STEREOCHEMISTRY OF POLYSACCHARIDE CHAIN CONFORMATIONS

1. INTRODUCTION

This chapter deals with the study on the conformation of polysaccharide chains from the stereochemical point of view. The main results obtained are: (a) the sugar residue, which is the monomer unit in the case of polysaccharide chains, has a rigid conformation; (b) the chain conformations in cellulose and chitin are very nearly the same as that found in the disaccharide cellobiose. The parameters describing the standard configuration of the sugar residue have also been worked out similar to the Pauling-Corey parameters for the peptide residue.

Polysaccharides are components of almost all living beings. They are high molecular weight carbohydrates in which the sugar residues are glycosidically linked to form a long chain polymer with the elimination of water according to the empirical equation

$$nC_6H_{12}O_6 \rightarrow [C_6H_{10}O_5]^n + (n-1)H_2O$$

In most cases the chain is linear and unbranched. For example, the polysaccharide chains in cellulose, chitin
and amyllose are linear and unbranched while those in amylepectin are branched. In the case of the most widely known polysaccharide cellulose, the monomer unit is $\beta$-D-glucose and the linkage is of the $\beta-D\, 1\rightarrow 4$ type. In the case of the polysaccharide chitin, the monomer unit is N-acetyl-$\beta$-D-glucosamine and the linkage is similar to that found in cellulose. But in the case of amyllose, the monomer unit is $\alpha$-D-glucose and the linkage is of the $\alpha-D\, 1\rightarrow 4$ type, while in laminaran, it is essentially $\beta-D\, 1\rightarrow 3$ linkage. Throughout this study, the interest has been mainly on the $\beta-D\, 1\rightarrow 4$ linkage and hence the results are directly applicable only to those polysaccharides in which the monomer units are linked glycosidically as the $\beta-1\rightarrow 4$ type. The two well-known polysaccharides cellulose and chitin, in fact, belong to this type. In addition xylan (monomer unit D-xylopyranose), mannan (monomer unit D-mannopyranose) and galactan, which is one of the components of pectin (monomer unit galactopyranose) are some of the other polysaccharides having $\beta-D\, 1\rightarrow 4$ linkage.

The problem of theoretically working out the possible conformations of a linear (unbranched) polysaccharide chain consists of two steps:
(i) finding out the configuration of the monomer unit i.e., the bond distances and bond angles between atoms in it; and

(ii) finding out the relative conformations of two linked units.

The approach made for studying the polysaccharide chain conformations is essentially similar to that made in the case of polypeptides, described in Chapters 1 and 2. While the problem of polypeptide chain conformation has been studied by quite a large number of workers for a very long time (c.f. section 2 of Chapter 1), such a study does not seem to have been made before in the case of polysaccharides. The first step, namely that of finding out the configuration of the backbone of the glucopyranose ring has been worked out and is described in the next section. By backbone is meant the six atoms, five carbons and one oxygen which form a closed ring. The extension to step (b) is also analogous to the case of polypeptides and is described in sections 3 and 4.

In the case of polysaccharides, the monomer units are linked at oxygen atoms, to be designated as the bridge oxygens (we shall denote the oxygen in the six membered ring as the ring oxygen) as shown below:
The linkage being of the 1→4 type, the two sugar residues linked at the bridge oxygen atom are capable of free rotation about the single bonds O₁—O₁ and O₄—O₄ respectively. Analogous to the case of polypeptides, these two rotations can conveniently be represented by two dihedral angles φ, φ', these being measured from a suitable initial conformation φ = φ' = 0°. Thus a polysaccharide chain can be completely specified by the parameters (φ₁, φ₁'), (φ₂, φ₂') ...... at the various bridge oxygen atoms. When the values of the pair (φ, φ') is the same at every bridge oxygen, then a regular helical structure would result in which each unit can be obtained from the previous one by a screw operation, integral or non-integral. Such a helix can be specified by two parameters, n, the number of residues per turn of the helix and h the height of the unit along the axis of the helix.

Just as in the case of polypeptides, here also restrictions arise on the values of φ and φ' due to the
possible short contacts between atoms in the adjacent residues. Also, there may be certain conformations which are preferred because of the formation of good hydrogen bonds to stabilise the chain. These points are considered in detail in sections 3 and 4 below.

2. STANDARD PARAMETERS FOR A SUGAR RESIDUE

Much of the great progress that has been made in the study of protein structure has been due to the very accurate knowledge of the amide residue which forms the backbone for the polypeptide chains. In this section the results of a similar study made in the case of sugars are described.

The basic unit in the case of polysaccharides is the monosaccharide unit or the sugar residue, e.g., the glucopyranose residue in cellulose. The backbone of the residue consists of five carbon atoms designated as C₁, C₂, C₃, C₄ and C₅ and an oxygen atom designated as O₅ forming a closed ring. The positions of the groups attached to the carbon atoms of the ring differs from one sugar to another. In the case of β-D-glucose, to the atoms C₁, C₂, C₃ and C₄ are attached H and OH groups and to the atom C₅, an H and a CH₂OH group are attached.
This is shown in Figure 4.1, wherein the designation of the various atoms in the usual way is also indicated. Unlike the peptide residue, which is planar, the configuration of the gluco-pyranose ring is not unique. It can take up either the so called boat form or the chair form. This point has been examined by Reeves (1950) and he has shown that the chair form is more likely to occur than the boat form. If a model is built of the six membered ring of the pyranose residue in the chair form, it would appear that there is a certain amount of flexibility for the ring. This is however misleading, for a careful analysis of the available crystal structure data on simple sugars which contain the pyranose ring shows that the ring skeleton is remarkably uniform in its configuration, in spite of the differences in the substituents at the various positions. The sugars whose structures are known to a good degree of accuracy are the monosaccharides $\beta$-D-glucose (Perrier, 1960), $\alpha$-D-glucose (McDonald and Beevers, 1952), $\alpha$-D-rhamnose monohydrate (McGeachin and Beevers, 1957), $\beta$-Arabinose (Hordvik, 1961), the two independent determinations of the structure of the disaccharide cellobiose (Jacobson et al., 1961; Brown, 1962a) and other related compounds like barium glucose ortho phosphate
β-D-GLUCOSE

