Carbohydrate Research 496 (2020) 108052

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

The conformation of the idopyranose ring revisited: How subtle O-substituent induced changes can be deduced from vicinal ^{1}H-NMR coupling constants

Cornelis A.G. Haasnoot[a,*], René de Gelder[b], Huub Kooijman[c], Edwin R. Kellenbach[d]

[a] N.V. Organon, P.O Box 20, 5340 BH, Oss, the Netherlands
[b] Radboud University, Institute for Molecules and Materials, Solid State Chemistry, Heyendaalseweg 135, 6525 AJ, Nijmegen, the Netherlands
[c] Bijvoet Center for Biomolecular Research, Crystal and Structural Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH, Utrecht, the Netherlands
[d] Aspen Oss B.V, Kloosterstraat 6, 5349 AB, Oss, the Netherlands

ARTICLE INFO

Keywords:
Six-membered ring puckering
Conformational equilibrium
X-ray data
Glycosaminoglycan
Heparin constituents
Sugar code

ABSTRACT

The idopyranose ring plays a pivotal role in the conformational, dynamical, and intermolecular binding aspects of glycosaminoglycans like heparin and dermatan sulfate and it was early on assigned a role in the Sugar Code governing biological recognition processes. There is consensus that next to the two canonical $^{1}C_{4}$ and $^{4}C_{1}$ chair conformations, the conformational space accessible to the idopyranose ring entails a $^{2}S_{O}$ skew-boat conformation, but the equilibrium between these three ring puckers has evaded satisfactory quantification. In this study a meta-analysis of X-ray solid-state data and vicinal NMR coupling constants is presented, based on the Truncated Fourier Puckering (TFP) formalism and the generalized Karplus (CAGPLUS) equation. This approach yields a model-free, granular and consistent reckoning of 159 idopyranose solution puckering equilibria studied by NMR and allows us to reproduce the involved 636 NMR vicinal couplings with an overall residual RMS(J_{obs}-J_{calc}) of 0.184 Hz. Our analyses show that for all ring systems examined, the idopyranosyl chair conformations take up the same ring pucker irrespective of the ring substituent pattern or a vast variety in experimental conditions. Instead, it is the (skew-)boat conformation that adapts to the substitution pattern of the idopyranose ring or a specific sulfation pattern of neighboring saccharides. All idopyranose rings are involved in conformational equilibria that subsume the aforementioned conformers which turn out to differ only a few kJ/mole in conformational energy. Thus, the plasticity and flexibility of idopyranose remains intact under practically all circumstances and, as the glycosidic linkages in heparin are considered to be relatively stiff, the iduronic moiety functions as the linchpin of heparin flexibility thereby being rather a “space(r)” than a “letter” in the alleged Sugar Code alphabet.

1. Introduction

Glycosaminoglycans (GAGs) are a group of structurally related polysaccharides [1,2] that play an important role in a wide range of physiological processes such as maintaining the structural integrity of connective tissues, regulation of protein secretion and gene expression, anticoagulation, but also in the pathophysiology of diabetes, atherosclerosis, *etc.* [3]. Their function in animals appears well conserved as is evidenced by their evolutionary lineage that traces back to quite primitive invertebrates [4].

GAGs are long, polydisperse unbranched polysaccharides, which are classified by the (major) disaccharide repeat unit that characterize their carbohydrate backbones. The disaccharide repeats typically consist of a uronic acid (α-L-iduronic acid [IdoA] and/or β-D-glucuronic acid [GlcA]) coupled to a hexosamine (β-D-glucosamine [GlcN] or β-D-galactosamine [GalN]). These basic, repetitive chains, however, acquire a considerable degree of structural variability by the regulated expression and action of multiple biosynthetic enzymes which modify the GAG backbone by means of sulfation, acetylation and epimerization. The microstructural heterogeneity thus created represents a high-capacity biological information storing and coding capacity for modulating biological functions that is referred to as the Sugar Code [5] (or sometimes more specifically dubbed Sulfation Code [6,7]) and on occasion even popularly addressed as the third alphabet of life

* Corresponding author.
E-mail addresses: cagh@planet.nl (C.A.G. Haasnoot), r.degelder@science.ru.nl (R. de Gelder), huub.kooijman@shell.com (H. Kooijman), ekellenbach@nl.aspenpharma.com (E.R. Kellenbach).

https://doi.org/10.1016/j.carres.2020.108052
Received 5 March 2020; Received in revised form 17 May 2020; Accepted 27 May 2020
Available online 16 July 2020

(nucleotides and amino acids forming the first two) [8].

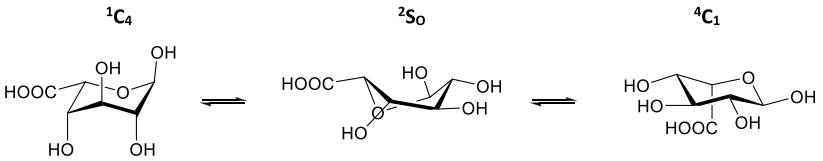
The actual mode of action is suggested [9] to be a combination of overall conformation, charge distribution and flexibility reflected by a GAG or a part of it, thereby determining the extent of binding, or activity with respect to its receptor. The functional diversity thus obtained allows GAGs to fulfill a variety of biological roles which translate into triggering specific biochemical cascades, assembling protein-protein complexes, cell adhesion, but also as entry points for pathogens, *etc.* [10].

Heparin/Heparan sulfate (HS) and dermatan sulfate (DS) distinguish themselves from the other GAG families by the presence of L-iduronic acid (IdoA). This particular uronic acid monomer is presumed to boost the flexibility of the saccharide chain (*vide infra*), thereby facilitating protein binding [11]. Moreover, the IdoA residue (or rather its depletion) has been linked to the regulation of biological processes that play a role in a wide range of pathologies, e.g. immunity [11], wound healing [11], cancer [12], developmental defects [13] – to name a few.

Historically, most GAGs were isolated and chemically identified before a biological function other than filling extracellular space could be assigned to them [14]. The discovery of heparin [15] was an exception in that the anticoagulation activity was unearthed [16] first, whereafter it still took almost half a century to establish the molecule as a highly sulfated, hexosamine containing polysaccharide with the last piece of the puzzle falling into place by the identification [17] of L-iduronic acid as its major uronic acid constituent.

In the meantime, heparin had found widespread application as a clinical anticoagulant [18] and in the 1970s structure-activity quests were started to understand and ultimately to improve on its effectivity. Thus, the all-important activation of antithrombin AT-III was nailed down to an essential IdoAS-containing pentasaccharide sequence [19] in the heparin polymer potentiating the anticoagulant action. This finding ultimately led to the development of the synthetic fondaparinux pentasaccharide, registered as an antithrombotic drug under the name Arixtra [20].

However, the conformation of the IdoA residue in GAGs remained controversial as ^{1}H NMR studies of the idopyranosyl rings in *e.g.* heparin [21] and dermatan sulfate [22] could not be interpreted in terms of the canonical $^{1}C_{4}$ and $^{4}C_{1}$ chair conformations common for carbohydrate rings. Casu et al. [23] settled the dissension by proposing the IdoA pyranose rings to participate in a conformational equilibrium that involves a third, $^{2}S_{O}$ skew conformer in addition to the two canonical chairs (*cf.* **1**). This proposal was furthered [24] in a 3 steps approach: (1) modelling the three constituent idopyranose conformations using Molecular Mechanics (MM), (2) calculating the endocyclic vicinal couplings from the models by means of the generalized [25] Karplus equation (CAGPLUS), and (3) extracting the conformer populations quantitatively from the experimental coupling constants. NMR data processed in this way for a number of IdoA[S] rings were taken to show that the three-state conformational equilibrium depends both upon its own substitution with 2-O-sulfate and on substitution of adjacent glucosamine residues. This then led to the concept [26] of the IdoA moiety increasing the conformational flexibility of heparin/HS, the biological relevance of which has long since been subscribed to [9,27,28].



1

*α-**L**-Idopyranosiduronic acid*

Although the idea of a three-state conformational equilibrium for IdoA was quickly adopted, there also emerged questions about the assumed constituent ring conformations which in turn led to receding to qualitative interpretations [29] or *ad hoc* conformational adaptations [30] in the interpretation of the observed experimental coupling constants. A plethora of computational chemistry studies to improve the underlying constituent ring conformations and coupling constants followed but although ever increasing levels of sophistication (MM, MD, DFT) were deployed, the match between calculated and experimental results (couplings, conformer populations) seemed to worsen rather than to improve.

Without disparaging the merits of these studies, they do suffer from a methodological weakness, *i.e.* experimental coupling constants are interpreted on the basis of calculated couplings derived from computed structures. As such, the underlying arithmetic procedure is model-induced as it heavily relies on the (semi-)empirical force fields used in these studies to forge the constituent idopyranose conformers. Unfortunately, the reliability of force fields used in MM/MD studies is notoriously questionable as was made clear by the comparative study of 20 force fields by Pérez et al. [31] from which the authors concluded rather euphemistically: "…a series of seven test cases provides a fairly nonuniform picture of the potentiality of these parameter sets for giving a consistent image of structure and energy of carbohydrate molecules." For MD calculations this is even worse as these will also be induced by the calculated energies for which the Pérez study [31] concluded that "…the consideration of these relative energies in further calculations is a highly risky exercise." And although since publication some time has passed, the situation has seemingly not improved much [32].

Perhaps somewhat less obvious is that most DFT calculations suffer from a related weakness. The DFT method itself is superior to the empirical MM/MD approach, but also DFT calculations tend to stick "close" to their starting input structures and the optimized structures therefore represent mostly local minima [33]. As the DFT starting structures are often generated by empirical MM methods, the DFT outcomes are bound to be heavily biased [32] by the classical force fields used. As shown by Kurihara and Ueda [34] this flaw may be undercut by applying a systematic conformational search, but so far their approach has not acquired many followers.

As an alternative to these force field induced approaches we explore in the present paper the application of a model-free approach, *i.e.* conformational analyses based on experimental solid-state X-ray structures and ^{1}H NMR coupling constant data only. We will describe a new application of the previously published [35] combination of the generalized [25] Karplus equation and the Truncated Fourier Puckering (TFP) formalism [36] that allows for the analysis of three-state idopyranose conformer equilibria in solution. In this setting the TFP formalism plays a crucial role as it prescribes that the endocyclic idopyranose torsions cannot adopt arbitrary values but instead are tightly coupled to one another. As such, every idopyranosyl ring conformation is characterized by a unique set of endocyclic torsions which by means of the generalized Karplus equation can be tied to a specific set of endocyclic vicinal coupling constants. Thus, it is the ensemble of the four idopyranose ring proton-proton couplings that are in unison bound to the same conformational element, *i.e.* the idopyranose ring pucker. As will be explained in Methods (*vide infra*), this creates an overdetermined system of equations from which conformational equilibria can be assessed, on the provision that a surplus of coupling constant

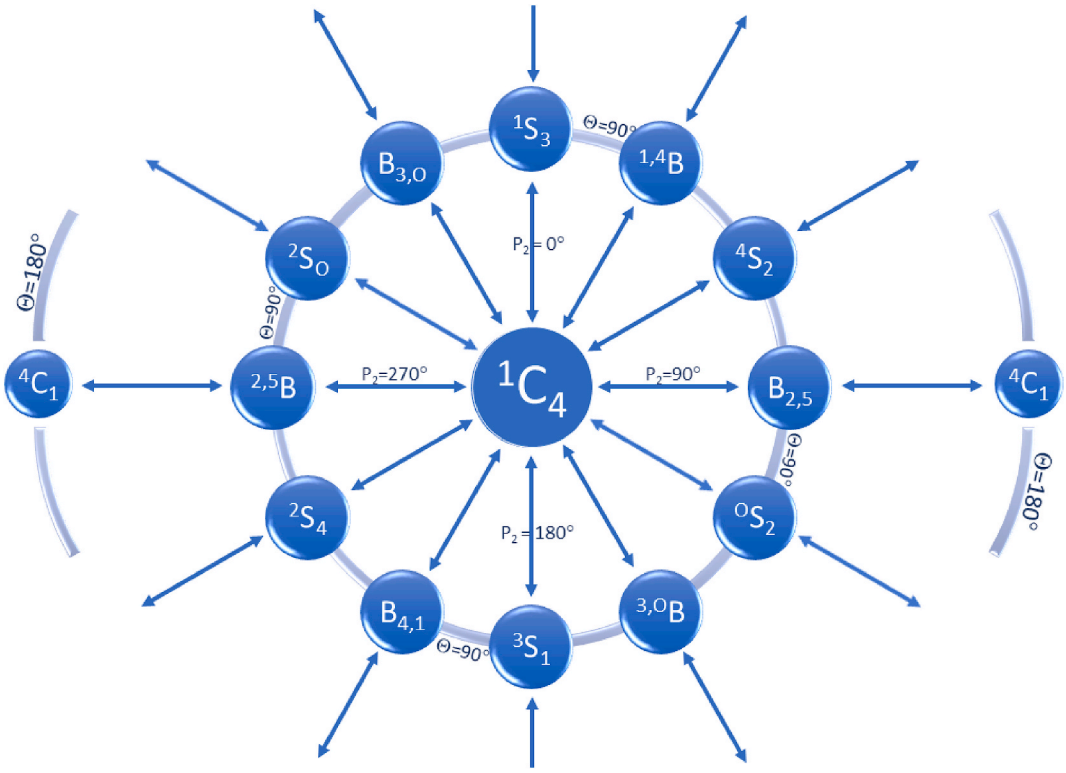
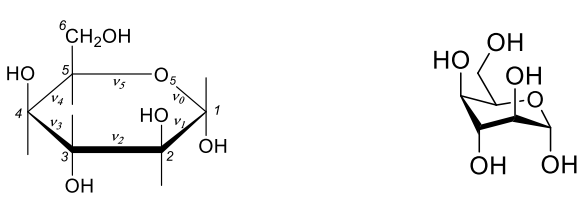


Fig. 1. Polar projection of the sphere depicting TFP-conformational space accessible to L-saccharide pyranose rings. The center position represents the $^1C_4$ chair at the "North pole" (Θ = 0°), whereas the outer position (Θ = 180°) is taken by the inverted $^4C_1$ chair. The circle in the middle maps the pseudorotational subspace at Θ = 90° that corresponds to the Boat/Skew itinerary.

ensembles is available. Fortunately, a cornucopia of idose related ^{1}H NMR studies have accumulated in the literature over the years and from these studies we were able to extract 159 complete ensembles of 4 vicinal endocyclic coupling constants per iduronic residue. It will be shown that this large body of experimental NMR couplings allows us to explore the intricate conformational behavior of the idopyranosyl ring at a granularity that allows us to determine whether the ring conformation is indeed coupled to the substituents pattern on and/or around the iduronic acid moiety and by such means is linked to the "Sulfation Code" that is supposed [37,38] to be a specific means whereby GAGs exert their regulatory role.

2. Methods



2
*α-**D**-Idopyranose*
Left: Haworth projection with atom and torsion numbering
Right: classical $^4C_1$ chair conformation

2.1. Solid-state data

Single-crystal X-ray structures of idopyranose (**2**, Ido) and idopyranosyluronic acid (**1**, IdoA) derivatives were extracted from the Cambridge Structural Database (CSD) [39]. Searches were performed on the basis of substructures as well as partial compound names (e.g. "idopyr*") which generated some 50-odd hits. These hits were all visually inspected for the presence of an α-type saccharide anomeric center (i.e. the C6–C5–O5–C1–O1 fragment in an α configuration) as well as the required stereoisomeric substitution pattern; both L- and D-configurations were encountered. Substituents other than oxygen were allowed for the non-anomeric carbons, however, bicyclic compounds were discarded as we are primarily interested in the conformational behavior of idopyranosyl rings that are not conformationally restricted by covalent links between ring atoms. These selection criteria left us with 14 compounds for which 18 solid state structures could be obtained.

