ELSEVIER

Polymer 43 (2002) 509–516

polymer

www.elsevier.com/locate/polymer

Conformational analysis of aqueous pullulan oligomers: an effective computational approach ☆

Jennifer H.-Y. Liu[a], Ken A. Brameld[b], David A. Brant[a,*], William A. Goddard III[b]

[a]*Department of Chemistry, University of California, Irvine, CA 92697-2025, USA*
[b]*Materials and Process Simulation Center, Beckman Institute (139-74), Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA*

Dedicated to Wayne L. Mattice in honor of his outstanding contributions to computational polymer science

Abstract

Three rotational isomeric state (RIS) models for pullulan, a polysaccharide defined as [→6)-α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→]$_n$, are presented. These RIS models are based upon simple molecular mechanics and solvation calculations and differ in the complexity of the energy evaluations used to determine Boltzmann weighting factors for each isomeric state. Each model is evaluated for its ability to predict correctly experimental small angle X-ray scattering data and the radius of gyration (R_g) for the hexameric pullulan oligomer $(G_3)_2$. A simple model based upon only disaccharide conformational energies in conjunction with a correction for solvation energies, as determined using a continuum Poisson–Boltzmann solvation method, was best and gave an excellent fit to experimental measurements.

Keywords: Pullulan; Conformational analysis; Small angle X-ray scattering

1. Introduction

Carbohydrates, proteins, nucleic acids, and lipids are four major classes of biomolecules. Unlike proteins and nucleic acids, which have linear sequences of amide bonds and 3′–5′ phosphodiester bonds, respectively, carbohydrates can have branched structures with diverse linkage isomers. Such structural variation and complexity allows carbohydrates to play a critical role in many biological functions [1–3]. In addition, carbohydrates contain considerable polar functionality including hydroxyl, hydroxymethyl, acetal, carboxyl, and amido groups that exert strong stereoelectronic and electrostatic effects on the conformational behavior. For these reasons, the development of computational methods for modeling properties of carbohydrates is particularly challenging and has been the focus of extensive study [4–20].

Pullulan, an extracellular polysaccharide produced by the fungus *Aureobasidium pullulans*, is a linear water soluble glucan with the chemical structure [→6)-α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→4)-α-D-glucopyranosyl-(1→]$_n$ [21]. Pullulan can also be produced extracellularly in a fermentation process to yield powdered pullulan that is odorless, tasteless, and non-hygroscopic under normal atmospheric conditions [22,23]. In principal, pullulan is co-polymeric in that both (1→4) and (1→6) glycosidic linkages are involved in the backbone in the ratio of 2:1 [24]. A hexameric segment of pullulan containing a single (1→6)-linkage is shown in Fig. 1. Disaccharides **1** and **2**, possessing (1→4)- and (1→6)-glycosidic linkages, respectively are shown in Fig. 2, as are the torsion angles ϕ_4 (O_5–C_1–O–C_4'), ψ_4 (C_1–O–C_4'–C_5') of the (1→4)-linkage and ϕ_6 (O_5–C_1–O–C_6'), ψ_6 (C_1–O–C_6'–C_5'), and ω_6 (O–C_6'–C_5'–O_5') of the (1→6)-linkage. Motions along these backbone torsional coordinates are the governing source of pullulan chain flexibility. In Figs. 2a and b, *gg* and *gt* refer to the torsional states of the methoxy group exocyclic torsion angle τ_6 (O_5–C_5–C_6–O_6). In this convention, the conformation of torsional angle (O_5–C_5–C_6–O_6) is stated first; then the torsional angle (C_4–C_5–C_6–O_6) is stated [25]. In terms of τ_6, the *gg* state refers to a torsional rotation of approximately −60° and *gt* refers to a torsional rotation of about +60°.

It is reported by Nakatani and co-workers that pullulan solutions can be used as a separation matrix for capillary electrophoresis for SDS-proteins [26]. The industrial and medical applications are reported extensively elsewhere

☆ This paper was originally submitted to *Computational and Theoretical Polymer Science* and received on 29 October 2000; received in revised form on 8 February 2001; accepted on 9 February 2001. Following the incorporation of *Computational and Theoretical Polymer Science* into *Polymer*, this paper was consequently accepted for publication in *Polymer*.

* Corresponding author. Tel.: +1-949-824-6019; fax: +1-949-824-8571.
E-mail address: dbrant@uci.edu (D.A. Brant).

0032-3861/02/$ - see front matter
PII: S0032-3861(01)00420-7

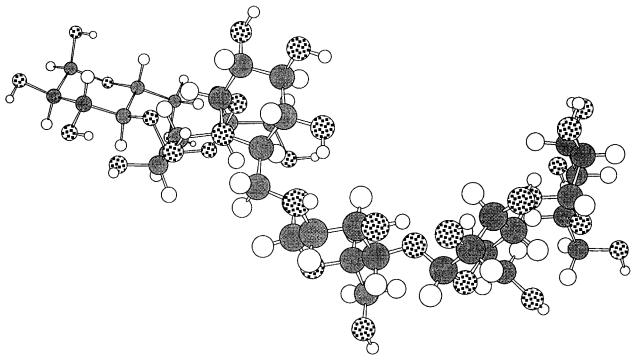

Fig. 1. Hexameric pullulan oligomer $(G_3)_2$ possessing four $(1 \rightarrow 4)$-linkages and one $(1 \rightarrow 6)$-linkage.

