International Tables for Crystallography (2006). Vol. F, Section 5.2.4, pp. 117–118.

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:

$$m_{bs} = \rho_{bs} V \varphi_{bs}, \qquad m_{fs} = \rho_{fs} V \varphi_{fs}. \tag{5.2.4.4}$$

The mass of the macromolecule in the cell can be defined either from its partial specific volume, $\overline{v}_m$, the unit-cell volume, V, and the molecules' volume fraction, φ_m, or from the molar weight, M, the number of molecular copies in the unit cell, n, and Avogadro's number, N_o:

$$m_m = V\varphi_m/\overline{v}_m = nM/N_o. \tag{5.2.4.5}$$

Now (5.2.4.3) may be rewritten as

$$\rho_c = \varphi_m/\overline{v}_m + \rho_{bs}\varphi_{bs} + \rho_{fs}\varphi_{fs}. \tag{5.2.4.6}$$

Define a mean solvent density, ρ_s:

$$\rho_s = \frac{\rho_{bs}\varphi_{bs} + \rho_{fs}\varphi_{fs}}{\varphi_{bs} + \varphi_{fs}}. \tag{5.2.4.7}$$

This allows (5.2.4.6) to be rewritten as

$$\rho_c = \varphi_m/\overline{v}_m + (1 - \varphi_m)\rho_s. \tag{5.2.4.8}$$

Upon rearrangement, this gives expressions for the volume fraction of a macromolecule and for the molecular-packing number:

$$\varphi_m = \frac{\rho_c - \rho_s}{(\overline{v}_m)^{-1} - \rho_s} = \frac{n\overline{v}_m M}{VN_o},$$
$$n = \frac{VN_o}{\overline{v}_m M}\frac{\rho_c - \rho_s}{(\overline{v}_m)^{-1} - \rho_s}. \tag{5.2.4.9}$$

In (5.2.4.9), all terms can be measured directly, except ρ_s. The treatment of ρ_s will be discussed in Section 5.2.7. (5.2.4.9) defines the total macromolecular mass in the unit cell, $m_m = nM/N_o$, from a measurement of the crystal density ρ_c. If M were known from the primary sequence of the molecule, this measurement determines the molecular-packing number, n, with considerable certainty. If the molar weight were not accurately known, it could be determined by measuring the crystal density.

5.2.5. Experimental estimation of hydration

During refinement of crystal structures, crystallographers must decide how many solvent molecules are actually bound to the macromolecule and for which refined coordinates are meaningful. The weight fraction of bound solvent to macromolecule in the crystal, w, is estimated for most protein crystals to be about 0.25 (Matthews, 1974, 1985). However, its true value can be derived experimentally in the following manner. Since all relevant studies identify the bound solvent as water, it is reasonable to set the density of bound solvent as $\rho_{bs} = 1.0$ g ml^{-1}. Therefore, w can be expressed algebraically as

$$w = \frac{m_{bs}}{m_m} = \frac{\varphi_{bs}\rho_{bs}\overline{v}_m}{\varphi_m} = \frac{\varphi_{bs}\overline{v}_m}{\varphi_m}. \tag{5.2.5.1}$$

For crystals in which the rules-of-thumb $w = 0.25$ and $\overline{v}_m = 0.74$ cm^3 g^{-1} are valid, (5.2.5.1) implies that bound solvent occupies about one-third of the volume occupied by protein.

The crystal density, ρ_c, changes linearly with the density of free solvent surrounding the crystal. Let ρ_o be defined as the density the crystal would have if all its solvent were pure water ($\rho_s = 1.0$ g ml^{-1}):

$$\rho_o = 1 + \varphi_m(1/\overline{v}_m - 1). \tag{5.2.5.2}$$

A plot of crystal density against density of the supernatant (free solvent) solution should be a straight line with an intercept (at $\rho_{fs} = 1.0$ g ml^{-1}) of ρ_o and a slope of φ_{fs}. Therefore, by making a few crystal-density measurements, each with the crystal first equilibrated in solutions of varying densities, experimental values for ρ_o and φ_{fs} can be derived. If the partial specific volume is known for this molecule, φ_m, φ_{bs} and w can be derived from the expressions above. This approach was used by Coleman & Matthews (1971) and Matthews (1974) to measure molecular weights of six crystalline proteins, assuming $w = 0.25$, but their measurements could alternatively have assigned more accurate values to w, had the molecular weights been previously known. Scanlon & Eisenberg (1975) measured w for four protein crystals by this method (values betwen 0.13 and 0.27 were observed) and also confirmed that bound solvent exhibited a density of 1.0 g ml^{-1}.

5.2.6. Methods for measuring crystal density

Density measurements of macromolecular crystals are complicated by their delicate constitution. These crystals tolerate neither dehydration nor thermal or physical shock or stress. Furthermore, since macromolecular crystals contain free solvent, their densities will change as the density of the solvent in which they are suspended is changed. They cannot be picked up with tweezers, nor rinsed with arbitrary solvents, nor placed out to dry on the table.

The other experimental problem with these crystals is that they are very small. Typically, their linear dimensions are 0.1–0.2 mm, volumes are 1–10 nl and weights are 1–10 µg. Molecular structures can now be determined from even smaller crystals (linear dimensions as small as 20 µm) using synchrotron radiation, so density-measurement methods compatible with very small crystals are required. With such small samples, it is far easier and more accurate to measure *densities* than to measure directly *volumes* and *weights*.

The physical properties of macromolecular crystals constrain the methods by which their densities can be measured accurately. In all circumstances, great care must be taken to avoid artifacts such as air bubbles or particulate matter which often adhere to these crystals. All measurements should be made at one tightly controlled temperature, since thermal expansion can change densities and thermal convection can corrupt density gradients. Because crystals contain solvent, it is bad to dry them, since this process usually disrupts them, changing all parameters in unpredictable ways. Yet many density-measurement methods require that all external solvent first be removed from the crystals, since the measured densities will be some average of crystal and any remaining solvent. This can be an almost insurmountable problem for crystals containing cavities and voids. Unfortunately, many crystallization mother liquors are viscous and difficult to remove, for example if they contain polyethylene glycol (PEG).

Richards & Lindley (1999) list six methods for measuring crystal densities: pycnometry, the method of Archimedes, volumenometry, the immersion microbalance, flotation and the gradient tube. The first three methods require direct weighing of the crystal and are therefore of limited value for crystals as small as those used in macromolecular diffraction, although these methods are used in various applications, such as mineralogy and the sugar industry. The latter three methods measure densities and density differences, and can therefore be used in macromolecular crystallography. A new method specifically for protein crystals has recently been described (Kiefersauer *et al.*, 1996), involving direct tomographic measurement of crystal volumes coupled with quantitative amino-acid analysis. Because the gradient-tube method remains the method of choice for most crystal-density measurements, it will be discussed last and most thoroughly here.

5.2.6.1. *Pycnometry*

Pycnometry measures the density of a liquid by weighing a calibrated volumetric flask before and after it is filled with the

references