ELSEVIER

PII: S0144-8617(96)00161-0

Carbohydrate Polymers 32 (1997) 255–266

0144-8617/97/$17.00 + .00

Conformational statistics of pectin substances in solution by a Metropolis Monte Carlo study

Bernard Boutherin[a], **Karim Mazeau**[a] & **Igor Tvaroska**[a,b]

[a]*Centre de Recherches sur les Macromolécules Végétales (CERMAV), CNRS, BP 53, 38041 Grenoble cedex 9, France*
[b]*Institute of Chemistry, Slovak Academy of Sciences, 84238 Bratislava, Slovak Republic*

(Received 20 November 1995; revised version received 12 September 1996; accepted 15 September 1996)

Using a Metropolis Monte Carlo algorithm, various average properties of several pectic polysaccharide models were calculated based on the conformational energies for parent disaccharides. The relaxed potential energy surfaces of disaccharides were calculated using the CHARMm force field. Solvent effects were evaluated by calculating a solvation energy for each conformational state by estimating contributions from a cavity formation, and from the electrostatic and dispersion interactions between solvent and solute molecules. The behavior of the mean characteristic ratio, the squared radius of gyration, and the persistence length versus chain length is discussed for various structural models, temperature, and solvent. It is found that the unrefined model of the alternating co-polymer polygalacto-galacturonic acid in vacuum is consistent with the experimentally measured dimension of pectin in salt excess. This model is used to generate computer images of the characteristic conformation of pectin chain. © 1997 Elsevier Science Ltd

INTRODUCTION

Polysaccharides are biological macromolecules currently attracting an enormous amount of interest. Investigation that can provide information on the three-dimensional structure of these molecules in solution is very important for understanding their physicochemical and biological properties and also their industrial applications. One of the most characteristic aspects of the physical chemistry of polysaccharides is the ability of the long polysaccharide chains to assume an enormous variety of spatial arrangements around the glycosidic linkages. This means that observable parameters describing the solution behavior of polysaccharides are averages of the properties of the individual conformations.

Theoretical polysaccharide models are based on studies of the relative abundances of the various conformations of a given polysaccharide together with the statistical mechanical theory of polymer chain configuration (Flory, 1969; Brant, 1976, 1980; Brant & Burton, 1981). Polysaccharide models, refined to various extent have been presented for different polysaccharides and the calculated properties have been compared with experimental results (Buliga & Brant, 1987; Brant, 1990; Ruggiero *et al.*, 1995). Calculations generally refer to the unperturbed state, which is the condition where the long-range repulsions are exactly compensated by long-range attractions. This corresponds to the experimental θ-conditions. A representation of potential energy surfaces is usually performed by calculating two-dimensional maps of the internal energy of an appropriately chosen skeletal segment of polysaccharide chain with respect to two glycosidic torsional angles, all the other internal parameters being kept fixed. This approach is usually based on the following assumptions: (a) that the distribution of internal energies for the various conformations of the chosen dimeric segment is not too different from that of the same segment in the polysaccharide chain; (b) that the torsional angles of any given glycosidic linkage, while strongly correlated with one another, are uncorrelated with the torsional angles of any other glycosidic linkage; and (c) that the presence of the solvent is neglected. Recently, it was suggested that the change in macroscopic properties of polysaccharides observed in different solvent are at least partially originated from variation in the short-range polymer–solvent interactions, which influence the abundances of the conformational isomers (Urbani & Cesàro, 1991).

In this paper we attempted to study the conformational distribution of different unperturbed pectin models. For this purpose a Metropolis algorithm of the

Monte Carlo technique is applied to generate a large sample of polymer conformations, distributed according to Boltzmann statistics at a given temperature. Conformational energy surfaces generated for this purpose are calculated as fully relaxed maps of proper oligomers and take into account solvent–saccharide interactions.

The basic backbone of pectin polysaccharide is formed by (1→4)-linked α-D-galacturopyranosyl residues, either free or in ester form, these homogalacturonan sequences may be interspersed at intervals with β-L-rhamnopyranosyl residues carrying the major part of neutral sugar side-chains, mainly arabinans, galactans or arabinogalactans (Walker, 1990). In the physicochemical analysis of pectins, the characterization of both their size and shape were studied by many techniques, such as osmometry measurement (Owens *et al.*, 1946; Fishman *et al.*, 1986), wide and low angle laser light scattering coupled with or with not size exclusion chromatography (Jordan & Brant, 1978; Plashchina *et al.*, 1985; Chapman *et al.*, 1987; Fishman *et al.*, 1991; Kravtchenko *et al.*, 1992a, b; Ousalem *et al.*, 1993; Malovikova *et al.*, 1993), and viscometry (Axelos *et al.*, 1987, 1991; Hourdet & Muller, 1987, 1991a, b; Berth, 1992), and low speed sedimentation equilibrium (Harding *et al.*, 1991). However, the heterogeneity of the primary structure and the presence of aggregates could affect the molecular state in solution and therefore disturb the accurate determination of molecular weight, thus explaining the divergent reported literature data on the solution characteristics of pectins. Depending on the authors and the measuring techniques used, pectins molecules have been reported to behave as rigid-rod particles (Owens *et al.*, 1946) or as coils of variable stiffness (Anger & Berth, 1985, 1986; Deckers *et al.*, 1986; Axelos *et al.*, 1987). Some authors have reported that the stiffness of chains could depend on the degree of esterification (Deckers *et al.*, 1986). The neutral sugar content has been reported to affect the conformation and it has been suggested that fractions rich in neutral sugars are responsible for the high molecular weight found in some cases (Berth, 1988). The present study is an attempt in development of the theoretical polysaccharide chain model of linear part of pectin molecules.

SYSTEMS AND METHOD OF CALCULATIONS

Conformational energy estimates

The calculations reported in this paper have been performed on various (1→4)-linked disaccharides of galactose derivatives which represent different segments occurring in the pectin chain. In order to evaluate the influence of the chemical structure of the group on the C5 atoms on polysaccharide models the following dimers were chosen:

O-α-D-Galactopyranosyl-(1→4)-α-D-galactopyranose (galabiose, **1**),
O-(α-D-Galactopyranosyluronic acid)-(1→4)-α-D-galactopyranose (**2**),
O-α-D-Galactopyranosyl-(1→4)-α-D-galactopyranuronic acid (**3**),
O-(α-D-Galactopyranosyluronic acid)-(1→4)-α-D-galactopyranuronic acid (digalactouronic acid, **4**),
O-(α-D-Galactopyranosyluronate)-(1→4)-α-D-galactopyranuronate (**5**), and
O-(α-D-Sodium galactopyranosyluronate)-(1→4)-α-D-sodium galactopyranuronate (**6**).

