J. Mol. Model. **1997**, *3*, 301 – 310

Molecular Modelling of Six-Ring Agarose Chains: Effects of Explicit and Implicit Solvent

Nicholas M.W.Haggett*, Rainer A.Hoffmann[1], Brendan J.Howlin, Graham A.Webb

Dept. of Chemistry, University of Surrey, Guildford, Surrey, GU2 5XH, UK (chp1nh@surrey.ac.uk)

[1]UNILEVER Central Research, Colworth Laboratory, Sharnbrook, Beds., UK

Received: 1 April 1997 / Accepted: 21 July 1997 / Published: 7 August 1997

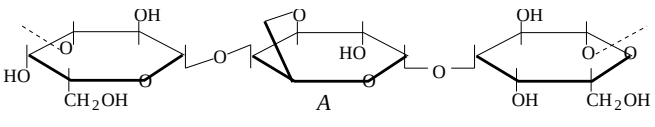
Abstract

A study on the molecular dynamics of agarose-type oligosaccharide chains in water is presented. In the field of molecular modelling one is able to use either explicit water molecules or a representative bulk dielectric constant: the two methods are compared here. The starting conformation was taken from X-ray fibre diffraction data for agarose double helices, which are neutral molecules. The resulting trajectories were analysed for the behaviour of dihedral linkages, and these were compared to systematic searches performed on isolated agarobiose and neo-agarobiose units. The results showed that when explicit water molecules are used the oligosaccharide chains are “stiffer” and have a reduced mobility.

Keywords: Oligosaccharides, Dihedral angles, Molecular Dynamics, Agarose
Abbreviations: MM: Molecular Mechanics; MD: Molecular Dynamics; AG: linkage between anhydrogalactose and galactose sugar rings; GA: linkage between galactose and anhydrogalactose rings.

Introduction

Agarose is composed of the alternating sugar units 4-linked β-D-galactopyranosyl (*G*) and 3-linked 3,6 anhydro-α-L-galactopyranosyl (*A*) [1]; see below:



Agarose. The light vertical lines represent the two atoms attached to a carbon.
A: anhydro-ring
G: galactose ring

* *To whom correspondence should be addressed*

Natural agarose (known simply as ‘agar’) occurs in the marine algal plants belonging to the *Gracilaria* group as a polysaccharide, usually randomly sulphated and/or methylated on O(2) of *A* and O(6) of *G*. It is extensively used in the food and medical industries as a thickening agent or as a supporting medium in the gel form [2].

Various examinations of agar have centred on obtaining structures of oriented fibres using X-ray crystallography [3] and also chemical examinations on the effect of methylation on the gelling characteristics observed [4]. Light scattering studies can provide some information on the size of aggregated structures within solutions, where more than one agarose helices have associated with each other.

These investigations give only indirect information on the nature of carbohydrate - carbohydrate interactions. Molecu-

lar modelling techniques using computers are unique in that molecules can be examined directly and spatial relationships between groups observed at first hand. Energies can be plotted and processes which normally occur almost instantaneously can be studied at length.

Most such investigations in the field to date have been concerned with very small parts of the agarose system; disaccharides or trisaccharides at most [5].

One way to begin the treatment of these molecules using computer models is to execute a systematic conformational search over the glycosidic linkage (being responsible for most of the shape changes in a chain), as has been done frequently by other workers [5, 6]. In this way the low energy (relaxed) conformations of each type of linkage can be identified and future trajectories compared with a standard. In the current report these so - called adiabatic or Ramachandran maps have been calculated for both types of linkage (taking disaccharide molecules) in the agarose chain using a dielectric constant of $\varepsilon = 80.0$. Possible effects of next - neighbour sugar residues (as in an infinite chain) were not taken into account, as the effect due to such remote residues would be small.

In the present work molecular dynamics experiments were conducted in the presence of both implicit and explicit solvent and the resulting time-dependent behaviour compared. The trajectories of the dihedral angles were compared to the Ramachandran plots constructed earlier. We believe the simulations will add to an understanding of the characteristics of solution behaviour for short chains of agarose, and these characteristics will also manifest themselves in models of extended agarose polymers.

Experimental

All the modelling work was carried out on a Silicon Graphics Iris Indigo XZ 4000 workstation running INSIGHT II version 2.3.0 for visualisation of molecules, and DISCOVER version 2.95 for the modelling calculations. The forcefield used was AMBER [7], with Homans' modification [8] for use with polysaccharides.

Conditions for a systematic search

The two possible glycosidic linkages in agarose were searched systematically and the energy *versus* dihedral linkage plotted [6, 9] (a Ramachandran plot).

This was done for both types of linkage (AG and GA) using a dielectric constant of $\varepsilon = 80.0$ to represent an aqueous medium.

A macro for DISCOVER was written which could assign names to the two torsion angles in the glycosidic linkages so that they could be fixed in space while the rest of the molecule was minimised. A diagram of atom definitions for the dihedral angles is reproduced in the Results and Discussion section below. These angles were systematically incremented through 360 degrees in steps of 12 degrees, the molecular energy being minimised using the steepest descents and conjugate gradients method for every combination. The result is a 30 × 30 grid of energy values, which can be viewed in three-dimensional mode using Insight II and exported to graph plotting programs.

Conditions for dynamics

For the MD experiments the initial X-ray conformations were relaxed using three minimiser algorithms: steepest descents for initial relaxation, conjugate gradients and a quasi Newton-Raphson method, VA09A when the structure was close to convergence. The convergence criterion was to meet a derivative of less than 0.1 $kcal \cdot mol^{-1} Å^{-1}$.

