Carbohydrate Polymers
January 2022, Volume 276 Pages 118732 (11p.)
https://doi.org/10.1016/j.carbpol.2021.118732
https://archimer.ifremer.fr/doc/00730/84248/

Three-dimensional structures, dynamics and calcium-mediated interactions of the exopolysaccharide, Infernan, produced by the deep-sea hydrothermal bacterium *Alteromonas infernus*

Makshakova Olga [1] , Zykwinska Agata [2], Cuenot Stéphane [3], Colliec-Jouault Sylvia [2], Perez Serge [4, *]

[1] Kazan Institute of Biochemistry and Biophysics, FRC Kazan Scientific Center of RAS, Lobachevsky Str., 2/31, 420111 Kazan, Russian Federation
[2] Ifremer, Laboratoire Ecosystèmes Microbiens et Molécules Marines pour les Biotechnologies, 44311 Nantes, France
[3] Université de Nantes, CNRS, Institut des Matériaux Jean Rouxel, IMN, Nantes, France
[4] Centre de Recherches sur les Macromolécules Végétales, Université de Grenoble Alpes, Centre National de la Recherche Scientifique, Grenoble, France

* Corresponding author : Serge Perez, email address : serge.perez@cermav.cnrs.fr

olga.makshakova@kibb.knc.ru ; Agata.Zykwinska@ifremer.fr ; Stephane.Cuenot@cnrs-imn.fr ; Sylvia.Colliec.Jouault@ifremer.fr

Abstract :

The exopolysaccharide Infernan, from the bacterial strain GY785, has a complex repeating unit of nine monosaccharides established on a double-layer of sidechains. A cluster of uronic and sulfated monosaccharides confers to Infernan functional and biological activities. We characterized the 3-dimensional structures and dynamics along Molecular Dynamics trajectories and clustered the conformations in extended two-fold and five-fold helical structures. The electrostatic potential distribution over all the structures revealed negatively charged cavities explored for Ca2+ binding through quantum chemistry computation. The transposition of the model of Ca2+complexation indicates that the five-fold helices are the most favourable for interactions. The ribbon-like shape of two-fold helices brings neighbouring chains in proximity without steric clashes. The cavity chelating the Ca2+ of one chain is completed throughout the interaction of a sulfate group from the neighbouring chain. The resulting is a ‘junction zone’ based on unique chain-chain interactions governed by a heterotypic binding mode.

Highlights

► Complete determination of the 3D structure of the anionic EPS Infernan by MM and MD calculations ► Clustering of the conformations into families having either extended two-fold or five-fold helical structures ► Identification, geometry and energetics of Ca^{2+} binding cavities via DFT calculations ► Five-fold helical structure as Ca^{2+} a binding reservoir ► Two-fold helices form “junction zones” with Ca^{2+}throughout a heterotypic mode of binding

Keywords : Exopolysaccharides, Alteromonas infernus, 3 dimensional structures, Gel forming, Molecular dynamics, Quantum chemistry, Calcium binding

Introduction

In 1977, discovering the first deep-sea hydrothermal vents on the Galapagos rift changed our view of biology. It raised the question of the ability of vent organisms to survive and thrive in such extreme pressures and temperatures in the presence of toxic mineral plumes. The conversion of mineral-rich hydrothermal fluid into energy is a crucial feature of these environments. Through chemosynthesis, which converts carbon dioxide into sugar in the absence of light, bacteria provide energy and nutrients to vent species without the need for sunlight. (Dick; 2019)

In 1991, the submersible « Nautile » collected samples of vent fluids, chimney rocks, invertebrate tissues, sediments and seawater at a depth of 2000 m in a rift system of the Guaymas basin (Guezennec 2002). The collection took place in the proximity of the active hydrothermal vent named « Claire ». The bacterial strain, labelled as GY785, was isolated from a diluted vent fluid among a dense population of tubeworms *Riftia pachyptila* (Raguénès et al., 1997). Phylogenetic analysis showed that the new bacterium belongs to the gamma subdivision of the *Proteobacterial* phylum. The bacterium belongs to the monophyletic unit that defines the true genus *Alteromonas.* The Gram-negative genus *Alteromonas* belonging to the Alteromonadaceae family (Baumann et al., 1972) was named *Alteromonas infernus* (Raguénès et al., 1997).

This mesophilic gram-negative bacterium GY785 secretes two different exopolysaccharides (EPS), a water-soluble EPS (EPS1) and a non-soluble EPS (EPS2) embedded in a gelatinous matrix containing bacterial cells (Raguénès et al., 1997). The GY785 soluble EPS is called Infernan. Infernan comes from the contraction of 'Infern' (from *infernus*) followed by the suffix 'an', which refers to its nature as glycan. Infernan in its native form is a high-molecular-weight, highly-branched anionic heteropolysaccharide. Structural analyses revealed that its repeating unit consists of nine osidic residues assembled in a complex structure (Figure 1; Roger et al., 2004). Three osidic residues compose the EPS main chain, glucose (Glcp), galacturonic acid (GalpA-2S) and galactose (Galp), which are covalently linked in the following sequence: →4)-β-D-Glcp-(1→4)-α-D-GalpA-2S-(1→4)- α-D-Galp-(1→. The GalpA residue of the main chain carries at O-2, one sulfate group. It is substituted at O-3 by a short side chain constituted of two glucuronic acids (GlcpA), Galp and Glcp linked in the sequence: β-D-Glcp-(1→6)-α-D-Galp-(1→4)-β-D-GlcpA-(1→4)-β-D-GlcpA-(1→. The two GlcpA of the side chain are each substituted by a terminal Glcp.

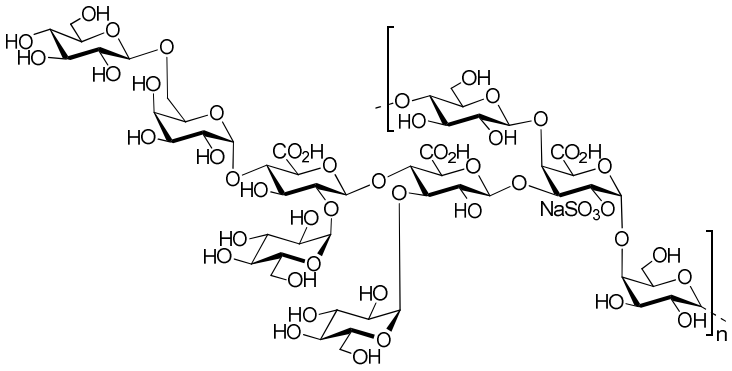

Figure 1. *The repeating unit of Infernan (GY785 EPS) (Roger et al., 2004).*

The chemical composition of Infernan, with uronic acids and sulfate groups and its particular structure, is unique in the occurrence of voluminous side chains linked to a backbone made up of a monosulfated trisaccharide repeat. Among the soluble and capsular EPS structures established from marine bacteria and archaebacteria, Infernan displays the most complex sequence of monosaccharides (Toukach, Egorova, 2016, Casillo et al., 2018; Birch et al., 2019). The occurrence of a cluster of negatively charged monosaccharides confers to this EPS several potential functional properties, such as glycosaminoglycan (GAG) activities, explored in tissue engineering and cancer treatment (Colliec-Jouault et al., 2001; Merceron et al., 2012; Heymann et al., 2016; Rederstorff et al., 2017) as well as gelling properties with calcium (Ca^{2+}), exploited to develop growth factor delivery systems (Zykwinska et al., 2019). The formation of gel from Internan chains in the presence of calcium has been observed and further characterized by AFM (Zykwinska et al., 2019). Also, the high content of uronic acids correlated to the metal-binding ability makes this EPS a good candidate for waste-water treatment, metal recovery and strontium therapy.

Recent advances in the computational modelling of polysaccharides have reached a point where highly complex structures can be investigated through a series of complementary methods (Perez et al., 2021). They allow establishing structures and dynamics as well as inter-macromolecular interactions. As for polysaccharides, it is challenging to sample the conformational state and transitions in solution experimentally because of their high flexibility. The computational methods offer a way to sample both long-living and short-living conformations via molecular dynamic simulations. It enables the description of the processes of chain interactions, which may result in polymeric entanglements and other soft matter state formation. The interactions of linear anionic

polysaccharides containing uronic residues, such as pectins or alginates, with Ca^{2+} and other ions with the eventual chain assembly in the form of "egg-box" have been a topic several computational studies. (Perez 2000, Hecht H, Plazinski 2011, Plazinski, Rudzinski 2012, Stewart, Gray, Vasiljevic, Orbell 2014, Srebnik 2016, Maire du Poset, Zitolo, Cousin, Assifaoui, Lerbret 2020, Cao, Lu, Mata, Nishinari , Fang 2020). Another type of anionic polysaccharide, carrageenans, carry sulfate groups. The carrageenans have a distinct type of chain complexation. Counter-ions mediate them by cross-linking helical chains as molecular dynamic simulations resolve (Perez, Claudio 2020). However, as reported for chondroitin sulfate, a linear polysaccharide from the GAG family that bears both carboxylate and sulfate functional groups, the subtle differences in force field parametrization may have a drastic effect on counter-ion binding and the consequent interactions mediated by them as seen in molecular dynamic simulations (Guvench, Whitmore 2021).