[22,23]. It also has been proposed that pullulan can be used in personal care products such as hand and facial lotions, shampoos, and cosmetics [27]. Furthermore, because of its ready water solubility and its commercial availability in samples of narrow molecular weight distribution covering a wide range of mean molar mass, pullulan serves as a prototypical non-ionic, water soluble polysaccharide and as a calibration standard for aqueous size exclusion chromatography (SEC) [28].

The measurable characteristics of aqueous pullulan have been investigated intensively with convincing agreement among the several studies [29–33]. Therefore, the asymptotic characteristic ratio and persistence length and the chain length dependence of such aqueous solution properties as the intrinsic viscosity, the root-mean-square radius of gyration, the diffusion coefficient, the sedimentation coefficient, and the osmotic second virial coefficient are known.

The equilibrium solution behavior of pullulan has been subjected to simulation at the atomistic level to produce satisfying agreement between the observed solution characteristics and the predictions of the model [34]. Both the absolute unperturbed polymer coil dimensions and the sign of the temperature coefficient of those dimensions are reproduced by the model, which implies a random coil chain with a weak tendency to follow a pseudohelical trajectory [34]. Nishinari and co-workers also confirmed that lower molecular weight ($MW = 5.3 \times 10^3$–1015×10^3) pullulan molecules behave as random coils in water [35]. Recently small angle X-ray scattering (SAXS) studies on a series of pullulan oligomers in aqueous solution have served as the basis for an improved pullulan chain model capable of fitting both the SAXS data on oligomers and the earlier data on high molecular weight pullulan [36].

The present work uses an alternative molecular mechanics approach in conjunction with property prediction methods to investigate the conformations of aqueous pullulan using a structurally realistic atomistic model. The CFF95 force field [37–39] is used in conjunction with a continuum solvation description [40], to compute conformational properties for the disaccharides **1** and **2**, shown in Fig. 2, that contain $(1 \rightarrow 4)$- and $(1 \rightarrow 6)$-glycosidic linkages, respectively. The predicted scattering function and radius of gyration for the pullulan hexasaccharide $(G_3)_2$, determined from a rotational isomeric state (RIS) model based upon the conformations and relative energies of only the disaccharides, is in excellent agreement with the experimentally measured SAXS data. Interestingly, a more computationally intensive model using optimized structures and energies as determined for the full hexasaccharide does not improve the fit to the experimental data. The shortcomings of each model system are discussed and a general strategy is presented that may be applied to any linear polysaccharide system.

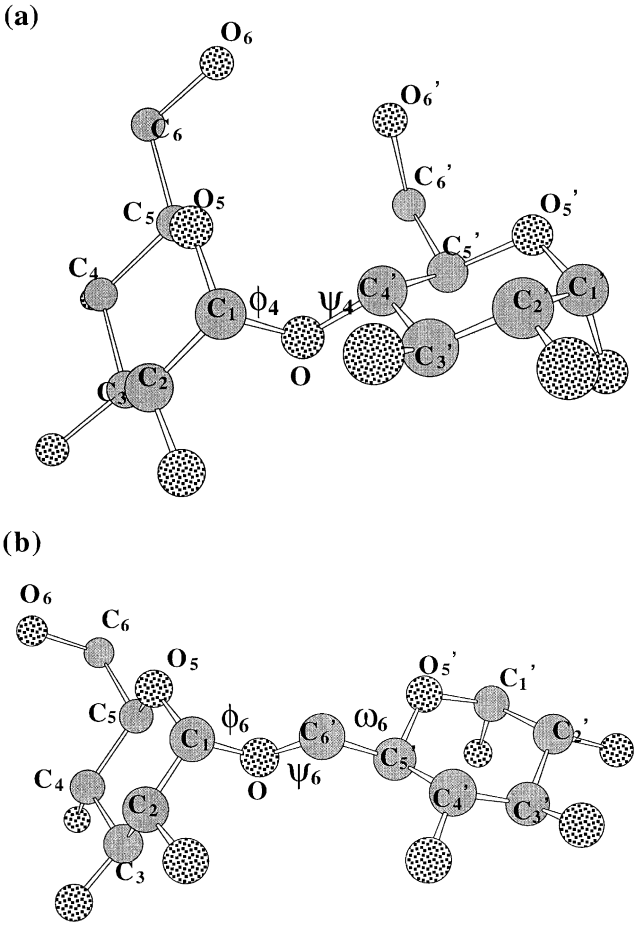

Fig. 2. (a) The $(1 \rightarrow 4)$-linked disaccharides **1**. Glycosidic torsion angles ϕ_4 (O_5–C_1–O–C_4') and ψ_4 (C_1–O–C_4'–C_5') are shown. Exocyclic torsion angle τ_6 (O_5–C_5–C_6–O_6) is in the *gt* state in the left residue and in the *gg* state in the right residue. Only carbon and oxygen atoms are illustrated. (b) The $(1 \rightarrow 6)$-linked disaccharides **2**. Glycosidic torsion angles ϕ_6 (O_5–C_1–O–C_6'), ψ_6 (C_1–O–C_6'–C_5'), and ω_6 (O–C_6'–C_5'–O_5') are shown. Exocyclic torsion angle τ_6 (O_5–C_5–C_6–O_6) is in the *gg* state in the left residue. Only carbon and oxygen atoms are illustrated.

2. Methods

All molecular mechanics calculations were performed using the CERIUS[2] software program, with implementation of the CFF95 force field [37–39] available in that program.