Minimize: in three stages of 500 iterations, using different convergence methods for each (one iteration = 1 femtosecond);

• *Dynamics:* warm to 350K in twelve stages of 200 iterations; run for 100,000 iterations at 350K with history file updated every 100 iterations,

• *Cutoff:* 16 Angstroms (switching over a distance of 1.5): calculations between atoms separated by more than this distance were ignored,

• *Relative Permittivities:* 1.0 when explicit water was specified; 80.0 for simulations with no specific water molecules.

Two dynamics simulations were run using exactly the same initial assembly of molecules: two strands of carbohydrate polymer arranged in double helical conformation, each strand having three anhydrogalactose rings and three galactose rings. This makes a total of 240 atoms. The first simulation employs a cell containing water molecules at a density corresponding to water at room temperature and pressure, and the double-stranded carbohydrate placed at the centre. The water molecules are responsible for a further 200 atoms and there are additional atomic interactions, both bonded and non-bonded. The second cell contains the same carbohydrate molecules but a dielectric of 80.0 instead of specific water molecules was used. This value has been used by other workers when representation of the electrostatic screening effect of water molecules is required [10]. In both cases Periodic Boundary Conditions (PBC) were applied to simulate a bulk solution case. It was expected that the lack of specific hydrogen bonds between solute and solvent in the second case would make significant differences to the behaviour of carbohydrate molecules.

Cell sizes

For the case when specific water was used, the cell was made large enough to enclose the oligosaccharide, leaving sufficient room between oligosaccharides for water molecules to diffuse. The dimensions are 21 × 15 × 15 $Å^3$. This system occupies a large proportion of the computing resources when running.

It should be noted here that the ratio of oligosaccharide to water is equivalent to an extremely concentrated solution, or to a region within the solution in which molecules are very densely packed, as in an aggregated particle (variations in density of molecular packing in gels have been observed using birefringence measurements [11]).

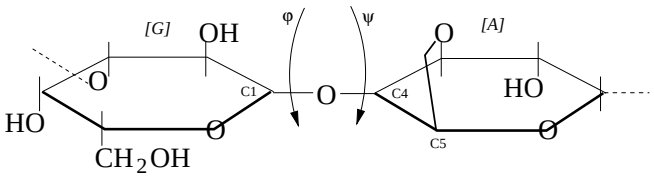
The cell size for use with no explicit water was chosen according to the concentration of agarose known to give gelation in water: greater than ~ 0.2 % *wt*/*wt*. (strong, elastic gels can be formed from solutions of just a few percent [12]). This concentration (assuming six-ring chains) is equivalent to a volume-per-double helix of 1.5 million Å^3.

To avoid having too big a cell dimensions of 24 × 48 × 48 Å^3 were used which represents a concentration increase of 27 times on the estimated figure above. This cell size could not be used in the explicit water simulation on account of constraints on the number of atoms generated.

Results and discussion

Ramachandran plots

The systematic plots for the *A-G* and *G-A* linkages can be seen in Figure 5 later in the paper, where they have been overlaid with trajectories from dynamics runs. The atoms chosen for measurement of dihedral angles were: Ring 1(ring oxygen → carbon 1 → linking oxygen) → Ring 2(carbon *n* → carbon (*n*+1)). An example of this is shown below for the case of a *G-A* linkage.



Path for dihedral angles:
ring-O/C1/bridge-O/C4/C5.

A dihedral angle of zero degrees means that all three bonds in it are equi-planar.

The potential energy surfaces for the two types of glycosidic linkage display several possible minimum energy conformations, some of which are close but not equal to X-ray derived values [3]. By inspection of *Orbit* models it can be seen that the minima found correspond to positions where groups on the rings are least hindered, especially those close to the glycosidic linkage.

Dynamics calculations: explicit water

The AMBER forcefield has terms for representing hydrogen bonds (in the non-bonding energy expression

$$\sum\left[\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right]$$

involving -OH or -NH groups) [7, 13], and these have large effects on the simulation. They enable hydroxy groups within 2 or 3Å of one another to form interactions which are stable for up to one or two pico-seconds: the average vibration period of an hydrogen bond is 0.2 picoseconds [14]. Several observations can be made:

• Water molecules are seen to interact with hydroxy-groups on the chain. All of the hydroxy-groups on the chain are in close proximity to water molecules except for O(2) on the galactose ring which points inside the double helix cavity. This group becomes exposed later on when the structure loosens, and all of the hydroxy groups are able to form hydrogen bonds to water molecules at any time during the simulation (see Figure 1 for images). In some situations one water molecule was able to bind for a period of the order of picoseconds with two -OH groups around an *A-G* glycosidic linkage: this could be expected to slightly reduce the freedom of the linkage. At all times in the simulation water molecules were in close proximity to -OH groups, and were sometimes "dragged" along with them during conformational changes.

• Water molecules are constantly making and breaking structures within their bulk, this being a consequence of their strongly hydrogen—bonding nature. Such structures are observed to take the form of water molecules arranged in clusters where three or four molecules are able to hydrogen bond to each other simultaneously, with the arrangement existing for a period of time beyond that where non interacting water molecules would have diffused away (1 → 2 picoseconds). It is a short ranging effect which was not observed in the region immediately adjacent to carbohydrate chains. Although the modelling software does not permit water molecules within the bulk assembly to be individually highlighted, Figure 1 shows snapshots of the water molecules during the simulation.

• Ends of carbohydrate chains have extra -OH groups and tend to interact with each other. The extra -OH group on a terminal sugar ring and the fact that the ring has greater conformational space than one inside a chain explains why ends of chains were often in contact with several water molecules or with another chain.