THE SCHEMATIC REPRESENTATION OF β-D-GLUCOSE. THE DESIGNATION OF THE VARIOUS ATOMS IS ALSO INDICATED.
Among these, the structure of the disaccharide cellobiosc contains two \( \beta -D-glucose \) units linked by the \( 1 \rightarrow 4 \) type of linkage and the compound barium glucose orthophosphate contains the monosaccharide glucose as the \( \beta \)-isomer. For the other substances, the backbone atoms are the same, while the nature and positions of the substituents (axial or equatorial) differ from one another. This is shown in Figure 4.2 as compared with \( \beta -D-glucose \). With the usual designation, \( \beta -D-glucose \) can be represented as 1e2e3e4e5e. That is, all the OH groups and the \( \text{CH}_2\text{OH} \) group are occupying the equatorial positions in \( \beta -D-glucose \). By referring to Figure 4.2, it can be seen that the OH group in position 1 in the case of \( \alpha -D-glucose \), in positions 1 and 2 in the case of \( \alpha -D-rhamnose \) and in positions 1 and 4 of \( \beta -arabinose \) occupy axial positions and so this point has to be specially considered while deducing the standard coordinates for the sugar residue with respect to a suitably chosen coordinate system.

The crystal structure data of these sugars can therefore be used to arrive at a standard configuration for the glucopyranose ring. The configuration observed
THE NATURE AND SUBSTITUENTS AT THE FIVE CARBON BACKBONE ATOMS IN THE VARIOUS SUGARS USED IN THE STUDY.
in all these sugars were reduced to a uniform coordinate system as follows. The atom C₁ (Figure 4.1) was taken to be the origin of coordinates and C₀C₅ as the X-axis with C₃C₅O₁ forming the XY-plane. The Z-axis was taken perpendicular to this plane, so that XYZ formed a right-handed system of coordinates.

By means of certain special techniques, involving the cyclographic projection, a modification of the stereographic projection method, described in Appendix 1, the coordinates of the atoms reported in the various structures were transformed in the chosen coordinate system.

The coordinates of the atoms for the different structures are given in Table 4.1 along with the mean coordinates of the various atoms with their probable errors. Table 4.1 contains in addition to the backbone atoms C₁, C₂, C₃, C₄, C₅ and O₅, the coordinates of the side group atoms O₁, O₂, O₃, O₄ and C₆ attached to the carbon atoms C₁ to C₅ respectively. The coordinates of these atoms correspond to the β-D-glucopyranose unit. The atom O₆ has not been included in the Table since the position of O₆ in a crystal need not be specific owing to the possible free rotation of O₆—O₆ about the bond C₅—C₆. This point is considered further towards the end of this section.
### Table 4.1

The coordinates of the atoms in different structures containing the glucopyranose ring. The origin is at atom C₁. The x-axis is along C₃-C₅ and C₃-C₅ is the xy-plane. The mean coordinates of the atoms with the probable errors are given in the last column.

<table>
<thead>
<tr>
<th>Atom</th>
<th>Structure 1</th>
<th>Structure 2</th>
<th>Structure 3</th>
<th>Structure 4</th>
<th>Structure 5</th>
<th>Structure 6</th>
<th>Structure 7</th>
<th>Structure 8</th>
<th>Structure 9</th>
<th>Structure 10</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.45</td>
<td>1.30</td>
<td>1.36</td>
<td>1.39</td>
<td>1.31</td>
<td>X</td>
<td>1.32</td>
<td>1.27</td>
<td>1.34</td>
<td>1.28</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.07</td>
<td>0.05</td>
<td>0.09</td>
<td>0.10</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.07</td>
<td>0.00</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1.24</td>
<td>1.22</td>
<td>1.32</td>
<td>1.23</td>
<td>1.24</td>
<td>1.23</td>
<td>1.24</td>
<td>1.25</td>
<td>1.22</td>
<td>1.23</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>2.47</td>
<td>2.54</td>
<td>2.56</td>
<td>2.56</td>
<td>2.53</td>
<td>2.48</td>
<td>2.51</td>
<td>2.49</td>
<td>2.41</td>
<td>2.50</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>2.48</td>
<td>2.47</td>
<td>2.41</td>
<td>2.46</td>
<td>2.47</td>
<td>2.46</td>
<td>2.43</td>
<td>2.47</td>
<td>2.37</td>
<td>2.46</td>
<td>2.45</td>
</tr>
<tr>
<td></td>
<td>1.52</td>
<td>1.37</td>
<td>1.40</td>
<td>1.52</td>
<td>1.39</td>
<td>1.38</td>
<td>1.33</td>
<td>1.48</td>
<td>1.43</td>
<td>1.46</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>1.24</td>
<td>1.23</td>
<td>1.39</td>
<td>1.28</td>
<td>1.26</td>
<td>1.23</td>
<td>1.19</td>
<td>1.24</td>
<td>1.22</td>
<td>1.23</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>-1.05</td>
<td>-1.09</td>
<td>-1.05</td>
<td>-1.05</td>
<td>-1.06</td>
<td>-1.04</td>
<td>-1.16</td>
<td>-1.04</td>
<td>-1.07</td>
<td>-1.06</td>
<td>-1.06</td>
</tr>
<tr>
<td></td>
<td>-1.17</td>
<td>-1.20</td>
<td>-1.19</td>
<td>-1.16</td>
<td>-1.02</td>
<td>-1.15</td>
<td>-1.19</td>
<td>-1.18</td>
<td>-1.02</td>
<td>-1.15</td>
<td>-1.15</td>
</tr>
<tr>
<td></td>
<td>3.76</td>
<td>3.85</td>
<td>3.86</td>
<td>3.77</td>
<td>3.78</td>
<td>3.80</td>
<td>3.83</td>
<td>-</td>
<td>3.70</td>
<td>-</td>
<td>3.80</td>
</tr>
<tr>
<td></td>
<td>2.07</td>
<td>2.08</td>
<td>2.08</td>
<td>2.08</td>
<td>2.06</td>
<td>2.05</td>
<td>2.15</td>
<td>2.10</td>
<td>2.07</td>
<td>2.04</td>
<td>2.08</td>
</tr>
<tr>
<td></td>
<td>1.47</td>
<td>1.41</td>
<td>1.40</td>
<td>1.49</td>
<td>1.37</td>
<td>1.46</td>
<td>1.47</td>
<td>1.44</td>
<td>1.42</td>
<td>1.44</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