Solvation energies, including a cavity correction term, were calculated using the Poisson–Boltzmann solvation program (PBF) available as part of the JAGUAR quantum chemistry package [40]. Solvation energies calculated with PBF are comprised of two parts. First, the change in electrostatic energy associated with the transfer of the solute from vacuum into solution is calculated by solving the linearized Poisson–Boltzmann equation [41]. The solute is assumed to have a dielectric constant of 4, while the bulk solvent region beyond the solvent-accessible surface of the solute has a dielectric constant of 80. The second term provides an energy estimate for creation of a cavity within the bulk solvent (i.e. solvent displacement) and is dependent upon solute volume [42].

An exhaustive conformational search of the (1 → 6)-linked disaccharide (**2**) was carried out using CERIUS^2 as follows: Torsion angles ϕ_6, ψ_6, and ω_6 were varied from 0 to 360° in 30° increments. All hydroxyl and hydroxymethyl group exocyclic torsion angles were varied from 0 to 360° in 120° increments. Each conformation was energy minimized twice, first with a harmonic potential to constrain the torsion angles being varied and second with no constraints so as to allow for full relaxation. Only those conformations within 5 kcal/mol of the global minimum were retained for further study, and all duplicate conformations, defined as having a coordinate root mean square difference <0.01 Å, were removed. This resulted in a total of 12 unique conformations that could be grouped into four families (A–D) based on the ϕ_6, ψ_6, and ω_6 torsion angles.

The 12 conformations of **2** were then used to develop a RIS model for pullulan $(G_3)_2$. A hexasaccharide chain was built by connecting two (1 → 4)-linked trisaccharides through a single (1 → 6)-linkage. Only a single conformational state was permitted to each (1 → 4)-linkage, because there is a single narrowly confined accessible energy minimum on the corresponding energy surface. Thus, most of the pullulan chain flexibility arises from the variety of conformations available to the (1 → 6)-linkage [36]. Correspondingly, the ensemble of conformations considered for the hexasaccharide $(G_3)_2$ arises entirely from the 12 allowed conformations of **2**. Averages over this ensemble then produced a result suitable for comparison with experimental results on $(G_3)_2$. It would, of course, be inappropriate to choose only a single state for the (1 → 4)-linkage, if it were our purpose to treat the properties of the (1 → 4)-linked α-D-glucan homopolymer. In this work, however, the pullulan chain conformation is sensitive primarily to the details of the (1 → 6)-linkage [36].

Three methods were used to determine the relative energies and thus the statistical weights for the 12 conformers. For Model 1, each conformation was weighted by a Boltzmann factor calculated using the minimized molecular mechanics energy (Table 1) of the isolated (1 → 6)-linked dimers; contributions from the (1 → 4)-linked disaccharides in the molecule were ignored because they make a constant energy contribution in the approximation of Model 1. In Model 2, the relative energies of the 12 conformations were further refined by adding a solvation energy correction, again calculated only for the isolated (1 → 6)-linked disaccharides. Solvation energies were calculated as described above and are listed in Table 1. Finally, each of the 12 full pullulan $(G_3)_2$ chains was energy minimized and solvation energies calculated (Table 2) to create Model 3.

Predicted SAXS data (Kratky plots) were generated using the Diffraction-Amorphous module of CERIUS^2. The total

Table 1
Conformations, relative energies (kcal/mol) and statistical weights for disaccharide **2**

Conformation [ϕ_6, ψ_6, ω_6][a]	Relative gas phase E[b]	Relative solvation E[c]	Relative cavity E[d]	Model 1 σ[e]	Model 2 σ[e]
A1 (*gg*) [64.7, −163.2, 68.2]	2.85	3.37	0.10	0.007	0.030
A2 (*gg*) [69.4, 170.7, 66.4]	2.90	2.89	0.16	0.006	0.057
A3 (*gg*) [64.8, 170.2, 67.7]	3.27	1.97	0.18	0.003	0.140
A4 (*gt*) [63.2, −165.9, 68.2]	2.98	3.56	0.09	0.005	0.018
A5 (*gt*) [61.8, −172.6, 66.9]	3.84	1.99	0.13	0.001	0.056
B1 (*gt*) [92.0, −94.8, 73.0]	0.00[f]	7.18	0.01	0.831	0.007
B2 (*gt*) [94.0, −92.8, 71.4]	1.38	5.21	0.02	0.081	0.019
B3 (*gt*) [97.3, −89.7, 69.1]	3.95	3.03	0.03	0.001	0.009
C1 (*gt*) [64.8, 175.5, 178.2]	5.08	0.00[g]	0.13	0.000	0.198
D1 (*gg*) [62.2, −178.7, −66.2]	1.74	3.01	0.08	0.044	0.381
D2 (*gt*) [60.2, 179, −66.4]	2.52	3.26	0.08	0.012	0.067
D3 (*gt*) [78.8, −146.5, −65.2]	2.76	3.87	0.00[h]	0.008	0.018

[a] Conformations are described by groups A–D and by torsion angles (°).
[b] Relative gas phase energies (kcal/mol) calculated using CFF95 force field.
[c] Relative solvation energies (kcal/mol) determined with Poisson–Boltzmann continuum method.
[d] Relative cavity energy (part of solvent correction terms).
[e] Statistical weight calculated at 298.15 K.
[f] Absolute gas phase energy = 28.99 kcal/mol.
[g] Absolute solvation energy = −52.96 kT.
[h] Absolute cavity energy = 5.95 kT.

