N         :N
Finished with this calculation? :Y
Make another calculation? Y/N   :N
```

Scheme 1. Sample run of K35S (continued).

the experimental spectrum differ by 5.3 p.p.m., calculated as the $\Delta\delta$ sum, *i.e.*, an average of 0.22 p.p.m. for each of the 24 signals.

The second ranked structure has the same sequence of sugar residues as the first structure, but contains a β-D-galactopyranosyl instead of an α-D-galactopyranosyl residue. The $\Delta\delta$ sum is 12.2 p.p.m., *i.e.*, approximately twice the best value. The next proposed structure has a different sequence of the sugar residues, as well as an α-D-glucopyranosyl residue, compared to the second suggestion. In the fourth structure, one of the mannose residues is β and the remaining sugars are α.

The next step is a comparison between the experimental and simulated spectra. For the first suggested structure, there are four signals for which the chemical shift differences are $\geq$0.5 p.p.m. This value is arbitrarily chosen and used only to point out large deviations. The signals are found in the experimental spectrum at δ 102.8, 83.0, 79.5, and 70.8. The first represents an anomeric carbon, and next are two glycosylated carbons. As the glycosylation shifts are largest by far for the resonances of anomeric and glycosylated carbons, the deviations can also be expected to be correspondingly large.

The simulated non-sorted spectrum is also useful. The chemical shifts of the ^{13}C resonances of the four sugar residues then appear as four sets for the monomers presented in the same order as they appear in the printed structure. The signal in the simulated spectrum at, for example, δ 83.5 can be assigned immediately to the linkage carbon in the fourth residue. This makes a more detailed analysis possible.

Because no signals are assigned, there may be a larger real $\Delta\delta$ sum as comparison is performed by superposition only.

A closer look at the check number sum for the first structure indicates that the substituent shift for the resonances of one of the four disaccharide elements was present in the database as such, and that the other three were taken from similar compounds. Thus, the following approximations were introduced.

Present disaccharide	*Used disaccharide and corrections*	
α-D-Gal*p*-(1→3)-α-D-Man*p*-(1→	α-D-Fuc*p*-(1→3)-α-D-Man*p*-(1→	
α-D-Man*p*-(1→3)-α-D-Man*p*-(1→	α-D-Fuc*p*-(1→3)-α-D-Man*p*-(1→	$\delta_{C\text{-}3}$ −0.6, $\delta_{C\text{-}2'}$ −0.7
α-D-Man*p*-(1→3)-β-D-Glc*p*-(1→	α-D-Rha*p*-(1→3)-α-D-Glc*p*-(1→	
β-D-Glc*p*-(1→3)-α-D-Gal*p*-(1→	β-D-Glc*p*-(1→3)-α-D-Gal*p*-(1→	

Each of the other three proposed structures has substituent shifts with a check number of 0.1. Consequently, no alternative to the first structure would be expected, even if better model substances had been available.

```
        KL-K8S
No.     Polysaccharide.
2       3BDGLCP     -3BDGALP    -3ADGALP
4       3ADGLCP     -3BDGALP    -3BDGALP
6       3BDGLCP     -3ADGALP    -3BDGALP
8       3BDGLCP     -3BDGALP    -3BDGALP

No.     Delta-sum   SSS         Check≠
2        2.5        0.9         0.30
4       11.7        5.1         0.30
6       18.1        6.3         0.21
8       18.4        7.1         0.30

Experimental spectrum.
104.7  104.3  99.9  83.4  82.9  79.9  76.4  75.5  73.0  71.3
 71.2   70.6  69.9  69.1  68.6  61.7  61.7  61.4

Spectrum number 2.
104.9  104.4  100.0  83.8  83.5  80.1  76.4  75.6  73.1  71.5
 71.1   70.7   69.7  69.2  68.5  61.8  61.6  61.3

Spectrum number 4.
105.0  104.1  95.9  83.5  83.4  78.8  75.8  75.7  72.6  71.6
 71.2   70.5  69.2  69.0  65.9  61.6  61.6  61.6

Non-sorted spectrum number 2.
104.4  73.1  83.8  70.7  76.4  61.3
104.9  71.1  83.5  69.2  75.6  61.6
100.0  68.5  80.1  69.7  71.5  61.8