...Contd
<table>
<thead>
<tr>
<th>Atom</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4</td>
<td>-0.73</td>
<td>-0.84</td>
<td>-0.74</td>
<td>-0.74</td>
<td>-0.79</td>
<td>-0.77</td>
<td>-0.73</td>
<td>-0.78</td>
<td>-0.73</td>
<td>-0.76±0.03</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>O5</td>
<td>1.27</td>
<td>1.33</td>
<td>1.38</td>
<td>1.31</td>
<td>1.31</td>
<td>1.32</td>
<td>1.41</td>
<td>1.37</td>
<td>1.31</td>
<td>1.33±0.03</td>
<td></td>
</tr>
<tr>
<td>O1</td>
<td>3.33</td>
<td>3.38</td>
<td>3.42</td>
<td>3.37</td>
<td>3.34</td>
<td>3.24</td>
<td>3.38</td>
<td>3.39</td>
<td>3.33</td>
<td>3.35±0.03</td>
<td></td>
</tr>
<tr>
<td>O4</td>
<td>-2.15</td>
<td>-2.04</td>
<td>-2.02</td>
<td>-1.99</td>
<td>-2.03</td>
<td>-2.09</td>
<td>-2.05</td>
<td>-2.11</td>
<td>-2.07</td>
<td>-2.06±0.03</td>
<td></td>
</tr>
<tr>
<td>O2</td>
<td>2.25</td>
<td>2.21</td>
<td>2.06</td>
<td>2.15</td>
<td>2.17</td>
<td>2.19</td>
<td>2.19</td>
<td>2.22</td>
<td>2.11</td>
<td>2.17±0.04</td>
<td></td>
</tr>
<tr>
<td>O3</td>
<td>-0.61</td>
<td>-0.61</td>
<td>-0.59</td>
<td>-0.54</td>
<td>-0.75</td>
<td>-0.72</td>
<td>-0.51</td>
<td>-0.62</td>
<td>-0.65</td>
<td>-0.62±0.05</td>
<td></td>
</tr>
<tr>
<td>O6</td>
<td>-0.73</td>
<td>-0.71</td>
<td>-0.52</td>
<td>-0.61</td>
<td>-0.74</td>
<td>-0.66</td>
<td>-0.67</td>
<td>-</td>
<td>-0.59</td>
<td>-</td>
<td>-0.64±0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>s</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>O2</td>
<td>0.37</td>
<td>0.55</td>
<td>0.36</td>
<td>0.53</td>
<td>0.55</td>
<td>0.53</td>
<td>0.44</td>
<td>0.51</td>
<td>0.39</td>
<td>0.47</td>
<td>0.47±0.06</td>
</tr>
<tr>
<td>O3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>O4</td>
<td>0.44</td>
<td>0.27</td>
<td>0.37</td>
<td>0.38</td>
<td>0.36</td>
<td>0.48</td>
<td>0.50</td>
<td>0.49</td>
<td>0.52</td>
<td>0.38</td>
<td>0.42±0.06</td>
</tr>
<tr>
<td>O5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>O6</td>
<td>0.52</td>
<td>0.41</td>
<td>0.50</td>
<td>0.57</td>
<td>0.50</td>
<td>0.46</td>
<td>0.35</td>
<td>0.46</td>
<td>0.53</td>
<td>0.55</td>
<td>0.48±0.04</td>
</tr>
<tr>
<td>O1</td>
<td>0.65</td>
<td>0.47</td>
<td>0.35</td>
<td>0.52</td>
<td>0.52</td>
<td>0.52</td>
<td>0.60</td>
<td>0.52</td>
<td>0.63</td>
<td>0.49</td>
<td>0.53±0.05</td>
</tr>
</tbody>
</table>

