Developing Analytical Tools for Saccharides in Condensed Phases

A thesis submitted to the
UNIVERSITY OF CAPE TOWN
in fulfilment of the requirements for the degree of

MASTER of SCIENCE

by
Michelle Kuttel
B.Sc(Hons)(University of Cape Town)

March 13, 1999

Abstract

Carbohydrates are conformationally very complex molecules. It is this complexity that lies at the basis of the important roles that these molecules play in many biochemical and biomaterial systems. Moreover, the unusual response of these macromolecules to their environment allow them to play these often critical roles. This is particularly true for solvated carbohydrates. A knowledge of the molecular structure of carbohydrates is essential for an understanding of their function and the molecular basis of their macroscopic properties. The details of solution structure have proven difficult to probe experimentally, but computer simulations are a means for examining solvent structure directly.

In this thesis we develop various computational methods for analysing saccharides in solution and in the solid state. These methods are applied to molecular dynamics simulations of maltose, hexa-amylose and a series of cyclodextrins in solution, in order to investigate the effects of water on these polysaccharides. Maltose is investigated because of its potential as a model for larger polysaccharides comprising $\alpha(1 \rightarrow 4)$-linked glucose monomers.

Solvation was found to effect the conformations of the saccharides studied considerably. In particular, the range of motion around the glycosidic linkage is increased. Comparison of the dynamics around the glycosidic linkages for the various simulation show that oligosaccharides linked via $\alpha(1 \rightarrow 4)$ glycosidic linkages have similar behaviour around this linkage. The saccharides studied were found to impose considerable anisotropic structure on the surrounding water which may give insights into their solution properties.

In addition to the studies in solution, a recently developed method for analysing the close contacts in crystal structures is applied to crystal structures of cyclodextrin inclusion compounds. It shown to be a useful tool for investigating hydrogen-bonding patterns in the cyclodextrins.

Acknowledgments

I would like to thank:

- My supervisor, Dr. Kevin Naidoo, for his unstinting assistance and advice thoughout the project.
- My co-supervisor, Dr. Susan Bourne, for her enthusiastic help with all things crystallographical.
- Prof. L. R. Nassimbeni for discussions of many aspects of the crystal structure analysis.
- The Foundation for Research And Development for their financial support.
- The USDA-ARS for their support in providing the computers on which the simulations were performed.

Presentations

Parts of this work have been presented at the following conferences:

1. CCP5 Annual Meeting, Edinburgh, Scotland, 7-9 September 1998
2. The 34th SACI Convention/ 7th International Conference on Chemistry in Africa, Durban, South Afica, 6-10 July, 1998
3. IUCr Structural Chemistry Indaba II, Kruger Park, South Africa, 3-7 August 1997

List of Abbreviations

Å	angstrom
CHARMM	Chemistry at Harvard Macromolecular Mechanics
CD	cyclodextrin
DMSO	dimethyl sulphoxide
fs	femotosecond
MD	molecular dynamics
NMR	nuclear magnetic resonance
NVT	canonical ensemble
ORD	optical rotation diffraction
ps	picosecond

Contents

Chapter 1

Introduction

Though carbohydrates are the most abundant class of biological compounds, they are familiar chiefly as storage molecules, in the form of starches and sugars, or natural structural materials, such as cellulose. However, carbohydrates have been found to perform a multitude of intricate functions, from cryoprotectants in desert plants to the mediators of molecular recognition. Currently the potential of the mundane starches for use as biodegradeble polymers is generating considerable interest.

Cellulose is the most common carbohydrate. This polymer, along with lignin and chitin, performs a structural function. Cellulose and lignin form the supporting cell walls and woody tissue of plants and trees, while chitin is found in the body shells of many invertebrates, such as crustaceans and insects.

Carbohydrates are found on the surfaces of membranes in the form of glycoproteins, glycolipids and proteoglycans. These macromolecules are essential components in various cell functions such as cell recognition by carbohydrate-binding proteins, cell-cell interactions and adhesion [1]. Many pathogens, such as viruses, bacteria and parasites, bind to the surface of host cells with the aid of these specific carbohydrate structures. Carbohydrates on the surface of blood cells specify the different human blood groups and are intimately involved in the immunochemistry of blood.

Thus, carbohydrates are of scientific interest for a variety of reasons, from a need to design biodegradable polymers to a desire to understand the mechanisms of cell recognition and immunology. Investigations into these diverse fields have a common goal to correlate the complex structure of a carbohydrate to its specific function, whether that be as a structural material or a cell-wall glycoprotein. However, carbohydrates are complex molecules and, compared to proteins and nucleic acids, relatively under-investigated. Much remains to be discovered.

1.1 Classification of the Carbohydrates

Carbohydrates are composed of carbon, hydrogen and oxygen atoms and have the general composition $C_xH_{2y}O_y$. The simplest carbohydrates are the saccharides, the smallest of these being the monosaccharides, or sugars. Monosaccharides form the basic building blocks of the more complex polysaccharides. They are subdivided according

to the number of carbon atoms they contain, with *tetroses*, *pentoses* and *hexoses* containing 4, 5 and 6 carbon atoms respectively. Most of the monosaccharides' carbon atoms are asymmetric, they are optically active and their various stereo-isomers possess distinct physical and biological properties.

Monosaccharides can exist either as straight chains or ring-shaped molecules. The open chain forms of monosaccharides are quite flexible, as they can rotate about each of their single carbon-carbon bonds. This flexibility often leads to the formation of a ring molecule when the carbonyl group reacts with one of the hydroxyl groups from the other end of the molecule. The resulting rings tend to contain five or six atoms, depending on the point of attack. Also, the closing of the linear monosaccharide to form a ring creates a new chiral atom at $C1$, called the anomeric carbon. As a result of this chirality, two different anomeric forms of the particular ring are possible, designated the α and β anomers. The cyclisation of a monosaccharide is an easily reversible reaction and inter-conversions between the different forms of the sugar will occur at room temperature in aqueous solution in a phenomenon known as mutarotation. Thus, in water an equilibrium between the various forms of the sugar will exist. The different forms of glucose are illustrated in figure 1.1.

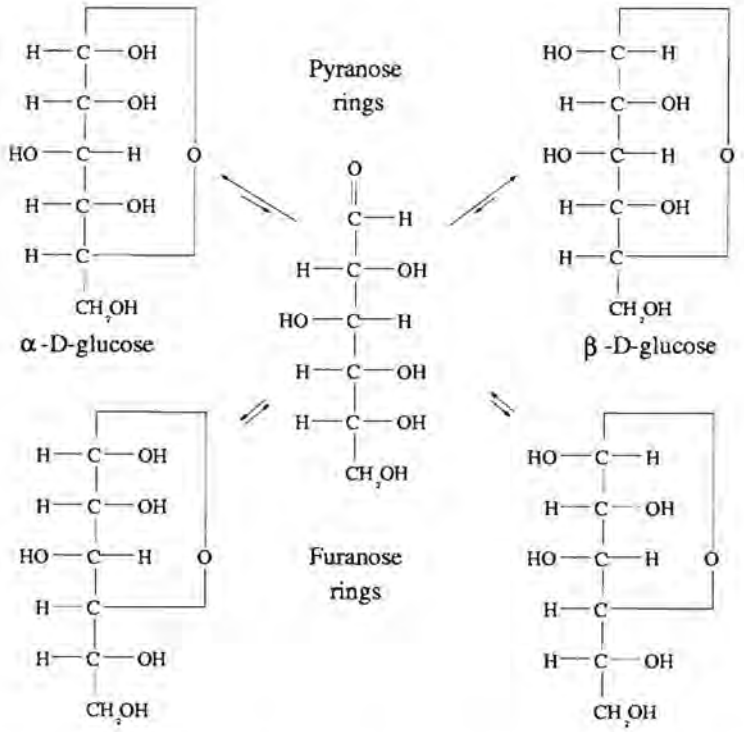

Figure 1.1: Fischer projections of the various open chain forms of glucose.

Monosaccharides may be joined together as the basic building blocks for larger molecules. *Disaccharides* can be formed via a condensation reaction between two monosaccharides with the elimination of a water molecule. Disaccharides are linked together by a glycosidic bond between the anomeric carbon and any one of the hy-

droxyls of the second monosaccharide. Glucopyranose disaccharides may have linkages that connect to $C1$, $C2$, $C3$, $C4$ or $C6$ in either the α or β form, therefore there are 10 different possibilities. For example, cellobiose is a glucopyranose disaccharide connected from the $C1$ carbon in one monosaccharide to the $C4$ carbon in the second monosaccharide. As the anomeric carbon in the glycosidic bond in cellobiose is in the β form, the monomer units are said to be joined via a $\beta(1 \rightarrow 4)$ linkage. In this thesis, we will consider D-glucopyranose (a six-membered ring form of glucose) units joined via the $\alpha(1 \rightarrow 4)$ glycosidic linkage exclusively.

In turn, many disaccharides can be joined together to form *polysaccharides*. Polysaccharides composed of only a few sugar residues are known as *oligosaccharides*. Polysaccharides are complex molecules with their constituent sugar residues arranged in periodic sequences (such as in the polysaccharides cellulose and amylose), in interrupted sequences, which contain periodic sequences with occasional departures from the pattern, or in aperiodic sequences with no regular pattern of monomer units at all. They may be linear, as in amylose, or branched, as in amylopectin.

1.2 Starch

Starch is the storage polysaccharide for glucose in plants and thus is the basis of the staple diets of the world - rice, maize, potatoes, wheat. It is abundant, cheap and familiar. Yet starch is a surprisingly complex and ill-understood compound.

Starch is a mixture of two polysaccharides, amylose (a linear polymer) and amylopectin (a branched polymers) comprising $\alpha(1 \rightarrow 4)$ linked glucose units. Amylose is found in algae and other lower forms of plants. Amylopectin is the dominant form of starch in the higher plants. For example, amylose comprises approximately 30% of cornstarch and has molecular weights of 200 000 to 700 000, while the amylopectin molecules have molecular weights as high as 100-200 million.

Starch exists in its native state as compact, microscopic granules that vary their shape, size and composition according to the biological origin. These granules are composed of concentric rings of alternating crystalline lamella (mostly amylopectin) and amorphous regions (containing the branch points of amylopectin plus amorphous amylose) [2]. Starch granules have the advantage of being insoluble in cold aqueous solutions, and therefore will not cause any osmotic imbalance when accumulated in large amounts.

There is a significant difference in the mechanical properties of starch from different sources. This is a result of their different ratios of amylose to amylopectin and also the varying amounts of crystallinity within the granule, as well as the size of the starch granules.

1.2.1 Amylose Structure

The direct relations between the structure and properties of starch and modified starches are extremely complex and not very well understood. In this study, we are

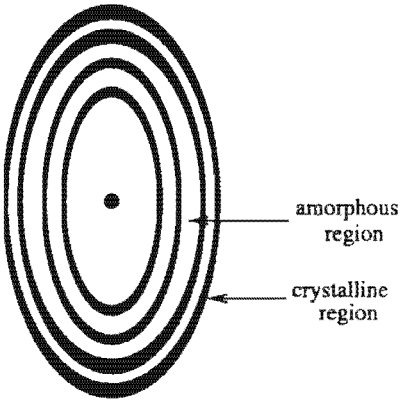

Figure 1.2: A schematic starch granule.

exclusively concerned with the properties and structure of amylose. Amylopectin, as a branched macromolecule, is the more complex of the two components of starch and the logical approach is to start with the simpler molecule.

Amylose crystallises in different helical forms depending on its environment (figure 1.3). In starch granules, amylose chains are naturally crystalline in forms known as the A form (in cereals) and the B form (in tubers). These are not yet fully characterized, and there has been some debate about the structures [3], but it seems from X-ray diffraction studies that these forms are both left-handed double helices [3, 4, 5] which mainly differ in the water content of the unit cells. In the A-type the double helices are closely packed, probably without water molecules, whereas in the B-type, packing is more loose and water molecules fill the spaces [5].

V-amylose is the only fully characterized crystalline form of amylose. It forms in the presence of suitable small molecules, such as iodine. These molecules are intercalated within the amylose helix and this process accounts for the blue-black colour of starch treated with iodine. V-amylose is a single stranded, left-handed helix with 6-glucose units per turn.

1.2.2 The Potential of Amylose for Use as a Biodegradable Polymer

Starch and starch derivatives are primarily used as food additives, particularly as thickening agents [3]. However, there has been increased interest recently in the use of starch, be it amylose or amylopectin, as a biodegradable plastic [7, 8]. Starch is attractive material to be used for this purpose, as it is abundantly available, cheap and rapidly degraded by biological enzymes.

Starch is different to synthetic polymers in that it contains a mixture of polysaccharides (amylose and amylopectin) and has a granular structure. However, in common with synthetic, semi-crystalline polymers, starch undergoes a glass transition in addition to regular phase transitions and is plasticised by small molecules such as water [9, 10, 11].

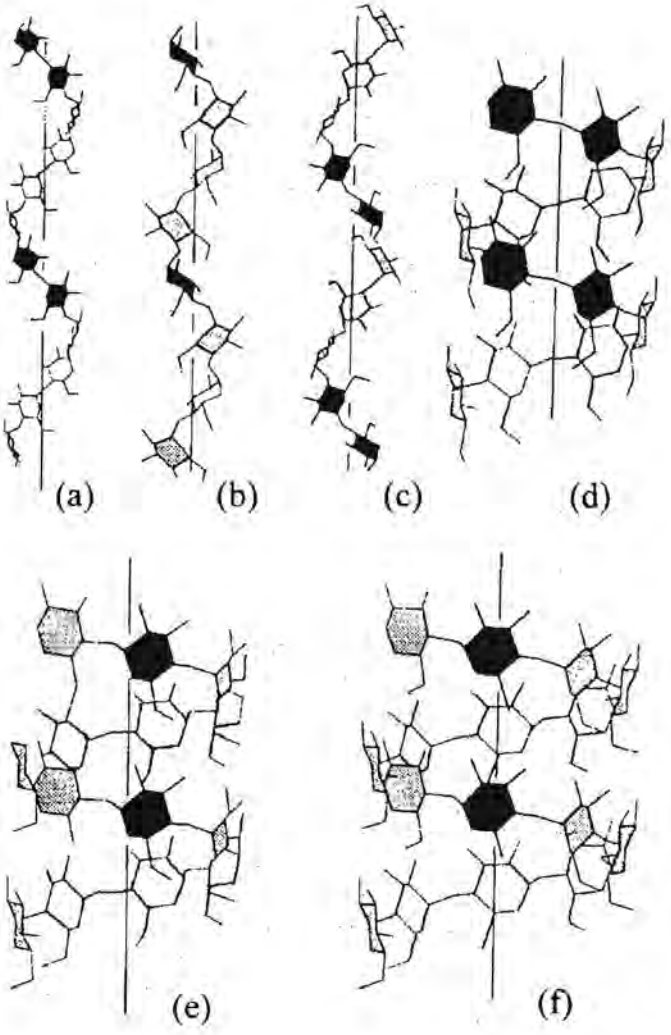

Figure 1.3: *Conformational forms of amylose, all of which can be regarded as the same left-handed helix in various states of extension. **(a)** Double helix of amylose in the A form, **(b)** and **(c)** extended forms of the helix which forms in the presence of potassium bromide (4 glucose units per turn) **(d)** V form of amylose (6 glucose units per turn) **(e)** and **(f)** compressed forms of the helix with eight units per turn obtained by crystallizing with larger guests.

*from [6]

Gelatinised or melted starch has been proposed to replace non-biodegradable polymers for use in disposable, single-use commercial products such as eating utensils, plates and planting pots. A biodegradable thermoplastic starch would serve an important function in reducing the total amount of plastic waste. The potential for starch to be used for compression-molded materials has been investigated [11, 12], and moulded starch golf tees and foamed packing material are already being made. Generally starches are mostly made up of amylopectin (approximately 80%). However, high amylose starches (amylomaize has 50-70% amylose content) have been shown to have superior mechanical strength, greater flexibility and slower physical aging. Hence, plastic materials prepared from gelatinised or melted high-amylose cornstarch are predicted to have greater commercial desirability compared to products produced from normal cornstarch. However, starch as a material has a few undesirable properties, such brittleness as a result of rapid physical aging [12] and swelling on contact with water. Lack of information about starch structure and the response of its constituent molecules to environmental stresses largely hold up research in this area. Therefore, we need to understand the structure of starch more thoroughly before the design of a polymer with the ideal properties of a controllable service life followed by rapid degradation is possible.

1.3 $\alpha(1 \rightarrow 4)$-Linked Polysaccharides

The ultimate goal of this research is to investigate the conformations and properties of large $\alpha(1 \rightarrow 4)$ linked polysaccharides, particularly amylose. However, amylose itself is far too big a molecule to be simulated satisfactorily on computer using the methods currently available. Therefore, it can be useful to study some smaller oligosaccharides as models for starch in the hope that the results can be extrapolated to larger systems. Where these molecules are themselves of scientific interest, the results can be useful in their own right. The oligosaccharides under study in this thesis are all composed of *D*-glucopyranose monomers linked via $\alpha(1 \rightarrow 4)$ glycosidic linkages, and thus can all be considered in some sense models for amylose. As the constituent sugar residues and glycosidic linkages are identical, they are expected to exhibit some similar physical properties.

A study of a simplified system is a good starting point for a general investigation into a particular class of polysaccharides, as a thorough understanding of the basic repeat unit is a necessary first step in the investigation into the structure of a polymer. Therefore, the approach taken for our investigation is to study the simplest monomer unit initially, the maltose disaccharide, and then extend the investigation to larger oligosaccharides: an amylose fragment and three cyclodextrins. In this way, the dependency of the properties and conformation of the polysaccharides on the dynamics of the $\alpha(1 \rightarrow 4)$ glycosidic linkage in the constituent disaccharide may be investigated.

1.3.1 Maltose

Maltose is a disaccharide composed of two D-glucopyranose monomers linked by an $\alpha(1 \rightarrow 4)$ glycosidic linkage (figure 1.4).

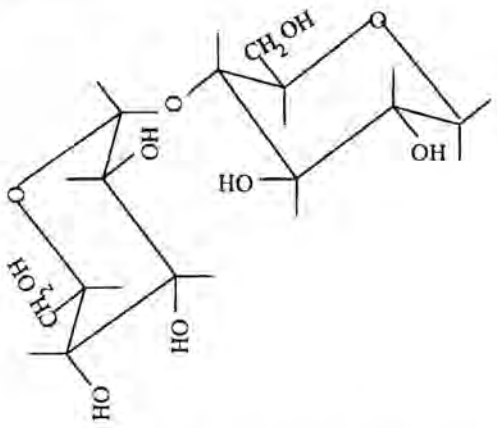

Figure 1.4: The maltose disaccharide.

This molecule serves as a general conceptual model for $\alpha(1 \rightarrow 4)$ linked polysaccharides in the same way that the alanine dipeptide, *N-acetylalkyl-N-methylamide*, does for proteins [13, 14]. Thus, through the study of maltose, we expect to gain insight into the conformation of polysaccharides composed of the maltose dimer.

1.3.2 A Model for Amylose: a Hexa-Amylose Strand

To extend the maltose model, we investigated a strand comprising six glucose monomers linked by $\alpha(1 \rightarrow 4)$ glycosidic linkages (figure 1.5).

Figure 1.5: The hexa-amylose strand.

This strand could be considered to be a model for a long amylose chain, as it

constitutes a single turn of a helix in both the A-, B- and the V-helical forms of amylose.

1.3.3 Cyclodextrins

The cyclodextrins, or cyclo-amyloses, are a group of torus-shaped oligosaccharides formed as degradation products from starch. They consist of $\alpha(1\rightarrow4)$ linked D-glucopyranose units. The most common of the cyclodextrins are alpha-, beta- and gamma-cyclodextrins, comprising 6,7 and 8 glucose units respectively.

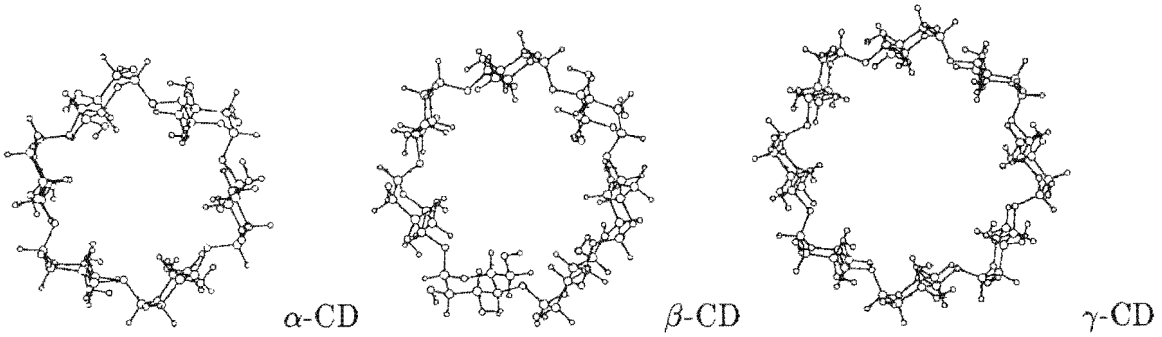

Figure 1.6: α-, β- and γ-cyclodextrin.

The cyclodextrins have a hydrophilic exterior and a hydrophobic central cavity. In addition, the larger opening in the cyclodextrin, containing the 2-OH and 3-OH groups, is intensely hydrophilic; the opposite, narrower opening, ringed by the CH_2OH groups is considerably less hydrophilic [15]. The hydrophilic and hydrophobic regions in the cyclodextrin torus are illustrated in figure 1.7.

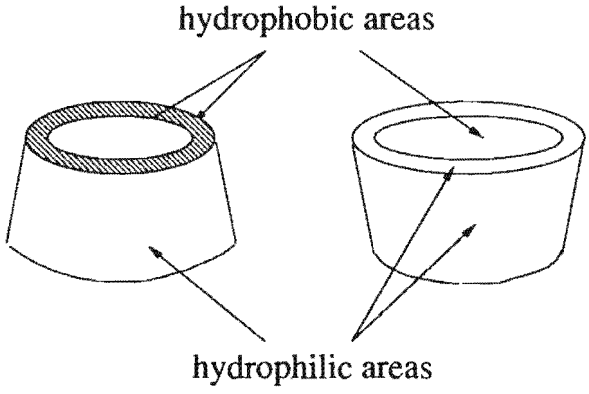

Figure 1.7: Schematic representation of the hydrophilic and hydrophobic regions of a cyclodextrin torus.

Cyclodextrins represent closed turns detached from the starch helix, therefore they can be thought of as models for starch. Thus, our six-unit amylose strand is similar to α-cyclodextrin.

However, the cyclodextrins are certainly not only models for starch. They are of pharmaceutical interest as they are non-toxic and can form inclusion compounds with a variety of guest molecules. A guest molecule of the appropriate size and shape and a lower polarity than water can be accommodated in the cyclodextrin cavity in aqueous solution.

As a result of their ability to form inclusion compounds, cyclodextrins have been extensively used for the separation, protection and solubilization of various kinds of compounds. They have applications in the pharmaceutical, cosmetic, pesticide and food industries [16]. The inherent chirality of cyclodextrin enables them to form diastereomeric complexes with enantiomeric compounds. This has resulted in cyclodextrins being widely used as chiral stationary phases in gas chromatography [17]. In addition to host-guest selectivity, cyclodextrins display some catalytic properties related to those of enzymes [18].

1.4 Carbohydrate Structure

The chemical function of a molecule is intimately related to its three-dimensional structure, known as the conformation. Therefore, it is clear that carbohydrates, with their multitude of biological roles, will show great structural variation. This is in fact the case, and so an understanding of the operation of a carbohydrate is largely dependent upon a knowledge of its preferred conformations.

1.4.1 Conformational Descriptors for $\alpha(1 \rightarrow 4)$-Linked Polysaccharides

In this thesis, we are principally concerned with $\alpha(1 \rightarrow 4)$-linked polysaccharides. Thus, a general outline of the widely-accepted conformational descriptors for these molecules would be useful. Only those aspects of carbohydrate structure necessary for a discussion of the $\alpha(1 \rightarrow 4)$-linked polysaccharides have been included here.

Pyranoid Ring Conformation

As we have discussed, monosaccharides in aqueous solution exist in an equilibrium between a variety of of different forms, including the open chain form and the α and β anomers of both five-membered (furanoid) and six-membered (pyranoid) ring structures. However, the most common situation for the aldohexoses is for the equilibrium to be primarily between the α and β anomers of the pyranoid ring.

However, each of the ring forms has a variety of different conformations. A six-membered pyranose ring has two possible chair forms and several boat and skew forms. In order to classify the ring shapes, a notation is used in which the approximate conformation of the ring is indicated with an italic, capital letter, which designates the ring shape, and numerals, which distinguish between the variant forms of each shape. For example, ${}^4C_1$ denotes a chair conformation in which the $C4$ lies above and the $C1$

below the plane of the pyranose ring when it is viewed so that the numbering appears clockwise from above.

The chair conformations are preferred, as the boat and skew forms are significantly higher in energy. In particular, the $^4C_1$ conformation of α-D-glucose, which has all the ring substituents in the equatorial position, is preferred to the $^1C_4$ conformation in which all substituents are in the axial orientation. X-ray diffraction and NMR studies have shown this to be the conformation that exists in both the solid state and aqueous solution, both in simple sugars and larger polysaccharides. [6].

A general definition of ring puckering coordinates has been defined by Cramer and Pople [19]. For six-membered rings, the puckering may be described by a *spherical polar set* (Q, θ, ϕ), where Q is the total puckering amplitude and θ specifies the conformational form. This coordinate system permits the mapping of all types of puckering (for a given amplitude Q) onto the surface of a sphere (Figure 1.8).

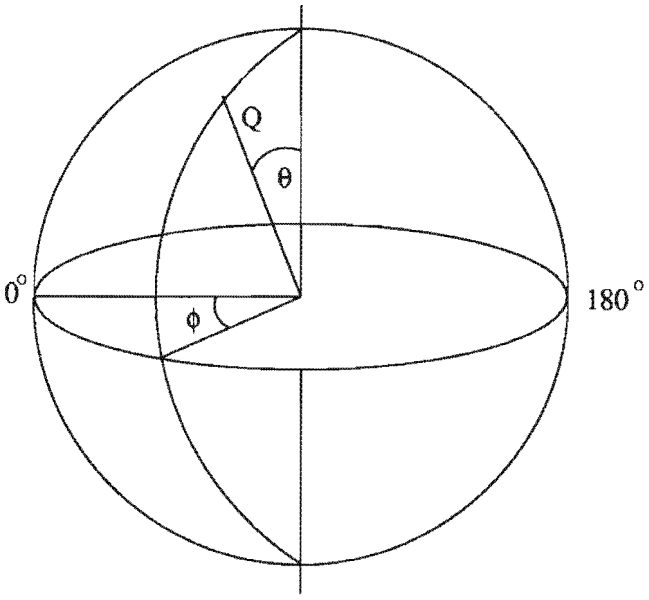

Figure 1.8: Puckering parameters for pyranoid rings.

The polar positions ($\theta = 0$ or $180°$) correspond to a chair conformation. Values of θ around $90°$ correspond to various twist-boat forms.

Glycosidic Torsion Angles

The glycosidic bonds form the backbone of polysaccharides and are the source of most of the molecules' flexibility. The overall conformation of a pyranose polysaccharide is chiefly determined by rotations about these bonds, as the $^4C_1$ chair form of the pyranose ring is rather rigid [6]. Two torsion angles, ϕ and ψ, are used to characterise these bonds (see figure 1.9).

There are several alternative definitions for the ϕ, ψ angles used in the literature, but all are equivalent in their description of the torsional motion about the glycosidic linkage. We used the definitions:

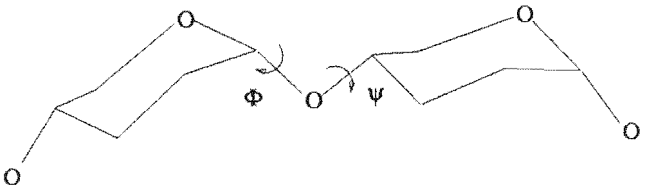

Figure 1.9: Notation for the torsion angles specifying the orientation of glycosidic bonds.

$$\phi = H_1 - C_1 - O'_4 - C'_4$$
$$\psi = C_1 - O'_4 - C'_4 - H'_4$$

However, for comparison purposes, it was sometimes necessary to use the definitions:

$$\phi^{a1} = O_5 - C_1 - O'_4 - C'_4$$
$$\psi^{a1} = C_1 - O'_4 - C'_4 - C'_5$$
or
$$\phi^{a2} = O_5 - C_1 - O'_4 - C'_4$$
$$\psi^{a2} = C_1 - O'_4 - C'_4 - C'_3$$

Our definition generally has a slightly larger standard deviation, due to the greater mobility of the hydrogen atoms.

Primary Alcohol Orientations

The orientation of the 6-hydroxymethyl group in relation to the pyranoid ring in hexoses is defined by two dihedral angles, $O_5 - C_5 - C_6 - O_6$ (ω) and $C_4 - C_5 - C_6 - O_6$ ($\omega 2$). In principal, these dihedrals can adopt any value between $-180°$ and $+180°$, but generally one of three staggered orientations occurs. These orientations are designated *tg*, *gg* and *gt*, respectively (Fig. 1.10), where the first letter refers to the *trans* or *gauche* orientation of the first dihedral and the second the orientation of the second dihedral.

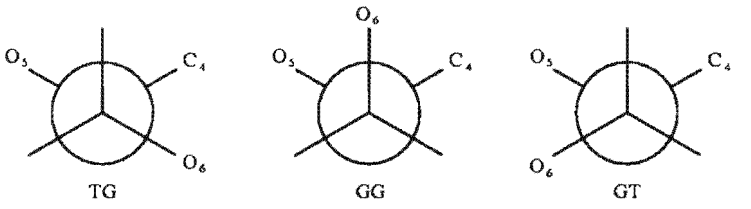

Figure 1.10: Definitions of the primary alcohol positions for a sugar ring.

The orientations of the primary alcohol positions affect the energy of the different ring conformations.

Polysaccharides

The chain is numbered from the reducing glucose residue to the non-reducing glycosyl group. Thus i refers to a particular saccharide unit in the polymer chain, $(i-1)$ to

the adjacent unit in the direction away from the non-reducing end and $(i+1)$ in the direction of the non-reducing end.

1.4.2 Environmental Effects

For a complete study of a biological function of a molecule, it is important to know its structure and properties in its natural environment as well as the effects on the molecule of changes in the environment, such as an increase in temperature or a decrease in pH. This is particularly true of carbohydrates, which are very sensitive to changes in their surroundings. Polysaccharides are intolerant of severe heat treatment and extremes of pH.

Most carbohydrates are present in an aqueous environment. As they play a significant role in many important biological processes, it is important to understand the aqueous solution behaviour of carbohydrate polymers. Carbohydrates interact strongly with water via their numerous hydroxyl groups, and can have a structuring affect on the solution molecules around them. This ability of some carbohydrates to structure the water around them is exploited by a variety of drought and cold-resistant plants. Maltose, along with trehalose, fructose and sucrose, has been shown to to inhibit dehydration, though trehalose is the most effective[20].

The structural details of specific carbohydrates, particularly their flexibility and the location of the hydroxyl groups, determine exactly how they interact with water and the degree of structuring they impose on the surrounding solvent. A small change in solute structure can have a large effect on the solvent structuring. For example, the solution properties of cyclic and linear oligosaccharides differ dramatically. While the α-, β- and γ-cyclodextrins are of limited solubility in water, with β-cyclodextrin the least soluble [16], their linear analogs are almost indefinitely soluble [5]. Evaporation of aqueous solutions of linear oligosaccharides yields glasses, whereas cyclodextrins crystallise.

An analysis of the structural properties of carbohydrate systems depends on a detailed understanding of the effect of solvent on the stability of various conformations as the phenomenon of solvation can radically alter both the geometry and the relative stabilities of various conformations.

The Effects of Environment on Starch

The principal hurdle to the use of starch as a biodegradeble polymer is its sensitivity to water. Dry starch does not melt, but rather dehydrates, cross-links and then burns. In contrast, when starch granules are heated in water to $> 90°C$ they gelatinise. During this process, the texture in the crystalline regions of native starch is destroyed as a result of irreversible changes (hydrogen-bond breaking, water uptake, melting of double helices) that take place. This can be detected as a glass transition. The glass transition temperature, T_g, is the temperature above which segmental motion of polymer molecules takes place. Water plasticises the polymer, decreasing the T_g and allowing it to melt and be processed as a conventional plastic. The effects of the amount of water

and other plasticisers on the mechanical properties of thermoplastic starches have been determined by several researchers [11, 10, 9]. It has been found that the mechanical properties of moulded starch are very sensitive to water content and, as the plasticising water is lost, the moulded piece of starch is likely to shrink and become brittle. Absorption of water and an increase in temperature will also affect a moulded item; lowering the Tg and consequently changing the mechanical properties of the polymer.

1.4.3 Experimental Investigations into Carbohydrate Structure

There have been various attempts to probe the structures of carbohydrates experimentally, using methods such as X-ray diffraction, NMR and optical diffraction. We will give an outline of the information obtained from these methods and their drawbacks below.

- **X-ray diffraction** is principally of use when analysing the structures of polysaccharides in the solid state. However, the crystallization of non-cyclic and branched oligosaccharides is problematic, as the molecules are flexible and adopt a variety of conformations that are stabilized by hydrogen bonds. Though the crystal structure of maltose [21] and other disaccharides have been reported, crystal structures of large oligosaccharides remain elusive. Results from fiber diffraction studies of polysaccharides have often proved ambiguous, requiring additional information for their interpretation [3]. In contrast, the cyclic cyclodextrins produce good crystals as they are far more rigid molecules. Consequently, there are a wealth of crystal structures of cyclodextrins and their inclusion compounds. Fiber diffraction can also be used to analyse polysaccharides, but the results are often ambiguous and require extra information for their interpretation.

  Crystal structures are of limited usefulness when one wishes to gain information on the solution structure of carbohydrates, as it is by no means certain that the conformation adopted by the polysaccharide in the solid state will be similar to that in solution. Packing effects as opposed to the effects of solvation may lead to widely differing structures. Another consideration is that static crystal structures do not show the dynamic effects of rotations about torsion angles that significantly affect the properties of polysaccharides.

- **NMR** is usually the method of choice for investigating the conformations of polysaccharides in solution. Glucose and polyglucoses have been investigated by NMR in DMSO [22]. NMR studies have been performed on amylose in DMSO [23] and water [24]. NMR has been used to study the structures of cyclodextrins and their complexes in solution [25]. In addition, NMR has been the primary means of investigating the conformations of the glucose hydroxymethyl group in solution [26].

  In principle, NMR could provide much of the desired information about oligosaccharide conformations. However, it unfortunately has two primary drawbacks as a method for investigating solution conformations of saccharides. Firstly, as

water is an inconvenient solvent for this analysis, solvents which do not interfere with carbohydrate NMR experiments, such as dimethyl sulphoxide (DMSO), have been principally used in NMR structure determinations [23, 22, 24]. However, there is no guarantee that the conformations observed in DMSO are the same as would be observed in a more polar solution. It has been shown in optical rotation experiments that a change in solvent has a marked effect on the solution conformation in polysaccharides [27].

Secondly, conformational flexibility may complicate the picture further as rotation about glycosidic linkages in general are fast on the NMR time-scale, resulting in average conformations predicted by NMR for multiply bonded glucose monomers. Thus NMR studies produce no certainty that an actual conformation is being observed, instead of a pseudo-structure comprising the dynamic average of several structures.

The limitations of experimental methods used to resolve the solution structure of disaccharides and more so polysaccharides have led to a successful collaboration between experimentalists and computational chemists. In particular, Molecular Dynamics (MD) simulations have been used to interpret information NMR experiments to elucidate the conformation of oligosaccharides [28].

- **ORD** As polysaccharides are optically active, various optical rotation experiments have been used to investigate their solution structures [27, 29, 30]. An innovative approach involving the correlation of optical rotation measurements with linkage conformation has been useful in analysing dissacharide solution structures [30]. However, chiroptical data generally cannot be used to make an unambiguous structure determination without additional information. This is particularly true in cases involving conformational averaging.

1.4.4 Computational Investigations into Carbohydrate Structure

Computational techniques are highly suited to investigate molecules with a large degree of conformational flexibility. The primary focus of computational investigation into biological molecules in water has been on proteins, however, a growing number of calculations are being performed on saccharides.

Since 1987, various simulations of carbohydrates in vacuum [31] and in aqueous solution [32, 33, 34, 35, 36, 37, 38] have been published. In the case of cyclodextrins, simulations have been performed for the crystalline state and for cyclodextrins in aqueous solution [39, 40], as well as several studies of cyclodextrin inclusion compounds in vacuum [41, 42, 43, 15].

Though quantum mechanical calculations have been performed on glucose [44], the large size and structural complexity of carbohydrates usually precludes the use of this method.

The majority of dissacharide modelling studies have been molecular mechanics calculations. Static conformational energy calculations in particular have been widely used

to characterize the conformational behaviour of polysaccharides in vacuum [45, 31]. These are usually interpreted in the form of two-dimensional Ramachandran maps of the conformational energy as a function of the glycosidic torsion angles ϕ and ψ. However, these maps are often unsatisfactory, as no effects arising from the dynamic motion of the carbohydrates is taken into account.

Molecular dynamics studies of saccharides in solution and in vacuo have have recently increased in number [46]. Though vacuum studies can be useful, of principal interest is the simulation of carbohydrate molecules in solution. However, because of the increased complexity of a solvated system, molecular dynamics simulations exploring the water-carbohydrate relationship have been less numerous. The reasons are firstly that only a few accurate atomistic force fields for carbohydrates have been published [47, 48, 49] and secondly that the solute-solvent interactions are complex and require a detailed knowledge of the carbohydrate-water conformational space before a general theoretical model can be proposed.

Monosaccharide-water stereo-chemistries are more easily investigated than those for polysaccharides, since the general conformational space is significantly less complex. Molecular dynamics simulations in solution have been performed for the the monosaccharides α-D-glucose [32], β-D-manose, and (α,β)-D-xylose [33] as well as β-D-galactose and (α,β)-D-talose [38]. A few disaccharides such as maltose [34, 50, 51], trehalose [36], sucrose and carrabiose [52] have been studied computationally in aqueous solution.

Cyclodextrins have been the focus of a large number of computational studies in the last few decades [53]. However, the large majority of these have been molecular mechanics calculations aimed at predicting the minimum energy structures of cyclodextrin inclusion compounds. Far fewer molecular dynamics simulations have been performed and, of these, only a handful have simulated cyclodextrins in water. In 1988, Koehler, Sanger and van Gunsteren performed a short (90ps) simulation of α-cyclodextrin in water [39] in a truncated octahedron containing 611 water molecules. Recently, a study of β-CD in water has been published [40]. However, this study was run in a simulation box containing just 512 SPC water molecules in a cubic cell of 26.0 $\AA$ in length (compare our 4040 molecules). As the length of β-CD is 15.7 $\AA$ (table 5.1), one would expect that spurious edge effects would be quite dominant in this simulation. To our knowledge, there have been no simulations of γ-cyclodextrin in water.

1.5 Objectives

The objective of this thesis is to develop and demonstrate the utility of various computational tools for probing the structure of carbohydrates in condensed phases. We particularly intend to explore the structures and dynamics of various $\alpha(1 \rightarrow 4)$-linked carbohydrates in solution. This analysis may give insights into the molecular basis behind the macroscopic properties of polysaccharides comprising $\alpha(1 \rightarrow 4)$-linked glucose monomers.

In Chapter 2 we discuss in general the molecular dynamics simulation methods we

used for carbohydrates, while Chapter 3 gives an overview of the analytical techniques employed to investigate data provided by these simulations.

Molecular dynamics simulations were performed for maltose and a hexa-amylose strand in solution. The results and subsequent analysis of these simulations are given in Chapter 4. Chapter 5 contains comparative results for simulations of the α, β and γ cyclodextrins in solution.

In addition to simulations in the liquid state, we also investigate close contacts in saccharide crystal structures briefly in Chapter 6. Finally, concluding remarks and comments on future work to be done appear in Chapter 7.

Chapter 2

Computer Simulations

In order to completely understand the macroscopic effects of changes in environment on chemical compounds, it is necessary to investigate the environmental effects on the microscopic scale. Unfortunately, if one is concerned with solvated molecules, experimental studies that yield detailed information on the microscopic scale for larger molecules are very difficult, if not often impossible, to perform. Computer simulations which provide detailed data on the microscopic scale can therefore be a very useful tool for investigating the microscopic motions and interactions of molecules and relating this information to their macroscopic behaviour.

There are two main categories of chemical simulations: quantum mechanical calculations (semi-empirical and *ab-initio*) and molecular mechanics calculations. Carbohydrate problems are generally too large and complex to be investigated by the more sophisticated quantum mechanical methods, especially if one explicitly includes solvent molecules. However, calculations based on the simpler physical models in molecular mechanics and dynamics simulations can give important insights into the internal motions of macromolecules and how these motions affect the molecules' properties.

2.1 Dynamics Simulations

Biological macromolecules are inherently dynamic systems - they vibrate, rotate, diffuse and bend - and carbohydrates are no exception. In fact, in many instances, a static picture of molecular conformation is inadequate for explaining the biological function of molecules and it is necessary to take the motion of molecules into account in order to rationalise their chemical properties. For example, oxygen can only reach the binding site in the proteins myoglobin and haemoglobin if a temporary pathway is opened by structural fluctuations within the molecule.

Molecular dynamics, the science of simulating the motions of particles, plays an important role in the investigation of the conformational fluctuations of molecules. A significant advantage of this method is that it provides a microscopically detailed history of the behaviour of individual particles over time of a type which cannot be acquired from experimental methods. This history can be probed to answer specific

questions about the microscopic properties of a molecular system, from which it is possible to infer, or rationalise, the macroscopic properties.

There are many introductory texts on the theory of molecular dynamics and its application to biological systems [54, 55, 56, 57]. The aim of this section is to provide a brief overview of the subject and its relevance to the study of carbohydrates.

The two essential elements for a dynamics simulation are a knowledge of the interaction potential for the molecules involved in the system, from which the forces can be calculated, and the equations of motion governing the dynamics of the system.

2.1.1 Molecular Mechanics Potential Functions

Empirical potential functions specify how the energies of atomic and molecular systems vary as a function of the atomic coordinates. These energy functions determine the relative stabilities of the various possible stable and meta-stable structures of a system.

Most potential energy functions for carbohydrates and other biological molecules treat the molecular potential energy as a sum of valence terms and non-bonded interactions. A typical energy function for a molecular system has the form:

$$E = E_{bonds} + E_{angles} + E_{dihedrals} + E_{improper-dihedrals} + E_{vanDerWaals} + E_{electrostatic} \quad (2.1)$$

Conceptually this force field can be divided into three parts: intermolecular interactions, vibrational modes and torsional flexibility.