• Water molecules placed inside the helix at the start of the calculation were able to escape during "breathing modes" displayed by the helix loops, which could open and close. This happened on a scale of tens of picoseconds. It is much less likely that the reverse process will occur due to the associated loss of entropy when a water molecule becomes restricted in a confined site. Therefore if water molecules are to exist inside the macromolecular helices there must be sufficiently favourable interactions to make the process thermodynamically viable. In this simulation such a favourable site does not appear to exist. On the agar double helix there

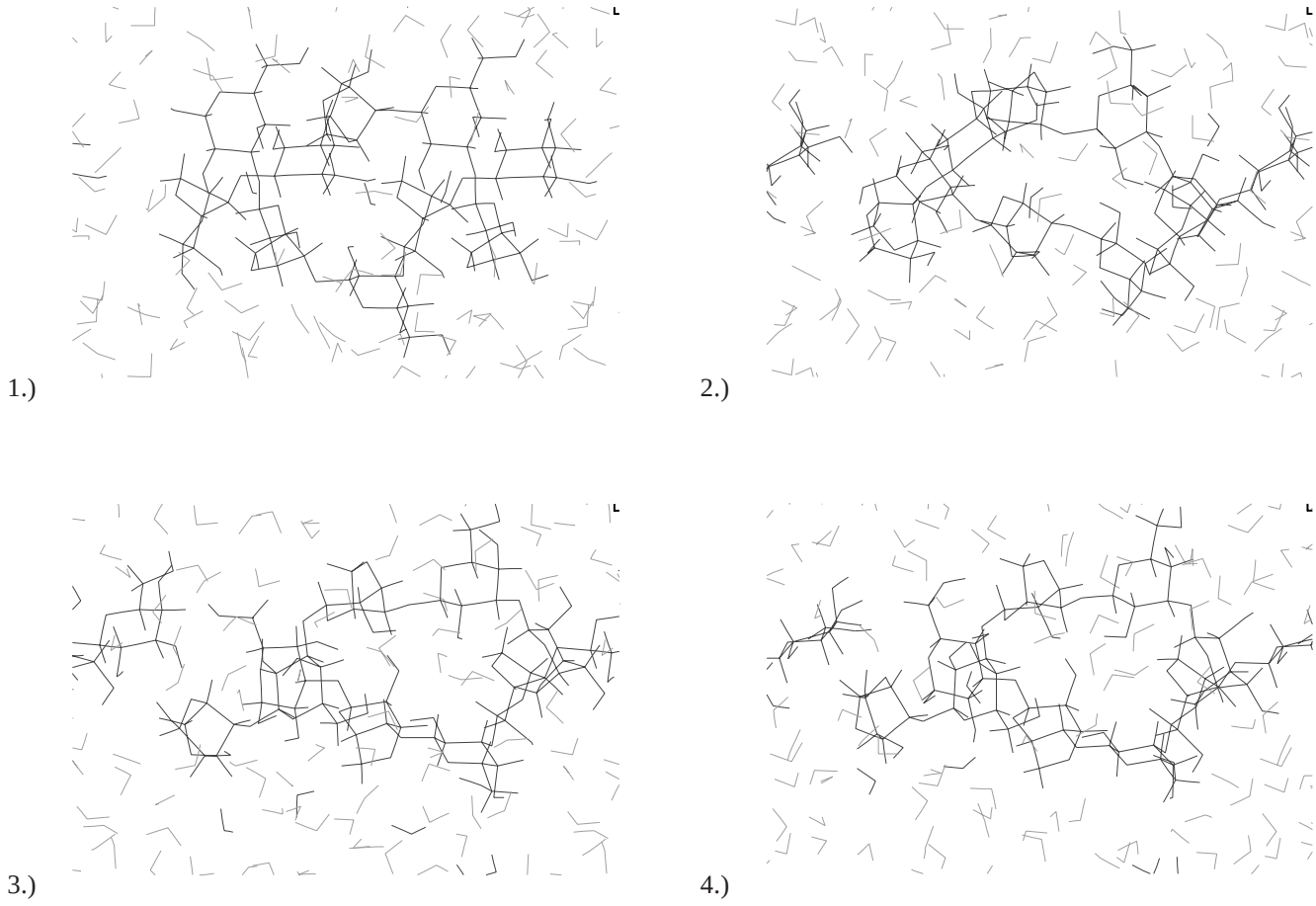


Figure 1. *Images from the dynamics calculation with explicit water molecules.1) After minimisation, but before dynamics; 2) after 0.03ns of dynamics calculations; 3) after 0.07ns of simulation. Notice that (unlike the simulation with no explicit water) the carbohydrate molecules remain adjacent to each other.; 4) The final image in the simulation shows very little change in the carbohydrate molecules.*

are a greater number of -OH groups on the outside than in the core, and in addition internal water molecules would have greatly restricted degrees of freedom. So, although there may be water molecules trapped in crystalline or fibrous forms [3], we propose that *internal* waters are not part of double helices in solutions, following the work described above.

• Intramolecular hydrogen bonding is observed intermittently during the simulation. Common arrangements are hydroxy-groups interacting on the same ring, for example O(6)H···O(4) on galactose rings. Hydroxy-groups on rings can also interact with glycosidic bridging oxygen atoms, and these arrangements were observed to persist for up to one picosecond.

• Scatter plots of dihedral angles visited show a more restricted behaviour than that seen without explicit water (compare Figure 2 with Figure 5), although the variation in total energy is similar. In addition the most visited parts of the potential surface are slightly different when explicit water molecules are present in the calculation; this is also demonstrated in the plots. In other words, the potential energy surface for the full explicit water treatment must be slightly different to that calculated using a simple dielectric constant of 80.