2.2. Pyranose ring puckering

The conformation of a six-membered ring is characterized by the endocyclic torsion angles ν_j (j = 0–5) which are interrelated *via* the Truncated Fourier Puckering (TFP) formalism [36]:

$$\gamma_j = \Phi_2 \cos(P_2 + 4\pi j/6) + \Phi_3 \cos(\pi j) \quad (1)$$

This defines a given six-membered ring conformation in terms of its three puckering parameters, *viz.* Φ_2, P_2 and Φ_3 which can be deduced [36] from the six endocyclic torsions by Fourier inversion. The puckering parameters may be rewritten [36] as a set of spherical polar coordinates Q, Θ and P_2, where Q is the total puckering amplitude

$$Q = \sqrt{\Phi_2^2 + \Phi_3^2} \quad (2)$$

and

$$\Theta = \arctan(\Phi_2/\Phi_3) \quad (3)$$

with $0 \le \Theta \le \pi$.

Thus, conformational space accessible to the six-membered saccharide ring may be envisaged as the surface of a (distorted [40]) sphere [41,42] onto which the individual conformations of the

pyranose ring can be mapped in terms of its three ring puckering coordinates (radius Q, inclination Θ and azimuthal P_2).

For a unique description of the six-membered ring it is necessary to define the endocyclic torsion angle numbering scheme. In this paper we will adhere to the IUPAC Nomenclature of Carbohydrates (Recommendations 1981 [43] and Recommendations 1996 [44]) which assigns the anomeric carbon the locant 1 (i.e. C1) and hence the origin torsion angle ν_0 is formed by C5–O5–C1–C2 with subsequent torsions to follow the standard numbering, i.e. ν_1 is the endocyclic torsion about the bond between C_1 and C_2; and so forth (see *e.g.* **2**).

Following this numbering scheme, the $^1C_4$ conformation for L-monosaccharides is found at the "North pole" of the sphere (Θ = 0°) whereas the $^4C_1$ conformation occupies the "South pole" (Θ = 180°). The "equator" of the sphere (Θ = 90°) denotes the pseudorotational subspace on which the canonical B(oat) and S(kew) conformations are found. The phase angle P_2 designates the conformation at play, e.g. P_2 = 300° represents the ^2S_O conformation (*cf.* Fig. 1).

However, pyranose conformations will rarely occur in one of the abovementioned "pure" canonical forms. Instead, they will adopt (slightly) distorted ones that may be viewed as linear combinations of the basis chair, boat and/or skew forms. The TFP-formalism eschews pictorial but in essence qualitative descriptions [45] of such "intermediate" conformations and instead allows for a quantitative description of the ring conformation. And notably, because the TFP-parameters are directly related to (internal) torsions they are consistent with angular characteristics of the six-membered ring [36]: a particular linear combination quantifies the "basic" conformations involved and to what extent each of them shapes the ensuing six-membered ring in terms of local flattening, etc.

While quantitative, the TFP parameters are not suited to invoke a mental image of the conformation at play: the three parameter values define a precise position on the "Q-sphere", but, commonly, Fig. 1 is needed to "translate" these values into a tangible conformation. Of course, this limits such envisagements to easily discernible variants like the canonical boat (B), chair (C) and skew (S) conformers [44], and therefore we will use in this paper the descriptors shown in Fig. 1 in a "loose" way, i.e. describing a range or family of structures similar to the corresponding canonical structure. For example, $^4C_1$ is defined as the classical chair (Φ_3 = Q, Φ_2 = 0 and P_2 = irrelevant), but will in this paper denote all structures with Θ larger than, say, 165°; in other words: not strictly the south pole, but rather the polar cap. Similarly, the descriptor ^2S_O should here not be taken to designate the single conformer described by Φ_3 = 0, Φ_2 = Q and P_2 = 300°, but rather the family of conformations ranging Φ_3 = 0 ± 0.15Q, Φ_2 = Q and P_2 = 300° ± 15°, with the canonical ^2S_O as its eponym. Thus, in this paper the ontology of B, C and S descriptors are to be taken as rather broad classifications of the conformations at play, whereas particular single conformations will be specified in terms of their three TFP parameters.

Finally, we note in passing that for the enantiomeric D-saccharides the locations of the ring conformations on the sphere are mirrored, *i.e.* $^4C_1$ is now "North" (Θ = 0°) and $^1C_4$ is "South" (Θ = 180°); moreover, the meridian P_2 is also mirrored, *i.e.* shifted by 180°.

2.3. The joint TFP/CAGPLUS approach

An outstanding quality of the TFP formalism is that it is based on internal coordinates and therefore does not rely on a mean [42] or some other best-fit [45] external plane to describe the puckering of the six-membered ring. As such it does not require the upfront availability of the Cartesian coordinates of *all* six ring atoms which in practice can only be delivered by X-ray or computer modelling techniques. Contrastingly, by foregoing the need of a reference plane the TFP formalism is geared to dealing with six-membered rings in solution, all the more so because the TFP-formalism can also work with partial datasets as eq. (1) can be solved [35] when minimally three of the torsions ν_j are known.

These characteristics have been shown [35] to open up the opportunity to derive the TFP coordinates (Q, Θ, P_2) from the endocyclic vicinal NMR proton-proton coupling constants measured for a given six-membered ring. To this end ring torsions ν_j are related to their corresponding proton-proton torsion angle (ν_{HH}) by using trigonal projection symmetry

$$\nu_j = a + \nu_{HH} \tag{4}$$

where *a* will be 0° for *cis* protons and ± 120° for *trans* protons. Combining eqs. (1) and (4) with the generalized [25] Karplus equation (CAGPLUS) will then allow the vicinal proton-proton coupling constants from the six-membered ring to be expressed as a function of the TFP coordinates

$$J_{HH} = f(\Phi_2, P_2, \Phi_3) \tag{5}$$

This function can be put to work in two ways: (a) calculate endocyclic H–H coupling constants for a given (Φ_2, P_2, Φ_3) ring conformation which are to be compared with observed couplings; or, (b) calculate the ring puckering coordinates directly from the observed coupling constants using an iterative least-squares procedure (*cf.* Procedure I in Ref. [35]). Both approaches will be deployed in the following analyses.

Another, slightly complicating factor in the conformational analysis of medium-sized ring systems is that the ring might be involved in fast multi-state conformational equilibria. As such pyranose interconversions are rapid (estimated [46] to be in the single-digit μs range) on the NMR time-scale, the observed coupling constants are in that case time-averaged so that they reflect the couplings of the underlying conformers (J_i) and their relative populations (mole fractions ρ_i), *i.e.*

$$\langle J_{HH}\rangle_{exp} = \sum \rho_i J_i \tag{6}$$

The TFP/CAGPLUS approach and computational procedures for analyzing two-state conformational equilibria have been described *in extenso* (*cf.* Procedures II and III in Ref. [35]); here we will pursue the methodology by detailing the approach for a three-state equilibrium involving conformations C′, C″ and C‴:

$$C' \overset{K_1}{\rightleftharpoons} C'' \overset{K_2}{\rightleftharpoons} C'''$$

In this case each of the experimentally observed coupling constants will be time-averaged over its three constituent couplings partitioned according to the two equilibrium constants K_1 and K_2. Representing the six-membered ring conformations C′-C‴ in terms of their TFP coordinates (eqs. (1)–(3)) and combining these with eqs. (5) and (6) we get:

$$\langle J_{HH}\rangle_{av} = \rho' f(\Phi_2', P_2', \Phi_3') + \rho'' f(\Phi_2'', P_2'', \Phi_3'') + \rho''' f(\Phi_2''', P_2''', \Phi_3''') \tag{7}$$

It follows that each of the experimental endocyclic vicinal coupling constants is a function of 11 parameters: 3*3 TFP coordinates and 2 equilibrium constants, which plainly far exceeds the number of pyranose endocyclic couplings that can be extracted from a single NMR experiment. For instance, an idopyranosyl derivative will only give rise to maximally four endocyclic couplings, which limits the number of conformational parameters that can be extracted from them drastically. However, we can improve on this limitation if for the given derivative the two equilibrium constants can be influenced by changes in measurement conditions, *e.g.* temperature, pH, solvent, *etc.* In that case the ratio between observables *vs.* conformational parameters can be improved as every NMR measurement under such a changed condition will bring on board four extra ring couplings, albeit at the cost of two

new parameters K_1' and K_2'. As such, the number of observables will exceed the number of parameters to be determined when five (but preferably more) independent NMR measurements on the same three-state conformational equilibrium are available so that the overall set of equations will no longer be underdetermined and can – at least in theory – be solved by iterative least-squares minimization techniques [47]. We implemented such least-squares techniques by adapting and expanding the Turbo Pascal programs previously developed [35] for analyzing two-state equilibria.

There are a few caveats to be borne in mind when applying this iterative technique though. First of all, the values of the observed coupling constants must span a substantial range as a function of the induced changes in equilibrium constants; if not, the minimization will be performed on what is effectively a single set of couplings in which case it is doomed to fail. Another, more calculational type of pitfall may be encountered in the Newton-Raphson minimization procedure which involves a large set of partial differential equations leading to a sparse matrix that may breakdown during inversion. And, as was observed [35] earlier for the two-state equilibrium case, the minimization is rather easily trapped in some local minimum.

These impediments led us to look into grid (or *brute force*) searching as a supplementary technique for deriving the best TFP parameters and equilibrium constants from experimental coupling constants. Conceptually, our grid search starts at the back end of the problem, *i.e.* assigning values to the nine TFP coordinates in eq. (7), use these to calculate endocyclic proton-proton torsion angles for the three conformations by combining eqs (1) and (4), and calculate the endocyclic coupling constants ($J_{k,k+1}$) for the three conformers by applying the CAGPLUS equation.

As the four experimental idopyranosyl ring couplings are time-averaged over the three constituent conformers, the following set of equations emerge

$$\langle J_{k,k+1}\rangle_{exp} = \rho' J'_{k,k+1} + \rho'' J''_{k,k+1} + \rho''' J'''_{k,k+1} ...(k = 1...4) \quad (8)$$

with

$$\rho' + \rho'' + \rho''' = 1 \quad (9)$$

Thus, we arrive at an overdetermined linear system of 4 equations from which two independent ρ's may be determined [47] by calculating the least squares solution for the system by solving the matrix equation

$$\mathbf{A^T A}\,\underline{\rho} = \mathbf{A^T}\,\underline{\mathbf{J_{exp}}} \quad (10)$$

This is a very efficient way of calculating the three unknown ρ's and by determining the rms-deviation between the matching set of $< J_{i,i+1} >_{calc}$ and $< J_{i,i+1} >_{exp}$ a performance metric is obtained for the start set of TFP parameters. By systematically varying the 9 TFP input parameters for eq (7) we can perform an exhaustive grid search over a specific (sub)set of conformational space in order to get at the best performance metric (*i.e.* lowest rms-difference).

Although very straightforward, the grid search technique also has its idiosyncrasies. Replacing the two equilibrium constants by three (interdependent) mole fractions opens up the possibility that in some cases a negative value is calculated for one of the mole fractions. Although such a false solution might be prevented by applying constraints, doing so would considerably complicate the algorithm; here we elect to simply discard such solutions as fixing it will make its performance metric go up (and hence discard it for that reason). Another weakness is that grid searches inflate exponentially with dimensionality and so a full-fledged 9 variables search is prone to create a sizeable computer workload (albeit an embarrassingly parallel one).

Both the iterative and the grid search technique are most decisive when the number of TFP coordinates to be extracted from the set of experimental couplings is reduced, *e.g.* by fixing some of them to the values of a reference conformation such as an X-ray or Molecular Mechanics derived structure.

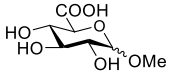
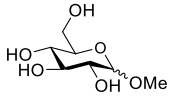
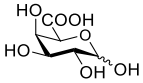
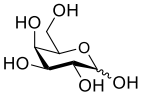
Finally, at the end of the analysis it is to be verified that all conformational data indeed concern one identical three-state equilibrium, *i.e.* no inexplicable outliers or drifts in the calculated parameters that might undercut the assumed invariancy of the constituent conformers. For an example we refer to the note added to the figure in Table S5 (*cf.* Supporting Information) where we outline such a crosscheck.

As a corollary we note that although the conformations found by the TFP/CAGPLUS approach are represented by TFP coordinate sets that seem to pin the rings to single points on the sphere describing conformational space (Fig. 1), it is taken for granted that in reality such TFP coordinate sets denote the center of essentially symmetric low-energy wells [48] in the conformational landscape in which the six-membered rings will enjoy at least some pseudolibrational freedom. It has been argued [35] that the torsional motions that arise from such allowed pseudolibrational motions will give rise to time-averaged coupling constant values that correspond to the single point TFP-minimum conformation. This assertion was independently corroborated by Förster and Mulloy [49] who showed that the averaged CAGPLUS couplings calculated for the pseudorotational ensemble found in their MD study for the iduronate ^2S_O conformer were "reasonably close" to the couplings calculated for the energy-minimized ^2S_O structure.

2.4. NMR coupling constants

Experimental vicinal NMR coupling constants for the four endocyclic H–H pairs in Ido and IdoA derivatives were gathered from the

Table 1
Influence of -COOH vs. -CH_2OH on J_{H4H5} in Glucopyranose and Galactopyranose derivatives.

 **3** *methyl-* ***D****-gluco-pyranosiduronic acid*	 **4** *methyl-* ***D****-glucopyranose*	 **5** ***D****-galacturonic acid*	 **6** ***D****-galactopyranose*
J_{H4H5} (Hz)	J_{H4H5} (Hz)	J_{H4H5} (Hz)	J_{H4H5} (Hz)
β-OMe 9.6 [51]	β-OMe 9.7 [53]	α-OH[b] 1.27 [54]	α-OH 1.23 [55]
α-OMe[a] 10.0 [52]	α-OMe[a] 10.0 [52]	β-OH[b] 1.28 [54]	β-OH 1.1 [55]
β-OMe[a] 9.9 [52]	β-OMe[a] 10.0 [52]		

[a] pD = 2.0.
[b] pH = 4.0.

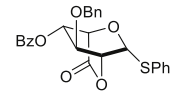
literature. Only complete sets, *i.e.* with four endocyclic couplings for the molecule at hand, were accepted and in case of small variations in the ^{3}J's in a series of measurements of the same molecule reported in one publication, only median values were added to our dataset.

Generally speaking, the quality of the obtained NMR data is good to very good: spectra were recorded at high fields (220–950 MHz) and a considerable number of the found coupling constant sets were derived from simulated spectra. Simulation is indeed desirable, especially in the case of small couplings (which are easily underestimated [50] in case of line-widths that are comparable to the coupling constant value), strong couplings and/or spectra of oligosaccharides with overlapping spin systems. As line-broadening and overlapping of resonances strongly correlate with molecule size it is not surprising that the accuracy of the reported coupling constant data appears to lessen when going from monosaccharides to oligosaccharides.

Vicinal coupling constants are interpreted in terms of the corresponding H–H torsion angles by means of the generalized Karplus equation (CAGPLUS) [25] which takes into account the electronegativities of the substituents along the C–C bond in question. However, it was noted by Sanderson et al. [30] that the electronegativity value to be used for the carboxylate group in the CAGPLUS equation is ill-defined. In order to settle this issue we compared the $^3J_{H4H5}$ of gluco- and galactopyranosides with their uronic acid derivatives (*cf.* Table 1). The coupling constant data from the table show that both in the *anti* as the *syn* orientation of the coupled protons $^3J_{H4H5}$ is practically invariant to changing C6 from a carboxyl into a hydroxymethyl group. As it was also shown [52] that $^3J_{H4H5}$ in α-/β-glucuronic acid is unaffected by ionization of the COOH group, we conclude that in terms of the CAGPLUS equation [25] the electronegativity of a COOH group can be taken equal to that of a CH_2OH group (in the present case: $\Delta\chi = 0.152$).

2.5. Applicability and accuracy of the TFP/CAGPLUS approach

As the reliability of the generalized Karplus equation applied to (charged) carbohydrates has been questioned [56–59] and even characterized [55] as "self-defeating", we put the CAGPLUS equation with the above derived $\Delta\chi$ for the carboxyl group to the test by analyzing the coupling constant data reported [60] for 4-O-benzoyl-3-O-benzyl-1-thiophenyl-β-L-idopyranoside-Urono-2,6-lactone (**7**). To this end we deployed the CAGPLUS equation in combination with the TFP formalism (*vide supra*), an approach that was earlier shown [35] to provide a superior analysis of six-membered ring conformations.