The present article describes for the first time the conformational flexibility of the highly-branched anionic EPS, Infernan, which is characterized by diverse monosaccharides in the backbone and sidechains and bears both carboxylate and sulfate groups. It establishes the calcium-binding sites and the ability for Infernan to form gels in the presence of divalent cations such as Ca^{2+}, which is the key to biological functions and technological applications.

Material and Methods

Conformational descriptors of the constituting monosaccharides. The conformational descriptors of the monosaccharides include the shape of the rings and the orientations of chemical groups. The Φ and Ψ torsion angles dictate the orientation of the two successive monosaccharides linked by a glycosidic bond in a disaccharide. In crystallography, Φ is defined as Φ = O5-C1-O1- Cx, and Ψ as Ψ = C1-O1-Cx-Cx+1, where x is the number of the carbon atom of the second monosaccharide with which the 1-x glycosidic bond is formed. An alternate definition used in NMR spectroscopy refers to the hydrogen atoms with Φ^H = H1-C1-O1- Cx and Ψ^H = C1-O1-Cx-Hx

Molecular mechanics calculations and potential energy surfaces (PES) The geometry optimization of the monosaccharides and the eight disaccharides was performed with the MM3 force field with proper parametrization for carbohydrates containing sulfate groups (Allinger et al. 1989, Le Questel et al. 1995). The calculations and plots of the conformational maps (PHI/PSI maps) were performed using the Conformational Analysis Tool (CAT) software (http://www.md-simulations.de/CAT/, Martin Frank, martin@glycosciences.org, May 2012) with the TINKER (Harger et al., 2017) program to minimize, optimize and run molecular dynamics calculation usually for a 50 ns simulation at 700 K.

Initial models of oligosaccharide models. Four types of oligosaccharides were constructed:

1. The linear backbone of two repeats of the elementary block **(**→4)-β-D-Glcp-(1→4)-α-D-GalpA-(1→4)- α-D-Galp-(1→)$_2$;

2. The backbone of two repeats, branched with one linear side chain α-D-Galp-(1→4)-β-D-GlcpA-(1→4)-β-D-GlcpA-(1→ attached to the O-3 of GalpA ;

3. The side chain branched with terminal glucose residues α-D-Galp-(1→4)-β-D-GlcpA[2 α-D-Glcp](1→4)-β-D-GlcpA[3 α-D-Glcp]-(1→ 3)-α-D-GalpA-2S ;

4. Branching point extended with branched side chain β-D-Glcp- α-D-GalpA-2S[α-D-Galp-(1→4)-β-D-GlcpA[2 α-D-Glcp](1→4)-β-D-GlcpA[3 α-D-Glcp]-(1→ 3)]-α-D-Galp. Such a variety of initial structures is aims at unravelling the long-distance conformational influence of branching.

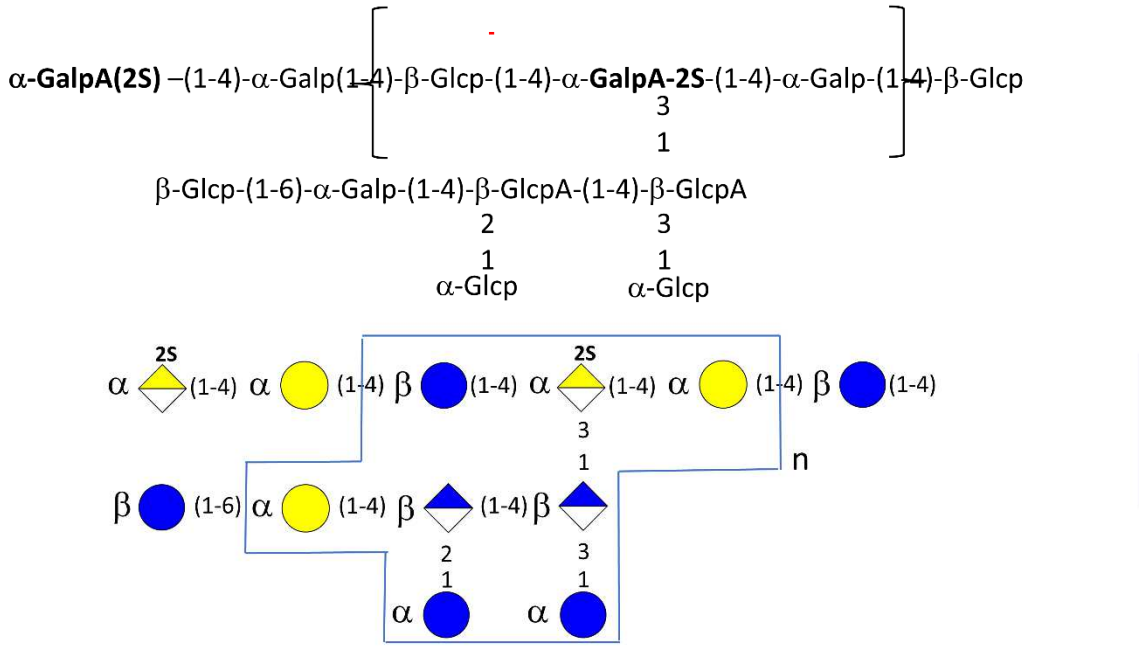

Figure 2. *Schematic representation of the repeating unit of Infernan. The structural elements considered in the present study use the Symbol Nomenclature for Graphical Representations of Glycans (Varki et al., 2015).*

Molecular Dynamics and Mechanics calculations. Molecular Dynamics (M.D.) simulations started from the initial structures immersed in a water box with periodic boundary conditions. Na^+ ions were added to the system to neutralize the negative charges of Infernan. The simulations were carried out using program package AMBER12 (Case et al., 2012) in the isotherm isobar thermodynamic ensemble at 300K, using the force-field parameters for GLYCAM06 for saccharides (Kirschner et al., 2008). The AMBER force field was used for ions and the TIP3 model for water molecules. A time step of 2 fs was applied with the SHAKE algorithm (Ryckaert, Ciccotti, & Berendsen, 1997) constrained bonds involving hydrogen atoms. Particle Mesh Ewald (PME) (Essmann et al., 1995) was used to handle long-range electrostatic interactions. A cut-off value for electrostatic interactions was used 8Å in simulations; a rectangular water box was selected, having dimensions commensurate to the size of the molecule under investigation. The temperature and the pressure were kept constant using a Langevin thermostat with a collision frequency of 2 ps^{-1} and a weak coupling anisotropic algorithm with a relaxation time of 2 ps^{-1}, respectively. The system was optimized stepwise, with restraints on all atoms of the complex gradually decreasing from 20 to 5 kcal mol^{-1} $Å^{-2}$. Each minimization step was followed by 20 ps of molecular dynamics equilibration (NVT, 100K) with the same restraints. The system was then heated from 100 to 300K using a Langevin thermostat with a collision frequency of 5 ps^{-1}. Initial velocities were derived from a Maxwellian distribution at 100K using a random seed for each simulation. Then, ten ns equilibration phases in the NVT and NPT ensemble were then run. Finally, the production phases were performed on totally relaxed systems in the NPT ensemble at 300K with trajectories in the range of 100 ns and output frequency of 10 ps. The Schrödinger Release 2021-3**:** Maestro, Schrödinger, LLC, New York, NY, 2021, was used for Molecular Mechanics calculations,

