5.2. Crystal-density measurements

BY E. M. WESTBROOK

5.2.1. Introduction

Crystal-density measurements have traditionally been a valuable and accurate (±4%) method for determining molecular weights of proteins (Crick, 1957; Coleman & Matthews, 1971; Matthews, 1985). But since exact chemical compositions of proteins being crystallized today are usually known from DNA sequences, crystal densities are rarely used for this purpose. Rather, crystal-density measurements may be necessary to define a crystal's molecular-packing arrangement, particularly when a crystal has an unusual packing density (very dense or very open); when there are a large number of subunits in the crystallographic asymmetric unit; when the structure consists of heterogeneous subunits, so the molecular symmetry or packing is uncertain; and for crystals of nucleic acids, nucleic acid/protein complexes and viruses.

5.2.2. Solvent in macromolecular crystals

Crystals of biological macromolecules differ from crystals of smaller molecules in that a significant fraction of their volume is occupied by solvent (Adair & Adair, 1936; Perutz, 1946; Crick, 1957). This solvent is not homogeneous: a part binds tightly to the macromolecule as a hydration shell, and the remainder remains free, indistinguishable from the solvent surrounding the crystal.

Hydration is essential for macromolecular stability: bound solvent is part of the complete macromolecule's structure (Tanford, 1961). Diffraction-based studies of macromolecular crystals verify the presence of well defined bound solvent. Typically, 8–10% of the atomic coordinates in each Protein Data Bank file are those of bound water molecules. The consensus observation of protein hydration (Adair & Adair, 1936; Perutz, 1946; Edsall, 1953; Coleman & Matthews, 1971; Kuntz & Kaufmann, 1974; Scanlon & Eisenberg, 1975) is that every gram of dry protein is hydrated by 0.2–0.3 g of water: this is consistent both with the presence of a shell of hydration, the thickness of which is about one water molecule (2.5–3 Å), and with the rule-of-thumb that approximately one water molecule is found for every amino-acid residue in the protein's crystal structure. Matthews (1974) suggests setting this hydration ratio, w, to 0.25 g water per gram of protein as a reasonable estimate for typical protein crystals.

Crystallographic structures also exhibit empty regions of 'free' solvent. Such voids are to be expected: closely packed spheres occlude just 74% of the space they occupy, so to the extent that proteins are spherical, tight packing in their crystals would leave 26% of the crystal volume for free solvent. Although the distinction between free and bound solvent is not sharp (solvent-binding-site occupancies vary, as do their refined B factors), it is a useful convention and is consistent with many observed physical properties of these crystals.

5.2.3. Matthews number

In an initial survey of 116 crystals of globular proteins (Matthews, 1968) and in a subsequent survey of 226 protein crystals (Matthews, 1977), Matthews observed that proteins typically occupy between 22 and 70% of their crystal volumes, with a mean value of 51%, although extreme cases exist, such as tropomyosin, whose crystals are 95% solvent (Phillips *et al.*, 1979). The volume fraction occupied by a macromolecule, φ_m, is reciprocally related to V_M, the Matthews number, according to

$$V_M = \overline{v}_m/N_o\varphi_m = V/nM, \qquad (5.2.3.1)$$

where $\overline{v}_m$ is the partial specific volume of the macromolecule (Tanford, 1961), N_o is Avogadro's number, V is the volume of the crystal's unit cell, n is the number of copies of the molecule within the unit cell and M is the molar weight of the macromolecule (grams per mole). V_M is the ratio between the unit-cell volume and the molecular weight of protein contained in that cell. The distribution of V_M (2.4 ± 0.5 Å^3 Da^{-1}) is asymmetric, being sharply bounded at 1.7 Å^3 Da^{-1}, a density limit consistent with spherical close packing. The upper limits to V_M are much less distinct, particularly for larger proteins. Matthews observed a slight tendency for V_M to increase as the molecular weight of proteins increases. V_M values below 1.9, or above 2.9, can occur but are relatively rare (beyond a 1σ cutoff).

The unit-cell volume, V, is determined from crystal diffraction. The partial specific volume of a macromolecule, $\overline{v}_m$, is the rate of change in the volume of a solution as the (unhydrated) macromolecule is added. It can be measured in several ways, including by ultracentrifugation (Edelstein & Schachman, 1973) and by measuring the vibrational frequency of a capillary containing a solution of the macromolecule (Kratky *et al.*, 1973). $\overline{v}_m$ typically has a value around 0.74 cm^3 g^{-1} for proteins and around 0.50 cm^3 g^{-1} for nucleic acids (Cantor & Schimmel, 1980). Values of $\overline{v}_m$ are tabulated for all amino acids and nucleotides, and $\overline{v}_m$ of a macromolecule can be estimated with reasonable accuracy as the mean value of its monomers. Commercial density-measuring instruments are available to determine $\overline{v}_m$ by the Kratky method. Because M is usually well known from sequence studies, n – the number of copies of the macromolecule in the unit cell – can be calculated thus:

$$n = V/V_M M = VN_o\varphi_m/\overline{v}_m M. \qquad (5.2.3.2)$$

For proteins, evaluating this expression with $V_M = 2.4$ usually provides an unambiguous integer value for n – which must be a multiple of the number of general positions in the crystal's space group! Setting n to its integer value then provides the actual value for V_M. If the calculated V_M value lies beyond the usual distribution limits, if n has an unexpected value or a large value, or if the crystal contains unusual components or several different kinds of molecular subunits, the crystal density may need to be measured accurately.

5.2.4. Algebraic concepts

Let V be the volume of one unit cell of the crystal. Let m_c be the total mass within one unit cell, and m_m, m_{bs} and m_{fs} be the masses, within one unit cell, of the macromolecule, bound solvent and free solvent, respectively. Let ρ_c, ρ_m, ρ_{bs} and ρ_{fs}, respectively, be the densities of a complete macromolecular crystal, its unsolvated macromolecule, its bound-solvent compartment and its free-solvent compartment. Let φ_m, φ_{bs} and φ_{fs}, respectively, be the fractions of the crystal volume occupied by the unsolvated macromolecule, the bound solvent and the free solvent. By conservation of mass,

$$m_c = m_m + m_{bs} + m_{fs}. \qquad (5.2.4.1)$$

The volume fractions must all add to unity:

$$\varphi_m + \varphi_{bs} + \varphi_{fs} = 1. \qquad (5.2.4.2)$$

The density of the crystal is the total mass divided by the unit-cell volume:

$$\rho_c = \frac{m_c}{V} = \frac{m_m}{V} + \frac{m_{bs}}{V} + \frac{m_{fs}}{V}. \qquad (5.2.4.3)$$

The mass in each solvent compartment is the product of its density and the volume it occupies:

references