CHAPTER I

INTRODUCTION TO X-RAY CRYSTAL STRUCTURE DETERMINATION

1.1 INTRODUCTION

Structure analysis by X-ray diffraction technique is one of the most powerful methods of unambiguously determining the complete three-dimensional structure of crystalline substances. Besides confirming the connectivity and stereochemistry of the molecule, X-ray diffraction method is uniquely capable of providing precise information concerning bond lengths, bond angles, torsion angles and molecular dimensions. With the advent of direct methods, the widespread availability of electronic computers and the efficient program packages coupled with the development of computer controlled automatic diffractometers to measure X-ray intensities, X-ray crystallography has emerged as the main source of unambiguous information on crystal and molecular structure of compounds. In order to understand the nature of chemical bond, the functions of molecules in biological contexts and also to understand the mechanics and dynamics of reactions, structure analysis by X-ray diffraction techniques is the unique method of determining the complete three-dimensional representation of the atoms in a crystal. Hence the crystal structure determination by X-ray diffraction was undertaken to obtain a detailed picture of the contents of the crystals presented in the thesis. This chapter briefly outlines the methods adopted for studying single-crystal X-ray structure analysis. For detailed information, one may refer to a number of books contributed by various authors (Stout & Jenson, 1968; Sherwood, 1976; Dunitz, 1979; Giacovazzo, 1980; Woolfson, 1980; Hauptman, 1988; Giacovazzo et al., 1992).
1.2 UNIT CELL PARAMETERS AND INTENSITY DATA COLLECTION

Determination of unit cell parameters and the three-dimensional intensity data collection were carried out using a four circle Enraf-Nonius diffractometer with \( \omega/2\theta \) scan mode and CCD area detector with \( \phi \) and \( \omega \) scan mode and monochromatic radiation (MoK\(_\alpha\)). A crystal of suitable size was mounted and unit cell parameters were obtained and refined by least-squares fit.

1.3 STRUCTURE SOLUTION

The structure factor is represented by the following equation

\[
F_{hkl} = \sum_{j=1}^{N} f_j \exp \left[ 2\pi i (hx_j + ky_j + lz_j) \right]
\]

(1.1)

where \( f_j \) is the scattering factor for the \( j^{th} \) atom. \( F_{hkl} \) can be represented by

\[
F_{hkl} = |F_{hkl}| e^{i\phi_{hkl}}
\]

(1.2)

where \( |F_{hkl}| \) is the structure amplitude and \( \phi_{hkl} \) is the associated phase. The structure amplitude can be obtained directly from the square root of the observed intensity. But there is no direct method to find out the associated phase values experimentally. In order to locate the position of atoms one usually computes the electron density at various parts. The position, at which the electron density is maximum, gives the position of an atom.

The general expression for electron density function \( \rho(x, y, z) \) is given by

\[
\rho(x,y,z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} F_{hkl} \exp[-2\pi i (hx+ky+lz)]
\]

(1.3)
In order to get the structure it appears from the above equation as if we simply have to map $\rho(x,y,z)$ and locate the maxima in it, but the process is not so straightforward; for to sum of series $\rho(x,y,z)$, we have to find out the complex structure factor $F_{hkl}$.

Using equation 1.2, $\rho(x,y,z)$ becomes

$$
\rho(x,y,z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} |F_{hkl}| \exp \left[ i\phi_{hkl} - 2\pi i (hx + ky + lz) \right] \quad (1.4)
$$

where $V$ is the volume of the unit cell and $x$, $y$, $z$ are the fractional coordinates in the unit cell. To evaluate equation (1.4) one needs the exact values of the complex quantity $F_{hkl}$ in both magnitudes and phases. Our experimental data however gives us only the real quantity $|F_{hkl}|$. Thus the phase $\phi_{hkl}$ is necessary if one wants to compute the electron density to locate the positions of atoms. This is called “phase problem” in crystallography.

Several methods are available to solve the phase problem and some of them are:

i) Direct methods

ii) Heavy atom method

iii) Isomorphous replacement method

iv) Anomalous dispersion method

The above methods can be successfully applied to locate the approximate positions of all the atoms (trial structure of a molecule) in the unit cell.

1.3.1 DIRECT METHODS

The structure determination of all the compounds presented in this thesis has been carried out using Direct methods.

