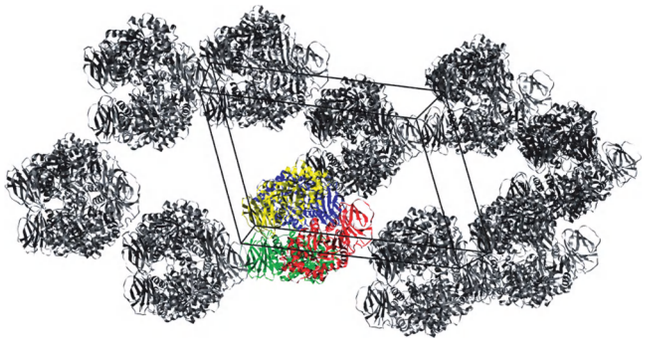


Protein Crystals Contain Solvent-Filled Channels

Figure 2.2. Because protein molecules are generally of irregular shapes, packing in the crystal leads to spaces and channels between molecules that are filled with solvent from the crystallization mother liquor. Consequently, small molecules, such as substrates and inhibitors, can diffuse into the crystal and reach the protein surface. Shown are several molecules of glucocerebrosidase, showing the arrangement of the packing, the unit cell, and the asymmetric unit. Data were taken from PDB code 1OSG.

crystal and the higher the resolution of the resulting electron density, thus, the higher the indices of the reflections that can be observed, the higher the resolution of the diffraction pattern, and the more precise the resulting electron density (Figure 2.3).

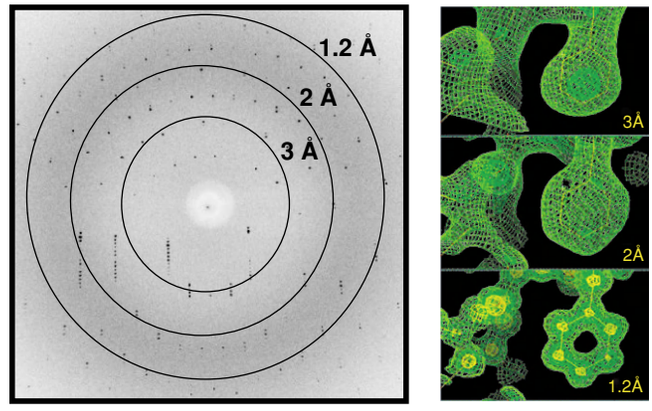
Ultimately, the quality of the diffraction pattern, in terms of the intensities of the reflections and the resolution of the data set, will determine the quality of the electron density map that can be obtained. A number of criteria are used to determine the quality of a data set: I(ntensity), R(adiation damage), O(verlap), Rm(erge), C(ompleteness) (Table 2.1) (a useful discussion of these parameters may be found in Wlodawer et al., 2008).[7]

(I) The intensities of the reflections will clearly influence the quality of the data. Intensity depends on a number of factors, mainly the size and quality of the crystal, the length of exposure to the x-ray beam, and the intensity of the x-ray beam. The quality of the intensity relative to background is given as the signal-to-noise ratio: $I/\sigma(I)$. Because proteins are subject to interactions with x-rays that lead to chemical changes, the latter two factors are counterproductive relative to the intensity of the reflection. Therefore, the $I/\sigma(I)$ criterion is sometimes used to define the resolution limits of the data, the diffraction limit defined where this value decreases to 2.0.

(R) Radiation damage is a significant factor for data quality and is usually dealt with by reducing the temperature of the crystal during data collection. The most common temperature is the "cryo" range, achieved by using liquid nitrogen to cool to approximately 100 K. Such a temperature requires special treatment of the crystal, because it contains water, which can freeze with disastrous results for the crystal. Flash freezing, with and without additives to prevent crystallization of water, is used.

(O) Beyond intensity of a reflection, the ability to distinguish individual reflections from each other is also essential. The separation of reflections in a diffraction pattern depends primarily on the size of the unit cell: the larger the unit cell axes, the closer the reflections are to each other, and overlap of reflections leads to inaccurate intensity determination. Disorder in the crystal, either of packing or from mechanical damage (e.g., through freezing), can lead to a broadening of the reflections, and excessive broadening will produce overlap.

(R_m) Because the diffraction pattern contains elements of symmetry, and because of the method used to measure reflections, most reflections are measured more than once. Consequently, the reproducibility of these measurements is a measure of the precision to which the reflections can be determined. Statistically, the more often a reflection is measured, and the closer those measurements are to each other, the better the data set. The redundancies for the data are given in terms of an average, overall redundancy. The reproducibility of the reflection measurements is given in terms of a residual factor, R_{merge} (R_m, sometimes called R_{sym}), the difference between a measured intensity and the average intensity, divided by the average intensity, for all related reflections. This factor will change with resolution, so it should be given for all data and for data in the highest resolution shell for



Resolution vs electron density

Figure 2.3. The resolution to which a structure can be determined depends on the reflections that can be measured. Shown at left is a diffraction pattern of a typical protein. The rings indicate levels of resolution. At right are electron densities for the same residue of a protein, calculated from the data within a resolution range. Thus, if only the data within the 3Å resolution circle are used, the electron density map lacks detail at the atomic level. On the other hand, if all of the data within the 1.2Å resolution circle are used, the electron density resolves the positions of individual atoms (shown are two levels of the electron density). Note that the reflections become weaker as the resolution gets higher (measured as $I/\sigma(I)$), and the total number of possible reflections measured becomes less (completeness).