Kumar, A., & Pederson, T. (1975) *J. Mol. Biol. 96*, 353–365.
Lebleu, B., Marbaise, G., Huez, G., Temmerman, J., Burny, A., & Chautrenne, H. (1971) *Eur. J. Biochem. 19*, 264–269.
Liautard, J. P., Setyono, B., Spindler, E., & Kohler, K. (1976) *Biochim. Biophys. Acta 425*, 373–383.
Lindberg, V., & Sundquist, B. (1974) *J. Mol. Biol. 86*, 451–468.
Morel, C., Gander, E. S., Herzberg, M., Dubochet, J., & Scherrer, K. (1973) *Eur. J. Biochem. 36*, 455–464.
Olsnes, S. (1970) *Eur. J. Biochem. 15*, 464–471.
Perry, R. P., & Kelley, D. E. (1968) *J. Mol. Biol. 35*, 37–59.
Quinlan, T. J., Billings, P. B., & Martin, T. E. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 2632–2636.
Roberts, B. E., & Paterson, B. M. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 2330–2334.
Schwartz, H., & Darnell, J. E. (1976) *J. Mol. Biol. 104*, 833–851.
Spirin, A. S. (1969) *Eur. J. Biochem. 10*, 20–35.
Spohr, G., Granboulan, N., Morel, C., & Scherrer, K. (1970) *Eur. J. Biochem. 17*, 296–318.
Spohr, G., Kayibanda, B., & Scherrer, K. (1972) *Eur. J. Biochem. 31*, 194–208.
Strohman, R. C., Moss, P. M., Micou-Eastwood, J., Spector, D., Przybyla, A., & Paterson, B. (1977) *Cell 10*, 265–273.
Van der Marel, P., Tasseron-deJong, J. G., & Bosch, L. (1975) *FEBS Lett. 51*, 330–334.
Van Venrooij, W. V., van Eekelen, C. A. G., Jansen, R. T. P., & Princeu, J. M. G. (1977) *Nature (London) 270*, 189–191.
Williamson, R. (1973) *FEBS Lett. 37*, 1–6.

Measurement of Macromolecular Equilibrium Binding Constants by a Sucrose Gradient Band Sedimentation Method. Application to Protein–Nucleic Acid Interactions†

David E. Draper‡ and Peter H. von Hippel*

ABSTRACT: A method is described for determining equilibrium binding constants using sucrose gradient band sedimentation data. The method is particularly well suited to measuring the binding of small ligands or proteins to much more rapidly sedimenting macromolecules or macromolecular complexes (e.g., large DNA molecules or ribosomes). Only very small quantities of radioactively labeled binding proteins are needed for each determination, and association constants (K) as large as 10^{11} M^{-1} can be measured. Over a wide range of sedimentation conditions the only information necessary for calculation of the association constant is the initial concentration of potential binding sites for protein, the distance of sedimentation of the fast-sedimenting peak, and the fraction of total binding protein migrating with the peak. Results are presented demonstrating the applicability of this method to three different systems: (1) the interaction of bovine pancreatic ribonuclease with double-stranded DNA, (2) the interaction of *Escherichia coli* ribosomal protein S1 with single-stranded DNA, and (3) the interaction of ribosomal protein S1 with 30S ribosomal subunits. In each case the binding constant obtained reproduces (within a factor of two) that measured on the same system by a completely independent method. Common errors or misinterpretations in the application of this technique are described, as well as the extension of this approach to proteins with multiple binding sites, to cooperatively binding proteins, and to the determination of binding site stoichiometry.

The determination of binding parameters for interacting macromolecules (for example, in protein–nucleic acid complex formation) is often of central importance to the quantitative analysis of a biological control system. However this can be an experimentally difficult problem and one for which straightforward procedures of general utility and interpretability are still largely not available.

Standard equilibrium dialysis techniques generally cannot be used for such systems because both macromolecular components may be too large to dialyze; in addition binding constants are often so large ($K > 10^7$ M^{-1}) that concentrations required for the reasonably accurate determination of optical changes upon complex formation (at least 10^{-7} M components) often cannot be employed. Futhermore, for many systems only small quantities of material are available.