Direct methods are mathematical techniques used to derive the phase information directly from the observed intensities in arriving at a final map to locate the atoms. These methods determine the values of some linear combination of phases from which individual phases can be determined.
The assumptions and various steps involved in the above direct methods procedures are summarized:

**Basic assumptions**

The direct methods work on three important basic assumptions. They are:

(i) Positivity (ii) atomicity and (iii) randomness.

The electron density is assumed to be positive or zero everywhere. In Direct methods, a real crystal with continuous electron density is replaced by an idealized one, the unit cell of which consists of N discrete, equal non-vibrating point atoms i.e., the structure factors $F_{hkl}$ is replaced by the normalized structure factors $E_{hkl}$. In the initial stage, atoms are assumed to be the point type (|F| now being replaced by |E| the normalized structure factors, which do not vary with $\sin \theta/\lambda$). Although this method has some limitations in its applications when the number of atoms in a molecule is (non hydrogen) < 100, for smaller structures the success rate of this method is close to 100%. With user intervention even more complex structures could be solved. But this needs the user’s intuitions and expertise.

The various steps involved in the direct methods are:

**Step I** Conversion of observed structure factors to normalized structure factors $|E_{hkl}|$ which are independent of $\theta$.

**Step II** Setting up of phase relations triple phase relations (triplets) and four phase relations (quartets).

**Step III** Selection of a few reflections, the phases of which are assigned a prior.

**Step IV** Phase propagation and refinement using tangent formula (Karle & Hauptman, 1956).

**Step V** Calculation of best phase sets and expressing the reliability of the phases in terms of Combined Figure of Merit (CFOM)
Step VI  Calculation of electron density map (E-map) with $|E_{hkl}|$ as the Fourier coefficient.

Patterson Method

Another method of intensity distribution analysis of the diffraction pattern comes from the Patterson map. This is based on the Fourier series where the indices and $F^2$ value of each diffracted beam are the needed quantities that are directly derivable from the experimental procedure. The Patterson function $P(u, v, w)$ is defined as

$$P(u,v,w) = \frac{1}{V} \sum_{hkl} |F_{hkl}|^2 \cos [2\pi (hu + kv + lw)]$$

(1.5)

The function $P(u, v, w)$ is always centrosymmetric and defines a map, not of atomic positions, but of inter atomic distances plotted from one point: origin. A peak in the Patterson map at $u, v, w$ implies that there are two atoms in the crystal structure at $x_1, y_1, z_1$ and $x_2, y_2, z_2$ such as $u = x_2 - x_1$, $v = y_2 - y_1$, $w = z_2 - z_1$. The height of the peak is proportional to the product of the number of electrons in each of the two atoms involved.

Direct methods in practice

When part of the structure is known, the unknown part of the structure can be solved by using Direct methods. For phase extension and for the refinement of input phases and amplitudes, the difference structure factors, phased by the partial structure are used as input to a weighted tangent-refinement process. The method is referred to as DIRDIF (Beurskens et al., 1998 & Beurskens et al., 1999). This method is useful if the known part is only marginally sufficient to solve the structure. It can also be used if the known atoms lie in special or pseudo-special position (origin ambiguity), or if, for non-centrosymmetric structures, the known atoms form a centrosymmetric arrangement (enantiomorph ambiguity). The observed structure amplitudes and positional
parameters of the known atoms are used by the computer program DIRDIF, to produce a greatly improved electron density map.

### 1.4 STRUCTURE REFINEMENT

Structure refinement consists of obtaining the best fit between a set of observed measurements and quantities calculated from a model postulated to explain them. Differences between the observed and the calculated values can arise due to random errors (statistical fluctuations) in the observations and defects in the model (systematic errors). The errors introduced into the calculation of the electron density function have inaccuracies firstly in the magnitude and phases of the structure factors $F_{hkl}$. So the trial structure has to be refined to get a more accurate set of structure factors. Numbers of structure refinement processes are in vogue, the full-matrix least-squares refinement technique is widely used in determining the small molecular structure determination. The least-squares refinement consists of using the squares of the differences between the observed and calculated values as a measure of their disagreement and one would expect that the more closely the calculated and observed structure factors agree, then the more closely does the trial structure represent the true structure. The discrepancy between the observed and calculated structure factors is therefore and indication of the degree of confidence, which may be ascribed to any given trial structure.