7

4-O-Benzoyl-3-O-benzyl-1-thiophenyl-β-L-idopyranoside-Urono-2,6-lactone

The idopyranosyl ring in **7** is part of a [2.2.2] lactone system which effectively locks the sugar ring in a $B_{2,5}$ conformation, as is evidenced by the X-ray structure determination obtained [60] for this compound (*cf.* Table 2). The X-ray structure was subjected to an energy minimization by means of molecular mechanics using the MMFF94 force field [61]. The solid state and the MM optimized structures show very similar pyranose ring conformations in terms of their TFP parameters (*cf.* Table 2) implying that crystal packing forces have minimal impact, albeit that the small difference in P2 (94.4° *vs.* 86.3°) might be taken to indicate that some pseudolibrational freedom of movement is still available to this highly restrained molecule.

The TFP/CAGPLUS analysis (Procedure I [35]) of the experimental [60] NMR coupling constants of **7** yields FTP coordinates which are in excellent agreement (roughly within 1σ, *cf.* Table 2) with their solid state and MM counterparts.

Apart from the finding that the solution conformation of this rigid pyranose ring as extracted from the NMR couplings is in striking conformity with the solid state and MM six-membered rings, it may also be inferred from these calculations that the methodology and hence the CAGPLUS is very well capable of translating vicinal coupling constants into torsion angles from which sensible six-membered ring conformations may be deduced. Moreover, the present analysis answers another question raised by Sanderson et al. [30] as it is shown that the TFP/CAGPLUS approach works well for Boat-like conformations and an overall rms difference between 4 observed vs. calculated coupling constants of 0.21 Hz is indeed attainable.

Table 2
Conformation analysis of 4-O-benzoyl-3-O-benzyl-1-thiophenyl-β-L-idopyranoside-urono-2,6-lactone, **7**.

TFP parameters	X-ray (IGUYEQ[a,b])	MMFF94[c]	NMR[d] (st. dev.)
Q	70.3°	70.1°	68.8° (2.9°)
Θ	92.2°	92.6°	88.4° (5.4°)
P_2	94.4°	86.3°	90.5° (3.9°)
Couplings (Hz)	J_{obs}[b]	J_{calc}[d]	Δ J
H1–H2	1.0	0.8	0.2
H2–H3	4.4	4.2	0.2
H3–H4	1.8	1.5	0.3
H4–H5	4.4	4.3	0.1
Rms dev			0.21

[a] CSD refcode.
[b] Ref. [60].
[c] Ref. [61].
[d] Results obtained from coupling constant analysis, *cf.* Procedure I ex ref. [35].

3. Results and discussion

3.1. Conformations of the idopyranosyl ring in the solid-state

Our search of the CSD [39], as described in Methods (*vide supra*), yielded 18 X-ray α-idose structures which could be broken down into 8 α-idopyranosyl (α-Ido, **2**) and 10 α-idopyranosyluronic acid (α-IdoA, **1**) derivatives, *cf.* Table 3. All IdoA as well as half of the Ido derivatives possess the L-configuration, but four Ido derivatives have a D-configuration. In order to allow for a useful comparison we made use of the enantiomeric qualities of these structures and mirrored the D specimens so that conformation-wise they fall in line with the L-structures. We could thus calculate commensurable TFP coordinates for the structures and classify them according to their ring conformation ($^1C_4$, ^2S_O, $^4C_1$).

The TFP coordinates for each conformer class show a very small spread, *i.e.* a standard deviation of just a few degrees for each of the parameters. Notably, the Ido and IdoA structures display the same conformational behavior as is evinced from the finding that where both are available the average TFP parameters for both groups of structures are within each other's standard deviation. From these findings we conclude that as far as its solid-state conformation is concerned, the six-membered pyranose ring seems indifferent to the oxidation state of C6.

Inspection of the entries in Table 3 shows that the $^1C_4$ conformation by far outnumbers the other idopyranosyl X-ray structures, especially when it is kept in mind that the four SAHZOS structures are to some extent duplicates [62]. The TFP coordinates calculated for the $^1C_4$ chairs all fall in a narrow range, both within and between each derivative class (Ido *vs.* IdoA). As such, we take these observations to

Table 3
TFP parameters for X-ray structures of idopyranose fragments.

	α-L-idopyranosyluronic acid						α-L-idopyranosyl					Average
CSD id	IGUYAM	JEXRIN (1)	JEXRIN (2)	KUTZIJ	RAKFEP	WEKNAE	AWOMII	TUDWAS	TUDWEW	PAIDOP[a]	VERTER[a]	$^1C_4$
Q	54.2	52.5	52.5	49.4	51.7	53.7	53.0	54.1	53.9	50.4	54.4	52.7
Θ	4.1	11.5	11.0	7.8	12.6	14.7	13.3	17.7	17.2	10.0	4.9	11.4
P_2	289.8	304.4	300.0	293.6	298.4	296.0	300.9	307.5	299.0	303.1	302.1	299.5
CSD id	SAHZOS	SAHZOS01	SAHZOS02	SAHZOS03			HOXZUQ[a]					^2S_O
Q	64.0	64.7	64.0	65.0			68.1					65.2
Θ	80.7	82.5	85.1	82.9			86.0					83.5
P_2	307.8	296.3	296.6	295.6			305.8					300.4
CSD id							BIKWOH10	MABIDP[a]				$^4C_1$
Q							59.7	57.8				58.7
Θ							174.6	173.1				173.8
P_2							226.1	155.2				190.7

[a] Originally reported as the D-enantiomer, for comparison purposes in this table mirrored into the L-configuration.

indicate that a single well-defined local minimum in the potential energy profile of the idopyranosyl ring delineates the $^1C_4$ conformation of both Ido and IdoA derivatives. It is noted that this contrasts starkly with the puckering or energy landscapes computed [63–65] by Molecular Dynamics techniques for idose derivatives in which the vast spread of recorded $^1C_4$ conformations might be taken to suggest that this chair conformer can adopt a very broad range of distorted chair-like conformations for an appreciable part of the time.

Table 3 lists a ^2S_O conformation for Ido as well as IdoA – the two pyranose rings are practically undistinguishable, again indicating their close conformational relationship. Unfortunately, for the $^4C_1$ conformation no such a common ground between Ido and IdoA can be established as for this particular conformation only two Ido derivatives have been reported.

Taken together, the apparent interchangeability of Ido and IdoA structures in terms of their six-membered ring conformations allows us to propose the averaged TFP coordinates in the last column of Table 3 as a set of generalized conformations for the idopyranosyl ring in the solid-state. However, the $^4C_1$ conformation is admittedly the least well established one as the two Ido structures do not have an IdoA solid-state $^4C_1$ counterpart and, moreover, the two given Ido structures display quite a large spread in their P_2 coordinates.

We also looked into another source of experimentally obtained (solid-state) structural data, *viz.* the Protein Data Bank (PDB) [66], the single worldwide archive of structural data of biological macromolecules. About eighty entries in the PDB database were found to contain data of relevant ligands (2,3-di-O-methyl-α-L-idopyranuronic acid, L-iduronic acid, 2-O-sulfo-α-L-idopyranuronic acid, methyl 2-O-sulfo-α-L-idopyranosiduronic acid, heparin heptasaccharide, trisulfoamino heparin pentasaccharide and heparin pentasaccharide). However, the structural data of the ligands from which ring puckering parameters are deduced were deemed unsatisfactory as they are based on X-ray diffraction experiments lacking atomic resolution, or electron microscopy or solution NMR or other experimental techniques. This evaluation was bolstered by Smart et al. [67] who argued that the quality with which especially small-molecule ligands have been modelled in PDB-entries is a matter of concern. More specifically, a warning was issued [68] that at the typical resolution of 3 Å "…features such as ring pucker may eventually be set as a consequence of the restraint dictionary and fitting procedures used, rather than on the basis of the X-ray data." It was therefore decided not to use the information from the PDB database in the present analyses of the conformational space accessible to idopyranose derivatives, in order to avoid structural bias and/or errors.

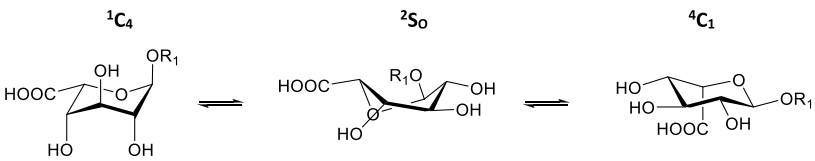
3.2. Conformation of α-idose pyranose rings in solution

As was noted in the Introduction, there is a longstanding consensus [23,29,30,69–71] that the six-membered L-iduronic acid ring is involved in a three-state equilibrium that encompasses $^1C_4$, ^2S_O and $^4C_1$ type of conformations (*cf.* **1**). This conformational equilibrium is thought to depend on the substitution pattern of the idopyranose ring itself [30] and, specifically in the case of linear heparin-like oligosaccharides, on that of adjacent [26,72] glucosamine units. We adhere to these notions and in the following sections we have grouped our data and analyses accordingly.

However, as a preamble to the analyses in the next sections we further note that the NMR data we assembled regard idose derivatives with a D- as well as an L-configuration. NMR cannot distinguish between enantiomers and so any consideration given to one of the enantiomorphs applies equally to the other. This makes the interpretation of *e.g.* coupling constants in terms of specifically the signs of the torsion angles involved solely dependent on the declared configuration of the sample compound. Throughout this paper we shall make use of this characteristic by putting all of our conformational analyses on a single footing, *i.e.* all results will be reported in terms of the L-enantiomer, irrespective of the stated configuration. In this way the solution TFP coordinates can be compared directly with each other as well as to the solid-state conformations (*cf.* Table 3). As such, this will make conformational correspondences (and differences) between six-membered rings instantly prominent since all necessary inversion/mirroring has already been applied.

3.2.1. O1-substituted α-L-Idopyranosiduronic acid (IdoA-1R) monosaccharides

A body of 57 α-IdoA-1R derivatives **8** with their four vicinal pyranose H–H couplings was extracted from the literature. The compounds are characterized by being substituted at O1 only, roughly half of them carry a methyl group (**8,** R_1 = CH_3), the remainder are mostly non-reducing terminal units of desulphated di-/tri-saccharides prepared from polymeric glycosaminoglycans such as heparin, dermatan sulfate, and the like. The NMR spectra of these compounds were recorded under widely varying conditions of temperature, pH and solvents.



8

O1-substituted α-L-Iduronic acid (IdoA-1R)

Our analysis started with the averaged TFP-coordinates of the three types of α-L-IdoA conformations observed in the solid-state (*cf.* Table 3). Following the procedures explained in the Methods section, these TFP-coordinates were used to generate the four endocylic proton-proton torsions for each of the three conformations, which in turn were translated into a set of four coupling constants for each of the three idopyranose conformations by means of the CAGPLUS equation [25] (*cf.* table in Fig. 2). Applying Eqs. (8)–(10) (*cf.* Methods) to these three sets of couplings in combination with the experimental coupling constants allowed us to determine the mole-fractions for the three constituent conformations and the concomitant averaged $^3J_{calc}$'s for each of the 57 IdoA-1R derivatives.

The results for this exercise are summarized in Fig. 2 where the calculated couplings of the aforementioned 57 IdoA-1R derivatives are plotted *versus* the experimentally observed ones. The datapoints are color-coded to identify the particular endocyclic coupling proton pair (H1H2, H2H3, *etc.*) they belong to. Moreover, the calculated mole-fractions (sorted by the population of the $^1C_4$ conformation) are abstracted in the column diagram at the bottom-right corner of the figure.

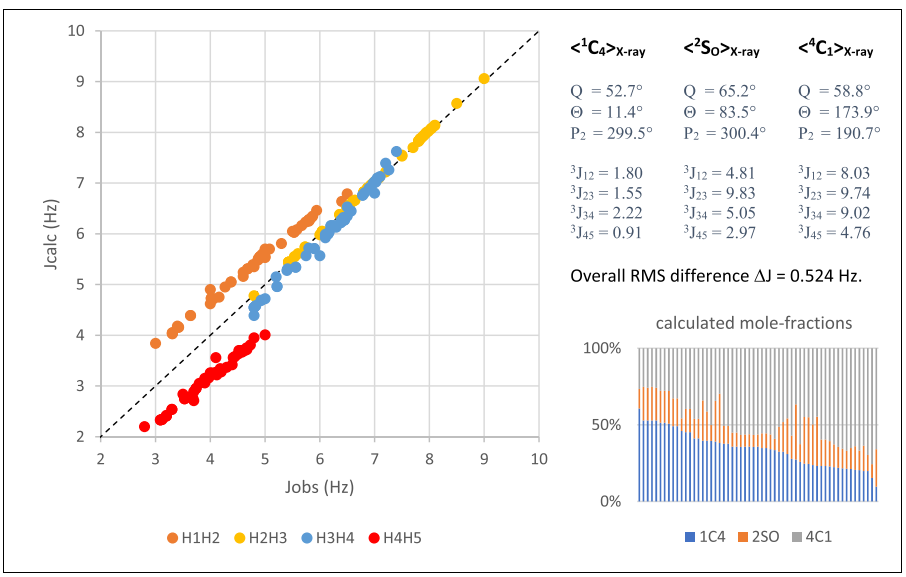


Fig. 2. Calculated vs. observed coupling constants for the endocyclic vicinal proton pairs in 57 IdoA-1R derivatives **8** using the average X-ray structures *ex* Table 3 for deriving the H–H torsion angles used in the CAGPLUS equation in order to calculate the couplings of the conformers that make up the three-state conformational equilibrium.

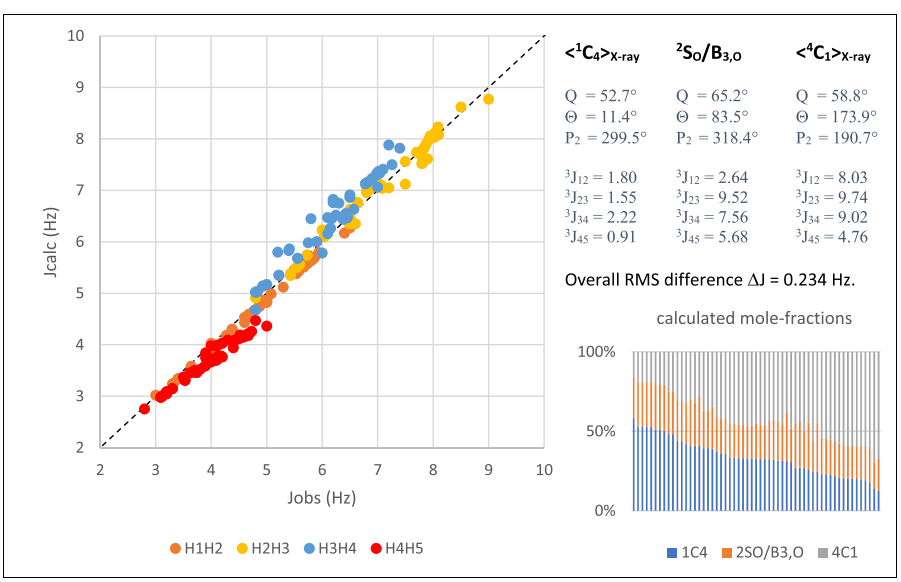


Fig. 3. Calculated vs. observed coupling constants for the endocyclic vicinal proton pairs in 57 IdoA-1R derivatives **8** after optimizing the pseudorotational phase angle P_2 of the ^2S_O conformation.