...contd
<table>
<thead>
<tr>
<th>Atom</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_4$</td>
<td>0.04</td>
<td>-0.25</td>
<td>-0.32</td>
<td>-0.37</td>
<td>-0.32</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.06</td>
<td>0.07</td>
<td>-0.30</td>
<td>-0.13±0.14</td>
</tr>
<tr>
<td>$O_2$</td>
<td>-0.03</td>
<td>0.24</td>
<td>-0.23</td>
<td>-0.13</td>
<td>0.19</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
<td>-0.35</td>
<td>-0.05</td>
<td>-0.00±0.10</td>
</tr>
<tr>
<td>$O_3$</td>
<td>0.53</td>
<td>0.50</td>
<td>0.57</td>
<td>0.52</td>
<td>0.39</td>
<td>0.72</td>
<td>0.42</td>
<td>0.49</td>
<td>0.38</td>
<td>0.56</td>
<td>0.50±0.06</td>
</tr>
<tr>
<td>$O_6$</td>
<td>0.40</td>
<td>0.39</td>
<td>0.65</td>
<td>0.61</td>
<td>0.64</td>
<td>0.38</td>
<td>0.48</td>
<td>-</td>
<td>0.37</td>
<td>-</td>
<td>0.49±0.09</td>
</tr>
</tbody>
</table>

1. $\beta$-D-Glucose (Ferrier, 1960)
2.3. Cellobiose (Jacobson et al., 1961)
4.5. Cellobiose (Brown, 1962a)
6. $\alpha$-D-Glucose (McDonald and Beevers, 1952)
7. $\alpha$-D-Rhamnose monohydrate (McGeachin and Beevers, 1957)
8. $\beta$-Arabinose (Hordvik, 1961)
10. $\beta$-Methyl Xyloside (Brown, 1962b)
The probable error $p$ given in the last column of Table 4.1 was calculated from the expression

$$p = 0.8455 a$$

where $a$ is the average deviation given by

$$a = \frac{\sum_{i=1}^{n} |\Delta y_i|}{n}$$

$\Delta y$ being the deviation of the corresponding coordinate ($\Delta x, \Delta y, \Delta z$) and $n$ is the total number of values.

It will be seen from the probable errors given in Table 4.1, that the coordinates of the atoms forming the backbone (ring) are highly specific. The maximum probable error is only $\pm 0.06 \text{ Å}$ in the $y$ coordinate i.e., normal to the plane of the ring. Further, the three atoms $C_2$, $C_4$ and $O_5$ also lie in a plane very nearly parallel to the $XY$-plane containing the atoms $C_1$, $C_3$ and $C_5$. The probable errors in the coordinates of the atoms $O_4$ and $O_2$ are fairly large, $\pm 0.14 \text{ Å}$ and $\pm 0.10 \text{ Å}$ respectively while the probable error in the other two coordinates $x$ and $y$ is only about $\pm 0.04 \text{ Å}$. Thus these two atoms can have a slight wagging motion of the order of $\pm 0.1 \text{ Å}$ in a direction perpendicular to the plane of the ring. The bond lengths and bond angles
calculated from the mean coordinates given in Table 4.1 are given in Table 4.2

As mentioned earlier in this section, the oxygens in position 1 in the case of \( \alpha \)-D-glucose, in positions 1 and 2 in the case of \( \alpha \)-D-rhamnose monohydrate and in positions 1 and 4 in the case of \( \beta \)-arabinose occupy axial positions. So the coordinates of these oxygen atoms in these cases, even after conversion into the chosen coordinate system cannot be used as such in the evaluation of standard coordinates. Hence in these cases the position of the oxygen atom that will correspond to the \( \beta \)-D-glycopyranose unit, has been determined as follows.

By means of the stereographic projection technique, the direction of the bond corresponding to the equatorial position was determined at the ring carbon under consideration. This equatorial direction has been taken to be the one which is very nearly tetrahedral to the other three directions at the carbon atom (e.g., \( C_1O_2 \), \( C_1O_5 \) and \( C_1O_1 \) corresponding to the axial position of \( O_1 \) given in the structure, at the carbon atom \( C_1 \)). The position of the oxygen atom was fixed at the same distance from the carbon atom as the axial oxygen but in
**Table 4.2**

Bond lengths and bond angles in the gluco trunose ring calculated from the mean coordinates of the atoms given in Table 4.1

<table>
<thead>
<tr>
<th>Bond</th>
<th>Length (Å)</th>
<th>Bond angle</th>
<th>Value (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁-O₂</td>
<td>1.52</td>
<td>O₅-O₁-O₂</td>
<td>110</td>
</tr>
<tr>
<td>O₂-O₃</td>
<td>1.51</td>
<td>O₁-O₂-O₃</td>
<td>109</td>
</tr>
<tr>
<td>O₃-O₄</td>
<td>1.52</td>
<td>O₂-O₃-O₄</td>
<td>111</td>
</tr>
<tr>
<td>O₄-O₅</td>
<td>1.53</td>
<td>O₃-O₄-O₅</td>
<td>110</td>
</tr>
<tr>
<td>O₅-O₆</td>
<td>1.42</td>
<td>O₄-O₅-O₆</td>
<td>110</td>
</tr>
<tr>
<td>O₁-O₅</td>
<td>1.42</td>
<td>O₅-O₅-O₁</td>
<td>114</td>
</tr>
<tr>
<td>O₁-O₄</td>
<td>1.38</td>
<td>O₅-O₁-O₄</td>
<td>108</td>
</tr>
<tr>
<td>O₄-O₂</td>
<td>1.41</td>
<td>O₁-O₄-O₂</td>
<td>109</td>
</tr>
<tr>
<td>O₂-O₃</td>
<td>1.41</td>
<td>O₁-O₂-O₃</td>
<td>110</td>
</tr>
<tr>
<td>O₃-O₂</td>
<td>1.40</td>
<td>O₂-O₂-O₃</td>
<td>112</td>
</tr>
<tr>
<td>O₅-O₆</td>
<td>1.52</td>
<td>O₂-O₃-O₅</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O₃-O₅-O₄</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O₃-O₄-O₅</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O₄-O₅-O₁</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O₄-O₅-O₆</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O₆-O₅-O₅</td>
<td>109</td>
</tr>
</tbody>
</table>
the equatorial direction and the coordinates of this atoms in these position only have been included in Table 4.1