The vibrational part of the force field is represented by harmonic potentials for internal coordinates: bond lengths, bond angles and out-of-plane distortions. The bond-stretching and angle-bending terms are often regarded as *hard* degrees of freedom, in that quite substantial energies are required to cause significant deformations from their reference values. At normal temperatures and pressures the deformations are sufficiently small for the harmonic approximations to apply (Hooke's Law). The functional forms of these terms are:

- Bond potential:

$$E_b = \sum_{bonds} K_b(b - b_0)^2 \quad (2.2)$$

  The ideal bond length is represented by b_0. The force constant K_b determines the elasticity of the bond, which varies with the particular type of bond. K_b can be evaluated from infrared stretching frequencies or quantum mechanical calculations, while b_0 can be inferred from crystal structures or microwave spectroscopy data.

- Bond angle potential:

$$E_\theta = \sum_{bond\ angles} K_\theta(\theta - \theta_0)^2 \quad (2.3)$$

  Here K_θ is the force constant for angle-bending and θ_0 is the ideal bond angle.

- Improper torsions:

$$E_\omega = \sum_{improper\ torsions} K_\omega(\omega - \omega_0)^2 \tag{2.4}$$

The improper torsion angles are intended to model linearity about a tetrahedrally extended heavy atom and to maintain planarity about certain planar atoms.

The bond, bond angle and improper torsion terms essentially represent the vibrational frequencies as experimentally observed.

The intermolecular interactions are represented by Coulombic and Lennard-Jones(12,6) pairwise interactions. These interactions are essential for modelling the interaction of solute molecules with their solvent.

- Van der Waals interactions: Van der Waals interactions are calculated using a Lennard-Jones (12-6) potential.

$$E_{vdw} = \sum_{nb\ pairs} (\frac{A}{r^{12}} - \frac{B}{r^6}) \tag{2.5}$$

This has a minimum at an inter-atomic separation equal to the van der Waals radii of the two interacting atoms. Parameters A and B depend on the atoms involved and have been determined by a variety of measurements, including non-bonding distances in crystals and gas-phase scattering measurements.

- Electrostatic potential:

$$E_{el} = \sum_{nb\ pairs} \frac{q_i q_j}{Dr} \tag{2.6}$$

The electrostatic interactions are represented by a Coulomb potential where D is the effective dielectric function used for the medium and r is the distance between the two charges.

The third and last of the contributions to the potential energy function are the dihedral angles, which represent most of the molecular flexibility. They are described by a Fourier term:

- Dihedral angles:

$$E_\phi = \sum |K_\phi| - K_\phi \cos(n\phi) \tag{2.7}$$

where n=1,2,3,4,6

This is a four-atom term based on the dihedral angle about an axis defined by the middle pair of atoms.

The potential energy function discussed above is the form of the CHARMM [58] potential energy function used in our simulations. However, there are a large number of variants for the expressions for the different parts of the empirical energy function in use. For instance, although hydrogen bonding can be adequately modelled with the van der Waals interactions and electrostatic potential terms above [56], some force fields include special hydrogen bonding-terms to ensure that this interaction is correctly modelled.

2.1.2 Force Fields

A molecular mechanics force field comprises the set of potential functions, the idealised bond and angle geometries (r_0, θ_0, ω_0 etc.) and any constants required (k_b, $A_{i,j}$, etc.). Force fields must be parametrized for every class of system and carefully parametrized force fields are necessary for reproducing the properties of molecules accurately. The various constants are determined either experimentally, by fitting calculated properties of small molecules to measured values, or obtained from high-level *ab initio* calculations. Geometric constraints are usually derived from crystallographic data. As potential functions are empirical, the individual parameters for one force field are not generally transferable to other force fields.

Carbohydrate Force Fields

Most of the force fields currently in use were originally developed for proteins, but later extended to treat various other kinds of molecules. Force fields specifically parametrized for carbohydrates include: CHARMM [58, 47], AMBER [59, 60, 48, 49] and GROMOS [61, 51]. For our simulations, we used a CHARMM-like force field parametrized specifically for carbohydrates [62]. This force field is a further development of one used in several other studies of carbohydrates [34, 33, 37, 36] designed by Brady et. al [47].

Unfortunately, however, we have found that there are still some deficiencies in the new force field, specifically in the parameterization barriers to rotation about the hydroxyl torsion angles. This is particularly true in the parameterization of the primary alcohol torsion angle O_5-C_5-C_6-O_6 (ω, definition in section 1.4.1). The previous version of the force field [47] favoured the *tg* conformer excessively, and was thought to be corrected in the new version of the force field [62]. However, we found that in all our simulations that the barrier to rotation between the *gg* and *gt* rotameric forms was parametrized to be too high: transitions between these conformations were never observed.

The experimental distribution of the rotamers in simple sugars and their derivatives has been the subject of some disagreement. From NMR measurements of proton chemical shifts and coupling constants, it has been reported that for glucose in D_2O solution the *tg* conformer is not present to any great extent [26], with the *gg*:*gt* ratio for the β anomer of D-glucose being in the range 7:3 to 5:5. In the crystal structure of maltose [21], the primary alcohol of the non-reducing residue is in the *gt* conformation, while that of the reducing residue is in the *gg* conformation. However, it has been

reported from optical rotation calculations that the *gt* conformer is dominant in aqueous solution [27]. Quantum mechanical calculations on the conformations of glucose in aqueous solution [44] have also predicted that the *tg* conformer is much less stable than the other two. Furthermore, the *gt* conformer is predicted to be more stable in water that the *gg* conformer, due the more favourable solvation around the hydroxymethyl group when in the *gt* position.

In our simulations the hydroxymethyl groups were primarily in the *gt* conformation, with some excursions into the *tg* conformation. The *gg* conformation was never observed.

To date, various force fields for carbohydrates have been able to produce results that are in agreement with the available data for the the carbohydrate ring structure, but fail to agree on which orientation of the exocyclic hydroxymethyl group is favoured [32, 45, 34, 50, 52]. It can therefore be expected that the primary alcohol configuration has little effect on the ring structure. Differences have been attributed to insufficient equilibration and sampling, the different water models used and differences in the truncation method employed. However, inaccuracy in the relative energies of the primary alcohol positions does give rise to some uncertainty in the glycosidic angle conformations, to the extent that the position of this functional group affects the energy of the remainder of the molecule.

Water model

Water remains one of the most interesting and difficult liquids to study. Water is an associated liquid, wherein linear hydrogen-bonding produces a local tetrahedral structure that is distinct from what would be predicted if only the size and shape of the water molecule were taken into account.

Computational solution studies on the atomic level can take two forms: one where the solvent is included explicitly as separate solvent molecules that interact with the solute, and another where the solvent is implicitly included in parameterization of the terms of the force field. Simulations where solvent is implicitly included are attractive as they are computationally far cheaper to perform. *CHEAT95* is an example of one such force field in use for carbohydrates [63]. However, because of the structural role that specific hydrogen-bonds play in water-polysaccharide solutions, implicit models may be inadequate for modelling aqueous states of carbohydrates.

Water models usually take the form of effective pair potentials. Effective pair potentials incorporate the average many-body interactions in a system into interactions between pairs of atoms. Many effective pair potentials have been used for water, such as the ST2 model, the TIPS (transferable intermolecular potential functions) [64] model with its variants [65] and the SPC (simple point charge) model. We use an improvement on the SPC potential – SPC/E [66]. This potential was used in our simulations as it exhibits many of the structural features experimentally determined via neutron diffraction for bulk water. It is considered to be one of the most accurate water models, as it features localised tetrahedral structuring similar to ice structures. However, differences between the TIP3P and SPC/E water model in monosaccharide

simulations have been shown to be insignificant [35].

2.1.3 Dynamic Simulation Methods

The two chief simulation techniques for dynamics are Molecular Dynamics, where Newton's equations of motion are integrated over time, and Stochastic Dynamics, in which the Langevin equation of motion for Brownian motion is integrated over time. However, "molecular dynamics", or "MD", is often used as a blanket term to refer to either of these two methods.

Molecular Dynamics simulations

In the Molecular Dynamics approach, Newton's equations of motion, (2.8) and (2.9), are integrated simultaneously for all the atoms in the system.

$$\frac{d^2 r_i(t)}{dt^2} = m_i^{-1} F_i \tag{2.8}$$

$$F_i = \frac{-\partial V(r_i, ...r_N)}{\partial r_i} \tag{2.9}$$

The force on atom i is denoted by F_i and t denotes the time. The gradient of the potential energy $U(r)$ are the forces, and therefore $U(r)$ must be a differentiable function of the atomic coordinates r_i. The integration of equation (2.9) is performed in small time steps, typically 1 to 10 fs.

Stochastic Dynamics

Stochastic dynamics is an extension of molecular dynamics [67]. It involves integrating the stochastic Langevin equation of motion (2.10).

$$\frac{d^2 r_i(t)}{dt^2} = m_i^{-1} F_i + M_i^{-1} R_i - \gamma_i \frac{dr_i(t)}{dt} \tag{2.10}$$

This equation has added two terms to equation (2.9): a stochastic force, R_i, that models random collisions with solvent molecules and a frictional drag force (proportional to γ_i) that dampens the solute's motion through the solvent. The stochastic term introduces energy and the frictional term removes energy from the system.

Stochastic dynamics can be used to mimic the effect of solvent, without having to include all of the solvent particles explicitly or to establish a coupling of the individual atom motion to a heat bath. We used the later method, as this enables the energy of the system to be maintained without the usual velocity rescaling.

Periodic Boundary Conditions

Periodic boundary conditions enable macroscopic properties to be calculated from simulations using relatively small numbers of particles. This problem with dynamics simulations is the large number of molecules that lie on the surface: these molecules experience forces very different to the forces in a bulk fluid. This problem can be overcome by implementing periodic boundary conditions. In this way a simulation can be performed using a relatively small number of particles in such a way that the particles experience forces as if they were in bulk fluid.

The atoms of the system are put into a box of a shape that will be space-filling. The cube and the truncated octahedron are the most widely-used shapes. If a cube is used, the box is treated as if it was surrounded by 26 $(= 3^3 - 1^3)$ identical images of itself. When a particle moves in the central box, each of it's images will move in the same way. When a particle leaves the box, it enters it with identical velocity at the opposite side in the translated image position (figure 2.1.3).

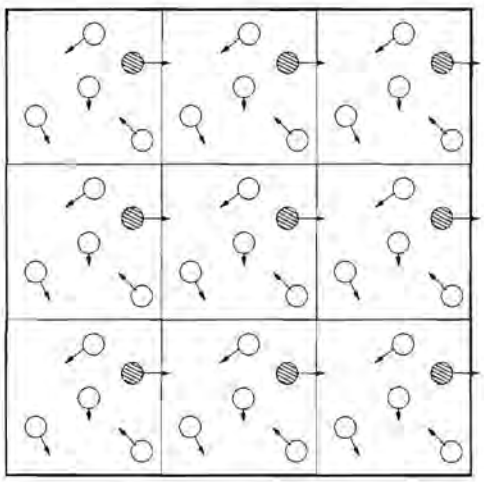

Figure 2.1: Two-dimensional illustration of periodic boundary conditions.

Truncating the long-range forces and the Minimum Image Convention

For large systems, the non-bonded interactions represent the most time-consuming part of the simulation. For a pairwise interaction model, such as the CHARMM empirical energy function, the number of non-bonded interaction terms increases with the square of the number of atoms and thus the computation time is proportional to N^2, rather than N. In order to speed up the calculations, it is necessary to truncate the potential at some distance (usually around 7.5 $\mathring{A}$).

The usual method of treating the non-bonded interactions is to use a non-bonded cutoff and to apply the minimum image convention. The minimum image convention is a natural consequence of periodic boundary conditions. It requires that the interactions for a molecule i in the simulation are only calculated between that molecule and the

closest image of every other molecule in the system. The cut-off radius usually has a value between 6 $\mathring{A}$ and 9 $\mathring{A}$.

In order that an atom does not simultaneously interact with another atom and that atoms periodic image, it is necessary that the length of the periodic box should exceed twice the cutoff radius of the non-bonded interactions.

The truncation of the intermolecular potential at the cut-off distance creates some problems in defining a consistent potential and force for use in the MD simulation: the cut-off introduces a discontinuity in the potential function at the cut-off distance, which artificially increases the kinetic energy and thus the temperature of the system. This can be avoided by using smoothing functions. Two of the most widely used methods are "shifting" (equation 2.11) and "switching" (example in equation 2.12) functions.

$$\begin{aligned} S_f(r_{ij}) = &\quad (1 - \tfrac{2r_{ij}^2}{r_{cut}^2} + \tfrac{r_{ij}^4}{r_{cut}^4}) && r_{ij} < r_{cut} \\ = &\quad 0 && r_{ij} > r_{cut} \end{aligned} \tag{2.11}$$

$$\begin{aligned} S_w(r_{ij}) = &\quad 1 && r_{ij} \leq r_{on} \\ = &\quad \tfrac{(r_{off}^2 - r_{ij}^2)(r_{off}^2 + 2r_{ij}^2 - 3r_{on}^2)}{(r_{off}^2 - r_{on}^2)^3} && r_{on} < r_{ij} \leq r_{off} \\ = &\quad 0 && r_{ij} > r_{off} \end{aligned} \tag{2.12}$$

Switching functions are polynomials in the separation distance that smoothly "turn off" the interaction potential over a given range between a "cut-on" distance and a "cut-off" distance. The interaction potential is unaffected for distances less than the cuton distance. In contrast, a shifting function modifies the interaction potential over its entire range, going to zero at the cutoff distance.

Smoothing and truncation procedures can be based on an atom-by-atom basis or can be applied to predefined neutral groups of molecules. Switching functions, applied on a group-by-group basis have been shown to be the superior method [68].

2.1.4 Dynamics Protocol

All molecular dynamics simulations we carried out using the CHARMM program and a CHARMM-like force field specifically parametrized for carbohydrates. For each simulation, the saccharide in question was super-imposed on the center of a previously equilibrated box of water molecules. Those water molecules with oxygen atoms that overlapped with any of the solute heavy atoms were removed. Boxes of two different sizes were needed for the simulations, depending on the size of the saccharides (table 2.1.4). In each case, the box length was then adjusted to correct the solution density to 1.01 g/cm^3.

The solvated systems were subjected to simple cubic, minimum-image periodic boundary conditions to remove edge effects. Long-range interactions were adjusted

Box size ($\AA$)	No. water molecules
24.6433	512
49.7232	4096

Table 2.1: The two different boxes sizes used in the simultions

smoothly to zero using a switching function applied on a group-by-group basis between 12.0 and 14.0 $\AA$ [68]. The groups correspond to whole water molecules and electrically neutral collections of adjacent atoms in the solute. The SHAKE algorithm [69] was applied to constrain the bond lengths between heavy atoms and hydrogens. We used a canonical ensemble (NVT) at 300K. Simulations were performed using Langevin dynamics with heat bath (T=300K, friction coefficient on non-hydrogen atoms = 62.5) and a 1 fs integration time step.

Chapter 3

Analytical Methods

Molecular dynamics simulations typically produce large quantities of data, particularly if the simulation time is long (on the order of nanoseconds or hundreds of picoseconds) or the system is large. Once a simulation has been performed, analysis is done to make sense of the data produced. We will concentrate in this section on analytical techniques for investigating structural and thermodynamic properties.

Typically, molecular dynamics data are produced at regular time intervals during a simulation in the form of a configuration file. Each stored configuration generally contains the vectors describing the positions of the atoms, the velocities and forces for each molecules, as well as the instantaneous values of all calculated properties. It would be inappropriate to store every configuration from a simulation, as neighbouring configurations are highly correlated, so usually the configuration at the end of every 5th or 10th time step is stored.

In the following two sections we discuss the analysis of carbohydrates in condensed phases, first the liquid phase (with which we are primarily concerned) and second the solid phase. The final two sections discuss the more generally applicable techniques of enthalpy analysis and cluster analysis.

3.1 Analysis in the Solvated State

Water forms extensive hydrogen-bonded networks in the solid state and still exhibits considerable local structure when a liquid. Specific information on the water-structuring abilities of certain saccharides would thus be useful in explaining their solution properties. However, though it is sometimes possible to characterise solvent structuring from diffraction studies of hydrated crystals of carbohydrates, it is much more difficult to obtain the details of solvent structure for dilute liquid solutions, particularly for non-spherically symmetrical solutes. Such information is required for a complete understanding of the origins of the phenomena associated with the interaction of water with saccharides. This problem is often more easily approached computationally, using molecular dynamics simulations. Below we discuss some computational analytical methods for probing the characteristics of liquid systems as simulated by molecular dynamics.

3.1.1 The Pair Distribution Function

The structure of molecular systems can be investigated by a set of distributions functions for the atomic positions, the simplest of which is the pair distribution function (often called the pair correlation function or radial distribution function). This function, $g(r)$, gives the probability of finding a pair of atoms a distance r apart, relative to the probability expected for a completely random distribution at the same density.

The pair distribution function, $g(r)$, for water molecules around a sugar solute molecule is defined [70] in equation 3.1:

$$g(r) = \frac{1}{4\pi\rho r^2}\frac{dN(r)}{dr} \tag{3.1}$$

In this equation, r is the inter-atomic distance from selected atoms in the sugar molecule, ρ is the bulk water number density and $N(r)$ is the number of water molecules within a sphere of radius r around the solute. With the factor $(4\pi\rho r^2)^{-1}$, $g(r)$ is normalised to one at positions in the bulk water far from the solute.

The pair distribution function can be experimentally determined, by evaluating the diffraction of X-rays, neutrons or electrons, or computed from the system trajectories of a molecular dynamics simulation.

Gases, liquids and solids have very different pair distribution functions. In a crystal, the pair distribution function has an infinite number of sharp peaks whose separation and heights are characteristic of the crystal structure. In contrast, the pair distribution function of a liquid has a small number of sharp peaks at short distances and then rapidly decays (figure 3.1).

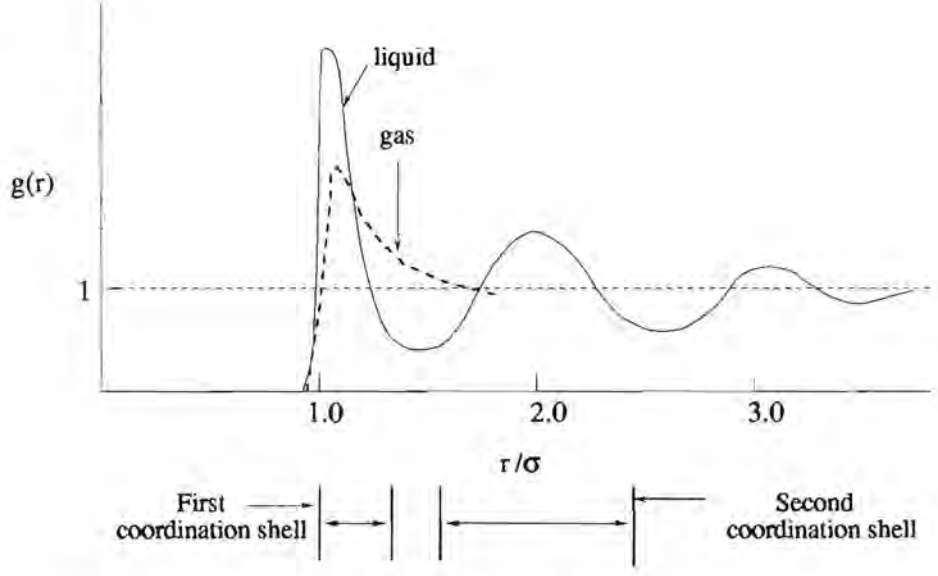

Figure 3.1: A typical pair distribution function for a simple fluid.

Pair distribution functions can give valuable insight into the structuring of liquids. However, they have disadvantages in that they are radially averaged, and thus are

unsuitable for investigating anisotropic liquid structuring. Therefore, other techniques must be employed in order to investigate non-spherically symmetrical systems.

3.1.2 Water Probability Density Calculations

Non-radially-averaged analysis methods are needed to elucidate the water structuring about carbohydrates, as radially averaged methods obscure the anisotropic detail. Analytical methods to map the average three-dimensional, anisotropic water structuring about carbohydrates have been developed in some extensive computational investigations [33, 37, 36, 35]. We use a method similar to that employed by Lui et al. for trehalose [36], with some adaptations to make the method more suitable for studying flexible carbohydrates. This method produces a 3D water probability density matrix - a picture of the non-uniform distribution of water molecules - which may be contoured using standard three dimensional graphing packages. Hereafter we refer to the probability density as simply the density.

The Method

To generate a map of the average anisotropic structure of water about a solute molecule, it is first necessary to remove the rotational and translational diffusion of the solute molecule through the simulation. In order to do this, the instantaneous position and orientation of the solute molecule in each stored coordinate set is translated and rotated to achieve the best least-squares overlap with a reference frame. The coordinate transformation is applied to all the atoms in the system. If two or more separate dynamics runs are to be compared, the same original frame is used in order that the resultant density maps will have the same relative orientations. Particular care must be taken of the atoms chosen for the least-squares fit. Choosing atoms that are likely to move around too much relative to their neighbours could result in an unsatisfactory overlay of solute molecules and a concomitant blurring of the water densities.

Previous studies using this method (cited above) were done on relatively inflexible carbohydrate molecules, such as monosaccharides. As there was little internal solute motion in these cases, all the frames from the dynamics simulations could be used to generate the map of water density. However, the carbohydrates we investigated are all polysaccharides, linked via flexible $\alpha(1 \rightarrow 4)$ glycosidic linkages. Using all the dynamics frames together, regardless of the ϕ, ψ conformations, results in an unacceptable blurring of the very anisotropic water structuring one wishes to investigate. Therefore, we considered it necessary to use some *selection criteria* for choosing frames with solute conformations similar enough to provide a detailed picture of the water structure. The obvious choice are the ϕ, ψ angles defining the glycosidic linkage orientations, as these are the source of most of the internal molecular motions and thus have the greatest effect on the overall molecular conformations. Then, when generating a map of the average water densities around a polysaccharide molecule, angles and ranges about these angles are decided upon for all the ϕ, ψ dihedrals and only frames falling within these criteria are used to calculate the water density map. (When choosing a range of ϕ, ψ

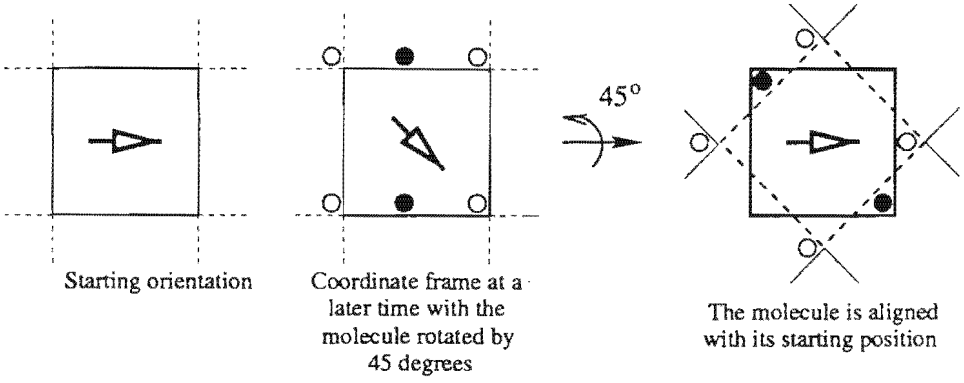

Figure 3.2: Illustration of the coordinate transformation of the frames from the molecular dynamics simulation.* The solute molecule is represented by an arrow and specific water molecules by circles.
* adapted from [36]

angles, care must be taken in the choice of the range. If too much molecular motion is allowed, the water densities will be blurred.). This approach has the advantage that the water-structuring ability of different conformations of a polysaccharide molecule can be compared. This may help in explaining conformational preferences of polysaccharides in water: conformations that impose greater structure on the surrounding solvent may be relatively entropically disfavoured. We found that rather long simulations are required to generate sufficient frames for a satisfactory water density map around a particular conformation, thus sufficient data is only available for conformations that form a significant part of the trajectory.

It is important that equal number of frames are used when one desires to compare the water densities around different conformations, as the water densities calulated are very sensitive to the number of frames used in the calulation. Using more frames essentially allows for more motion and the possibility that the densities will be blurred.

Only the oxygen water atoms were used to calculate the electron densities. In order to do this, a Gaussian distribution function centered on each water oxygen atom was used to approximate the distribution of electrons for each oxygen atom. The Gaussian function used is of the form:

$$G(r) = elec \times (\frac{a}{\pi})^{1.5} \times e^{-a \times r^2} \tag{3.2}$$

where *elec* is the total number electrons for the atom and a is calculated so that the function drops to 10% of its maximum value at the atom's van der Waals radius (1.4 $\mathring{A}$ for O).

The simulation boxes are divided into bins approximately 0.5$\mathring{A}$ wide and the densities in each box summed for all the selected frames from the dynamics run The final water density matrix was normalized using equation 3.3 so that the density of bulk water corresponds to a value of 1, and 50 per cent above bulk density corresponds to a value of 1.5 etc.

$$dens_{norm}(i,j,k) = dens(i,j,k) \times \frac{n_{xbin} \times n_{ybin} \times n_{zbin}}{n_{electrons} \times n_{atom} \times n_{frames}} \tag{3.3}$$

In this equation, n_{xbin}, n_{ybin} and n_{zbin} are the number of divisions in the x,y and z directions in the simulation box respectively. These are usually adjusted so that the box size is $0.5\mathring{A}$. The number of atoms whose density is being calculated is represented by n_{atom} (in our case the water oxygens) and $n_{electrons}$ is the number of electrons for that atom (i.e. 8 for oxygen, 5 for carbon etc.). This method of calculating water densities is essentially the same as a three-dimensional pair correlation function without radial averaging.

The resultant electron densities are contoured in three dimensions to produce diagrams of the water structuring about the solute molecule.

3.1.3 Vornoi Analysis

Vornoi analysis is a purely geometric approach to the problem of characterizing the structure of disordered systems such as liquids. Vornoi analysis has been used in fields as diverse as fluid mechanics and astrophysics, but also increasingly for the microstructural description of disordered systems.

The Vornoi polyhedron associated with a given point i, in an assembly of N points, is defined as that volume of space containing all points closer to i than any other point j. Thus, the bonding surfaces are planes drawn perpendicular to the intercenter ij at their midpoints. The intersection of these planes forms the polyhedron edges and vertices.

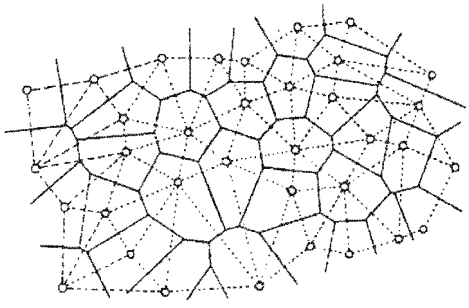

Figure 3.3: Example of a Vornoi polyhedra in a disordered two-dimensional system*.
*·adapted from [71]

Vornoi analysis provides the number of faces and vertices for each point. As we use Vornoi analysis to investigate the structure of water molecules around a saccharide solute, the "points" correspond to water oxygen atoms and the faces give a measure of the number of nearest neighbours around each atom. The average number of faces, calulated over a simulation, gives an indication of how structured a liquid system is: a higher number of faces for atoms in a area indicates a higher density. An ice-like

structure would have 16 faces, while for thermalized water the number of faces ranges between 24 and 9 [72].

3.2 Analysis in the Solid State

Comparisons are an essential element of analysis, as one often wishes to relate differences in structure between chemical compounds to differences in their physical properties. This is true not only for molecules in solution, but also for the solid phase. Differences in crystal structure can lead to widely different physical properties, even if the constituent molecules of the crystal are exactly the same. There are many examples of substances that can exist in more than one crystalline form (a phenomenon known as *polymorphism*), carbon being the obvious instance. Also, packing forces within a crystal may lead to crystals where the constituent molecules are in a different conformation to that which they would have in solution or the liquid phase. Subtle differences in the packing of two isomer guests within a host molecule may lead to one complex being more stable than the other.

X-ray-diffraction studies of crystals generally provide the specific positions of most, if not all, of the atoms in the unit cell. However, crystal structures often contain so much geometrical data that the essential or important differences between similar structures are obscured. Therefore, analysis methods that simplify the data and highlight particular differences in packing and orientation can be of use in justifying the relative stabilities of polymorphic crystals, or the enatioselectivity of host molecules such as cyclodextrins (section 1.3.3).

3.2.1 Non-bonded Interaction Pattern Matrix Analysis

A recently developed graphical method for probing crystal structures is non-bonded interaction pattern matrix analysis (NIPMAT) [73]. This analysis involves the use of a shaded distance matrix that represents the nearest neighbour, non-bonded, contacts between each atom i and every other atom j in a crystal. Thus, the matrix element ij represents the *shortest intermolecular* contact $i \ldots j$ in the crystal. The distances are normalised to compensate for the differences in atomic sizes by subtracting the sum of the van der Waals radii of the two atoms involved from the total distance between them. Each matrix element is shown in terms of a grey scale: the shorter the contact, the darker the square which represents that particular contact.

As previously developed, *NIPMAT* analysis displays a symmetric map of all the atoms in each row and column - i.e. the matrix is a visual representation of all the intermolecular interactions simultaneously. The plot thus gives a complete quantitative representation of the packing environment within a crystal. Desiraju recently used this method to compare the packing differences in the crystals of naphthalene and terephthalic acid [73]. The matrices produced are shown in figure 3.4.

NIPMAT analysis can be very useful in showing clear differences between the packing of very similar structures. A plot showing a wide range of grey areas indicates

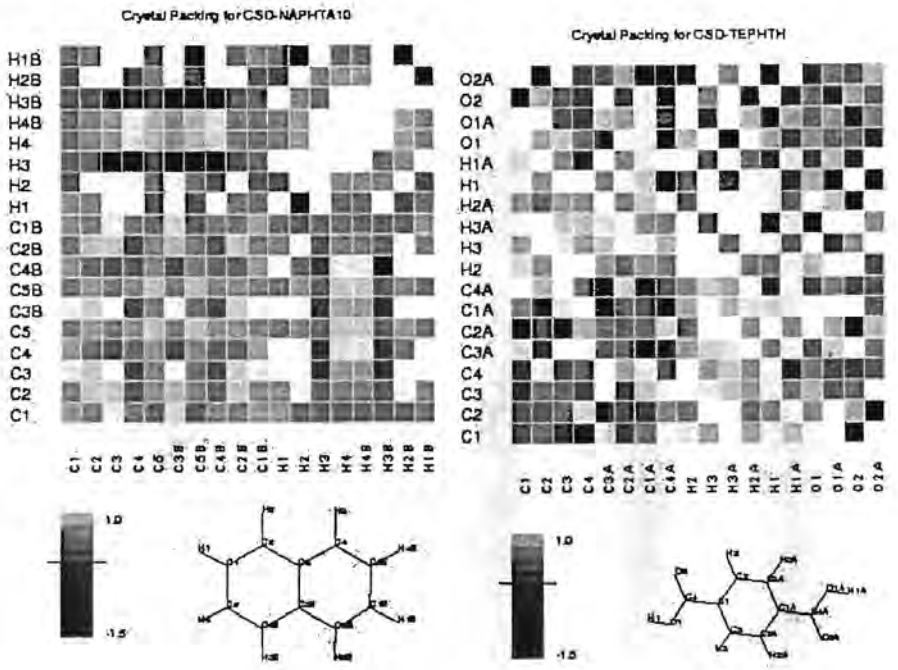

Figure 3.4: NIPMAT plots of napthalene and terephthalic acid *.
* from [73]

isotropic packing, whereas a more black-and-white plot indicates anisotropicity and directional forces in the packing. Thus, the map gives a general impression of overall types of non-bonded interactions that predominate in a crystal structure.

However, viewing all the interactions simultaneously (i.e. via a symmetrical distance matrix) is obviously only practical for crystal structures involving small numbers of atoms, as an overly large map would obscure necessary information rather than highlighting it. Therefore, we took different approach to this analysis. Our method allows not only symmetrical comparison maps, but non-symmetrical maps as well. Non-symmetrical comparison matrices have different atoms listed on the row and column of the matrix. They allow the packing within specific areas of a crystal structure to be investigated and thus have a number of potential uses. For instance, they can be used to compare host-guest interactions for two similar guest within the same host and thus highlight the packing similarities and differences between two inclusion compounds.

3.3 Enthalpy Analysis

Molecular structures are actually determined not by the relative potential energies of the different conformations of the molecule , but rather by the free energy differences. These free energies contain enthalpic as well as entropic terms. However, an analysis of the enthalpic energy surface of a system will give an indication of which conformations are enthalpically preferred.

The energy of a solute molecule in solution as simulated by molecular dynamics is a sum of the internal energy of the solute (bond and angle deformations and dihedral angle terms etc.) and the energy of its non-bonded interactions with the solvent. Though the internal energy of a solute is easy to calculate using CHARMM [58], there is no simple method for separating the solute-solution interaction energy from the rest of the non-bonded interactions. Therefore, this was calculated in a round-about manner by subtracting the non-bonded interaction energy for the water molecules and the solvent molecules from the total non-bonded interactions energy, yielding the solute-solvent non-bonded interaction energy.

Maps of the enthalpy as a function of the ϕ, ψ angles for maltose were generated in the following manner: The total energy for the solute (internal energy plus solute-solvent interaction energy) was calculated for all the frames in the simulation trajectory, along with the ϕ, ψ angles from each frame. The map was prepared by averaging the ten lowest energy values for each ϕ, ψ pair, using a grid with increment of 5°. This grid was then contoured at intervals of 2 kcal/mol to produce the enthalpy map.

3.4 Cluster Analysis

Many of the molecular conformations generated by a dynamics run are similar to each other. It is useful to have some measure of how "different" molecular conformations are, and some means of classing conformations into groups of similar types. Clustering analysis is one approach to this problem.

A cluster analysis requires a measure of "similarity" between pairs of objects. This measure can be a distance, an angle or any combination of geometric measures.

There are a large number of algorithms for cluster analysis. We used the method provided by the CHARMM [58] package, $ART-2'$. This algorithm clusters time series data. The assignment of conformations to a cluster is optimized subject a constraint on the cluster radius, such that no member of a cluster is more than a specified distance from the cluster center. This optimization is carried out as an iterative minimization procedure that minimizes the Euclidean distance between the cluster center and its members.

Chapter 4

Molecular Dynamics Simulations of Maltose and a Hexa-Amylose Strand in Aqueous Solution

The maltose molecule plays an important role as a model for larger polysaccharides containing $\alpha(1 \rightarrow 4)$ glycosidic linkages. In particular, it has been previously argued that oscillatory motions within the maltose disaccharide unit constitute the local internal motions of amylose [24]. Therefore, knowledge of the relationship between water and the conformations of maltose is pivotal to understanding the molecular basis behind macroscopic properties of amylose.

In this chapter, we explore the relationship between water molecules and the conformation of a single maltose molecule, the repeat unit for amylose. Our approach was first to generate data on the relation between water and the $\alpha(1 \rightarrow 4)$ glycosidic linkage via molecular dynamics simulations of maltose in aqueous solution. For comparison purposes, we also ran simulations for maltose in vacuum. General analyses were performed on all the simulations. We then employed the method discussed in section 3.1.2 to investigate time-averaged water densities around maltose at selected points in ϕ-ψ space. The results for maltose are then compared to those for a short amylose strand (hexa-glucose), calculated from a 0.5 ns molecular dynamics simulation in aqueous solution.

4.1 Adiabatic Maps for Maltose

An adiabatic, or relaxed energy, map for maltose as a function of the ϕ, ψ angles gives an indication of the preferred regions of ϕ, ψ space. Generally, these maps show large areas of ϕ, ψ space to be excluded because of steric clashes. Madsen and Brady calculated an adiabatic map for β-maltose [45] using a forcefield for polysaccharides [47]. This map was recently recalculated by Naidoo [74] for the altered carbohydrate forcefield ([62] and section 2.1.2) using a simulated annealing technique [75].

This map (figure 4.2) is similar to the Madsen-Brady map in its general features.

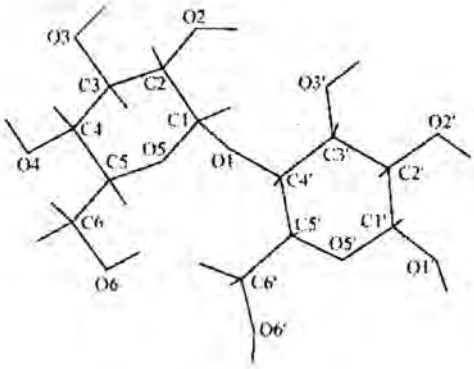

Figure 4.1: Atom numbering scheme for β-maltose

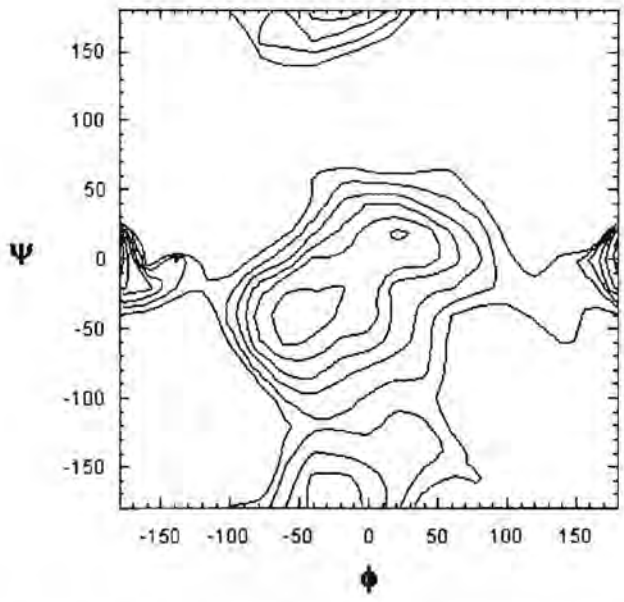

Figure 4.2: The relaxed adiabatic conformational energy map for β-D-maltose in vacuum, calculated on a $20° \times 20°$ grid, contoured at 2 kcal intervals.

There is a broad minimum around $\phi = -40°$, $\psi = -40°$, encompassing two small minima previously present (designated the A and B wells). Also, a small minimum at $\phi = 20°$, $\psi = 20°$ replaces the rather larger C well in the Madsen-Brady study. Adiabatic maps are calculated from vacuum conformations and thus can be assumed to be an approximation of conformational space in vacuum , albeit not at the simulation temperature. This assumption is supported by the qualitative similarity between the free energy surface of maltose at 300 K [76] and the adiabatic energy surface for the molecule in vacuum. However, vacuum adiabatic maps cannot be expected to give a complete description of the conformational space of a molecule in solution.

4.2 Simulation Procedure

The molecular dynamics simulations were performed as discussed in section 2.1.4. In total, four dynamics simulations were performed for maltose, two in vacuum, designated $V1$ and $V2$, and two in water, designated $W1$ and $W2$. The simulations were started from two separate conformations with ϕ,ψ angle values of (0,0) ($W1$ and $V1$) and (20,20) ($W2$ and $V2$) respectively. Our reasons for concentrating on this area of conformational space are twofold. Firstly, experimental studies of maltose in the solid state, for example X-ray [77] and neutron diffraction [78] work, show that ϕ,ψ values for the glycosidic linkage are generally located in this region. NMR [79] and optical rotation experiments [30] performed in aqueous solution predict values around ϕ = -70, ψ = -35 for maltose. Secondly, inspection of the adiabatic map reveals that the energy barriers separating the central region from the rest of conformational space are very high due to a number of steric clashes. These barriers are difficult to traverse and it is likely that traversal will not be possible without the component glucose monomers "ring-flipping" from the experimentally vastly preferred $^4C_1$ chair conformation to a boat conformation. The (0,0) conformation was chosen as corresponding to the saddle point between the two minima on the adiabatic map, and the position (20,20) is a small valley on the adiabatic map.

The $W1$ simulation was 1.8 ns in length (20 ps equilibration and 1780 ps data collection). The $W2$ simulation was first run for 15 ps, constraining the dihedral angles to the (20,20) conformation. Subsequently a simulation of 1.8 ns in length was performed (20 ps equilibration and 1780 ps data collection). Both the vacuum simulations were run for 1.8 ns.

In addition to these simulations of maltose, two dynamics simulations, each 500ps (20ps equilibration and 480ps data collection) in length, were performed on the amylose strand in vacuum and in water respectively. The amylose strand was started from a left-handed helical conformation, as X-ray diffraction experiments suggest that amylose adopts a helical structure with a left-handed sense [4, 5, 3]. All ϕ angles in the strand were set to 30° and all ψ angles set to 25°. These angles were chosen as they were found to be popular angles for the $W1$ and $W2$ maltose simulations, and they represent conformations within the large central valley on the adiabatic map.

A box of 488 SPC/E [66] water molecules was used to solvate the maltose molecule,

while the amylose strand required a larger box (4040 SPC/E water molecules). In both cases, the box length was adjusted in order to get a solution density of $1.013 g/cm^3$, the approximate experimental density of an aqueous maltose solution of this percentage weight (24.8616 $\AA$ for maltose and 49.6105 $\AA$ for the amylose strand). The solute molecule was placed in the center of the box. The terminal glucose monomer was modelled in the β from as this is known to be the form preferred for glucose in solution [6]. For the same reason, the primary alcohols were placed in the *gt* conformation (see section 1.4.1).

In all the simulations, configurations of the molecules were stored at intervals of 0.5 ps. This relatively large interval ensures that subsequent frames are sufficiently independent of each other and reduces the total amount of data stored for each simulation. The trajectories were subsequently analysed, excluding the first 20 ps of equilibration time.

4.3 Dynamics of the $\alpha(1 \rightarrow 4)$ linkage

The 1.8 ns ϕ,ψ time series extracted from the maltose simulations, $V1$, $V2$, $W1$ and $W2$, are shown super-imposed upon the adiabatic map in figure 4.3.

These trajectories for maltose explore the large central minimum well on the adiabatic map. However, it is apparent that the simulations in vacuum, $V1$ and $V2$, explore a smaller area of conformational space compared to their counterparts in water, $W1$ and $W2$. This is despite the damping affects of water on the motion of the maltose solute. The simulations in solution not only cover a broader region of conformational space as a result of the interactions with water, but also show a shift towards lower values of ϕ and ψ. Similar results of a shift in the preferred values of ϕ and ψ in solution as opposed to vacuum have been reported for other simulations of maltose [34, 51]. NMR studies of maltose in aqueous solution have predicted an equilibrium between two favoured conformations for maltose in aqueous solution: conformation A with $\phi = -70°$, $\psi = -40°$ and conformation B with $\phi = -30°$, $\psi = -20°$ [80]. In addition, optical rotation experiments have indicated that maltose in solution distributes between the states $\phi = -30°$, $\psi = -15°$ and $\phi = -70°$, $\psi = -40°$ [27].