The best agreement is obtained for

$$D = \sum_{i=1}^{m} w_i (|F_o| - |F_c|)^2$$  \hspace{1cm} (1.6)

The residual $R$ or the Discrepancy index is given as

$$R = \frac{\sum_{i} |F_o| - |F_c|}{\sum_{i} |F_o|}$$  \hspace{1cm} (1.7)
The weighted R value (wR2) used by the SHELXL97 (Sheldrick, 1997a) program is

\[
wR2 = \{ \sum w (|F_o|^2 - |F_c|^2)^2 \} / \sum w (|F_o|^2)^2 \}^{1/2}
\]

(1.8)

After completing the refinement of the structure, it can be interpreted and a number of computing programs may be very useful for calculating bond lengths, bond angles, torsion angles, asymmetry parameters and stereo views.

1.5 MOLECULAR CONFORMATIONS

The internal parameters of a molecule are the parameters that characterize the molecular conformation and hence they are known as conformational parameters. They are bond length, bond angle and torsion angle. Of these three, the torsion angle is the most important one, since, a variation of the torsion angles leads to different conformations of a molecule. The torsion angle is defined as follows:

The torsion angle otherwise known as the dihedral angle of four atoms A,B,C,D with a chemical bond between AB, BC,CD, (Figure 1.1) is defined as the angle between the two planes through A, B, C and B, C, D. As it varies from -180 to 180º, its sign is defined according to the convention of Klyne and Prelog (1960).

The initial conformation corresponding to \( \tau = 0 \) is that in which all the atoms are coplanar and the end atoms A and D are cis with respect to BC. In any orientation of CD, the torsion angle \( \tau \) is given by the rotation in degrees it has undergone from the \( (\tau = 0) \) original position. The torsion angle is considered positive when it is measured clockwise from the front substituent D and negative when it is measured anti-clockwise. The \( \tau(D-C-B-A) \) has the same sign and magnitude as \( \tau(A-B-C-D) \) (Luger, 1980).
1.6 SYMMETRIES IN SIX MEMBERED RINGS

A quantitative evaluation of the conformation of rings of any size from which a comparative analysis can be easily made is obtainable from a mathematical combination of torsion angles. This analysis is based upon consideration of torsion angles. This analysis is based upon consideration of approximate symmetry possessed by most rings (Duax et al., 1976). The two types of symmetry that define the ring conformation are mirror planes perpendicular to the dominant ring plane and two fold axes lying in the ring plane. Either of these symmetry elements may be present at any of the three locations described in Figure 1.2. The location of symmetry in a ring depends on the number of atoms comprising that ring. In rings containing even number of atoms, symmetry element may pass through two ring atoms located directly across the ring from each other (Fig 1.2 a, b) or bisect two opposite ring bonds (Fig 1.2 c, d). The most commonly observed conformations of six-membered rings are shown in Figure 1.3, along with the mirror and two fold rotational symmetries for the six membered rings (Duax et al., 1976).

Six-membered rings possess twelve potential symmetry elements that must be considered in order to determine the ring’s conformation. The planar ring is highly symmetric and contains all possible symmetry elements (a mirror plane and a twofold axis of symmetry at each of the six possible locations). The chair conformation possesses the next highest symmetry, having three mirror planes of symmetry and three two fold axis of symmetry. The boat and twist-boat conformation each have two mutually perpendicular symmetry elements. The sofa and half-chair conformations each have only a single symmetry element (Duax et al., 1976).

1.7 CONFORMATION IN RING STRUCTURES

The conformational analysis of interest in crystallography varies from molecule to molecule. It mainly depends on the type of molecule. If the molecules contain rings, the planarity or otherwise of the ring is an important
conformational feature. The non-planar rings are described in terms of well-known objects in common use like the chair, the boat, the envelope, the sofa etc.