The six studied dimers differ from each other by the nature of their pendant groups: hydroxymethyl, carboxyl, carboxylate or sodium carboxylate. The recommendations and symbols proposed by the Commission on Nomenclature IUPAC-IUB (1971) are used throughout this paper. Representations of the six disaccharide compounds along with the labeling of the atoms are given schematically in Fig. 1. The notation of individual conformers is based on the values of torsional angles Φ (O5′—C1′—O4—C4) and Ψ (C1′—O4—C4—C5) that describe the relative orientation of galactopyranosyl residues. The orientation of pendant groups is defined by additional torsion angles ω (O5—C5—C6—O6) and ω' (O5′—C5′—C6′—O6′). The initial internal geometry of **1** is based on its X-ray crystal structure (Svensson *et al.*, 1986). The coordinates of the hydrogen atoms were determined using a C—H bond

	R1	R2
(1)	CH_2OH	CH_2OH
(2)	COOH	CH_2OH
(3)	CH_2OH	COOH
(4)	COOH	COOH
(5)	COO^-	COO^-
(6)	COO^-Na^+	COO^-Na^+

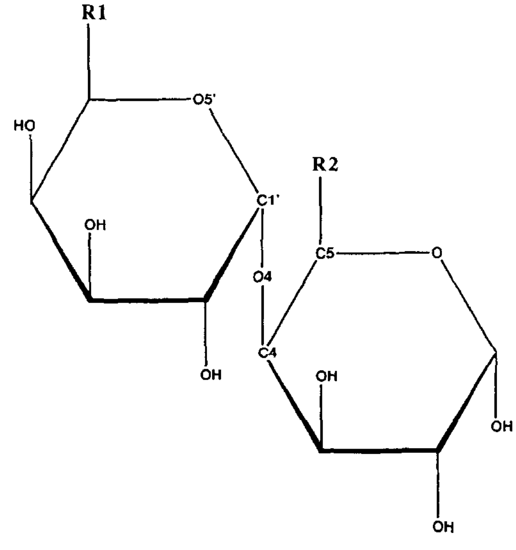

Fig. 1. Schematic representation of the disacharides **1–6**.

length of 110 pm and an O—H bond length of 99 pm, with bond vector related appropriately to the C—C and C—O bond vectors. The geometry for **2**–**6** were the same as for **1** except for the replacement of hydroxymethyl groups by appropriate carboxyl, carboxylate or sodium carboxylate groups whose geometries were taken from crystallographic coordinates of acid compounds (Thanomkul *et al.*, 1976; Newton *et al.*, 1977; DeLucas *et al.*, 1978).

The method of calculation of the energies is identical to that described in our previous papers (Gouvion *et al.*, 1994, 1995) which were focused on the conformational analysis of the above disaccharides using NMR in conjunction with molecular mechanics. Relaxed (Φ,Ψ) maps were prepared by rotating the rigid residues of **1**–**6**, to each combination of Φ and Ψ on a 10° grid. At each of the Φ_i, Ψ_j grid point the energy was minimized allowing the cartesian coordinates of each atom to vary. Geometry optimization was carried out with 500 steps of a conjugate gradient procedure in cartesian coordinates. At each step, potential energy and energy gradients were calculated. Minimization was carried out until a r.m.s. value of less than $0.1\,\mathrm{kcal/mol\,\AA^2}$ of the calculated gradient is obtained. Energy calculations were performed with the CHARMm program (Brooks *et al.*, 1983) using an original force field as interfaced in QUANTA (Molecular Simulations Inc).

In calculating the energy, the solvation energy contribution is evaluated for each conformation state Φ_i, Ψ_j using a method based on the scaled particle theory (Reiss *et al.*, 1959). This theory can give a rigorous description for liquid system consisting of hard spheres. The applicability of the method to a wider class of systems such as aqueous solution has already been demonstrated for the solubility of small non-polar molecules in water (Pierotti, 1976). The application of the modified version of this method to the solvation of saccharide conformers shown remarkable agreement with the experimentally observed effect of solvent on the axial–equatorial equilibrium of 2-methoxytetrahydropyran (Tvaroska & Kozar, 1980) and glucose (Tvaroska & Kozar, 1986). The method was also successfully used to predict the effect of solvent on the conformational properties of various oligosaccharides (Tvaroska, 1989) and recently to describe the effect of solvent on a dimension of amylose and cellulose chains (Urbani & Cesàro, 1991). Also this method correctly described solution properties of two relevant oligosaccharides, namely digalacturonic acid and sodium digalacturonate (Gouvion *et al.*, 1994). In this procedure, the dissolution of the solute is decoupled into two steps by experiment: (1) creation of a cavity in solvent, which has right size to accommodate the solute conformer, (2) establishing the attractive interaction between the solvent and solute. The free energy of cavity formation can be estimated from the scaled particle theory formula. To calculate the contribution from the attractive interactions we assumed Lennard–Jones type dispersion interactions and electrostatic interactions between solvent and solute. The model and relevant equations have been described in detail in our previous papers (Tvaroska & Kozar, 1980; Tvaroska, 1982). The resulting energy values were used to calculate the partition function of the molecule being studied and Boltzmann factors needed in carrying out Monte Carlo generation of statistical chains as described in the following section.

Metropolis Monte Carlo generation of chain

A few studies utilizing Monte Carlo methods for the investigation of the configurational and conformational properties of polysaccharides have been published (Jordan *et al.*, 1978; Cesaro & Brant, 1981; Gagnaire *et al.*, 1982; Kitamura *et al.*, 1987; Urbani & Cesàro, 1991; Yui *et al.*, 1994). In all these investigations, a normalized probability has been associated with each Φ_i, Ψ_i state. Then polysaccharide chain of *n* residues have been created using a sequence of $n-1$ random numbers with the configuration of the *i*-th residue determined by the *i*-th random number. We present a Monte Carlo method in which the orientation of the *i*-th segment is determined using the Metropolis algorithm (Metropolis *et al.*, 1953). The algorithm comprises the following steps. In the first, the conformer in (Φ,Ψ) space is randomly selected. In the next step, the new orientation of Φ and Ψ dihedral angles is selected at random. Then the energy difference ΔE between the newly generated state and the previous state is calculated. If a new state has lower or equal energy than the old one, the state is always accepted, otherwise the Metropolis test is applied. The new conformation is then accepted or discarded according to the outcome of a statistical test in which the random number is generated in the interval 0–1 and compared to the Boltzmann factor $\exp(-\Delta E/RT)$, where R is the gas constant and T the temperature. If the random number is smaller than the Boltzmann factor, the new state is accepted. If it is greater, the old conformation is restored and counted again in the Monte Carlo sample. With respect to the generated values of Φ and Ψ the polysaccharide chain is built using geometry corresponding to these conformations of the two residues. Optimized or rigid geometries can be used in this step. Program METROPOL (METROpolis POLymer generator) is written in C.

Configurational statistic

Polymer chain dimensions are usually calculated using the rotational isomeric state model (Flory, 1969) which has been adopted for calculation of observable configuration-dependent polysaccharide properties (Brant, 1976, 1980). In the present case all ensemble-average chain properties are determined by Metropolis Monte Carlo (MMC) configurational sampling.

From generated chains, several configurational characteristics are evaluated. The size of the polysaccharide chain is characterized by the mean-square end-to-end distance $<r^2(x)>$

$$<r^2(x)>=<(\mathbf{r}_{x+1}-\mathbf{r}_1)^2> \tag{1}$$

and by the mean square radius of gyration $R_g{}^2$

$$R_g^2(x) = 1/x^2\Sigma < \mathbf{r}_{ij}^2 > \tag{2}$$

where x is the degree of polymerization, $\mathbf{r}_i$ is the position vector locating the i-th bead of the chain, $\mathbf{r}_{ij}$ is the vector from bead i to bead j. Angle brackets $<...>$ denote an ensemble average. In order to facilitate the comparison of the effect of the structure and the effect of a solvent on the intrinsic extension of flexible chain with differing degree of polymerization, it is useful to use the dimensionless quantity $C_x = <r^2>/xL^2$. Here L is the so called 'virtual' bond connecting a pair of adjacent glycosidic oxygens. In the flexible residue model, this distance is not constant but varied with conformation. However, a single geometry was used in constructing the polymer chains because of computational speed and memory problems. This simplification has in fact no consequences on the resulting chain dimensions because ring puckering fluctuations are compensated along the chain. Therefore, in such a case the average value of L over all residues is used. In the limit of infinite chain C_x becomes equal to the characteristic ratio

$$C_\infty = \lim C_x = \lim <r^2>/xL^2 \tag{3}$$

As an alternative way to describe a chain extension we have calculated the persistence length L_∞, which is originally defined as the average projection of the end-to-end distance vector r on the first bond of the chain, in the limit of infinitive chain length, for a chain of degree polymerization x. Directional correlations of the terminal residue are described by a correlation function which measures the mean projection of a unit vector along the x-th residue onto the direction of a residue vector aligned with the first residue

$$<\cos\gamma_{i1}>=<\mathbf{u}_i\mathbf{u}_1> \tag{4}$$

where $\mathbf{u}_i$ is the unit vector between the i and $(i+1)$th beads.