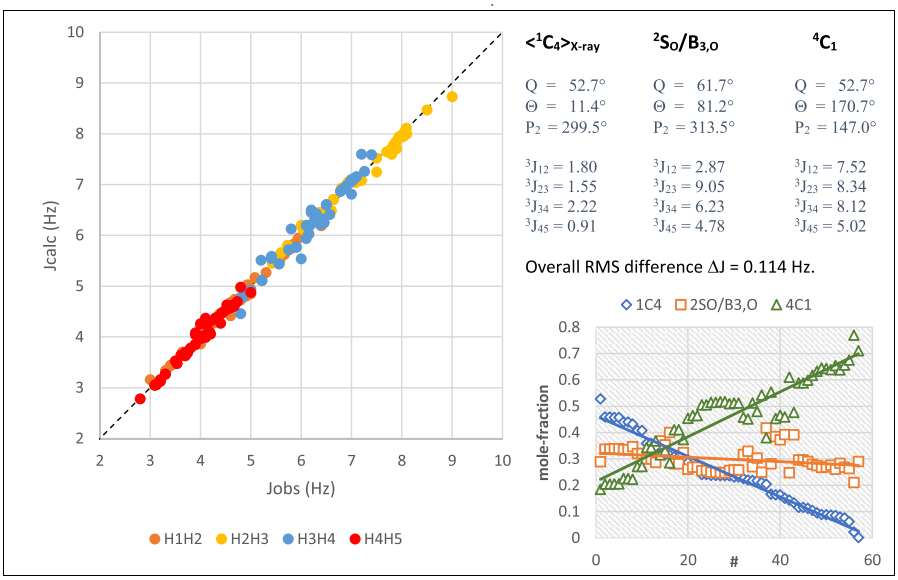


Fig. 4. Calculated vs. observed coupling constants for the endocyclic vicinal proton pairs in 57 IdoA-1R derivatives **8** after optimizing the TFP coordinates of both the ^2S_O and the $^4C_1$ conformations. Optimized mole-fractions are sorted according to decreasing fractions of the $^1C_4$ conformation in the three-state equilibria and plotted in the bottom-right chart; note that the X-axis has no correlation with any experimental condition. Solid lines are linear trendlines for the 3 mol-fractions, they have no physical bearing but are drawn for facilitating the discussion.

As Fig. 2 demonstrates there is an overall accordance between $^3J_{obs}$ and $^3J_{calc}$ whereby the calculated couplings deviate from the observed ones by 0.524 Hz on rms basis. Although such a residual rms value is not directly invalidating [35] the underlying conformational model (*i.e.* the 3 constituent conformers translated into the model coupling constants sets), the rms difference is quite substantial compared to the accuracy [25] of the CAGPLUS for the type of couplings at play (rms ≈ 0.36Hz). However, more disturbing is the distribution of the ($^3J_{obs}$, $^3J_{calc}$) datapoints in Fig. 2: realizing that in case of a perfect fit the datapoints would all lie on the dotted 45° diagonal line, the plotted data show that while the H2–H3 and H3–H4 couplings are practically spot on, the calculated H1–H2 couplings are roughly 0.6 Hz larger than the observed ones, whereas the calculated H4–H5 couplings are systematically *ca.* 0.8 Hz smaller than the reported experimental couplings.

Such discrepancies have been noted earlier in analyses of much smaller couplings ensembles on basis of canonical $^1C_4$, $^4C_1$ and ^2S_O conformations, but were left unremedied [24,73] or were redressed [30,74] somewhat haphazardly by assigning a single unusually high $^3J_{H4H5}$ value to a "slightly distorted $^4C_1$ IdoA conformer".

To address the systematic deviations apparent from Fig. 2 we reasoned that the lines formed by both the H1–H2 as well as the H4–H5 couplings are practically parallel to the 45° degrees line, thereby indicating that the deviations are more or less constant over the full scale of variances the couplings undergo. The mole-fractions diagram in Fig. 2 shows that of the three conformers only the ^2S_O displays a reasonable constant population, therefore the populations of the two chairs are more or less inversely related. Disregarding the remote possibility that corresponding couplings in the two chair conformations would exactly compensate each other upon changes in the equilibrium constants, we opted for looking into the ^2S_O conformation first. To that end the pseudorotation phase angle P_2 of the skew/boat was added to the least-squares optimization procedure which, after rapid convergence, led to the results summarized in Fig. 3.

The optimization led to a P_2 value of 318.4°, *i.e.* a small pseudorotational shift that slightly twists the starting ^2S_O conformer into a neighboring conformation roughly halfway between ^2S_O and $B_{3,O}$. But however small the conformational change, the impact on the fit of $^3J_{calc}$ *vs.* $^3J_{obs}$ is confounding: Fig. 3 demonstrates that the rms difference is more than halved (to 0.234 Hz) and, even more remarkably, the systematic deviations noted earlier for $^3J_{H1H2}$ and $^3J_{H4H5}$ have virtually disappeared. Comparison of the mole-fractions diagrams in Figs. 2 and 3 shows that the optimized equilibrium constants are also quite susceptible to the change in P_2 leading to a less choppy look for the mole-fractions diagram in Fig. 3 (although that is not necessarily a statistical feat).

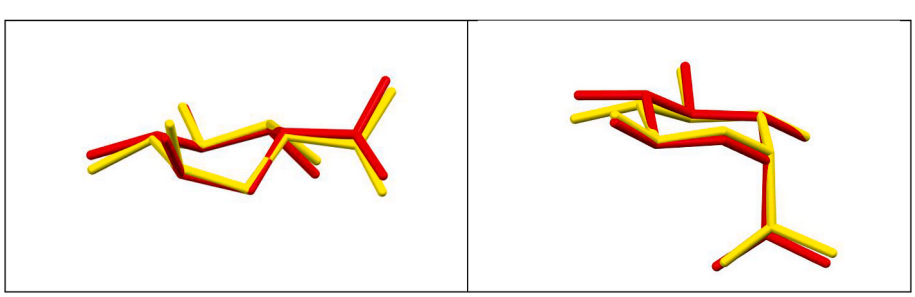


Fig. 5. Images of superimposed IdoA-1R models for ^2S_O-like (left panel) and $^4C_1$-like (right panel) conformations. Averaged solid-state conformations (*cf.* TFP coordinates Table 3) are colored red; solution structures (*cf.* TFP coordinates in Table of Fig. 4) are colored yellow. R-substituent and hydrogens not shown.

Table 4
Conformational energies calculated from the mole-fractions obtained for IdoA-1Me and IdoA-(1 → 4)-anManOH at pH7 and 22–25 °C, *cf.* entries #11 and #16, respectively, in Table S1, Supporting Information.

	IdoA-1Me	IdoA-(1 → 4)-anManOH
mole-fraction $^1C_4$	0.237	0.116
mole-fraction $^2S_O/B_{3,O}$	0.249	0.296
mole-fraction $^4C_1$	0.514	0.588
-ΔE($^1C_4$–$^2S_O/B_{3,O}$) *(kJ/mol)*	0.12	2.32
-ΔE($^1C_4$–$^4C_1$) *(kJ/mol)*	1.92	4.02

As such, we appear to have obtained a satisfactory description of the O1-substituted IdoA pyranose rings but for the notion that the $^3J_{calc}$ *vs.* $^3J_{obs}$ correlation for H3–H4 seems to have deteriorated somewhat: the calculated couplings for this proton pair are now slightly but systematically overestimated. Moreover, scrutiny of the coupling constant data shows that the deviations in $^3J_{H3H4}$ as well as the remaining deviations in $^3J_{H4H5}$ are directly related to the calculated occurrence (mole-fraction) of the $^4C_1$ conformation in the three-state equilibria. These findings then triggered an extra round of optimization which involved the TFP coordinates of specifically the $^4C_1$ as well as the ^2S_O conformer and obviously the mole-fractions of the 57 equilibria involved.

The results from this final iteration are depicted in Fig. 4 (for a full account, *cf.* Table S1, Supporting Information), as can be seen from the figure there is an excellent correspondence between calculated and observed coupling constants (rms-difference = 0.114 Hz). Interestingly, of the three constituent six-membered ring conformations the well-defined solid-state $^1C_4$ structure appears to be preserved in solution whereas the remaining solid-state ^2S_O and $^4C_1$ conformers adopt in solution ring geometries that display a change in TFP coordinates. The data show that in solution the ^2S_O conformer is slightly less puckered ($< Q >_{X\text{-ray}}$ = 65.2° *vs.* 61.7° in solution) and has pseudorotated *ca.* 15° towards an intermediate $^2S_O/B_{3,O}$-like conformation.

The final value calculated for the P_2 coordinate of the $^4C_1$ conformer appears to differ quite substantially from the average solid-state value (190.7° *vs.* 147° in solution), however, as was noted before the solid-state average is derived from just two datapoint encompassing quite a large spread in P_2. We take this large spread as an indication for a diversity in crystal packing effects, especially as the optimized solution P_2 coordinate practically mirrors the solid-state P_2 value of one of the entries (*cf.* MABIDP, Table 3). Moreover, it should be borne in mind that the effect of the P_2 coordinate on the overall ring geometry is governed by the (relative) amplitude of Φ_2 – which is rather small in the case of chair-like conformations. As such, morphologically the differences between the corresponding solid-state *vs.* solution conformers are a matter of degree, not of kind. This is visualized in Fig. 5 where the core structures are shown in a superimposed fashion.

But notwithstanding their overall conformity, at the more detailed level of individual torsion angles the structures differ considerably. For instance, in the ^2S_O-like conformations the largest change is in the endocyclic O5–C1–C2–C3 torsion angle (ν_1) which tightens from 24° in the solid-state to 8° in solution. It goes without saying that the corresponding H1–C1–C2–H2 follows suit and changes from 144° to 128°, thereby lowering the $^3J_{12}$ from 4.8 Hz to 2.9 Hz. A similar, but opposite change (39° → 27°) is noted for ν_4 (*i.e.,* C3–C4–C5–O5 torsion) which increases the corresponding $^3J_{45}$ by 1.8 Hz. These two concomitant changes underly the aforementioned removal of the systematic deviations demonstrated in the graph of Fig. 2. For the sake of completeness we note that in the $^4C_1$-like conformations the C1–C2–C3 flap of the chair "flattens" somewhat in solution, instigating a lowering of $^3J_{23}$ and $^3J_{34}$ which in turn caused the final "perfect" fit in Fig. 4 to be achieved.

The conformational analysis of the IdoA-1R derivatives is completed by considering the involved equilibrium constants, *i.e.* the relative populations of the three constituent conformers. Remarkably, the mole-fractions plot in Fig. 4 shows the population of the $^2S_O/B_{3,O}$-conformer to remain more or less constant at roughly 30% over a wide variety of experimental conditions and O1-substituents, whereas the populations of the two chair conformations vary in an almost inversely proportional way: lowering of the pH will favor the $^1C_4$ form, whereas the $^4C_1$ conformer benefits from increasing temperatures and the addition of organic solvents or cations. However, although the mole-fractions for the 57 samples fluctuate substantially, the associated shifts in conformational energies is modest: for IdoA-1Me (**8**, R_1 = CH_3) at ~ pH 2 the relative populations $^1C_4$: $^2S_O/B_{3,O}$: $^4C_1$ ≈ 0.5 : 0.3: 0.2 corresponds [75] to the $^1C_4$-conformer being around 1.25 kJ/mol *lower* and the $^4C_1$-conformer being 1.25 kJ/mol *higher* in energy than the $^2S_O/B_{3,O}$-conformation. At the other end of the trendlines we typically find the non-reducing terminal IdoA 's of di-/tri-saccharide derivatives (**8**, R_1 = furanose or pyranose derivative) with distributions that end up around $^1C_4$: $^2S_O/B_{3,O}$: $^4C_1$ ≈ 0.05 : 0.25: 0.70 which translate into a reversed situation in which the $^1C_4$-conformation is now ~4.0 kJ/mol *higher* and the $^4C_1$-conformer is ~2.5 kJ/mol *lower* in energy than the $^2S_O/B_{3,O}$-form.

Thus, the trendlines also appear to indicate that, compared to a methyl group, the bulkier saccharide substituents occupying the right-hand side of the mole-fractions plot, have a larger preference for the (quasi-)equatorial orientation of O1 as it occurs in the $^2S_O/B_{3,O}$ and the $^4C_1$ conformations. This preference can even be quantified by comparing *e.g.* the conformer distributions for IdoA-1Me and IdoA-(1 → 4)-anManOH (**8**, R_1 = 2,5-anhydro-D-mannitol) at pH 7 and ambient temperature: as can be gleaned from the ΔE data given in Table 4, the bulky anManOH-group disfavors the axial O1-position in the IdoA-1R $^1C_4$-conformer by *ca.* 2 kJ/mol when compared to an axial methyl group in the corresponding conformation.

So far in this section we have sported a deductive line of reasoning that led to the final set of conformers underpinning the three-state equilibria of IdoA-1R derivatives (*cf.* Fig. 4). For the sake of the discussion we also explored some alternate, more systematic approaches. For instance, instead of starting the structure optimization with the ^2S_O conformer (ultimately leading to the situation summarized in Fig. 3) we also ran comparable optimizations that started with either the $^1C_4$ or the $^4C_1$ conformer. Allowing the same degrees of freedom the thus optimized $^1C_4$ conformer was virtually identical to the averaged X-ray structure for the $^1C_4$ conformer (*cf.* Table 3): a negligible change in P_2 (299.5° → 305.0°) was noted that did not have any effect on the rms difference ΔJ. For the $^4C_1$ conformer the optimization did not converge, but the lowest rms difference ΔJ attained was 0.469 Hz. Even more relevant than the small (if any) rms improvements is the finding that

Table 5
Conformational equilibria of two idopyranose derivatives analyzed in terms of the conformational parameters determined for IdoA-1R derivatives (*cf.* Table S1, Supporting Information).

	J1-2	J2-3	J3-4	J4-5	% $^1C_4$	% $^2S_O/B_{3,O}$	% $^4C_1$	rmsdiff
α-idopyranose (1)								
J_{obs} [77]	6.0	8.1	7.9	5.0				
J_{calc}[a]	6.19	8.34	7.54	4.86	0.024	0.232	0.744	
Diff	−0.19	−0.24	0.36	0.14				0.247
Me-α-idopyranose								
J_{obs} [76]	4.3	6.7	6.1	3.6				
J_{calc}	4.39	6.71	5.86	3.81	0.275	0.335	0.389	
Diff	−0.09	−0.01	0.24	−0.21				0.164

[a] Coupling constant calculations corrected for non-substituted O1.

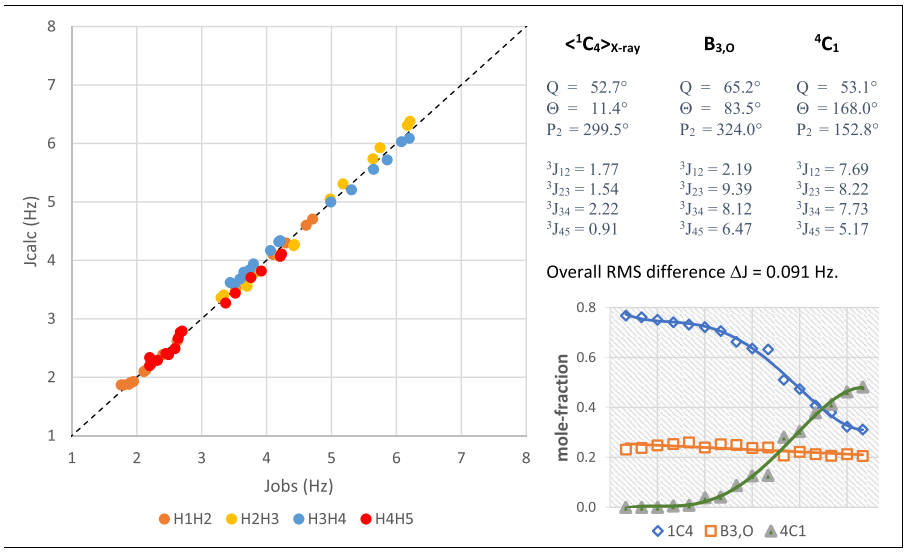


Fig. 6. Calculated *vs.* observed coupling constants for the endocyclic vicinal proton pairs in 16 IdoA2S-1R derivatives **9** involved in a three-state equilibrium. Only the pseudorotational phase angle P_2 of the ^2S_O conformation and all three TFP parameters of the $^4C_1$ conformation were optimized. Optimized mole-fractions are sorted according to decreasing fractions of the $^1C_4$ conformation in the three-state equilibria and plotted in the bottom-right chart; note that the X-axis has no correlation with any single experimental condition, hence the polynomial solid trendlines for the 3 mole-fractions bear no physical meaning.

optimization of either chair conformation does hardly influence the systematic deviations displayed by $^3J_{H1H2}$ and $^3J_{H4H5}$ in Fig. 2. In other words, it is indeed only a change in the ^2S_O puckering that can remedy the mentioned systematic deviations and bring about a sizeable decrease in the rms ΔJ. We further note that also at later stages of the optimization procedure (*i.e.* after optimizing the ^2S_O conformer as summarized in Fig. 3) no $^1C_4$ puckering could be found that surpasses the fit summarized in Fig. 4.