When we extrapolate this result for maltose to amylose in an aqueous environment, it is expected that the starch macromolecule will rotate more readily about its component ϕ,ψ torsional angles in water than it does in a vacuum. This will lead to a lowering of the glass transition temperature, T_g, a result which is consistent with known experimental findings [9].

The history of the ϕ-ψ torsion angles for the simulations in vacuum and water are shown in figures 4.4 and 4.5 respectively. The simulations in vacuum do not show any dramatic, long-lasting transitions and, though they were started from different conformations, after 100ps the $V1$ and $V2$ simulations show very similar histories. In contrast, the simulations in water, though sampling the same general ϕ, ψ region, show quite different histories. (The $W1$ simulation can be seen to traverse from the saddle point (0,0) to the vacuum global minimum (-40, -40) and beyond on to (-60,-60).

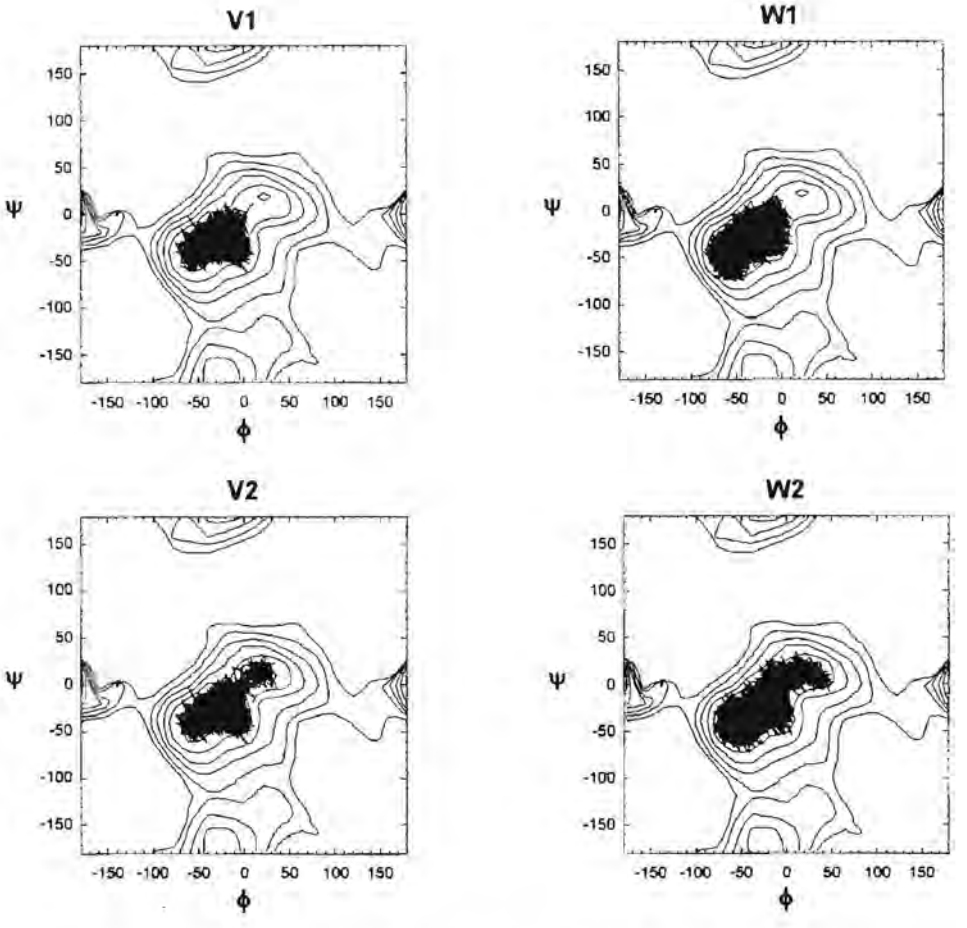

Figure 4.3: Trajectories for the ϕ and ψ angles for all the simulations ($V1$, $V2$, $W1$ and $W2$) projected onto the adiabatic map.

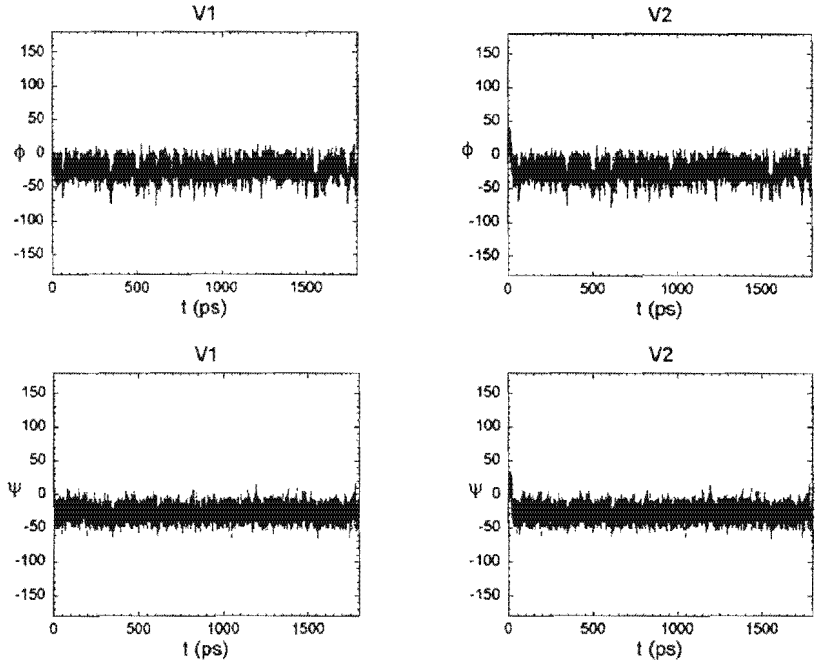

Figure 4.4: Time series for the ϕ and ψ angles of the $V1$ and $V2$ simulations.

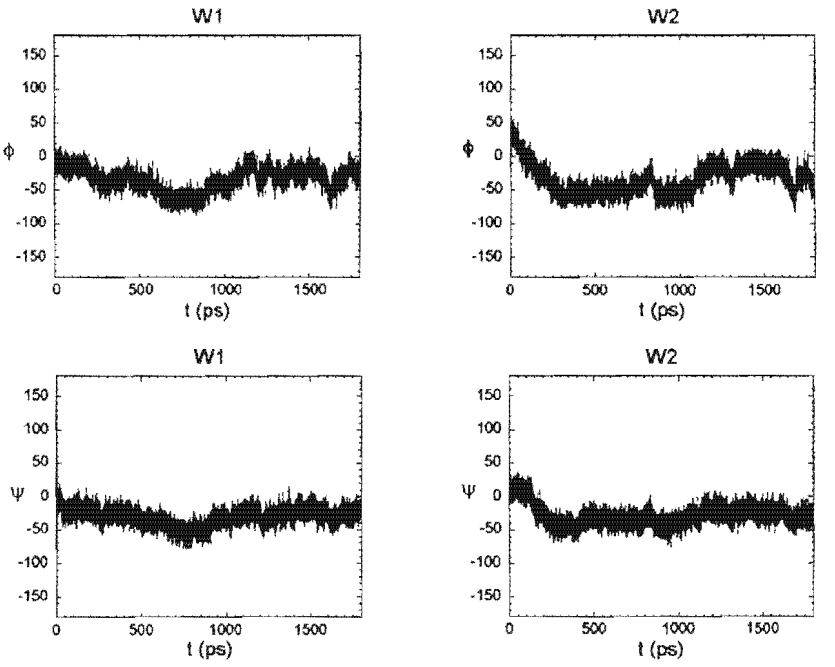

Figure 4.5: Time series for the ϕ and ψ angles of the $W1$ and $W2$ simulations.

After sampling the area round (-60,-60) for approximately 200 ps the molecule returns to the (-40, -40) region. The $W2$ simulation starts from the (20,20) conformation, then within 300 ps moves to the (-60,-60) region where it remains for about 500 ps, subsequently moving to the (-20,-20) region for most of the rest of the simulation.) This discrepancy between the behaviour of the simulations in water and in vacuum illustrates the frictional damping effect of solvent water molecules on a solute. This effects both the overall motions of maltose and its internal conformational changes. Previous investigations have found solvation to have a significant damping effect on the frequency and magnitude of disaccharide conformational fluctuations [34].

The ϕ,ψ time histories differ for $W1$ and $W2$, as these simulations were started from different regions of conformational space. However, the two simulations explore approximately the same ϕ,ψ regions (figure 4.3). However, a complete sampling of ϕ,ψ space was certainly not achieved. Incomplete sampling of conformational space is a common complaint when analysing MD simulations. These simulations have to be judged against knowledge of conformational space for the environment under which the simulations have taken place. In the case of disaccharides in vacuum, the representative nature of a simulation is more easily verified since it can be compared to an adiabatic map such as the one in figure 4.2. The case for disaccharides in water is much more complicated. Here the only measure of sufficient sampling by an MD simulation in a defined ϕ-ψ space is a direct comparison to the free energy surface constructed at that simulation temperature as a function of ϕ,ψ, of the molecule in water. This is a complex calculation which has not as yet been performed for maltose in solution, though it has been recently completed for the related dixylose disaccharide in water [81].

4.4 Hydrogen-Bonding Analysis

Hydrogen-bonding analysis of the maltose molecule serves to highlight the types of hydrogen bonds that are formed during molecular dynamics simulations.

Internal hydrogen bonds for a maltose molecule can be classified into intra-ring hydrogen bonds between hydroxyl groups in the same monomer unit, inter-ring hydrogen bonds between hydroxyls in different monomer units and bonds from a hydroxyl group in either of the adjacent glucose monomers to the glycosidic oxygen. Intra-ring hydrogen bonds form between the hydroxyls on $C2$, $C3$ and $C4$ of the nonreducing ring, and $C2'$, $C3'$ and $C4'$ of the reducing ring. In vacuum, as there are no other potential hydrogen-bonding partners, internal hydrogen bonds satisfy the hydrogen-bonding requirements of the maltose hydroxyl groups. In general, each hydroxyl group orients to make a hydrogen bond to an adjacent hydroxyl on the ring, with each of the hydroxyl groups pointing in the same direction around the sugar rings. Many of the internal hydrogen bonds have significantly distorted bond angles due to the constraints imposed by the relative positions of the hydroxyl groups in the structure of the sugar molecule [21].

Solute-solvent hydrogen bonds between maltose and water may have several differ-

ϕ,ψ	$W1$	$W2$
-60,-60	22.41	19.90
-50,-50	21.81	20.52
-40,-40	22.33	20.13
-30,-30	22.07	19.73
-20,-20	21.18	18.65
-10,-10	20.90	18.39
0,0	20.67	20.59
10, 10	20.64	20.64

Table 4.1: Average number of solute-solvent bonds form for a series of ϕ,ψ values in the $W1$ and $W2$ simulations.

ent arrangements, illustrated in figure 4.6.

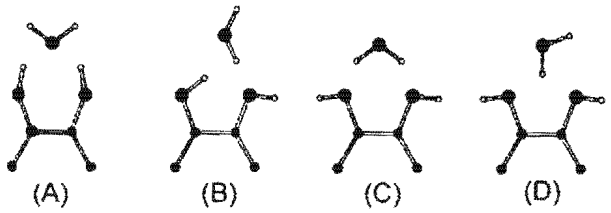

Figure 4.6: Schematic description of four possible hydrogen bond arrangements between solute hydroxyls and a water molecule*: **(A)** two solute donor-water acceptor hydrogen bonds; **(B)** one solute donor-water acceptor and one solute acceptor-water hydrogen bond; **(C)** two solute acceptor-water donor hydrogen bonds (involving two different water hydrogens); **(D)** two solute acceptor-water donor hydrogen bonds (involving only one water hydrogen)
* from [34]

Solute-solvent hydrogen bonds were defined as interactions having donor-oxygen acceptor-oxygen distances less than $3.4\mathring{A}$ and donor-proton-acceptor angles $\geq 120°$. Using these criteria, the average number of solute-solvent hydrogen bonds across all conformations formed for the $W1$ and $W2$ simulations were calculated to be 21.74 and 19.73 respectively. The number of solute-solvent bonds for a series of ϕ,ψ conformations is compared in table 4.1. The discrepancy between the average number of solute-solvent hydrogen bonds in the two simulations can be accounted for by the difficulty with rotations about the hydroxyl dihedral angles in the force field used (section 1.10. Brady and Schimdt have previously performed analysis of hydrogen-bonding in maltose [34]. They reported an average of 22.7 bonds for maltose in a TIP3P [65] water solution [34]. However, none of the simulations analysed were longer than 100 ps.

It is interesting to compare the number of intra- and inter-ring hydrogen bonds formed, not only for the different simulations, but also for selected conformations within a simulation. However, the angle criterion used for solute-solvent hydrogen bonds is too strict for defining the more distorted internal hydrogen bonds. Instead, we used

an angle criterion of $\geq 100°$. The statistics for these bonds in the vacuum and water simulations are displayed in tables 4.2 and 4.3.

ϕ,ψ	inter		intra		gly	
	$V1$	$V2$	$V1$	$V2$	$V1$	$V2$
-60,-60	-	-		-	-	
-50,-50	0.00	0.00	3.73	3.50	0.67	0.47
-40,-40	0.00	0.00	3.69	3.75	0.49	0.28
-30,-30	0.38	0.37	3.72	3.69	0.29	0.20
-20,-20	0.96	0.93	3.62	3.63	0.29	0.24
-10,-10	1.04	1.00	3.65	3.62	0.24	0.50
0,0	1.00	1.00	3.18	3.52	0.45	1.00
10,10	-	1.00	-	4.25	0.91	0.87

Table 4.2: Internal hydrogen-bond statistics for the maltose simulations in vacuum.

ϕ,ψ	inter		intra		gly	
	$W1$	$W2$	$W1$	$W2$	$W1$	$W2$
-60,-60	0.00	0.00	1.56	1.83	0.00	0.10
-50,-50	0.00	0.00	1.73	1.31	0.01	0.12
-40,-40	0.00	0.00	1.28	1.55	0.02	0.06
-30,-30	0.12	0.10	1.16	1.63	0.04	0.08
-20,-20	0.71	0.85	1.14	1.46	0.26	0.34
-10,-10	0.93	0.97	1.65	1.25	0.61	0.54
0,0	1.00	1.06	1.57	0.71	0.71	0.73
10,10	-	1.0	-	1.05	-	0.87

Table 4.3: Internal hydrogen-bond statistics for the maltose simulations in water.

Comparison of the number of internal bonds formed in vacuum and water will give an indication of whether internal hydrogen bonding is preferred over bonds with water molecules.

The crystal conformation of maltose is stabilized by the hydrogen bond between the $O2$ and $O3'$ hydroxyl group [21, 78], but is destabilized by torsion strain within the glycosidic bridge [27]. The presence of water provides alternative hydrogen-bonding partners, and thus the torsional strain can be relieved by moving to lower values of ϕ and ψ. This allows the maltose molecule to explore a wider range of conformational space in solution and has a considerable affect on the solution structure of maltose. We found, however, that the $O2$-$O3$ internal hydrogen bond is favoured over the water-solute competitive hydrogen bonds in regions where these bonds can form (ϕ,ψ values between 30,30 and -30,-30). In favourable bond-forming conformations, there is only a slight reduction in the frequency with which the bond forms when comparing the vacuum and solution simulations. Brady and Schmidt had a similar result [34], using the older carbohydrate forcefield for CHARMM [47] and TIP3P water. In contrast,

a study of maltose in solution by Ott and Meyer [51] using the GROMOS force field reported the opposite finding: the $O2$-$O3$ hydrogen bond, though present in vacuum, was rarely present in solution.

The histories of the $O2$-$O3'$ hydrogen bond distance for the two vacuum simulations, $V1$ and $V2$, and the two water simulations, $W1$ and $W2$, are displayed in figure 4.7. These show the history of formation and breaking of the internal $O2$-$O3'$ bond in

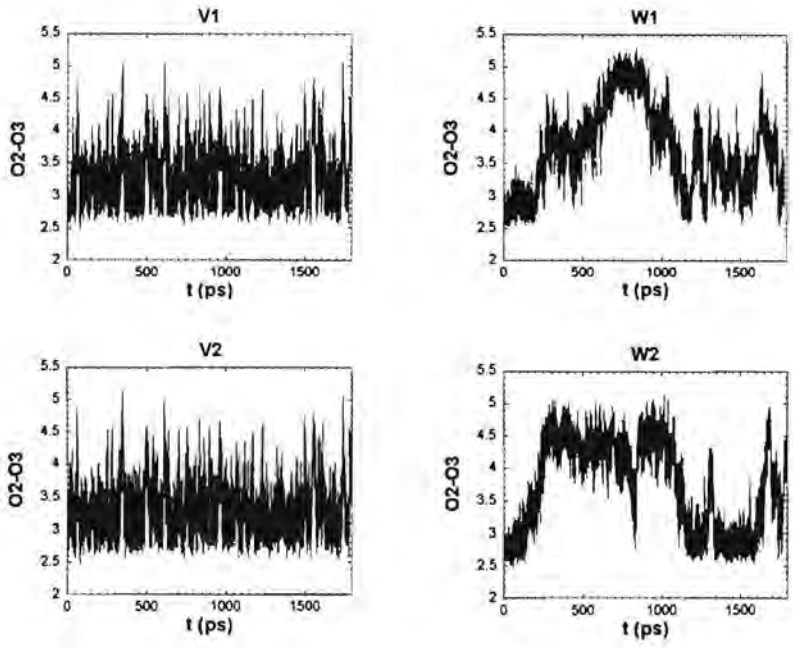

Figure 4.7: Time series for the $O2$-$O3'$ distance in the $V1$, $V2$, $W1$ and $W2$ simulations.

the two simulations. It is clear that in the vacuum simulations, the $O2$-$O3'$ breaks only briefly, rapidly returning to conformations where the bond reforms. In contrast, the water simulations show long periods with conformations that do not exhibit this internal bond.

When comparing the average number of intra-ring bonds for the simulations in solution and vacuum (tables 4.2 and 4.3), it is apparent that the average number of these bonds formed is considerably reduced when moving from vacuum to solution. As these bonds are quite distorted, it is to be expected that they will form less frequently if alternate hydrogen-bonding partners are provided.

4.5 Primary Alcohol Analysis

A matter of some concern was that neither the vacuum nor the water simulation started from the (0,0) conformation explored the small valley on the adiabatic map centered at (20,20). It was expected that a transition between the valley centered at (-30,-30) and the (20,20) valley would be made, as the adiabatic map indicates that the energy

barrier between them is low (about 2 kcal). In fact, after some analysis, we discovered that this seems to be a result of the problem with the primary alcohol parameterization mentioned earlier (section 2.1.2). It seems, from inspection of the minimum-energy conformations used to calculate the adiabatic map, that the (20,20) minimum corresponds to a primary alcohol conformation of *gg* in both residues of the maltose molecule. The barrier to rotation between the *gg* and *tg* conformations has been parametrized to be too high to allow transitions which are consistent with experimental estimations of the *gg*:*gt* ratios (section 1.10). This can be seen in the plot of the $O_5 - C_5 - C_6 - O_6$ dihedral angle (ω) for the simulations in vacuum (figure 4.8) and in water (figure 4.9). All these simulations were started from the *gt* conformation and, though transitions to the *tg* conformation do occur, there are none to the *gg* conformation. This effectively prevents the maltose molecule from moving to the (20,20) valley, as it seems that this valley does not occur if only the *tg* and *gt* conformations are taken into account.

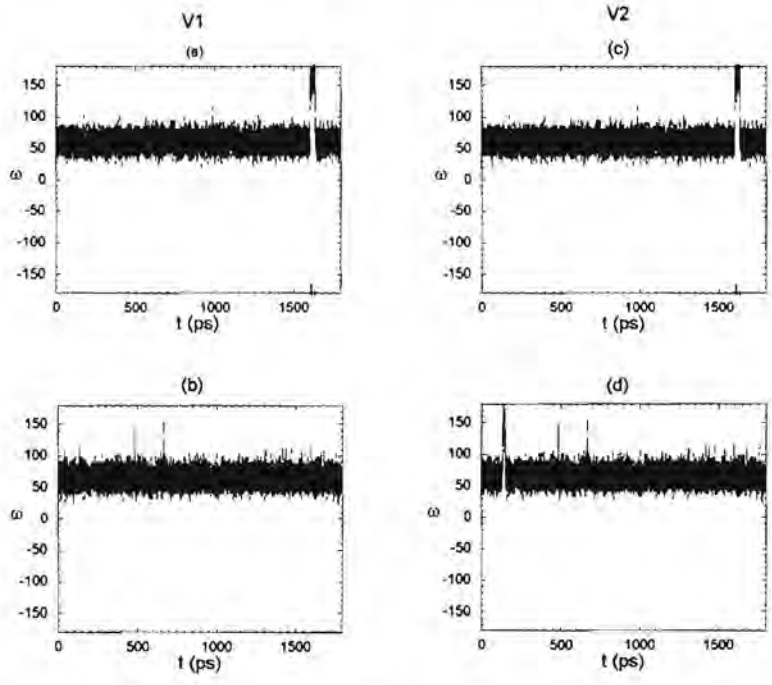

Figure 4.8: Time series of the angle ω ($O5$-$C5$-$C6$-$O6$) for the simulations in vacuum. (a) and (c) : non-reducing residue (b) and (d) : reducing residue

To further investigate this issue, we ran a 1800ps simulation in vacuum, starting from the (0,0) conformation with both the primary alcohols in the *gg* conformation. The plot of the ω dihedral angle for this simulation is shown in figure 4.10. This plot shows transitions occurring from the *gg* to the *tg* conformation, but again no *gg*-*gt* transitions.

The calculation of the adiabatic map for maltose is not affected by the problem with the hydroxymethyl parameterization, as a simulated annealing method [82] was used to generate the map [75]. This method heats the molecule up to 600K, which gives it enough energy to overcome the rotational barriers.

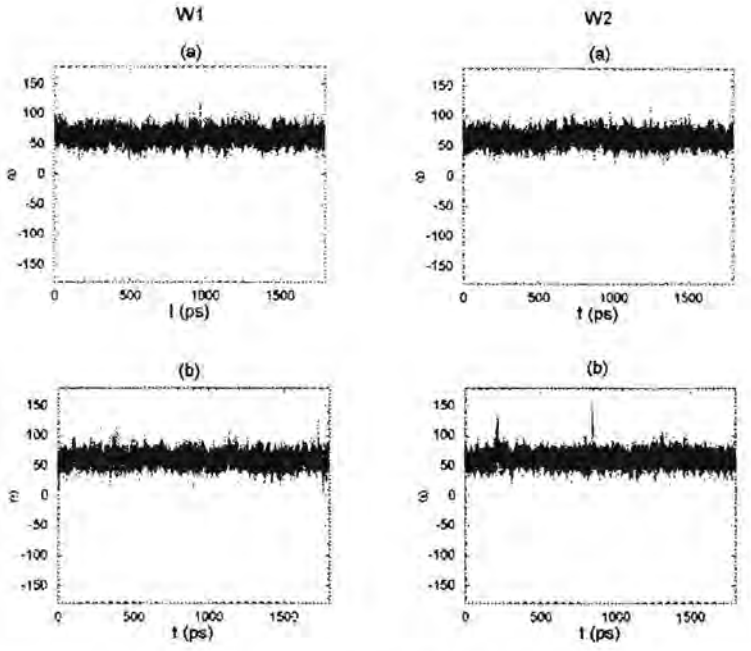

Figure 4.9: Times series of the angle ω ($O5$-$C5$-$C6$-$O6$) for the simulations in water. (a) and (c) : non-reducing residue (b) and (d) : reducing residue

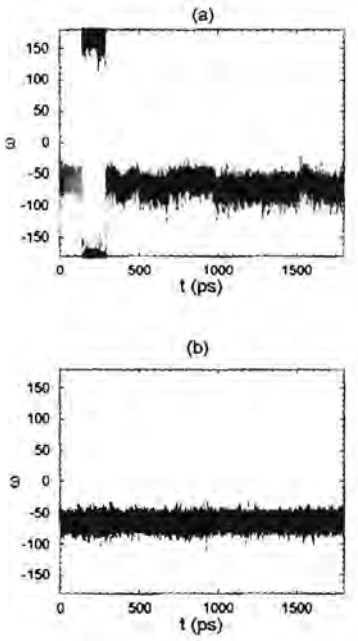

Figure 4.10: Times series of the angle ω ($O5$-$C5$-$C6$-$O6$) for simulation in vacuum started from *gg* primary alcohol conformations. (a) : reducing residue (b) non-reducing residue.

4.6 Statistical Analysis of the Trajectories in Vacuum and Water

We analysed the ϕ, ψ trajectory statistics for all the maltose simulations. The statistics for the simulations in vacuum, $V1$ and $V2$ show insignificant differences, therefore we combined the statistics data for these two simulations. As previously mentioned, the $W1$ and $W2$ simulations sample somewhat different regions of conformational space. This is clearly seen in the ϕ,ψ statistics for the simulations, and therefore we report the statistics for these two simulations separately. The statistics for the simulations are illustrated in the form of a ϕ,ψ contour plot, showing contours of equal sampling (figure 4.11).

	% time	
$\phi \pm 7.5, \psi \pm 7.5$	$W1$	$W2$
-15, -25	13.26	11.84
-25, -25	18.19	10.15
-50, -35	5.03	16.95
-60, -50	6.69	7.38

Table 4.4: Comparison of the percentage of simulation time spent in areas around popular conformations in the $W1$ and $W2$ simulations.

From an inspection of these these plots, it is again clear that in vacuum maltose covers a more limited area of the ϕ-ψ conformational space compared to the two water simulations. However, of more interest are the rather different popular conformations shown for the two simulations . The two vacuum simulations show the most popular conformation to be around ϕ = -20, ψ =-30. In contrast, the two water simulations each exhibit two popular conformations. The most popular conformation in the $W1$ simulation is around ϕ = -25, ψ =-25, which is quite close to the vacuum popular conformation. There is a second less popular conformation at (-60,-50). In contrast, the $W2$ most popular conformation is around (-50, -35), with another at (-15,-25). These differences show the effects that the initial conditions have on the simulation trajectory and allude to incomplete conformational sampling in the water simulations. However, it is clear that solvation of the maltose molecule leads to the creation of at least another low energy well, roughly in the (-55, -45) region of ϕ,ψ space. Table 4.4 compares the percentage time spent around the various popular conformations mentioned above for the $W1$ and $W2$ simulations.

4.7 Energy Analysis of the Maltose Trajectories

We calculated the internal energy of the maltose molecule for each point visited in ϕ,ψ space during the vacuum simulations, as described in section 3.3. The contour plot for this is shown in figure 4.12, where energies from the lowest 20 kcal conformations are

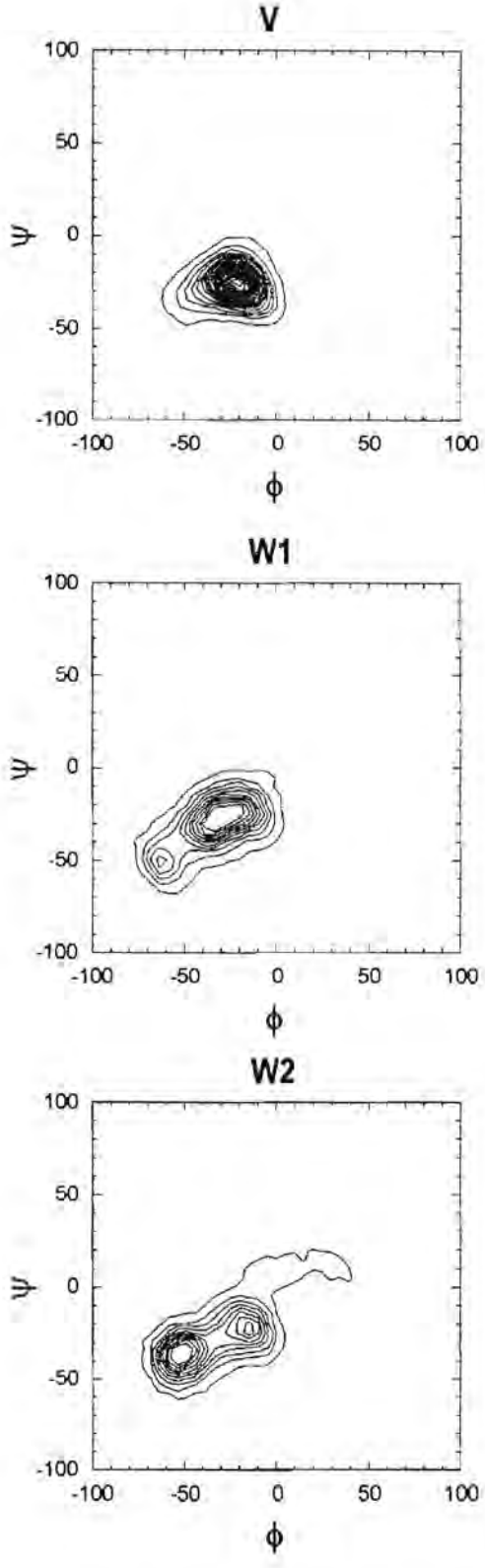

Figure 4.11: Plots of the combined statistics for the $V1$ and $V2$ simulations (labeled V) and the $W1$ and $W2$ simulations. The contours are plotted at intervals of 0.2% probability.

shown in contours of 4 kcals. Comparing this data to that of the V (ϕ,ψ) probability surface shown in figure 4.6, we found an exact correlation. Therefore, our MD results indicate that the most popular conformation for maltose in vacuum is enthalpy dominated. This result is consistent with the previously calculated free energy surface for maltose [76], albeit using the older force field [45].

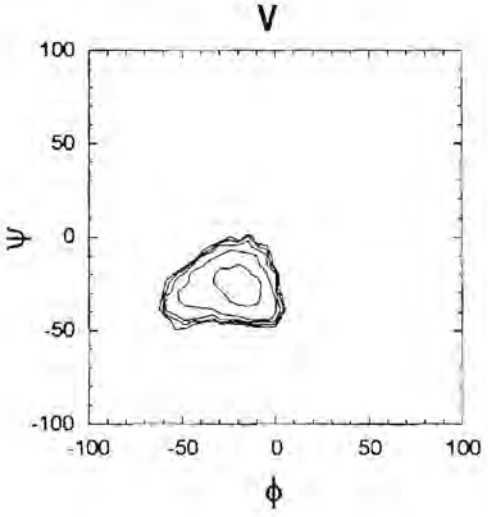

Figure 4.12: The internal energy contour plot for the vacuum simulations, contoured at intervals of 4 kcal.

As a free energy surface for maltose in water at 300K has not as yet been calculated, an analysis of the enthalpic energy surface of maltose in water based on the 1.8 ns simulations may give some important insights into the relationship between water and maltose at the molecular level. It is particularly interesting to compare the mechanical energy of maltose in solution compared to vacuum. In order to calculated the effective enthalpy of maltose in water, we combined the energy data from the two simulations. It was hoped that this would give us better sampling of conformational space. The effective enthalpy of maltose in water is shown as a function of the glycosidic linkage conformations in figure 4.13.

4.8 Analysis of the Water Densities About Maltose

Water density analyses of the type described in section 3.1.2 have been performed for several pentoses [37, 33] and the relatively rigid disaccharide α,α-trehalose [36]. However, to our knowledge, this is the first time that this type of calculation has been performed on a disaccharide as flexible as maltose.

In order to investigate the water structuring about maltose we calculated the water densities around four conformations chosen from the broad region of ϕ, ψ conformational space sampled during the simulations. The conformations chosen were (-15,-20), (-30,-30), (-55,-40) and (-60,-50). They are shown super-imposed on the adiabatic map

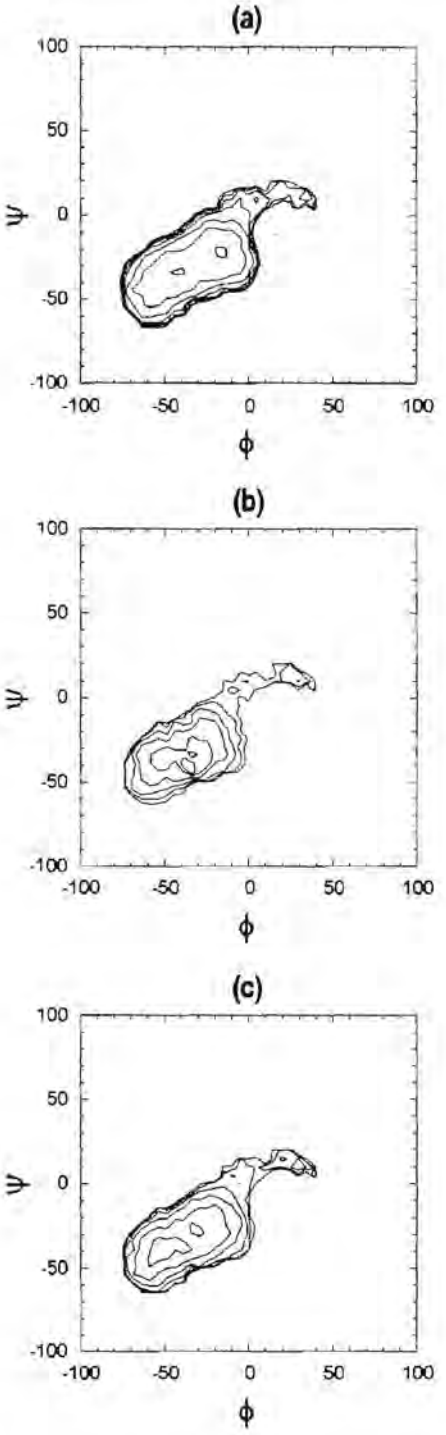

Figure 4.13: Various enthalpy plots, using the combined data from the two water simulations. **(a)** map of the internal maltose energy, contoured at 4 kcal. **(b)** map of the maltose-water interaction energy, contoured at 4 kcal. **(c)** map of the total maltose energy, contoured at 4 kcal.

in figure 4.14 as A, B, C and D respectively. Ranges of ± 3 degrees about each of the dihedral angles were used to select frames for analysis from the simulations. 180 frames were used for each density calculation, with the stipulation that frames are no closer together than 0.5 ps. This helps to avoid artificially high water densities cause by consecutive frames being too highly correlated with each other. The first 300ps of the simulations were not considered for the water density analysis in order to ensure that the data is fully equilibrated.

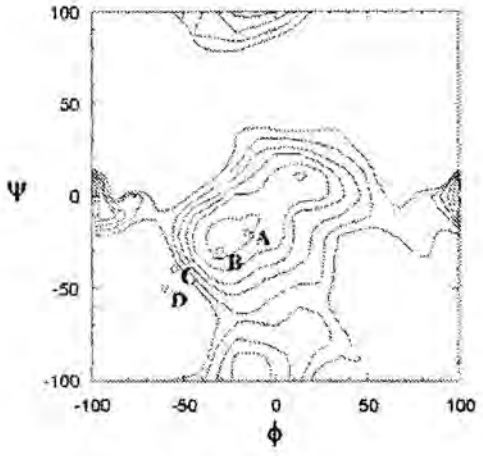

Figure 4.14: ϕ, ψ conformations selected for water density analysis

The water density pictures for the conformations chosen are shown in figure 4.15. The dark grey central masses indicate no water density (i.e. the maltose molecule) and the light grey regions show water volumes contoured at 65% above bulk density. A comparison of the water densities for the two simulations $W1$ and $W2$ reveals that the densities for the different simulations are not greatly different. Differences that there are can be ascribed to the insufficient sampling of conformational space. Also, water densities for the individual conformations of the maltose molecules are similar, showing no dramatic differences. This was to be expected.

For these diagrams, it can be seen that the water molecules preferentially occupy specific sites located close to the hydroxyl substituents on the maltose molecule.

4.9 Vornoi Analysis

Vornoi analysis, as described in section 3.1.3, was performed on the $W1$ and $W2$ simulations in order to investigate whether maltose imposes greater structure on the surrounding water than it would otherwise exhibit. Vornoi analysis was performed separately in two circular shells around the solute molecule, one close to the molecule and one further away. The shells were defined as in figure 4.16, with the inner shell being the volume between a radius of 6 $\AA$ and 9 $\AA$ from the central glycosidic oxygen,

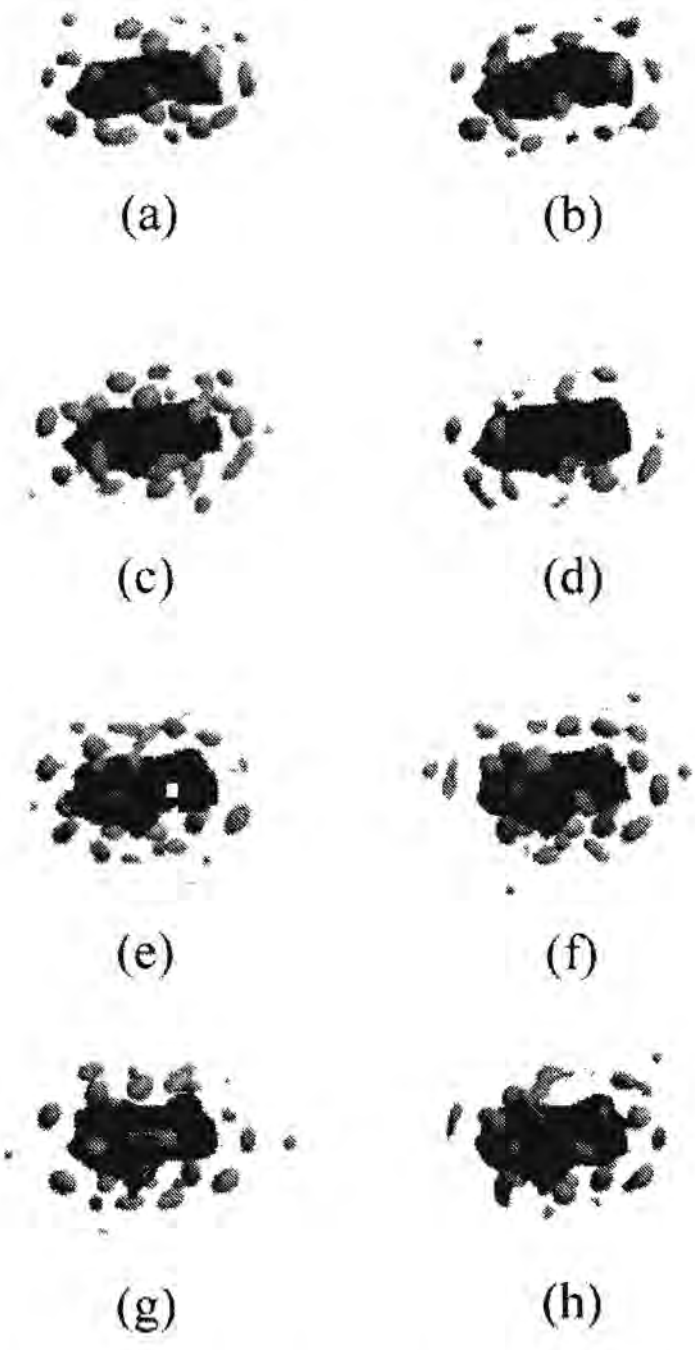

Figure 4.15: Average water densities about maltose for various conformations from the $W1$ and $W2$ simulations, contoured at 65% above bulk density. **(a)** and **(b)**: ϕ = -15, ψ = -20; **(c)** and **(d)** : ϕ = -30, ψ = -30; **(e)** and **f)** : ϕ = -55, ψ = -40; **(g)** and **(h)** : ϕ = -60, ψ = -50

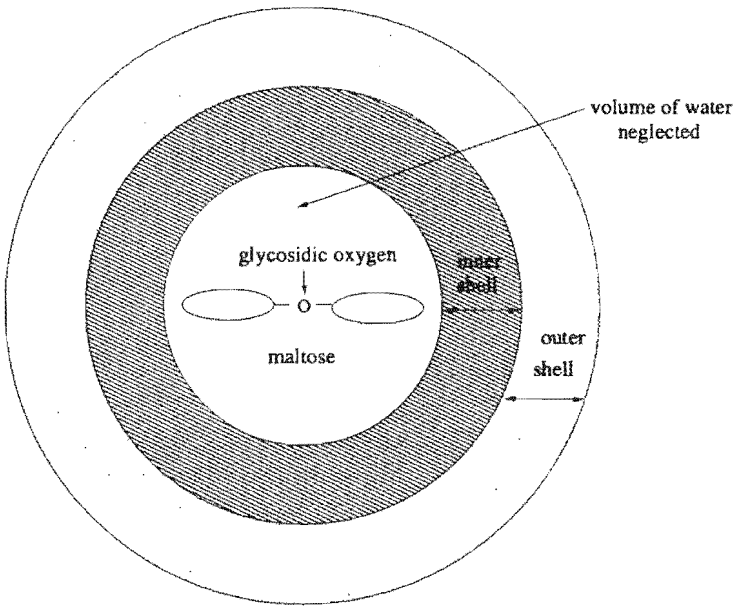

Figure 4.16: Diagram illustrating the inner and outer shells chosen for Vornoi analysis.

and the outer shell being between 9 and 14 $\mathring{A}$. The use of spherical shells in which to perform the Vornoi analysis is not ideal, as a considerable portion of the water close to the maltose molecule is ignored (indicated on figure 4.16) - an ellipsoidal shell around the solute molecule should rather be used. However, an ellipsoidal shell that would mimic the maltose shape would be is difficult to define, and the spherical shell approximation should give us a rough indication of the comparative water structuring close to and further away from the maltose molecule. The number of faces was only calculated for the water oxygen atoms and the analysis was performed over the entire trajectory in each case, with intervals of 0.5 ps between successive frames analysed.

	Average no. faces for the Vornoi polyhedra.	
Shell	$W1$	$W2$
inner	15.69(1.97)	15.68(1.97)
outer	15.51(1.94)	15.50(1.94)

Table 4.5: The average number of faces for Vornoi polyhedra around each oxygen atom in the $W1$ and $W2$ simulations.

The average number of faces found for the two shells in both simulations are listed in table 4.5. As expected, the values for the two simulations are very similar. The inner shell shows in each case a slightly higher average number of faces compared to the outer shell, indicating increased structure. In addition to the analysis of the simulations of maltose in water, we performed Vornoi analysis on 30ps of a simulation of the large water box alone. The average number of faces was calculated to be 15.51 for the entire box. Schnitker and Mausbach performed a Vornoi analysis study of 40 configurations of

a molecular dynamics simulation of 216 ST2 [83] water molecules of density $1gm/cm^3$ [71]. They report the average number of faces for the Vornoi polyhedra around each oxygen to be 15.97. Considering the fact that a different water model was used in our case, the two values are not significantly different.