• No lubrication of the flexible chains was obvious, as this can only happen when water molecules are holding the Van der Waals surfaces of the carbohydrate apart. In our simulation only the ends of the chains were separated enough for water molecules to move between the strands (see Figure 1).

Dynamics calculations: no water molecules

The simulation without explicit water displayed much more structural variation and movement. Images from the 0.1 nanoseconds of motion are reproduced in Figure 3. Graphs of total energy variation with time (Figure 4) for both simulations show that evolution of the systems is stable.

The straight lines indicate the boundaries of the original cell; all material outside this has been generated using Periodic Boundary Conditions. After minimisation the original regular helical conformation has already been able to relax significantly, even though the values of the dihedral angles

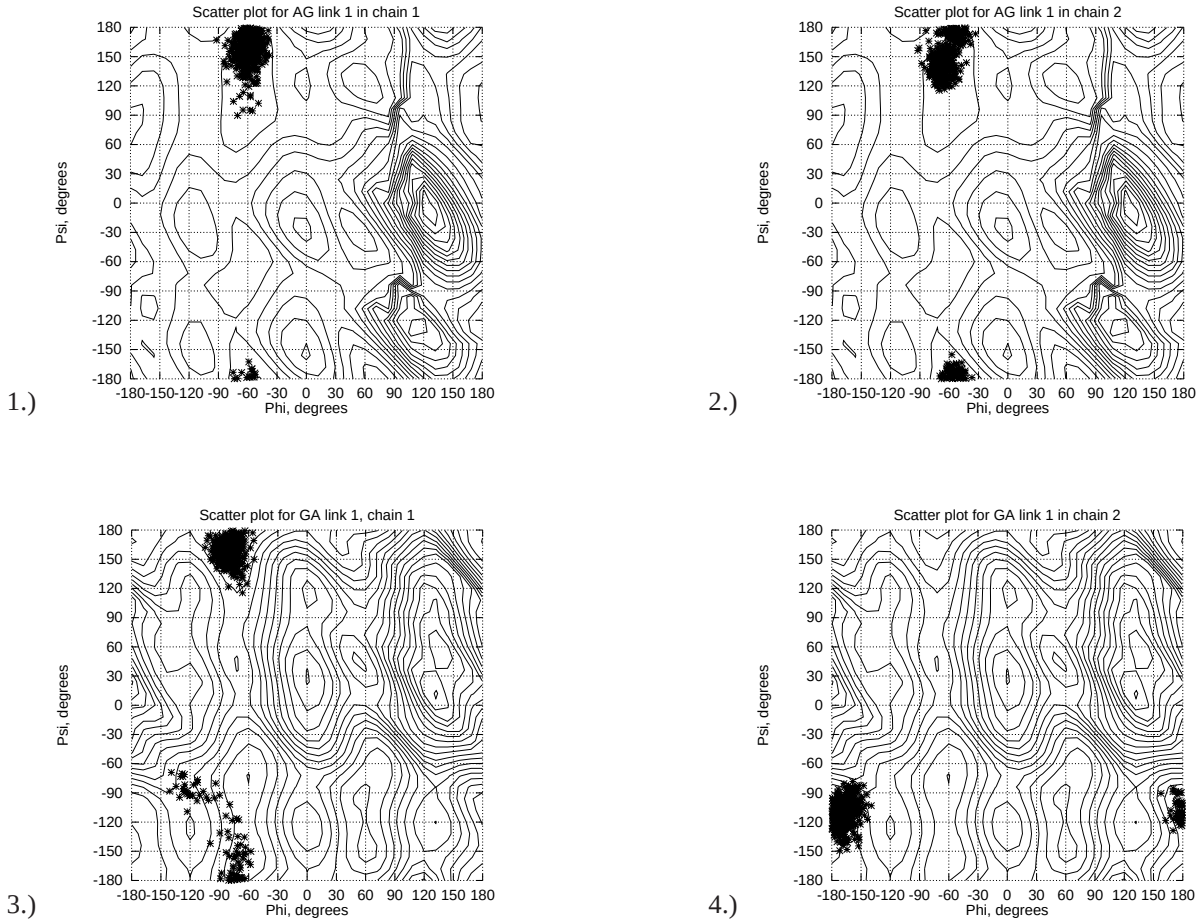


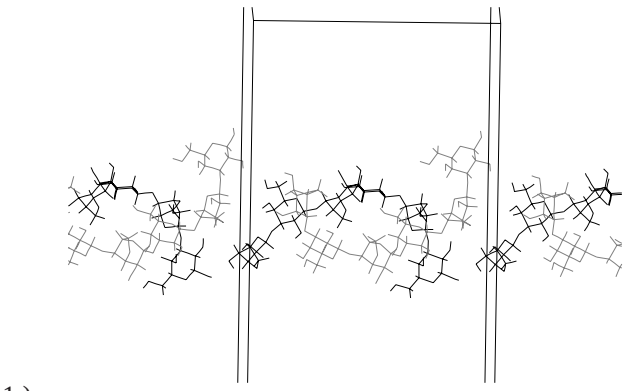
Figure 2. *Typical patterns of movement for dihedral linkages in the experiment with explicit water, shown by black dots (superimposed onto Ramachandran plots for isolated glycosidic linkages). The two angles phi and psi were defined using five atoms across the glycosidic linkage: [ring]O–C[1]–[bridge]O–C[n]–C[n+1]. 1) and 2):* A-G *links; 3) and 4):* G-A *links. Contour intervals are 1.3 and 1.2 kcal respectively.*

have only changed slightly. Upon warming up to 350K the two molecules are seen to unwind over a period of less than 0.01ns (1/10 of the simulation); this is clearly a rapid process, resulting in the velocity of molecule 1 being roughly equal in size and opposite in direction to that of molecule 2. By this stage in the simulation it is already apparent that the lack of specific water molecules has a large effect on the stability of the initial assembly. This is not surprising, and serves to confirm the expectation that hydrogen bonds between solvent and carbohydrate have a great influence on the freedom of chains.