As such, we conclude that the intermediate $^2S_O/B_{3,O}$-like conformation (P_2 = 313.5°) is essential for obtaining a satisfactory fit between experimental and calculated couplings (*i.e.* without laterally shifted subsets). The final overall best fit confirms the applicability of the well-defined averaged X-ray $^1C_4$ structure as one of the two chair conformations that complete the present conformational equilibria.

Our study shows that in this way the TFP/CAGPLUS approach is able to reproduce 228 experimental IdoA-1R coupling constants from various sources in a single overarching model with an overall rms difference of 0.114 Hz. This excellent correspondence appears to transcend the stated accuracy (rms ≈ 0.36Hz) of the CAGPLUS equation [25]; by the way an occurrence that in itself is not isolated, *cf. e.g.* the 0.21 Hz rms difference calculated for compound **7** (*vide supra*) and other showcases [35].

Incongruous as this may seem at first sight, it is good to bring to mind that the stated CAGPLUS accuracy is based on the rms difference between observed and calculated coupling constants determined from a "training" set of 170 couplings used in parametrizing the CAGPLUS. It was estimated [25] that experimental (random) errors in the observed couplings of the "training" set accounted for 0.2–0.3 Hz of this overall rms difference whereas systematic errors (*e.g.* neglect of extra cross-terms by the CAGPLUS) could be held responsible for the remainder (here 0.05–0.15 Hz), in other words, the latter (ISO 5725) *trueness* of the CAGPLUS is much better than the stated overall accuracy. As indicated in Section 2.4 (*vide supra*) the quality of the NMR data used in the current paper is very good, therefore the reported vicinal couplings are definitely more precise than the couplings on which the CAGPLUS was gauged. This goes especially for the small monomeric units studied in this section and hence it might be expected that the rms differences get closer to the *trueness*-range of the CAGPLUS. In line with this rationale is the deliberation that for the bigger oligomeric IdoA units the experimental couplings are expected to be slightly less accurate (*cf.* Section 2.4), a condition that is indeed reflected in the higher overall rms differences reported in Tables S5–S7 (Supporting Information).

Taken together, the data manifest that the conformational behavior of the IdoA-1R pyranose ring can be described in terms of a three-state equilibrium in which the three constituent ring conformations are the same for all 57 measurements notwithstanding the wide range of experimental conditions they represent in terms of solvent, salt, pH, temperature and/or substituent variations. It is shown that in IdoA-1R derivatives **8** only the equilibrium constants appear to vary as a corollary of the experimental conditions which invoke limited range of energy differences (4–5 kJ/mol) between the constituent conformers.

3.2.2. α-Idopyranose conformation in solution

Having established the conformational features of the six-membered O1-substituted idopyranosiduronic acid (IdoA-1R) ring it is of interest to see whether the proclaimed conformity between IdoA and Ido derivatives in the solid-state (*vide supra*) also holds in solution. To that end we analyzed the coupling constant data for two idopyranose derivatives, *i.e.* **1** and its O1-methylated form.

In view of the rather limited set of Ido data we settled for a cross-validation approach in which we treat the IdoA conformations as the "training" data and check whether this IdoA model performs equally well on the two idopyranose derivatives. Thus, we applied the least squares approach described by Eqs. (8)–(10) as implemented in the Excel workbook SixPuckTF_CarbRes.xlsx, tab LSQ IdoA-1R (*cf.* Supporting Information) and used the coupling constants derived for the

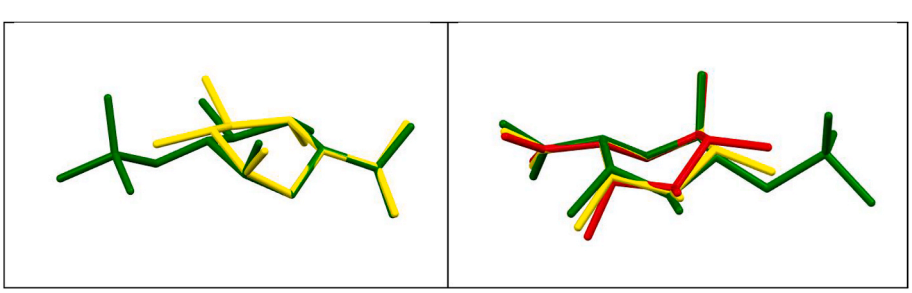


Fig. 7. Images of superimposed $^2S_O/B_{3,O}$-like vs. $B_{3,O}$-like conformations for IdoA-1R and IdoAS-1R models, respectively. Averaged solid-state conformation (^2S_O, *cf.* TFP coordinates Table 3) is colored red; solution IdoA-1R structure (*cf.* TFP coordinates in Table of Fig. 4) is colored yellow, solution IdoAS-1R structure (*cf.* TFP coordinates in Table of Fig. 6) is colored green; R-substituent and hydrogens not shown. Left panel: C1–O5 in front. Right panel: C4–C3–C2 flap up front.

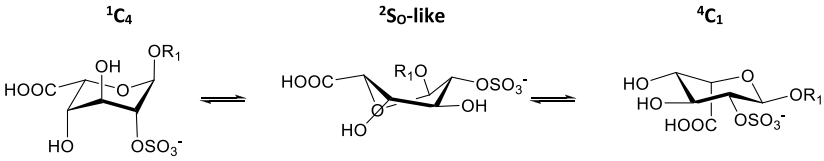
three constituent conformers of IdoA-1R derivatives (*cf.* Fig. 4) as the model for quantifying the conformational equilibria the Ido derivatives are involved in.

Table 5 lists the results for this analysis and it is shown that a three-state conformational equilibrium identical to the one delineating IdoA-1R derivatives is indeed able to deliver a good correspondence between observed and calculated coupling constants for the two idopyranose rings (rmsdiff $\Delta(J_{obs} - J_{calc})$ of 0.16–0.25 Hz). By contrast, our analyses do not corroborate the simplifying assumption [76] of a two-state $^1C_4/^4C_1$ equilibrium for this type of Ido derivatives as such an assumption gives rise to an unsatisfactory rms-difference $\Delta(J_{obs} - J_{calc})$ of 0.6–0.75 Hz.

Taken together, we conclude that the TFP coordinates determined for IdoA-1R derivatives serve the two Ido derivatives equally well. As the three-state equilibrium for Ido displays the same ~30% population for the skew/boat-like conformer as IdoA turns out to have, we conclude that the conformational features of IdoA and Ido in solution are interchangeable, like they are in the solid-state (*vide supra*).

3.2.3. O1-substituted 2-O-sulfo-α-L-iduronic acid (IdoAS-1R) derivatives

Sulfation at O2 is a common variety of the iduronic acid moiety found in heparin, but although the monomeric unit has been the subject of many semi-empirical and/or theoretical studies [51,78], experimental data on this basic IdoAS-1R moiety (**9**) have been reported less abundantly.



9
2-O-sulfo-α-L-Iduronic acid

Fortunately, the study by Sanderson et al. [30] provided us with 13 coupling constants sets for the IdoA2S-(1,4)-anManOH(6S) disaccharide recorded under a spread of experimental conditions. The dataset was augmented with coupling constants for one more disaccharide and two methyl derivatives (*cf.* Table S2, Supporting Information). Thus, a grand total of 16 sets of IdoA2S-1R (**9**) coupling constants could be analyzed along the same lines as described above for the non-sulfated IdoA2-1R derivatives (**8**). After optimization of the ^2S_O and $^4C_1$ conformations we ended up with an excellent fit between 64 observed and calculated couplings (rms difference = 0.091 Hz, *cf.* Fig. 6).

The TFP coordinates reported in Fig. 6 show that next to the preserved solid-state $^1C_4$ structure also the optimized $^4C_1$ conformer of **9** is virtually identical to the corresponding structure deduced for the unsulfated analogues **8** (*vide supra*). It is for the skew/boat form that the key change is found: in the IdoA2S-1R derivatives (**9**) this conformation has shifted further along the pseudorotational equator to take up a $B_{3,O}$-like conformation (P_2 = 324°); Fig. 7 gives an impression of how this

Table 7
Optimized TFP coordinates for the constituent pyranose conformations in O1,O4-disubstituted-α-L-iduronate (IdoA-1,4R) derivatives. Chair conformations are invariant, only the ^2S_O conformations are different in the three types of derivatives studied (DS-type, DS-red refer to dermatan species; *vide infra*).

IdoA-1,4R	$^1C_4$ (X-ray)	^2S_O (Hp-type)	^2S_O (DS-type)	^2S_O (DS-red)	$^4C_1$
Q	52.7°	54.7°	60.1°	56.1°	52.9°
Θ	11.4°	83.9°	82.6°	82.4°	169.7°
P_2	299.5°	303.0°	302.6°	293.3°	144.5°

$B_{3,O}$-form aligns with the solid-state and IdoA-1R structures discussed in the preceding section.

In this $B_{3,O}$ conformation the sulfate and carboxyl groups each take up a boat-equatorial (Be) position [79] that are situated diagonally across the six-membered ring, which is arguably the farthest apart the two anionic groups can get.

Inspection of the mole-fractions plot in Fig. 6 shows again a remarkable constancy of the relative population of the $B_{3,O}$-like conformer (20–25%) over quite a spread of experimental conditions, albeit at a somewhat lower level than noted before for **8**. Energy differences [75] between the $^1C_4$ and the second stable conformation at play are quite small and roughly limited to a range of −3 to +1 kJ/mol. The $^1C_4$ conformer appears to prevail in aqueous solution at low pH and low temperature, whereas organic solvent mixtures tend to favor the $^4C_1$ conformation. Moreover, competition between these trends is evidenced by the IdoA2S-(1 → 4)-anManOH(6S) disaccharide in DMSO at pH 2.3 [30]: entries #12 (60 °C) and #13 (90 °C) in Table S2 (*cf.* Supplementary Information) show the idopyranosyl equilibria to parallel those in D_2O at pH2 (entries #4 and #5), thereby suggesting that the effect of the COOH ionization state on the conformational equilibrium [76] transcends the DMSO solvation effects.

Ushered by our findings on **8** we looked whether the nature of substituent R_1 of **9** influences the three-state equilibrium similarly. Unfortunately, the data in Table S2 (*cf.* Supplementary Information) in itself are inconclusive as to whether a bulky saccharide substituent affects the equilibria constants in a way different of that of *e.g.* a methyl substituent. However, a comparison of specific IdoA-1R derivatives with their corresponding IdoAS-1R derivatives does lend some insight into the effect a 2-*O*-sulfo substituent has on the idopyranose ring

Table 6
Conformational energies calculated [75] from the mole-fractions obtained for IdoA-1R and IdoAS-1R derivatives selected to estimate the effect of IdoA sulfation.

	IdoA-1Me[a]	IdoAS-1Me[b]	IdoA-(1 → 4)-GlcNS[c]	IdoAS-(1 → 4)-GlcNS[d]
mole-fraction $^1C_4$	0.210	0.768	0.315	0.741
mole-fraction $B_{3,O}$	0.249	0.232	0.400	0.254
mole-fraction $^4C_1$	0.542	0	0.285	0.005
$-\Delta E(^1C_4–B_{3,O})$ *(kJ/mol)*	0.42	−2.96	0.59	−2.65
$-\Delta E(^1C_4–^4C_1)$ *(kJ/mol)*	2.35	≤-20	−0.25	−12.4

[a] Entry #14 *ex* Table S1, Supporting Information.
[b] Entry #14 *ex* Table S2, Supporting Information.
[c] Entry #47 *ex* Table S1, Supporting Information.
[d] Entry #16 *ex* Table S2, Supporting Information.

Table 8
Conformational equilibria for **13** heparin-type IdoA-1,4R derivatives sorted by the $^1C_4$ mole fraction of the IdoA units.

Entry[a]	Length[b]	IdoA pos[c]	#Sulf NR–R[d]	fragment	mole-fraction		
					$^1C_4$	^2S_O	$^4C_1$
#9	hexa	3	3–2	GlcNS3S6S-(1 → 4)-IdoA-(1 → 4)-GlcNS6S	0.67	0.33	
#8	hexa	3	2–2	GlcNS6S-(1 → 4)-IdoA-(1 → 4)- GlcNS6S	0.63	0.37	
#3	tetra	1	3–0	GlcNS3S6S-(1 → 4)-IdoA-O1Me	0.62	0.38	
#4	penta	2	3–1	GlcNS3S6S-(1 → 4)-IdoA-(1 → 4)- GlcNS	0.58	0.42	
#5	penta	2	3–3	Glc2S3S6S-(1 → 4)-IdoA-(1 → 4)-Glc2S3S6S[e]	0.55	0.45	
#12	hexa	2	1–0	GlcNS-(1 → 4)-IdoA-(1 → 4)- GlcNAc	0.54	0.46	
#13	poly		1–1	GlcNS-(1 → 4)-IdoA-(1 → 4)-GlcNS	0.54	0.28	0.18
#6	hexa	3	1–1	GlcNS-(1 → 4)-IdoA-(1 → 4)- GlcNS	0.51	0.49	
#7	hexa	3	1–1	GlcNS-(1 → 4)-IdoA-(1 → 4)- GlcNS	0.49	0.51	
#2	tri	2	1–1	GlcNS-(1 → 4)-IdoA-(1 → 4)-GlcNS	0.49	0.51	
#11	hexa	1	2–0	GlcNS6S-(1 → 4)-IdoA-O1-iPr	0.49	0.51	
#1	tri	2	0–0	GlcNAc-(1 → 4)-IdoA-(1 → 4)-GlcNAc	0.32	0.68	
#10	hexa	3	1–1	GlcNAc6S-(1 → 4)-IdoA-(1 → 4)-GlcNS	0.30	0.70	

[a] Entry number corresponds with Table S3, Supporting Information.
[b] Chain length of the oligosaccharide.
[c] Position of the IdoA unit proceeding from the reducing group in the saccharide oligomer, *cf* [44].
[d] Sulfation pattern expressed as the number of sulfate groups on the saccharide flanking IdoA at its non-reducing (NR) end, followed by the number of sulfate groups of the saccharide residing at the IdoA reducing (R) end.
[e] Note that in this entry (SanOrg 34006/Idraparinux) the flanking GlcN saccharides are replaced by Glc units; moreover, O2 and O3 of IdoA are both methylated.

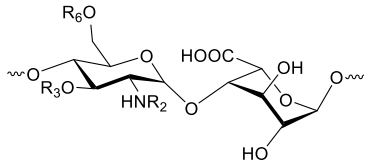
conformers. Table 6 compares the conformational equilibria in D_2O at ambient temperature derived for two pairs of IdoA derivatives which exemplify the effect of sulfation of the idose-O2: both pairs of derivatives show that the conformational energy difference between the $^1C_4$ and $B_{3,O}$-like conformers change by 3–3.5 kJ/mol, *i.e.* the $B_{3,O}$ conformer becoming less favorable by that amount upon sulfation. The $^4C_1$ conformer suffers an even larger loss as sulfation appears to induce at least a 12 kJ/mol change in conformational energy in favor of the $^1C_4$ conformation [80].

Of course, it is tempting to rationalize the above findings in terms of clear-cut and well-established (conceptual) forces, but the reality is that our data cannot differentiate between the possible causes (*e.g.* steric [81], stereoelectronic [82], coulombic [83], hydrophilic [84], entropic [81], *etc.*) for our findings and as such do not speak to the etiology of the observed phenomena.

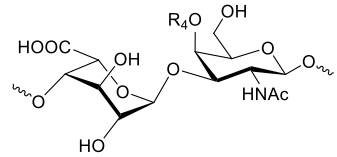
We therefore rest this case with the finding that O2 sulfation of the IdoA-1R framework has practically no consequences for the chair conformations of the resulting IdoA2S-1R derivatives, but it does mold the O2 sulfated skew form into a $B_{3,O}$-like conformation. Together these conformers are the constituents of a three-state equilibrium for which only the equilibrium constants vary as a function of experimental conditions.