The study described in this Chapter was made chiefly during the early months of 1962 and the results were presented at the International Symposium on Protein Structure and Crystallography held in Madras during January, 1963. This has since been published in "Aspects of Protein Structure" edited by G.N. Ramachandran

Since then, six more compounds containing the glucopyranose unit have been reported. These are $\alpha$-D-glucose monohydrate (Killeen et al., 1962), sinigrin ($1-\beta$-D-thioglucooside; Waser and Watson, 1963), casimidine (Ram et al., 1963), glucosamine hydrobromide (Chandrasekaran and Mallikarjuman, 1964), in addition to the refinement of the structure of $\beta$-D-glucose (Perrier, 1963) and the refinement of the structure of sucrose by neutron diffraction (Brown and Levy, 1963). In this connection, it may be mentioned that the coordinates of the atoms in the glucopyranose unit of the di-saccharide sucrose ($\alpha$-D-glucopyranosyl-$\beta$-D-fructofuranoside) by Beevers et al. (1952) were converted to
our system. But since the positions of the atoms were far from the specificity shown by the coordinates of the atoms in other sugars, they were not included for calculating the mean coordinates. A similar case is that of the crystal structure of α-D-glucosamine hydrobromide (Cox and Jeffrey, 1939) in which the refinement was not complete. This structure has been refined recently in this laboratory (Chandrasekaran and Mallikarjunan, 1964). Hence an interesting study would be to compare the coordinates of the atoms constituting the glucopyranose unit in these cases, converted to our system of coordinates, with the mean values given in Table 4.1. This is given in Table 4.3 in which columns 1 and 2 give the coordinates of the atoms in the preliminary and refined structures of β-D-glucose (Ferrier, 1960, 1963), columns 3 and 4 those of sucrose (Beavers et al., 1952; Brown and Levy, 1963), columns 5 and 6 those of glucosamine hydrobromide (Cox and Jeffrey, 1939; Chandrasekaran and Mallikarjunan, 1964). The coordinates of the atoms in the structure of α-D-glucose monohydrate (Kilean et al., 1962), sinigrin (Vaser and Watson, 1963), and esamidine (Haman et al., 1963) are given under columns 7, 8 and 9 of Table 4.3. The last column simply gives the mean value as obtained in Table 4.1. The coordinates of the atoms in sucrose were kindly made
### Table 4.3

The coordinates of atoms in earlier and refined structures of sugars and also for some more structures containing glucopyranose ring, in addition to those listed in Table 4.1

<table>
<thead>
<tr>
<th>Atom</th>
<th>Structure</th>
<th>1°</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Mean*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>1.45</td>
<td>1.36</td>
<td>1.21</td>
<td>1.37</td>
<td>1.56</td>
<td>1.36</td>
<td>1.39</td>
<td>1.39</td>
<td>1.37</td>
<td>1.34</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>0.03</td>
<td>0.02</td>
<td>0.10</td>
<td>0.09</td>
<td>0.28</td>
<td>0.06</td>
<td>0.05</td>
<td>0.09</td>
<td>0.10</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>1.24</td>
<td>1.24</td>
<td>1.26</td>
<td>1.20</td>
<td>1.36</td>
<td>1.26</td>
<td>1.28</td>
<td>1.29</td>
<td>1.29</td>
<td>1.24</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>2.47</td>
<td>2.49</td>
<td>2.52</td>
<td>2.49</td>
<td>2.67</td>
<td>2.52</td>
<td>2.50</td>
<td>2.59</td>
<td>2.60</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>2.48</td>
<td>2.44</td>
<td>2.38</td>
<td>2.49</td>
<td>2.65</td>
<td>2.49</td>
<td>2.50</td>
<td>2.47</td>
<td>2.50</td>
<td>2.45</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>1.52</td>
<td>1.41</td>
<td>1.45</td>
<td>1.45</td>
<td>-</td>
<td>1.44</td>
<td>1.45</td>
<td>-</td>
<td>1.52</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>1.24</td>
<td>1.28</td>
<td>1.13</td>
<td>1.25</td>
<td>1.22</td>
<td>1.36</td>
<td>1.28</td>
<td>1.34</td>
<td>1.27</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>-1.03</td>
<td>-1.04</td>
<td>-</td>
<td>-</td>
<td>-0.86</td>
<td>-1.07</td>
<td>-1.04</td>
<td>-1.05</td>
<td>-1.04</td>
<td>-1.06</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>-1.17</td>
<td>-1.17</td>
<td>-1.19</td>
<td>-1.25</td>
<td>-1.08</td>
<td>-1.13</td>
<td>-1.15</td>
<td>-1.12</td>
<td>-1.07</td>
<td>-1.15</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>3.76</td>
<td>3.81</td>
<td>3.60</td>
<td>3.76</td>
<td>3.58</td>
<td>3.79</td>
<td>3.86</td>
<td>3.81</td>
<td>3.85</td>
<td>3.80</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>2.07</td>
<td>2.13</td>
<td>2.16</td>
<td>2.14</td>
<td>2.08</td>
<td>2.12</td>
<td>2.10</td>
<td>2.02</td>
<td>2.11</td>
<td>2.08</td>
<td></td>
</tr>
</tbody>
</table>