DFT calculations. The complexes of Infernan with Ca^{2+} ions were calculated to estimate the location of the binding site with a maximum calcium coordination number. The location of putative Ca^{2+}

binding sites derived from the examination of the electrostatic potential of Infernan was calculated using the Avogadro program (Hanwell et al. 2012) (see Figure Supplemental Material). Then, one Ca^{2+} ion was placed in the vicinity of the sulfate group or carboxylate group of αGalpA-2S residue in tetrasaccharide and further optimized. The calculations were carried out at the DFT level using B3LYP hybrid density functional (Becke, 1983; Lee, Yang, & Parr, 1988), widely employed for the description of complexes of biological molecules (Makshakova et al. 2020, Bekri, Zouaoui-Rabah, Springborg, & Sekkal Rahal, 2018; Makshakova, Chachkov, & Ermakova, 2011; Yousfi, Sekkal-Rahal, Sayede, & Springborg, 2010 Cao L, Lu W, Mata A, Nishinari K, Fang Y 2020). Two types of calculations were implemented for all the systems studied: 1. with the 6-31G(d,p) basis set in vacuum and 2. 6-31G+(d,p) with the polarized continuum conditions with SMD variation of integral equation formalism variant (IEFPCM) was used to take the solvation effects into account implicitly (Marenich, Cramer, Truhlar, 2009). All calculations were carried out using the Gaussian09 program (Frisch et al., 2009)

Results

Building and displaying the three-dimensional components of Infernan. A prerequisite to explore and build the three-dimensional structures of Infernan requires the computation of the potential energy surfaces of each constituting disaccharide component. The determination of conformational preferences of the constituting disaccharides was first performed by characterizing their preferred conformations on potential energy surfaces as a function of their glycosidic torsional angles Φ and Ψ. Due to its occurrence as a terminal residue of the side chain, we did not consider the β 1-6 linked Glcp. Its location at the periphery of the repeating unit does not impact the 3D features of Infernan.

The potential energy surfaces of the eight disaccharides occurring in Infernan were built. A typical conformational map displaying the shape of the low energy region (up to 10 kcal/mol concerning the lowest energy minimum), the positions of energy minima, the routes for interconversion between conformers and the heights of the transitional barriers were obtained for each disaccharide. These maps display similar features, i.e., the occurrence of a low energy basin embedding one or two low-energy conformations. (Figure 3)

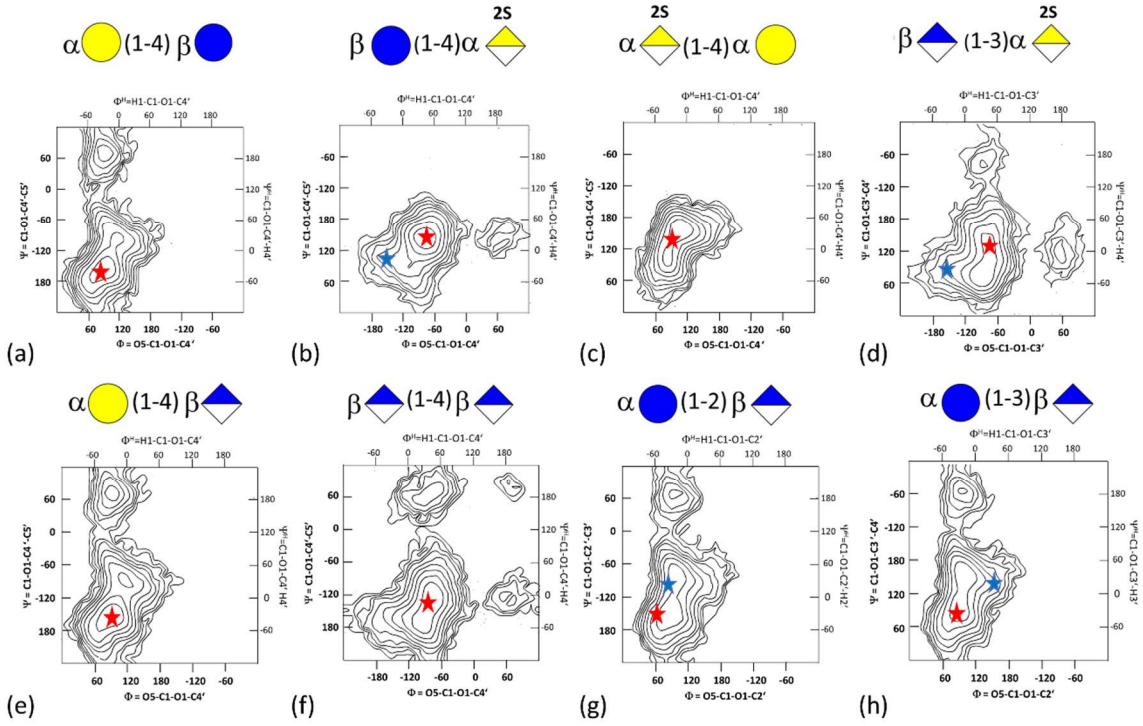

Figure 3. *Potential Energy Surfaces of the constituting disaccharides of Infernan,* (a) α-D-Galp-(1→4)-β-D-Glcp; (b) β-D-Glcp-(1→4)-α-D-GalpA-2S; (c) α-D-GalpA-2S-(1→4)- α-D-Galp; (d) β-D-GlcpA-(1→3)-α-D-GalpA-2S; (e) α-D-Galp-(1→4)-β-D-GlcpA; (f) β-D-GlcpA-(1→4)-β-D-GlcpA; (g) β-D-Glcp-(1→2)-α-D-GalpA; (h) β-D-Glcp-(1→3)-α-D-GalpA. The iso-contours are drawn at 1 kcal intervals with respect to the energy minimum of the corresponding map. The stars indicate the location of the low energy conformers as determined from Molecular Dynamics performed on the whole repeating unit of Infernan (see text for explanation).

Building and display of the three-dimensional structural models of Infernan. The construction of three-dimensional structures of Infernan was performed using the polysaccharide builder program POLYS (Engelsen et al. 2014; Perez & Rivet, 2021). It is based on an algorithm designed to use optimized monosaccharide structures and connect them using established energetically favourable geometries. This approach was previously used and efficiently built reliable complex carbohydrate structures (Perez and Tvaroska 2014). Steric conflicts rarely occur when using energetic minima derived from conformational analysis of disaccharides. To construct the initial structure of Infernan's fragments, the values corresponding to the global minima on the were Potential Energy Surfaces were assigned to the glycosidic linkages.

Conformational variability of polysaccharides. The results confirm that two of the glycosidic junctions**:** αGalpA-2S(1→4)αGalp and αGalp(1→4)βGlcp), within the backbone, do not significantly change during the length of the trajectories. The oscillations around their average values do not exceed 10-20°. In contrast, the βGlcp(1→4)αGalpA-2S linkage oscillates between two conformational states (Figure 4 and Figures S1; S2). The transitions occur in a concerted way, with the correlation coefficient of 0.7 (the correlation coefficient may vary from -1 to +1, the larger the value of the modulus of magnitude is, the more concerted two events are). Therefore, we will consider two conformational states. The most populated state (-74±16; 153±13) has a longer lifetime of about a dozen nanoseconds, occurring 70±15% of the time. The less populated state (-148±12; 114±12) lives for 2-3 ns and a 30±15% occupancy. They will be referred to as Model 1 and Model 2.

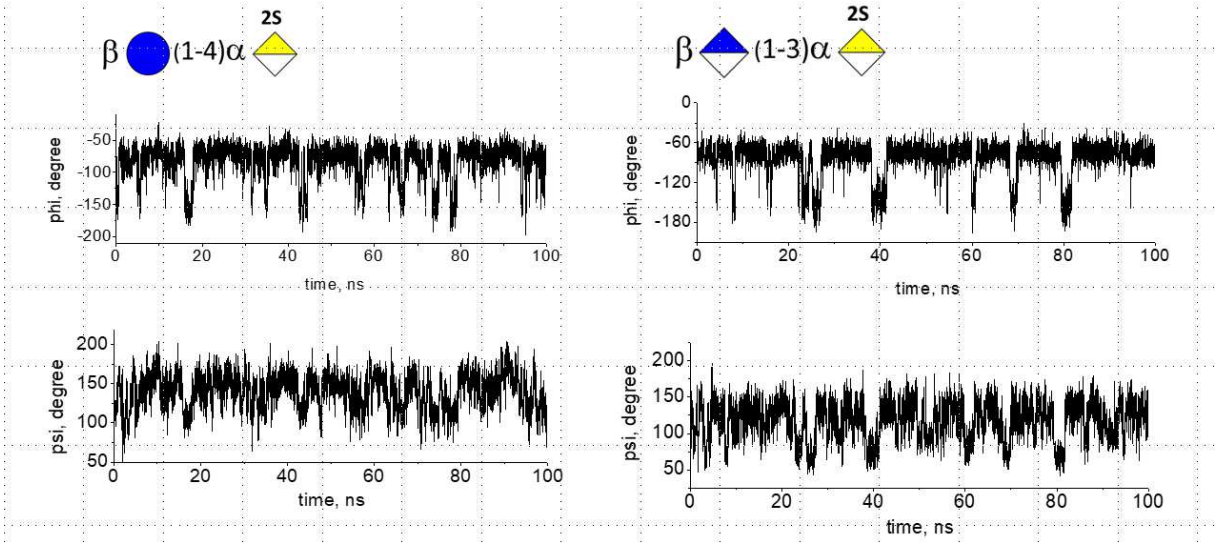

Figure 4. *Evolution of the βGlcp(1-4)αGalpA-2S and βGlcpA(1-3)αGalpA-2S linkages over a 100 ns trajectory.*