3.2.4. O1,O4-disubstituted-α-L-iduronic acid (IdoA-1,4R) derivatives

Non-sulfated IdoA residues are found in two types of glycosaminoglycans (GAGs), *i.e.* the heparin (**10**) and dermatan (**11**) species [1]. As the IdoA conformation is thought to be correlated [26,28,72] to the nature of its neighboring saccharide units we treated the two types of GAG derivatives separately.



10
*Main repeating GlcN-IdoA disaccharide unit making up **heparin-type** structures; R may be H, Ac or sulfate group*



11
*Main repeating IdoA-GalNAc disaccharide unit making up **dermatan-type** structures; R_4 may be H or sulfate group*

3.2.4.1. IdoA-1,4R in heparin-type oligosaccharides. In the GAG heparin species the saccharide moieties are all (1 → 4) linked to each other. We extracted from the literature 13 sets of coupling constants for oligomers of this type with an IdoA saccharide within the chain. The couplings were analyzed in terms of a conformational equilibrium in a similar way as described above, yielding in a very good correspondence between observed and calculated coupling constants (rms = 0.256 Hz); the results of this analysis are paraphrased in Table S3, *cf.* Supporting Information.

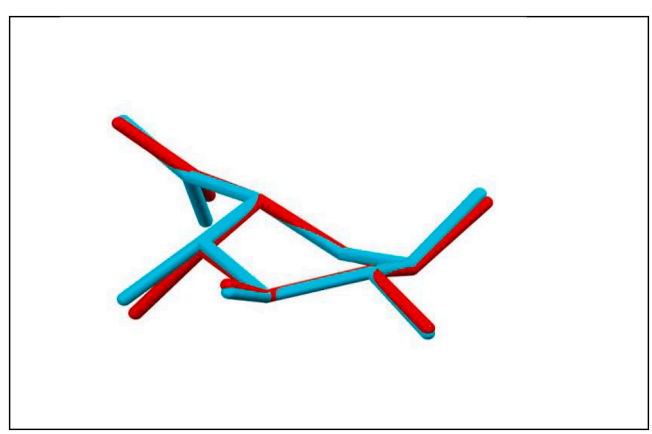


Fig. 8. Image of the superimposed ^2S_O-like conformations calculated for the two dermatan-type IdoA-1,4R derivatives. The optimized structure for the "DS-type" cluster is colored cyan (P_2 = 302°), the structure for the "DS-red" derivatives (P_2 = 293°) is colored red.

As Table 7 points out, both chair conformations in the optimized dataset are invariable and as such virtually identical to the corresponding conformations in earlier analyses (*vide supra*). Instead, it is again the skew/boat conformation that bears the brunt of adapting to this particular idose O1,O4 substitution pattern by taking up an almost pure $^{2}S_{O}$ form ("Hp-type": Q = 54.7°, Θ = 83.9°, P_2 = 303.0°). In pseudorotational terms it is very much a look-alike of the corresponding averaged solid-state conformation reported in Table 3, albeit that the solution form appears somewhat less puckered. This being the case, the three superimposed models in the right-hand box of Fig. 7 can be used to grasp the effect that an added O4-substituent has on the IdoA conformation: the red $^{2}S_{O}$ structure is "favored by the absence of axial interactions" [78a] and hence preferred by disubstituted IdoA-1,4R derivatives.

Table 8 characterizes the analyzed heparin-type IdoA-1,4R fragments and the conformational equilibria they are involved in. Our optimizations purport to show that the population of the $^{4}C_{1}$ conformation is zero for all fragments, with the exception of Entry #13, an O- and N-desulfated re-N-sulfated heparin [85], which appears to be engaged in a three-state equilibrium (but see also the Note included in Table S3, Supporting Information). There is no clear-cut explanation for the deviant behavior of this modified heparin polymer other than the notion that the couplings for this singular measurement were recorded at 70 °C, a condition from which the $^{4}C_{1}$ conformation seems to benefit (*vide supra*).

The remaining 12 oligosaccharides adopt two-state equilibria for which the range in calculated mole-fractions is rather small: the $^{1}C_{4}$-conformer population scale runs from 0.67 down to 0.30, *i.e.* the conformational energy difference -ΔE stays within ± 2 kJ/mol.

In Table 8 we also recorded the structural/chemical environment of each of the IdoA moieties, *i.e.* the length of the oligomer and the position the IdoA takes up in it, as well as the sulfate pattern originating from the Glc neighbors. Scrutiny of the collected data shows that the conformation of IdoA moieties in heparin-type oligosaccharides is not markedly influenced by these surrounding characteristics: the length of the oligomer nor the place of the IdoA moiety in the oligomer sequence appears to appertain to the mole-fractions distribution. Also, the sulfation pattern of the IdoA's flanking saccharides shows no obvious correlation apart from the rather indistinctive notion that the fragments with the highest degree of sulfation (> 3 sulfate groups) appear at the top of the table where the conformational equilibria displaying the highest $^{1}C_{4}$ mole-fractions are found, a finding that corroborates a similar suggestion made by Hsieh et al. [28].

3.2.4.2. IdoA-1,4R in dermatan-type oligosaccharides. Dermatan sulfate (DS) belongs to a second family of iduronic acid-containing GAGs and is typified by a repeating IdoA-GalNAc disaccharide unit with an (α-1→3) linkage between IdoA and the galactosamine moiety at its reducing locant 1. Similar to IdoAS's purported role in heparin-type GAGs, the non-sulfated IdoA units in DS are believed to allow for a more flexible saccharide chain with an increased potential for binding [12,86].

Flexibility of the IdoA moiety is indeed borne out by our optimizations for the conformational equilibria of the dermatan-type oligomers. But notably, the collective optimization of all DS oligomers also made clear that the calculated coupling constants for IdoA saccharides with an N-acetyl-D-galactosaminitol ("H-ol") substituent at its O1 position deviated in a systematic way from the observed couplings. This discrepancy was remedied by carving out the affected IdoA moieties and subject them to a separate minimization. Thus we ended up with two ensembles of conformational equilibria in which only the constituent $^{2}S_{O}$ conformers differ (*cf.* Table 7, columns marked DS-type for the normal "DS-type" oligomers and "DS-red" for the oligomers with a reduced, alditol-type terminus; full details for the minimizations are listed in Table S4, Supporting Information).

Table 7 shows that also for the "DS-type" IdoA-1,4R derivatives the constituent $^{1}C_{4}$ and $^{4}C_{1}$ chairs are practically invariant with respect to the above described analyses of Ido[A][S] variants, and anew it is only the skew/boat conformer that needs to adapt in the optimizations. In the present case we find the optimized $^{2}S_{O}$ conformer for DS-type oligomers to be similar to the corresponding solid-state structure as well as to the heparin-type $^{2}S_{O}$ conformers described in the previous paragraph: the largest difference in TFP coordinates is noted for the Q parameter with the DS-type $^{2}S_{O}$ conformer displaying a Q that is practically in the middle between the other two.

Table 9
Comparison of the two-state equilibrium analyses for the three clusters of O1,O4-disubstituted-2-O-sulfo-α-L-iduronic acid (IdoAS-1,4R) derivatives sorted out on basis of the number of sulfate substituent on flanking saccharides (*cf.* text).

IdoAS-1,4R		(≤4 sulfates)[a]	(4+ sulfates)[b]	(6 sulfates)[c]
# sets ^{3}J		45	14	4
rms ΔJ (Hz)		0.280	0.224	0.248
TFP coordinates:	$^{1}C_{4}$	$^{2}S_{O}$	$^{2}S_{O}$	$^{2}S_{O}$
Q	52.7°	55.6°	55.9°	55.2°
Θ	11.4°	81.1°	81.1°	86.2°
P_2	299.5°	305.3°	301.2°	296.5°

[a] *Cf.* Table S5, Supporting Information.
[b] *Cf.* Table S6, Supporting Information.
[c] *Cf.* Table S7, Supporting Information.

Remarkably, in the reduced oligosaccharides ("DS-red") the IdoA moieties next to the terminal alditol residue display a $^{2}S_{O}$ conformation that has shifted to a slightly smaller P_2 coordinate value (*cf.* Table 7). Morphologically such a change in the overall $^{2}S_{O}$ ring conformation is almost negligible (see Fig. 8), but the changes in endocyclic coupling constant values are sizeable [87]. It is suggested that the pseudorotational shift is a consequence of a (steric?) interaction of the open-chain N-acetyl-D-galactosaminitol residue at O1 and the GalNac-pyranose substituent at the IdoA-O4, as the corresponding IH-ol disaccharides, *i.e.* IdoA with the GalNac-alditol at O1 but without a substituent at O4, do not display a conformational divergence from the collective IdoA-1R derivatives (*cf.* entries #55 and #56 in Table S1, Supporting Information).

The equilibrium constants deduced for these DS-type oligomers show that the two compounds in which the idopyranose is on both ends flanked by O4-sulfated GalNAc moieties (entries #1 and #2 in Table S4, Supporting Information) are engaged in a two-state $^{1}C_{4}/^{2}S_{O}$ equilibrium, with the $^{2}S_{O}$ being overrepresented (55–80%). This is also the case for the desulfated DS oligomers, albeit that *e.g.* at higher temperature also the $^{4}C_{1}$ conformer comes into play again, indicating that in these uncharged molecules the $^{4}C_{1}$ is in energetic terms not far off. The IdoA-pyranose rings next to the galactosaminitol moieties ("DS-red") reinforce this notion as these rings all adopt a three-state conformational equilibrium with a considerable $^{4}C_{1}$ mole-fraction. This behavior is akin to the IH-ol disaccharides (*cf.* Table S1, Supporting Information) which massively favor the $^{4}C_{1}$ conformer.

In summary, our analyses show that the IdoA-1,4R conformational features of heparin-type and dermatan-type oligomers are very similar: the constituent idopyranose rings are engaged in "classical" $^{1}C_{4}/^{2}S_{O}/^{4}C_{1}$ equilibria differing in conformer mole-fractions only. The IdoA moieties neighboring a terminal N-acetyl-D-galactosaminitol residue form an idiopathic exception in that their $^{2}S_{O}$ conformation takes up a slightly less skewed form that has pseudorotated slightly into the direction of a $^{2,5}B$ conformation. No other idopyranose conformations were needed for a satisfactory interpretation of the observed NMR coupling constants and as such our analyses do not support the suggested presence [57,88,89] in DS of grossly different conformations like $^{O}S_{2}$. Finally, as the spread in conformer populations cannot be correlated to specific features of the oligomers the IdoA moieties are part of, we conclude that the small conformational energy differences observed

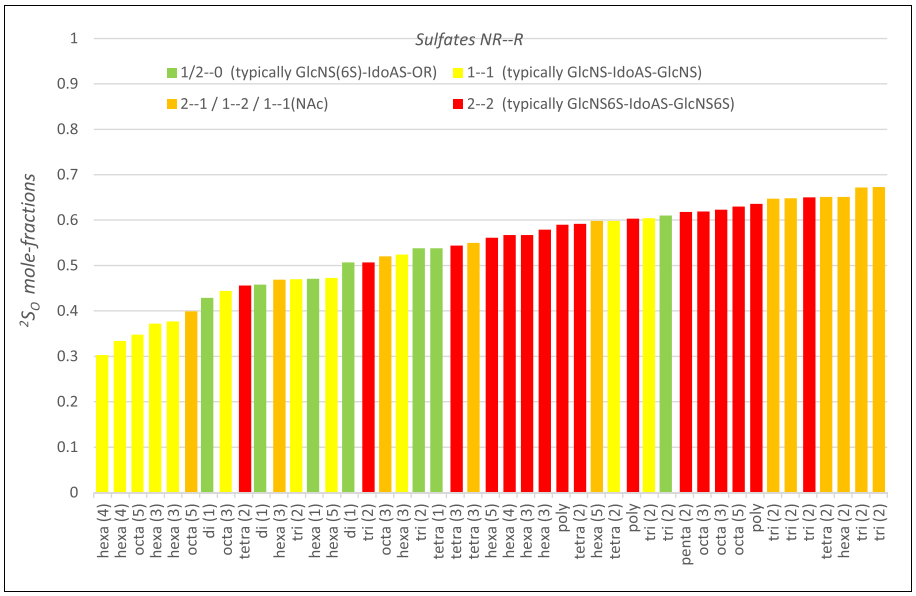


Fig. 9. Optimized mole-fractions for the "≤ 4 sulfates" cluster of IdoAS-1,4R derivatives sorted according to increasing fractions of the ^2S_O conformation in the two-state equilibria (ex Table S5, *cf.* Supporting Information) - note that the X-axis in this graph labels the chain length of the oligosaccharide with the sugar position of the IdoAS unit in brackets (proceeding from the reducing group in the saccharide oligomer). Bars are color-coded to indicate the IdoAS's neighbors sulfation pattern (expressed as the number of sulfate groups on the saccharide flanking IdoAS at its non-reducing (NR) end, followed by the number of sulfate groups of the saccharide at the IdoAS reducing (R) end).

for the fragments in *e.g.* Table 8 and Table S4 (Supporting Information) are a consequence of a subtle mix of non-distinct steric, electrostatic, and/or solvation effects.

3.2.5. O1,O4-disubstituted-2-O-sulfo-α-L-iduronic acid (IdoAS-1,4R) derivatives

A corpus of 63 coupling constants sets for the idopyranosyl ring in heparin-type oligosaccharides was obtained from the literature. Anticipating the broadly accepted [26,28,71] premise that the IdoAS conformational equilibrium depends *inter alia* on the sulfate substituent patterns of (Glc) saccharide units flanking the IdoAS-moiety of interest at both the reducing (O1) and non-reducing (O4) end, the coupling constants were split into smaller clusters on basis of the number of sulfate groups on the IdoAS's direct neighbors. Somewhat unexpectedly, the distinctive clusters of IdoAS derivatives with up to four neighboring sulfate substituents all came up with two-state equilibria encapsulating virtually identical $^1C_4$ and ^2S_O conformations (deviations in TFP coordinates 1°–2°). This finding ushered us in amalgamating these clusters and re-optimize the resulting collection of 45 coupling constant sets. This compounded cluster is dubbed "≤4 sulfates"; full details of its optimization are listed in Table S5, *cf.* Supporting Information.

The remaining 18 coupling constant sets comprise IdoAS derivatives with four or more neighboring sulfates with at least one located on either neighboring O3. Indicative optimizations showed that these coupling constant sets were best analyzed into two separate clusters: one capturing the IdoAS derivatives for which the neighboring saccharides carry 4 or more sulfate groups with one sulfate on either of the O3's (designated "4+ sulfates", *cf.* Table S6, Supporting Information) and a second cluster hosting the IdoAS derivatives flanked by neighbors carrying 3 sulfate groups each ("6 sulfates", *cf.* Table S7, Supporting Information).