Typical calculations involve averaging for 1000 molecules containing 2000 residues each. Calculations have been performed at various temperatures, 27°, 227°, 427°, and 727°C. Calculations were performed on the IRIS Indigo 2 workstation at CERMAV, molecular drawings were produced by using Quanta software (Molecular Simulations Inc.). Isoenergy contour lines were generated using a home made program CARTO.

RESULTS AND DISCUSSION

In order to estimate the average properties of various models of pectic polymer chains it is necessary to know the conformational energy surfaces for corresponding dimeric units. Some fully relaxed conformational energy surfaces calculated for disaccharides **1**–**6** are shown in Fig. 2 for dimethyl sulfoxide and water solutions. The maps calculated for isolated molecule are not shown here as they were already presented in our previous paper (Gouvion *et al.*, 1995).

The noteworthy features of the relaxed conformational potential energy surfaces for various (1→4)-linked diaxial disaccharide systems are the presence of several low energy domains where all minima appear. These regions are centered at (Φ = 80°, Ψ = 100°), (150°, 150°), (70°, −100°), (140°, −30°), (150°, −60°), and (−60°, 140°), respectively. The occurrence and the relative energy of these areas noticeably depend on the chemical structure of substituents at C5 and C5′. The minimum energy conformation in all studied compounds corresponds to the one of the (80°, 100°) or (150°, 150°) regions. As transitions between these two regions generally cost less energy than departures to another area, conformations belonging to these domains are significant in the evaluation of equilibrium properties. A substitution of carboxyl group for hydroxymethyl group appreciably decreases the flexibility of molecules which is even enhanced by dissociation of this group. A comparison of the relaxed conformational energy surfaces in solution shows that the studied compounds display a quite different solution behavior. The reasonable agreement between the calculated and experimental data (Gouvion *et al.*, 1994) for digalacturonic acid and sodium digalacturonate suggests that the force field with the incorporation of the solvent effect used for the calculation of relaxed Φ,Ψ surfaces correctly describes the solution behavior of non-charged as well as charged compounds.

The sensitivity of average characteristics to the chemical structure, solvent and temperature has been investigated for four homopolymers and one copolymer: polygalactan (**11**), undissociated polygalacturonic acid (**44**), ionized polygalacturonic acid (**55**), sodium polygalacturonate (**66**), and a perfectly alternating copolymer, the undissociated polygalacto-galacturonic acid (**23**). The virtual bond length used for the calculations was 4.5 Å. The MMC simulations were carried out using corresponding conformational energy surfaces of disaccharides without any refinement of the force field with respect to experimental data. With respect to this, evaluated configurational properties refer to unrefined models of polysaccharide chains (Brant & Christ, 1990). Figure 3 shows C_x as a function of the degree of polymerization x. Table 1 summarizes the relevant configurational characteristics (equations 1–4) for all the pectin substances. Note that the C_∞ and L_∞ values

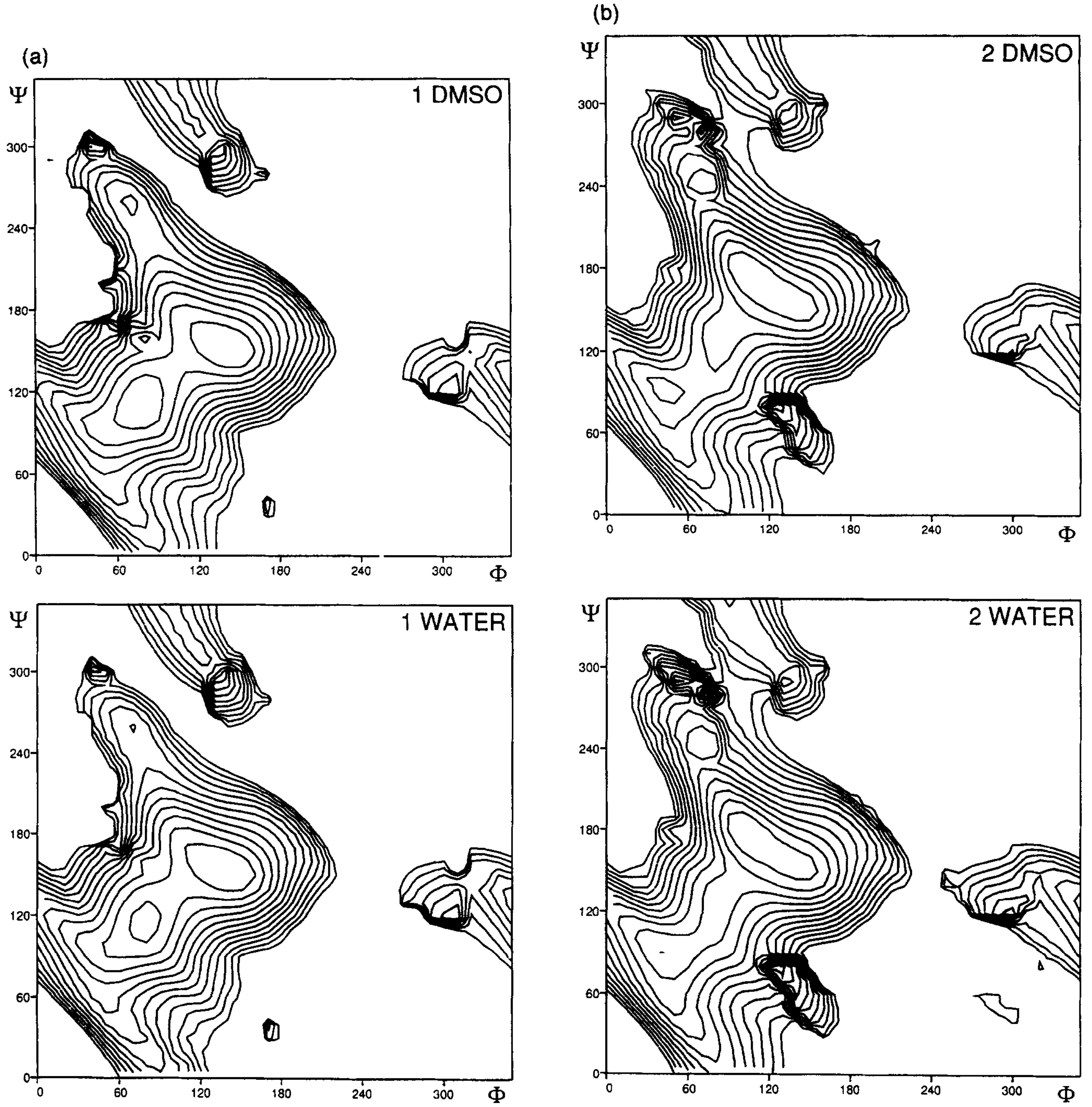

Fig. 2. Relaxed conformational energy surfaces computed for disaccharides 1–6 for dimethyl sulfoxide and aqueous solutions.

reported in the table are calculated from the C_x and L_x after fitting to a exponential equation in order to describe properly the asymptotic behavior of these functions. A comparison of calculated radii of gyration R_g with the experimental values is given in Fig. 4.