The conformations of two of the glycosidic junctions within the side chain (αGalp(1→4)βGlcpA and βGlcpA(1→4)βGlcpA) remain stable and oscillate around their conformation, no more than 10-20°. The conformation of the glycosidic linkage at the branching point (βGlcpA(1-3)αGalpA-2S) experiences transitions between two states (Φ, Ψ)=(-75±12; 133±13) and (Φ, Ψ)= (-149±16; 85±19) that occur with the correlation coefficient 0.7; the respective lifetime being 10ns and 2-3ns (Figure 4). They are referred to as Model A and Model B.

The conformations of the two αGlcp in side chains αGlcp(1→2)βGlcpA and αGlcp(1→3)βGlcpA oscillate between two low energy regions (Figure 3g, 3h), without any influence on the conformational equilibrium of the backbone. Each of them can adopt two sets of Φ,Ψ torsion angles (deduced from correlated changes of Φ and Ψ angles. The population of the two states for both linkages relates as 70±15% to 30±15% (Table S1). The αGlcp(1→3)βGlcpA linkage demonstrates the highest lifetimes of each state compare to other linkages in this polysaccharide. The lifetime of these conformations amounts up to 10 ns. It results from competitive hydrogen bonds between OH-2 and OH-4 groups of αGlcp with one of the O6 of the neighbouring βGlcpA. As the conformational transitions in the backbone and the branching points occur independently, the overall 3D structures of Infernan can be described based on only four models from Models 1 and 2, namely M1A, M1B, M2A and M2B.

Further propagation of the polysaccharide backbone in a repetitive fashion reveals two types of helical conformations. These are two-fold and five-fold helices (Figure 5) based on the results derived from the molecular dynamics simulation. The former is more populated, resulting in 70±15% occupancy over the trajectory. The latter is less occupied presenting only in 30±15% of simulation time. Overall the side chains are oriented, either in the direction of the backbone or perpendicular to it.

The end-to-end distance in the backbone extended in the five-fold helix motif of the five-mer is 67.5±2.2 Å. It gives an axial rise of 13.6±0.4 Å per repeating unit (αGalp(1→4)βGlcpA-2S(1→4)βGlcpA) or 4.5±0.14 Å per monosaccharide residue. It indicates an almost total elongation. The overall backbone extension of 67.5±2.2 Å does not change when the conformational transition brings the 5-fold helix to the 2-fold helix. In the 2-fold conformation, the overall structure has the shape of a brush. The side chains alternate their orientation by 180°, separated by a distance of

27.7±0.7 Å. The height of such polysaccharide calculated as a distance on the projection between distant α-Galp residues is around 33 Å. In this conformation, the width of the polysaccharide is about 8 Å.

Without influencing the conformations of Model 1 to 4, the four possible combinations resulting from the two α linked terminal Glcp monosaccharides enhance a volume of 20 Å diameter the space occupied at the branching point βGlcpA(1→3)αGalpA-2S (Figure S3).

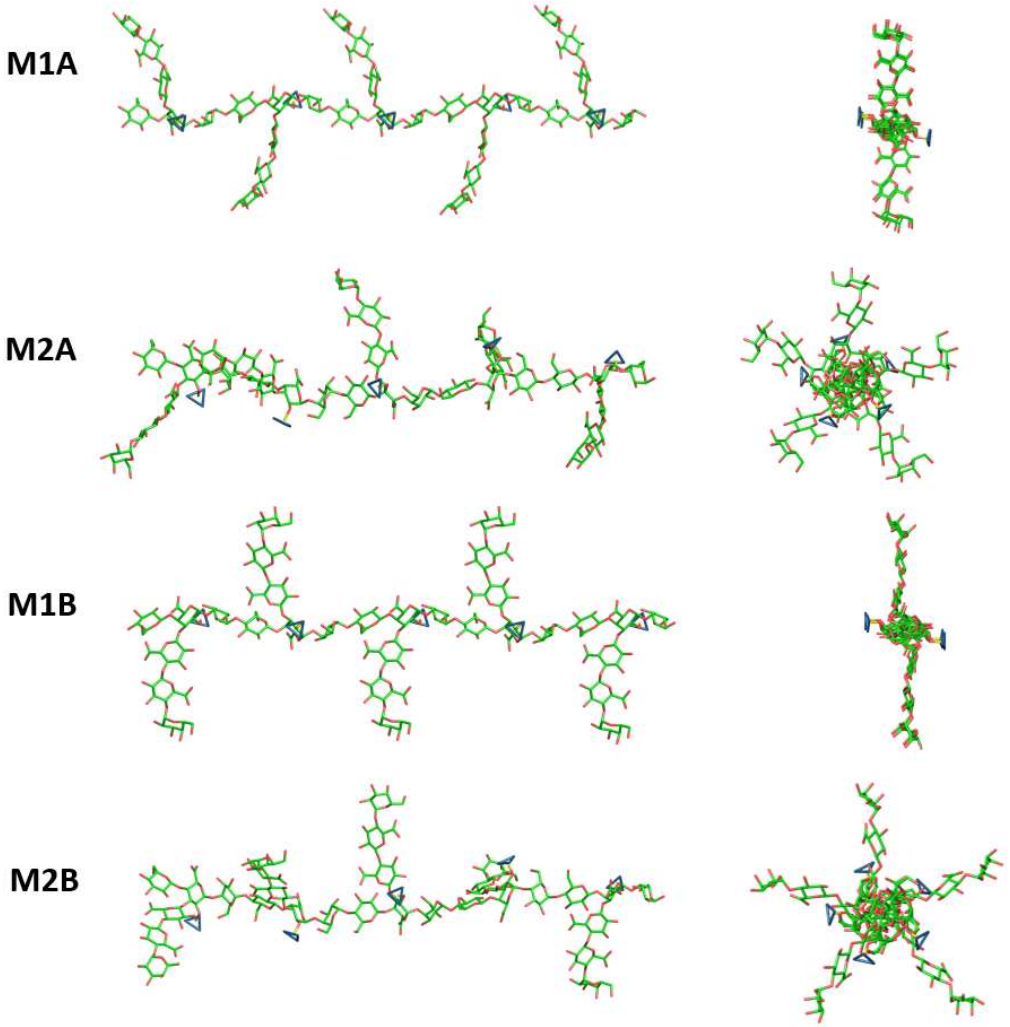

Figure 5. *Three-dimensional representation of the four structures of Infernan.*

Ca^{2+} binding at the branching point. The computation of the electric potential of each low energy model, as calculated from the Avogadro program, gives clear indications of the occurrence of a negatively charged region around the branching points (Figure S4). Such a feature results from the vicinity of sulfate and several carboxylate groups of Infernan and indicates regions for the location of Ca^{2+}. The calcium ion was placed next to COO^- or $SO3^-$ fragments of αGalpA-2S of tetramer in the backbone branching point. The detailed structural organization of Ca^{2+} in the combining site was optimized throughout quantum mechanics methods following the protocol described in the Material and Methods. To avoid potential effects of the monosaccharide units at extremities, we restained the Φ,Ψ angles of the backbone linkages βGlcp(1→4)αGalpA-2S and αGalpA-2S(1→4)αGalp to those in two-fold and five-fold conformation (which are respectively (-55, 153); (84, 131) and (-148, 114);

(84, 131)). Only βGlcpA(1→3)αGalpA-2S linkage connecting the side chain and the backbone was allowed to adjust to accommodate fully Ca^{2+} binding. The results show that Ca^{2+} can interact, either with the two carboxylate groups belonging to βGlcpA and αGalpA-2S or with the carboxylate group from βGlcpA and a sulfate group from αGalpA-2S (Figure 6).