For protein–nucleic acid interactions, nitrocellulose filter binding assays have been successfully used (Riggs et al., 1970; Yarus & Berg, 1970; Hinkle & Chamberlin, 1972). This assay depends on the preferential retention of protein and protein–nucleic acid complexes by the filter; artifacts can arise if the filter is rinsed too thoroughly after trapping the complex or if the complex has a short half-life relative to the time of filtering (Hinkle & Chamberlin, 1970). In addition a number of factors affect the retention of complexes by filters, including temperature, ionic strength, and presence of divalent metal ions; this often complicates a survey of binding conditions by this technique. Furthermore, protein–protein interactions cannot, of course, be measured by the filter assay technique.

A boundary sedimentation method worked out by Jensen & von Hippel (1977; see also Revzin & von Hippel, 1977) can be used to study a variety of interactions. However it generally cannot be applied to systems with binding constants greater

† From the Institute of Molecular Biology and the Department of Chemistry, University of Oregon, Eugene, Oregon 97403. *Received August 15, 1978; revised manuscript received December 11, 1978.* This work was supported in part by U.S. Public Health Service Research Grant GM-15792 as well as by a Predoctoral Traineeship (to D.E.D.) from U.S. Public Health Service Training Grant GM-00444.

‡ Present address: Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80302.

0006-2960/79/0418-0753$01.00/0

than $\sim 10^7$ M^{-1} and requires fairly large amounts of material. Therefore we have developed, and describe here, a sucrose gradient band sedimentation technique for measuring association constants (K) with values up to at least 10^{11} M^{-1}. The method is rapid, easy to apply to most systems, and requires little material.

The principle of the technique is very straightforward. If a small ligand (e.g., a binding protein) and an excess of a faster sedimenting component (e.g., a nucleic acid) to which the protein can bind are centrifuged as a band through a sucrose gradient, the amount of protein cosedimenting with the nucleic acid after some distance will be a function of the binding affinity and the concentration of nucleic acid. At an appropriate initial concentration of nucleic acid, some portion of the protein will be free at the start of the run and therefore will remain at or near the top of the gradient. As the band containing the complex moves down the tube, equilibrium will be continuously reestablished and more unbound protein will be left behind, leaving a smear of protein through the gradient.

To describe this process quantitatively, we first use a very simple model that ignores spreading (due to diffusion, heterogeneity, or convection) of the sedimenting peak. Application of the equation obtained to several protein–nucleic acid interaction systems (for which independent measurements of K exist) shows that accurate values of K are obtained. A more detailed treatment, which takes band spreading into account, demonstrates that we are justified in ignoring this complication over a wide range of sedimentation conditions. The various limitations and technical problems that may arise are discussed, and finally we show how this simple treatment can be extended to situations involving multiple binding sites, cooperativity, and large initial ratios of binding protein to nucleic acid binding sites.

Materials and Methods

All buffers were made up in doubly distilled water. Sucrose was ribonuclease-free grade (Sigma). Tritium-labeled bovine pancreatic ribonuclease was obtained from Worthington. Tritium-labeled *Escherichia coli* ribosomal protein S1 was prepared as described by Draper et al. (1977), and 30S *Escherichia coli* ribosomal subunits devoid of S1 protein were prepared by a modification of the method of Tal et al. (1972). Phage λ DNA was purified by standard phenol extraction of λ bacteriophage.

All sucrose gradients were 5–20% in sucrose concentration (w/v) in the indicated buffer, with a total volume of 5 mL. The macromolecule-containing band (usually 0.2 mL) was layered on top of the gradient with a micropipet; sedimentation was carried out in a Beckman (L2-65B) preparative ultracentrifuge using an SW 50.1 rotor. The ribonuclease–DNA experiments were run at 35 000 rpm for 100 min at 18 °C, the S1–DNA experiments at 35 000 rpm for 240 min at 5 °C, and the S1–30S ribosomal subunit experiments at 40 000 rpm for 90 min at 18 °C.