Influence of the chemical structure

The characteristic ratio serves as a key quantity in describing the conformational properties of polymers. The plots of C_x for non-ionized (1→4)-linked diaxial polysaccharides at 27°C in the isolated state, computed as a function of the degree of polymerization x for polymers **11**, **23** and **44** are shown in Fig. 3a. The extrapolated asymptotic values of the characteristic ratio C_∞ for infinitely long chains as well as the values of X_{95}, which is the degree of polymerization at which C_x reaches 95% of its value in the gaussian limit (Brant & Burton, 1981), together with the predicted values of the persistence length L_∞ are reported in Table 1. It can be seen that the calculated characteristic ratios noticeably depend on the character of the monomer pendant group. The values $C_\infty = 62$ and 47 that have been obtained for **11** and **44** are much higher in comparison

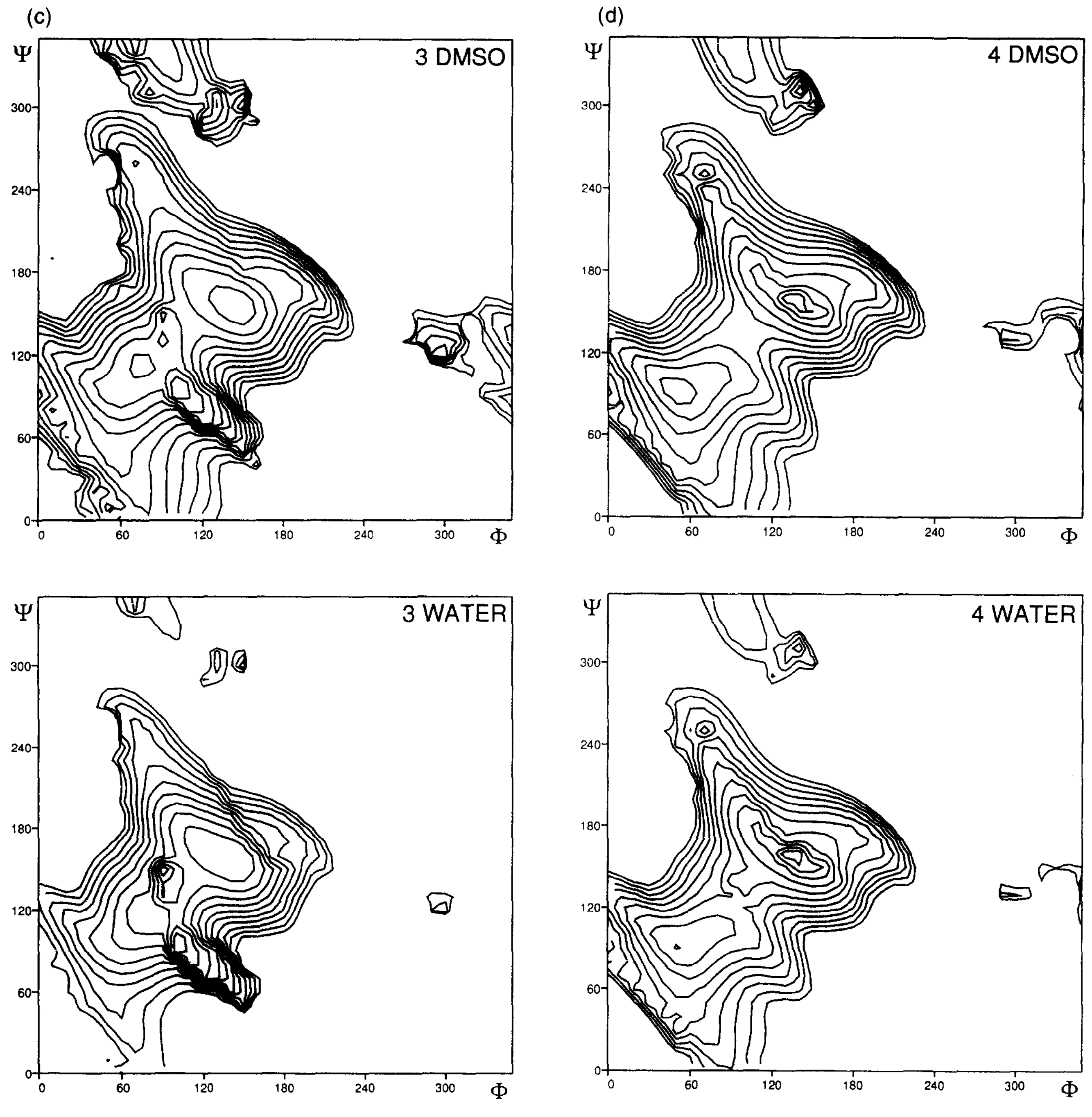

Fig. 2. *Continued.*

to $C_\infty = 24$ calculated for **23**. This suggests larger spatial extension and stiffness of the former polysaccharide chains. The calculated results also show that the C_x of **23** converges to the asymptotic value at lower degree of polymerization than that for homopolymers having either carboxylic (**44**) or hydroxymethyl (**11**) substituents. It is clear from the relaxed (Φ,Ψ) potential energy surfaces of the disaccharides that flexibility about the diaxial glycosidic bond is restricted. The large difference between the low value of C_∞ computed for the co-polymer and those obtained for homopolymers suggests a considerable sensitivity of calculated configurational properties to the relative energy of individual conformers or in other words to the shape of the (Φ,Ψ) map. Monte Carlo samples of homopolymers are generated using only one potential energy surface. The map of the parent disaccharide **1** for polygalactan is characterized by the presence of two minima with similar energy levels. The map for disaccharide **4** which corresponds to polygalacturonic acid also has one dominant conformer. On the contrary, a co-polymer is generated by using two different maps alternatively. For polymer **23** these maps correspond to the maps for the **2** and **3** disaccharides. The dominant conformer is at