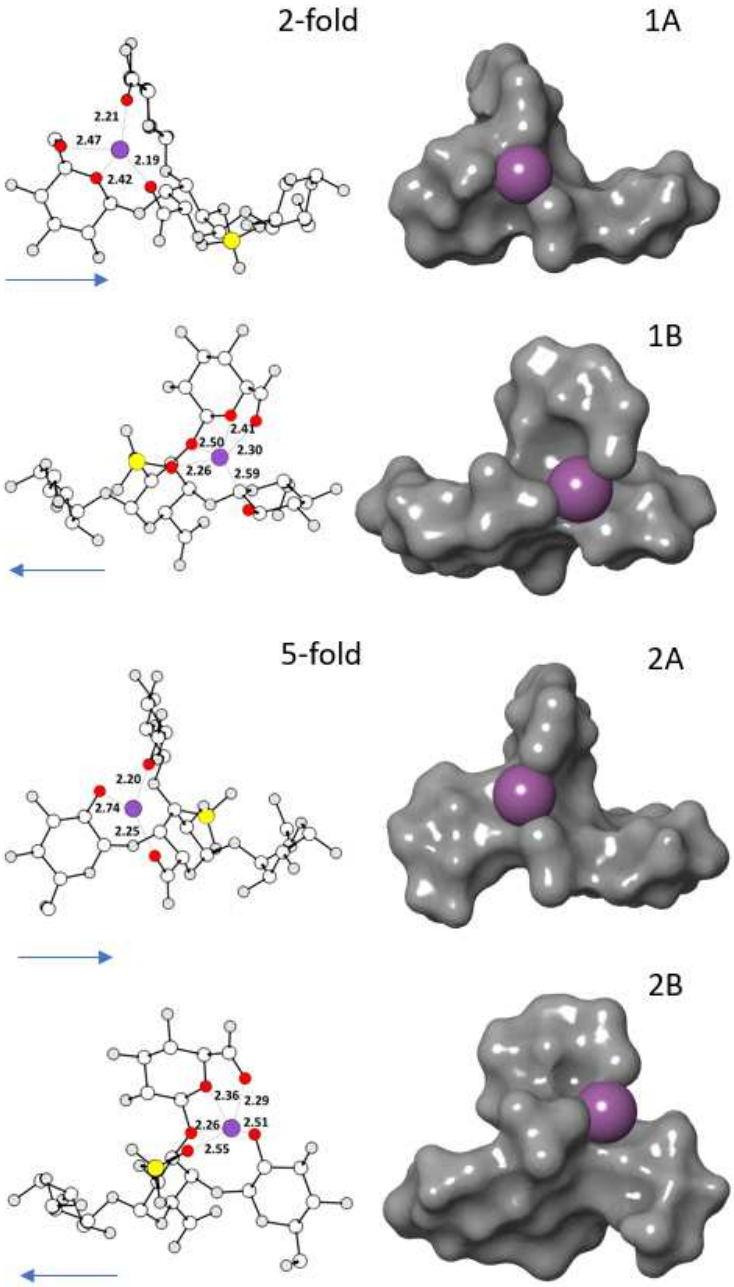

Figure 6. *Depiction of the complexation schemes of Ca^{2+}, combining sites of the four low energy conformations of Infernan; as results of quantum chemistry computation using B3LYP/6-31(d,p). Colour coding: Red: oxygen; Yellow: sulfur; Purple: calcium. The arrow indicates the direction of the backbone, from non-reducing to reducing end.*

The B3LYP functional with the Pople's split-valence basis sets has been extensively used for the calculation of related systems, including uronic acids (Plazinski and Drach, 2013; Stewart, Gray, Vasiljevic, Orbell, 2014; Zhao et al., 2018; Ardiles and Rodriguez, 2021). In the current study, the

minimization of system has been performed both in the conditions of vacuum and accounting water implicitly. These give slightly different results (Table S2 and S3). The use of 6-31(d,p) in vacuum provides Ca2+…O distances comparable with those reported for relative systems experimentally (Sommer et al., 2015; Novoa et al. 2014; Goldberg et al., 2018; Goldberg et al., 2021). Interestingly, the use of diffuse function on heavy atoms shortened the distances. In contrast, the implementation of implicit water makes the Ca^{2+}…O distances longer than experimentally observed (Table S2, S3).

In the four types of complexes, the Ca^{2+} ion interacts with up to five oxygen atoms at distances ranging from 2.2 to 2.6 Å (Figure 6) βGlcpA, αGalpA-2S and βGlcp form the binding site. In three complexes, O2, belonging to βGlcp from the backbone, interacts with Ca^{2+}. In contrast, in the first M1A complex, the βGlc offers its O6 and O5 atoms. For all the optimized complexes, the Φ,Ψ values of βGlcpA(1→3)αGalpA-2S remain in the low energy region. They slightly deviate from those populated in the course of molecular dynamics simulations. In all models, the coordination sphere of the Ca^{2+} could be completed by interaction with other oxygen atoms to form a seven-fold up to nine-fold coordination shell (Dheu & Perez, 1992). This is where the oxygen atoms of a sulfate group can complete the coordination to six; (Figure 7), leaving room for water molecules. (Figure 7). The atomic coordinates of the four models are given in Table S3.

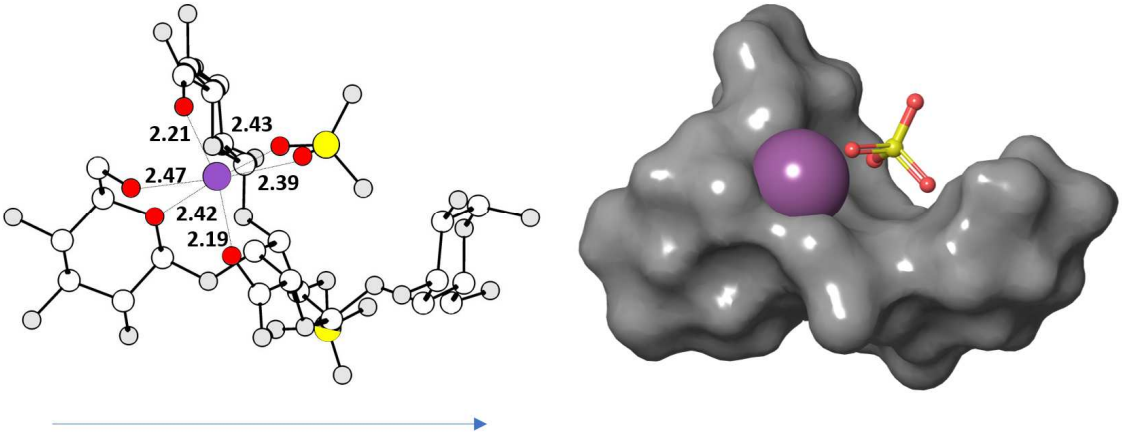

Figure 7. *Depiction of the complexation schemes of Ca^{2+}, (M1A) showing as the oxygen atoms of a neighbouring sulfate group complete the chelating site of Ca^{2+}. Colour coding: Red: oxygen; Yellow: sulfur; Purple: calcium. The arrow indicates the direction of the backbone, from non-reducing to reducing end.*

Extension to the polysaccharide chains. The transposition of these four models of calcium complexation to the four models of Infernan does not alter the overall chain conformation for the previously established models. Figure 8 displays the four polysaccharide models in their interactions with Ca^{2+} ions. The five-fold helical model 2B is the most favourable for interactions with Ca^{2+}. The two-fold helical models, M1A and M1B offer two types of Ca^{2+} sites of similar binding energy.