Table 2
Relative energies (kcal/mol) and statistical weights for fully geometry optimized pullulan $(G_3)_2$

Conformation[a]	Relative gas phase E[b]	Relative solvation E[c]	Relative cavity E[d]	Model 3 σ[e]
A1	4.39	4.16	0.09	0.596
A2	14.13	0.08	0.68	0.000
A3	3.64	5.13	0.58	0.178
A4	3.65	5.11	0.58	0.182
A5	4.82	6.16	0.53	0.005
B1	2.21	13.03	0.10	0.000
B2	0.94	12.98	0.00[f]	0.000
B3	5.49	12.41	0.10	0.000
C1	0.00[g]	9.51	0.75	0.039
D1	16.11	0.00[h]	0.86	0.000
D2	16.30	0.53	0.83	0.000
D3	16.91	0.52	0.77	0.000

[a] Conformations are described by the same descriptors used in Table 1 and indicate the starting geometry before energy minimization.
[b] Relative gas phase energies (kcal/mol) calculated using CFF95 force field.
[c] Relative solvation energies (kcal/mol) determined with Poisson–Boltzmann continuum method.
[d] Relative cavity energy (part of the solvent correction terms).
[e] Statistical weight calculated at 298.15 K.
[f] Absolute cavity energy = 9.87 kT.
[g] Absolute gas phase energy = 64.78 kcal/mol.
[h] Absolute solvation energy = −106.65 kT.

scattering is calculated using the Debye formula:

$$I(Q) = \sum_j \sum_k \frac{f_j(Q) f_k(Q) \sin Q r_{jk}}{Q r_{jk}}$$

where $j \neq k$, Q is the scattering vector, f_j and f_k are the scattering factors for atoms j and k, respectively, and r_{jk} is the distance between atoms j and k. Default values were used for all parameters except for the following: $Q_{min} = 0.00$, $Q_{max} = 0.5$, interval = 0.01 Å^{-1}. This limits the scattering vector to the experimental range (0.0–0.5 Å^{-1}), and points are calculated at 0.01 Å^{-1} intervals. Scattering functions for each of the 12 $(G_3)_2$ conformers were calculated and weighted according to Boltzmann factors for Models 1, 2, and 3. Because the experimental $(G_3)_2$ fraction actually contained 23.8% (w/w) G_3, the scattering function for the (1 → 4)-linked trisaccharide was also calculated and averaged with proper weighting with the function computed for $(G_3)_2$ [36]. To allow direct comparison to the experimental data, the calculated and experimental intensities were scaled such that $I(Q) = 1$ at $Q = 0$.

3. Results

Using disaccharides **1** and **2** (Fig. 2) that contain a (1 → 4)- and (1 → 6)-glycosidic linkage, respectively, three RIS models for pullulan were developed, as outlined above. When constructing an RIS model to be used for property prediction, two basic steps are required. First, a suitable sampling method is necessary to explore fully the conformational space available to the system of interest. Second, given an ensemble of conformations, the relative energy of each conformation must be calculated so as to determine the statistical weight each conformation contributes to the ensemble. These two steps are addressed separately in the following sections. Finally, the ability of the RIS models to predict the radius of gyration and scattering function (Kratky plot) for pullulan $(G_3)_2$ is evaluated.

3.1. Conformational searching

Disaccharide **1** contains a single (1 → 4)-glycosidic linkage that is defined by two torsion angles: ϕ_4 (O_5–C_1–O–C_4') and ψ_4 (C_1–O–C_4'–C_5'). Using the CERIUS^2 software package and the CFF95 force field [37–39], the potential energy surface determined by these two torsions was examined. Only one energy minimum was found; it is centered around $\phi_4 = 105°$ and $\psi_4 = -155°$. This single minimum compares favorably to results obtained using the AMBER* force field and in crystal structures [14,36]. For this reason, only a single conformation for the (1 → 4)-glycosidic linkage was considered in all three RIS models. By removing variability in the (1 → 4)-glycosidic linkage, the RIS models are simplified and the chain ensemble properties are determined largely by the conformations about the (1 → 6)-glycosidic linkage. This same approach was taken by Liu et al. [36].

The accessible conformations for disaccharide **2**, which possesses a (1 → 6)-glycosidic linkage, are considerably more diverse than for **1** and require a more exhaustive search method to assess fully. Using the conformer search options available in CERIUS^2, the three torsion angles that define the (1 → 6)-glycosidic linkage, ϕ_6 (O_5–C_1–O–C_6'), ψ_6 (C_1–O–C_6'–C_5'), and ω_6 (O–C_6'–C_5'–O_5'), were examined carefully. In addition, the exocyclic torsion angles formed by both hydroxyl and hydroxymethyl groups were

also varied, although by larger increments than the backbone torsion angles. The details of the search method are presented in Section 2.

From this exhaustive conformational search, 12 unique conformations of **2** were identified that had energies within 5 kcal/mol of the global minimum, as determined by gas phase calculations using the CFF95 force field. Table 1 lists the energies and ϕ_6, ψ_6, and ω_6 torsion angles for each of these conformations. Based on the torsion angles listed in Table 1, it is evident that the different conformations of **2** may be subdivided into four groupings that are labeled A–D. Within each group, there are conformational differences associated with RISs of the exocyclic groups that give rise to different gas phase energies.