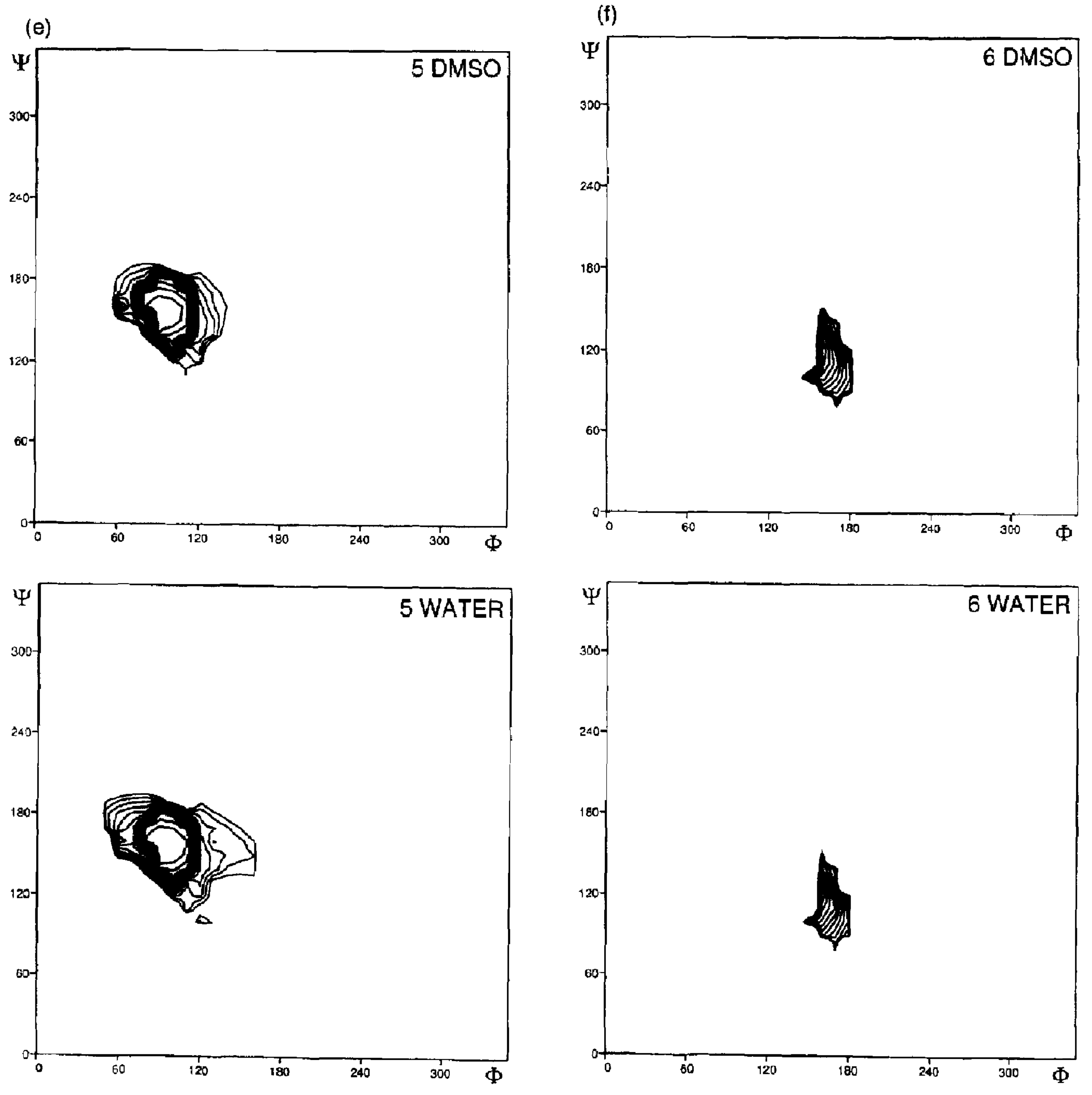

Fig. 2. *Continued.*

$\Phi = 110°$ and $\Psi = 170°$ for **2**, whereas two different conformations occurs approximately at the same extent for **3** at (100°, 90°) and (50°, 90°). Since the spatial arrangement of two contiguous monosaccharide residues in the lowest energy minima of **2** and **3** is different, the directional tendencies inherent in the linkage geometry of **2** and **3** disaccharides also differ. As a result polymer **23** is evenly bent (Fig. 5b), and thus yields the considerably smaller $C_\infty = 24$. This suggests that the low value of C_∞ reflects a tortuosity of the **23** polymer.

Theoretical studies on (1→4)-diaxially linked homopolysaccharides have also been made (Yathindra & Rao, 1972; Whittington *et al.*, 1973; Walkinshaw & Arnott, 1981; Burton & Brant, 1983; Ruggiero *et al.*, 1995). All of them illustrate the important stiffness of these chains. Using standard matrix methods for statistical mechanical averaging, different values of the characteristic ratio, C_∞ between 142 and 280 have been found for (1→4)-α-D-galactans. The resulting values of C_∞ depended on the magnitude of the glycosidic bond angle and whether or not the electrostatic interactions were included in the calculations of (Φ,Ψ) maps. These values are much higher than those calculated in the present study. We assume that this discrepancy is a consequence of the rigid residues approach used for the calculation of potential energy surfaces. This approach

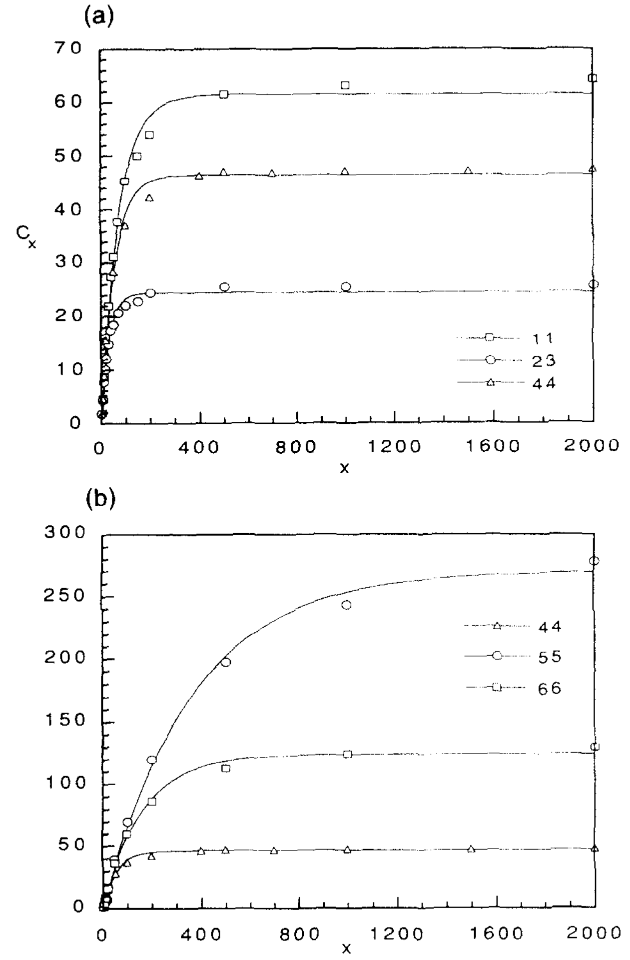

Fig. 3. Characteristic ratio versus degree of polymerization for (a) polygalactan (**11**), polygalacto-galacturonic acid (**23**), and polygalacturonic acid (**44**); and (b) for polygalacturonic acid (**44**), ionized polygalacturonic acid (**55**) and sodium polygalacturonate (**66**) in vacuum at 27°C.

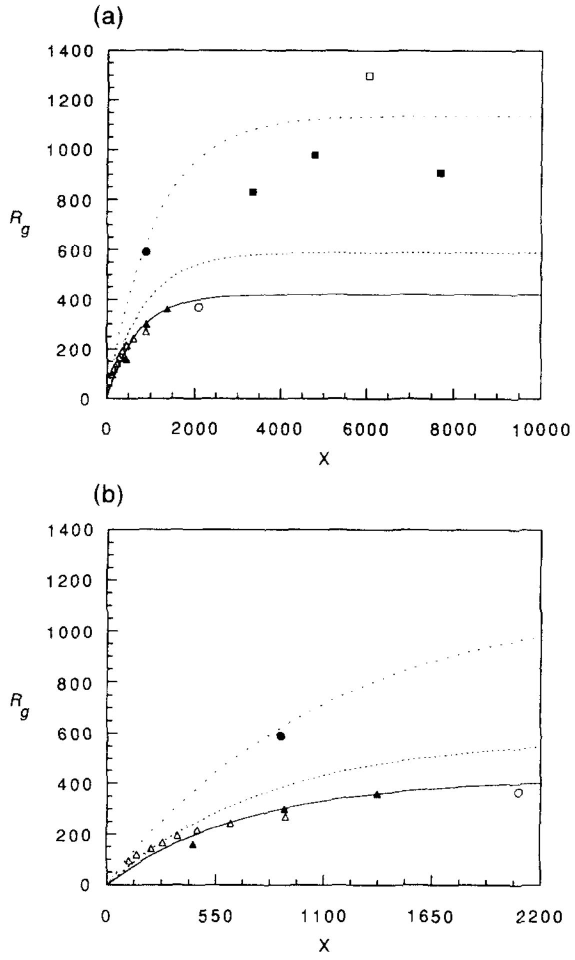

Fig. 4. Comparison between experimental (□ Chapman *et al.*, 1987; ■ Jordan & Brant, 1978; ▲ Ousalem *et al.*, 1993; ● Plashchina *et al.*, 1985; △ Malovikova *et al.*, 1993; ○ Axelos *et al.*, 1987) and calculated R_g values. The lines connect the theoretical values for polygalacto-galacturonic acid (**23**) in vacuum (continuous line), and polygalacturonic acid (**44**) in water (dotted line) and in vacuum (dashed line).

usually produces unrealistically restricted available conformational space and as a result, the C_∞ values are larger.