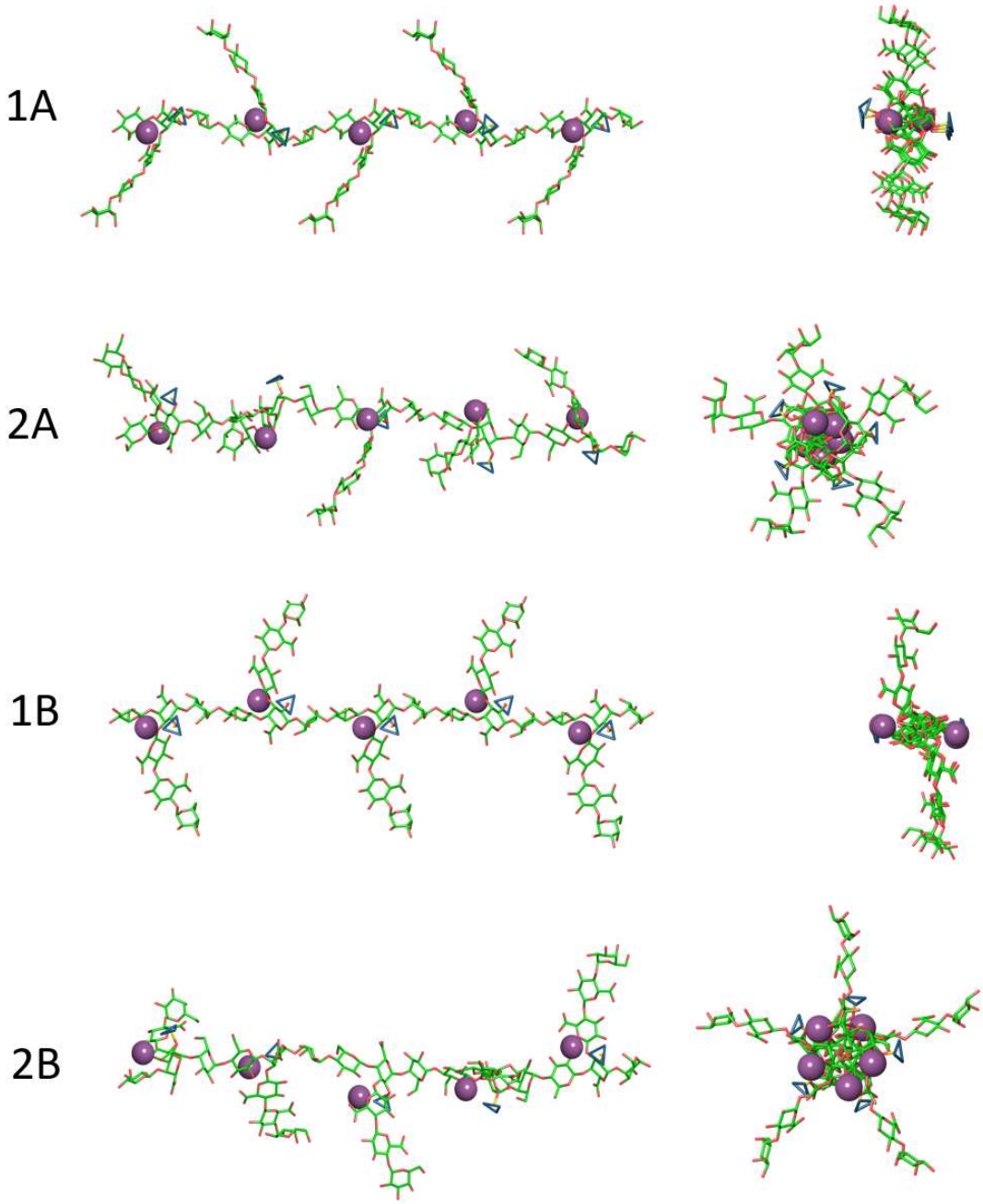

Figure 8. *The calcium-binding models within the helical structures of Infernan. The lowest energy arrangement occurs for model 2B. Model 1A and 1B display similar binding energy.*

Junctions zones. The occurrence of sulfate groups and uronic acid residues confers to ionic polysaccharides' ability to form gels in the presence of divalent cations. The molecular basis of Ca^{2+} induced gelation in alginates and pectins, among others, are well established and documented (Braccini & Perez, 2001). As Infernan forms a gel in the presence of Ca^{2+} (Zykwinska et al., 2019), we explored the modes of interactions that would lead to form junction zones and propose some gelation models. Only those corresponding to a two-fold helical structure (M1A and M1B) must be considered among the four existing structural models. Their ribbon-like structures allow bringing neighbouring chains in proximity without steric clashes. Moreover, the orientation of the side chains in model M1A, would not prevent two polymer chains from coming together.

To evaluate the favourable Inferans chain arrangements in the presence of calcium ions, we used the chain-pairing procedure developed previously to probe the Ca^{2+} induced gelation in alginates and pectins (Braccini & Perez, 2001). In this procedure, an inter-helical distance is computed for all the relative translational and rotational positions of one chain with respect to the other. Not surprisingly, and because of the structural features of the infernan chain described above, there is a single way for two infernan chains in the 2-fold conformation to interact in a parallel arrangement. The results of the chain-pairing procedure define a spatial organization of the interacting molecular partners. The ensemble was submitted to an energy optimization at the molecular mechanics level, using the OPLS force-field available in the Schrodinger Maestro 11.0 version.

The resulting is a 'junction zone' based on chain-chain interactions that lead to strong dimer associations with essential contributions from the van der Waals and hydrogen bonding interactions. Calcium ions and sulfate groups, and uronic acids have specific positions in well-adapted cavities. The feature of the interaction is the heterotypic mode of binding. Its nature contrasts with the so-called « egg-box » feature that most frequently characterizes the binding mode of ionic polysaccharides.

The present model is established following a logical succession of the structural elements of increasing complexity, with a limited amount of experimental data: calcium induced formation of gel, parallel arrangements of chain assemblies of polysaccharides as seen from AFM. Needless to state that further confrontations with available experimental data are needed to ascertain the model's validity. Nevertheless, the model makes it possible to represent some grossly observed features and to make predictions. The availability of a structural model of complex associations constitutes a provisional and partial structural elucidations, represents a tool for thinking and guidelining new experiments.

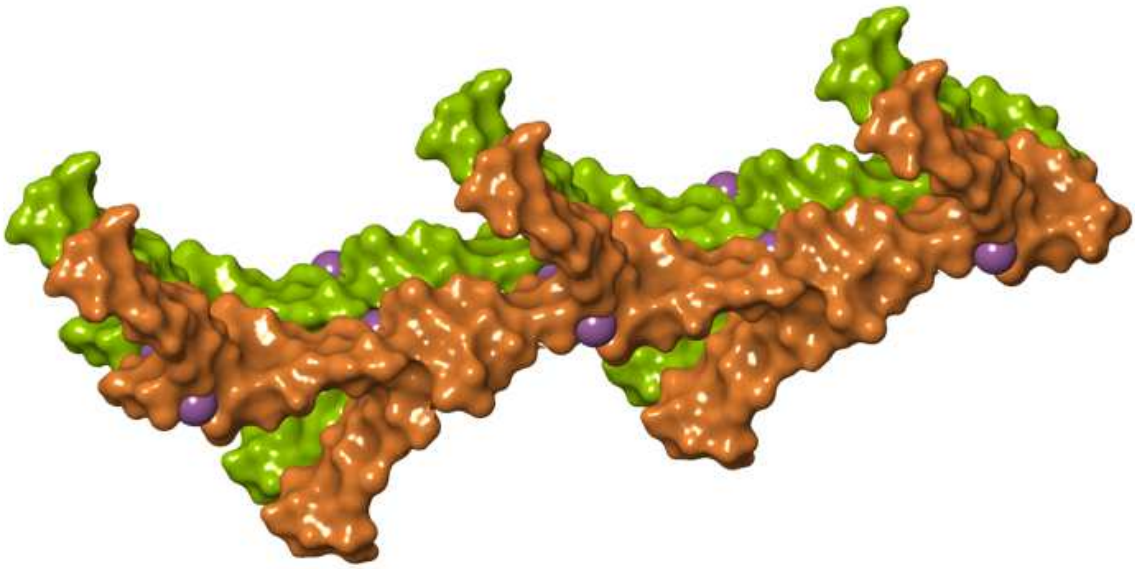

Figure 9. *Schematic representation of two parallel chains of Infernan (M1A) throughout Ca^{2+} complexation, illustrating the formation of junction zones of the physical gel.*

Discussion & Conclusion

Infernan, the EPS produced by the deep-sea hydrothermal bacterium *Alteromonas infernus*, exhibits an unusually complex primary structure, with the main backbone wearing two levels of sidechains. The nature of the monosaccharides and their glycosidic linkages give rise to a linear backbone that propagates with an almost fully extended elongation. Nevertheless, some conformational flexibility about glycosidic linkages exists. The flexibility of βGlcp(1-4)αGalpA-2S linkage determines the conformation of the Infernan's backbone between a 2-fold or a 5-fold helical structure. The conformational transition in the polysaccharide backbone occurs through a low energy barrier. The side chains do not influence the overall backbone conformation, unlike other branched polysaccharides such as rhamnogalacturonan I (Makshakova et al., 2017). However, the side chains of Infernan are essential for creating the Ca^{2+} chelating sites in the region of the branching point. They involve the three adjacent residues: βGlcp and αGalpA-2S from the backbone and βGlcpA from the side chain. Two types of chelating sites can be formed depending on the orientation of the side chain to the backbone. This holds true for both conformations of the backbone, thus, resulting in four types of binding sites in total.