Histograms comparing the average number of faces for the Vornoi polyhedra in the inner and outer shells for the $W1$ and $W2$ simulations are shown in figure 4.17. It is clear from these graphs that the water molecules in the inner shell show a shift towards higher numbers of faces compared to those in the outer shell. Thus we conclude that the maltose solute does impose a greater degree of structure on the surrounding water.

4.10 Amylose Strand

A comparison of the behaviour of a hexa-amylose strand in water with that of maltose should give an indication of whether the behaviour of the maltose dimer can be extrapolated to account for the effects of water on amylose. However, one must bear in mind that inter-monomer interactions in amylose, such as the interactions between different monomers in the chain that typically stabilise helical conformations, are not present in the maltose simulation. Therefore, of principal interest is whether the effects of water on the dynamics around the $\alpha(1 \rightarrow 4)$ glycosidic bond in maltose are comparable to those for the hexa-amylose strand and, ultimately, a large amylose molecule.

4.10.1 Analysis of ϕ, ψ Dihedral Angles

The glycosidic angles in the hex-amylose strand are numbered from the non-reducing end to the reducing end of the strand. The average values of the ϕ, ψ angles from the simulations of amylose in water and in vacuum are listed in table 4.6.

Angle	Water	Vacuum
$\phi 1$	-44.31(9.62)	-13.1(7.99)
$\phi 2$	-28.84(10.04)	-13.96(8.22)
$\phi 3$	-42.09(16.96)	-16.37(9.00)
$\phi 4$	-36.50(9.41)	-16.95(8.26)
$\phi 5$	-39.60(13.32)	-50.74(12.11)
$\psi 1$	-42.10(11.19)	-21.53(9.80)
$\psi 2$	-28.58(8.57)	-23.10(9.23)
$\psi 3$	-34.85(11.08)	-25.77(8.75)
$\psi 4$	-25.47(7.54)	-26.45(8.08)
$\psi 5$	-31.18(-12.28)	-41.66(9.63)

Table 4.6: Comparison of the average ϕ,ψ values for the simulations of the hexa-amylose strand in vacuum and in water.

Scatter plots of the ϕ, ψ angle pairs for the hexa-amylose strand simulation in vacuum and water are presented in figures 4.18 and 4.19 respectively.

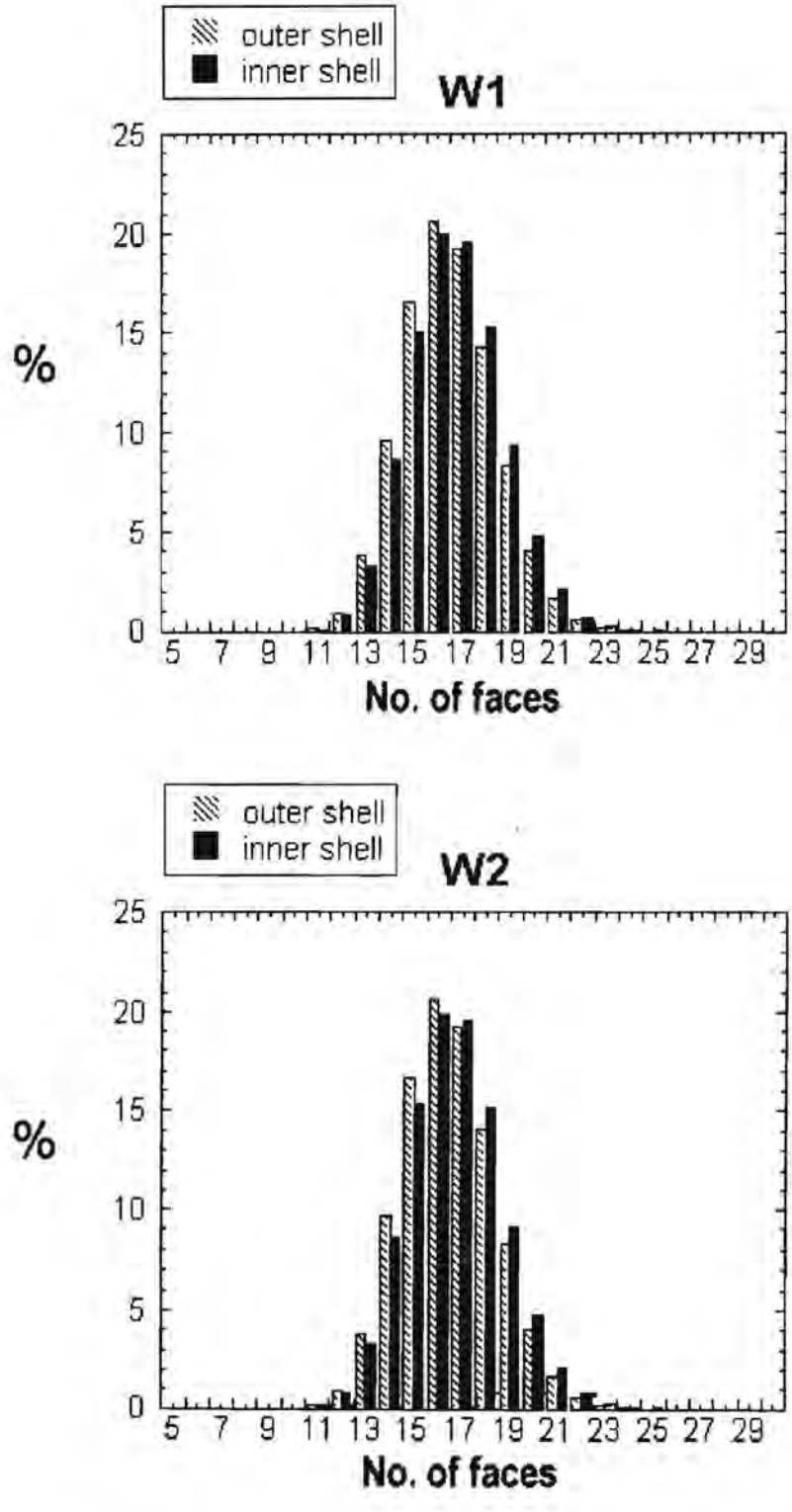

Figure 4.17: Histograms of the average number of faces for the $W1$ and $W1$ simulations. (**a**) and (**b**) inner shell, (**c**) and (**d**) outer shell.

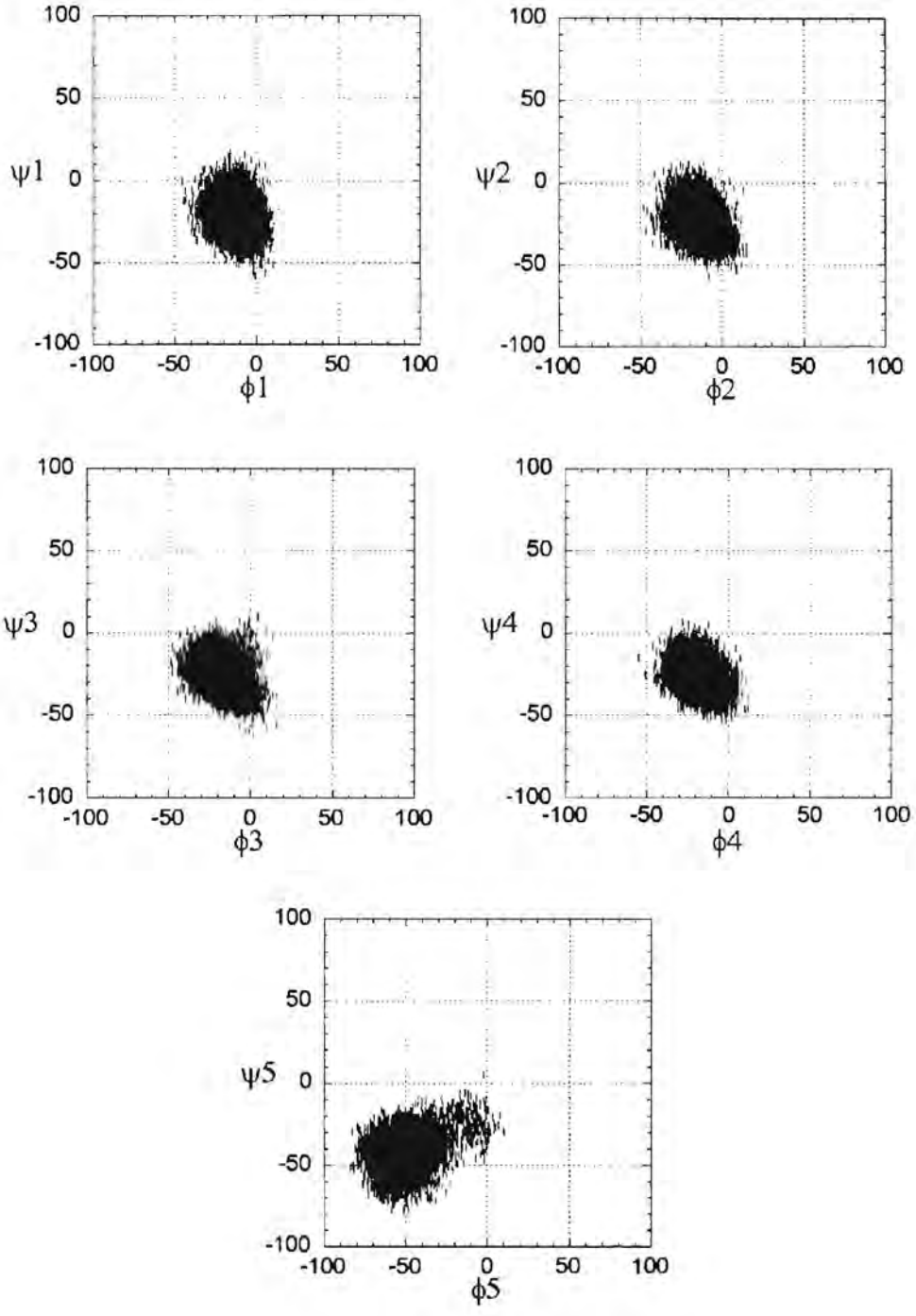

Figure 4.18: Scatter plots for the ϕ, ψ angle pair for the simulation of the hexa-amylose strand in vacuum.

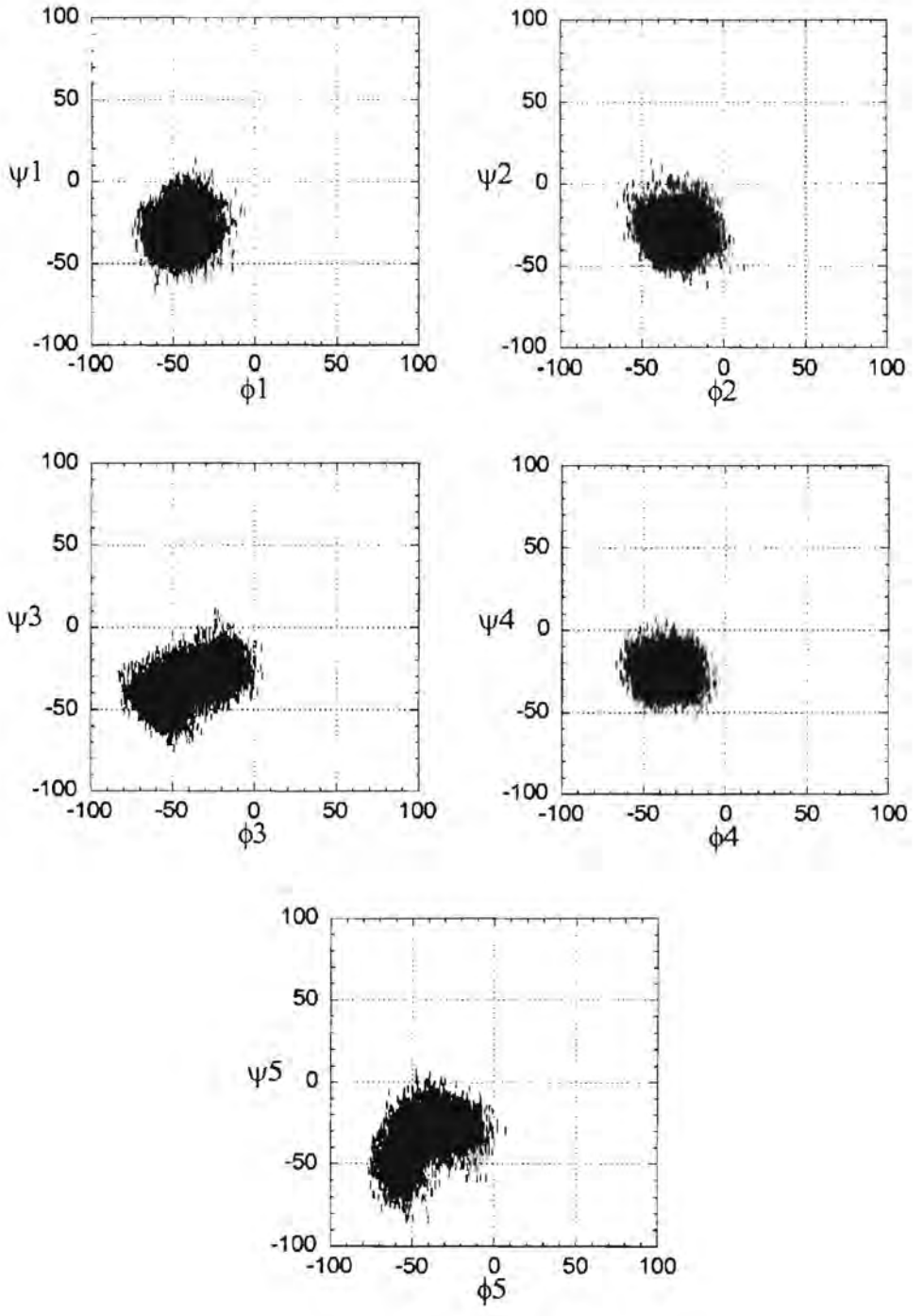

Figure 4.19: Scatter plots for the ϕ, ψ angle pair for the simulation of the hexa-amylose strand in water.

These figures show that the simulations of the hexa-amylose strand in vacuum and in water sample same region of ϕ,ψ space as is shown in the analogous maltose simulations. The addition of solvent once more results in the sampling of a wider area of phase space as compared to vacuum. The area sampled by the individual ϕ,ψ angles is , in general, not as broad as that sampled in the maltose simulations $W1$ and $W2$, but this is to be expected for two reasons. Firstly, the simulation time for the amylose strand simulations is shorter and so a smaller region of conformational space is explored. Secondly, the larger size of the amylose strand results in an increased restriction of movement in the individual dihedrals. The special case is the terminal reducing residue in both the vacuum and water simulations: here the $\phi 5$,$\psi 5$ angles explore a wider area than expected. This is probably a result of interactions between the relatively mobile terminal glucose monomer with other monomers in the amylose strand. The time series plot for $\phi 5$,$\psi 5$ dihedral angles (figure 4.20) show that a transition to lower values of ϕ and ψ is made fairly late in the simulation.

This study thus supports the assumption that the motions of the ϕ,ψ torsion angles in amylose can be extrapolated from the motions of the maltose dimer.

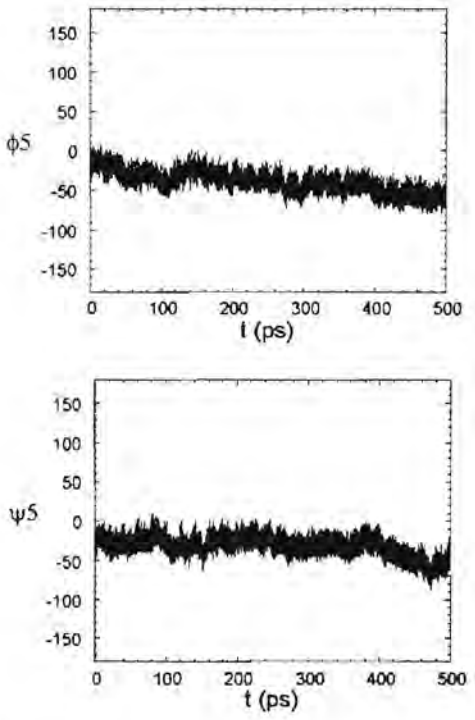

Figure 4.20: Time series for the $\phi 5$ and $\psi 5$ angles for the simulation of hexa-amylose in water.

4.10.2 Water Densities About Amylose

The hexa-amylose strand has a wide range of conformational flexibility and thus a simulation time of 500 ps is certainly not adequate to completely sample the conformational space of this molecule. This presents a problem for the water density analysis, as too specific a definition of the ϕ, ψ dihedral angles values for a particular conformation results in too few frames for a satisfactory analysis of the water density. Therefore, in order to select a conformation for the water density analysis, we used only the six dihedral angles specifying the inner glycosidic linkages as selection criteria. The values chosen are listed in table 4.7. These correspond to positions roughly in the center of the ϕ, ψ scatter plots for the strand (figure 4.19). The $\phi 3$, $\psi 3$ dihedral angle pair, being in the center of the strand, has the greatest effect on the overall strand conformation and is thus restricted the most.

Dihedral angle	Value
$\phi 2$	-30.0 ± 8.0
$\psi 2$	-29.0 ± 8.0
$\phi 3$	-55.0 ± 4.0
$\psi 3$	-40.0 ± 4.0
$\phi 4$	-37.0 ± 8.0
$\psi 4$	-26.0 ± 8.0

Table 4.7: Values chosen for ϕ, ψ dihedral angles defining conformation of the amylose stand for water density analysis.

Three different views of the resultant three-dimensional water density maps, contoured at 70% above bulk water density, are displayed in figure 4.21. Once again, the solute molecule is shown in dark grey, while the contoured water density regions are the lighter grey areas.

The amylose strand can be seen in these pictures to have a number of well-defined areas of high water density located around the strand.

Figure 4.22 show a view of the strand with the water densities contoured at 230%. From this picture, one can see that the highest water densities form a chain up the center of the central cavity in the amylose helix. Figure 4.23 shows that these areas of extremely high water density correspond to hydrogen-bonded water molecules situated in the central helical cavity.

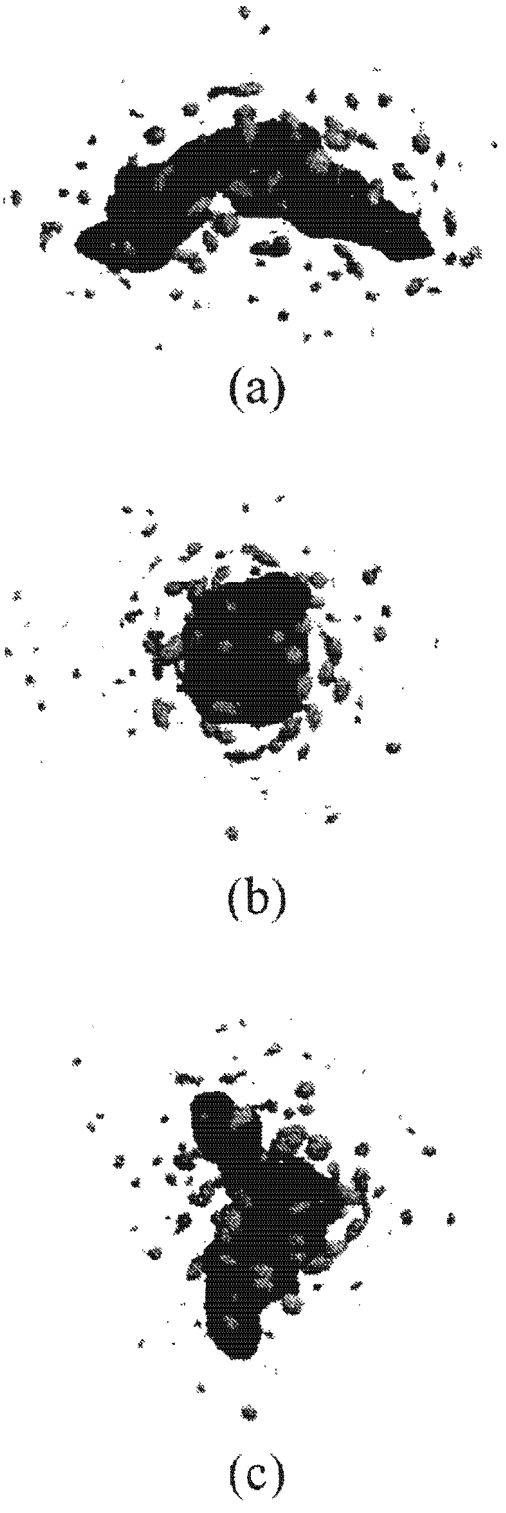

Figure 4.21: Different views of the average water densities about the amylose strand, contoured at 70% above bulk density. (**a**): view along the length of the strand; (**b**): view straight down the helix; (**c**): oblique view, showing water density around the helix.

Figure 4.22: Views of areas of high average water densities about the amylose strand, contoured at 230% above bulk density.

Figure 4.23: Snapshot of the hexa-amylose helix, showing hydrogen-bonded water molecules situated in the central helical cavity.

Chapter 5

Molecular Dynamics Simulations of Solvated Cyclodextrins

Cyclodextrins are chiefly of interest because of their ability to form inclusion complexes with a variety of guest molecules. The complexation process is generally brought about in aqueous solution. Here the hydrophobic central cavity of the cyclodextrins in occupied by water molecules, which are readily displaced by a less polar guest molecule. In the cyclodextrin rings, a total of 3 hydroxyl groups per glucose residue provide donors for hydrogen-bonding. Thus, it can be surmised that cyclodextrins interact strongly with water, and that the use of water as a solvent would have a significant effect on cyclodextrins. As water is the medium in which the complexation process takes place, a knowledge of the effects of water on the cyclodextrin structures and, correspondingly, the effects of the cyclodextrins on the water structure, should give insights into the mechanism of guest complexation.

cyclodextrin	α	β	γ
no. of glucose units	6	7	8
molecular weight	972	1135	1297
solubility in water $(g/100ml)$	14.5	1.85	23.2
cavity diameter $\AA$	4.7-5.3	6.0-6.5	7.5-8.3
torus diameter $\AA$	14.6 $\pm$4	15.4 $\pm$4	17.5 $\pm$4
torus height $\AA$	7.9 $\pm$1	7.9 $\pm$1	7.9 $\pm$1

Table 5.1: Characteristics of α-, β- and γ-cyclodextrin.

5.1 Conformational Descriptors for Cyclodextrins

In addition to the standard geometric descriptors for the $\alpha(1 \rightarrow 4)$ linked polysaccharides already described (see section 1.4.1), there are some additional descriptors that are particularly useful for characterising the cyclodextrin macrocycle. These are described below.

5.1.1 Cross-Ring Distances

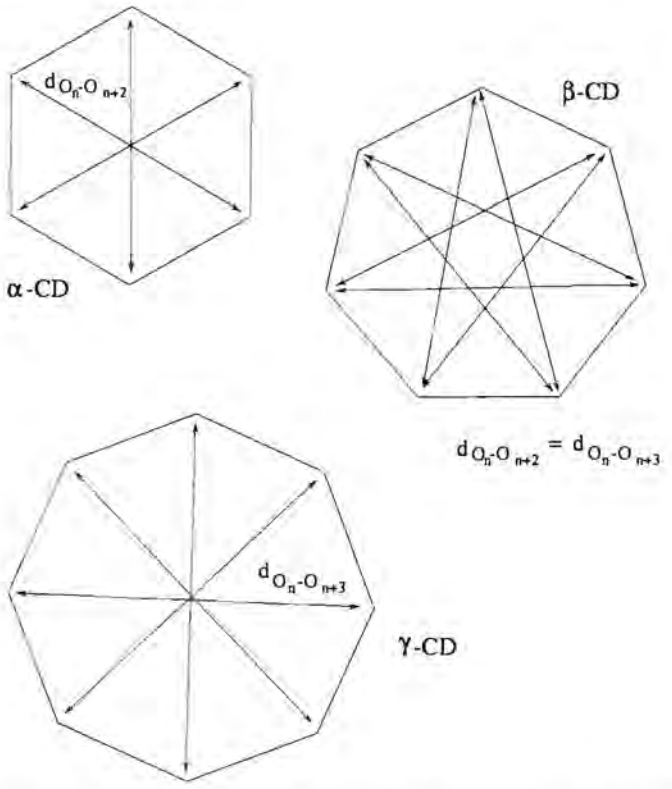

Figure 5.1: Schematic of the cross-ring distance definitions in the three cyclodextrins.

The cross-ring distances are useful for characterizing shape of the cyclodextrin cavities. These distances are defined as the distance from one inter-saccharide oxygen, $O1$, to the inter-saccharide oxygen diagonally across the ring. In the case of α-CD, this is the distance from O_n to O_{n+2}, while in γ-CD it is the distance from O_n to O_{n+3}. In β-CD these two distances are equivalent (see figure 5.1).

5.1.2 Tilt Angles

The tilt angle τ denotes the inclination of the pyranose rings towards the macrocyclic ring perimeter. This is defined as the angle between the $O4$ plane and the plane through $C1$, $C4$, $O4$ and $O4'$ of each glucose residue. Tilt angles provide a clear illustration of the inclination of the glucose monomers with respect to the macrocycle of the cyclodextrin. If the glucose moiety is inclined with its $O6$ side towards the center of the cavity, then the tilt angle is positive, but if the glucose moiety is inclined away from the cavity, then the tilt angle is assigned a negative value.

5.2 Simulation Procedure

The molecular dynamics simulations were performed as discussed in section 2.1.4. Dynamics runs of 520ps (20ps equilibration and 500ps data collection) in length were performed on the $\alpha-$, $\beta-$ and γ-cyclodextrins in water. All the cyclodextrins were simulated in a cubic box containing one cyclodextrin molecule and 4040 SPC/E [66] water molecules, using periodic boundary conditions. The box lengths were altered in order to get a density of $1.003 g/cm^3$ in each case. These lengths are listed in table 5.2.

CD	Box dimension ($\AA$)
α	49.61050
β	49.64682
γ	49.68309

Table 5.2: Box dimensions for the cyclodextrins

The starting structures for the α-, β- and γ-cyclodextrins were generated by building the structures in CHARMM [58] and then minimizing the resulting structures while constraining the torsion angles in the glucose rings to values that maintain the chair conformation in each ring. The dihedral angles of the carbohydrates rings were constrained using the potential $V_{\phi con} = \sum k_i(\phi_i - \phi_{i0})^2$ to maintain the $^4C_1$ ring conformations during minimization. No restraints were applied during the MD simulation. Configurations were stored at intervals of 0.5 ps.

5.3 Dynamics of the $\alpha(1 \rightarrow 4)$ Linkages

Scatter plots for the different ϕ,ψ angle pairs in α-, β- and γ-CD are shown in figures 5.2, 5.3 and 5.4. These figures show the range of ϕ,ψ conformational space explored in each glycosidic linkage during the simulation. As previously mentioned, the ϕ, ψ torsion angles are the principal determinators of molecular conformation in the polysaccharides. Thus, the conformational space sampled by the glycosidic bond dihedral angles gives an indication of the extent of deformation of the cyclodextrin ring during the simulation.

The scatter plots for α-CD show that the different glycosidic linkages in this molecule explore similar broad regions of ϕ, ψ conformational space, centered roughly on the (0,0) conformation. Interestingly, a considerable amount of the simulation time is spent in the region around (-20,30). This region is not explored in either the maltose simulation (figure 4.3) or the simulation of the amylose strand (plots on pages 55 and 56) and corresponds to a higher-energy region in the adiabatic map for maltose (figure 4.2). This can be explained by the fact that α-CD is sterically strained [15]. The inter-pyranose $O2$-$O3$ hydrogen bonds present a less than ideal geometry for α- as compared to β- or γ-cyclodextrin, resulting in a relatively strained macrocycle.

As a result of the increased strain energy, the dihedral angles exhibit higher-energy conformations than are seen in the essentially unstrained maltose and hexa-amylose

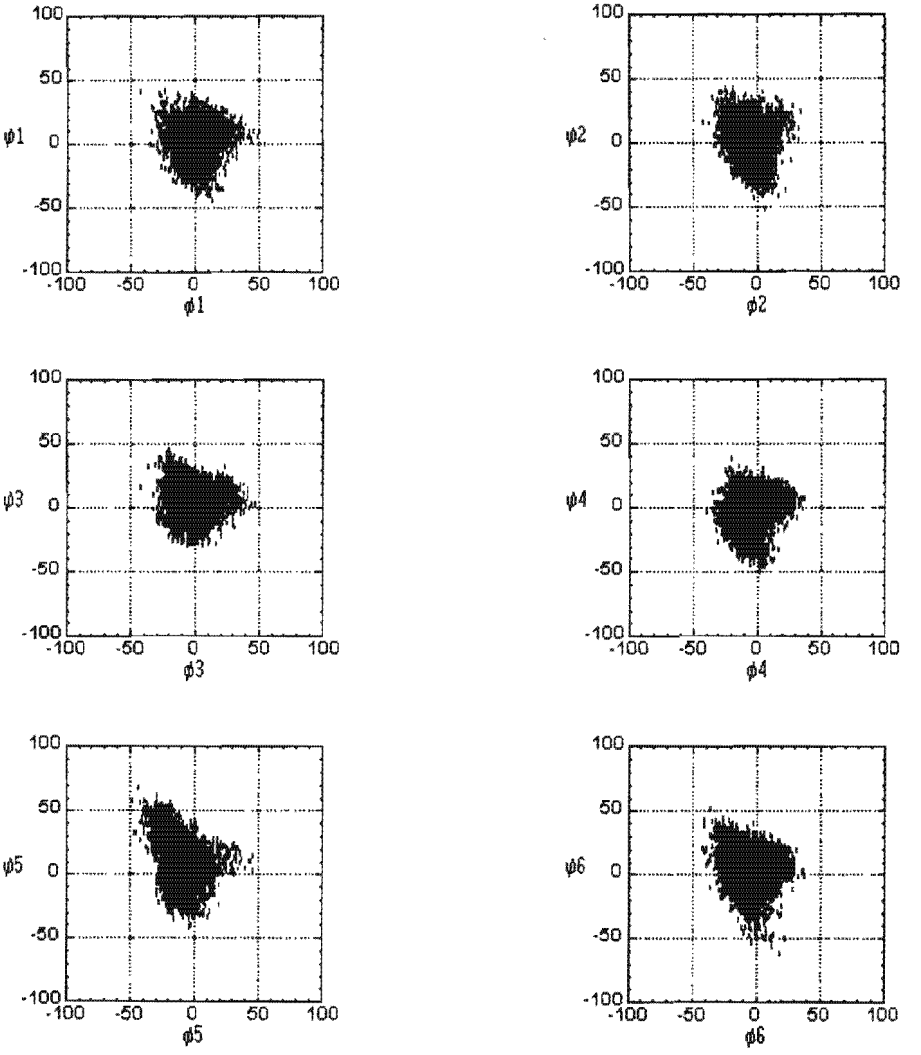

Figure 5.2: Scatter plots for the ϕ,ψ dihedral angle pairs in α-CD.

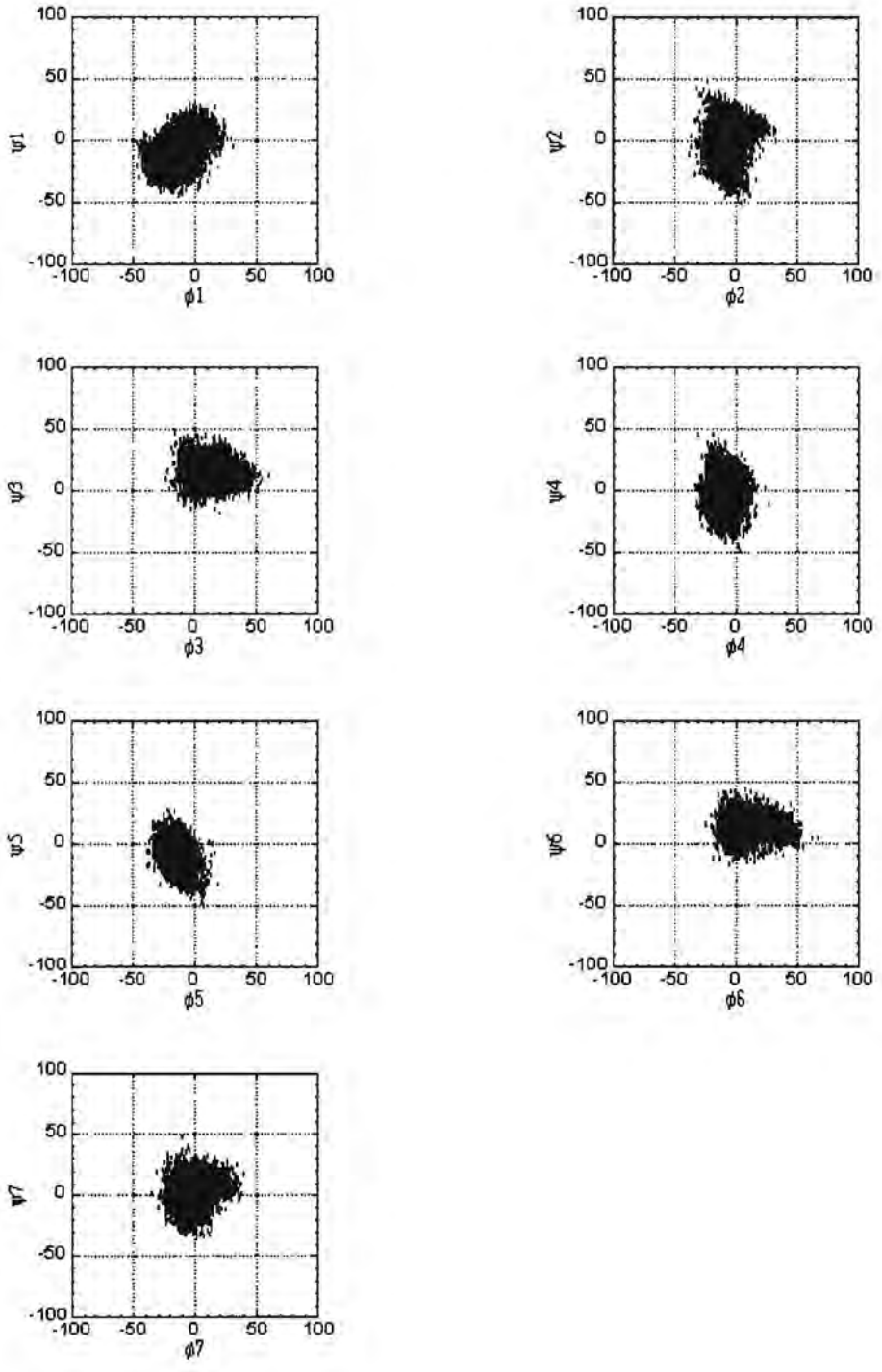

Figure 5.3: Scatter plots for the ϕ,ψ dihedral angle pairs in β-CD.

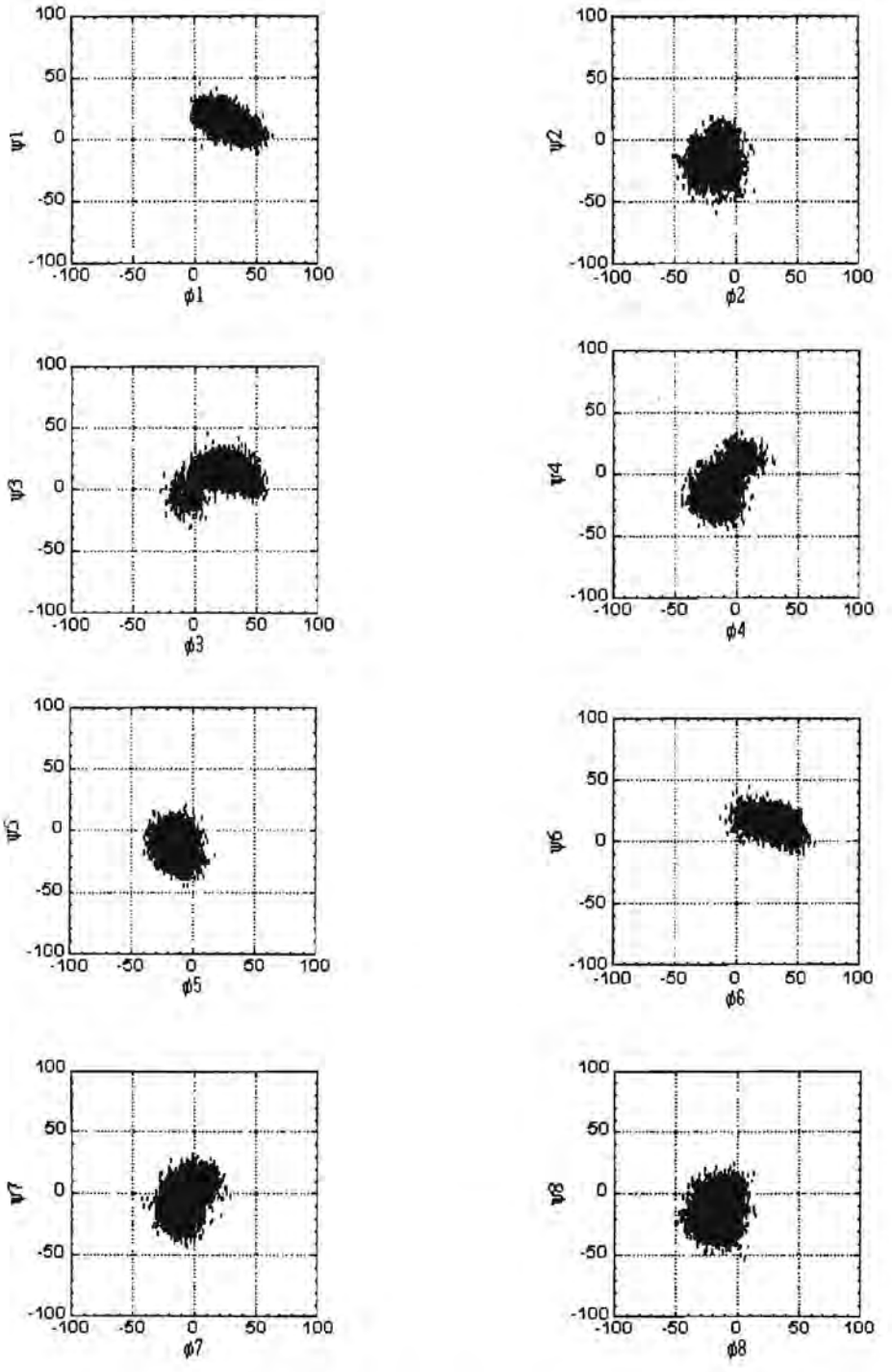

Figure 5.4: Scatter plots for the ϕ,ψ dihedral angle pairs in γ-CD.

molecules. This explanation is supported by the fact that β-CD, which is less strained than α-CD, spends far less time in the (-20,30) region than α-CD. In addition, γ-CD, which is the least strained of the three, does not visit this region at all.

In contrast to α-CD, the scatter plots for β- and γ-CD show different glycosidic linkages sampling different areas of ϕ, ψ conformational space during the simulations. This is particularly apparent in the γ-CD scatter plots, which show very clearly two separate populations for the different glycosidic linkages. The population centered on (-20,-20) corresponds rather well to the area explored in the simulations of maltose and amylose. The other population, centered approximately on (20,15), corresponds to the small minimum on the maltose adiabatic map and thus can reasonably be assumed to be another low-energy region.

Some of the dihedral angles make transitions between the two populations. The scatter plots for γ-CD show that the $\phi 5$, $\psi 5$ angle pair spends the entire simulation in the (-20,-20) area, while the $\phi 3$, $\psi 3$ pair makes a transition to the (20,15) region. This transition can be clearly seen in the time series graphs of theses dihedral angles, shown in figure 5.5.

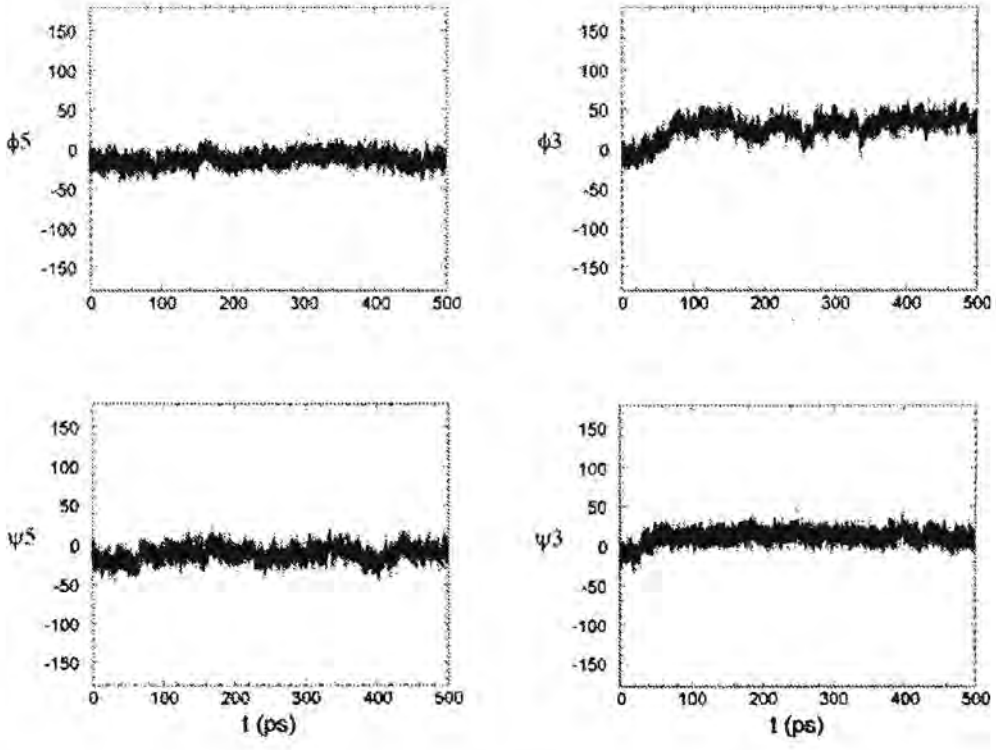

Figure 5.5: Graphs of the variation with time of the $\phi 5$, $\psi 5$, $\phi 3$ and $\psi 3$ dihedral angles in γ-CD.