Pullulan G_3 possesses three glucose residues and only two (1 → 4)-glycosidic linkages. As described in Section 2, only one conformation for the (1 → 4)-glycosidic linkage was considered. Thus in our RIS model for pullulan G_3, there is only one possible chain conformation. In contrast, pullulan $(G_3)_2$ contains four (1 → 4)-glycosidic linkages and one (1 → 6)-glycosidic linkage. Since the (1 → 6)-glycosidic linkage was found to have up to 12 low energy conformations, the entire ensemble of low energy pullulan $(G_3)_2$ chains contains 12 distinct members. For larger pullulan oligomers, the number of conformationally distinct chains increases exponentially with each (1 → 6)-glycosidic linkage (e.g. $(G_3)_3$ has 144 distinct conformers and $(G_3)_4$ has 1728 distinct conformers), many of which are unpopulated and contribute very little to the overall ensemble. Given this diverse set of chain conformations, the remaining difficulty is to correctly assign the statistical weight of each chain.

3.2. Conformational energies

We have used two general methods to determine the conformational energies and statistical weights for the polymer chains within an ensemble. In the simplest case, each polymer chain is subdivided into independent dimers, and the backbone linkage within a dimer is approximated as an isolated system. Subsequently, a sum over all dimers, i.e. linkages, gives the total chain energy, and only nearest neighbor intramolecular interactions contribute to the total energy. A more complex method is to evaluate the energy of each backbone linkage in the presence of the complete polymer chain, i.e. the entire oligomer is considered simultaneously. For much larger systems, computation of the energy of the entire polymer chain is not tractable. In those cases, it is frequently assumed for polysaccharides that relative conformational energies of individual linkages between monomers provide a sufficient basis on which to generate an ensemble of polymer chain conformations [8]. For the pullulan $(G_3)_2$ system, both methods are available, and they are compared here for their ability to reproduce experimentally observed properties.

Three models were used to calculate the relative chain energies for pullulan $(G_3)_2$, as described in Section 2. Models 1 and 2 are based upon energetics from conformational searching of only disaccharide **2**. While the 12 energy minimized conformations of the isolated (1 → 4)- and (1 → 6)-glycosidic linkages are known to be of low energy, there is no certainty that these same backbone geometries will remain energetically favorable in the presence of the other sugar residues which form the entire pullulan $(G_3)_2$ molecule. However, visual inspection indicated that none of the 12 chain conformations resulted in gross steric clashes. In Model 1, the statistical weights for the 12 chains are calculated from the gas phase energies of disaccharide **2** and are listed in Table 1.

The high density of polar functional groups found in polysaccharides suggests that gas phase calculations will offer a poor description of pullulan $(G_3)_2$ in aqueous solution. Model 2 improves upon Model 1 by including two corrections to the gas phase energies for the effects of solvent. First, the solvation energy is calculated using the Poisson–Boltzmann continuum method and, second, a cavity term is added to correct for water displacement. Both of these quantities are calculated using the PBF solvation program that is available as part of the JAGUAR quantum chemistry program [40]. Solvation energies and the Model 2 statistical weights for each conformation of disaccharide **2** are also listed in Table 1. Both Models 1 and 2 assume that conformations and energies from disaccharides may be extended to polysaccharides. In contrast, Model 3 considers the entire pullulan $(G_3)_2$ chain. Starting from the 12 conformations used for Models 1 and 2, each chain is energy minimized using the CFF95 force field. The global conformations change little upon geometry optimization. However, the relative energies calculated for the full hexamers are considerably different from those determined by summing the energies of constituent disaccharides (Models 1 and 2). To these gas phase energies are added corrections for solvation as determined for the entire pullulan $(G_3)_2$ molecule. Any additional conformational rearrangements that might be provoked by solvation effects are ignored. Table 2 lists these energies and the resulting statistical weights. Because Model 3 makes fewer approximations, it is of considerable interest to see if this model can achieve better correspondence with experimental data than Models 1 and 2.

3.3. Comparison to SAXS experiments

SAXS experiments offer insight to the local and global conformation of polysaccharides in aqueous solution. Subtle differences in backbone conformation can give rise to significant changes in the radius of gyration and scattering function. Fig. 3 compares the Kratky plots for representative pullulan $(G_3)_2$ structures from structural groups A–D. In the region above $Q = 0.2$ Å^{-1}, which is most diagnostic for specific chain conformations, the more extended conformers (i.e. Group D) show a monotonic increase in

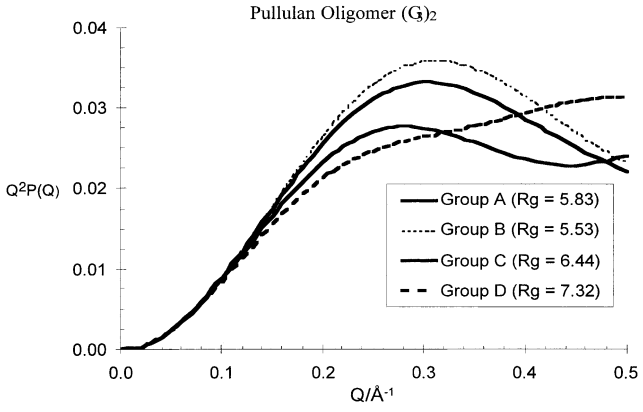

Fig. 3. Kratky plots for representative members from families A–D for pullulan $(G_3)_2$.

$Q^2P(Q)$, whereas the three more compact conformer families display a maximum. The oligomer G_3 is the simplest member of the pullulan oligomer family. It contains only two (1 → 4)-glycosidic linkages, and in RIS Models 1–3 only one conformation for each (1 → 4)-glycosidic linkage is allowed. Because the experimental data reported for $(G_3)_2$ in Fig. 5 refer to a sample with some contamination from G_3, the contribution from this contaminant must be taken into account when considering the scattering function of $(G_3)_2$. Fig. 4 shows the experimentally measured and predicted Kratky plots for pullulan G_3. The fit is extremely good with the predicted plot reproducing the experimental values closely. This encouraging result for G_3 lends confidence to the model for $(G_3)_2$, which provides a more demanding test of the theory owing to the ensemble of molecular shapes expected for the hexamer containing a rather flexible (1 → 6)-linkage.