The configurations of the binding sites within one Infernan branching point have different binding energy values for Ca^{2+} complexation. The difference of energy of calcium complexation is the determinant factor to the polysaccharide backbone transition. In other words, Ca^{2+} binding may induce and stabilize a particular conformational state of the polysaccharide. Even though the two-fold helix being the most populated state of the backbone, as revealed in molecular dynamics in the absence of specifically bound counter-ions, the binding of Ca^{2+} may contribute to stabilizing the five-fold helix conformation. As such, the five-fold helix may be considered as a "reservoir" for calcium ions. Then, the transition from five-fold to two-fold helices occurs due to a rotation over the βGlcp(1→4)αGalpA-2S linkage. The Ca^{2+} binding sites are located at a distance of 13.6±0.4 Å from each other along the backbone. The two-fold helical conformation of the backbone, presenting the appropriate Ca^{2+} cavity, may induce the formation of the polysaccharide chain pairing, providing that the chains are favourably oriented. The coordination of Ca^{2+} within one chain reaches completion with the oxygen atoms of a sulfate group belonging to the neighbouring chain, with a significant gain in enthalpy. Such chain pairing would result in the formation of "junction zones", yielding to gel formation. A parallel arrangement of the Infernan chains is energetically more favourable than an antiparallel one. Such a feature is not favourable in terms of entropy. An eventual contribution of the side chain, which also may play a role in the extensive gel network formation (Mikshina 2017, Makshakova 2018), is still to be unravelled. For a physical gel to occur, both junction zones and regions that act as "gel breakers" to stop the propagation of the junction zones are required. The five-fold conformation of the backbone where the side chains are fully deployed is likely to act as such.

The present results suggest the occurrence of a novel ionic chelating mechanism of polysaccharides. Previously characterized mechanisms involve a monotypic type of interactions underlying the « egg-box » binding (Braccini & Perez, 2000). Another type of complexation occurs in the case of κ-carrageenan gelation. The κ-carrageenan chains form thermo-reversible gel based on double helices stabilized by a counter-ion binding in junction zones. It was experimentally proven that ion binding precedes helix formation. Upon cooling, the counter-ions being unbound start to occupy, first, the weaker binding site and the stronger binding sites. When the strong binding sites are occupied, the transition from random coil to helix starts, and gel forms (Makshakova et al., 2020). It remains to be seen whether the heterotypic mode of pairing Infernan chains can occur in other polysaccharides interactions. The methodology developed in the present work could provide a way how to understand or predict the formation of physical gels in complex polysaccharide structures.

Acknowledgements. The calculations were carried out using the equipment of the High Performance Computing centre of University Grenoble Alpes, Grenoble, France and the shared research facilities of high-performance computing resources at Lomonosov Moscow State University, Moscow, Russian Federation. The CrossDisciplinary Program Glyco@Alps supported this work within the framework 'Investissement d'Avenir' program [ANR-15IDEX-02]. The study was partly funded by RFBR & CNRS according to the research project № 21-54-15008 (in part of DFT calculations). O.M. acknowledges (in part of M.D. simulations of branched oligosaccharide fragments) funding from the government assignment for FRC Kazan Scientific Center of (RAS).

Author Contributions

ONM and S.P. coordinated the paper. ONM performed all M.D. calculations, analysis of trajectories and writing the paper. S.P. – general leading of work, PES calculations, generation of initial structures, writing, and revising the manuscript. AZ formulated the task in terms of Infernan structure, S.C. and SC-J consulted about previously obtained experimental data.

References

Allinger, N. L., & Yan, I. Q. (1993). Molecular mechanics (MM3). Calculations of furan, vinyl ethers, and related compounds. *Journal of the American Chemical Society*, *115,* 11918–11925.

Ardiles, C.S., Rodríguez, C.C. (2021) Theoretical study for determining the type of interactions between a GG block of an alginate chain with metals Cu2+, Mn2+, Ca2+ and Mg2+. *Arabian Journal of Chemistry,* 14 (10), 103325

Baumann P., Baumann L., Bowditch, R.D. & Beaman, B. (1984). Taxonomy of *Alteromonas: A. nigrifaciens* sp. nov., nom. rev.; *A. macleodii*; and *A. haloplanktis*. *Int J Syst Bacteriol 34*, 145–149.

Becke, A.D. (1983). Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.,* 98, 5648–5652.

Bekri, L., Zouaoui-Rabah, M., Springborg, M., & Sekkal Rahal, M. (2018). A structural DFT study of MM, G.G., M.G., and G.M. alginic acid disaccharides and reactivity of the M.G. metallic complexes. *J. Mol. Model., 24*. 312.

Birch, J., Van Calsteren, M.R., Perez, S. & Svensson, B. (2019). The exopolysaccharide properties and structures database: EPS-DB. Application to bacterial exopolysaccharides. *Carbohydr. Polym., 205*, 565–570.

Cao L, Lu W, Mata A, Nishinari K & Fang Y. (2020). Egg-box model-based gelation of alginate and pectin: A review. *Carbohydr Polym.*, *242*, 116389.

Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., et al. (2012). AMBER 12. San Francisco: University of California.

Casillo, A., Lanzetta, R., Parrilli, M., Corsaro, M. (2018). Exopolysaccharides from Marine and Marine Extremophilic Bacteria: Structures, Properties, Ecological Roles and Applications. Marine Drugs 16(2), 69.

Colliec-Jouault, S., Chevolot, L., Helley, D., Ratiskol, J. Bros, A., Sinquin, C., Roger, O., & Fischer, A.-M. (2001). Characterization, chemical modifications and in vitro anticoagulant properties of an exopolysaccharide produced by *Alteromonas infernus*. *Biochimica et Biophysica Acta, 1528,* 141-151.

Dheu, M.L. & Perez, S. (1983). Geometrical features of calcium-carbohydrate interactions. *Carbohydr. Res.*, *124*, 324-332.

Dick, G.J. (2019). The microbiomes of deep-sea hydrothermal vents: distributed globally, shaped locally. *Nature Reviews Microbiology, 17*, 271-283.

Engelsen, S. B., Hansen, P. I., & Perez, S. (2014). POLYS 2.0: An open-source software package for building three-dimensional structures of polysaccharides. *Biopolymers, 101*, 733–743.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics, 103*, 8577–8593.

Frisch, M. J., Trucks, G. W., Schlegel H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, N. J., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., & Fox, D. J., (2009) Gaussian 09, Revision A.02 Gaussian Inc., Wallingford, C.T.

Goldberg, M, Gafurov, M, Makshakova, O, Smirnov, V., Komlev V., Barinov, S., Kudryavtsev, E., Sergeeva, N., Achmedova, S., Mamin, G., Murzakhanov, F., Orlinskii, S. (2019) Influence of Al on the Structure and in Vitro Behavior of Hydroxyapatite Nanopowders. *J. Phys. Chem. B*. 123(43). 9143-9154.

Goldberg, M., Gafurov, M., Murzakhanov, F., Fomin, A., Antonova, O., Khairutdinova, D., Pyataev, A., Makshakova, O., Konovalov, A., Leonov, A., Akhmetova, S., Sviridova, I., Sergeeva, N., Barinov, S., Komlev, V. (2021) Mesoporous Iron(III)-Doped Hydroxyapatite Nanopowders Obtained via Iron Oxalate. *Nanomaterials*. 11(3), 811.

Guezennec, J. (2002). "Deep-sea hydrothermal vents: A new source of innovative bacterial exopolysaccharides of biotechnological interest?" *Journal of Industrial Microbiology & Biotechnology* **29**, 204-208.

Guvench, O., Whitmore E.K. (2021). Sulfation and Calcium Favor Compact Conformations of Chondroitin in Aqueous Solutions, *ACS Omega 20,* 13204–13217.

Hanwell, M.D., Curtis, D.E., Lonie, D.C., Vandermeersch, T., Zurek, E., Hutchison, G.R. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform *J. Cheminf*. 4, 17.

Harger, M., Li, D., Wang, Z., Dalby, K., Lagardere, L, ; Piquemal, J-P.; Ponder, J.W. & Ren, P. (2017). Tinker-OpenMM : Absolute and Relative Alchemical Free Energies using AMOEBA on GPUs. *J. Comp. Chem.,* **38** (23): 2047–2055.

Hecht, H., Srebnik, S. (2016). Structural Characterization of Sodium Alginate and Calcium Alginate. *Biomacromolecules. 2016*, 17, 2160-2167.