The differences between the ranges of ϕ,ψ conformational space sampled in the α-, β- and γ-CD simulations can be further clarified by comparing a graph of all the scatter

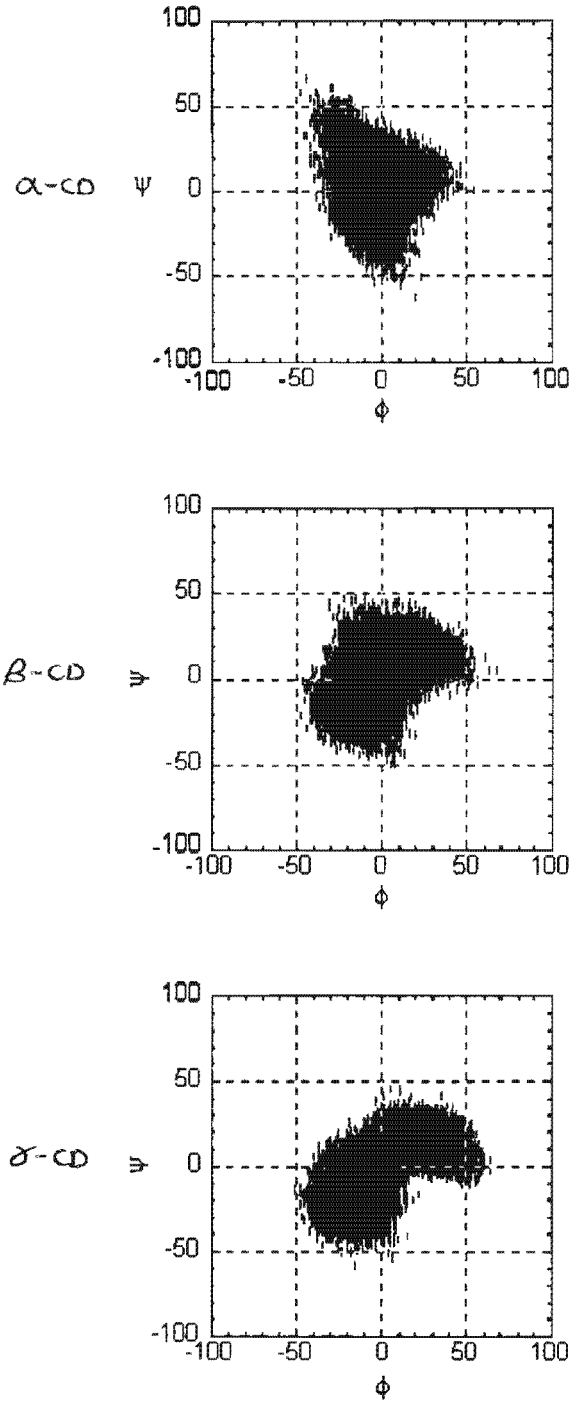

Figure 5.6: Total scatter plots for the ϕ,ψ angles in each of the cyclodextrins.

plots for the ϕ, ψ angles super-imposed on one another (figure 5.6). From these plots, it can be clearly seen that the ϕ, ψ dihedrals in α-CD sample a rather different region to either β- or γ-CD, which are more similar. The shape of the scatter plot for β-CD simulation can be seen to be intermediate between the α-CD plot and the γ-CD plot. The latter is clearly separated into two distinct populations, and occupies a smaller region than either of the other two.

5.4 Mean Molecular Parameters for Each of the Cyclodextrins

It has been a matter of some discussion whether the cyclodextrin crystal structures are representative of their average solution conformation, or not [15]. Some insight into this issue may be obtained by comparing some mean molecular parameters obtained from a statistical analysis of the cyclodextrin solid state structures [84, 15] with the mean parameters obtained from the MD simulations. All the parameters were calculated ignoring the first 20ps equilibration period.

5.4.1 ϕ, ψ Torsion Angles

Table 5.4.1 compares the mean values for the ϕ, ψ torsion angles with those obtained from the hydrate crystal structures [84]. (The various alternative definitions for the ϕ, ψ dihedral angles are defined in section 1.4.1).

	Glycosidic linkage torsion angles					
	ϕ	ψ	ϕ^{a2}		ψ^{a2}	
CD	sim.	sim.	sim.	crys.	sim.	crys
α	-3.2(12.0)	3.2(14.4)	115.4(10.8)	109.2	122.0(13.0)	128.8
β	-0.9(15.5)	0.5(15.4)	112.1(14.1)	109.8	119.6(14.0)	127.6
γ	3.2(23.4)	-3.0(16.4)	120.94(22.43)	108.9	116.47(15.38)	127.1

Table 5.3: Mean values for torsion angles from crystal structures and our MD simulations.

The average ϕ, ψ values for the simulations differ significantly from those of the crystal structures. The difference is particularly marked in γ-CD. The deviation can be explained, however, by considering the different populations of ϕ,ψ angles. The crystal structures correspond to ϕ, ψ angles in the (-20,-20) region.

The standard deviation of the ϕ,ψ angle values can give an indication of the relative mobilities of the cyclodextrin rings. However, as the ϕ,ψ angles fall into separate populations, the standard deviation of all the ϕ or ψ angles together gives an incorrect impression of this mobility. A better measure is the average of the individual standard deviations of each ϕ or ψ angle displayed in table 5.4.1. From this table, it is clear that α-CD shows the large standard deviations in its ϕ, ψ angles, and hence the greatest

CD	Average standard deviations	
	ϕ	ψ
α	11.584	13.457
β	10.814	11.470
γ	10.067	9.453

Table 5.4: Averages of the individual ϕ, ψ standard deviations.

mobility. The mobility of the cyclodextrins decreases in the order α-CD > β-CD > γ-CD.

5.4.2 Glycosidic Angles

The mean values for the glycosidic angles for the CD's in the MD simulations and those from a survey of crystal structures [15] are reported in table 5.5.

CD	$\angle C1\text{-}O1\text{-}C4$ φ	
	sim.	crys
α	117.6(2.8)	118.4(2.0)
β	116.8(2.8)	117.7(2.6)
γ	116.3(2.9)	115.0(3.1)

Table 5.5: Mean values for the glycosidic linkage angles.

The angles, though generally smaller than in the crystal structures, follow the same trend of decreasing size in the order α-CD > β-CD > γ-CD. This, once again, shows decreasing strain in the cyclodextrin rings. The standard deviations are of similar size, showing that the angles have roughly the same mobility in each of the rings.The larger size of the crystal structure average can be explained by the fact that a high proportion of these were inclusion compounds, and thus can be expected to have slightly stretched cavities.

5.4.3 Inter-Atomic Distances

The Cross-Ring Distances

The average cross-ring distances are smaller in the simulation structures than in the crystal structure analysis, which is fairly predictable considering that the crystal structure averages are mostly averages of cyclodextrin inclusion compounds which can be expected to increase the ring size.

Time series plots of the cross-ring distances for α, β and γ are shown in figures 5.7, 5.8 and 5.9 respectively.

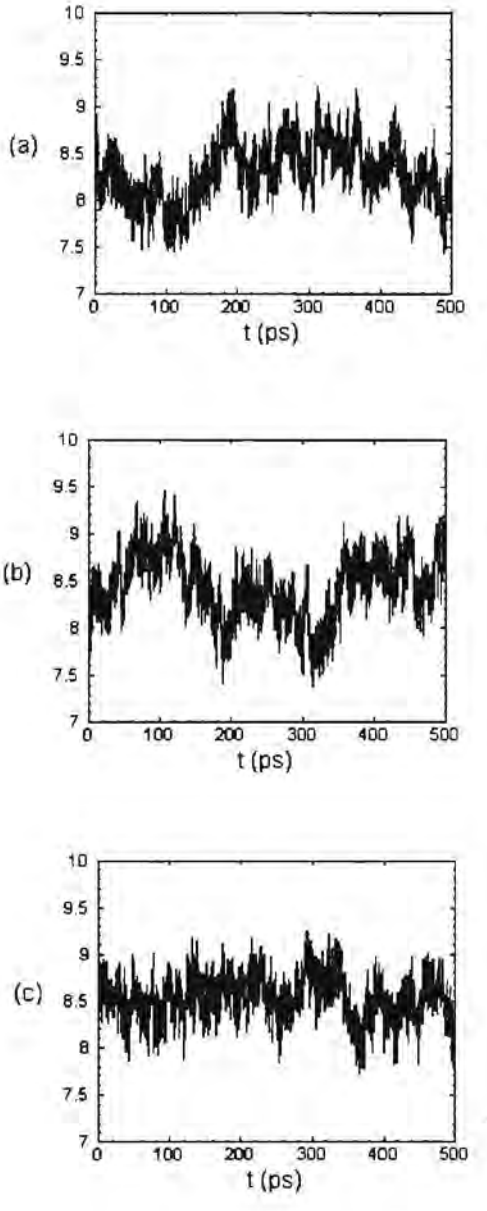

Figure 5.7: Time series plots of the cross-ring distances in α-CD. **(a)** $O1$-$O4$, **(b)** $O2$-$O5$, **(c)** $O3$-$O6$.

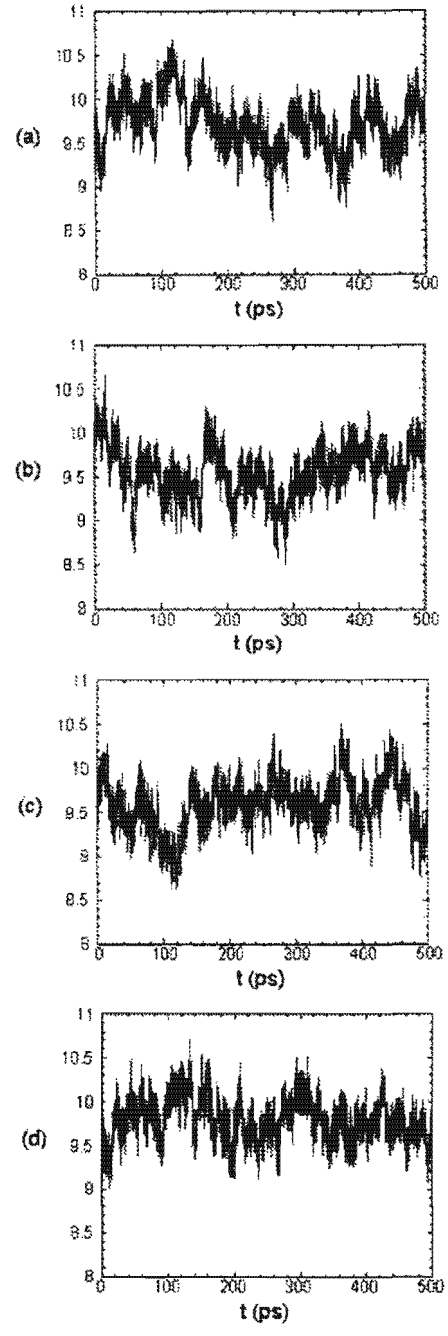

Figure 5.8: Time series plots of the cross-ring distances in β-CD. **(a)** $O1$-$O5$, **(b)** $O2$-$O6$, **(c)** $O3$-$O7$, **(d)** $O4$-$O1$.

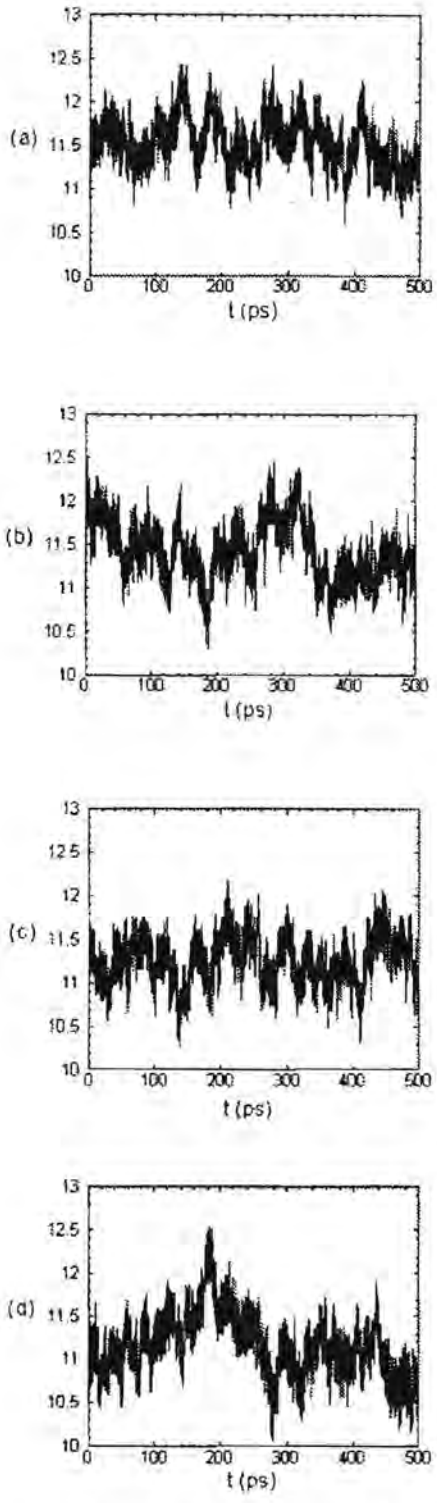

Figure 5.9: Time series plots of the cross-ring distances in γ-CD. **(a)** *O*1-*O*5, **(b)** *O*2-*O*6, **(c)** *O*3-*O*7, **(d)** *O*4-*O*8.

CD	Atomic distances (in $\AA$)					
	$O_1 - O_{1^n}$		$O_2 - O_{3'}$		$C_6 - C_{6'}$	
	sim.	crys.	sim.	crys.	sim.	crys.
α	8.42(0.32)	8.50(0.22)	2.96(0.27)	3.05(0.55)	4.82(0.39)	4.47(0.21)
β	9.65(0.31)	9.83(0.24)	2.91(0.21)	2.92(0.27)	4.86(0.50)	4.64(0.34)
γ	11.38(0.32)	11.76(0.07)	2.90(0.22)	2.84(0.06)	5.00(0.69)	4.39(0.30)

Table 5.6: Mean values for atomic distances from crystal structures and our MD simulations.

The cross-ring distance are discussed further in the section on clustering analysis.

The $O2 - O3'$ Distances

The cyclodextrins' toroidal shape stems from a characteristic hydrogen bonding pattern and the conformation of the torsion angles in the $\alpha(1 \rightarrow 4)$ glycosidic linkages. α-CD is more sterically strained than its 7- and 8-unit counterparts, and thus cannot adopt as close, favourable hydrogen bonds. The steric strain within the cyclodextrin toroid diminishes in the series $\alpha > \beta > \gamma$. This sequence is reflected in the hydrogen bonding: the average $O2 - O3'$ distances in α-CD are significantly longer than in either β-CD or γ-CD (table 5.6) and the standard deviation is also considerably greater.

From the time series of the the $O2$-$O3'$ distances for each cyclodextrin (figure 5.10) we can see that the hydrogen-bonding distances fluctuate far more in α-CD than in the other two cyclodextrins. Bridging hydrogen bonds with water molecules form more often in order to relieve the steric strain within α-CD.

5.4.4 Pyranose Ring Puckering

CD	Cremer-Pople puckering parameters						gluc. conf.
	Q		θ		ϕ		
	sim.	crys.	sim.	crys.	sim.	crys.	
α	0.559(0.029)	0.574(0.034)	11.96(5.29)	5.5(4.7))	77.915	118.7	$^4C_1$
β	0.562(0.029)	0.579(0.038)	9.23(4.66)	5.0(2.9)	108.67	173.5	$^4C_1$
γ	0.566(0.035)	0.613(0.053)	7.96(4.27)	8.5(4.1)	121.213	231.2	$^4C_1$

Table 5.7: Average Cremer-Pople puckering parameters for the three cyclodextrins from crystal structures and the MD simulations.

The average pyranose ring puckering parameters (defined in section 1.4.1) for each cyclodextrin are listed in table 5.7. Individual puckering parameters for each ring in each cyclodextrin may be found in the appendices (section A.2).

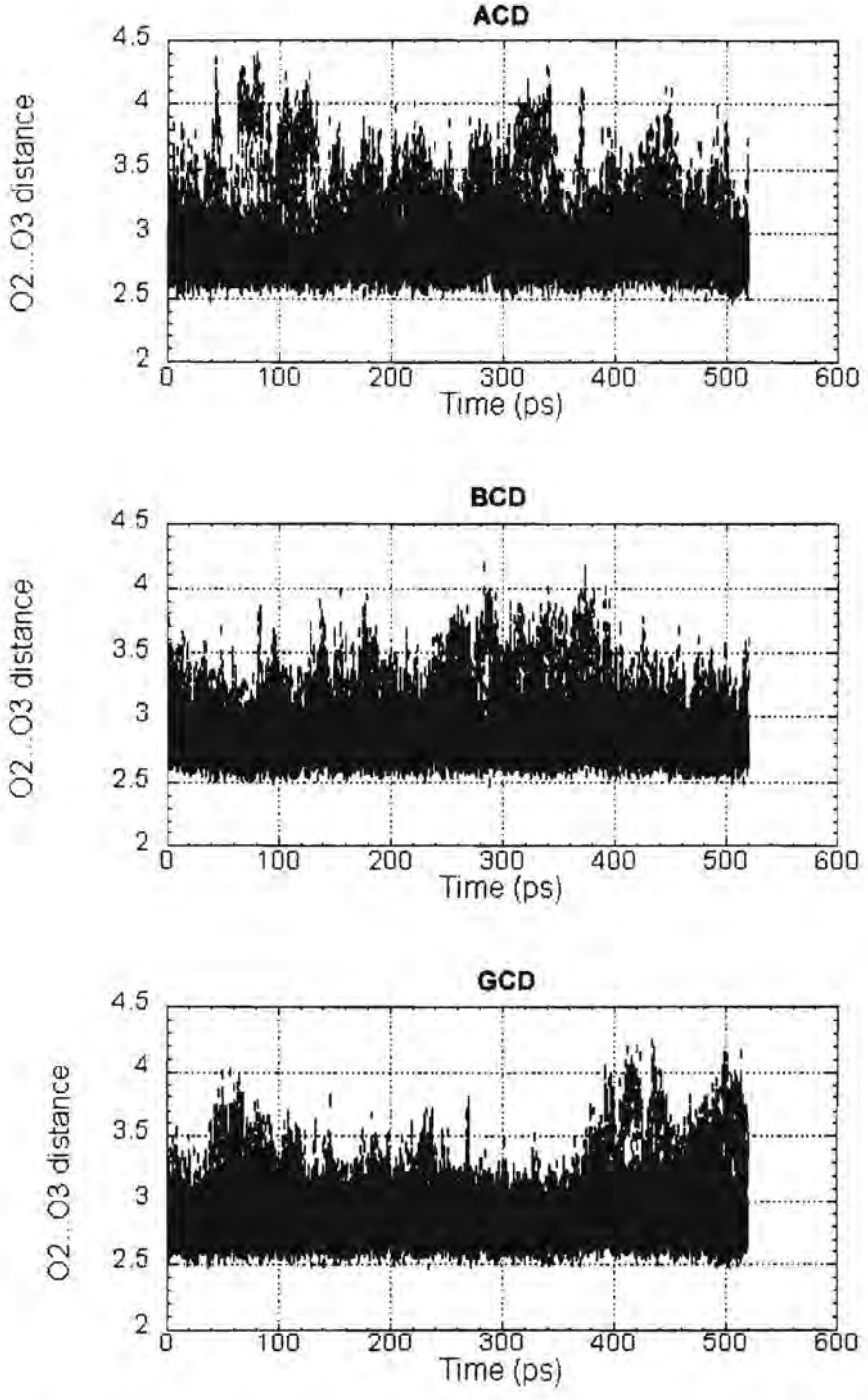

Figure 5.10: Time series for the $O2$-$O3'$ distances in α-, β- and γ-CD.

5.5 Primary Alcohol Conformations

The hydroxymethyl substituents on the glucose residues of the cyclodextrins suffer from the same rotational problem described earlier (sections 2.1.2 and 4.5). Experimentally it has been found that in crystal structures of the cyclodextrins, the *gg* conformation is favoured over the *gt* conformation by a ratio of 3.5:1 [15]. This is the conformation where the 6-*OH* points towards the outside of the ring, as opposed to the inside of the ring in the *gt* conformation. The *tg* conformation is extremely disfavoured: in an extensive analysis of the crystal structures of cyclodextrins available [15], not a single case of a *tg* conformation was encountered. However, in the cyclodextrins, the primary alcohol conformation oscillated between the *gt* and *tg* conformations, with no transitions to the the *gg* conformation occurring in any of the simulations. This once more confirms that in the force field used, the barrier to rotation about the primary alcohol dihedral angle is too high.

5.6 Cluster Analysis

Cluster analysis, as described in section 3.4, was performed on the conformations generated for α-CD and γ-CD during the MD simulations. The cross-ring distances (figure 5.1) were used as the similarity measures, as these give an indication of the extent of deformation of the macrocyclic rings. Thus one is able to group the conformations into circular or varying degrees of ellipsoidal classifications. As is often assumed that cyclodextrins are basically symmetrical circles in aqueous solution, it is instructive to investigate how far this assumption holds in our simulations. Because β-CD has an odd number of glycosidic linkages, the cross-ring distances are less conveniently defined and give a less useful description of the molecular shape. Therefore this simulation was excluded from our analysis. The clustering for both α- and γ-CD was done according to the cross-ring distances using a clustering radius of 0.35. A cutoff of 0.2 $\AA$ was used to fit clusters into classes. Clustering is by its nature an approximation, so some judgement must be used when deciding into which conformational class a cluster falls.

5.6.1 α-cyclodextrin

Thirty clusters were generated in the analysis of α-CD. These were were divided into groups and classes. The classes were defined, as in figure 5.11, into circular, ellipsoidal and deformed classes.

Conformations in the circular class (C) have the cross-ring distances essentially the same length. Ellipsoidal classes (E) have two distance parameters the same length, and one different. They are subdivided into two sub-classes; one with one length shorter than the other two (E_1) and one with one length longer than the other two (E_2). The deformed class has all the cross-ring distances of different lengths.

The classes were then subdivided into groups of the same shape. There are three distinct length measurements in a cluster and, as the axis chosen first is arbitrary, a group of different clusters will describe the same shape.

Clust. no.	Group	% sim. time	Cross-ring distances (Å)			Class.	Group time %
			O_{11}-O_{14}	O_{12}-O_{15}	O_{13}-O_{16}		
1	1	4.21	8.5	8.4	8.5	C	4.21
2	2	3.46	8.3	8.4	8.3	C	3.46
3	3	2.24	8.7	8.4	8.3	E_2	
4		5.04	8.3	8.3	8.6		
5		3.56	8.3	8.8	8.3		8.84
6	4	1.02	8.1	8.0	9.0	E_2	
7		2.54	8.1	9.0	8.1		3.56
8	5	3.45	8.6	8.2	8.4	E_2	
9		4.05	8.4	8.2	8.8		7.50
10	6	4.51	8.1	8.3	8.7	E_2	
11		4.00	8.1	8.7	8.3		
12		2.66	8.3	8.7	8.1		11.17
13	7	4.19	8.1	8.4	8.5	E_1	4.19
14	8	1.93	8.6	8.7	8.0	E_1	
15		5.05	8.0	8.6	8.6		
16		3.17	8.5	8.0	8.6		10.15
17	9	3.12	8.5	8.6	8.3	E_1	
18		3.53	8.6	8.2	8.6		6.65
19	10	2.61	8.9	7.7	8.7	E_1	
20		2.86	7.7	8.8	8.7		
21		3.57	8.7	7.9	8.7		9.04
22	11	4.67	8.2	8.5	8.7	E_1	
23		4.69	8.3	8.7	8.5		9.36
24	12	2.75	7.8	9.1	8.3	D	
25		3.45	7.8	8.8	8.5		
26		2.58	8.5	7.8	8.9		8.78
27	13	2.90	8.9	8.4	8.0	D	
28		3.20	8.8	8.0	8.5		
29		3.41	8.1	8.9	8.5		
30		1.08	8.0	8.4	8.9		10.59

Table 5.8: Clusters generated for α-CD cross-ring distances.

Conformation class	%time
C	7.76
E_1	39.39
E_2	33.07
D	19.73

Table 5.9: % time spent in each conformational class for the α-CD simulation.

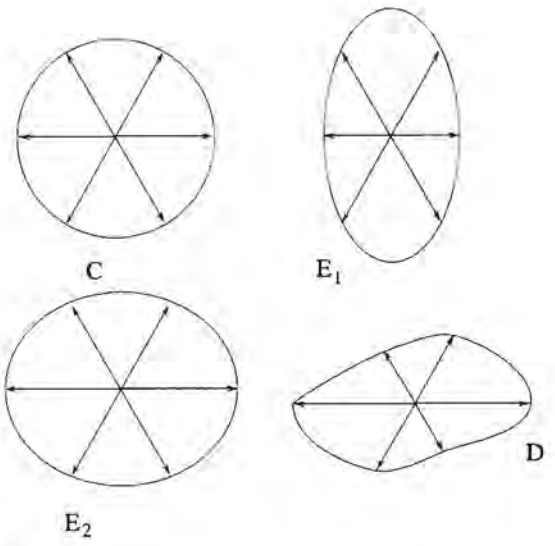

Figure 5.11: Conformation classes as defined for α-CD.

Table 5.9 gives a summary of the total simulation time spent in each of the different conformational classes. By far the majority of the simulation time was spent with cyclodextrin in an ellipsoidal conformation.

5.6.2 γ-cyclodextrin

Twenty-two clusters were generated for γ-CD, using the same clustering radius of 0.35 as was used for α-CD. As evidence from the analysis of the ϕ,ψ angle distribution shows, γ-CD exhibits less motion in solution than α-CD. Thus it is to be expected that fewer clusters will be generated for this molecule.

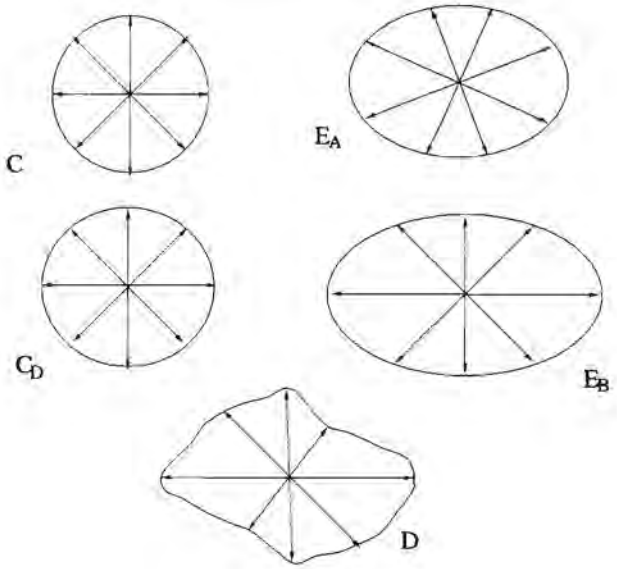

Figure 5.12: Conformation classes as defined for γ-CD.

As γ-CD has four different cross-ring distances, the division into classes is slightly different to α-CD. The classes for this molecule were defined, as in figure 5.6.2, into circular (all lengths the same), slightly deformed circular (one length slightly different to the other three), ellipsoidal A (two pairs of distances of the same length), ellipsodal B and deformed classes (all lengths different). E_B is a more elongated ellipse than E_A.

Clust. no.	Group	% sim. time	Cross-ring distances ($\AA$)				Class.	Group time %
			O_{11}-O_{15}	O_{12}-O_{16}	O_{13}-O_{17}	O_{14}-O_{18}		
1	1	5.58	11.6	11.6	11.5	11.5	C	5.58
2	2	6.96	11.6	11.3	11.2	11.2	C_D	
3		4.07	11.7	11.2	10.8	11.0		
4		3.82	11.5	10.9	11.2	11.4		
5		4.90	11.2	11.3	11.6	10.9		19.49
6	3	7.11	11.3	11.1	11.5	11.4	C_D	7.11
7	4	4.58	11.1	11.1	11.4	11.0	C_D	
8		5.56	11.4	11.0	11.1	11.0		10.14
9	5	5.46	11.4	11.3	11.2	10.7	C_D	5.46
10	6	5.42	11.9	11.9	11.0	11.0	E_A	
11	.	3.14	11.9	10.8	11.0	12.2		8.56
12	7	7.43	11.3	11.6	11.5	11.2	E_A	7.43
13	8	6.85	11.8	11.6	11.1	11.2	E_A	
14		9.24	11.5	11.7	11.3	11.1		
15		5.06	11.8	11.2	10.9	11.5		
16		4.13	11.2	11.3	11.8	11.6		
17		6.51	11.6	11.8	11.1	10.8		31.79
18	9	5.00	11.5	11.9	11.4	10.9	E_B	
19		4.44	11.5	11.1	11.4	11.8		9.44
20	11	2.60	12.1	11.5	10.7	11.3	E_B	2.60
21	12	4.96	11.8	12.1	11.0	10.6	D	4.96
22	13	3.50	11.1	11.5	11.4	10.5	D	3.50

Table 5.10: Clusters generated for γ-CD cross-ring distances.

Conformation class	%time
C and C_D	47.78
E_A and E_B	51.26
D	8.96

Table 5.11: % time spent in each conformational class for the γ-CD simulation.

Table 5.11 gives a summary of the total simulation time spent in each of the different conformational classes. Compared to α-CD, γ-CD spends a large amount of time in circular, or nearly circular conformations. This is corroborated by the crystal structures of γ-CD, which are in general perfectly rounded and represent by far the

most symmetrical structures encountered in the entire series of crystal structures of cyclodextrin hydrates [15].

5.7 Water Structuring About the Cyclodextrins

Water density analysis was performed on each of the cyclodextrins as described in section 3.1.2. All the ϕ, ψ angles in each cyclodextrin were specified as selection criteria for the water density analysis. The angles were chosen so as to select angles in the middle of the ϕ,ψ scatter plots with ranges of $\pm 15°$. The actual angles used are list in table 5.12.

Torsion angle	Cyclodextrin		
	α-CD	β-CD	γ-CD
$\phi 1$	0 ± 15	$-15(\pm 15)$	$25(\pm 15)$
$\phi 2$	$-5(\pm 15)$	$-5(\pm 15)$	$-15(\pm 15)$
$\phi 3$	$0(\pm 15)$	$15(\pm 15)$	$25(\pm 15)$
$\phi 4$	$-5(\pm 15)$	$-5(\pm 15)$	$-15(\pm 15)$
$\phi 5$	$-10(\pm 15)$	$-15(\pm 15)$	$-15(\pm 15)$
$\phi 6$	$-5(\pm 15)$	$15(\pm 15)$	$25(\pm 15)$
$\phi 7$	-	$-5(\pm 15)$	$-10(\pm 15)$
$\phi 8$	-	-	$-15(\pm 15)$
$\psi 1$	$0(\pm 15)$	$-15(\pm 15)$	$15(\pm 15)$
$\psi 2$	$0(\pm 15)$	0 ± 15	$-20(\pm 15)$
$\psi 3$	$0(\pm 15)$	$15(\pm 15)$	$15(\pm 15)$
$\psi 4$	$-5(\pm 15)$	0 ± 15	$-20(\pm 15)$
$\psi 5$	$10(\pm 15)$	$-15(\pm 15)$	$-15(\pm 15)$
$\psi 6$	$-5(\pm 15)$	$15(\pm 15)$	$15(\pm 15)$
$\psi 7$	-	0 ± 15	$-10(\pm 15)$
$\psi 8$			$-15(\pm 15)$

Table 5.12: Angles used to define the conformations used for water density analysis.

In order to ensure that each frame is sufficiently independent from its neighbours, we stipulated that dynamics frames selected for analysis had to be at least 1.5 ps apart. The resultant three-dimensional water density surfaces for each cyclodextrin,contoured at 65% above bulk density, are shown in figure 5.13. The dark grey rings are the cyclodextrins and the light grey areas the water surfaces. The solubilities of the three cyclodextrins are listed in table 5.1. α-CD and γ-CD are considerably more soluble than β-CD. The water structuring reflects this order, in that β-CD shows considerably more areas of high water density further away from the molecule. It is tempting to conclude that β-CD is less soluble than the other two cyclodextrins because of the increased structuring it imposes on the surrounding solvent. However, because of the sensitivity of this analysis method to the mobility of the solutes, a comparison of the

3D water structuring surfaces for the three cyclodextrins is by no means conclusive.

The central cavity of the cyclodextrins is the most hydrophobic region of the molecule. Thus, hydrophobic interactions are the main driving force behind guest complexation in the cyclodextrins. However, in aqueous solution, with the absence of less polar guests, water occupies the central cavity. It has been suggested that 2 to 3 water molecules occupy the central cavity in α-CD, about 5 in β-CD and 8 to 9 in γ-CD [15].

Figure 5.14 shows the water density structuring in cross-sections through the center of the cyclodextrin rings. The first hydration shell of water is clearly visible as a ring of water around the cyclodextrin toruses. α-CD shows simply a central sphere of water in the cavity. β-CD exhibits a rather more disordered cluster of water molecules. γ-CD exhibits a ring of water molecules within the central cavity. Of interest is that the cavity in β-CD appears to be more ellipsoidal than either α or γ-CD. This may indicate that β-CD exhibits a more ellipsoidal configuration in solution. Figure 5.15 shows water structuring in cross-sections at right angles to the cyclodextrin rings. Rings of high water density passing through the center of the cyclodextrin torus can be seen.

Lichtenthaler et al. performed a computational study analysing the molecular lipophilicity patterns of the cyclodextrins, using solid-state data [15]. This study revealed the $O2$,$O3$ side of the macrocycles (the wider rim of the torus) to be distinctly hydrophilic, while the opposite, narrower opening made up of the 6-CH_2OH groups was shown to be intensely hydrophobic. However, to our knowledge, no computational analysis of the water structuring around the two rims of the cyclodextrin has been performed. Of interest is the effect of the hydrophobic and hydrophilic regions of the cyclodextrins on the average density and structure of surrounding solvent water.

Figure 5.16 illustrates the different water densities on around the two rims of the cyclodextrins. It is clear from this picture that the 02,03 rim of this cyclodextrin (the wider rim) is considerably more hydrophilic than the opposite rim. This contrast is most pronounced in γ-CD and least in α-CD.

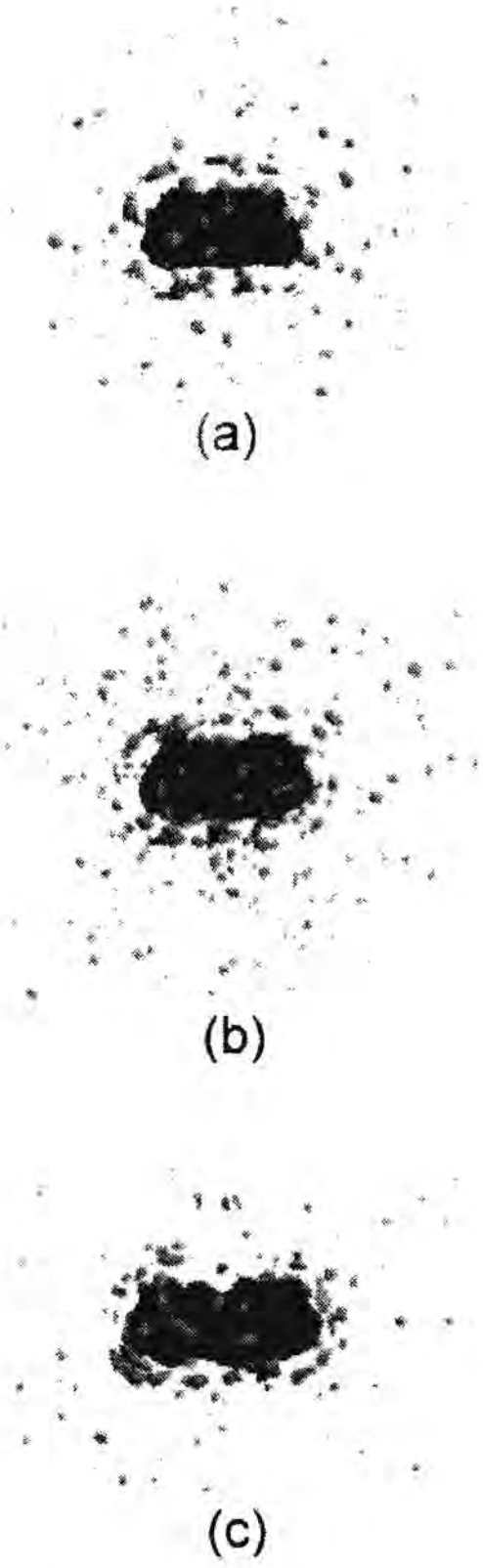

Figure 5.13: Water density surfaces contoured at 65% above bulk density for **(a)** α- , **(b)** β- and **(c)** γ-CD .

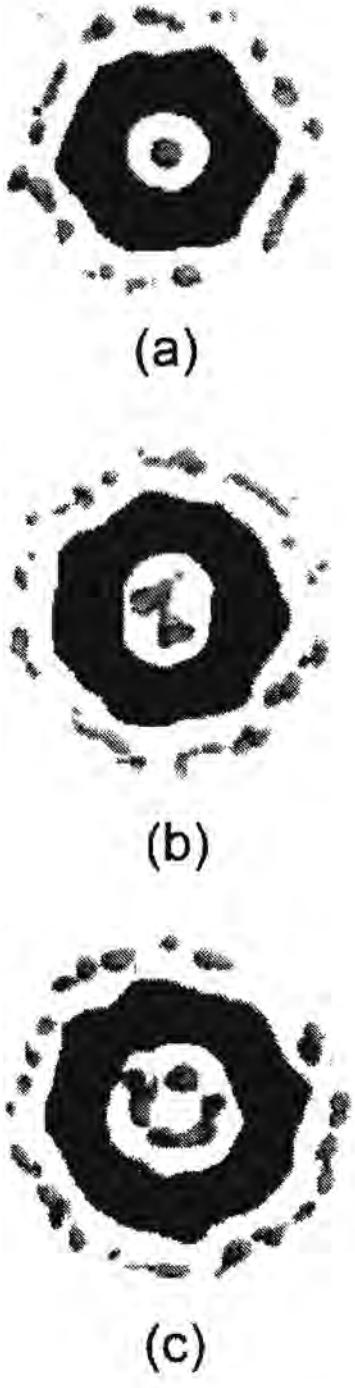

Figure 5.14: Cross-sections through the center of the cyclodextrin cavities, showing water molecules within the cavity. The water areas are contoured at 50% above bulk density. **(a)** α-CD, **(b)** β-CD, **(c)** γ-CD

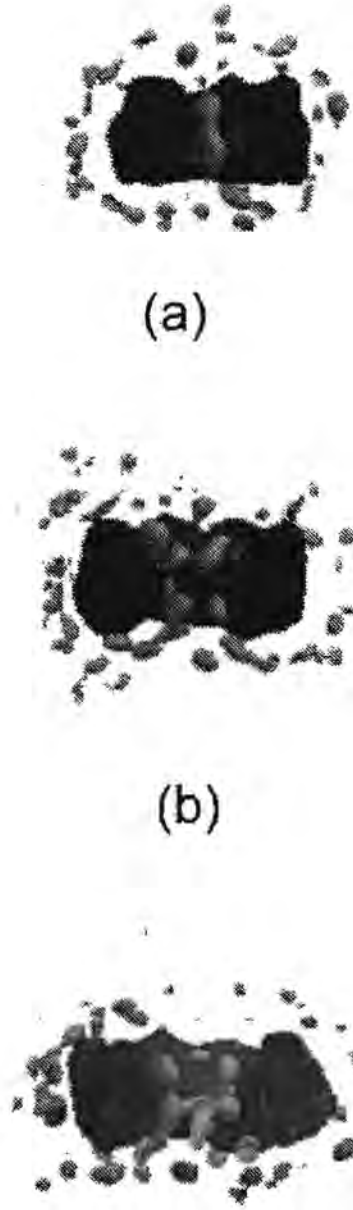

Figure 5.15: Cross-sections at right angles to the cyclodextrin cavities, showing water molecules within the cavity. The water areas are contoured at 60% above bulk density. **(a)** α-CD, **(b)** β-CD, **(c)** γ-CD

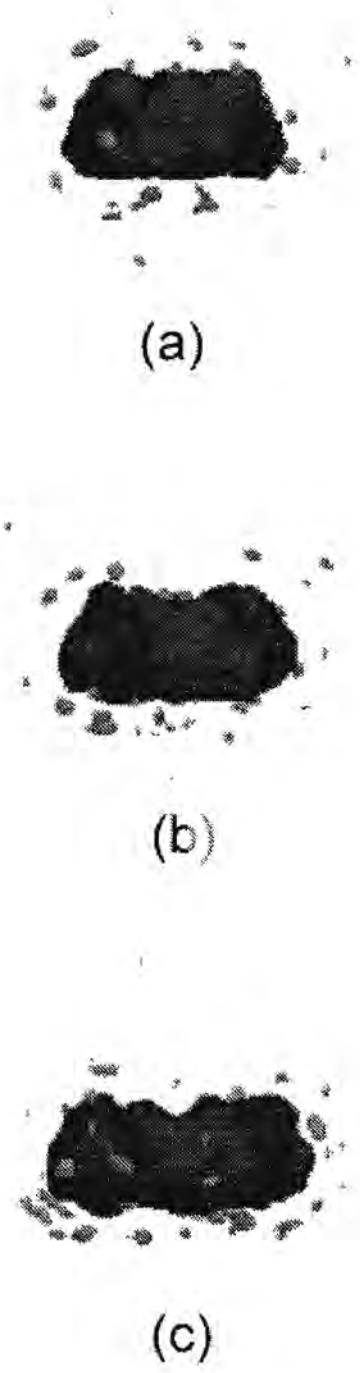

Figure 5.16: Cyclodextrins viewed from the side with the $O3$,$O3$ side facing downwards. The water areas are contoured at 70% above bulk density.

Chapter 6

Close Contacts Analysis of α and β-Cyclodextrin in the Solid State.

Non-bonded interactions dominate in deciding the relative stabilities of two inclusion complexes. Thus, non-bonded interaction matrix analysis (described in section 3.2.1), which shows a matrix of the non-bonded interactions between selected atoms, can potentially be useful in analysing the placements of guests within host compounds.

The cyclodextrins are chiral molecules and have the potential for use as resolving agents for racemic mixtures of chiral guest molecules. If a 5-10 fold excess of a cyclodextrin solution is mixed with a racemate, one of the cyclodextrin-guest complexes precipitates out preferentially, resulting in a partial racemate resolution [85]. The resolving potential of cyclodextrins and their derivatives has been exploited in the use of cyclodextrins as chiral stationary phases in capillary columns for gas chromatography [17].

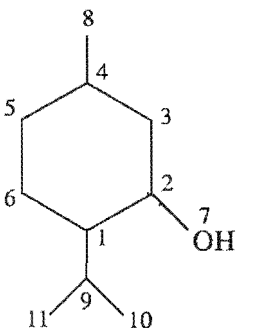

Figure 6.1: Numbering scheme for menthol.