Fig. 5 compares the experimental Kratky plot and radius of gyration of $(G_3)_2$ to those predicted by each RIS model. The models take into account the correct weighting of each pullulan $(G_3)_2$ chain conformation and the presence of 23.8% (w/w) pullulan G_3 in the experimental sample.

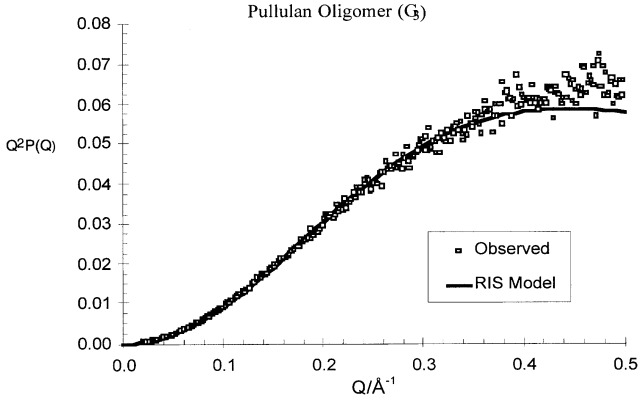

Fig. 4. Comparison of experimental and model-based Kratky plots for pullulan G_3.

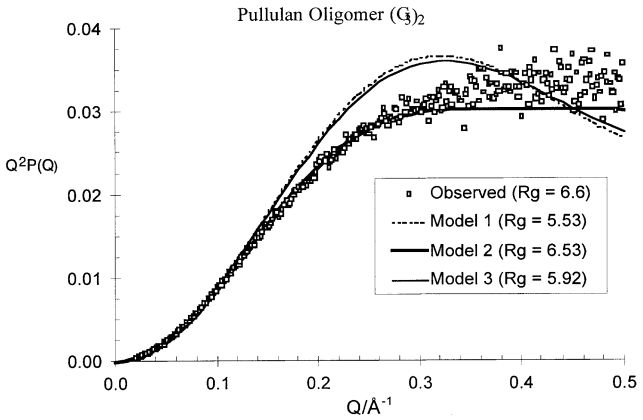

Fig. 5. Comparison of experimental and model-based (Models 1–3) Kratky plots for $(G_3)_2$. The calculated radius of gyration is 5.53 Å for Model 1, 6.53 Å for Model 2, 5.92 Å for Model 3, respectively.

Model 2 is far better than either Model 1 or 3 in correlating the experimentally measured radius of gyration (6.53 Å, predicted, versus 6.57 ± 0.16 Å, observed) and the scattering function $Q^2P(Q)$. Indeed, the fit to the experimental R_g is almost quantitative and indicates that the global conformational characteristics of pullulan $(G_3)_2$ in aqueous solution are well reproduced by RIS Model 2. Moreover, the satisfactory fit at higher Q suggests that the model also reproduces the shorter range structural features well.

4. Discussion

These results for pullulan suggest a method for development of a RIS model for polysaccharides that is both simple and accurate. An exhaustive conformational search using the CFF95 force field can generate low energy conformations (RISs) for the relevant disaccharide segments of the polymer. The resulting gas phase relative energies of these disaccharides are then corrected for solvation effects using a Poisson–Boltzmann continuum solvent description. All conformations of a polysaccharide chain of the desired length are generated, consistent with a given set of RISs for the dimeric segments. Finally, the statistical weights for each chain conformation can be computed using the assumption of independent contributions to the configuration energy of the chain from the dimeric segments [8]. Given the statistical weights, mean conformational properties for the polymer can be computed directly from the ensemble of chain conformations.

It is interesting to observe that RIS Model 2, which is based upon conformational energies of disaccharides only, including solvent effects, is better than Model 3 that attempts to incorporate the longer range intramolecular energies of the entire pullulan $(G_3)_2$ oligomer. Is this simply a result of a fortuitous cancellation of errors for Model 2, or are there evident flaws within the other two model systems to account for their failure? Clearly, RIS Model 1 is

Table 3
Comparison of torsion angle locations found by CFF95 force field with Poisson–Boltzmann continuum solvation with those found using MM3 (92) by Dowd et al. [43] and AMBER* models by Liu et al. [36]

CFF95 model conformation [ϕ_6, ψ_6, ω_6]	MM3 (92) assignment[a](*gt*)	AMBER* assignment (gt)
A1 (*gg*) [64.7, −163.2, 68.2]		
A2 (*gg*) [69.4, 170.7, 66.4]	[74.5, 185.4, 66.5]	B_6 [80, 200, 60]
A3 (*gg*) [64.8, 170.2, 67.7]	[78.2, 82.0, 60.9]	A_6 [80, 80, 60]
A4 (*gt*) [63.2, −165.9, 68.2]		
A5 (*gt*) [61.8, −172.6, 66.9]		
B1 (*gt*) [92.0, −94.8, 73.0]		
B2 (*gt*) [94.0, −92.8, 71.4]	[95.5, 284.2, 73.7]	C_6 [80, 300, 80]
B3 (*gt*) [97.3, −89.7, 69.1]		
C1 (*gt*) [64.8, 175.5, 178.2]	[75.3, 179.7, 177.6]	D_6 [80, 160, 180]
D1 (*gg*) [62.2, −178.7, −66.2]	[77.5, 186.4, 294]	F_6 [80, 200, 300]
D2 (*gt*) [60.2, 179, −66.4]	[75.7, 90.9, 281.9]	E_6 [80, 100, 300]
D3 (*gt*) [78.8, −146.5, −65.2]		