Influence of the charge, solvent and temperature

The chain dimensions expand upon charging the chain due to electrostatic repulsions. This is documented by calculated C_∞ and L_∞ values. In Fig. 3b the characteristic ratio is plotted as a function of x for **44**, **55**, and **66** molecules. It can be seen from the comparison with the neutral polygalacturonan chain (**44**) that the ionized chain (**55**) assumes a highly extended conformation. The calculated C_∞ value increases from 47 for **44** to 266 for **55**. This may be explained by strong short-range repulsive electrostatic interactions that occur between the charged acids groups. These interactions are responsible for a restriction of the rotational motion to an extremely small area of the potential energy surface for this molecule. The presence of sodium decreases these repulsion interactions and increases the flexibility around the glycosidic linkage in **5**. As a result, the C_∞ value decreases to 124 for **66**.

In order to access the influence of the solvent on the unperturbed chain dimensions, simulations were carried out for all polymers in the isolated state, in DMSO and in water solution. The results listed in Table 1 clearly show the influence of the solvent on the behavior of polymer chain. A change of solvent from DMSO to water leads to an increase of the limiting chain dimensions C_∞. Comparison of the (Φ,Ψ) maps for both solvents suggests the following explanation of these effects. The two most stable areas of the (Φ,Ψ) map computed in the isolated state for the disaccharide **4** are approximately at the same energy levels, whereas solvation effects lead to a strong stabilization of the area where conformers display the largest dipole moment values. In water solution, these conformers are exclusively present and the conformations within the second lowest energy area disappear from the equilibrium

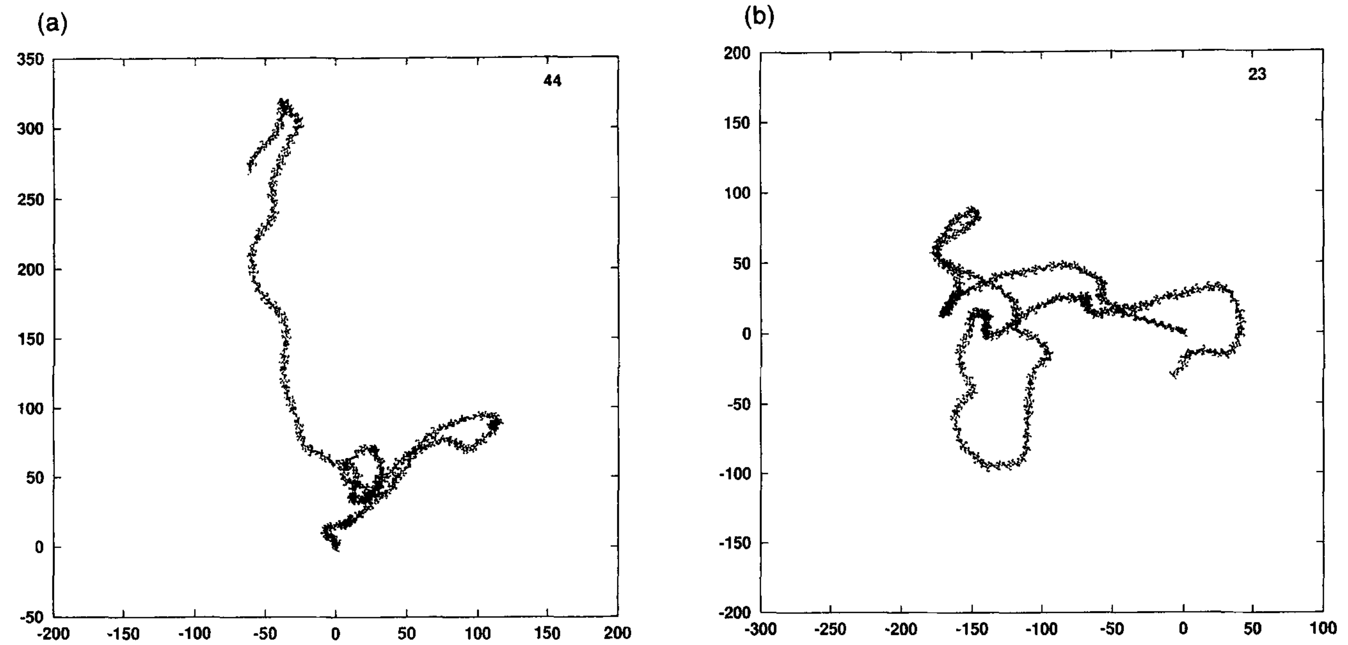

Fig. 5. Perspective drawing of a 300 residue segment of (a) polygalacturonic acid (**44**); and (b) polygalacto-galacturonic acid (**23**) in vacuum, chosen as a representative from 1000 MMC samples. View is perspective to the *xy* plan and the scale is measured in Å. All the carbon and oxygen atoms are represented by points.

Table 1. Monte Carlo calculated (equations 1–4) mean characteristic ratio, C_∞, radius of gyration, R_g (Å), persistence length, L_∞, L_x (Å) and X_{95} of several pectins substances at 27°C

Polymer	Solvent	$<C_\infty>$	$<X_{95}>$	$R_g{}^a$	$<L_x>^a$	$<L_\infty>$
11	vacuum	62	235	642	152	146
	DMSO	67	246	673	152	152
	water	71	264	684	196	175
23	vacuum	24	87	415	67	58
	DMSO	20	99	369	42	28
	water	29	108	445	79	72
44	vacuum	47	180	558	139	129
	DMSO	75	248	715	182	190
	water	146	480	972	402	370
55	vacuum	266	1029	1239	641	645
	DMSO	257	990	1202	635	676
	water	235	898	1172	570	581
66	vacuum	124	486	885	287	300
	DMSO	553	1520	1688	1325	1413
	water	693	1653	1826	1662	1812

[a]Calculated for $x = 2000$.

mixture. On the contrary, solvent effect on the behavior of the charged polymer **55** is reversed as compared to the neutral and the sodium forms. It is clear from the potential energy maps of the parent disaccharide that the solvent increases the freedom of rotation about the glycosidic torsion angles which make the **55** chain much less extended than in vacuum. Increasing temperature increases the accessible conformational space and thus increases the flexibility of the chain. This tends to reduce the calculated C_∞ values.

Calculated solvent specific surfaces for **5** and **6** disaccharides successfully described their solution behavior (Gouvion *et al.*, 1994). However, we are aware of approximations in applying these surfaces to polysaccharides. These calculations are restricted to an isolated chain, ignoring the influence of long-range electrostatic interactions and the effect of salt and ionic strength which are screening the electrostatic interactions. For charged polymers these effects are very important. As a result, the short-range electrostatic interactions are unrealistically high with the consequence that the **55** and **66** configurations of unrealistically large extension are predicted. Nevertheless, the calculated chain dimensions for **55** might be assumed to be dimensions at very low salt concentration and to represent a maximal limit of extension. On the other hand, dimensions calculated for **44** represent the other extreme, the dimensions that in reality may represent the high salt concentration behavior for homopolymer linked with the α-(1→4) linkage.