Menthol ((5-methyl-2-(1-methyl-ethyl)-cyclohexanol) (figure 6.1) is a molecule with three chiral centers and therefore eight stereo-isomers. (L)-menthol (1R,2S,5R) is used externally as an analgesic in rheumatism and by inhalation in the alleviation of nasal congestion. (L)-menthol has previously been shown to form inclusion compounds with β-cyclodextrin [86] and a patent has been taken out on the menthol-β-cyclodextrin inclusion complex for use as an inhalant in the treatment of respiratory ailments [87].

Menthol has historically been produced from *Mentha arvensis* (Japanese mint) oil by cooling and separating the crystals by centrifugation. This is an inherently unpredictable source and usually insufficient to meet the world demand for menthol. This has led to the development of many synthetic and semi-synthetic methods. The disadvantage of synthetic methods is that they invariably produce racemic products. Thus, it would be advantageous if β-cyclodextrin formed an inclusion complex preferentially with one of the menthol isomers.

We used two cyclodextrin-menthol inclusion complexes as a convenient model on which to test the NIPMAT analysis method.

6.1 Crystal structure of the β-cyclodextrin-(D)-menthol inclusion complex

As a first step in investigating the potential of β-cyclodextrin as a resolving agent for menthol (and as an instructive exercise) we solved the crystal structure of the β-cyclodextrin-(D)-menthol inclusion complex. The crystal structure of the β-cyclodextrin-(L)-menthol host has been solved previously [86]. Crystals were obtained by slow evaporation of a 1:1 solution of β-cyclodextrin and menthol in water. A crystal was mounted on a glass fiber. X-ray diffraction data were obtained on a Enraf-Nonius CAD4 diffractometer using graphite-monochromated MoK_α radiation ($\lambda = 0.7107\mathring{A}$) in ω-2θ mode. The data collection was at 278K, in order to reduce the thermal motion within the structure. Three reference reflections were monitored periodically to check crystal orientation and stability. The data reduction included correction for Lorentz and polarization effects, but not for absorption.

The structure was solved with the SHELXL97 program [88], using isomorphous replacement with the previously-solved β-cyclodextrin-(L)-menthol crystal structure [86] , and refined to an R1-value of 0.1068. The final model included anisotropic refinement of all non-hydrogen atoms in the cyclodextrins. Hydrogen atoms of the glucose units were placed in geometrically calculated positions with common temperature factors. The hydroxyl hydrogens were not located. The menthol molecules display very high thermal motion and so were refined isotropically. Distance restraints were applied to some bond lengths. Thermogravimetric analysis of the (D)-menthol - β-cyclodextrin inclusion complex showed a weight loss corresponding to approximately 27 waters. These water molecules are disordered across 32 sites. Crystal data, experimental and refinement parameters for the solved crystal structure are listed in table 6.1. Supplementary data appears in appendix B.

The (L)-menthol (1R,2S,5R) and (D)-menthol (1S,2R,5S) complexes are similar, as can be seen from table 6.2. Both crystallise in the $P2_1$ space group, with two cyclodextrin and two menthol molecules in the asymmetric unit. There is a host:guest ratio of 1:1 in both complexes. As in the majority of β-cyclodextrin complexes, the two menthol complexes are dimeric. Two cyclodextrin molecules form a head-to-head unit stabilised by hydrogen bonding between the secondary hydroxyl groups. The menthol molecules in the central cavities have a head-to-tail arrangement and are

Identification code	CDMENT
Molecular formula	$C_{42}H_{70}O_{35}.C_{10}H_{20}O.13.5H_2O$
M_r (g.mol^{-1})	1534.69
Crystal system	Monoclinic
Space group	$P2_1$
Z	4
Unit cell dimensions	a = 15.394(3)Å α = 90 deg
	b = 32.699(9)Å β = 102.36(2) deg
	c = 15.450(4)Å γ = 90 deg
Volume	7597(3) Å^3
Density (calculated)	1.342 Mg/m^3
Absorption coefficient	0.120 mm^{-1}
F(000)	3300
Crystal size	0.44 x 0.44 x 0.41 mm
Temperature of data collection	278(2) K
Wavelength	0.71070 Å
Theta range for data collection	1.25 to 24.97 deg
Index ranges	$-18 <= h <= 17$, $0 <= k <= 38$, $0 <= l <= 18$
Reflections collected / unique	13588 / 13588
Completeness to 2theta = 24.97	99.9%
Independent reflections	13588 [R(int) = 0.0000]
Refinement method	Full-matrix-block least-squares on F^2
Data / restraints / parameters	13588 / 1 / 1607
Goodness-of-fit on F^2	0.950
Final R indices $[I > 2\sigma(I)]$	R1 = 0.1068, wR2 = 0.2726
R indices (all data)	R1 = 0.1497, wR2 = 0.3063
Absolute structure parameter	0.0(19)
Largest diff. peak and hole	0.951 and -0.431 e.Å^{-3}
Scan type	$\omega - 2\theta$
Scan width	$0.75 + 0.35tan\theta$
Apeture width (mm)	$1.12 + 0.5tan\theta$
Decay	< 5%

Table 6.1: Crystal data, experimental and refinement parameters for CDMENT (structure refined using SHELXL97)

	(D)-menthol complex	(L)-menthol complex
Temperature	278 K	278 K
Space group	$P2_1$	$P2_1$
Z	4	4
a	15.394	15.342
b	32.699	32.54
c	15.450	15.324
β	102.36	102.44

Table 6.2: Comparisons of the (D)-menthol- and (L)-menthol β-cyclodextrin complexes

crystallographically distinct (figure 6.2).

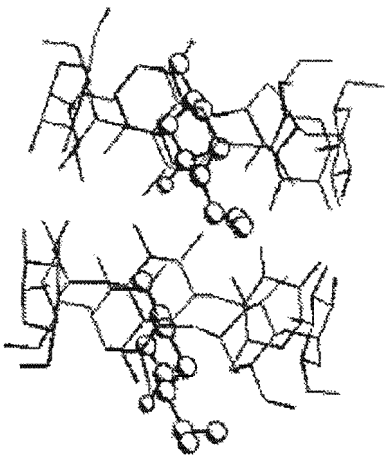

Figure 6.2: The two crystallographically distinct molecules in the (D)-menthol-β-CD inclusion complex, comprising the dimer.

These dimers pack in the unit cell along a screw axis parallel to the b axis. This is shown in figure 6.3. Figure 6.4 shows the top view of the dimer.

Equivalent analysis of the (L)-menthol complex has shown it to contain 29 waters. There is, as is usual for cyclodextrin complexes, extensive hydrogen-bonding between the cyclodextrin hosts and the water molecules that fill the inter-molecular spaces.

Figure 6.1 shows the numbering scheme used for the menthol guest. The same numbering scheme for the guest as appears in [86] was used, in order to facilitate comparison between the two menthol structures.

The glucose moieties of the two crystallographically distinct β-CD molecules were identified using a three digit numbering scheme. The first digit represents the cy-

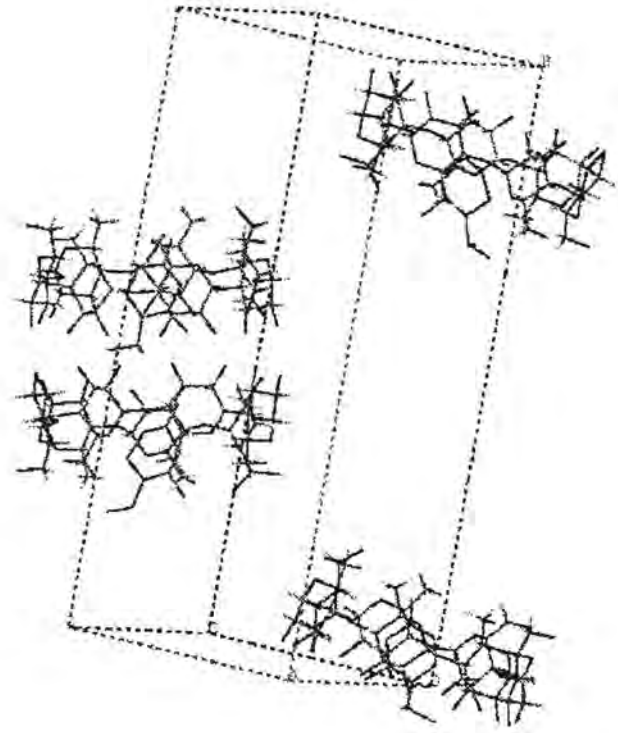

Figure 6.3: Packing diagram of the crystal structure. The long axis is (010).

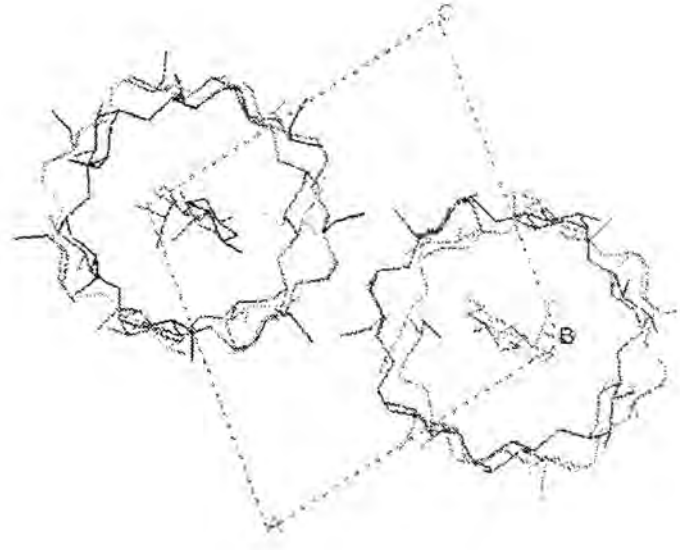

Figure 6.4: Section through the unit cell down [010], showing the top view of the dimer.

clodextrin in question, the second digit the residue number and the third digit the atom number. Thus, C213 is carbon number three in the first residue of the second cyclodextrin in the dimer. There is a slight discrepancy between this numbering scheme and the terminology used in the previous chapter. There the glycosidic linkage oxygens were termed $O1$, whereas in this numbering scheme the linkage oxygens are termed $O4$.

6.2 Comparison of the cyclodextrin-menthol complexes using NIPMAT analysis.

Various non-bonded interaction matrices for the (L)-menthol and (D)-menthol complexes were prepared using atomic radii of C = 1.75 $\AA$, O = 1.40 $\AA$ and H = 1.0 $\AA$. These maps are of assiatnce in clarifying the similarities and differences between the two menthol-cyclodextrin inclusion complexes.

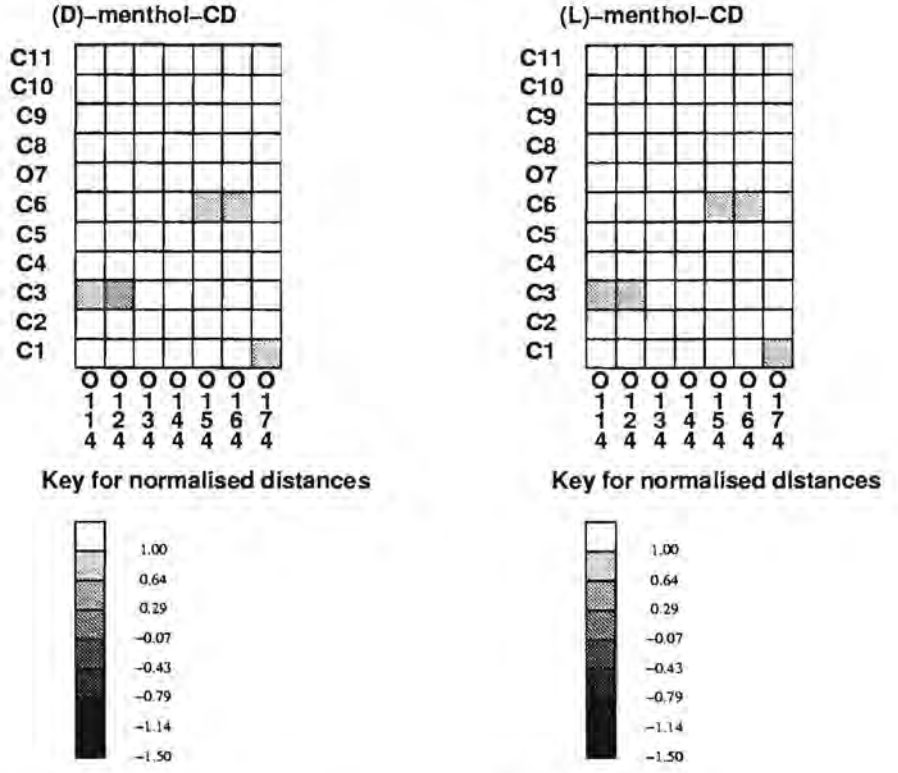

Figure 6.5: Non-bonded interaction matrices for the (D)- and (L)-menthol-cyclodextrin inclusion complexes, showing the non-bonded interactions between the glycosidic linkage oxygen atoms (O114, O124 etc.) and the menthol atoms for the first cyclodextrin.

Figures 6.5 and 6.6 show matrices of the non-bonded interactions between each of the two crystallographically distinct menthol guests with the glycosidic linkage oxygens, 04, of the respective cyclodextrin host molecules for both the menthol inclusion compounds. Darker grey blocks correspond to closer non-bonded interactions. In this

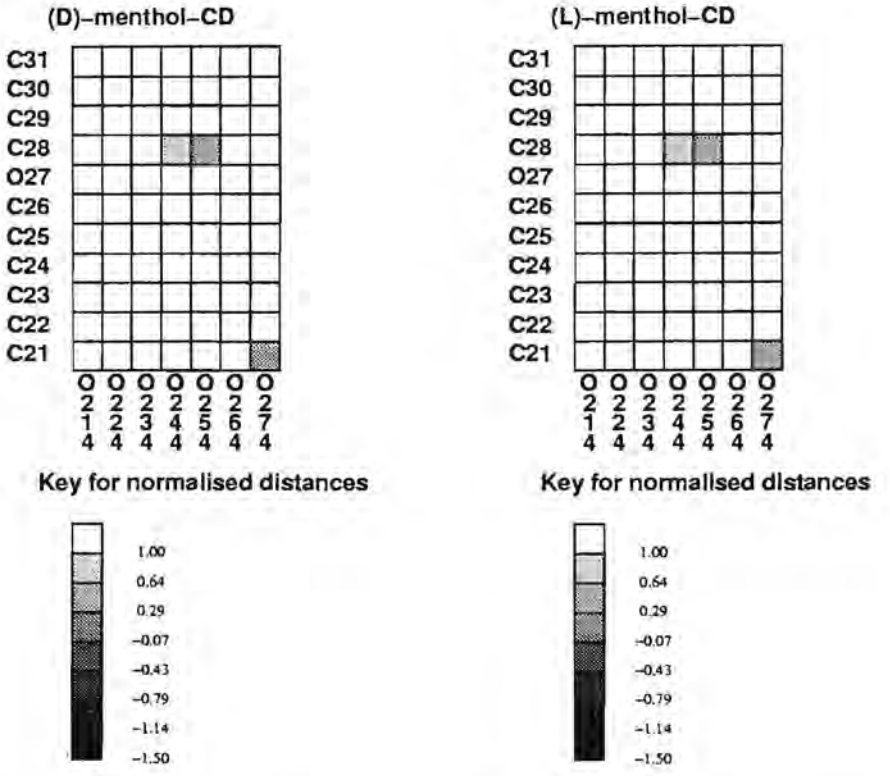

Figure 6.6: Comparison maps for the (D)- and (L)-menthol-cyclodextrin inclusion complexes, showing the non-bonded interactions between the ring oxygens and the guest for the second cyclodextrin.

case, NIPMAT analysis highlights not so much general packing features as the situation of the two menthol guests within the host. It is clear from these figures that, when comparing the inclusion compounds of (L)- and (D)- menthol, the corresponding menthol molecules are oriented very similarly within the cyclodextrin central cavity. From this one can conclude that it is unlikely that β-cyclodextrin discriminates between the two menthol isomers and thus will probably not be of use as a a resolving agent in this case.

The circle of $O2$-$O3$ hydrogen bonds is the primary stabilizing force in the cyclodextrin torus. Anomalies in the hydrogen-bonding pattern are clearly shown by NIPMAT analysis. Figure 6.7 shows two symmetric NIPMAT matrices of the non-bonded interactions between the $O2$ and $O3$ substituents for the second cyclodextrin in the (L)- and (D)-menthol complexes respectively. The (D)-menthol complex shows a regular bonding pattern, while an obvious anomaly in the h-bonding pattern of the second cyclodextrin in the L-menthol complex occurs at the third glucose residue. Here the $O3$ hydroxyl does not make the usual bond with the $O2$ hydroxyl on the same residue, but is rather oriented so that is makes hydrogen bonds with both the $O2$ and $O3$ hydroxyls on the fourth residue.

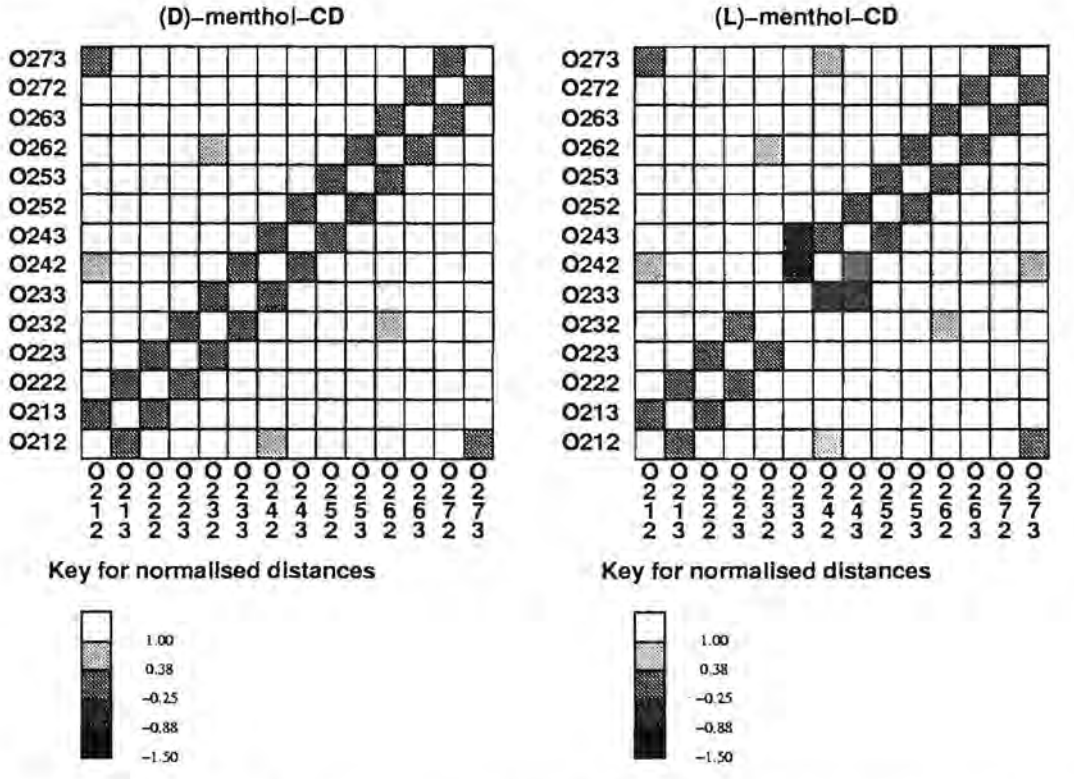

Figure 6.7: Symmetric NIPMAT matrices for the second cyclodextrin in the two complexes

This anomaly has a corresponding effect on the hydrogen-bonding between the 02 and $O3$ hydroxyls at the interface between the two crystallographically distinct

cyclodextrins, as can be seen in figure 6.8. Once more, the (D)-menthol show a regular hydrogen-bonding pattern, while a disruption occurs at the third glucose residue on the second cyclodextrin in the (L)-menthol complex. The $O3$ hydroxyl exhibits only one rather distant contact with an $O3$ hydroxyl on the opposite cyclodextrin.

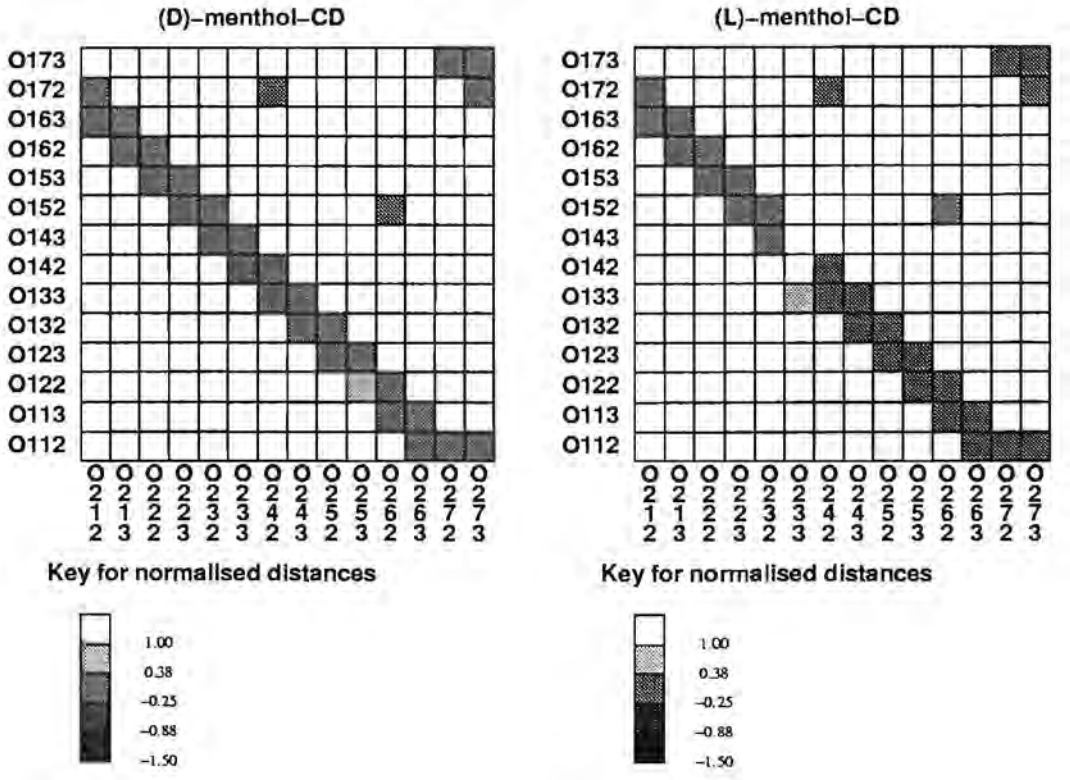

Figure 6.8: Non-symmetric NIPMAT matrices for the interaction between the O2 and O3 hydroxyls in the two crystallographically distinct cyclodextrins in the two complexes.

6.3 Non-bonded Interaction Matrix comparison of the "empty" structures of α- and β-CD.

As a further example of the use of NIPMAT analysis, hydrate crystal structures of α-CD and β-CD were analysed. The crystal structures of β-cyclodextrin dodecahydrate clathrate [89] and α-cyclodextrin hexahydrate [90] (form I) were selected from the Cambridge Crystallographic Database [91].

Figure 6.9 shows symmetric NIPMAT matrices of the non-bonded interactions between the $O2$ and $O3$ substituents in the α- and β-CD hydrates. These maps clearly illustrate that the $O2$-$O3$ hydrogen-bonding network in α-CD is far more disrupted than in β-CD. This is a result of the increased strain in the six-member α-cyclodextrin as compared to β-cyclodextrin.

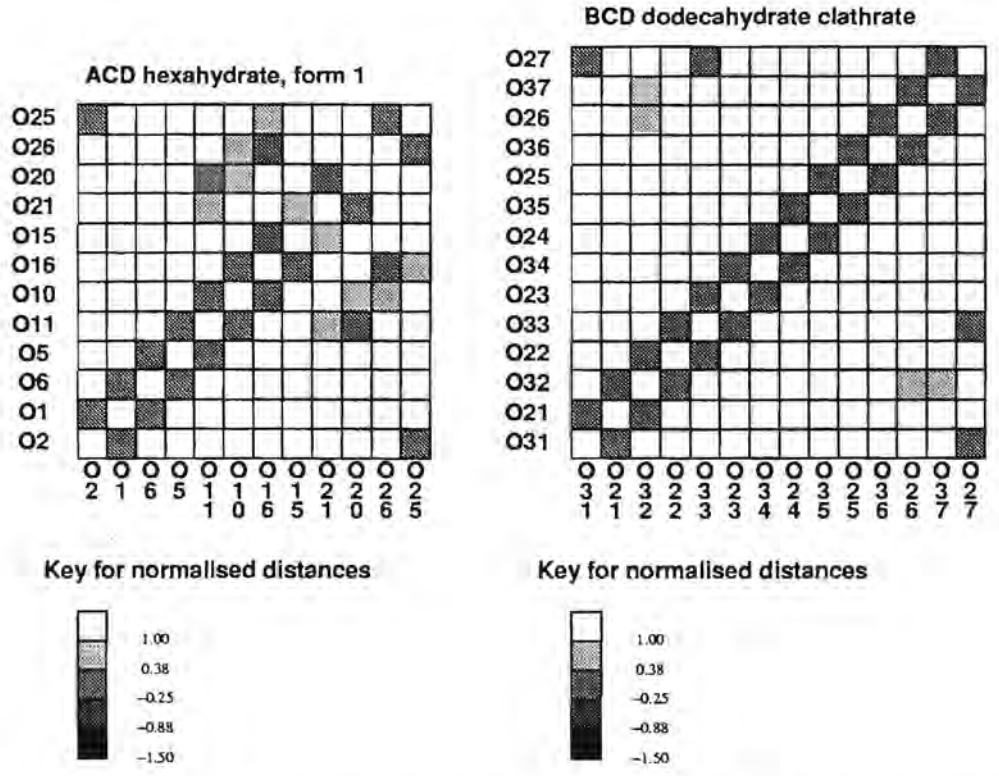

Figure 6.9: Symmetric matrices of the non-bonded interactions between the $O2..O3$ oxygens for α-cyclodextrin hexahydrate and β-cyclodextrin dodecahydrate clathrate

Chapter 7

Conclusions and future work

The primary objective of this work was to develop tools for analysing saccharides in solution and condensed phases. This objective has been achieved and some interesting results have been obtained.

Our results for maltose in water show a preference for a broader region of ϕ,ψ conformational space when compared to simulations in vacuum. Hydrogen-bonding analysis of maltose showed that the increased area of conformational space sampled in solution is largely a result of the internal $O2$-$O3$ hydrogen bond in maltose being replaced by hydrogen bonds with the solution water molecules.

Water-density studies revealed that maltose imposes anisotropic structuring on the surrounding water molecules. This was corroborated by Vornoi analysis of the maltose simulations, which demonstrated that the water molecules closer to maltose have a larger number of nearest neighbours compared to those further away.

The practice of studying maltose as a model for larger polysaccharides was supported, as the motions of the hexa-amylose strand about its glycosidic linkages were found to be very similar to those of maltose. If this result is extrapolated to the amylose polymer, the lowering of the T_g for starch in the presence of water can be explained by increased flexibility around the glycosidic linkage in aqueous solution.

The cyclodextrins exhibited decreasing mobility in solution in the order $\alpha < \beta < \gamma$ -cyclodextrin. The areas of ϕ,ψ space explored by the various dihedral angles in the cyclodextrins were shown to be a superset of those explored by maltose.

The cyclodextrins were shown via water density analysis to structure the surrounding water. A first hydration shell around the macrocycle was clearly visible in all cases. The primary hydroxyl rim of the cyclodextrins was show to exhibit less water density than the opposite $O2$,$O3$ rim. This is in agreement with the general designation of the $O2$,$O3$ rim to be the "hydrophilic" rim and the CH_2OH rim the "hydrophobic" rim. Interestingly, the effect was most marked in γ-cyclodextrin.

Clustering analysis of α-cyclodextrins predicted that this molecule spends the majority of the time in solution in an ellipsoidal conformation. In contrast, the equivalent analysis of γ-cyclodextrin showed that it had a more spherical shape.

The NIPMAT analysis tool was demonstrated to be particularly useful as a quick method for probing the hydrogen-bonding networks in cyclodextrin crystal structures.

Rotations about the primary alcohol dihedral angle were found to be restricted in all the simulations performed. Although the *gg:gt* ratio in solution is still a matter of some debate, transitions between these two conformations should be observed during a simulation. Therefore, the re-parametrization of this dihedral should be addressed before any further work is undertaken using this force field.

The analytical methods developed in this thesis are shown to complement experimental methods, as well as provide information that is not accessible experimentally. The tools of Vornoi analysis and, particularly, the water density analysis tool are shown to be useful for investigating water structure about carbohydrates. However, a water structuring investigation is incomplete without an investigation of the free energy surface of saccharides in solution. This is an extremely complex problem and could be the subject of a larger, more extensive study.

Appendices

Appendix A

Tables of Cyclodextrin Values

A.1 Torsion Angles and Atomic Distances

Cross-ring distance $O_n \rightarrow O_{n+3}$	value
$1 \rightarrow 4$	8.297(0.306)
$2 \rightarrow 5$	8.432(0.341)
$3 \rightarrow 6$	8.519(0.240)

Table A.1: Average of each of the cross-ring distances for α-CD.

Cross-ring distance $O_n \rightarrow O_{n+4}$	value
$1 \rightarrow 5$	11.577(0.261)
$2 \rightarrow 6$	11.477(0.355)
$3 \rightarrow 7$	11.241(0.266)
$4 \rightarrow 8$	11.209(0.363)

Table A.2: Average of each of the cross-ring distances for γ-CD.

A.2 Additional puckering parameters

Glucose unit	α-CD	β-CD	γ-CD
1	2.878(0.193)	3.O37(0.317)	2.819(0.138)
2	2.943(0.217)	2.906(0.196)	3.027(0.304)
3	2.921(0.215)	2.879(0.183)	2.805(0.138)
4	2.887(0.186)	2.910(0.178)	2.942(0.218)
5	3.126(0.359)	2.898(0.167)	2.922(0.179)
6	3.016(0.318)	2.858(0.167)	2.826(0.138)
7		2.858(0.162)	2.850(0.139)
8			3.026(0.287)

Table A.3: Average $O2..O3$ distances for each glucose ring.

Glucose unit	α-CD	β-CD	γ-CD
1	4.929(0.401)	4.593(0.318)	5.769(0.333)
2	4.731(0.322)	4.672(0.288)	4.534(0.263)
3	4.969(0.416)	5.365(0.419)	5.739(0.412)
4	4.766(0.364)	4.584(0.245)	4.524(0.273)
5	4.684(0.331)	4.525(0.250)	4.487(0.258)
6	4.822(0.408)	5.396(0.535)	5.889(0.319)
7		4.863(0.332)	4.560(0.272)
8			4.465(0.233)

Table A.4: Average $C6..C6$ distances for each glucose ring.

Torsion angle	Cyclodextrin		
	α-CD	β-CD	γ-CD
$\phi 1$	1.200(11.729)	-11.494(12.986)	28.528(10.914)
$\phi 2$	-5.401(9.949)	-5.626(9.317)	-16.287(9.533)
$\phi 3$	1.496(12.507)	14.875(11.837)	28.844(12.873)
$\phi 4$	-4.585(11.067)	-8.053(8.179)	-12.608(10.201)
$\phi 5$	-8.054(11.648)	-11.526(7.583))	-12.770(8.036)
$\phi 6$	-3.736(12.601)	16.144(15.014)	33.040(10.253)
$\phi 7$	-	0.157(10.784)	-7.794(9.431)
$\phi 8$		- -	-15.524(9.294)
$\psi 1$	3.077(12.817)	-9.614(13.146)	14.246(7.019)
$\psi 2$	0.113(12.951)	-1.361(16.053)	-19.705(9.470)
$\psi 3$	0.315(10.657)	15.931(8.020)	12.497(7.694)
$\psi 4$	-5-540(13.731)	-4.532(12.679)	-13.259(12.322)
$\psi 5$	9.709(16.595)	-12.548(10.532)	-11.480(9.123)
$\psi 6$	5.358(13.993)	13.774(8.075)	13.688(6.861)
$\psi 7$	-	1.307(11.784)	-5.713(11.855)
$\psi 8$	-	-	-14.368(11.231)

Table A.5: Averages of the inter-saccharide torsion angles.

Torsion angle	Cyclodextrin		
	α-CD	β-CD	γ-CD
$\phi^{2a}1$	119.43(10.59)	106.96(12.15)	145.53(10.01)
$\phi^{2a}2$	113.22(8.64)	112.62(8.15)	102.04(8.51)
$\phi^{2a}3$	119.89(11.40)	132.10(10.83)	145.95(12.11)
$\phi^{2a}4$	114.04(9.98)	110.55(6.69)	105.50(9.06)
$\phi^{2a}5$	110.58(10.43)	106.97(6.15)	105.52(6.67)
$\phi^{2a}6$	114.82(11.36)	133.28(14.09)	149.90(9.17)
$\phi^{2a}7$	-	118.19(9.50)	110.22(8.29)
$\phi^{2a}8$	-	-	102.56(8.35)
$\psi^{2a}1$	112.03(11.49)	109.81(12.03)	133.17(5.38)
$\psi^{2a}2$	118.93(11.66)	117.94(14.69)	100.22(8.46)
$\psi^{2a}3$	124.96(9.45)	134.10(6.61)	131.64(6.13)
$\psi^{2a}4$	114.93(12.44)	114.96(11.30)	106.61(11.08)
$\psi^{2a}5$	128.67(15.10)	107.32(9.44)	108.23(7.84)
$\psi^{2a}6$	123.89(12.65)	132.12(6.63)	132.66(5.24)
$\psi^{2a}7$	-	120.33(10.44)	113.83(10.46)
$\psi^{2a}8$			105.39(10.22)

Table A.6: Averages of the alternative inter-saccharide torsion angles.

Torsion angle	Cyclodextrin		
	α-CD	β-CD	γ-CD
$\varphi 1$	117.658(2.751)	116.281(2.757)	117.646(2.882)
$\varphi 2$	117.391(2.740)	116.539(2.757)	115.202(2.705)
$\varphi 3$	117.989(2.784)	117.488(2.763)	117.441(2.805)
$\varphi 4$	117.306(2.774)	116.783(2.731)	115.278(2.705)
$\varphi 5$	117.930(2.859)	116.134(2.735)	115.579(2.727)
$\varphi 6$	117.708(2.756)	117.253(2.753)	117.699(2.796)
$\varphi 7$		117.051(2.697)	115.833(2.715)
$\varphi 8$			115.394(2.711)

Table A.7: Average values for each of the intersaccharide $C1$-$O1$-$C4$ angles.

Glucose unit	Puckering parameters		
	Q	θ	phi
1	0.558(0.035)	12.109(5.818)	81.321
2	0.568(0.029)	11.354(4.947)	80.016
3	0.560(0.028)	12.353(5.168)	70.017
4	0.558(0.029)	12.677(5.307)	72.254
5	0.563(0.030)	10.460(4.604)	88.394
6	0.556(0.030)	12.823(5.440)	75.958

Table A.8: Average puckering parameters for each of the glucose rings in α-CD.

Glucose unit	Puckering parameters		
	Q	θ	phi
1	0.560(0.028)	9.857(4.661)	89.41
2	0.562(0.030)	9.826(4.574)	101.764
3	0.564(0.029)	7.972(4.332)	145.836
4	0.561(0.029)	10.570(5.049)	89.149
5	0.563(0.028)	8.919(4.391)	105.268
6	0.564(0.028)	8.193(4.017)	125.097
7	0.562(0.029)	9.259(4.959)	104.162

Table A.9: Average puckering parameters for each of the glucose rings in β-CD.

Glucose unit	Puckering paramters		
	Q	θ	phi
1	0.564(0.035)	8.242(4.619)	126.54
2	0.565(0.036)	8.272(4.282)	126.88
3	0.566(0.035)	8.514(4.733)	116.33
4	0.567(0.034)	7.290(3.977)	163.53
5	0.565(0.035)	7.485(4.020)	138.46
6	0.565(0.035)	8.775(4.287)	107.29
7	0.568(0.035)	7.667(4.040)	146.03
8	0,568(0.035)	7.405(3.893)	173.42

Table A.10: Average puckering parameters for each of the glucose rings in γ-CD.

Appendix B

Supplementary Data for the crystal structure of the β-Cyclodextrin - D-Menthol Complex

B.1 Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($Å^2 \times 10^3$).

$U_{(eq)}$ is defined as one third of the trace of the orthogonalized U_{ij} tensor

.