..... contd
<table>
<thead>
<tr>
<th>Atom</th>
<th>1*</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Mean*</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁</td>
<td>0.65</td>
<td>0.61</td>
<td>0.61</td>
<td>0.53</td>
<td>-</td>
<td>0.56</td>
<td>0.56</td>
<td>-</td>
<td>0.57</td>
<td>0.53</td>
</tr>
<tr>
<td>O₄</td>
<td>0.04</td>
<td>0.01</td>
<td>0.13</td>
<td>0.18</td>
<td>0.02</td>
<td>0.01</td>
<td>0.04</td>
<td>-0.17</td>
<td>-0.11</td>
<td>-0.13</td>
</tr>
<tr>
<td>O₂</td>
<td>-0.03</td>
<td>0.23</td>
<td>-</td>
<td>-</td>
<td>0.05</td>
<td>0.03</td>
<td>0.07</td>
<td>-0.08</td>
<td>-0.13</td>
<td>0.00</td>
</tr>
<tr>
<td>O₃</td>
<td>0.53</td>
<td>0.61</td>
<td>0.92</td>
<td>0.55</td>
<td>0.63</td>
<td>0.54</td>
<td>0.73</td>
<td>0.57</td>
<td>0.70</td>
<td>0.50</td>
</tr>
<tr>
<td>O₆</td>
<td>0.40</td>
<td>0.43</td>
<td>-1.01</td>
<td>0.53</td>
<td>0.57</td>
<td>0.61</td>
<td>0.49</td>
<td>0.57</td>
<td>0.50</td>
<td>0.49</td>
</tr>
</tbody>
</table>

* The coordinates are the same as given in Table 4.1

1) β -D-Glucose (Perrier, 1960)
2) β -D-Glucose (Perrier, 1963)
3) β -D-Glucosamine hydrobromide (Cox and Jeffrey, 1939)
4) β -D-Glucosamine hydrobromide (Chandrasekaran and Mallikarjuna, 1964)
5) Sucrose (Beavers et al., 1952)
6) Sucrose (Brown and Levy, 1964)
7) α -D-Glucose monohydrate (Killeen et al., 1962)
8) Sinigrin (Waser and Watson, 1963)
9) Casimidine (Raman et al. 1963)
made available to Professor Ramachandran by Dr.
and these have been used in this study.

The main feature of Table 4.3 is that in those
cases in which the structure has been refined, the co-
ordinates of the atoms have come closer to the standard
cordinates obtained in this study. Especially, in the
case of glucosamine hydrobromide, the $y$ and $z$ coordinates
of the atoms $O_3$ and $C_6$ have changed from $-0.31$, $-0.44$,
$-0.92$ and $-1.01$ to $-0.53$, $-0.62$, $0.55$ and $0.53$ and the
latter are closer to our standard values $-0.62$, $-0.64$,
$0.50$ and $0.59$. So also, the $z$ coordinates of the back-
bone atoms $C_2$ and $C_5$ have improved, but not the $x$ coorde-
nate of the atom $C_4$. In the case of $\beta$-D-glucose, there
is a general improvement in the agreement between the
refined values and the mean values compared to the ear-
lier data. In the case of sucrose, again, there is
better agreement with the refined value, especially in
the refined $y$ coordinates of $O_3$ and $C_6$, in which the
earlier values are $-0.14$ and $-1.20$, refined values are
$-0.67$ and $-0.59$ which are nearer to the mean values $-0.62$
and $-0.64$. In the remaining three cases, $\alpha$-D-glucose
monohydrate, casimidine and sinigrin, it can be seen that
the coordinates obtained are fairly close to the mean value obtained in this study.

In the case of the glucopyranose residue, the oxygen atom O₆ is capable of free rotation about the single bond C₅—O₆ and hence the position of O₆ need not be specific in the actual structures. But a study of the positions of this oxygen atom in the various structures show that this is not so and in fact there are two distinct positions which these atoms assume in the actual structure. This is given in Table 4.4 along with the mean coordinates and the probable error. In the structure of the polysaccharide chitin proposed by Carlström (Carlström, 1957) and described further in the next chapter, the position of O₆ is the one which corresponds to the latter group containing cellulbiose.