Theory, Results, and Discussion

Idealized Treatment. We first consider band sedimentation, through a sucrose gradient, of a rapidly sedimenting nucleic acid with a more slowly sedimenting protein; a very small initial ratio of protein relative to nucleic acid binding sites is assumed. The volume of the protein–nucleic acid sample layered on top of the gradient is taken as the volume of one "fraction", and the sample is numbered as the 0th fraction (see diagram in Figure 1a). Initially we neglect diffusion and other sources of band spreading and consider the sedimentation of this one-fraction-wide band down the gradient. Before sedimentation begins, the concentration of *bound* protein in the 0th fraction is (by the law of mass action)

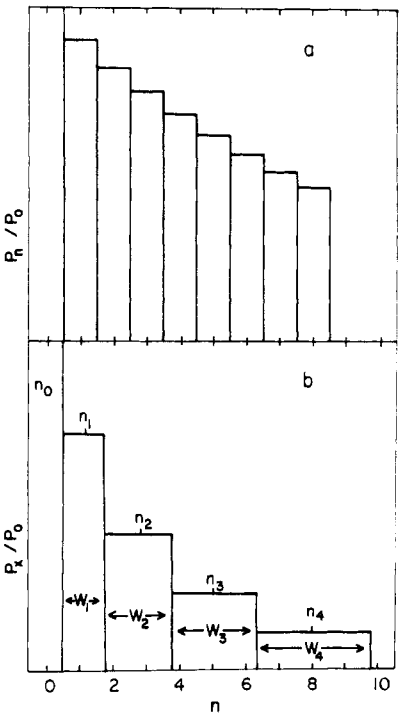

FIGURE 1: Illustration of band sedimentation experiments. (a) Sedimentation without band spreading. The bars represent the fraction of binding protein remaining with the fast-sedimenting peak (P_n/P_0) at each point in the gradient, calculated according to eq 2, with KL = 10. (b) Sedimentation with band spreading. The boxes represent the broadening of each successive peak transposition predicted by eq 4–9; the area of each box is the fraction of binding protein remaining with the fast-sedimenting peak (P_x/P_0). Calculated for KL = 10, W_f = 3.5, n_f = 8. Note that the area of the peak centered at n = 8 (n_4) is nearly the same as the area of the peak at n = 8 in Figure 1a (0.40 and 0.47, respectively).

$$P_1 = \left(\frac{KL}{1 + KL}\right)P_0 \qquad (1)$$

where P_0 is the total initial concentration of protein, K is the association constant, and L (for ligand) is the concentration of *free* nucleic acid binding sites. Since the concentration of protein is small relative to the number of nucleic acid binding sites, we make the approximation that $L_{total} \simeq L_{free}$ and in the following calculations take L to be the *total* concentration of potential polynucleotide binding sites. As the nucleic acid sediments into the volume labeled fraction one, the amount of protein carried with it will be only the protein bound in the 0th fraction, i.e., P_1.

The sedimentation on sucrose gradients of large nucleic acids or nucleic acid–protein complexes (e.g., phage DNAs, ribosomal RNAs, ribosomal subunits, etc.) requires $\sim$10 min for the band to move through one fraction (usually 0.1–0.2 mL on a 5.0-mL gradient). Using the Debye–Smoluchowski equation we may calculate that the forward rate constant (k_f) for the diffusion-controlled *nonspecific* binding of a protein of $\sim$50 000 daltons to a large nucleic acid will be $\sim 10^{10}$ M^{-1} s^{-1} (for example, see von Hippel, 1978). The half-time for the dissociation of such a complex will be $\sim$15 s for $K = 10^{12}$ M^{-1}. Since this time is considerably smaller than the time required to move the band by one fraction, effective equilibrium between free and bound protein will be maintained throughout the run for binding constants (K) smaller than $\sim 10^{12}$ M^{-1}. Of course, considerably smaller values of K will represent the upper limit for the method if specific site binding of protein to nucleic acid