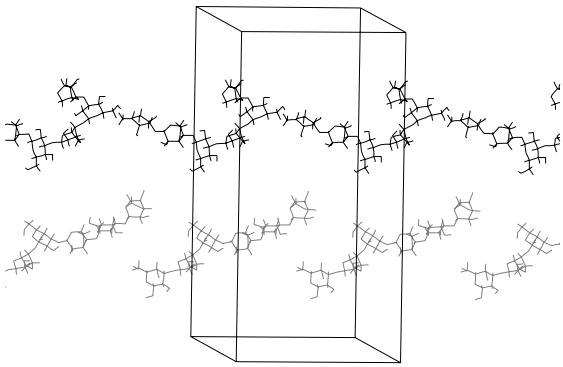
Following the initial dissociation away from a double helix there is a drift of molecule 1's against 2's in the array, with minor changes in molecular shape constantly occurring. The lack of explicit water molecules clearly permits much greater diffusion rates than were seen in the previous experiment. We suspect that the type of dimers formed in the early stages of a calculation (the first 0.1 nanoseconds) may be determined by initial orientation inside the Periodic Cell: more calculations are needed to investigate whether the system eventually behaves the same irrespective of assemblies being lined up end to end (when cells are reproduced) or placed at an angle inside each cell.

After 0.07 nanoseconds of dynamics, molecules from different cells collide and begin to interact by partially enveloping each other (see Figure 3, image 3)). This interaction continues until the end (0.03 nanoseconds) and a longer calculation may show that it can last for greater periods of time. The chains constantly alter their shape between extended, arc-shaped and coiled forms. The final situation is one of pairs of molecules, but not in a double helical conformation as at the start. Nevertheless on the 0.01 nanosecond timescale stable dimers of molecules do form rather than a completely dispersed solution.

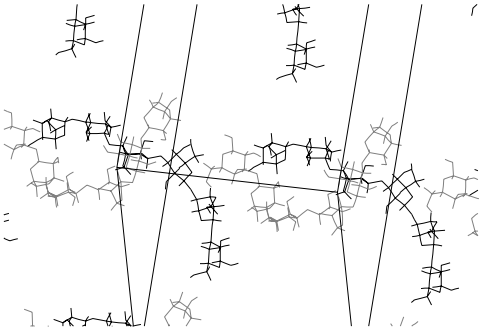
This second simulation shows that even very small chains are in a state of continual association into dimers (the current simulation does not provide evidence for or against the existence of double helices in macromolecular sized systems) and dissociation into a dispersed solution. In a bulk solution with explicit water this process would happen much more



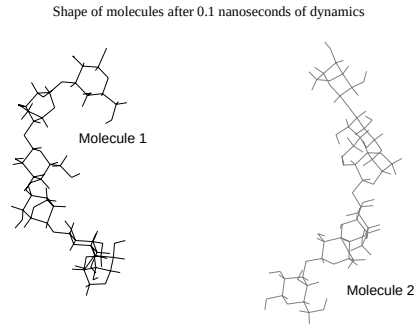
1.)



2.)



3.)



4.)

Figure 3. *Images from the dynamics calculation with no explicit water (dielectric = 80.0). 1) After minimisation but before dynamics; 2) after 0.02ns of simulation. The two molecules are progressing away from each other. 3) After 0.07ns molecules from different cells collide; 4) Two molecules in arc-shaped conformations which are the most suitable for intermolecular binding, as seen near the end of the simulation.*

slowly, but there would be a proportion of molecules associated all the time.

If longer chains are used the balance of enthalpy, ΔH, which favours dimerisation *versus* entropy, ΔS, which favours dispersion into solution would be different. With small molecules the benefit of total dissolution is greater, while the enthalpy of dimerisation for a long chain would be sufficient