Atom	x	y	z	U_eq
C(111)	-3542(7)	5703(3)	1729(7)	39(2)
C(112)	-4124(7)	5377(3)	1181(7)	42(2)
C(113)	-4019(7)	5392(3)	229(7)	43(3)
C(114)	-4286(6)	5817(3)	-137(7)	39(2)
C(115)	-3740(7)	6146(3)	466(7)	39(2)
C(116)	-4083(8)	6573(3)	252(8)	49(3)
O(112)	-3912(5)	4986(2)	1548(6)	55(2)
O(113)	-4560(5)	5097(2)	-303(5)	53(2)
O(114)	-4117(4)	5850(2)	-995(4)	39(2)
O(115)	-3793(5)	6088(2)	1386(5)	42(2)
O(116)	-4971(6)	6619(2)	353(6)	60(2)
C(121)	-4782(7)	6004(3)	-1683(7)	40(2)
C(122)	-5012(6)	5689(3)	-2415(7)	41(3)
C(123)	-4200(8)	5603(3)	-2811(7)	45(3)
C(124)	-3897(7)	6003(3)	-3153(7)	35(2)
C(125)	-3732(7)	6316(3)	-2397(7)	38(2)
C(126)	-3530(7)	6745(3)	-2725(8)	48(3)
O(122)	-5310(5)	5313(2)	-2058(5)	52(2)
O(123)	-4433(6)	5314(2)	-3499(5)	58(2)
O(124)	-3089(5)	5924(2)	-3414(5)	44(2)
O(125)	-4512(4)	6363(2)	-2021(5)	40(2)
O(126)	-3207(6)	7011(3)	-2009(7)	69(2)
C(131)	-3027(7)	6005(3)	-4291(7)	44(3)
C(132)	-2716(8)	5626(3)	-4676(7)	43(3)

C(133)	-1785(7)	5517(3)	-4143(7)	42(2)
C(134)	-1170(7)	5867(4)	-4163(7)	49(3)
C(135)	-1512(8)	6251(4)	-3852(8)	51(3)
C(136)	-984(8)	6617(4)	-4003(9)	59(3)
O(132)	-3329(5)	5306(2)	-4707(5)	55(2)
O(133)	-1477(5)	5160(2)	-4524(5)	49(2)
O(134)	-339(5)	5757(2)	-3580(5)	47(2)
O(135)	-2409(5)	6336(2)	-4310(5)	51(2)
O(136)	-985(7)	6650(3)	-4943(7)	82(3)
C(141)	472(7)	5819(3)	-3910(9)	51(3)
C(142)	931(7)	5418(3)	-3842(7)	41(2)
C(143)	1250(8)	5291(3)	-2879(7)	44(3)
C(144)	1795(8)	5622(3)	-2390(8)	47(3)
C(145)	1290(8)	6026(3)	-2492(8)	49(3)
C(146)	1871(12)	6388(4)	-2087(11)	79(5)
O(142)	405(5)	5110(2)	-4360(5)	52(2)
O(143)	1742(6)	4917(2)	-2823(6)	61(2)
O(144)	2070(4)	5523(2)	-1480(5)	40(2)
O(145)	996(5)	6129(2)	-3404(5)	54(2)
O(146)	2722(9)	6393(4)	-2312(9)	107(4)
C(151)	2976(6)	5516(3)	-1079(8)	42(3)
C(152)	3194(6)	5110(3)	-658(7)	39(2)
C(153)	2691(7)	5039(3)	75(7)	39(2)
C(154)	2912(6)	5394(3)	744(7)	38(2)
C(155)	2784(6)	5811(3)	291(7)	37(2)
C(156)	3112(8)	6151(4)	871(9)	54(3)
O(152)	2985(5)	4777(2)	-1290(5)	46(2)
O(153)	2936(5)	4668(2)	535(5)	48(2)
O(154)	2320(4)	5352(2)	1330(5)	42(2)
O(155)	3209(4)	5828(2)	-443(5)	44(2)
O(156)	4071(6)	6096(3)	1302(6)	69(2)
C(161)	2692(8)	5361(4)	2269(8)	53(3)
C(162)	2379(7)	4985(4)	2659(8)	46(3)
C(163)	1390(7)	4996(3)	2562(7)	42(2)
C(164)	1129(6)	5380(3)	2956(7)	35(2)
C(165)	1475(8)	5744(3)	2552(8)	48(3)
C(166)	1345(10)	6149(4)	3032(13)	77(5)
O(162)	2673(5)	4622(2)	2299(5)	54(2)
O(163)	1105(5)	4632(2)	2985(6)	54(2)
O(164)	182(4)	5407(2)	2768(4)	41(2)
O(165)	2433(4)	5710(2)	2653(5)	48(2)
O(166)	1667(9)	6137(4)	3916(10)	119(5)
C(171)	-265(6)	5441(3)	3510(7)	40(2)
C(172)	-907(7)	5106(3)	3448(7)	42(3)
C(173)	-1624(6)	5144(3)	2583(7)	39(2)
C(174)	-2060(6)	5552(3)	2586(6)	35(2)
C(175)	-1352(6)	5900(3)	2712(7)	40(2)
C(176)	-1742(8)	6306(4)	2840(8)	54(3)
O(172)	-486(4)	4707(2)	3482(5)	49(2)
O(173)	-2261(5)	4821(2)	2543(6)	56(2)
O(174)	-2660(4)	5608(2)	1744(5)	43(2)
O(175)	-676(5)	5822(2)	3492(5)	47(2)
O(176)	-2179(6)	6312(3)	3575(6)	68(2)

C(1)	-814(18)	5043(9)	-328(18)	139(8)
C(2)	-1110(2)	5221(10)	-1200(2)	179(11)
C(3)	-1650(2)	5623(10)	-1220(2)	187(12)
C(4)	-901(17)	5926(8)	-773(18)	134(8)
C(5)	-730(3)	5815(13)	150(3)	238(16)
C(6)	-270(3)	5365(12)	220(3)	222(15)
O(7)	-1726(17)	4949(8)	-1855(18)	219(9)
C(8)	-1310(2)	6354(10)	-660(2)	187(12)
C(9)	-450(3)	4596(13)	-530(3)	261(19)
C(10)	260(3)	4404(13)	270(3)	216(15)
C(11)	170(3)	4527(14)	-1190(3)	241(17)
C(21)	-1685(19)	2481(9)	-1120(19)	156(10)
C(22)	-1700(2)	2694(11)	-2050(2)	183(12
C(23)	-2240(3)	3059(11)	-2170(2)	204(14)
C(24)	-1820(18)	3391(8)	-1560(19)	142(8)
C(25)	-1670(3)	3199(12)	-660(3)	213(15)
C(26)	-1113(18)	2787(9)	-546(19)	148(9)
O(27)	-2180(2)	2361(11)	-2680(2)	285(15)
C(28)	-2260(2)	3786(10)	-1650(2)	178(11)
C(29)	-1240(2)	2041(9)	-1060(2)	165(10)
C(30)	-340(3)	2024(14)	-1290(3)	236(18)
C(31)	-1100(3)	1870(13)	-120(3)	213(15)
C(211)	-208(7)	3128(3)	2830(7)	44(3)
C(212)	479(7)	3475(3)	3019(8)	46(3)
C(213)	858(6)	3545(3)	2216(7)	37(2)
C(214)	1244(6)	3158(3)	1953(7)	38(2)
C(215)	585(7)	2803(3)	1843(7)	41(2)
C(216)	953(9)	2402(4)	1719(10)	62(4)
O(212)	109(5)	3830(2)	3304(5)	47(2)
O(213)	1567(5)	3844(2)	2415(6)	56(2)
O(214)	1501(5)	3214(2)	1139(5)	43(2)
O(215)	177(5)	2769(2)	2600(5)	48(2)
O(216)	1718(6)	2304(3)	2386(7)	74(3)
C(221)	2388(7)	3125(3)	1032(8)	43(3)
C(222)	2751(7)	3510(4)	735(8)	50(3)
C(223)	2214(7)	3626(3)	-167(7)	38(2)
C(224)	2234(6)	3278(3)	-798(7)	38(2)
C(225)	1931(7)	2874(3)	-404(7)	39(2)
C(226)	2101(8)	2497(3)	-954(8)	50(3)
O(222)	2804(5)	3834(2)	1351(5)	53(2)
O(223)	2587(5)	3978(2)	-490(5)	49(2)
O(224)	1631(5)	3368(2)	-1622(5)	47(2)
O(225)	2416(5)	2801(2)	463(5)	43(2)
O(226)	2973(5)	2482(2)	-1065(5)	54(2)
C(231)	1934(7)	3314(3)	-2410(7)	44(3)
C(232)	1864(7)	3715(3)	-2896(7)	45(3)
C(233)	937(7)	3844(3)	-3212(7)	39(2)
C(234)	417(6)	3506(3)	-3778(7)	41(2)
C(235)	503(8)	3108(3)	-3230(8)	50(3)
C(236)	53(9)	2755(4)	-3797(10)	69(4)
O(232)	2401(5)	4013(2)	-2351(5)	54(2)
O(233)	869(5)	4215(2)	-3707(6)	56(2)
O(234)	-511(4)	3621(2)	-4021(4)	39(2)

O(235)	1420(5)	3010(2)	-2943(5)	49(2)
O(236)	334(7)	2719(3)	-4592(7)	80(3)
C(241)	-904(7)	3621(3)	-4924(7)	41(2)
C(242)	-1335(8)	4032(4)	-5150(7)	48(3)
C(243)	-2072(8)	4083(3)	-4685(7)	48(3)
C(244)	-2726(7)	3744(3)	-4887(8)	46(3)
C(245)	-2246(7)	3329(4)	-4686(9)	52(3)
C(246)	-2863(9)	2971(4)	-4980(11)	70(4)
O(242)	-698(5)	4353(2)	-4983(5)	51(2)
O(243)	-2535(5)	4469(2)	-4928(6)	54(2)
O(244)	-3349(5)	3778(2)	-4300(5)	47(2)
O(245)	-1535(5)	3306(2)	-5130(5)	52(2)
O(246)	-2427(12)	2608(6)	-4599(12)	80(7)
O(247)	-3188(13)	3019(6)	-6044(14)	64(8)
C(251)	-4276(7)	3782(3)	-4701(7)	44(3)
C(252)	-4665(7)	4173(3)	-4340(8)	46(3)
C(253)	-4559(6)	4147(3)	-3381(7)	44(3)
C(254)	-4963(6)	3761(3)	-3107(7)	36(2)
C(255)	-4593(8)	3378(3)	-3483(7)	47(3)
C(256)	-5084(10)	3006(4)	-3375(9)	63(4)
O(252)	-4282(5)	4526(2)	-4631(5)	54(2)
O(253)	-4961(6)	4503(2)	-3059(5)	58(2)
O(254)	-4766(4)	3723(2)	-2174(4)	41(2)
O(255)	-4691(5)	3440(2)	-4462(5)	50(2)
O(256)	-6010(9)	3032(4)	-3681(9)	110(4)
C(261)	-5481(6)	3682(3)	-1741(7)	40(2)
C(262)	-5411(6)	4022(3)	-1061(7)	41(2)
C(263)	-4549(7)	3970(3)	-367(7)	40(2)
C(264)	-4542(6)	3542(3)	50(7)	39(2)
C(265)	-4672(7)	3210(3)	-652(7)	41(2)
C(266)	-4816(8)	2796(3)	-307(7)	43(3)
O(262)	-5477(5)	4411(2)	-1460(5)	44(2)
O(263)	-4491(6)	4276(2)	306(5)	56(2)
O(264)	-3689(4)	3487(2)	644(5)	39(2)
O(265)	-5460(5)	3294(2)	-1330(5)	43(2)
O(266)	-5548(6)	2791(2)	125(5)	58(2)
C(271)	-3657(6)	3379(3)	1515(7)	40(2)
C(272)	-3133(7)	3699(4)	2129(7)	48(3)
C(273)	-2174(6)	3698(3)	2044(7)	32(2)
C(274)	-1811(6)	3272(3)	2273(7)	39(2)
C(275)	-2357(7)	2967(3)	1661(9)	46(3)
C(276)	-2091(9)	2528(4)	1925(12)	77(5)
O(272)	-3526(5)	4087(2)	1945(5)	55(2)
O(273)	-1674(5)	3993(2)	2632(5)	53(2)
O(274)	-905(4)	3266(2)	2116(5)	40(2)
O(275)	-3279(4)	2993(2)	1711(5)	46(2)
O(276)	-2166(9)	2455(4)	2863(11)	127(5)
W(1)	3214(6)	2448(3)	7207(7)	76(3)
W(2)	6192(8)	4433(4)	3416(8)	96(3)
W(3)	6030(8)	1876(4)	1152(8)	95(3)
W(4)	1914(8)	2335(4)	5616(8)	96(3)
W(5)	4584(9)	4648(4)	1824(9)	110(4)
W(6)	3913(9)	1515(4)	6491(9)	110(4)

W(7)	4666(8)	2441(4)	1767(8)	105(4)
W(8)	5272(10)	945(5)	6939(10)	126(5)
W(9)	4700(9)	1962(5)	7913(9)	119(4)
W(10)	5025(14)	3035(7)	3211(14)	82(6)
W(11)	4501(12)	3782(6)	2339(13)	72(5)
W(12)	2533(13)	2583(6)	4101(13)	75(5)
W(13)	3103(14)	4385(7)	6372(14)	82(6)
W(14)	4134(13)	287(6)	6371(13)	78(5)
W(15)	1776(13)	4791(6)	4925(13)	78(5)
W(16)	1440(14)	3988(7)	4714(14)	83(6)
W(17)	8872(14)	2361(7)	4414(14)	85(6)
W(18)	9338(15)	1642(7)	5305(16)	93(6)
W(19)	6648(19)	118(9)	3448(18)	116(8)
W(20)	4345(17)	4645(8)	3770(17)	103(7)
W(21)	3230(2)	4012(12)	3850(3)	74(10)
W(22)	4175(19)	1977(10)	5040(2)	50(7)
W(23)	4270(2)	2742(11)	4480(2)	63(9)
W(24)	5890(3)	175(13)	5080(3)	86(12)
W(25)	6390(3)	1052(14)	5630(3)	90(12)
W(26)	3290(2)	-268(10)	1430(2)	54(8)
W(27)	6240(3)	2335(15)	3160(3)	98(14)
W(28)	2850(3)	3457(16)	4940(3)	101(14)
W(29)	6310(3)	1987(14)	7290(3)	86(12)
W(30)	7230(3)	1377(15)	3990(3)	98(14)
W(31)	6000(4)	1972(19)	4780(4)	122(17)
W(32)	810(2)	-571(10)	-870(2)	139(10)

B.2 Bond lengths ($\AA$).

C(111)-O(174)	1.388(12)
C(111)-O(115)	1.387(12)
C(111)-C(112)	1.529(16)
C(112)-O(112)	1.406(13)
C(112)-C(113)	1.515(15)
C(113)-O(113)	1.417(13)
C(113)-C(114)	1.523(15)
C(114)-O(114)	1.409(12)
C(114)-C(115)	1.545(14)
C(115)-O(115)	1.454(13)
C(115)-C(116)	1.505(15)
C(116)-O(116)	1.416(14)
O(114)-C(121)	1.402(13)
C(121)-O(125)	1.385(12)
C(121)-C(122)	1.515(15)
C(122)-O(122)	1.462(13)
C(122)-C(123)	1.531(15)
C(123)-O(123)	1.411(13)
C(123)-C(124)	1.520(15)
C(124)-O(124)	1.411(12)
C(124)-C(125)	1.532(14)
C(125)-O(125)	1.449(12)

C(125)-C(126)	1.546(15)
C(126)-O(126)	1.411(14)
O(124)-C(131)	1.404(13)
C(131)-O(135)	1.446(13)
C(131)-C(132)	1.496(15)
C(132)-O(132)	1.402(13)
C(132)-C(133)	1.534(16)
C(133)-O(133)	1.435(13)
C(133)-C(134)	1.490(16)
C(134)-O(134)	1.444(14)
C(134)-C(135)	1.481(16)
C(135)-O(135)	1.438(14)
C(135)-C(136)	1.493(17)
C(136)-O(136)	1.455(17)
O(134)-C(141)	1.459(13)
C(141)-O(145)	1.422(14)
C(141)-C(142)	1.481(15)
C(142)-O(142)	1.426(13)
C(142)-C(143)	1.522(15)
C(143)-O(143)	1.432(13)
C(143)-C(144)	1.473(15)
C(144)-O(144)	1.415(13)
C(144)-C(145)	1.526(15)
C(145)-O(145)	1.425(14)
C(145)-C(146)	1.533(17)
C(146)-O(146)	1.43(2)
O(144)-C(151)	1.400(12)
C(151)-O(155)	1.408(13)
C(151)-C(152)	1.486(15)
C(152)-O(152)	1.452(13)
C(152)-C(153)	1.520(15)
C(153)-O(153)	1.415(12)
C(153)-C(154)	1.543(14)
C(154)-O(154)	1.421(12)
C(154)-C(155)	1.528(15)
C(155)-O(155)	1.427(12)
C(155)-C(156)	1.448(16)
C(156)-O(156)	1.495(16)
O(154)-C(161)	1.442(14)
C(161)-O(165)	1.382(14)
C(161)-C(162)	1.494(17)
C(162)-O(162)	1.423(13)
C(162)-C(163)	1.497(15)
C(163)-O(163)	1.469(13)
C(163)-C(164)	1.487(14)
C(164)-O(164)	1.427(11)
C(164)-C(165)	1.496(15)
C(165)-O(165)	1.454(13)
C(165)-C(166)	1.550(17)
C(166)-O(166)	1.35(2)
O(164)-C(171)	1.462(12)
C(171)-O(175)	1.393(13)
C(171)-C(172)	1.465(15)

C(172)-O(172)	1.452(13)
C(172)-C(173)	1.545(15)
C(173)-O(173)	1.435(12)
C(173)-C(174)	1.493(14)
C(174)-O(174)	1.436(12)
C(174)-C(175)	1.559(14)
C(175)-O(175)	1.437(13)
C(175)-C(176)	1.487(16)
C(176)-O(176)	1.440(15)
C(1)-C(2)	1.45(3)
C(1)-C(6)	1.49(3)
C(1)-C(9)	1.62(4)
C(2)-O(7)	1.52(3)
C(2)-C(3)	1.55(3)
C(3)-C(4)	1.56(3)
C(4)-C(5)	1.44(4)
C(4)-C(8)	1.56(3)
C(5)-C(6)	1.62(4)
C(9)-C(11)	1.55(4)
C(9)-C(10)	1.60(4)
C(21)-C(26)	1.49(3)
C(21)-C(29)	1.59(3)
C(21)-C(22)	1.59(3)
C(22)-C(23)	1.44(4)
C(22)-O(27)	1.54(3)
C(23)-C(24)	1.49(3)
C(24)-C(28)	1.46(3)
C(24)-C(25)	1.50(3)
C(25)-C(26)	1.59(3)
C(29)-C(30)	1.50(4)
C(29)-C(31)	1.52(4)
C(211)-O(215)	1.395(13)
C(211)-O(274)	1.437(12)
C(211)-C(212)	1.535(15)
C(212)-O(212)	1.406(13)
C(212)-C(213)	1.497(15)
C(213)-O(213)	1.448(12)
C(213)-C(214)	1.492(14)
C(214)-O(214)	1.408(12)
C(214)-C(215)	1.526(14)
C(215)-O(215)	1.446(13)
C(215)-C(216)	1.459(16)
C(216)-O(216)	1.425(17)
O(214)-C(221)	1.440(12)
C(221)-O(225)	1.384(13)
C(221)-C(222)	1.489(16)
C(222)-O(222)	1.414(14)
C(222)-C(223)	1.509(16)
C(223)-O(223)	1.424(11)
C(223)-C(224)	1.504(14)
C(224)-O(224)	1.436(13)
C(224)-C(225)	1.567(14)
C(225)-O(225)	1.408(13)

C(225)-C(226)	1.551(15)
C(226)-O(226)	1.391(14)
O(224)-C(231)	1.406(13)
C(231)-O(235)	1.419(13)
C(231)-C(232)	1.503(15)
C(232)-O(232)	1.429(13)
C(232)-C(233)	1.467(15)
C(233)-O(233)	1.426(12)
C(233)-C(234)	1.525(15)
C(234)-O(234)	1.446(12)
C(234)-C(235)	1.543(16)
C(235)-O(235)	1.423(14)
C(235)-C(236)	1.523(17)
C(236)-O(236)	1.391(18)
O(234)-C(241)	1.397(12)
C(241)-O(245)	1.404(13)
C(241)-C(242)	1.507(16)
C(242)-O(242)	1.420(14)
C(242)-C(243)	1.478(16)
C(243)-O(243)	1.459(13)
C(243)-C(244)	1.485(16)
C(244)-O(244)	1.459(13)
C(244)-C(245)	1.543(17)
C(245)-O(245)	1.414(13)
C(245)-C(246)	1.516(17)
C(246)-O(246)	1.43(2)
C(246)-O(247)	1.62(3)
O(244)-C(251)	1.429(13)
C(251)-O(255)	1.375(13)
C(251)-C(252)	1.566(16)
C(252)-O(252)	1.412(13)
C(252)-C(253)	1.458(16)
C(253)-O(253)	1.455(13)
C(253)-C(254)	1.507(15)
C(254)-O(254)	1.414(12)
C(254)-C(255)	1.540(15)
C(255)-C(256)	1.459(17)
C(255)-O(255)	1.502(14)
C(256)-O(256)	1.405(19)
O(254)-C(261)	1.411(12)
C(261)-O(265)	1.416(13)
C(261)-C(262)	1.518(15)
C(262)-O(262)	1.408(13)
C(262)-C(263)	1.526(15)
C(263)-O(263)	1.433(13)
C(263)-C(264)	1.540(14)
C(264)-O(264)	1.443(11)
C(264)-C(265)	1.516(15)
C(265)-O(265)	1.448(13)
C(265)-C(266)	1.491(15)
C(266)-O(266)	1.427(14)
O(264)-C(271)	1.382(12)
C(271)-O(275)	1.394(13)

C(271)-C(272)	1.523(16)
C(272)-O(272)	1.409(14)
C(272)-C(273)	1.508(14)
C(273)-O(273)	1.430(12)
C(273)-C(274)	1.515(14)
C(274)-O(274)	1.466(12)
C(274)-C(275)	1.502(15)
C(275)-O(275)	1.440(12)
C(275)-C(276)	1.524(15)
C(276)-O(276)	1.50(2)

B.3 Bond angles (in degrees).

O(174)-C(111)-O(115)	113.6(8)
O(174)-C(111)-C(112)	107.9(8)
O(115)-C(111)-C(112)	109.8(8)
O(112)-C(112)-C(113)	111.0(9)
O(112)-C(112)-C(111)	110.6(8)
C(113)-C(112)-C(111)	110.0(9)
O(113)-C(113)-C(112)	111.6(9)
O(113)-C(113)-C(114)	109.2(9)
C(112)-C(113)-C(114)	108.4(9)
O(114)-C(114)-C(113)	109.2(8)
O(114)-C(114)-C(115)	109.1(8)
C(113)-C(114)-C(115)	110.0(8)
O(115)-C(115)-C(116)	104.2(8)
O(115)-C(115)-C(114)	111.2(8)
C(116)-C(115)-C(114)	113.1(9)
O(116)-C(116)-C(115)	112.2(9)
C(121)-O(114)-C(114)	119.3(7)
C(111)-O(115)-C(115)	115.0(7)
O(125)-C(121)-O(114)	111.2(8)
O(125)-C(121)-C(122)	109.6(8)
O(114)-C(121)-C(122)	109.7(8)
O(122)-C(122)-C(121)	109.4(8)
O(122)-C(122)-C(123)	110.2(9)
C(121)-C(122)-C(123)	110.1(8)
O(123)-C(123)-C(124)	111.3(9)
O(123)-C(123)-C(122)	109.1(9)
C(124)-C(123)-C(122)	108.6(8)
O(124)-C(124)-C(123)	107.3(8)
O(124)-C(124)-C(125)	109.1(8)
C(123)-C(124)-C(125)	108.8(8)
O(125)-C(125)-C(124)	111.1(8)
O(125)-C(125)-C(126)	106.0(8)
C(124)-C(125)-C(126)	111.6(8)
O(126)-C(126)-C(125)	111.3(10)
C(131)-O(124)-C(124)	119.2(8)
C(121)-O(125)-C(125)	113.7(7)
O(124)-C(131)-O(135)	109.9(9)
O(124)-C(131)-C(132)	108.8(9)

O(135)-C(131)-C(132)	110.2(8)
O(132)-C(132)-C(131)	111.1(8)
O(132)-C(132)-C(133)	113.0(9)
C(131)-C(132)-C(133)	108.6(9)
O(133)-C(133)-C(134)	110.5(8)
O(133)-C(133)-C(132)	108.7(9)
C(134)-C(133)-C(132)	109.6(9)
O(134)-C(134)-C(135)	109.7(9)
O(134)-C(134)-C(133)	106.1(9)
C(135)-C(134)-C(133)	111.9(9)
O(135)-C(135)-C(134)	112.0(10)
O(135)-C(135)-C(136)	105.4(9)
C(134)-C(135)-C(136)	112.7(10)
O(136)-C(136)-C(135)	109.3(10)
C(134)-O(134)-C(141)	117.2(8)
C(135)-O(135)-C(131)	114.3(8)
O(145)-C(141)-O(134)	109.5(9)
O(145)-C(141)-C(142)	112.7(9)
O(134)-C(141)-C(142)	106.5(9)
O(142)-C(142)-C(141)	112.3(9)
O(142)-C(142)-C(143)	112.4(8)
C(141)-C(142)-C(143)	111.1(9)
O(143)-C(143)-C(144)	111.0(9)
O(143)-C(143)-C(142)	110.3(9)
C(144)-C(143)-C(142)	109.3(9)
O(144)-C(144)-C(143)	111.0(9)
O(144)-C(144)-C(145)	109.7(9)
C(143)-C(144)-C(145)	111.1(9)
O(145)-C(145)-C(144)	110.7(9)
O(145)-C(145)-C(146)	104.8(10)
C(144)-C(145)-C(146)	112.7(10)
O(146)-C(146)-C(145)	113.4(12)
C(151)-O(144)-C(144)	120.0(8)
C(141)-O(145)-C(145)	113.5(8)
O(144)-C(151)-O(155)	112.3(8)
O(144)-C(151)-C(152)	108.6(8)
O(155)-C(151)-C(152)	110.0(9)
O(152)-C(152)-C(151)	112.3(8)
O(152)-C(152)-C(153)	108.1(8)
C(151)-C(152)-C(153)	111.0(8)
O(153)-C(153)-C(152)	112.4(8)
O(153)-C(153)-C(154)	108.1(9)
C(152)-C(153)-C(154)	108.0(8)
O(154)-C(154)-C(155)	109.8(8)
O(154)-C(154)-C(153)	106.0(7)
C(155)-C(154)-C(153)	112.1(9)
O(155)-C(155)-C(156)	107.6(8)
O(155)-C(155)-C(154)	111.2(8)
C(156)-C(155)-C(154)	114.1(9)
C(155)-C(156)-O(156)	111.3(9)
C(154)-O(154)-C(161)	117.9(7)
C(151)-O(155)-C(155)	115.7(7)
O(165)-C(161)-O(154)	111.1(10)

O(165)-C(161)-C(162)	111.0(9)
O(154)-C(161)-C(162)	107.2(9)
O(162)-C(162)-C(161)	111.8(9)
O(162)-C(162)-C(163)	112.5(9)
C(161)-C(162)-C(163)	110.4(9)
O(163)-C(163)-C(164)	111.7(8)
O(163)-C(163)-C(162)	109.3(9)
C(164)-C(163)-C(162)	109.7(8)
O(164)-C(164)-C(163)	108.9(8)
O(164)-C(164)-C(165)	107.7(8)
C(163)-C(164)-C(165)	110.6(8)
O(165)-C(165)-C(164)	109.7(9)
O(165)-C(165)-C(166)	104.4(9)
C(164)-C(165)-C(166)	112.8(10)
O(166)-C(166)-C(165)	113.6(13)
C(164)-O(164)-C(171)	118.5(7)
C(161)-O(165)-C(165)	113.4(8)
O(175)-C(171)-C(172)	111.7(8)
O(175)-C(171)-O(164)	109.9(8)
C(172)-C(171)-O(164)	108.4(9)
O(172)-C(172)-C(171)	112.3(8)
O(172)-C(172)-C(173)	109.6(9)
C(171)-C(172)-C(173)	110.0(8)
O(173)-C(173)-C(174)	110.7(8)
O(173)-C(173)-C(172)	109.4(8)
C(174)-C(173)-C(172)	108.0(8)
O(174)-C(174)-C(173)	108.3(8)
O(174)-C(174)-C(175)	108.5(8)
C(173)-C(174)-C(175)	110.4(8)
O(175)-C(175)-C(176)	106.7(9)
O(175)-C(175)-C(174)	109.7(8)
C(176)-C(175)-C(174)	112.1(8)
O(176)-C(176)-C(175)	112.9(10)
C(111)-O(174)-C(174)	118.7(8)
C(171)-O(175)-C(175)	114.8(8)
C(2)-C(1)-C(6)	106(3)
C(2)-C(1)-C(9)	104(3)
C(6)-C(1)-C(9)	125(3)
C(1)-C(2)-O(7)	114(3)
C(1)-C(2)-C(3)	115(3)
O(7)-C(2)-C(3)	103(2)
C(2)-C(3)-C(4)	101(2)
C(5)-C(4)-C(3)	104(3)
C(5)-C(4)-C(8)	96(3)
C(3)-C(4)-C(8)	110(2)
C(4)-C(5)-C(6)	106(3)
C(1)-C(6)-C(5)	115(3)
C(11)-C(9)-C(10)	93(3)
C(11)-C(9)-C(1)	123(4)
C(10)-C(9)-C(1)	114(3)
C(26)-C(21)-C(29)	112(2)
C(26)-C(21)-C(22)	98(2)
C(29)-C(21)-C(22)	112(2)

C(23)-C(22)-O(27)	109(3)
C(23)-C(22)-C(21)	112(3)
O(27)-C(22)-C(21)	100(3)
C(22)-C(23)-C(24)	112(3)
C(28)-C(24)-C(25)	116(3)
C(28)-C(24)-C(23)	117(3)
C(25)-C(24)-C(23)	104(3)
C(24)-C(25)-C(26)	116(3)
C(21)-C(26)-C(25)	106(3)
C(30)-C(29)-C(31)	105(3)
C(30)-C(29)-C(21)	115(3)
C(31)-C(29)-C(21)	111(3)
O(215)-C(211)-O(274)	111.0(9)
O(215)-C(211)-C(212)	110.9(9)
O(274)-C(211)-C(212)	106.6(8)
O(212)-C(212)-C(213)	112.8(9)
O(212)-C(212)-C(211)	111.1(8)
C(213)-C(212)-C(211)	109.3(9)
O(213)-C(213)-C(212)	109.7(8)
O(213)-C(213)-C(214)	107.8(8)
C(212)-C(213)-C(214)	110.1(8)
O(214)-C(214)-C(213)	109.6(8)
O(214)-C(214)-C(215)	107.9(8)
C(213)-C(214)-C(215)	112.4(8)
O(215)-C(215)-C(216)	106.6(9)
O(215)-C(215)-C(214)	111.3(8)
C(216)-C(215)-C(214)	115.5(9)
O(216)-C(216)-C(215)	112.9(11)
C(214)-O(214)-C(221)	122.0(8)
C(211)-O(215)-C(215)	115.0(8)
O(225)-C(221)-O(214)	112.8(8)
O(225)-C(221)-C(222)	112.7(9)
O(214)-C(221)-C(222)	107.0(8)
O(222)-C(222)-C(221)	113.3(9)
O(222)-C(222)-C(223)	112.0(9)
C(221)-C(222)-C(223)	109.3(9)
O(223)-C(223)-C(224)	108.6(8)
O(223)-C(223)-C(222)	110.2(8)
C(224)-C(223)-C(222)	109.0(8)
O(224)-C(224)-C(223)	108.7(8)
O(224)-C(224)-C(225)	109.2(8)
C(223)-C(224)-C(225)	109.7(8)
O(225)-C(225)-C(226)	105.7(8)
O(225)-C(225)-C(224)	111.8(8)
C(226)-C(225)-C(224)	111.0(9)
O(226)-C(226)-C(225)	112.2(9)
C(231)-O(224)-C(224)	118.1(7)
C(221)-O(225)-C(225)	113.1(8)
O(224)-C(231)-O(235)	110.1(9)
O(224)-C(231)-C(232)	108.7(8)
O(235)-C(231)-C(232)	110.5(9)
O(232)-C(232)-C(233)	113.6(9)
O(232)-C(232)-C(231)	109.2(9)

C(233)-C(232)-C(231)	112.2(9)
O(233)-C(233)-C(232)	112.2(9)
O(233)-C(233)-C(234)	109.8(8)
C(232)-C(233)-C(234)	109.7(9)
O(234)-C(234)-C(233)	109.2(8)
O(234)-C(234)-C(235)	109.1(8)
C(233)-C(234)-C(235)	108.9(9)
O(235)-C(235)-C(236)	107.9(10)
O(235)-C(235)-C(234)	108.9(9)
C(236)-C(235)-C(234)	110.4(10)
O(236)-C(236)-C(235)	112.5(11)
C(241)-O(234)-C(234)	116.9(8)
C(231)-O(235)-C(235)	114.6(8)
O(234)-C(241)-O(245)	111.0(9)
O(234)-C(241)-C(242)	107.8(8)
O(245)-C(241)-C(242)	110.7(8)
O(242)-C(242)-C(243)	113.6(9)
O(242)-C(242)-C(241)	111.4(9)
C(243)-C(242)-C(241)	109.5(9)
O(243)-C(243)-C(242)	110.9(9)
O(243)-C(243)-C(244)	108.7(9)
C(242)-C(243)-C(244)	111.7(10)
O(244)-C(244)-C(243)	108.6(9)
O(244)-C(244)-C(245)	106.9(9)
C(243)-C(244)-C(245)	109.9(9)
O(245)-C(245)-C(246)	108.2(10)
O(245)-C(245)-C(244)	109.7(9)
C(246)-C(245)-C(244)	112.1(10)
O(246)-C(246)-C(245)	108.1(13)
O(246)-C(246)-O(247)	121.0(14)
C(245)-C(246)-O(247)	105.5(12)
C(251)-O(244)-C(244)	117.4(8)
C(241)-O(245)-C(245)	114.9(8)
O(255)-C(251)-O(244)	110.8(9)
O(255)-C(251)-C(252)	109.5(8)
O(244)-C(251)-C(252)	106.0(8)
O(252)-C(252)-C(253)	114.1(10)
O(252)-C(252)-C(251)	109.8(9)
C(253)-C(252)-C(251)	110.1(9)
C(252)-C(253)-O(253)	109.7(9)
C(252)-C(253)-C(254)	111.9(9)
O(253)-C(253)-C(254)	110.0(8)
O(254)-C(254)-C(253)	110.5(8)
O(254)-C(254)-C(255)	107.3(8)
C(253)-C(254)-C(255)	111.5(8)
C(256)-C(255)-O(255)	106.6(9)
C(256)-C(255)-C(254)	113.1(10)
O(255)-C(255)-C(254)	107.7(9)
O(256)-C(256)-C(255)	115.1(11)
C(261)-O(254)-C(254)	118.2(7)
C(251)-O(255)-C(255)	115.4(8)
O(254)-C(261)-O(265)	110.6(8)
O(254)-C(261)-C(262)	108.2(8)

O(265)-C(261)-C(262)	110.7(8)
O(262)-C(262)-C(261)	111.7(9)
O(262)-C(262)-C(263)	112.3(8)
C(261)-C(262)-C(263)	109.1(8)
O(263)-C(263)-C(262)	109.9(8)
O(263)-C(263)-C(264)	109.8(8)
C(262)-C(263)-C(264)	108.8(8)
O(264)-C(264)-C(265)	109.3(8)
O(264)-C(264)-C(263)	107.6(8)
C(265)-C(264)-C(263)	111.3(8)
O(265)-C(265)-C(266)	105.3(8)
O(265)-C(265)-C(264)	110.1(8)
C(266)-C(265)-C(264)	113.6(9)
O(266)-C(266)-C(265)	111.7(9)
C(271)-O(264)-C(264)	119.3(7)
C(261)-O(265)-C(265)	115.4(7)
O(264)-C(271)-O(275)	111.7(9)
O(264)-C(271)-C(272)	109.5(8)
O(275)-C(271)-C(272)	110.2(8)
O(272)-C(272)-C(273)	111.7(9)
O(272)-C(272)-C(271)	110.3(8)
C(273)-C(272)-C(271)	110.1(8)
O(273)-C(273)-C(274)	110.5(8)
O(273)-C(273)-C(272)	110.4(8)
C(274)-C(273)-C(272)	107.5(8)
O(274)-C(274)-C(275)	108.1(8)
O(274)-C(274)-C(273)	106.9(8)
C(275)-C(274)-C(273)	109.7(8)
O(275)-C(275)-C(274)	110.6(9)
O(275)-C(275)-C(276)	104.7(9)
C(274)-C(275)-C(276)	112.1(10)
O(276)-C(276)-C(275)	109.6(13)
C(211)-O(274)-C(274)	117.3(8)
C(271)-O(275)-C(275)	114.4(7)

B.4 Anisotropic displacement parameters ($Å^2 \times 10^3$)

The anisotropic displacement factor exponent takes the form:
$-2\pi^2[h^2a*^2 U_{11} + ... + 2hka*b*U_{12}]$

	U11	U22	U33	U23	U13	U12
C(111)	43(6)	36(5)	42(6)	9(5)	19(5)	6(4)
C(112)	36(5)	54(7)	36(6)	6(5)	5(4)	9(5)
C(113)	43(6)	25(5)	60(7)	3(5)	12(5)	-1(4)
C(114)	31(5)	42(6)	45(6)	-8(5)	9(4)	4(4)
C(115)	53(6)	26(5)	34(6)	-9(4)	4(5)	-7(4)
C(116)	57(7)	44(6)	42(6)	-3(5)	5(5)	-4(5)

O(112)	51(4)	38(4)	70(5)	18(4)	3(4)	-1(3)
O(113)	62(5)	35(4)	54(5)	-10(4)	-7(4)	-11(4)
O(114)	45(4)	38(4)	33(4)	3(3)	8(3)	6(3)
O(115)	50(4)	30(4)	44(4)	-1(3)	6(3)	8(3)
O(116)	59(5)	52(5)	64(5)	-1(4)	4(4)	15(4)
C(121)	35(5)	38(6)	50(7)	5(5)	13(5)	4(4)
C(122)	28(5)	60(7)	37(6)	-3(5)	9(4)	-2(5)
C(123)	60(7)	36(6)	41(6)	-10(5)	16(5)	4(5)
C(124)	40(5)	31(5)	36(6)	2(4)	12(4)	11(4)
C(125)	43(6)	34(5)	37(6)	-2(5)	9(5)	-2(4)
C(126)	39(6)	37(6)	72(8)	-16(6)	23(6)	-12(5)
O(122)	55(5)	51(4)	58(5)	-2(4)	26(4)	-16(4)
O(123)	79(6)	44(4)	59(5)	-21(4)	30(4)	-13(4)
O(124)	46(4)	49(4)	40(4)	11(3)	16(3)	10(3)
O(125)	46(4)	36(4)	42(4)	-1(3)	17(3)	1(3)
O(126)	68(5)	50(5)	95(7)	-10(5)	30(5)	-15(4)
C(131)	42(6)	49(6)	41(6)	13(5)	7(5)	-1(5)
C(132)	62(7)	34(6)	41(6)	5(5)	27(5)	2(5)
C(133)	45(6)	43(6)	44(6)	6(5)	19(5)	6(5)
C(134)	48(6)	62(7)	39(6)	6(6)	16(5)	-3(6)
C(135)	57(7)	51(7)	44(7)	9(5)	8(6)	8(6)
C(136)	45(7)	49(7)	81(10)	-5(7)	11(6)	-4(5)
O(132)	58(5)	50(5)	64(5)	-15(4)	28(4)	-6(4)
O(133)	48(4)	39(4)	62(5)	-13(4)	19(4)	8(3)
O(134)	51(4)	41(4)	49(4)	13(3)	9(4)	10(3)
O(135)	56(5)	44(4)	57(5)	16(4)	22(4)	5(4)
O(136)	99(7)	57(6)	97(8)	11(5)	37(6)	-10(5)
C(141)	50(7)	43(6)	63(8)	2(6)	17(6)	4(5)
C(142)	37(5)	40(6)	46(6)	2(5)	9(5)	3(5)
C(143)	55(6)	24(5)	50(6)	0(5)	2(5)	-4(5)
C(144)	52(6)	33(6)	55(7)	-1(5)	7(6)	3(5)
C(145)	56(7)	39(6)	50(7)	-7(5)	8(6)	0(5)
C(146)	114(13)	34(7)	77(10)	-4(7)	-4(9)	-16(7)
O(142)	53(5)	56(5)	41(4)	-10(4)	-3(4)	2(4)
O(143)	63(5)	42(4)	67(5)	-8(4)	-8(4)	15(4)
O(144)	36(4)	42(4)	41(4)	-4(3)	6(3)	0(3)
O(145)	56(5)	37(4)	62(5)	6(4)	-2(4)	-7(4)
O(146)	110(9)	79(7)	121(10)	0(7)	-2(8)	-47(7)
C(151)	32(5)	40(6)	61(7)	-5(5)	23(5)	4(4)
C(152)	31(5)	47(6)	37(6)	-7(5)	1(4)	4(4)
C(153)	39(5)	20(5)	59(7)	-2(5)	13(5)	-8(4)
C(154)	27(5)	39(6)	49(6)	-6(5)	10(5)	-5(4)
C(155)	27(5)	30(5)	54(6)	-7(5)	8(5)	-6(4)
C(156)	64(8)	39(6)	64(8)	0(6)	23(6)	-7(6)

O(152)	47(4)	47(4)	42(4)	-8(4)	5(3)	6(3)
O(153)	59(5)	39(4)	49(5)	3(3)	17(4)	-6(3)
O(154)	29(3)	50(4)	46(4)	-5(3)	8(3)	-9(3)
O(155)	43(4)	41(4)	46(4)	-5(3)	7(3)	-9(3)
O(156)	70(6)	77(6)	58(5)	-12(5)	7(4)	-26(5)
C(161)	40(6)	65(8)	51(7)	-8(6)	4(5)	-10(6)
C(162)	32(5)	59(7)	45(6)	-5(5)	3(5)	2(5)
C(163)	39(6)	51(6)	36(6)	-6(5)	6(5)	-12(5)
C(164)	29(5)	40(6)	35(5)	-18(4)	5(4)	2(4)
C(165)	55(7)	32(6)	62(7)	-7(5)	24(6)	-12(5)
C(166)	60(8)	50(8)	127(15)	-21(9)	36(9)	-6(6)
O(162)	47(4)	55(5)	63(5)	10(4)	19(4)	15(4)
O(163)	49(4)	43(4)	78(6)	8(4)	32(4)	7(3)
O(164)	30(3)	54(4)	39(4)	-1(3)	9(3)	-2(3)
O(165)	28(4)	55(5)	60(5)	-16(4)	6(3)	-16(3)
O(166)	125(10)	120(10)	112(11)	-72(9)	27(9)	-17(8)
C(171)	29(5)	49(6)	41(6)	1(5)	6(4)	3(5)
C(172)	42(6)	46(6)	41(6)	11(5)	14(5)	10(5)
C(173)	33(5)	31(5)	53(7)	3(5)	11(5)	-9(4)
C(174)	32(5)	42(6)	30(5)	7(4)	5(4)	0(4)
C(175)	28(5)	38(6)	50(6)	3(5)	0(5)	-5(4)
C(176)	48(7)	47(7)	60(8)	-2(6)	-4(6)	-2(5)
O(172)	30(4)	56(5)	61(5)	8(4)	10(3)	11(3)
O(173)	42(4)	40(4)	84(6)	17(4)	6(4)	2(3)
O(174)	33(4)	53(4)	41(4)	3(3)	3(3)	-4(3)
O(175)	50(4)	44(4)	46(4)	-12(4)	8(4)	0(4)
O(176)	58(5)	69(6)	78(6)	-12(5)	17(5)	17(4)
C(211)	44(6)	37(6)	46(6)	14(5)	-1(5)	-5(5)
C(212)	37(6)	48(6)	51(7)	1(5)	4(5)	-4(5)
C(213)	28(5)	30(5)	52(6)	2(5)	8(5)	-8(4)
C(214)	32(5)	40(6)	43(6)	7(5)	12(5)	11(4)
C(215)	34(5)	44(6)	45(6)	-5(5)	8(5)	-8(5)
C(216)	67(8)	38(6)	93(10)	3(6)	42(8)	-5(6)
O(212)	41(4)	46(4)	54(5)	-3(4)	8(3)	-2(3)
O(213)	46(4)	44(4)	87(6)	-9(4)	30(4)	-9(4)
O(214)	46(4)	39(4)	44(4)	9(3)	11(3)	2(3)
O(215)	51(4)	33(4)	62(5)	16(4)	16(4)	2(3)
O(216)	77(6)	49(5)	97(7)	18(5)	20(6)	33(5)
C(221)	34(5)	46(6)	50(6)	1(5)	15(5)	2(5)
C(222)	29(5)	53(7)	67(8)	-2(6)	12(5)	4(5)
C(223)	36(5)	25(5)	59(7)	7(5)	22(5)	-5(4)
C(224)	27(5)	32(5)	56(7)	-7(5)	10(5)	2(4)
C(225)	44(6)	24(5)	49(7)	6(4)	7(5)	5(4)
C(226)	64(8)	28(5)	53(7)	-6(5)	3(6)	-6(5)

O(222)	68(5)	42(4)	51(5)	-12(4)	13(4)	-9(4)
O(223)	54(4)	34(4)	59(5)	1(4)	17(4)	-4(3)
O(224)	49(4)	44(4)	47(5)	3(3)	5(4)	12(3)
O(225)	48(4)	25(3)	61(5)	0(3)	20(4)	1(3)
O(226)	49(5)	43(4)	66(5)	-7(4)	5(4)	13(4)
C(231)	42(6)	34(6)	52(7)	-6(5)	3(5)	7(5)
C(232)	53(6)	42(6)	40(6)	-11(5)	13(5)	0(5)
C(233)	40(5)	35(5)	41(6)	8(5)	6(5)	-1(4)
C(234)	33(5)	45(6)	47(6)	-10(5)	8(5)	3(4)
C(235)	52(7)	46(6)	53(7)	1(6)	10(6)	3(5)
C(236)	68(8)	26(6)	97(12)	-4(7)	-16(8)	-7(6)
O(232)	50(4)	48(4)	60(5)	-5(4)	1(4)	1(4)
O(233)	60(5)	38(4)	63(5)	15(4)	-2(4)	-2(4)
O(234)	41(4)	36(4)	41(4)	1(3)	9(3)	4(3)
O(235)	49(4)	40(4)	53(5)	1(4)	-1(4)	11(3)
O(236)	86(7)	71(6)	82(7)	-34(5)	12(6)	-9(5)
C(241)	43(6)	41(6)	35(6)	-6(5)	2(5)	6(5)
C(242)	59(7)	57(7)	25(5)	-9(5)	3(5)	-2(6)
C(243)	64(7)	43(6)	34(6)	10(5)	2(5)	-4(5)
C(244)	45(6)	51(7)	42(6)	-8(5)	8(5)	-2(5)
C(245)	43(6)	59(7)	59(7)	-7(6)	21(6)	-8(5)
C(246)	67(8)	40(7)	104(11)	-3(7)	17(8)	-17(6)
O(242)	59(5)	39(4)	58(5)	13(4)	19(4)	-8(4)
O(243)	54(5)	38(4)	71(6)	8(4)	15(4)	11(3)
O(244)	40(4)	56(5)	44(4)	-9(4)	10(3)	-2(3)
O(245)	45(4)	40(4)	66(5)	-16(4)	4(4)	3(3)
C(251)	41(6)	48(6)	42(6)	3(5)	3(5)	-8(5)
C(252)	36(5)	46(6)	52(7)	23(5)	-1(5)	-4(5)
C(253)	28(5)	52(6)	51(7)	4(5)	9(5)	9(5)
C(254)	29(5)	36(5)	43(6)	2(5)	6(4)	4(4)
C(255)	55(7)	41(6)	44(7)	-5(5)	7(5)	-7(5)
C(256)	104(11)	40(7)	43(7)	-6(5)	12(7)	-4(7)
O(252)	57(5)	53(5)	58(5)	6(4)	27(4)	-4(4)
O(253)	83(6)	46(5)	50(5)	4(4)	31(4)	14(4)
O(254)	35(4)	51(4)	32(4)	0(3)	-1(3)	-2(3)
O(255)	57(5)	47(5)	44(4)	-11(4)	5(4)	-9(4)
O(256)	101(9)	110(9)	109(9)	21(7)	-2(7)	-55(8)
C(261)	22(5)	56(7)	40(6)	-4(5)	4(4)	-11(4)
C(262)	33(5)	42(6)	52(6)	-6(5)	19(5)	7(4)
C(263)	40(5)	31(5)	53(7)	4(5)	17(5)	1(4)
C(264)	22(5)	44(6)	47(6)	3(5)	-1(4)	7(4)
C(265)	43(6)	40(6)	43(6)	2(5)	12(5)	2(5)
C(266)	59(7)	29(5)	39(6)	1(5)	2(5)	0(5)
O(262)	52(4)	28(4)	50(4)	8(3)	5(3)	12(3)

O(263)	82(6)	36(4)	43(4)	-2(3)	-3(4)	0(4)
O(264)	30(3)	39(4)	49(4)	8(3)	9(3)	6(3)
O(265)	45(4)	39(4)	40(4)	4(3)	0(3)	-6(3)
O(266)	71(5)	49(5)	54(5)	11(4)	15(4)	-16(4)
C(271)	28(5)	50(6)	42(6)	16(5)	11(4)	-7(4)
C(272)	48(6)	64(8)	31(6)	-3(5)	8(5)	12(5)
C(273)	34(5)	27(5)	38(5)	0(4)	14(4)	-1(4)
C(274)	35(5)	39(6)	44(6)	2(5)	11(5)	5(4)
C(275)	35(5)	22(5)	79(8)	12(5)	5(5)	4(4)
C(276)	54(8)	31(6)	131(14)	10(8)	-14(8)	-10(6)
O(272)	62(5)	53(5)	44(4)	-7(4)	1(4)	27(4)
O(273)	41(4)	62(5)	49(4)	-11(4)	-7(3)	7(4)
O(274)	33(4)	42(4)	45(4)	2(3)	8(3)	0(3)
O(275)	33(4)	32(4)	72(5)	22(4)	11(3)	4(3)
O(276)	128(10)	78(8)	159(13)	74(9)	-5(9)	-20(7)

B.5 Hydrogen coordinates ($\times 10^4$) and isotropic displacement parameters ($Å^2 \times 10^3$).