The three clusters of IdoAS-1,4R coupling constants were optimized as a function of a two-state conformational equilibrium involving the single $^1C_4$ conformer used in all previous optimizations and a cluster-specific ^2S_O-like conformation. Table 9 specifies the thus obtained conformers in terms of their TFP coordinates. The table shows that conformers of the first two clusters are closely related, but where the Q and Θ values found for these clusters are within each other's calculated

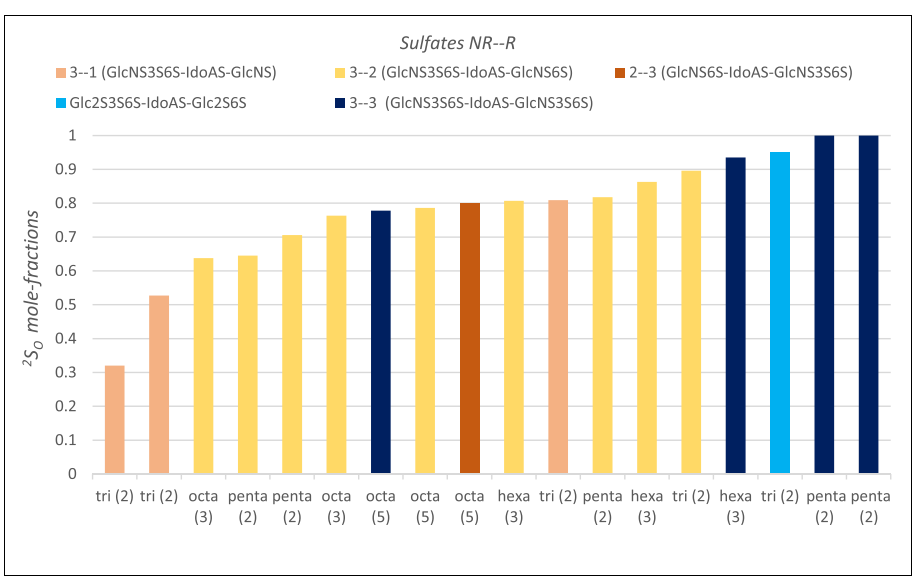


Fig. 10. Graphic representation of the optimized ^2S_O mole-fractions calculated for the "4+ sulfates" and "6+ sulfates" clusters (*cf.* Tables S6 and S7, *cf.* Supporting Information). For further details, see legend to Fig. 9.

Table 10
Relation between the sulfation pattern of the saccharides flanking the central IdoAS moiety and the two-state conformer equilibrium that central IdoAS is involved in: the three main oligomer clusters mentioned in the text are subdivided in groups according to the number of sulfate groups on the saccharides at the non-reducing (NR) end as well as the reducing (R) end of the IdoAS moiety.

NR—R sulfate pattern	# of entries	^2S_O mole-fraction		
		min	max	average
≤ 4 sulfates cluster:				
1/2–0 (typically GlcNS[6S]-IdoAS-OR)	7	0.43	0.61	0.507
1–1 (typically GlcNS-IdoAS-GlcNS)	11	0.30	0.60	0.441
2–1/1–2/1–1[NAc]	11	0.40	0.67	0.589
2–2 (typically GlcNS6S-IdoAS-GlcNS6S)	16	0.46	0.65	0.584
4+ sulfates cluster:				
3–1 (GlcNS3S6S-IdoAS-GlcNS)[a]	3	0.32	0.81[b]	0.552
3–2 (GlcNS3S6S-IdoAS-GlcNS6S)	9	0.64	0.90	0.769
2–3 (GlcNS6S-IdoAS-GlcNS3S6S)	1			0.800
3–2 (Glc2S3S6S-IdoAS-Glc2S6S)	1			0.949
6 sulfates cluster:				
3–3 (GlcNS3S6S-IdoAS-GlcNS3S6S)	4	0.78	1.0	0.928

[a] Span and average for these entries distorted by measurements under high salt conditions, see text.
[b] Determined from couplings recorded at low salt concentration.

errors (~0.8°), the 4.1° difference between the two skew conformers' P_2 is definitely outside the span of the calculated errors for this parameter (~0.5°). The calculated errors for the TFP coordinates of the third cluster ("6 sulfates") are slightly higher (1.0–1.3°) but nevertheless they mark this cluster as clearly distinct from both the other two.

Thus, it is inferred that the skew conformer in the two-state six-membered ring equilibria of IdoAS-1,4R derivatives can be typified as a ^2S_O-like conformation (P_2 = 305°), which creeps to P_2 = 301° upon additional O3 sulfation of one of the flanking saccharides and P_2 = 296° in case both neighboring sugar moieties are sulfated at O3. The overall differences between these three skew structures are quite modest (*cf.* Fig. 8 which offers a view of skews with a comparable difference in P_2), but bring about large enough perturbations in the endocyclic proton-proton torsion angles to be picked up in a robust but sensitive vicinal NMR coupling constants analysis as delivered by the TFP/CAGPLUS approach.

Fig. 9 and Fig. 10 provide insight into the two-state conformational equilibria in which the IdoAS-1,4R derivatives are involved. The bar charts in these figures depict the mole-fractions found for the data clusters at hand and they are sorted on increasing ^2S_O population: the oligomer saccharide chain length and the position of the IdoAS-moiety in that chain is given at the bottom of each bar; furthermore, the bars are colored in accordance to the sulfation pattern of the saccharides neighboring the IdoAS unit (*cf.* legend to Fig. 9).

The so-called "≤4 sulfates" cluster in Fig. 9 is noted for the 30%–70% spread that is observed for the ^2S_O mole-fractions, *i.e.* the equilibria vary within just ± 2 kJ/mol in terms of conformational energy. A cursory look at the X-axis may suggest the ^2S_O conformer having the upper hand in the shorter oligomers, but we found it too hard to forge that into a direct correlation between the ^2S_O population and the oligomer chain length and/or the IdoAS position therein.

The influence of the sulfate patterns of the flanking saccharides on the IdoAS pyranosyl conformer distribution is somewhat better to grasp: the pattern of the bar colors in Fig. 9 looks somewhat lopsided thereby indicating that a higher degree of sulfation is correlated with a higher ^2S_O mole fraction. This indication is to some extent substantiated by the quantifications presented in Table 10: the averaged ^2S_O mole-fractions show indeed variation for the four subclusters in which we categorized the IdoAS units collected in the "≤4 sulfates" cluster according to the sulfates on the neighboring saccharides. However, the differences in the averaged values are small and the ranges over which the equilibria are spread overlap considerably. In fact, the difference between the average for 3 and 4 neighboring sulfates is statistically insignificant (p 0.05) and the subcluster with 2 sulfates displays a lower average for the ^2S_O mole-fraction than the subcluster with 1 neighboring sulfate. As such, the pursued correlation between oligomer sulfate pattern and idopyranose conformer equilibrium is quite weak and hard to translate in terms of causal molecular features.

Contrastingly, Fig. 10 and Table 10 purport to show that in the case of additional O3 sulfation on either residue ("4+ sulfates") or both residues ("6 sulfates") flanking the IdoAS moiety not only provoke a slight pseudorotation in the ^2S_O structure (*cf.* P_2 values in Table 9), but also the equilibrium constants shift markedly towards favoring the ^2S_O side. Exception to this generalization are the two left-most entries in Fig. 10, but these were recorded for a specific trisaccharide under very high salt conditions that are presumed [72] to modulate the conformational equilibrium by screening electrostatic repulsion between the sulfated GlcN residues neighboring the central IdoAS moiety. This rationale is in line with the observation that the same trisaccharide measured at low salt concentrations (marked with superscript "b" in Table 10) is compliant with the remaining entries in the "4+ sulfates" cluster.

4. Conclusions

In the present study we have demonstrated that the application of the Truncated Fourier Puckering formalism [36] combined with the generalized [25] Karplus equation yields a superior quantitative picture of the solution conformations of the six-membered rings in idopyranosyl derivatives. The method is shown to be numerically robust and unbiased ("model-free") as it is able to detect and to quantitate multi-state six-membered ring equilibria from a large body of 636 vicinal NMR coupling constants without recourse to presumed constituent conformations. It is shown that especially the last quality makes the TFP/CAGPLUS approach particularly well suited to uncover subtle deformations in the six-membered idopyranose ring that remained hitherto undetected.

Starting from single-crystal X-ray data we were able to derive at a quantitative description of the constituent idopyranose conformers in solution in terms of their TFP coordinates. The two $^1C_4$ and $^4C_1$ chair conformations turned out to be virtually invariant throughout our analyses, but the third, boat-like conformation turned out to be essentially correlated to the primary substituent pattern of the IdoA ring by "rocking" (or more scientifically correct: pseudorotating) from a $B_{3,O}$-like conformation for O1,O2-disubstituted iduronic derivatives to a ^2S_O conformation for O1,O4-disubstituted idopyranose rings, with an in-between skew/boat conformation for the IdoA singly substituted at O1 (IdoA-1R). The $B_{3,O}$ conformation induced by O2 sulfation is overridden by additional substitution of the idopyranose ring at O4 as is evident from the identical ^2S_O conformations found for IdoA-1,4R and IdoAS-1,4R derivatives. Sulfation patterns of glucosamine residues flanking a central idopyranosyl ring do not influence any of the idopyranose conformations with the exception of sulfate substituents at O3 of one or both of the neighboring saccharides: only in that case the skew-boat conformer of the central IdoAS ring pseudorotates to slightly lower P_2 values, but the ring characterization remains in essence ^2S_O.

The corresponding conformational equilibria display a similar divide: the O1 substituted iduronic acid derivatives (IdoA-1R) are involved in a three-state equilibrium with the remarkable feature that the ^2S_O mole-fraction remains practically constant over a wide range of experimental conditions whereby the $^1C_4$ and $^4C_1$ mole-fractions numerically cross over to one another. This pattern does not change principally upon O2 sulfation albeit that over the whole range of experimental conditions the $^4C_1$ mole-fraction in these IdoAS-1R derivatives is definitely diminished. O4 substitution appears to raise the (energy) bar even further whereby the $^4C_1$ mole-fraction is effectively wiped out, leaving the iduronic acid residue in the middle (IdoA[S]-

1,4R) only to be engaged in a two-state $^1C_4/^2S_O$ equilibrium in which the O2 sulfation of the central IdoA appears to make little difference. As the glycosidic linkages in heparin appear to be relatively stiff [90], the in this paper determined small energy differences of a mere few kJ/mol between the constituent conformers set the iduronic moiety up as the linchpin of heparin flexibility.

From these analyses we conclude that sulfate patterns of saccharides flanking the IdoAS moiety influence the idopyranose conformational equilibrium only marginally. Only sulfate substituents at O3 of one or both of the neighboring glucosamines appear to have a distinct effect in that the ^2S_O-like idopyranose ring pseudorotates to slightly lower P_2 values and is significantly more populated. As such, the relatively rare, but biologically significant glucosamine O3 sulfation [91] appears to be the only sulfate modification in the heparin chain that exerts a distinct albeit small conformational change to the IdoA[S] pyranosyl ring. Overall the effects of sulfate groups on flanking glucosamine saccharides seem scant, both in structural and in equilibrium (*i.e.,* energy) terms and, moreover, any such conformational leeway thus created is expected to be easily absorbed in the overall shape of the heparin macromolecule [90].

Thus, on this basis it seems excessive to assign the idopyranose a steering or driving role in the Sugar Code or Sulfation Code other than its flexibility which remains intact under practically all circumstances. As such, the idopyranose might rather be likened to a "space(r)" than a "letter" [8,92] in the Sugar Code alphabet.

Dedication

In memory of Professor Dr. C. Altona (1931–2008).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Dr. L. van den Bos for his interest in this project.

Appendix A. Supplementary Information

Supplementary information to this article can be found online at https://doi.org/10.1016/j.carres.2020.108052

References

[1] U. Lindahl, M. Höök, Glycosaminoglycans and their binding to biological macromolecules, Annu. Rev. Biochem. 47 (1978) 385–417.
[2] In the present discussion we forego the notion that most GAGs are normally attached to a protein core to form proteoglycans.
[3] R.L. Jackson, S.J. Busch, A.D. Cardin, Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes, Physiol. Rev. 71 (1991) 481–539.
[4] S. Yamada, K. Sugahara, S. Özbek, Evolution of glycosaminoglycans, Commun. Integr. Biol. 4 (2011) 150–158.
[5] F.E. Poulain, H.J. Yost, Heparan sulfate proteoglycans: a sugar code for vertebrate development? Development 142 (2015) 3456–3467, https://doi.org/10.1242/dev.098178.
[6] C.I. Gama, L.C. Hsieh-Wilson, Chemical approaches to deciphering the glycosaminoglycan code, Curr. Opin. Chem. Biol. 9 (2005) 609–619.
[7] U. Lindahl, J.-P. Li, Interactions between heparan sulfate and proteins - design and functional implications, Int. Rev. Cell Mol. Biol. 276 (2009) 105–159.
[8] H.-J. Gabius, The sugar code: why glycans are so important, Biosystems 164 (2018) 102–111.
[9] T.R. Rudd, M.A. Skidmore, M. Guerrini, M. Hricovini, A.K. Powell, G. Siligardi, E.A. Yates, The conformation and structure of GAGs: recent progress and perspectives, Curr. Opin. Struct. Biol 20 (2010) 567–574.
[10] A.B. Souza-Fernandez, P. Pelosi, P.R.M. Rocco, Bench-to-bedside review: the role of glycosaminoglycans in respiratory disease, Crit. Care 10 (2006) 237, https://doi.org/10.1186/cc5069.
[11] A. Malmström, B. Bartolini, M.A. Thelin, B. Pacheco, M. Maccaran, Iduronic acid in chondroitin/dermatan sulfate: biosynthesis and biological function, J. Histochem. Cytochem. 60 (2012) 916–925.
[12] M.A. Thelin, B. Bartolini, J. Axelsson, R. Gustafsson, E. Tykesson, E. Pera, A. Oldberg, M. Maccarana, A. Malmstrom, Biological functions of iduronic acid in chondroitin/dermatan sulfate, FEBS J. 280 (2013) 2431–2446.
[13] J. Jia, M. Maccarana, X. Zhang, M. Bespalov, U. Lindahl, J.-P. Li, Lack of L-iduronic acid in heparan sulfate affects interaction with growth factors and cell signaling, J. Biol. Chem. 284 (2009) 15942–15950.
[14] M. Yanagishita, A brief history of proteoglyeans, Experientia 49 (1993) 366–368.
[15] *In casu* the polysaccharide discovered by W.H. Howell in the early 1920s, *cf.* J.A. Marcum, the Origin of the Dispute over the Discovery of Heparin, J. Hist. Med. 55 (2000) 37–66.
[16] J. McLean, The thromboplastic action of cephalin, Am. J. Physiol 41 (1916) 250–257.
[17] J.A. Cifonelli, A. Dorfman, The uronic acid of heparin, Biochem. Biophys. Res. Commun. 7 (1962) 41–45.
[18] P. Hedenius, O. Wilander, The influence of intravenous injections of heparin in man on the time of coagulation, Acta Med. Scan. LXXXVIII (1936) 443–449.
[19] L. Thunberg, G. Bäckström, U. Lindahl, Further characterization of the antithrombin-binding sequence in heparin, Carbohydr. Res. 100 (1982) 393–410.
[20] M. Petitou, C.A.A. van Boeckel, A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? Angew. Chem. Int. Ed *43* (2004) *3118–3133*.
[21] G. Gatti, B. Casu, G.K. Hamer, A.S. Perlin, Studies on the conformation of heparin by 1H and 13C NMR spectroscopy, Macromolecules 12 (1979) 1001–1007.
[22] D.A. Rees, E.R. Morris, J.F. Stoddart, E.S. Stevens, Controversial glycosaminoglycan conformations, Nature 317 (1985) 480.
[23] B. Casu, J. Choay, D.R. Ferro, G. Gatti, J.-C. Jacquinet, M. Petitou, A. Provasoli, M. Ragazzi, P. Sinaÿ, G. Torri, Controversial glycosaminoglycan conformations, Nature 322 (1986) 215–216.
[24] D.R. Ferro, A. Provasoli, M. Ragazzi, G. Torri, B. Casu, G. Gatti, J.-C. Jacquinet, P. Sinaÿ, M. Petitou, J. Choay, Evidence for conformational equilibrium of the sulfated L-iduronate residue in heparin and in synthetic heparin mono- and oligosaccharides: NMR and force-field studies, J. Am. Chem. Soc. 108 (1986) 6773–6778.
[25] C.A.G. Haasnoot, F.A.A.M. de Leeuw, C. Altona, The relationship between proton-proton NMR coupling constants and substituent electronegativities-I: an empirical generalization of the Karplus equation, Tetrahedron 36 (1980) 2783–2792.
[26] B. Casu, M. Petitou, M. Provasoli, P. Sinaÿ, Conformational flexibility: a new concept for explaining binding and biological properties of iduronic acid-containing glycosaminoglycans, TIBS 13 (1988) 221–225.
[27] N.S. Gandhi, R.L. Mancera, The structure of glycosaminoglycans and their interactions with proteins, Chem. Biol. Drug Des. 72 (2008) 455–482.
[28] P.-H. Hsieh, D.F. Thieker, M. Guerrini, R.J. Woods, J. Liu, Uncovering the relationship between sulphation patterns and conformation of iduronic acid in heparan sulphate, Sci. Rep. 6 (2016) 29602.
[29] C.A.A. van Boeckel, S.F. van Aelst, G.N. Wagenaars, J.-R. Mellema, H. Paulsen, T. Peters, A. Pollex, V. Sinnwell, Conformational analysis of synthetic heparin-like oligosaccharides containing α-L-idopyranosyluronic acid, Recl. Trav. Chim. Pays-Bas 106 (1987) 19–29.
[30] P.N. Sanderson, T.N. Huckerby, I.A. Nieduszynski, Conformational equilibria of α-L-iduronate residues in disaccharides derived from heparin, Biochem. J. 243 (1987) 175–181.
[31] S. Pérez, A. Imberty, S.B. Engelsen, J. Gruza, K. Mazeau, J. Jimenez-Barbero, A. Poveda, J.-F. Espinosa, B.P. van Eyck, G. Johnson, A.D. French, M.L.C.E. Kouwijzer, P.D.J. Grootenhuis, A. Bernardi, L. Raimondi, H. Senderowitz, V. Durier, G. Vergoten, K. Rasmussen, A comparison and chemometric analysis of several molecular mechanics force fields and parameter sets applied to carbohydrates, Carbohydr. Res. 314 (1998) 141–155.
[32] I.Y. Kanal, J.A. Keith, G.R. Hutchison, A sobering assessment of small-molecule force field methods for low energy conformer predictions, Int. J. Quant. Chem. 118 (2018), https://doi.org/10.1002/qua.25512 e25512.
[33] D.S. Sholl, J.A. Steckel, Density Functional Theory - A Practical Introduction, John Wiley & Sons, Inc, 2009, p. 211.
[34] Y. Kurihara, K. Ueda, An investigation of the pyranose ring interconversion path of α-L-idose calculated using density functional theory, Carbohydr. Res. 341 (2006) 2565–2574.
[35] C.A.G. Haasnoot, Conformational analysis of six-membered rings in solution: ring puckering coordinates derived from vicinal NMR proton-proton coupling constants, J. Am. Chem. Soc. 115 (1993) 1460–1468.
[36] C.A.G. Haasnoot, The conformation of six-membered rings described by puckering coordinates derived from endocyclic torsion angles, J. Am. Chem. Soc. 114 (1992) 882–887.
[37] C.I. Gama, S.E. Tully, N. Sotogaku, P.M. Clark, M. Rawat, N. Vaidehi, W.A. Goddard III, A. Nishi, L.C. Hsieh-Wilson, Sulfation patterns of glycosaminoglycans encode molecular recognition and activity, Nat. Chem. Biol 2 (2006) 467–473.
[38] D. Soares da Costa, R.L. Reis, I. Pashkuleva, Sulfation of glycosaminoglycans and its implications in human health and disorders, Annu. Rev. Biomed. Eng 19 (2017) 1–26.
[39] C.R. Groom, I.J. Bruno, M.P. Lightfoot, S.C. Ward, The Cambridge Structural Database, Acta Crystallogr. B72 (2016) 171–179.
[40] G.H. Petit, J. Dillen, H.J. Geise, Limitations of Cremer & Pople parameters as quantitative descriptors of ring shapes in comparative stereochemistry, Acta Crystallogr. B39 (1983) 648–651.