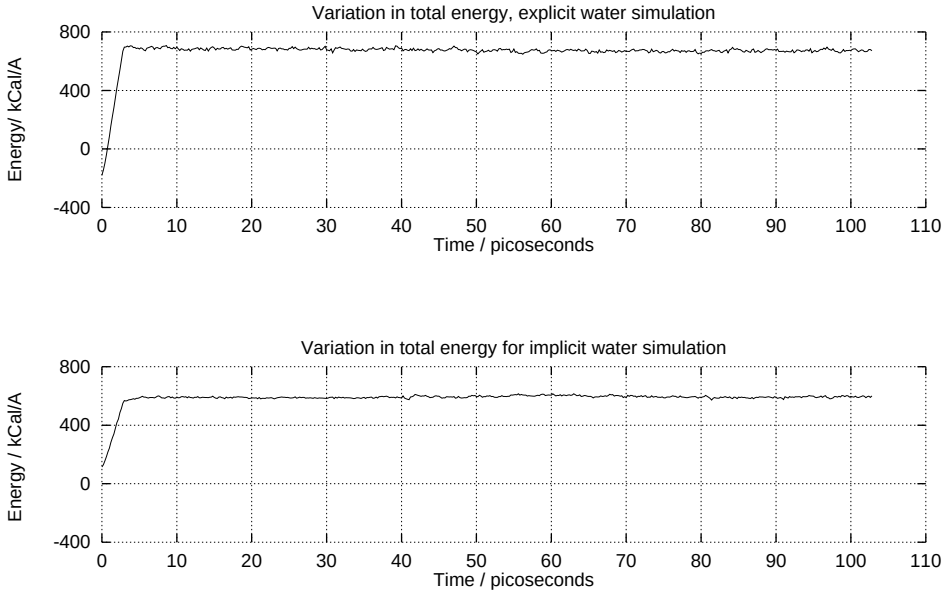


Figure 4. *Variation in total energy during simulations. Top: simulation with explicit water and dielectric of 1.0. Bottom: simulation with no water molecules and dielectric of 80.0.*

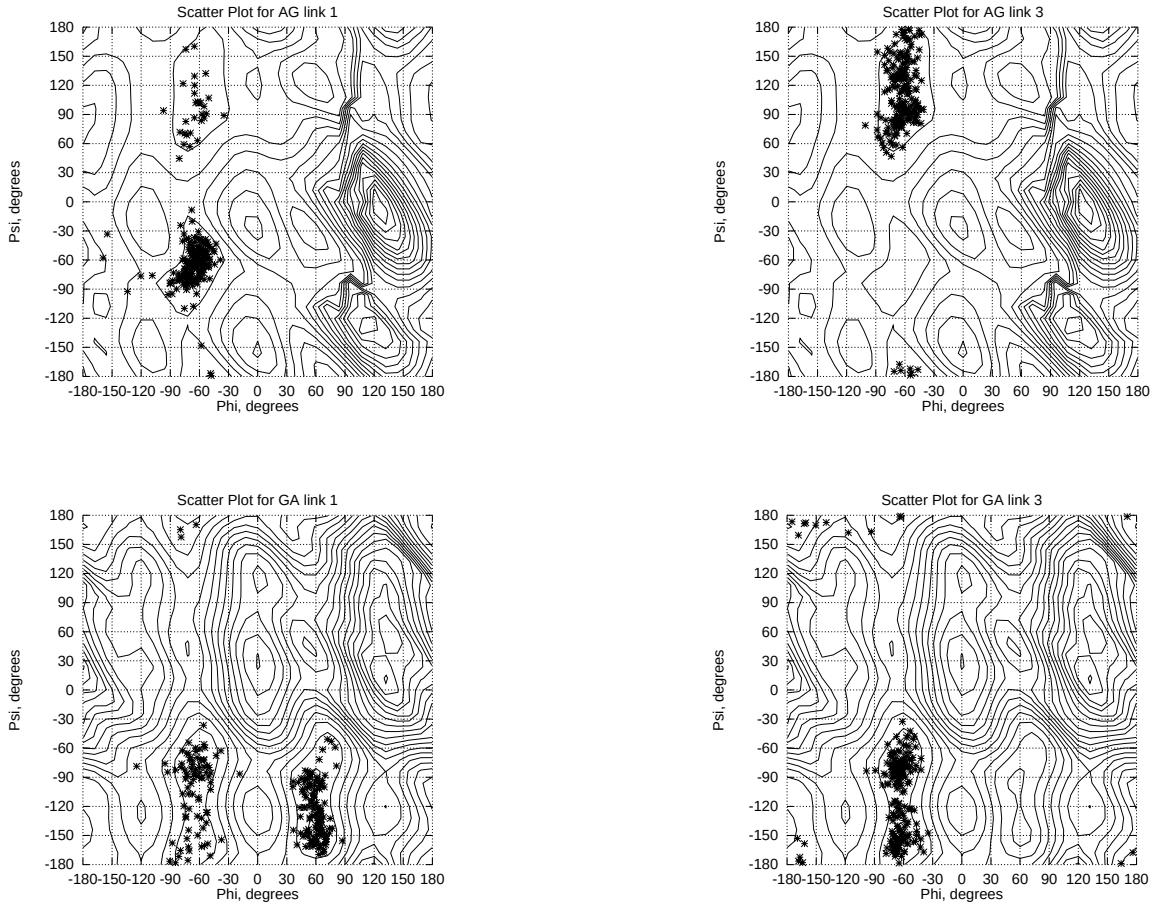


Figure 5. *Dihedral angles in the simulation with dielectric = 80 (no water) have been monitored and plotted onto energy surfaces previously obtained for isolated glycosidic linkages. The upper pair of plots are examples of A-G type linkages, while the lower pair are G-A scatter plots. Contour intervals are 1.3 kcal·mol^{-1} for A-G and 1.2 kcal·mol^{-1}for G-A. Different minima were preferred in the timescale examined for both cases, and more conformational space was explored than in the former simulation which used explicit water (this being shown by the more highly scattered nature of the data, and by correlation functions)*

to cause dimers to be the preferred form. The present simulation confirms the small molecule case. Ueda et al. [10] found that longer chains of the related polysaccharide β-carrageenan formed relatively stable dimers on a time - scale of 550 picoseconds.

Natural agars have considerable amounts of sulphate and methyl substitution on O(2) of the anhydrogalactose ring and/or O(6) of the galactose ring: these groups can be expected to affect dimerisation enthalpies by virtue of their bulk and charge. Some highly methylated agars are known to have elevated melting temperatures [15] indicating a large enthalpy of dimerisation or bigger junction zones. However small random amounts of substitution are more likely to disrupt junction zones. The general argument concerning enthalpy *versus* chain length is still valid, but predicting the effect of charged groups on water structure and molecular conformation is not trivial, and parameters for sulphate groups have yet to be fully evaluated.