Comparison with experimental results

The viscometric data on different sources of pectin chains shows poor reliability of the $[\eta]$–M relationship (Hourdet & Muller, 1991b). The reported values for the exponent a in the Mark–Houwink expression are in the range of 0.62–1.84 which can be explained by the presence of aggregates and the heterogeneity of the primary structure. The latter is dependent on both the origin and the extraction conditions of pectins. Based on the values of this exponent, it has been concluded that the solution behavior of pectin chain is either comparable to a flexible coil or rigid or semi-rigid rod. In analyzing the experimental data on pectin chains, the extreme models of a rigid rod or a gaussian coil provide equally poor fits to the experimental values. Most of the authors evaluated the conformational characteristics of those semi-rigid chain polymers by using the worm-like chain model as a first approximation. Considering this model which presupposes a homogeneous chemical structure, the persistence length L_∞ or the characteristic ratio C_∞ can be deduced in order to obtain the best fit for a set of experimental viscometric values. Most of the reported L_∞ values (Axelos & Thibault, 1991; Harding *et al.*, 1991; Hourdet & Muller, 1991a; Malovikova *et al.*, 1993) range from 67 to 90 Å. However, L_∞ values of 20–25 Å were reported for pectins with an extremely low anhydrogalacturonic acid content assuming a ramified structure (Hourdet & Muller, 1991b), as well as $L_\infty = 153$ Å for the most methylated and amidated samples (Axelos & Thibault, 1991). The calculated values for the **44** and **23** polymers are in this range. In contrast, the models with a localized charges (**55** and **66**) predict very large L_∞. The total persistence length L_∞ of the macroion is assumed to be a sum of the intrinsic persistence the length $L_{\infty,o}$ of the uncharged chain and the electrostatic persistence length $L_{\infty,e}$. Assuming this model (Odijk, 1977, 1978; Skolnick & Fixman, 1977) and being aware of approximations used in our calculations, a difference of the persistence lengths between **44** and **55** might represent the estimate of the maximal electrostatic persistence length.

Information about the shape and chain flexibility of pectins have been obtained by employing light scattering measurements. Figure 4 compares experimental R_g values (Jordan & Brant, 1978; Plashchina *et al.*, 1985; Axelos *et al.*, 1987; Chapman *et al.*, 1987; Ousalem *et al.*, 1993; Malovikova *et al.*, 1993) obtained by light scattering for different pectins with calculated values for polygalacturonic acid (**44**) in vacuum and in water, and polygalacto-galacturonic acid (**23**) in vacuum at 27°C. A quantitative comparison is complicated because pectin has structured features other than polygalacturonic acid. However, it is well known that homo-galacturonic acid sequences of lengths between 70–100 residues do occur in the pectin molecules. These models can lead to a better understanding of the contribution of these sequences to the global shape of the pectins. The insertion of rhamnose residues in the backbone as well as the degree of esterification of galacturonic acid units might influence the average shape of the polymer (Deckers *et al.*, 1986; Berth, 1988). Furthermore, the heterogeneity of the chemical structure, especially a difference in the extent of charged residues, between different sources of pectins make it difficult to make a straightforward comparison of the experimental results. In addition, the weight-average molecular weight measured by light scattering generally disagrees with that evaluated by viscometric measurements.

The experimental values of the degree of polymerization, x, that are reported in Fig. 4 are calculated from the experimental weight-average molecular weight assuming a molecular weight per monomer unit of 182 based on neutron scattering data (Axelos *et al.*, 1987). Considering that the experimental error in the measured values of the molar mass are estimated to be less than 20% (Ousalem *et al.*, 1993), this simplification has no serious consequence on the value of x. The lines represent theoretical values of R_g. For clarity, the values of R_g for $x < 2000$ have also been plotted on Fig. 4b.

A large variation of individual experimental results can be extracted from the Fig. 4. Pectins having almost the same chain length (Plashchina *et al.*, 1985; Ousalem *et al.*, 1993; Malovikova *et al.*, 1993) ($x = 901$, 880 and 907) show different R_g values of 300, 590 and 270 Å, respectively. On the contrary, pectins with the same values of R_g (Axelos *et al.*, 1987; Ousalem *et al.*, 1993), (360 and 367 Å) have completely different chain length ($x = 1373$ and 2088) residues. The main problem in the experimental study of the pectin solution behavior is its strong tendency to form aggregates which disturb the physical measurement. The shape of the scattering function shows negative virial coefficients, which can be interpreted as a concentration-dependent association and dissociation. The combination of two different approaches can lead to a better understanding of such behavior. First, the measured mean polymer chain dimensions have to be rationalized in terms of the presence of a very small proportion of aggregated molecules (Jordan & Brant, 1978; Brant & Goebel, 1972). Second, a clarification process should be developed in order to remove the interfering superstructures, most authors have developed more or less suitable procedures. Another source of uncertainty comes from the choice of the value of the projection of the monomer unit on the molecular axis. This value was assumed to be either 4.37 Å (Berth *et al.*, 1994) or 5.10 Å (Palmer & Hartzog, 1945).

Figure 4 clearly shows that the configurational statistics of polygalacturonic acid based on the water specific (Φ,Ψ) relaxed map of the parent disaccharide (**4**) leads to considerably higher values of R_g than most of the experimental data. The only exception is the one reported (Plashchina *et al.*, 1985) where the correlation is excellent but the presence of aggregates was clearly shown. Furthermore, the dependence of R_g on x based on the vacuum (Φ,Ψ) relaxed map of homopolygalacturonic acid is in better agreement with most of the experimental values (Axelos *et al.*, 1987; Ousalem *et al.*, 1993; Malovikova *et al.*, 1993). There is, however, a slight tendency to predict higher values of R_g. A possible explanation is that the unperturbed state is better described by the map computed for the isolated disaccharide. The solvent maps are clearly overestimating short-range interactions. However, electrostatic repulsion between carboxylic acid groups in the isolated state and in solution leads to conformational preferences that diverges from the Θ-conditions. These interactions are screened due to additional salt during experiment. On the other hand, agreement between calculated values for **23** in vacuum and experimental data is satisfactory for the whole range of $x < 2000$. This suggests that differences in both the charge distribution and bulkiness of the pendent groups such as hydroxymethyl or carboxyl ones in the **23** model reasonably well the structure of the pectin chain.

Figure 5 shows a perspective drawings of a typical conformation of a segment ($x = 300$) of $(1\rightarrow4)$-diaxially linked polysaccharides. All carbon and oxygen atoms are shown and represented by circles. These drawings were selected as representative examples from a survey of 1000 chains conformations presented in the MMC samples of chains having 2000 residues. In Fig. 5a, the conformation for polygalacturonic acid **44** in vacuum is presented. The influence of the C5 substituent in the polymer chain **23** on the character of $(1\rightarrow4)$-linked polysaccharides can be appreciated by a comparison of Fig. 5a with Fig. 5b where a typical conformation of **23** is shown.

CONCLUSION

Using a Monte Carlo simulation we have studied the conformational properties of five models of pectin chains, namely four homopolymers: polygalactan **11**, polygalacturonic acid **44**, ionized polygalacturonic acid **55**, sodium polygalacturonate **66**, and a copolymer polygalacto-galacturonic acid **23**. The effect of temperature and solvent was also evaluated.

The calculated mean characteristic ratio, squared radius of gyration, and persistence length as a function of degree of polymerization are discussed for various structural models, temperature, and solvent. The large difference in the values of C_∞ clearly indicate that the difference in the nature, as well as the ionic state, of the side group at C5 in the sugar ring considerably affects the unperturbed dimensions and thereby the spatial extension in solution of the polysaccharides chains. The high C_∞ values of all the studied polysaccharides suggest, however, that every polysaccharides having a $(1\rightarrow4)$-diaxial glycosidic linkage, irrespective of the nature of the side group, behaves as a semi-rigid rod.

A comparison of calculated and experimental R_g revealed that vacuum model of polygalacto-galacturonic acid **23** is consistent with the experimentally measured dimension of pectin in excess salt.

REFERENCES

Anger, H. and Berth, G. (1985) *Carbohydrate Polymers*, **5**, 241–250.

Anger, H. and Berth, G. (1986) *Carbohydrate Polymers*, **6**, 193–202.