Non-sorted spectrum number 4.
 95.9  71.6  83.4  69.0  72.6  61.6
105.0  70.5  78.8  65.9  75.7  61.6
104.1  71.2  83.5  69.2  75.8  61.6
```

Scheme 2. Sample run of K8S.

Calculation of the K8S structure. — The four highest ranked structures for K8S, the $\Delta\delta$-sums, SSS-values, and the check number sums together with the two best simulated spectra, both as the sorted and the non-sorted spectra, are given in Scheme 2.

The check number sum for each of the first four proposed structures is <1.0, and thus the simulated spectra are based upon good model substances. The first two proposed structures contain two β linkages and one α linkage, whereas the third contains only β linkages. The $\Delta\delta$ sum for the first structure is 2.5, corresponding to an average value of 0.14 p.p.m. per signal. For the second best proposal, the $\Delta\delta$ sum, 11.7 p.p.m., is more than four times higher, and the value for the third best proposal is more than seven times higher.

In the sorted spectrum for the first simulated structure, no differences in chemical shifts relative to the experimental values are >0.6 p.p.m. In the second best structure, the exchange of an α and a β linkage gives the disaccharide element →3)-α-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→. The calculated spectrum for this element contains signals for C-1′, C-3, and C-4 at 95.9, 78.8, and 65.9 p.p.m., respectively. No signals with chemical shifts close to these values are observed; thus, three large chemical shift differences responsible for the large $\Delta\delta$ sum are obtained. The following disaccharides were used as references for the calculation of the spectrum of the highest ranked structure.

Present disaccharide	*Used disaccharide*
β-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→	β-D-Glc*p*-(1→3)-α-D-Gal*p*-(1→
β-D-Gal*p*-(1→3)-α-D-Gal*p*-(1→	β-D-Glc*p*-(1→3)-α-D-Gal*p*-(1→
α-D-Gal*p*-(1→3)-β-D-Glc*p*-(1→	α-D-Glc*p*-(1→3)-β-D-Glc*p*-(1→

Conclusions and future aspects. — The potential of CASPER to predict polysaccharide structures has been demonstrated for two polysaccharides with known structures. The strong point in CASPER is the unbiased approach to the structure and that the chemical shifts of all of the ^{13}C-n.m.r. resonances are used. A complete analysis, including sugar and methylation analyses, determination of absolute configuration of the sugars, and calculation with CASPER, can be performed in less than 2 days. The structures proposed by CASPER cannot be regarded as conclusively proved. When there is a choice between two or more highly ranked structures, the correct one may be determined by specific degradations or by establishing connectivity of protons over the glycosidic bond through n.O.e. measurements, thereby obtaining sequence information. CASPER cannot be better than its database, which is presently being expanded on the basis of the synthesis and analysis of new disaccharides and by the addition of data from the literature.

The extension of CASPER to branched structures, which are more common than the unbranched ones, presents problems since studies of branched trisaccharides[11] show that, when the two non-reducing sugar residues can interact, dramatic changes in the ^{13}C-n.m.r. spectra occur. The prediction of spectra of branched structures will require a database of high quality.

The program can be extended to include ^{1}H-n.m.r. data. There are at least three groups of $J_{1,2}$ values, namely, when H-1 and H-2 are diaxial (~7–8 Hz; *e.g.*, β-glucosides), equatorial–axial (~4 Hz; *e.g.*, α-glucosides), and diequatorial or axial–equatorial (<2 Hz; *e.g.*, α- or β-mannosides). If the *J* values are known, the number of possible structures will be much diminished. Thus, for K35S, the fourth best alternative containing no β-glucose but a β-mannose residue could have been dealt with easily as only one $J_{1,2}$ value of 8 Hz was present in the ^{1}H-n.m.r. spectrum. As calculations of repeating units with 5–6 sugar residues will produce a large number of simulated structures, such discriminators are therefore important.

Other reporter groups may also be of interest for the proton part of the evaluation, *e.g.*, the chemical shift of the H-2 resonances which may be determined generally from a 2D COSY spectrum.

The ultimate aim is to use ^{1}H- and ^{13}C-n.m.r. data to determine the structures of polysaccharides without using chemical degradations or time-consuming 2D-n.m.r. experiments.

EXPERIMENTAL

Chemical shifts of the ^{13}C resonances for solutions in D_2O of mono-, di-, and poly-saccharides are expressed relative to that of internal 1,4-dioxane (δ 67.40). CASPER is written in Vax11-Pascal, and the program was run on a VAX 11/750 computer. For the three-dimensional models, CASPER produces a file consisting of a number of commands readable by the molecular modelling program CHEM-X. The oligosaccharides are made up from crystal structure data, and the conformations are allowed to take either standard values for dihedral angles or are optimised by a minimiser present in CHEM-X.

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Natural Science Council and the National Swedish Board for Technical Development. The *Klebsiella* polysaccharides were kindly provided by Professor G. G. S. Dutton (University of British Columbia).

REFERENCES

1 K. Bock and C. Pedersen, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 27–66.
2 K. Bock, C. Pedersen, and H. Pedersen, *Adv. Carbohydr. Chem. Biochem.*, 42 (1984) 193–225.
3 N. K. Kochetkov, O. S. Chizhov, and A. S. Shaskov, *Carbohydr. Res.*, 133 (1984) 173–185.
4 J. F. G. Vliegenthart, L. Dorland, and H. Van Halbeek, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 209–274.
5 N. A. B. Gray, *Prog. Nucl. Magn. Reson. Spectrosc.*, 15 (1982) 201–248.
6 H. L. Suprenant and C. N. Reilly, *ACS Symp. Ser.*, 54 (1977) 77–91.
7 H. Egli, D. H. Smith, and C. Djerassi, *Helv. Chim. Acta*, 65 (1982) 1898–1919.
8 E. Hounsell, D. J. Wright, A. S. R. Donald, and J. Feeney, *Biochem. J.*, 223 (1984) 129–143.
9 D. R. Anderson and W. J. Grimes, *Anal. Biochem.*, 146 (1985) 13–22.
10 M. Forsgren, P.-E. Jansson, and L. Kenne, *J. Chem. Soc., Perkin Trans. 1*, (1985) 2383–2388.
11 J. H. Bradbury and G. A. Jenkins, *Carbohydr. Res.*, 126 (1984) 125–156.
12 I. W. Sutherland, *Biochemistry*, 9 (1970) 2180–2185.
13 G. G. S. Dutton and A. V. S. Lim, *Carbohydr. Res.*, 145 (1985) 67–80.