Heymann, D., Ruiz-Velasco, C., Chesneau, J., Ratiskol, J., Sinquin, C., Colliec-Jouault, S. (2016). Anti-metastatic properties of a marine bacterial exopolysaccharide-based derivative designed to mimic glycosaminoglycans. *Molecules, 21,* 309.

Jannasch, H.W. &Taylor, C.D. (1984). Deep-sea microbiology. *Annu. Rev. Microbiol., 38*, 487–514.

Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., et al. (2008). GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *Journal of Comp. Chem., 29*, 622–655.

Le Questel, J.-Y., Cros, S., Mackie, W., & Perez, S. (1995). Computer modelling of sulfated carbohydrates: applications to carrageenans. *Int. J. Biol. Macromol., 17*, 161–175.

Lee, C., Yang, W., & Parr, R.G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B, 37*, 785–789.

Maire du Poset A, Zitolo A, Cousin F, Assifaoui A, & Lerbret A. (2020). Evidence for an egg-box-like structure in iron(ii)-polygalacturonate hydrogels: a combined EXAFS and molecular dynamics simulation study. *Phys Chem Chem Phys. 22*, 2963-2977.

Makshakova O., Faizullin D. & Zuev Y. (2020). Interplay between secondary structure and ion binding upon thermoreversible gelation of κ-carrageenan. *Carbohydr. Polym., 227*, 115342.

Makshakova O., Gorshkova T., Mikshina P., Zuev Y., Perez S. (2017). Metrics of rhamnogalacturonan I with β-(1→4)-linked galactan side chains and structural basis for its self-aggregation. *Carbohydr Polym., 158,* 93-101.

Makshakova, O., Chachkov, D., & Ermakova, E. (2011). Geometry and vibrational frequencies of the helical polypeptide complexes with ligand molecules. *Int. J. Quant. Chem., 111*, 2525–2539.

Makshakova, O.N., Faizullin, D.A., Mikshina, P.V., Gorshkova, T.A. & Zuev, Y.F. (2018). Spatial structures of rhamnogalacturonan I in gel and colloidal solution identified by 1D and 2D-FTIR spectroscopy. *Carbohydr. Polym., 192,* 231–239.

Marenich, A.V., Cramer, C.J. & Truhlar, D.G. "Universal solvation model based on solute electron density and a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions," *J. Phys. Chem. B*, **113** (2009) 6378-96.

Merceron, C., Portron, S., Vignes-Colombeix, C., Rederstorff, E., Masson, M., Lesoeur, J., Sourice, S., Sinquin, C., Colliec-Jouault, S., Weiss, P., Vinatier, C. & Guicheux, J. (2012). Pharmacological modulation of human mesenchymal stem cell chondrogenesis by a chemically oversulfated polysaccharide of marine origin: potential application to cartilage regenerative medicine. *Stem Cells, 30*, 471-480.

Mikshina P.V., Makshakova O.N., Petrova A.A., Gaifullina I.Z., Idiyatullin B.Z., Gorshkova T.A., & Zuev Y.F. (2017). Gelation of rhamnogalacturonan I is based on galactan side chain interaction and does not involve chemical modifications. *Carbohydr. Polym., 171*, 143-151.

Novoa, A., Eierhoff, T., Topin, J., Barluenga, S., Imberty, A., Römer, W., Winssinger, N. (2014) A LecA ligand identified from a galactoside-conjugate array inhibits host cell invasion by Pseudomonas aeruginosa. *Angew. Chem. Int. Ed. Engl.* 53(34), 8885-8889.

Perez S., Fadda, E., Makshakova O., (2021). Computational Modeling in Glycoscience. In: Barchi Jr., Joseph (ed.) *Comprehensive Glycoscience, 2nd edition*. vol. 1, pp. 374-404. Oxford: Elsevier.

Perez S.J.L.P. & Claudio G.C. (2020). Molecular dynamics simulations of two double-helical hexamer fragments of iota-carrageenan in aqueous solution. *J Mol Graph Model., 98*,107588. doi: 10.1016/j.jmgm.2020.107588.

Perez, S. & Rivet, A., (2021). Polys Glycan Builder: An online Application for Intuitive Construction of 3D Structures of Complex Carbohydrates, *Glycoinformatics, Methods in Molecular Biology.*

Perez, S. & Tvaroska, I., (2014). Protein-Carbohydrate interactions: Molecular Modeling Insights, *Advances in Carbohydrate Chemistry and Biochemistry, **71**, 9-137.*

529 Plazinski W. (2011). Molecular basis of calcium binding by polyguluronate chains. Revising the egg-
530 box model. *J Comput. Chem., 32*, 2988-95. doi: 10.1002/jcc.21880.

531 Plazinski, W. & Drach, M. (2013) Calcium-α-L-Guluronate Complexes: Ca2+ Binding Modes from DFT
532 MD Simulations *J. Phys. Chem. B*, 117, 12105−12112.

533 Plazinski W. & Rudzinski W. (2012). Molecular modeling of Ca2+-oligo(a-L-guluronate) complexes:
534 toward the understanding of the junction zone structure in calcium alginate gels. *Struct. Chem. 23*,
535 1409–1415.

536 Raguénès, G. H., Peres, A., Ruimy, R., Pignet, P., Christen, R., Loaëc, M., Rougeaux, H., Barbier, G., &
537 Guezennec, J. (1997). *Alteromonas infernus* sp. nov., a new polysaccharide-producing bacterium
538 isolated from a deep-sea hydrothermal vent. *Journal of Applied Microbiology, 82*, 422-430.

539 Rederstorff, E., Rethore, G., Weiss, P., Sourice, S., Beck-Cormier, S., Mathieu, E., Maillasson, M.,
540 Jacques, Y., Colliec-Jouault, S., Fellah, B. H., Guicheux, J., & Vinatier, C. (2017). Enriching a cellulose
541 hydrogel with a biologically active marine exopolysaccharide for cell-based cartilage engineering.
542 *Journal of Tissue Engineering and Regenerative Medicine, 11*, 1152-1164.

543 Roger, O., Kervarec, N., Ratiskol, J., Colliec-Jouault, S., & Chevolot, L. (2004). Structural studies of the
544 main exopolysaccharide produced by the deep-sea bacterium *Alteromonas infernus*. *Carbohydrate*
545 *Research, 339*, 2371-2380.

546 Ruiz Velasco, C., Baud'huin, M., Sinquin, C., Maillasson, M., Heymann, D., Colliec-Jouault,S.,
547 Padrines, M. (2011). Effects of a sulfated exopolysaccharide produced by *Alteromonas infernus* on
548 bone biology. *Glycobiology, 21*, 781-795.

549 Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1997). Numerical integration of the cartesian
550 equations of motion of a system with constraints; molecular dynamics of n-alkanes. *Journal of*
551 *Computational Physics, 23*, 327–341.

552 Sommer, R., Hauck, D, Varrot, A., Wagner, S., Audfray, A., Prestel, A., Möller, H.M., Imberty, A., Titz,
553 A. (2015) Cinnamide Derivatives of d-Mannose as Inhibitors of the Bacterial Virulence Factor LecB
554 from Pseudomonas aeruginosa. *ChemistryOpen,* 4(6), 756-767.

555 Stewart, M.B., Gray, S.R., Vasiljevic, T., Orbell, J.D. (2014) Exploring the molecular basis for the metal-
556 mediated assembly of alginate gels. *Carbohydr Polym*. 102, 246-253.

557 Toukach, P.V. & Egorova, K.S. (2016). Carbohydrate structure database merged from bacterial,
558 archaeal, plant and fungal parts. *Nucleic Acids Res., 44,* D1229–D1236.

559 Yousfi, N., Sekkal-Rahal, M., Sayede, A., & Springborg, M. (2010). Relaxed energetic maps of k-
560 carrabiose: A DFT study. *J. Comput. Chem., 31,* 1312–1320.

561 Zhao, H., Yang, Y., Shu, X., Wang, Y., Wu, S., Ran, Q., Liu, J. (2018) The binding of calcium ion with
562 different groups of superplasticizers studied by three DFT methods, B3LYP, M06-2X and M06.
563 *Computational Materials Science*, 152, 43–50.

564 Zykwinska, A., Marquis, M., Godin, M., Marchand, L., Sinquin, C., Garnier, C., Jonchère, C., Chédeville,
565 C., Le Visage, C., Guicheux, J., Colliec-Jouault, S., Cuenot, S. (2019). Microcarriers based on
566 glycosaminoglycan-like marine exopolysaccharide for TGF-β1 long-term protection. *Marine Drugs,*
567 *17*, 65.

568