Analysis of dihedral activity

Energy variation in calculations

All the glycosidic links in these simulations begin in the conformation

AG: (ϕ/ψ) = (-52.2/156.8)

GA: (ϕ/ψ) = (-123.9/-113.2)

(see earlier for definitions of these angles) but are minimised to positions which are closer to the bottom of the potential wells. For dynamics calculations therefore, all of the linkages start in practically the same position; close to the X-ray determined values [3] though not equal to them.

The simulations are conducted at 350K, and it is therefore not surprising that several minima are visited and con-

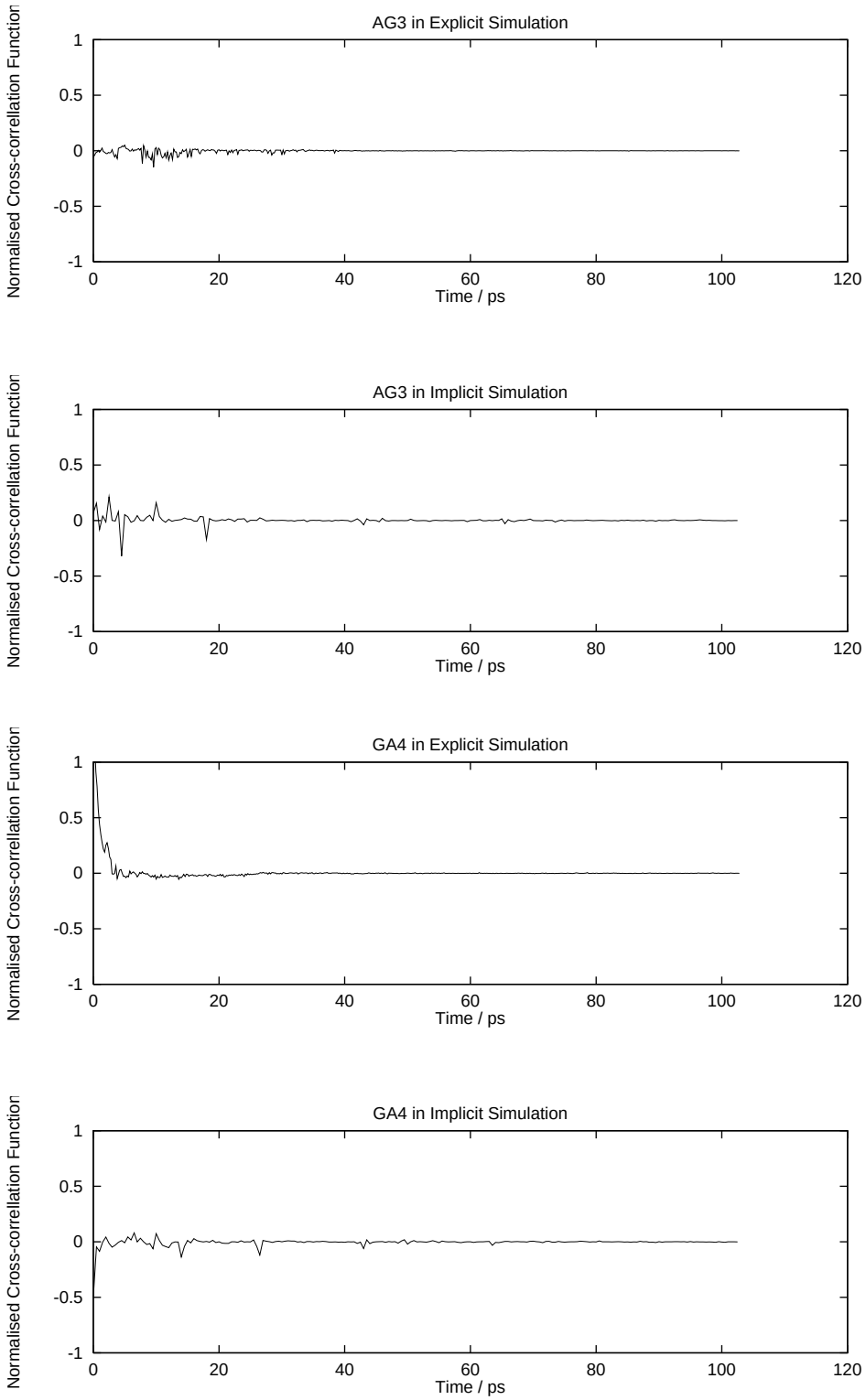


Figure 6. *Normalised Correlation Functions for dihedral angles. Top: for an A-G linkage in both simulations; bottom: a G-A linkage in each case. For method of calculation, see text.*

formational space is accessible to some extent up the sides of potential wells.

The total energy of a carbohydrate molecule in the simulation with no explicit water varies between 570 and 610 kcal·mol^{-1} (see Figure 4). Most of this variation is due to non-bonded interactions and torsional energies as the dihedral angles change. For the simulation involving explicit water molecules the variation is greater due to the large number of extra atoms, and cannot be analysed readily.

For the former case, taking the lower limit of the range to be the case where all the dihedral angles are relaxed at once (in the bottom of their potential wells), the variation in en-

ergy for one dihedral angle must be of the order of 4 kcal·mol^{-1} at 350K. ((610-570)/10; there are two molecules with five dihedral angles in each.)

Within potential wells the total energy oscillates inside a range of about 20 kcal·mol^{-1}, which equates to a maximum of 2 kcal·mol^{-1} for each dihedral angle. The extreme ends of the energy range reported above allow a change from one well to another or an exploration significantly away from the minimum position.