**Six-membered rings**

The six-membered rings assume varieties of conformations some of which are shown in Figure 1.3 (Duax et al., 1976). Simplest way of finding the conformations is to look at the torsion angle within the ring. In the case of chair conformation, the torsion angles will be alternating between $+60^\circ$ and $-60^\circ$. For a boat conformation, two of the non-consecutive torsion angles will be zero and these will be separated by the torsion angles around $+60^\circ$ and $-60^\circ$. In the half chair form, two consecutive torsion angles will be around zero. In a sofa conformation, 5 out of 6 atoms will be in one plane. The conformation of the rings can also be identified using the puckering degrees of freedom. In the six-membered rings there are three puckering degrees of freedom. These are described by a single amplitude-phase pair ($q_2$, $\phi_2$) and a single puckering coordinates may be replaced by a “spherical polar set” ($Q$, $\theta$, $\phi$) where $Q$ is the total puckering amplitude and $\theta$ is angle ($0 \leq \theta \leq \pi$) such that $q_2 = Q \sin \theta$ and $q_3 = Q \cos \theta$.

This coordinate system permits the mapping of all types of puckering (for a given amplitude $Q$) on the surface of a sphere (Figure 1.4). The polar positions ($\theta = 0$ or $180^\circ$) correspond to a chair conformation with $q_2 = 0$ and $q_3 = \pm Q$. The positions on the equator of the sphere (Figure 1.4) have $\theta = 90^\circ$ so that $q_3 = 0$ and $q_2 = Q$. As the phase angle $\phi$ varies the conformation traverses a series of six boat conformations ($\phi = 0, 60, 120, 180, 240, 300^\circ$) and six twist boat conformations ($\phi = 30, 90, 150, 210, 270, 330^\circ$). (Cremer and Pople 1975; Duax et al; Nardelli 1983).

In the case of rings molecule where the ring is known to be planar, the deviations of the atoms from the least squares plane is an important feature of molecular conformation and provides information about the effects of substitution on the ring planarity. Two possible representations are employed:
(i) the equation to the least-squares plane and the deviation of the atoms are given; (ii) the torsion angles about the various bonds of the ring are listed out. For completely planar molecules, all the torsion angles will be zero. Hence, the deviation of the observed torsion angle from this ideal value gives the amount of non-planarity present in the structure. The algebraic sum of the torsion angles will however be zero.

**Asymmetry Parameter**

The asymmetry parameters define the conformation of any ring relative to ideal conformations like chair, boat, twist etc., (Fig 1.3). The asymmetry parameters of non-ideal systems measure the degree of departure from ideal symmetry at any of the possible symmetry locations. Related torsion angles are compared in a way that will result in a value of zero if the symmetry in a question is present. Mirror related torsion angles are of some magnitude but of opposite in sign, and such torsion angles are of same magnitude but of opposite in sign, and such torsion angles are compared by addition. The two fold related torsions are directly related (same magnitude and sign) and are compared by subtraction. The root means square synthesis of these individual discrepancies then yields a measure of the ring’s deviation from ideal symmetry at the symmetry location in question. The two equations used to calculate the asymmetry parameters (Duax et al., 1976) are:

\[ \Delta C_S = \left[ \frac{\sum_{i=1}^{m} (\phi_i + \phi'_i)^2}{m} \right]^{1/2} \quad \text{(mirror asymmetry)} \]

\[ \Delta C_2 = \left[ \frac{\sum_{i=1}^{m} (\phi_i - \phi'_i)^2}{m} \right]^{1/2} \quad \text{(two-fold symmetry)} \]

where \( \phi_i, \phi'_i \)’s are symmetry related torsions and \( m \) is number of individual comparisons described in Figure 1.2a-d. The nomenclature defining the location of an asymmetry parameter includes the symmetry symbol and the lower numbered atom or bond intersected by the symmetry element and is schematically represented in Figure 1.5 (Duax et al., 1976).
1.8 HYDROGEN BONDING GEOMETRY

In the crystalline state, the molecules are stabilized by intra and intermolecular interactions like hydrogen bonds, van der Waal’s forces and some short contacts between the two atoms. Hydrogen bonding is the specific type of non-bonded interactions between two electronegative atoms, (donor and acceptor) where the hydrogen atom is bonded to them. The hydrogen bond can be characterized geometrically by the parameters $r$ and $d$, as shown in the Fig 1.6. The additional parameter $R$ is referred to as the hydrogen bond lengths. The geometrical criterion for the hydrogen bond between the groups D-H and A is that the distance $H…A$ be shorter than the van der Waals approach (Olovsson & Jonsson, 1976; Ramanathan & Chidambaram, 1978), as follows:

$$d_{H…A} < r_H + r_A$$

where D-H is the donor group and A is the acceptor group. $r_D$ and $r_A$ are the van der Waals radii for the donor and accepter atoms. $r_H = 1.20 \, \text{Å}$, $r_C = 1.75 \, \text{Å}$, $r_N = 1.55 \, \text{Å}$ and $r_O = 1.50 \, \text{Å}$ are the van der Waals radii used by the author (Taylor & Kennard, 1982), to calculate the H-bonding geometries in this thesis.