[41] J.B. Hendrickson, Molecular geometry. VII. Modes of interconversion in the medium rings, J. Am. Chem. Soc. 89 (1967) 7047–7061.
[42] D. Cremer, J.A. Pople, A general definition of ring puckering coordinates, J. Am. Chem. Soc. 97 (1975) 1354–1358.
[43] IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Symbols for specifying the conformation of polysaccharide chains (Recommendations 1981), Pure Appl. Chem. 55 (1983) 1269–1272.
[44] IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). NOMENCLATURE OF CARBOHYDRATES, Pure Appl. Chem. 68 (1996) 1919–2008.
[45] S. Makeneni, B.L. Foley, R.J. Woods, BFMP: a method for discretizing and visualizing pyranose conformations, J. Chem. Inf. Model. 54 (2014) 2744–2750.
[46] M.J. Hadad, W. Zhang, T. Turney, L. Sernau, X. Wang, R.J. Woods, A. Incandela, I. Surjancev, A. Wang, M.-K. Yoon, A. Coscia, C. Euell, R. Meredith, I. Carmichael, A.S. Serianni, CHAPTER 2. NMR spin-couplings in saccharides: relationships between structure, conformation and the magnitudes of JHH, JCH and JCC values, in: K. Kato, T. Peters (Eds.), NMR in Glycoscience and Glycotechnology, Royal Society of Chemistry, 2017, pp. 20–100.
[47] A.A. Clifford, Multivariate Error Analysis, Applied Science Publishers Ltd., 1973London.
[48] P.J. Hajduk, D.A. Horita, L.E. Lerner, Picosecond dynamics of simple monosaccharides as probed by NMR and molecular dynamics simulations, J. Am. Chem. Soc. 115 (1993) 9196–9201.
[49] M.J. Forster, B. Mulloy, Molecular dynamics study of lduronate ring conformation, Biopolymers 33 (1993) 575–588.
[50] Y. Chevalier, Measurement of coupling constants in broad multiplets, Magn. Reson. Chem. 24 (1986) 404–408.
[51] B.M. Sattelle, S.U. Hansen, J. Gardiner, A. Almond, Free energy landscapes of iduronic acid and related monosaccharides, J. Am. Chem. Soc. 132 (2010) 13132–13134.
[52] W. Zhang, X. Hu, I. Carmichael, A.S. Serianni, Methyl [^{13}C]glucopyranosiduronic acids: effect of COOH ionization and exocyclic structure on NMR spin-couplings, J. Org. Chem. 77 (2012) 9521–9534.
[53] C.A. Podlasek, J. Wu, W.A. Stripe, P.B. Bondo, A.S. Serianni, [^{13}C]-Enriched methyl aldopyranosides: structural interpretations of ^{13}C-^{1}H spin-coupling constants and ^{1}H chemical shifts, J. Am. Chem. Soc. 117 (1995) 8635–8644.
[54] M.L.D. Ramos, M.M.M. Caldeira, V.M.S. Gil, NMR Study of uronic acids and their complexation with molybdenum(VI) and tungsten(VI) oxoions, Carbohydr. Res. 286 (1996) 1–15.
[55] M.U. Roslund, K.D. Klika, R.L. Lehtilä, P. Tähtinen, R. Sillanpää, R. Leino, Conformation of the galactose ring adopted in solution and in crystalline form as determined by experimental and DFT ^{1}H-NMR and single-crystal X-ray analysis, J. Org. Chem. 69 (2004) 18–25.
[56] D.A. Horita, P.J. Hajduk, L.E. Lerner, Solution dynamics of the 1,2,3,4,6-penta-O-acetyl-α-D-idopyranose ring, Glycoconj. J. 14 (1997) 691–696.
[57] S. Ernst, G. Venkataraman, V. Sasisekharan, R. Langer, C.L. Cooney, R. Sasisekharan, Pyranose ring flexibility. Mapping of physical data for iduronate in continuous conformational space, J. Am. Chem. Soc. 120 (1998) 2099–2107.
[58] T. Klepach, H. Zhao, X. Hu, W. Zhang, R. Stenutz, M. Hadad, I. Carmichael, A.S. Serianni, Informing saccharide structural NMR studies with density functional theory calculations, in: T. Lütteke, M. Frank (Eds.), Glycoinformatics. Methods in Molecular Biology, vol. 1273, Humana Press, New York, NY, 2015, pp. 289–331.
[59] M. Hricovíni, F. Bízik, Relationship between structure and three-bond proton–proton coupling constants in glycosaminoglycans, Carbohydr. Res. 342 (2007) 779–783.
[60] S.U. Hansen, C.E. Dalton, M. Baráth, G. Kwan, J. Raftery, G.C. Jayson, G.J. Miller, J.M. Gardiner, Synthesis of L-iduronic acid derivatives via [3.2.1] and [2.2.2] L-iduronic lactones from bulk glucose-derived cyanohydrin hydrolysis: a reversible conformationally switched superdisarmed/rearmed lactone route to heparin disaccharides, J. Org. Chem. 80 (2015) 3777–3789.
[61] As provided by Spartan, Wavefunction, Inc., Irvine, CA.
[62] P. Ochsenbein, M. Bonin, K. Schenk-Joß, M. El-Hajji, The $^{2}S_{O}$ skew-boat conformation in L-iduronic acid, Angew. Chem. Int. Ed 50 (2011) 11637–11639.
[63] V. Babin, C. Sagui, Conformational free energies of methyl-α-L-iduronic and methyl-α-D-glucuronic acids in water, J. Chem. Phys. 132 (2010), https://doi.org/10.1063/1.3355621 104108.
[64] B.M. Sattelle, B. Bose-Basu, M. Tessier, R.J. Woods, A.S. Serianni, A. Almond, Dependence of pyranose ring puckering on anomeric configuration: methyl idopyranosides, J. Phys. Chem. B 116 (2012) 6380–6386.
[65] P. Oborský, I. Tvaroška, B. Králová, V. Spiwok, Toward an accurate conformational modeling of iduronic acid, J. Phys. Chem. B 117 (2013) 1003–1009.
[66] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.
[67] O.S. Smart, V. Horský, S. Gore, R.S. Vařeková, V. Bendová, G.J. Kleywegt, S. Velankara, Validation of ligands in macromolecular structures determined by X-ray crystallography, Acta Cryst. D74 (2018) 228–236.
[68] O.S. Smart, G. Bricogne, Chapter 13-achieving high quality ligand chemistry in protein-ligand crystal structures for drug design, in: G. Scapin, D. Patel, E. Arnold (Eds.), Multifaceted Roles of Crystallography in Modern Drug Discovery, NATO Science for Peace and Security Series A: Chemistry and Biology, Springer, 2015, pp. 165–181.
[69] B. Mulloy, M.J. Forster, C. Jones, D.B. Davies, N.m.r. and molecular-modelling studies of the solution conformation of heparin, Biochem. J. 293 (1993) 849–858.
[70] Z. Zhang, S. McCallum, J. Xie, L. Nieto, F. Corzana, J. Jiménez-Barbero, M. Chen, J. Liu, R.J. Linhardt, Solution structures of chemoenzymatically synthesized heparin and its precursors, J. Am. Chem. Soc. 130 (2008) 12998–13007.
[71] V.H. Pomin, Solution NMR conformation of glycosaminoglycans, Prog. Biophys. Mol. Biol. 114 (2014) 61–68.
[72] D.R. Ferro, A. Provasoli, M. Ragazzi, B. Casu, G. Torri, V. Bossenec, B. Perly, P. Sinaÿ, M. Petitou, J. Choay, Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences, Carbohydr. Res. 195 (1990) 157–167.
[73] J.C. Muñoz-García, F. Corzana, J.L. de Paz, J. Angula, P.M. Nieto, Conformations of the iduronate ring in short heparin fragments described by time-averaged distance restrained molecular dynamics, Glycobiology 23 (2013) 1220–1229.
[74] Y. Inoue, Y. Inouye, K. Nagasawa, Conformational equilibria of the L-iduronate residue in non-sulphated di-, tetra- and hexa-saccharides and their alditols derived from dermatan sulphate, Biochem. J. 265 (1990) 533–538.
[75] Conformational energy differences are calculated from conformer populations (p_i) using $\Delta E_{ij} = RT \ln (p_i/p_j)$.
[76] B. Bose-Basu, W. Zhang, J.L.W. Kennedy, M.J. Hadad, I. Carmichael, A.S. Serianni, ^{13}C-Labeled idohexopyranosyl rings: effects of methyl glycosidation and C6 oxidation on ring conformational equilibria, J. Org. Chem. 82 (2017) 1356–1370.
[77] J.R. Snyder, A.S. Serianni, D-Idose: a one- and two-dimensional NMR investigation of solution composition and conformation, J. Org. Chem. 51 (1986) 2694–2702.
[78] See e.g.
(a) M. Ragazzi, D.R. Ferro, A. Provasoli, A force-field study of the conformational characteristics of the iduronate ring, J. Comput. Chem. 7 (1986) 105–112;
(b) E. Scholtzová, P. Mach, M. Hricovíni, Structure of sulfated monosaccharides studied by quantum chemical methods, Molecules 8 (2003) 770–779;
(c) M. Hricovíni, B3LYP/6-311 + +G** study of structure and spin–spin coupling constant in methyl 2-O-sulfo-α-L-iduronate, Carbohydr. Res. 341 (2006) 2575–2580;
(d) L. Pol-Fachin, H. Verli, Depiction of the forces participating in the 2-O-sulfo-α-L-iduronic acid conformational preference in heparin sequences in aqueous solution, Carbohydr. Res. 343 (2008) 1435–1445;
(e) K.J. Murphy, N. McLay, D.A. Pye, Structural studies of heparan sulfate hexasaccharides: new insights into iduronate conformational behavior, J. Am. Chem. Soc. 130 (2008) 12436–12444;
(f) N.S. Gandhi, R.L. Mancera, Can current force fields reproduce ring puckering in 2-O-sulfo-α-L-iduronic acid? A molecular dynamics simulation study, Carbohydr. Res. 345 (2010) 689–695
[79] G.G. Lyle, R.E. Lyle, The boat form of cyclohexane as viewed by Midwestern sailors, J. Chem. Educ. 50 (1973) 655.
[80] Note that this discussion regarding the effect of sulfation is valid for solutions in D_2O at ambient temperature only – changes in experimental conditions (e.g. temperature, solvent, pH) give rise to vastly different outcomes.
[81] E.L. Eliel, N.L. Allinger, S.J. Angyal, G.A. Morrison, Conformational Analysis, Interscience Publishers, 1965, pp. 43–47.
[82] C. Altona, C. Romers, E. Havinga, Molecular structure and conformation of some dihalogenodioxanes, Tetrahedron Lett. 1 (1959) 16–20.
[83] K.B. Wiberg, W.F. Bailey, K.M. Lambert, Z.D. Stempel, The anomeric effect: it's complicated, J. Org. Chem. 83 (2018) 5242–5255.
[84] M.D. Walkinshaw, Variation in the hydrophilicity of hexapyranose sugars explains features of the anomeric effect, J. Chem. Soc. Perkin Trans. I (1987) 1903–1906.
[85] B. Mulloy, M.J. Forster, C. Jones, A.F. Drake, E.A. Johnson, D.B. Davies, The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modelling study, Carbohydr. Res. 255 (1994) 1–26.
[86] R. Raman, V. Sasisekharan, R. Sasisekharan, Structural insights into review biological roles of protein-glycosaminoglycan interactions, Chem. Biol. 12 (2005) 267–277.
[87] It is deemed unlikely that the pseudorotational shift is a calculational artefact as the $^{2}S_{O}$ conformer is dominant in the relevant equilibria and therefore carries a relatively large weight in the calculated time-averaged couplings. This is also borne out by the propagated errors [47] in the P_2 coordinates which are calculated to be 0.8°-1.4°.
[88] G. Venkataraman, V. Sasisekharan, C.L. Cooney, R. Langer, R. Sasisekharan, A stereochemical approach to pyranose ring flexibility: its implications for the conformation of dermatan sulfate, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 6171–6175.
[89] A. Raghuraman, P.D. Mosier, U. Desai, Understanding dermatan sulfate-heparin cofactor II interaction through virtual library screening, ACS Med. Chem. Lett. 1 (2010) 281–285.
[90] B. Mulloy, M.J. Forster, Conformation and dynamics of heparin and heparan sulfate, Glycobiology 10 (2000) 1147–1156.
[91] B.E. Thacker, D. Xu, R. Lawrence, J.D. Esko, Heparan sulfate 3-O-sulfation: a rare modification in search of a function, Matrix Biol. 35 (2014) 60–72.
[92] H. Kaltner, J. Abad-Rodríguez, A.P. Corfield, J. Kopitz, H.-J. Gabius, The sugar code: letters and vocabulary, writers, editors and readers and biosignificance of functional glycan-lectin pairing, Biochem. J. 476 (2014) 2623–2655.