[a] The torsion angles of the (1 → 6)-linkage were converted to conform to the definitions in the present paper to facilitate the comparison: ϕ_6 (O_5–C_1–O–C_6'), ψ_6 (C_1–O–C_6'–C_5'), and ω_6 (O–C_6'–C_5'–O_5').

expected to give significant errors due to lack of consideration of solvent effects in the energy function. Given that both Models 1 and 2 share the same 12 conformations for pullulan $(G_3)_2$, the use of conformations identified using a gas phase potential energy function is in itself not the source of error. However, using these gas phase energies to derive statistical weights for these conformations is a source of error. Indeed, examination of Table 1 reveals that the gas phase energies favor more compact structures while solvation energies are greatest for more extended structures. For this reason, the Kratky plot derived from RIS Model 1 corresponds to more compact structures than are observed experimentally.

The failure of Model 3 may be attributable to several factors. Geometry optimization of the full pullulan $(G_3)_2$ molecule using a gas phase potential energy function results in strong bias towards overly compact structures. This bias is considerably more pronounced for the pullulan $(G_3)_2$ case than for the disaccharides due to the large number of favorable electrostatic interactions possible between non-nearest neighbor residues. Ideally, the solvation energies should balance the gas phase energies by providing an equal bias in favor of more extended conformations. The results presented herein suggest that this is not the case, and examination of Table 2 reveals that conformations with favorable gas phase energies also have the largest statistical weights. In addition, the method used here assigns statistical weights using only the calculated minimum energies and does not take into account the breadth and possible anharmonicity of the potential energy wells.

The success of Model 2 suggests that the results of the disaccharide conformational searches are not particularly sensitive to solvation effects. If the conformations of **1** and **2** preferred in vacuum were greatly influenced by solvent effects, none of the RIS models presented here would have any predictive power. It is certainly true that relative energies for different conformations of **2** are greatly influenced by solvation, but the current results suggest that the detailed geometries of the preferred conformations are not. It is also of interest to compare the preferred conformations deduced here using CFF95 with the Poisson–Boltzmann solvation model with results obtained using other approaches. Preferred conformations for the (1 → 6)-linked dimer generated using MM3 (92) [43] in the gas phase and AMBER* with the GB/SA continuum solvent model for water [36] are listed in Table 3. Only the AMBER* states involving the *gt* conformer of exocyclic torson angle (τ_6 are shown in Table 3; note that all conformers were constrained to lie on a regular grid in torsion angle space in the work cited. The CFF95 families A–D are also selected by AMBER* and MM3, but these force fields also pinpoint two additional conformers (AMBER* A_6 and E_6) that do not appear as distinguishable minima in the CFF95 conformer search. Table 1 reveals that family D in CFF95 Model 2 (i.e. with solvation) has greater statistical weight than the other families. This agrees with the gas phase MM3 (92) calculations reported by Dowd et al [43]. The AMBER* search (with solvation), however, gives greatest weight to state B (AMBER* state C_6) in the *gt* conformation but to state A (AMBER* state B_6) in the *gg* conformation.

The method described here displays good success in fitting SAXS data for relatively small oligomeric polysaccharides. Larger oligomers and high polymeric material may present a greater challenge for the method, because small errors in describing the location and/or statistical weight of the RISs of the disaccharides can become amplified over the length of a longer chain. Thus, small errors that might overemphasize compact conformations may be difficult to detect in comparisons between theory and experiment for small oligomers. These same defects can, however, lead to serious underestimation of the dimensions of much longer chains calculated from the same model [8]. Caution should be used when extending this 'gas phase geometry + solvation' method to other polysaccharide

systems for which solvation effects may play a more significant role in determining geometry. In particular, intramolecular hydrogen bonds are typically overestimated in gas phase calculations due to strong electrostatic energies [44] and could lead to significant changes in geometry. These potential problems were not observed for the pullulan system presented in this work. In general, however, modeling of this type has not yet attained a level of reliability that allows it to be used without continual reference to appropriate experimental data.

5. Conclusion

A RIS model for pullulan is presented that is derived completely from theoretical methods and successfully predicts conformational properties for the trimeric and hexameric oligomers G_3 and $(G_3)_2$. The methods outlined here are simple, require no experimental input, and in principle may be applied to any linear polysaccharide system. However, this work only explores pullulan for which detailed aqueous properties have been reported. The argument may be made that this RIS model is too simple and that the forces due to solvation also must be taken into account. While solvation energies have been used to correct the gas phase calculations, the resulting forces due to solvation are expected to alter the gas phase structures. These structural changes are not accounted for in our model. Indeed, the development of potential energy functions and force field parameters that are consistent with a continuum solvent description is a worthy pursuit with the potential to impact a wide range of biological applications. Future work to investigate the ability of the current RIS model to predict conformational properties of higher order pullulan polymers and the effects of solvation forces on global conformation are part of ongoing efforts to understand polysaccharides in aqueous solution using realistic chain models.