Axelos, M. A. V., Lefebvre, J. and Thibault, J.-F. (1987) *Food Hydrocolloids*, **1**, 569–570.

Axelos, M. A. V. and Thibault, J.-F. (1991) *International Journal of Biological Macromolecules*, **13**, 77–82.

Berth, G. (1988) *Carbohydrate Polymers*, **8**, 105–117.

Berth, G. (1992) *Carbohydrate Polymers*, **19**, 1–9.

Berth, G., Dautzenberg, H. & Rother, G. (1994) *Carbohydrate Polymers*, **25**, 177–185.

Brant, D. A. (1976) *Quarterly Reviews in Biophysics*, **4**, 527–596.

Brant, D. A. (1980) in *The Biochemistry of Plants*, Vol. 3, ed. J. Preiss, pp. 425–472. Academic Press, New York.

Brant, D. A. and Burton, B. A. (1981) in *Solution Properties of Polysaccharides*, ACS Symposium Series, Vol 150, ed. D. A. Brant, pp. 81–99. American Chemical Society, Washington DC.

Brant, D. A. and Christ, M. D. (1990) in *Computer Modeling of Carbohydrates Molecules*, ACS Symposium Series, Vol. 430, eds A. D. French and J. W. Brady, pp. 42–68. American Chemical Society, Washington DC.

Brant, D. A. and Goebel, K. D. (1972). *Macromolecules*, **5**, 522–530.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. (1983) *Journal of Computational Chemistry*, **4**, 187–217.

Buliga, G. S. and Brant, D. A. (1987) *International Journal of Biological Macromolecules*, **9**, 71–76.

Burton, B. A. and Brant, D. A. (1983) *Biopolymers*, **22**, 1769–1792.

Cesaro, A. and Brant, D. A. (1981) *Solution Properties of Polysaccharides*, ACS Symposium Series, Vol. 150, ed. D.A. Brant, pp. 513–527. American Chemical Society, Washington DC.

Chapman, H. D., Morris, V. J. and Selvendran, R. R. (1987) *Carbohydrate Research*, **165**, 53–68.

Deckers, H. A., Olieman, C., Rombouts, F. M. and Pilnik, W. (1986) *Carbohydrate Polymers*, **6**, 361–378.

DeLucas, L. J., Gartland, G. L. and Bugg, C. E. (1978) *Carbohydrate Research*, **62**, 213–221.

Fishman, M. L., El-Atawy, Y. S., Sondey, S. M., Gillespie, D. T. and Hicks, K. B. (1991) *Carbohydrate Polymers*, **15**, 89–104.

Fishman, M. L., Pepper, L. and Pfeffer, P. E. (1986) *Advanced Chemistry Series*, **213**, 57–70.

Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules*. Interscience, New York.

Gagnaire, D., Pérez, S. and Tran, V. (1982) *Carbohydrate Polymers*, **2**, 171–191.

Gouvion, C., Mazeau, K., Heyraud, A., Taravel, F. R. and Tvaroska, I. (1994) *Carbohydrate Research*, **261**, 187–202.

Gouvion, C., Mazeau, K. and Tvaroska, I. (1995) *Journal of Molecular Structures*, **344**, 157–170.

Harding, S. E., Berth, G., Ball, A., Mitchell, J. R. and Garcia de la Torre, J. (1991) *Carbohydrate Polymers*, **16**, 1–15.

Hourdet, D. and Muller, G. (1987) *Carbohydrate Polymers*, **7**, 301–312.

Hourdet, D. and Muller, G. (1991a) *Carbohydrate Polymers*, **16**, 113–135.

Hourdet, D. and Muller, G. (1991b) *Carbohydrate Polymers*, **16**, 409–432.

IUPAC-IUB (1971) *Archives of Biochemistry and Biophysics*, **145**, 405–412.

Jordan, R. C. and Brant, D. A. (1978) *Biopolymers*, **17**, 2885–2895.

Jordan, R. C., Brant, D. A. and Cesàro, A. (1978) *Biopolymers*, **17**, 2617–2632.

Kitamura, S., Okamoto, T., Nakata, Y., Hayashi, T. and Kuge, T. (1987) *Biopolymers*, **26**, 537–548.

Kravtchenko, T. P., Berth, G., Voragen, A. G. J. and Pilnik, W. (1992) *Carbohydrate Polymers*, **18**, 253–263.

Kravtchenko, T. P., Voragen, A. G. J. and Pilnik, W. (1992) *Carbohydrate Polymers*, **19**, 115–124.

Malovikova, A., Rinaudo, M. and Milas, M. (1993) *Carbohydrate Polymers*, **22**, 87–92.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. M., Teller, A. H. and Teller, E. (1953) *Journal of Physical Chemistry*, **21**, 1087–1092.

Newton, M. D. and Jeffrey, G. A. (1977) *Journal of the American Chemical Society*, **99**, 2413–2421.

Odijk, T. (1977) *Journal of Polymer Science, Polymer Physics Edition*, **15**, 477–483.

Odijk, T. and Houwaart, A. C. (1978) *Journal of Polymer Science, Polymer Physics Edition*, **16**, 627–639.

Ousalem, M., Busnel, J. P. and Nicolai, T. (1993) *International Journal of Biological Macromolecules*, **15**, 209–213.

Owens, H. S., Lotzkar, H., Schultz, T. H. and Maclay, W. D. (1946) *Journal of the American Chemical Society*, **68**, 1628–1630.

Palmer, K. J. and Hartzog, M. B. (1945) *Journal of the American Chemical Society*, **67**, 2122–2127.

Pierotti, R. A. (1976) *Chemical Review*, **76**, 717–726.

Plashchina, I. G., Semenova, M. G., Braudo, E. E. and Tolstoguzov, V. B. (1985) *Carbohydrate Polymers*, **5**, 159–179.

Quanta 4.0 (1994) Molecular Simulations Inc., Burlington, MA, USA.

Reiss, H., Frisch, H. L. and Lebowitz, J. L. (1959) *Journal of Chemical Physics*, **31**, 369–380.

Ruggiero, J. R., Urbani, R. and Cesaro, A. (1995) *International Journal of Biological Macromolecules*, **17**, 213–218.

Skolnick, J. and Fixman, M. (1977) *Macromolecules*, **10**, 944–948.

Svensson, G., Albertsson, J., Svensson, C., Magnusson, G. and Dahmen, J. (1986) *Carbohydrate Research*, **146**, 29–38.

Thanomkul, S., Hjortas, J. and Soerum, H. (1976) *Acta Crystallographica*, **B32**, 920–922.

Tvaroska, I. (1982) *Biopolymers*, **21**, 1887–1897.

Tvaroska, I. (1989) *Pure and Applied Chemistry*, **61**, 1201–1216.

Tvaroska, I. and Bleha, T. (1989) *Advances in Carbohydrate Chemistry and Biochemistry*, **47**, 45–123.

Tvaroska, I. and Kozar, T. (1980) *Journal of the American Chemical Society*, **102**, 6929–6936.

Tvaroska, I. and Kozar, T. (1986) *Theoretica Chimica Acta*, **70**, 99–114.

Urbani, R. and Cesàro, A. (1991) *Polymer*, **32**, 3013–3020.

Walker, R. (1990) *The Chemistry & Technology of Pectin*. Academic Press, London.

Walkinshaw, M. D. and Arnott, S. (1981) *Journal of Molecular Biology*, **153**, 1075–1085.

Whittington, S. G., Glover, R. M. and Hallman, G. M. (1973) *Journal of Polymer Science*, **42**, 1493–1497.

Yathindra, N. and Rao, V. S. R. (1972) *Journal of Polymer Science*, **10**, 1369–1382.

Yui, T., Kobayashi, H., Kitamura, S. and Imata, K. (1994) *Biopolymers*, **34**, 203–208.