The polysaccharide chains are built up from their monomer units, by means of linkages through the bridge oxygen atoms. In the case of cellulose and chitin, as mentioned earlier, the linkage is of the β 1→ 4 type and so the angle C₅–O₆ at the bridge oxygen atom plays the same role as the angle τ at the α-carbon atom in the case of polypeptides. Among the various substances studied so far, only in the two
<table>
<thead>
<tr>
<th>Structure</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.87</td>
<td>-0.73</td>
<td>1.95</td>
</tr>
<tr>
<td>2</td>
<td>3.91</td>
<td>-0.67</td>
<td>1.86</td>
</tr>
<tr>
<td>3</td>
<td>3.83</td>
<td>-0.60</td>
<td>1.99</td>
</tr>
<tr>
<td>4</td>
<td>3.82</td>
<td>-0.42</td>
<td>2.01</td>
</tr>
<tr>
<td>4</td>
<td>3.77</td>
<td>-0.80</td>
<td>1.78</td>
</tr>
<tr>
<td>6</td>
<td>3.85</td>
<td>-0.78</td>
<td>1.90</td>
</tr>
<tr>
<td>7</td>
<td>4.85</td>
<td>-0.14</td>
<td>-0.17</td>
</tr>
<tr>
<td>8</td>
<td>4.90</td>
<td>0.21</td>
<td>0.40</td>
</tr>
<tr>
<td>9</td>
<td>4.85</td>
<td>0.37</td>
<td>0.18</td>
</tr>
<tr>
<td>10</td>
<td>4.98</td>
<td>0.00</td>
<td>-0.04</td>
</tr>
<tr>
<td>11</td>
<td>4.89</td>
<td>0.15</td>
<td>0.50</td>
</tr>
<tr>
<td>12</td>
<td>4.95</td>
<td>-0.01</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Mean (1) \(3.84 \pm 0.03\) \(-0.67 \pm 0.09\) \(1.92 \pm 0.06\)
Mean (2) \(4.90 \pm 0.03\) \(0.10 \pm 0.12\) \(0.16 \pm 0.17\)

1. \(\alpha\)-D-Glucose monohydrate (Killea et al., 1962)
2. \(\beta\)-D-Glucose (Ferrier, 1963)
3. \(\beta\)-D-Glucosamine HBr (Chandrasekaran and Mallikarjuna, 1964)
4. Sucrose (Brown and Levy, 1964)
5. Barium Glucose orthophosphate (Kartha, 1962)
6. Cadmium (Raman et al., 1963)
7. \(\alpha\)-D-Glucose (McDonald and Beever, 1952)
8,9. Cellobiose (Jacobson et al., 1961)
10,11. Cellobiose (Brown, 1962a)
12. Sinigrin (Waser and Watson, 1963)
disaccharides cellobiose and sucrose are there two units linked through an oxygen atom. The values of the angle at the bridge oxygen atom in these structures are 117.5°, 116.8° from the two determinations on cellobiose (Jacobson et al., 1961; Brown, 1962a) and 118.3° in sucrose* (Bevers et al., 1952) so that a mean value of 117.5° was taken to be the standard value for this angle. In addition, for a pair of linked sugar residues, another important parameter is the distance $O_1...O_4$. The actual distances found in the various sugars containing $\beta$-D-glucopyranose unit is given in Table 4.5 along with the mean value.

3. CONFORMATION OF THE POLYSACCHARIDE CHAIN

The study of the polysaccharide chain conformation has been made in a manner similar to that of the polypeptide chain conformation. In this case, the residues are linked through oxygen atoms and the conformation of the polysaccharide chain as a whole depends on the relative orientation the the two sugar residues.

In the recent structure determination of sucrose by neutron diffraction (Bown and Levy, 1963) the value of this angle is 114.6°.
**Table 4.5**

$O_1 \cdots O_4$ distances found in various sugars containing $\beta$-D-glucopyranose unit

<table>
<thead>
<tr>
<th>Structure</th>
<th>Distance (Å)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$-D-glucose</td>
<td>5.51</td>
<td>1</td>
</tr>
<tr>
<td>Cellobiose</td>
<td>5.46</td>
<td>2</td>
</tr>
<tr>
<td>Cellobiose</td>
<td>5.49</td>
<td>2</td>
</tr>
<tr>
<td>Cellobiose</td>
<td>5.46</td>
<td>3</td>
</tr>
<tr>
<td>Cellobiose</td>
<td>5.42</td>
<td>3</td>
</tr>
<tr>
<td>Barium glucose orthophosphate</td>
<td>5.44</td>
<td>4</td>
</tr>
<tr>
<td>Casimidine</td>
<td>5.47</td>
<td>5</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>5.47</strong></td>
<td></td>
</tr>
</tbody>
</table>

(1) Ferrier, 1963  
(2) Jacobson et al. (1961)  
(3) Brown, (1962a)  
(4) Kartha (1962)  
(5) Raman et al. (1963)
units linked at the bridge oxygen atom. Designating the atoms of the second sugar residue by a prime, as shown in Figure 4.3, the two residues are capable of free rotation about the two single bonds \( C_1'O \) and \( C_4'O \). The angles of rotation of the two residues can be denoted by \( \phi \) and \( \phi' \) respectively, the angles being measured from an initial standard conformation \( \phi = \phi' = 0^\circ \). This conformation \((\phi, \phi') = (0^\circ, 0^\circ)\) has been taken to be the one in which the other two terminal glycosidic oxygens i.e., \( O_4 \) and \( O'_4 \) lie in the plane \( C_1OC'_4 \) as shown in Figure 4.3. Just as in the case of polypeptides, the plane \( C_1OC'_4 \) was chosen for the initial conformation since it contains the axes of rotation and hence is not affected by the rotations \( \phi \) or \( \phi' \).

Analogous to the intramolecular \( N-H \ldots O-C \) bond of the polypeptide structures, in this case also, a hydrogen bond can be formed between the atom \( O_5 \) of one residue and the ring oxygen atom \( O_5' \) of the adjacent residue. The importance of the formation of such a bond has been dealt with by Carlstrom (1962) in discussing the bent chain structure of \( \alpha \)-chitin and by Jones (1960) in connection with the structure of cellulose. In addition, such a hydrogen bond has been observed
Figure 4.3: Two sugar residues linked through the bridge oxygen atom O. The initial conformation \((O^*, O')\) is shown along with the angles \((\phi, \phi')\) which define a general conformation.
in the two independent structure determinations of the disaccharide celllobiose (Jacobson et al., 1961; Brown, 1962a).