References

[1] Goochee CF, Gramer MJ, Andersen DC, Bahr JB, Rasmussen JR. Biotechnology 1991;9:1347–55.
[2] Helenius A. Mol Biol Cell 1994;5:253–65.
[3] Dwek RA. Chem Rev 1996;96:683–720.
[4] Brant DA. Q Rev Biophys 1976;9:527–96.
[5] Ha SN, Giammona A, Field M, Brady JW. Carbohydr Res 1988;180:207–21.
[6] French AD, Brady JW. Computer modeling of carbohydrate molecules, vol. 430. Washington, DC: American Chemical Society, 1990.
[7] Homans SW. Biochemistry 1990;29:9110–8.
[8] Brant DA, Christ MD. In: French AD, Brady JW, editors. Realistic conformational modeling of carbohydrates, vol. 430. Washington, DC: American Chemical Society, 1990. p. 42–68.
[9] Ha S, Gao J, Tidor B, Brady JW, Karplus M. J Am Chem Soc 1991;113:1553–7.
[10] Kouwijzer MLCE, Grootenhuis PDJ. J Phys Chem 1995;99:13426–36.
[11] Woods RJ, Dwek RA, Edge CJ, Fraser-Reid B. J Phys Chem 1995;99:3832–46.
[12] Brant DA. Pure Appl Chem 1997;69:1885–92.
[13] Cramer CJ, Truhlar DG, French AD. Carbohydr Res 1997;298:1–14.
[14] Senderowitz H, Still WC. J Org Chem 1997;62:1427–38.
[15] Hwang M-J, Ni X, Waldman M, Ewig CS, Hagler AT. Biopolymers 1998;45:435–68.
[16] Simmerling C, Fox T, Kollman PA. J Am Chem Soc 1998;120:5771–82.
[17] Woods RJ. Glycoconjugate J 1998;15:209–16.
[18] Pérez S, Imberty A, Engelsen SB, Gruza J, Mazeau K, Jiminez-Barbero J, Poveda A, Espinosa J-F, van Eyck BP, Johnson G, French AD, Kouwijzer MLCE, Grootenius PDJ, Bernardi A, Raimondo L, Senderowitz H, Durier V, Vergoten G, Rasmussen K. Carbohydr Res 1998;314:141–55.
[19] Jónsdóttir S, Welsh WJ, Rasmussen K, Klein RA. New J Chem 1999:153–63.
[20] Gregurick SK, Liu JH-Y, Brant DA, Gerber RB. J Phys Chem B 1999;103:3476–88.
[21] Colson P, Jennings HJ, Smith ICP. J Am Chem Soc 1974;96:8081–7.
[22] Yuen S. Process Biochem 1974;9:7–9.
[23] LeDuy A, Choplin L, Zajic JE, Luong JHT. Pullulan. Encyclopedia of Polymer Science and Engineering. In: Mark HF, Bikales NN, Overberger CG, Menges G, editors. 2nd ed. vol. 13. New York: Wiley, 1986. p. 650–60.
[24] Tsujisaka Y, Mitsuhashi M. Pullulan. Industrial Gums: Polysaccharides and Their Derivatives. In: Whistler RL, BeMiller JN, editors. 3rd ed. San Diego: Academic Press, 1993. p. 447–60.
[25] Pérez S, Marchessault RH. Biopolymers 1979;18:2369–74.
[26] Nakatani M, Shibukawa A, Nakagawa T. J Chromatogr A 1994;672:213–8.
[27] Robyt JF. Essentials of carbohydrate chemistry. New York: Springer, 1998.
[28] Yalpani M. Polysaccharides. Amsterdam: Elsevier, 1988.
[29] Kato T, Okamoto T, Tokuya T, Takahashi A. Biopolymers 1982;21:1623–33.
[30] Kato T, Katsuki T, Takahashi A. Macromolecules 1984;17:1726–30.
[31] Buliga GS, Brant DA. Int J Biol Macromol 1987;9:71–76.
[32] Nordmeier E. J Phys Chem 1993;97:5770–85.
[33] Pavlov G, Korneeva EV, Yevlampieva NP. Int J Biol Macromol 1994;16:318–23.
[34] Buliga GS, Brant DA. Int J Biol Macromol 1987;9:77–86.
[35] Nishinari K, Kohyama K, Williams PA, Phillips GO, Burchard W, Ogino K. Macromolecules 1991;24:5590–3.
[36] Liu JH-Y, Brant DA, Kitamura S, Kajiwara K, Mimura M. Macromolecules 1999;32:8611–20.
[37] Dinur U, Hagler AT. J Comput Chem 1990;11:1234–46.
[38] Maple JR, Dinur U, Hagler AT. Proc Natl Acad Sci USA 1988;85:5350–4.
[39] Maple JR, Hwang MJ, Stockfisch TP, Dinur U, Waldman M, Ewig CS, Hagler AT. J Comput Chem 1994;15:162–82.
[40] Ringnalda MN, Langlois J-M, Greeley BH, Murphy RB, Russo TV, Cortis C, Muller RP, Marten B, Donnelly RE, Mainz DT, Wright JR, Pollard WT, Cao Y, Won Y, Miller GH, Goddard WA III, Friesner RA. JAGUAR 3.0. Schrödinger Inc.: Portland, OR, 1997.
[41] Honig B, Nicholls A. Science 1995;268:1144–9.
[42] Tannor DJ, Marten B, Murphy R, Friesner RA, Sitkoff D, Nicholls A, Ringnalda M, Goddard III WA, Honig B. J Am Chem Soc 1994;116:11875–82.
[43] Dowd MK, Reilly PJ, French AD. Biopolymers 1994;34:625–38.
[44] Shan SO, Herschlag D. Proc Natl Acad Sci USA 1996;93:14474–9.