In crystal structure studies, the angle D-H…A is conventionally known as the hydrogen bond angle.

Very short hydrogen bonds

If the length of O…O hydrogen bond is around 2.40 Å the H-atom deviates significantly from the center of the O…O bond where the H-atoms is dynamically distributed between alternate positions very close to the center of the O…O bond (Olovsson & Jonsson, 1976). Another non-bonded interaction, which is of interest in recent times, is the C-H…O hydrogen bonds. The existence of C-H…O type hydrogen bonds in crystals is evident from the study of Taylor & Kennard (1982), Desiraju (1991), Desiraju and Steiner (1999). The
ability of a C-H group to act as a proton donor depends on the hybridization [C (sp)-H > C (sp$^2$)-H > C (sp$^3$)-H] and increases with the number of adjacent electron withdrawing groups (Steiner, 1996 & 1997). The crystal structures presented in this thesis are found to have O-H…O, N-H…O and C-H…O types of hydrogen bonds.

1.9 COMPUTER PROGRAMS USED BY THE AUTHOR

Data collection was done using CAD-4 (Enraf-Nonius, 1989), SMART (Bruker, 2004), CrysAlis CCD (Oxford Diffraction, 2007), COLLECT (Nonius, 2000) and APEX2 (Bruker, 2004) computer systems. Cell refinement was done using SAINT (Bruker, 2004), CrysAlis CCD (Oxford Diffraction, 2007), DENZO-SMN (Otwinowski & Minor, 1997), APEX2 (Bruker, 2004) and CAD-4 (Enraf-Nonius, 1989) computer systems. Data reduction was done using CrysAlis RED (Oxford Diffraction, 2007), SAINT (Bruker, 2004), DENZO-SMN and SCALE PACK (Otwinowski & Minor, 1997), SAINT NT (Bruker, 2004) and XCAD4 (Harms & wocadlo, 1997). The crystal structures of the compounds presented in the thesis were solved by the programs SHELXS97 (Sheldrick, 1997b), DIRDIF99 (Beurskens et al., 1999). Least-squares refinements of these compounds were carried out by means of the programs SHELXL97 (Sheldrick, 1997a). The programs PARST (Nardelli, 1996), PLATON (Spek, 2003b) and SHELXL97 (Sheldrick, 1997a) were used for geometry calculations and molecular diagrams were prepared using PLUTON (Spek, 2003a) and CHEMDRAW (Furst, 2000) software. The thermal ellipsoids plots were drawn employing the programs ORTEP-3 (Farrugia, 1997).
Figure 1.1
Figure 1.2 (a-d): The signs of torsion angles in six membered rings describe the symmetrical positioning of atoms related by symmetry operations. (a) & (b): Torsion angles related by mirror planes (---) have opposite signs. (c) & (d): Torsion angles related by 2-fold rotational axes (—) have the same sign.
Figure 1.3: The most commonly observed conformations of six-membered rings and the corresponding mirror and twofold rotational symmetries are indicated on the right (Duax et al., 1976).
Figure 1.4: One octant of the sphere on which the conformations of six-membered rings can be mapped (for a constant Q). Special conformations are indicated: C = chair for $\theta = 0^\circ$; B = boat for $\theta = 90^\circ$, $\phi = 0^\circ$; TB = twist boat for $\theta = 90^\circ$, $\phi = 90^\circ$; HB = half-boat; HC = Half-chair.
Figure 1.5: The nomenclature defining the location of an asymmetry parameter includes the symmetry and the lower numbered atom or bond intersected by the symmetry element (Duax et al., 1976).
Figure 1.6: Geometry of hydrogen bonds.