H(111)	-3633	5692	2337	53(5)
H(112)	-4747	5435	1190	53(5)
H(113)	-3395	5343	212	53(5)
H(114)	-4921	5861	-163	53(5)
H(115)	-3117	6131	415	53(5)
H(11A)	-3705	6764	640	53(5)
H(11B)	-4053	6638	-353	53(5)
H(121)	-5314	6061	-1450	53(5)
H(122)	-5494	5796	-2881	53(5)
H(123)	-3720	5493	-2350	53(5)
H(124)	-4345	6105	-3655	53(5)
H(125)	-3230	6226	-1934	53(5)
H(12A)	-3092	6721	-3090	53(5)
H(12B)	-4069	6859	-3088	53(5)
H(131)	-3614	6081	-4638	53(5)
H(132)	-2667	5687	-5285	53(5)
H(133)	-1817	5460	-3528	53(5)
H(134)	-1079	5905	-4767	53(5)
H(135)	-1496	6227	-3216	53(5)
H(13A)	-1242	6862	-3806	53(5)
H(13B)	-378	6591	-3665	53(5)
H(141)	314	5903	-4533	53(5)
H(142)	1462	5455	-4088	53(5)

H(143)	730	5248	-2618	53(5)
H(144)	2327	5655	-2636	53(5)
H(145)	776	6004	-2215	53(5)
H(14A)	1563	6641	-2288	53(5)
H(14B)	1952	6377	-1448	53(5)
H(151)	3320	5553	-1538	53(5)
H(152)	3833	5102	-397	53(5)
H(153)	2050	5037	-179	53(5)
H(154)	3527	5367	1078	53(5)
H(155)	2145	5852	65	53(5)
H(15A)	2765	6173	1324	53(5)
H(15B)	3039	6403	531	53(5)
H(161)	3342	5355	2368	53(5)
H(162)	2644	4987	3294	53(5)
H(163)	1110	4991	1930	53(5)
H(164)	1360	5379	3598	53(5)
H(165)	1189	5764	1923	53(5)
H(16A)	1637	6369	2782	53(5)
H(16B)	715	6212	2920	53(5)
H(171)	179	5415	4066	53(5)
H(172)	-1200	5128	3950	53(5)
H(173)	-1342	5126	2072	53(5)
H(174)	-2390	5563	3062	53(5)
H(175)	-1081	5910	2193	53(5)
H(17A)	-1274	6509	2936	53(5)
H(17B)	-2168	6379	2304	53(5)
H(211)	-452	3080	3357	38(4)
H(212)	965	3385	3500	38(4)
H(213)	391	3644	1727	38(4)
H(214)	1770	3086	2407	38(4)
H(215)	113	2861	1323	38(4)
H(21A)	501	2195	1721	38(4)
H(21B)	1112	2394	1144	38(4)
H(221)	2749	3053	1615	38(4)
H(222)	3358	3452	669	38(4)
H(223)	1599	3683	-128	38(4)
H(224)	2837	3244	-898	38(4)
H(225)	1296	2891	-402	38(4)
H(22A)	1972	2250	-658	38(4)
H(22B)	1700	2507	-1531	38(4)
H(231)	2558	3228	-2265	38(4)
H(232)	2118	3675	-3419	38(4)
H(233)	673	3888	-2696	38(4)
H(234)	652	3464	-4312	38(4)

H(235)	233	3144	-2716	38(4)
H(23C)	180	2502	-3467	38(4)
H(23D)	-586	2796	-3924	38(4)
H(241)	-442	3581	-5265	38(4)
H(242)	-1594	4033	-5787	38(4)
H(243)	-1825	4085	-4045	38(4)
H(244)	-3043	3756	-5508	38(4)
H(245)	-2008	3309	-4046	38(4)
H(251)	-4361	3796	-5347	38(4)
H(252)	-5304	4181	-4600	38(4)
H(253)	-3922	4146	-3112	38(4)
H(254)	-5609	3770	-3321	38(4)
H(255)	-3964	3342	-3202	38(4)
H(25C)	-4862	2785	-3687	38(4)
H(25D)	-4961	2935	-2751	38(4)
H(261)	-6042	3711	-2178	38(4)
H(262)	-5908	3992	-762	38(4)
H(263)	-4041	3996	-651	38(4)
H(264)	-5015	3523	384	38(4)
H(265)	-4152	3203	-924	38(4)
H(26C)	-4928	2602	-795	38(4)
H(26D)	-4282	2709	108	38(4)
H(271)	-4266	3372	1611	38(4)
H(272)	-3151	3626	2740	38(4)
H(273)	-2143	3764	1433	38(4)
H(274)	-1808	3207	2892	38(4)
H(275)	-2297	3017	1051	38(4)
H(27A)	-2476	2340	1535	38(4)
H(27B)	-1484	2480	1868	38(4)

B.6 Torsion angles [deg].

O(174)-C(111)-C(112)-O(112)	58.9(10)
O(115)-C(111)-C(112)-O(112)	-176.8(8)
O(174)-C(111)-C(112)-C(113)	-64.2(10)
O(115)-C(111)-C(112)-C(113)	60.2(10)
O(112)-C(112)-C(113)-O(113)	58.0(11)
C(111)-C(112)-C(113)-O(113)	-179.2(8)
O(112)-C(112)-C(113)-C(114)	178.3(8)
C(111)-C(112)-C(113)-C(114)	-58.9(10)
O(113)-C(113)-C(114)-O(114)	-63.4(10)
C(112)-C(113)-C(114)-O(114)	174.8(8)

O(113)-C(113)-C(114)-C(115)	176.8(8)
C(112)-C(113)-C(114)-C(115)	55.0(11)
O(114)-C(114)-C(115)-O(115)	-171.9(8)
C(113)-C(114)-C(115)-O(115)	-52.1(11)
O(114)-C(114)-C(115)-C(116)	71.3(11)
C(113)-C(114)-C(115)-C(116)	-168.9(9)
O(115)-C(115)-C(116)-O(116)	-60.4(11)
C(114)-C(115)-C(116)-O(116)	60.4(12)
C(113)-C(114)-O(114)-C(121)	131.3(9)
C(115)-C(114)-O(114)-C(121)	-108.4(9)
O(174)-C(111)-O(115)-C(115)	62.5(11)
C(112)-C(111)-O(115)-C(115)	-58.5(10)
C(116)-C(115)-O(115)-C(111)	177.4(8)
C(114)-C(115)-O(115)-C(111)	55.3(11)
C(114)-O(114)-C(121)-O(125)	117.4(9)
C(114)-O(114)-C(121)-C(122)	-121.2(9)
O(125)-C(121)-C(122)-O(122)	-179.1(8)
O(114)-C(121)-C(122)-O(122)	58.5(10)
O(125)-C(121)-C(122)-C(123)	59.7(11)
O(114)-C(121)-C(122)-C(123)	-62.7(11)
O(122)-C(122)-C(123)-O(123)	60.1(11)
C(121)-C(122)-C(123)-O(123)	-179.2(9)
O(122)-C(122)-C(123)-C(124)	-178.5(8)
C(121)-C(122)-C(123)-C(124)	-57.8(11)
O(123)-C(123)-C(124)-O(124)	-67.0(11)
C(122)-C(123)-C(124)-O(124)	173.0(8)
O(123)-C(123)-C(124)-C(125)	175.1(9)
C(122)-C(123)-C(124)-C(125)	55.0(11)
O(124)-C(124)-C(125)-O(125)	-171.8(7)
C(123)-C(124)-C(125)-O(125)	-55.1(11)
O(124)-C(124)-C(125)-C(126)	70.1(10)
C(123)-C(124)-C(125)-C(126)	-173.2(9)
O(125)-C(125)-C(126)-O(126)	70.7(10)
C(124)-C(125)-C(126)-O(126)	-168.2(8)
C(123)-C(124)-O(124)-C(131)	121.9(10)
C(125)-C(124)-O(124)-C(131)	-120.3(9)
O(114)-C(121)-O(125)-C(125)	61.0(11)
C(122)-C(121)-O(125)-C(125)	-60.5(10)
C(124)-C(125)-O(125)-C(121)	59.2(11)
C(126)-C(125)-O(125)-C(121)	-179.3(8)
C(124)-O(124)-C(131)-O(135)	113.4(9)
C(124)-O(124)-C(131)-C(132)	-125.9(9)
O(124)-C(131)-C(132)-O(132)	62.5(12)
O(135)-C(131)-C(132)-O(132)	-176.9(8)

O(124)-C(131)-C(132)-C(133)	-62.4(10)
O(135)-C(131)-C(132)-C(133)	58.2(11)
O(132)-C(132)-C(133)-O(133)	57.7(11)
C(131)-C(132)-C(133)-O(133)	-178.6(8)
O(132)-C(132)-C(133)-C(134)	178.5(8)
C(131)-C(132)-C(133)-C(134)	-57.8(11)
O(133)-C(133)-C(134)-O(134)	-65.6(11)
C(132)-C(133)-C(134)-O(134)	174.6(8)
O(133)-C(133)-C(134)-C(135)	174.7(9)
C(132)-C(133)-C(134)-C(135)	54.9(12)
O(134)-C(134)-C(135)-O(135)	-169.6(8)
C(133)-C(134)-C(135)-O(135)	-52.1(12)
O(134)-C(134)-C(135)-C(136)	71.8(12)
C(133)-C(134)-C(135)-C(136)	-170.7(10)
O(135)-C(135)-C(136)-O(136)	-64.7(12)
C(134)-C(135)-C(136)-O(136)	57.8(13)
C(135)-C(134)-O(134)-C(141)	-106.1(10)
C(133)-C(134)-O(134)-C(141)	132.9(9)
C(134)-C(135)-O(135)-C(131)	53.5(12)
C(136)-C(135)-O(135)-C(131)	176.4(9)
O(124)-C(131)-O(135)-C(135)	62.6(11)
C(132)-C(131)-O(135)-C(135)	-57.4(12)
C(134)-O(134)-C(141)-O(145)	113.6(10)
C(134)-O(134)-C(141)-C(142)	-124.3(10)
O(145)-C(141)-C(142)-O(142)	-179.3(8)
O(134)-C(141)-C(142)-O(142)	60.6(12)
O(145)-C(141)-C(142)-C(143)	53.8(12)
O(134)-C(141)-C(142)-C(143)	-66.2(11)
O(142)-C(142)-C(143)-O(143)	57.0(11)
C(141)-C(142)-C(143)-O(143)	-176.2(9)
O(142)-C(142)-C(143)-C(144)	179.3(9)
C(141)-C(142)-C(143)-C(144)	-53.8(12)
O(143)-C(143)-C(144)-O(144)	-60.9(12)
C(142)-C(143)-C(144)-O(144)	177.2(9)
O(143)-C(143)-C(144)-C(145)	176.7(9)
C(142)-C(143)-C(144)-C(145)	54.8(13)
O(144)-C(144)-C(145)-O(145)	-179.2(8)
C(143)-C(144)-C(145)-O(145)	-56.0(13)
O(144)-C(144)-C(145)-C(146)	63.7(13)
C(143)-C(144)-C(145)-C(146)	-173.1(11)
O(145)-C(145)-C(146)-O(146)	-74.9(14)
C(144)-C(145)-C(146)-O(146)	45.7(16)
C(143)-C(144)-O(144)-C(151)	123.1(10)
C(145)-C(144)-O(144)-C(151)	-113.7(10)

O(134)-C(141)-O(145)-C(145)	62.7(11)
C(142)-C(141)-O(145)-C(145)	-55.6(12)
C(144)-C(145)-O(145)-C(141)	55.5(12)
C(146)-C(145)-O(145)-C(141)	177.4(10)
C(144)-O(144)-C(151)-O(155)	113.2(10)
C(144)-O(144)-C(151)-C(152)	-125.0(9)
O(144)-C(151)-C(152)-O(152)	58.1(11)
O(155)-C(151)-C(152)-O(152)	-178.7(7)
O(144)-C(151)-C(152)-C(153)	-63.1(11)
O(155)-C(151)-C(152)-C(153)	60.1(10)
O(152)-C(152)-C(153)-O(153)	60.9(10)
C(151)-C(152)-C(153)-O(153)	-175.4(8)
O(152)-C(152)-C(153)-C(154)	-179.9(8)
C(151)-C(152)-C(153)-C(154)	-56.3(11)
O(153)-C(153)-C(154)-O(154)	-67.6(9)
C(152)-C(153)-C(154)-O(154)	170.5(8)
O(153)-C(153)-C(154)-C(155)	172.6(8)
C(152)-C(153)-C(154)-C(155)	50.7(10)
O(154)-C(154)-C(155)-O(155)	-166.5(7)
C(153)-C(154)-C(155)-O(155)	-49.0(10)
O(154)-C(154)-C(155)-C(156)	71.5(11)
C(153)-C(154)-C(155)-C(156)	-171.0(9)
O(155)-C(155)-C(156)-O(156)	-68.2(11)
C(154)-C(155)-C(156)-O(156)	55.7(12)
C(155)-C(154)-O(154)-C(161)	-108.9(10)
C(153)-C(154)-O(154)-C(161)	129.9(9)
O(144)-C(151)-O(155)-C(155)	61.7(11)
C(152)-C(151)-O(155)-C(155)	-59.4(10)
C(156)-C(155)-O(155)-C(151)	179.5(9)
C(154)-C(155)-O(155)-C(151)	53.8(11)
C(154)-O(154)-C(161)-O(165)	112.0(10)
C(154)-O(154)-C(161)-C(162)	-126.5(9)
O(165)-C(161)-C(162)-O(162)	-177.4(8)
O(154)-C(161)-C(162)-O(162)	61.1(11)
O(165)-C(161)-C(162)-C(163)	56.5(12)
O(154)-C(161)-C(162)-C(163)	-65.0(11)
O(162)-C(162)-C(163)-O(163)	56.8(12)
C(161)-C(162)-C(163)-O(163)	-177.5(9)
O(162)-C(162)-C(163)-C(164)	179.6(9)
C(161)-C(162)-C(163)-C(164)	-54.7(12)
O(163)-C(163)-C(164)-O(164)	-65.3(10)
C(162)-C(163)-C(164)-O(164)	173.4(9)
O(163)-C(163)-C(164)-C(165)	176.6(9)
C(162)-C(163)-C(164)-C(165)	55.2(12)

O(164)-C(164)-C(165)-O(165)	-174.5(8)
C(163)-C(164)-C(165)-O(165)	-55.6(11)
O(164)-C(164)-C(165)-C(166)	69.6(12)
C(163)-C(164)-C(165)-C(166)	-171.5(11)
O(165)-C(165)-C(166)-O(166)	-66.4(14)
C(164)-C(165)-C(166)-O(166)	52.5(15)
C(163)-C(164)-O(164)-C(171)	122.3(9)
C(165)-C(164)-O(164)-C(171)	-117.8(9)
O(154)-C(161)-O(165)-C(165)	60.3(12)
C(162)-C(161)-O(165)-C(165)	-58.9(12)
C(164)-C(165)-O(165)-C(161)	58.5(12)
C(166)-C(165)-O(165)-C(161)	179.6(11)
C(164)-O(164)-C(171)-O(175)	114.1(9)
C(164)-O(164)-C(171)-C(172)	-123.5(9)
O(175)-C(171)-C(172)-O(172)	-179.0(8)
O(164)-C(171)-C(172)-O(172)	59.8(11)
O(175)-C(171)-C(172)-C(173)	58.7(11)
O(164)-C(171)-C(172)-C(173)	-62.5(10)
O(172)-C(172)-C(173)-O(173)	57.6(10)
C(171)-C(172)-C(173)-O(173)	-178.4(8)
O(172)-C(172)-C(173)-C(174)	178.2(8)
C(171)-C(172)-C(173)-C(174)	-57.8(11)
O(173)-C(173)-C(174)-O(174)	-66.3(10)
C(172)-C(173)-C(174)-O(174)	173.9(8)
O(173)-C(173)-C(174)-C(175)	175.1(8)
C(172)-C(173)-C(174)-C(175)	55.3(11)
O(174)-C(174)-C(175)-O(175)	-172.4(8)
C(173)-C(174)-C(175)-O(175)	-53.9(11)
O(174)-C(174)-C(175)-C(176)	69.2(11)
C(173)-C(174)-C(175)-C(176)	-172.3(9)
O(175)-C(175)-C(176)-O(176)	-63.4(11)
C(174)-C(175)-C(176)-O(176)	56.7(12)
O(115)-C(111)-O(174)-C(174)	115.2(9)
C(112)-C(111)-O(174)-C(174)	-122.8(9)
C(173)-C(174)-O(174)-C(111)	125.9(9)
C(175)-C(174)-O(174)-C(111)	-114.2(9)
C(172)-C(171)-O(175)-C(175)	-59.0(11)
O(164)-C(171)-O(175)-C(175)	61.4(10)
C(176)-C(175)-O(175)-C(171)	176.8(8)
C(174)-C(175)-O(175)-C(171)	55.2(11)
C(6)-C(1)-C(2)-O(7)	-177(3)
C(9)-C(1)-C(2)-O(7)	50(4)
C(6)-C(1)-C(2)-C(3)	-57(4)
C(9)-C(1)-C(2)-C(3)	170(3)

C(1)-C(2)-C(3)-C(4)	71(3)
O(7)-C(2)-C(3)-C(4)	-164(2)
C(2)-C(3)-C(4)-C(5)	-74(3)
C(2)-C(3)-C(4)-C(8)	-176(3)
C(3)-C(4)-C(5)-C(6)	68(3)
C(8)-C(4)-C(5)-C(6)	-179(3)
C(2)-C(1)-C(6)-C(5)	48(4)
C(9)-C(1)-C(6)-C(5)	168(3)
C(4)-C(5)-C(6)-C(1)	-58(4)
C(2)-C(1)-C(9)-C(11)	46(5)
C(6)-C(1)-C(9)-C(11)	-75(5)
C(2)-C(1)-C(9)-C(10)	157(3)
C(6)-C(1)-C(9)-C(10)	36(5)
C(26)-C(21)-C(22)-C(23)	-72(3)
C(29)-C(21)-C(22)-C(23)	170(3)
C(26)-C(21)-C(22)-O(27)	173(3)
C(29)-C(21)-C(22)-O(27)	55(3)
O(27)-C(22)-C(23)-C(24)	179(3)
C(21)-C(22)-C(23)-C(24)	69(4)
C(22)-C(23)-C(24)-C(28)	176(3)
C(22)-C(23)-C(24)-C(25)	-54(4)
C(28)-C(24)-C(25)-C(26)	-175(3)
C(23)-C(24)-C(25)-C(26)	55(4)
C(29)-C(21)-C(26)-C(25)	-178(2)
C(22)-C(21)-C(26)-C(25)	64(3)
C(24)-C(25)-C(26)-C(21)	-67(4)
C(26)-C(21)-C(29)-C(30)	-57(4)
C(22)-C(21)-C(29)-C(30)	51(4)
C(26)-C(21)-C(29)-C(31)	62(4)
C(22)-C(21)-C(29)-C(31)	171(3)
O(215)-C(211)-C(212)-O(212)	-176.3(8)
O(274)-C(211)-C(212)-O(212)	62.8(11)
O(215)-C(211)-C(212)-C(213)	58.5(11)
O(274)-C(211)-C(212)-C(213)	-62.4(11)
O(212)-C(212)-C(213)-O(213)	61.2(11)
C(211)-C(212)-C(213)-O(213)	-174.6(8)
O(212)-C(212)-C(213)-C(214)	179.7(8)
C(211)-C(212)-C(213)-C(214)	-56.1(11)
O(213)-C(213)-C(214)-O(214)	-67.4(11)
C(212)-C(213)-C(214)-O(214)	172.9(8)
O(213)-C(213)-C(214)-C(215)	172.6(8)
C(212)-C(213)-C(214)-C(215)	52.9(12)
O(214)-C(214)-C(215)-O(215)	-170.4(8)
C(213)-C(214)-C(215)-O(215)	-49.4(12)

O(214)-C(214)-C(215)-C(216)	67.9(12)
C(213)-C(214)-C(215)-C(216)	-171.1(10)
O(215)-C(215)-C(216)-O(216)	-71.6(11)
C(214)-C(215)-C(216)-O(216)	52.6(14)
C(213)-C(214)-O(214)-C(221)	125.8(9)
C(215)-C(214)-O(214)-C(221)	-111.4(10)
O(274)-C(211)-O(215)-C(215)	60.7(11)
C(212)-C(211)-O(215)-C(215)	-57.5(11)
C(216)-C(215)-O(215)-C(211)	179.4(9)
C(214)-C(215)-O(215)-C(211)	52.6(11)
C(214)-O(214)-C(221)-O(225)	113.7(10)
C(214)-O(214)-C(221)-C(222)	-121.8(10)
O(225)-C(221)-C(222)-O(222)	-174.4(8)
O(214)-C(221)-C(222)-O(222)	61.0(11)
O(225)-C(221)-C(222)-C(223)	60.0(11)
O(214)-C(221)-C(222)-C(223)	-64.6(11)
O(222)-C(222)-C(223)-O(223)	57.5(11)
C(221)-C(222)-C(223)-O(223)	-176.1(8)
O(222)-C(222)-C(223)-C(224)	176.6(8)
C(221)-C(222)-C(223)-C(224)	-57.0(10)
O(223)-C(223)-C(224)-O(224)	-67.3(10)
C(222)-C(223)-C(224)-O(224)	172.6(8)
O(223)-C(223)-C(224)-C(225)	173.4(8)
C(222)-C(223)-C(224)-C(225)	53.4(10)
O(224)-C(224)-C(225)-O(225)	-171.3(8)
C(223)-C(224)-C(225)-O(225)	-52.3(11)
O(224)-C(224)-C(225)-C(226)	71.1(10)
C(223)-C(224)-C(225)-C(226)	-170.0(8)
O(225)-C(225)-C(226)-O(226)	-68.9(11)
C(224)-C(225)-C(226)-O(226)	52.4(12)
C(223)-C(224)-O(224)-C(231)	132.3(9)
C(225)-C(224)-O(224)-C(231)	-108.0(9)
O(214)-C(221)-O(225)-C(225)	61.8(11)
C(222)-C(221)-O(225)-C(225)	-59.5(11)
C(226)-C(225)-O(225)-C(221)	175.6(8)
C(224)-C(225)-O(225)-C(221)	54.8(10)
C(224)-O(224)-C(231)-O(235)	117.5(9)
C(224)-O(224)-C(231)-C(232)	-121.3(9)
O(224)-C(231)-C(232)-O(232)	60.2(11)
O(235)-C(231)-C(232)-O(232)	-178.8(8)
O(224)-C(231)-C(232)-C(233)	-66.6(11)
O(235)-C(231)-C(232)-C(233)	54.3(12)
O(232)-C(232)-C(233)-O(233)	58.2(12)
C(231)-C(232)-C(233)-O(233)	-177.3(9)

O(232)-C(232)-C(233)-C(234)	-179.5(9)
C(231)-C(232)-C(233)-C(234)	-55.0(12)
O(233)-C(233)-C(234)-O(234)	-61.2(11)
C(232)-C(233)-C(234)-O(234)	175.0(8)
O(233)-C(233)-C(234)-C(235)	179.8(8)
C(232)-C(233)-C(234)-C(235)	56.1(11)
O(234)-C(234)-C(235)-O(235)	-176.2(8)
C(233)-C(234)-C(235)-O(235)	-57.2(11)
O(234)-C(234)-C(235)-C(236)	65.5(12)
C(233)-C(234)-C(235)-C(236)	-175.5(10)
O(235)-C(235)-C(236)-O(236)	-67.4(13)
C(234)-C(235)-C(236)-O(236)	51.4(14)
C(233)-C(234)-O(234)-C(241)	125.3(9)
C(235)-C(234)-O(234)-C(241)	-115.8(9)
O(224)-C(231)-O(235)-C(235)	62.6(11)
C(232)-C(231)-O(235)-C(235)	-57.5(12)
C(236)-C(235)-O(235)-C(231)	179.4(10)
C(234)-C(235)-O(235)-C(231)	59.6(12)
C(234)-O(234)-C(241)-O(245)	111.3(9)

Bibliography

[1] N. Sharon and H. Lis. Lectins as cell recognition molecules. *Science*, 246:227–231, October 1989.

[2] T. A. Waigh. Analysis of the native structure of starch granules with x-ray microfocus diffraction. *Macromolecules*, 30:3813–3820, 1997.

[3] A. M. Stephen and H. F. Zobel. Starch: Structure, analysis and application. In A. M. Stephen, editor, *Food Polysaccharides and their applications*. Marcel Dekker Inc., 1995.

[4] W. Hindricks, G. Büttner, M. Steifa, Ch. Betzel, V. Zabel, B. Pfannemüler, and W. Saenger. An amylose antiparallel double helix at atomic resolution. *Science*, 238:205–208, 1987.

[5] W. Hinrichs and W. Saenger. Crystal and molecular structure of the hexasaccharide complex (p-nirophenyl α-maltohexaoside)$_2$ · Ba(i_3)$_2$ · $27H_2O$. *J. Am. Chem. Soc.*, 112:2789–2796, 1990.

[6] D. A. Rees. *Polysaccharide Shapes*. Outline Studies in biology. Chapman and Hall Ltd., 11 New Fetter Lane, London EC4P 4EE, 1977.

[7] W. L. Doane. USDA research on starch-based biodegradable plastics. *Starch*, 44(8):293–295, 1992.

[8] C. L. Swanson, R. L. Shogren, G. F. Fanta, and S. H. Imam. Starch-plastic materials - preparation, physical properties, and biodegradability (a review of recent USDA research). *Journal of Environmental Polymer Degradation*, 1(2):155–165, 1993.

[9] R. L. Shongren. Effect of moisture content on the melting and subsequent physical aging of cornstarch. *Carbohyd. Poly.*, 19:83–90, 1992.

[10] S. H. D. Hulleman, F. H. P. Janssen, and H. Feil. The role of water during plasticization of native starches. *Polymer*, 39(10):2043–2048, 1998.

[11] A. R. Kirby, S. A. Clark, R. Parker, and A. C. Smith. The deformation and failure behaviour of wheat starch plasticized with water and polyols. *Journal of Materials Science*, 28:5937–5942, 1993.

[12] R. L. Shogren and B. K. Jasberg. Aging properties of extruded high-amylose starch. *Journal of Environmental Polymer Degradation*, 2(2):99–109, 1994.

[13] H. Yu, B. M. Pettitt, and M. Karplus. Aqueous solvation of *n*-methylacetamide conformers: Comparison of simulations and integral equation theories. *J. Am. Chem. Soc.*, 113:2425–2434, 1991.

[14] M. Mezei, P. K. Mehrotra, and D. L. Beveridge. Monte carlo determination of the free energy and internal energy of hydration for the ala dipeptide at 25 deg c. *J. Am. Chem. Soc.*, 107:2239–2245, 1985.

[15] Frieder W. Lichtenthaler and Stephan Immel. On the hydrophobic characteristics of cyclodextrins: Computer-aided visualization of molecular lipophilicity patterns. *Liebigs. Ann.*, pages 27–37, 1996.

[16] J. Szejtli. *Cyclodextrin Technology.* Topics in Inclusion Science. Kluwer Academic Publishers, 1988.

[17] V. Schurig and H. Nowotny. Gas chromatographic separation of enantiomers on cyclodextrin derivatives. *Angew. Chem. Int. Ed. Engl.*, 29(9):939–1076, 1990.

[18] R. Breslow and S. D. Doug. Biomimetic reactions catalyzed by cyclodextrins and their derivatives. *Chem. Rev.*, 98:1997–2011, 1998.

[19] D. Cremer and J. A. Pople. A general definition of ring puckering coordinates. *J. Am. Chem. Soc.*, 97(6):1354–1358, 1975.

[20] J. L. Green and C. A. Agnell. Phase relations and vitrification in saccharide-water solutions and the trehalose anomaly. *J. Phys. Chem.*, 93:2880–2882, 1989.

[21] G. J. Quigley, A. Sarko, and R. H. Marchessault. Crystal and molecular structure of maltose monohydrate. *J. Am. Chem. Soc.*, 92(20):5834–5839, 1970.

[22] B. Casu, M. Reggiani, G.G. Gallo, and A. Vigevani. Hydrogen bonding and conformation of glucose and polyglucoses in dimethylsulphoxide solution. *Tetrahedron*, 22:3061–3068, 1966.

[23] M. St-Jacques, P. R. Sundararajan, and R. H. Marchessault K. J. Taylor. Nuclear magnetic resonance and conformational studies on amylose and model compounds in dimethyl sulfoxide solution. *J. Am. Chem. Soc.*, 98(15):4386–4391, 1976.

[24] P. Dais. Carbon-13 magnetic relaxation and local chain motion of amylose in dimethyl sulfoxide. *Carbohyd. Res.*, 160:73–93, 1987.

[25] H. Schneider, F. Hacket, and V. Rüdiger. NMR studies of cyclodextrins and cyclodextrin complexes. *Chem. Rev.*, 98:1755–1785, 1998.

[26] Y. Nishida, H. Ohrui, and H. Meguro. ^{1}H-NMR studies of (6R)- and (6S)-deuterated D-hexoses: Assignment of the preferred rotamers about C6-C6 bond of D-glucose and D-galactose derivatives in solutions. *Tetrahedron Lett.*, 25(15):1575–1578, 1984.

[27] D. A. Rees and D. Thorn. Polysaccharide conformation. Part 10. Solvent and temperature effects on the optical rotation and conformation of model carbohydrates. *J. C. S. Perkin II*, pages 191–201, 1977.

[28] S. B. Engelsen, C. H. du Penhoat, and S. Pe'rez. Molecular relaxation of sucrose in aqueous solution: How a nanosecond dynamics simulation helps to reconcile NMR data. *J. Phys. Chem.*, 99:13334–13351, 1995.

[29] D. A. Rees. Conformational analysis of polysaccharides. Part v. The characterisation of linkage conformations (chain conformations) by optical rotation at a single wavelength. *J. Chem. Soc. B*, pages 877–887, 1970.

[30] E. S. Stevens and B. K. Sathyanarayana. Potential energy surfaces of cellobiose and maltose in aqueous solution: A new treatment of disaccharide optical rotation. *J. Am. Chem. Soc.*, 111(12):4149–4154, 1989.

[31] Kevin. J. Naidoo and J. W. Brady. Calculation of the Ramachandran potential of mean force for a disaccharide in aqueous solution. *J. Am. Chem. Soc.*, submitted 1998.

[32] J. W. Brady. Molecular dynamics simulations of α-D-glucose in aqueous solution. *J. Am. Chem. Soc.*, 111:5155–5165, 1989.

[33] R. K. Schmidt, M. Karplus, and J. W. Brady. The anomeric equilibrium in d-xylose: Free energy and the role of solvent structuring. *J. Am. Chem. Soc.*, 118:541–546, 1996.

[34] J. W. Brady and R. K. Schmidt. The role of hydrogen bonding in carbohyrates: Molecular dynamics simulations of maltose in aqueous solution. *J. Phys. Chem.*, 97:958–966, 1993.

[35] Q. Lui and J. W. Brady. Model dependence of the aniostropic structuring of solvent water around sugars in molecular dynamics simulations. *J. Phys. Chem. B*, 101(8):1817–1312, 1997.

[36] Q Lui, R. K. Schmidt, B. Teo, P. A. Karplus, and J. W. Brady. Molecular dynamics studies of the hydration of α,α-trehalose. *J. Am. Chem. Soc.*, 119(33):7851–7862, 1997.

[37] Q. Lui and J. W. Brady. Anisotropic solvent structuring in aqueous sugar solutions. *J. Am. Chem. Soc.*, 118:12276–12286, 1996.

[38] S. A. Galema, E. Howard, J. B. F. N. Engberts, and J. P. Grigeria. The effect of stereochemistry upon carbohydrate hydration. A molecular dynamics simulation of β-D-galactopyranose and (α,β)-D-talopyranose. *Carbohyd. Res.*, 265:215–225, 1994.

[39] J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren. Conformational differences between α-cyclodextrin in aqueous solution and in crystalline form. *J. Mol. Biol.*, 203:241–250, 1988.

[40] B. Manunza, S. Deiana, M. Pintore, and C. Gessa. Structure and internal motion of solvated beta-cyclodextrin: a molecular dynamics study. *Journal of Molecular Structure (Theochem)*, 419:133–137, 1997.

[41] Delbert R. Black, Craig G. Parker, S. Scott Zimmerman, and Milton R. Lee. Enantioselective binding of α-pinene and of some cyclohexanetriol derivatives by cyclodextrin hosts: A molecular modelling study. *Journal of Computational Chemistry*, 17(8):931–939, 1996.

[42] Kenny B. Lipkowitz, Greg Pearl, Bob Coner, and Michael A. Peterson. Explanation of where and how enantioselective binding takes place on permethylated β-cyclodextrin, a chiral secondary phase used in gas chromatography. *Journal American Chemical Society*, 119:600–610, 1997.

[43] Mathias Göschl, Serge Crouzy, and Yves Chapron. Molecular dynamics study of an α-cyclodextrin-phosphatidylinositol inclusion complex. *Eur Biophys J*, 24:300–310, 1996.

[44] C. J. Cramer and D. G. Truhlar. Quantum chemical conformational analysis of glucose in aqueous solution. *J. Am. Chem. Soc.*, 115:5745–5753, 1993.

[45] S. N. Ha, L. J. Madsen, and J. W. Brady. Conformational analysis and molecular dynamics simulations of maltose. *Biopolymers*, 27:1927–1952, 1988.

[46] J. W. Brady. Molecular dynamics simulations of carbohydrate molecules. *Ad. Biophys. Chem.*, pages 155–202, 1990.

[47] S. N. Ha, L. J. Madsen, and J. W. Brady. *Biopolymers*, 27:1927–1952, 1988.

[48] Hanoch Senderowitz, Carol Parish, and W. Clark Still. Carbohydrates: United atom amber parametrization of pyranoses and simulations yielding anomeric free energies. *J. Am. Chem. Soc*, 118:2078–2086, 1996.

[49] Honoch Senderowitz and W. Clark Still. A quantum mechanically derived all-atom force field for pyranose oligosaccharides. amber parameters and free energy simulations. *J. Org. Chem.*, 62:1427–1438, 1997.

[50] K. Ott and B. Meyer. Molecular dynamics simulations of maltose in water. *Carbohyd. Res*, 281:11–34, 1996.

[51] K. Ott and B. Meyer. Parameterization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations. *J. Comp. Chem.*, 17(8):1068–1084, 1996.

[52] K. Ueda and J. W. Brady. Molecular dynamics simulations of carrabiose. *Biopolymers*, 41:323–330, 1997.

[53] K. B. Lipkowitz. Applications of computational chemistry to the study of cyclodextrins. *Chem. Rev.*, 98:1829–1873, 1998.

[54] A. R. Leach. *Molecular Modelling. Principles and Applications.* Addison Wesley Longman Limited, 1996.

[55] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Oxford University Press, 1987.

[56] W. F. van Gunsteren and H. J. C. Berendsen. Computer simulation of molecular dynamics: Methodology, applications and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.*, 29:992–1023, 1990.

[57] M. Karplus and G. A. Petsko. Molecular dynamics simulations in biology. *Nature*, 347:631–639, October 1990.

[58] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4(2):187–217, 1983.

[59] T. M. Glennon, Y. Zheng, S. M. le Grand, B. A. Shutzberg, and K. M. Jnr. Merz. A force field for monosaccharides and $(1 \rightarrow 4)$ linked polysaccharides. *J. of Comp. Chem.*, 15(9):1019–1040, 1994.

[60] Robert J. Woods, Raymond A. Dwek, and Christopher J. Edge. Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1. GLYCAM_93 parameter development. *J. Phys. Chem*, 99:3832–3846, 1995.

[61] W.F van Gunsteren. GROMOS: Groningen molecular simulation program package. Technical report, University of Groningen, The Netherlands, 1987.

[62] G. Liang and J. W. Brady. CHARMM carbohydrate force field.

[63] M. L. C. E. Kouwijzer and Grootenhuis. P. D. J. Parametrization and application of CHEAT95 an extended atom force field for hydrated (oligo)saccharides. *J. Phys. Chem.*, 99:13426–13436, 1995.

[64] W. L. Jorgensen. Transferable intermolecular potential functions for water, alcohols and ehters. application to liquid water. *J. Am. Chem. Soc.*, 103:335–340, 1981.

[65] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulations of liquid water. *J. Chem. Phys*, 79(926), 1983.

[66] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91:6269–6271, 1987.

[67] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72(4):2384–2393, 1980.

[68] K. Tasaki, S. McDonald, and J. W. Brady. Observations concerning the treatment of long-range interactions in molecular dynamics simulations. *Journal of Computational Chemistry*, 14(3):278–284, 1993.

[69] W. F. van Gunsteren and H. J. C. Berendsen. *Mol. Phys.*, 34:1311–27, 1977.

[70] P. J. Rossky and M. Karplus. Solvation. A molecular dynamics study of a dipeptide in water. *J. Am. Chem. Soc.*, 101(8):1913–1937, 1979.

[71] J. Schnitker and P. Mausbach. Vornoi polyhedra analysis and order-disorder transition of a tetrahedrally ordered liquid. *Z. angew. Math. Mech.*, 70(4):T 266 – T269, 1990.

[72] Jing-Ping Shih, Sheh-Yi Sheu, and Chung-Yuan Mou. A Vornoi polyhedra analysis of structures of liquid water. *J. Chem. Phys*, 100(3):2202–2212, February 1994.

[73] Gautum R. Desiraju. Designer crystals: intermolecular interactions, network structures and supramolecular synthons. *Chemical Communications*, pages 1475–1482, 1997.

[74] J. W. Brady and K. J. Naidoo. Adiabatic map for maltose. 1998.

[75] K. J. Naidoo and J. W. Brady. The application of simulated annealing to the conformational analysis of disaccharides. *Chem. Phys.*, 224:261–273, 1997.

[76] R. K. Schmidt, B. Teo, and J. W. Brady. Use of umbrella sampling in the calculation of the potential of mean force for maltose in vacuum from molecular dynamics simulations. *J. Phys. Chem*, 99:11339–11343, 1995.

[77] F. Takusagawa and R. A. Jacobson. The crystal and molecular structure of α-maltose. *Acta Cryst.*, B34:213–218, 1978.

[78] M. E. Gress and G. A. Jeffrey. A neutron diffraction refinement of the crystal structure of β-maltose monohydrate. *Acta Cryst.*, B33:2490–2495, 1977.

[79] A. S. Shashkov, G. M. Lipkind, and N. K. Kochetkov. Nuclear Overhauser effects for methyl β-maltoside and the conformational states of maltose in aqueous solution. *Carbohyd. Res.*, 147:175–182, 1986.

[80] A. S. Shashkov, G. M. Lipkind, and N. K. Kochetkov. Nuclear Overhauser effects for methyl β-maltoside and the conformatinal states of maltose in aqueous solution. *Carbohydrate Research*, 147(175-182), 1986.

[81] K. J. Naidoo and J. W. Brady. Calculation of the Ramachandran potential of mean force for a disaccharide in aqueous solution. *J. Am. Chem. Soc.*, 1998.

[82] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983.

[83] A Stillinger, F. H. ad Rahman. *J. Chem. Phys.*, 60:1545, 1974.

[84] W. Saenger, J. Jacob, K. Gessler, T. Steiner, D. Hoffmann, H. Sanbe, K. Koizumi, S. M. Smith, and T. Takaha. Structures of the common cyclodextrins and their larger analogues - beyond the doughnut. *Chem. Rev.*, 98:1787–1802, 1998.

[85] W. Saenger. Cyclodextrin inclusion compounds in research and industry. *Angew. Chem. Int. Ed. Engl.*, 19:344–362, 1980.

[86] M. R. Caira, V. J. Griffith, L. R.assimbeni, and B. van Oudtshoorn. X-ray structures of 1:1 complexes of (L)-menthol with β-cyclodextrin and permethylated β-cyclodextrin. *Supramolecular Chemistry*, 7:119–124, 1996.

[87] G. Kocza and A. Hung. *Chemical Abstracts*, 117(448), 1992.

[88] G. M. Sheldrick. SHELX-97 program suite. Technical report, University of Göttingen, 1998.

[89] K. Linder and W. Saenger. Crystal and molecular structure of cyclohepta-amylose dodecahydrate. *Carbohyd. Res.*, 99:103–115, 1982.

[90] B. Klar, B. Hingerty, and W. Saenger. *Acta Crys.,Sect.B*, 36:1154, 1980.

[91] F.H. Allen and O. Kennard. 3D search and research using the Cambridge Structural Database. *Chemical Design Automation News*, 8(1):1 & 31–37, 1993.