In our notation, the hydrogen bond occurs around \((30^\circ, 210^\circ)\) and so the interest was mainly concentrated on conformations around this region. In fact, the study was done to cover a range of \(0^\circ\) to \(\pm 30^\circ\) of \(\beta\) and \(150^\circ\) to \(240^\circ\) of \(\beta'\). However, all the conformations in this range may not be permitted due to the possible short contacts between the atoms in the adjacent residues. Using the normally allowed and outer limit minimum contact distances given in Table 1.2 of Chapter 1, the contact distances between the atom \(C_1\) and those of the second residue and between the atom \(C_4'\) and those of the first residue were examined. Since the atoms \(C_3, C_5\) and \(C_4\) of the first residue are fairly away from \(C_4'\) and similarly since the atoms \(C_1', C_2'\) and \(O_5'\) of the second residue are also fairly away from \(C_1\), the contact distances which are of importance are \(C_1C_3, C_1C_5', C_4O_5',\) and \(C_4'C_2\). The boundaries of the normally allowed and the outer limit regions thus obtained are shown in Figure 4.4.
FIG. 4.4. THE NORMALLY ALLOWED AND THE OUTER LIMIT REGION AROUND THE CONFORMATION (-30°, 210°) (MARKED BY A CIRCLE) FOR A PAIR OF SUGAR RESIDUES. THE REGION WHERE A REASONABLY GOOD HYDROGEN BOND CAN BE FORMED IS THE ONE BETWEEN THE CONTOURS MARKED 2.5 AND 2.8. THE COUNTOURS OF \( \eta \) AND \( \eta' \) ARE ALSO SHOWN.
The intramolecular hydrogen bond that is of importance is of the type $O_3^2 - H \cdots O_5$. Making use of techniques involving stereographic projections, the hydrogen bond distance around this region was evaluated. Taking 2.5 Å and 2.8 Å as the lower and upper limits for hydrogen bond length, the region within which this bond is well formed is also shown in Figure 4.4, as the region between the two curves marked 2.5 and 2.8.

In a polysaccharide chain, if the conformation $(\phi, \phi')$ is the same at every linkage, then a regular helical structure results and this can be specified by two parameters $n$, the number of residues per turn, which need not be integral, and $h$, the unit translation of the residue along the helix axis. By using the matrix method described in section 3 of Chapter 2, these helical parameters were evaluated in this range. The contours of constant $n$ and constant $h$ are also given in Figure 4.4.

4. DISCUSSION OF RESULTS

The configurations likely for polysaccharide chains are those which occur within the strip outlined by the hydrogen bond lengths 2.5 Å and 2.8 Å and the outer limit boundary. It will be noticed from Figure 4.4,
that the number of residues per turn \( n \) in the helix can vary between 2 and 3 for these conformations. If the chains have to form a regular lattice, then \( n \) has to be either 2 or 3. On the other hand, it is quite likely that \( n \) may be equal to 2.5 and the structure may have 5 residues in 2 turns. In this case the repeat distance will obviously be 5 times the unit height \( h \). It can also be seen from Figure 4.4, that the residue height varies only very little over the whole range, and is around 5 Å. The above results are in good agreement with observation. Thus, Meyer (1950) has divided the various crystalline modifications of cellulose into three families, having an observed repeat spacing along the fibre axis of about 10 Å (2 x 5 Å), 15 Å (3 x 5 Å) and 25 Å (5 x 5 Å). Further the actual repeat spacings found for cellulose (Meyer, 1950) and chitin (Carlström, 1962) are close to 10.3 Å i.e., for these cases \( n = 2 \) and \( h = 5.15 \) Å. This agrees perfectly with \( \phi = -30^\circ \), \( \phi' = 210^\circ \) (marked by a circle in Figure 4.4). Although the contact \( C_1 \cdots C_3 \) is only in the outer limit (3.10 Å), the occurrence of the strong intrachain hydrogen bond \( O_3^v \cdots \cdots O_5 \) in its neighbourhood offsets this. In fact, the natural conformation in cellulobiose itself is also
close to this, which again shows that the cellulose and chitin structures should be based on the celllobiose conformation. The attempt to refine the structure of chitin on this basis has been made and is described in next chapter.

SUMMARY

The stereochemistry of polysaccharide chain conformation is analogous to the polypeptide case. The two main steps in the evaluation of possible chain conformations are (1) the determination of the dimensions of the monomer unit which is a sugar residue and (2) the conformations at the linkage of two such residues. The crystal structure data on various sugars containing glucopyranose unit have been analysed to arrive at a standard configuration for the monomer unit. The coordinates of the various atoms given in the structural reports are converted into a suitable system of coordinate axes fixed in the glucopyranose unit. The coordinates of the backbone atoms, forming the ring, are found to be highly specific while the side groups attached to the backbone have a slight wagging motion perpendicular to the plane of the ring. Representing the rotation of the two linked sugar residues by two dihedral angles $\phi$ and $\phi'$,
similar to the case of polypeptides, the stereochemically allowed conformations have been worked out for a limited range of conformations close to that observed in structures like cellobiose. The parameters defining the polysaccharide helix have also been worked out in this range. The important result that is obtained in this study is that the conformation at each bridge oxygen atom in the case of cellulose and chitin is very nearly the same as in the disaccharide cellobiose.