This should be compared with the height of energy barriers between two minima as seen on detailed Ramachandran plots. Indeed it can be seen from the plots that this barrier can sometimes be traversed see (Figure 5, plot 1).

Flexibility of carbohydrates and correlation functions

It has already been mentioned that in the two simulations (with explicit and implicit water) the carbohydrate chain displays different flexibility. This was concluded following observations on the movement of the chain during animated replays of the calculation, and also by studying the dihedral angle scatter plots (Figures 2 and 5), in which the more restricted dihedral angles in the explicit simulation resulted in a more concentrated scatter of data.

To provide some measure of this "stiffness factor" normalised correlation functions have been calculated[16, 17] for dihedral angles in both cases, using the statistical relation for the correlation coefficient between two variables *A* and *B* evaluated at each reaction coordinate:

$$c_{A,B} = \frac{\langle \delta A \, \delta B \rangle}{\sigma(A)\,\sigma(B)}$$

where

$$\sigma(A) = \left[\left\langle A^2 \right\rangle_{\text{ens}} - \langle A \rangle_{\text{ens}}^2 \right]^{\frac{1}{2}}$$

and $\delta A = A - \langle A \rangle_{\text{ens}}$.

These formulae were evaluated using a short fortran program for both simulations and results are given for two glycosidic linkages in each case in Figure 6. The same trends were displayed by all other dihedral angles.

All the plots begin with large fluctuations in the time correlation function and tend towards a limiting value later on. The increased stiffness of the carbohydrates in the explicit water calculation is reflected in a faster decay of the correlation function and a steadier limiting value as time progresses towards infinity. In the plots from the implicit simulation variations in the time correlation function continue to manifest themselves up to the end of the data.

Conclusion

From the above dynamics calculations it is seen that the helical nature of an agarose chain is tolerant to small changes in dihedral angles at the glycosidic linkage. However during dynamics calculations of a small chain the "perfect" X-ray conformation is soon lost (though less quickly in bulk water) as all the angles are independently varying, and the rings themselves can flex. It is a vanishingly small probability that all the torsions will be correct at the same time, but the average values may persist if the helical dimer is stable.

In real solutions helices may exist because of the following reasons:

• As the chain length increases the enthalpy of dimerisation will eventually be enough to favour association of separate strands, possibly into the helices proposed for strong gels and oriented fibres.

• For long chains there is a much lower proportion of terminal sugar rings, termed "ends", in relation to the number of internal sugar rings. Ends are more mobile so that short helices are more likely to be disrupted.

• Water molecules assist by binding to the outside of helices, encouraging their association and restricting their flexibility to some degree.

• Although the conformational space for the dihedral links is centred mainly about one or two minima, significant flexibility of the chains is possible by twisting the linkage to plus or minus 10 $\rightarrow$ 20 degrees away from the minimum. This allows dimers to adjust their conformation on a local scale without having to unwind.

The difference between the full explicit water treatment and the dielectric of 80 manifests itself in reduced conformational flexibility (less space is explored), and slower relaxation and tumbling of molecules. In other words the system takes longer to evolve when surrounded by explicit solvent molecules. Increased mobility of chains due to lubrication by solvent molecules appears to be very much smaller than the loss of mobility due to reduced diffusion rates, when the two dynamics calculations are compared. In addition the minimum energy conformations in explicit water are slightly different to those in a dielectric of 80, and further work is to be done in this area.

Acknowledgements: Unilever Research for financial support.

References

1. Craigie, J.S.; Jurgens, A. *Carbohydrate Polymers* **1989**, *11*, 265.
2. Rees, D.A. *Advances in Carbohydrates and Biopolymers* **1969**, *24*, 266.
3. Arnott, S.; Fulmer, A.; Scott, W.E.; Dea, I.C.M.; Moorhouse, R.; Rees, D.A. *J. Mol. Biol.* **1974**, *90*, 269.
4. Miller, I.J.; Falshaw, R.; Furneaux, R.H. *Carb. Res.* **1994**, *262*, 127.

5. Jimenez-Barbero, J.; Bouffar-Roupe, C.; Rochas, C.; Perez, S. *Int. J. Biol. Macromol.* **1989**, *11*, 265.
6. Asensia, J.L.; Martin-Pastor, M.; Jimenez-Barbero, J. *Int. J. Biol.Macromol.* **1995**, *17*, 137.
7. Weiner, P.K.; Kollman, P.A. *J. Comp. Chem.* **1981**, *2*, 287.
8. Homans,S. *Biochemistry* **1990**, *29*, 9110.
9. Larwood, V.L.; Howlin, B.J.; Webb, G.G. *J. Mol. Model.* **1996**, *2*, 175.
10. Ueda, K.; Ochiai, H.; Imamura, A.; Nakagawa, S. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 95.
11. Stellwagen, J.; Stellwagen, N.C. *J. Phys Chem.* **1995**, *99*, 4247.
12. Djabourov, M.; Clark, A.H.; Rowlands, D.W.; Ross-Murphy, S.B. *Macromol.* **1989**, *22*, 180.
13. DISCOVER Theory and Methodology, In: Biosym Manuals 1993, 1.
14. Vinogradov, S.N.; Linnell, R.H. In: *Hydrogen Bonding*; Van Nostrand: 1971.
15. Guiseley, K.B. *Carb. Res.* **1970**, *13*, 247.
16. Allen, M.P.; Tildesley, D.J. In: *Computer Simulation of Liquids*; Oxford Science: 1996.
17. Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.