Table I

interaction	K (M^{-1}) sucrose gradient method	K (M^{-1}) independent method	reference
RNase–native DNA[a]	$(0.9 \pm 0.2) \times 10^5$ [d]	1.7×10^5 (velocity sedimentation)	Jensen & von Hippel (1976)
S1–denatured DNA[b]	$(3.0 \pm 0.6) \times 10^6$ [d]	$(3.8 \pm 1.1) \times 10^6$ (fluorescence titration)	Draper et al. (1977)
S1–30S subunits[c]	$(1.7 \pm 0.8) \times 10^8$ [e]	$(4.8 \pm 1.0) \times 10^8$ [f] (competition)	Draper & von Hippel (1979)

[a] Determined in 1.0 mM NaCl, 0.1 mM Na_2EDTA, 1 mM Na_2HPO_4, pH 7.7. Phage λ DNA was used. [b] Determined in 0.1 M NaCl, 0.1 mM Na_2EDTA, 10 mM Na_2HPO_4, pH 7.7, 1.0 mM β-mercaptoethanol. Phage λ DNA was used. [c] Determined in 0.3 mM $Mg(OAc)_2$, 0.1 M NH_4Cl, 20 mM Tris, pH 7.7, 1.0 mM β-mercaptoethanol. 30S subunits are from *E. coli* MRE 600 ribosomes and contain less than 0.01 copies of S1 per subunit. [d] Estimation of error is based on 5% error in determination of P_n/P_0, L, and n_f. [e] Average of five determinations. [f] Average of determinations with poly(dA) and poly(dC).

(or protein to protein) is involved or if the binding interaction is slower than diffusion controlled (see following sections).

The concentration of *bound* protein (after equilibrium is established) in fraction one will therefore be

$$P_2 = \left(\frac{KL}{1 + KL}\right)P_1 = \left(\frac{KL}{1 + KL}\right)^2 P_0$$

The general equation for the fraction of protein carried into fraction n [i.e., the amount of protein bound in fraction $(n\text{-}1)$] is thus simply

$$\frac{P_n}{P_0} = \left(\frac{KL}{1 + KL}\right)^n \qquad (2)$$

This can be rearranged to give KL as a function of the distance of sedimentation and the fraction of protein remaining with the nucleic acid peak

$$KL = \left[\left(\frac{P_n}{P_0}\right)^{-1/n} - 1\right]^{-1} \qquad (3)$$

Equation 2 is similar to the equation derived by Silhavy et al. (1975) to describe the retention of small ligands by binding proteins inside a dialysis bag (the time of dialysis being analogous to the distance of sedimentation) and identical with the equation developed by deHaseth et al. (1977) to account for the retention of binding protein by a DNA-cellulose affinity column. The principle behind these procedures and that described here is the same; i.e., as a protein–ligand complex is "washed" with a volume of buffer (e.g., by dialysis of the complex against a large volume of buffer or by sedimenting the complex through buffer), the amount of buffer required to remove a certain fraction of the ligand is a function of the binding affinity. We emphasize that, though each method gives an apparent "rate" of ligand release from the complex, this rate depends solely on *equilibrium* considerations. Only in the case of exceptionally tight binding does the actual first-order dissociation rate constant of the complex begin to affect the apparent rate of ligand release. Misunderstanding of this point has led to estimates (based on band sedimentation experiments) of protein–nucleic acid dissociation rate constants which are too small by many orders of magnitude (e.g., see Reuben & Gefter, 1974; Molineaux et al., 1975).

The analysis presented above is, of course, very simplistic; in particular it ignores band spreading by assuming that the peak remains one-fraction wide throughout the sedimentation run. Therefore we were suprised to find that association constants calculated using eq 3 come very close to values determined by other methods. The reason for this, as will be shown below, is that when band spreading is taken into account two modifications to eq 2 must be applied, and these modifications very nearly cancel one another.

To illustrate the application of eq 3, the sedimentation of [^{3}H]ribonuclease ([^{3}H]RNase) with native λ DNA is shown

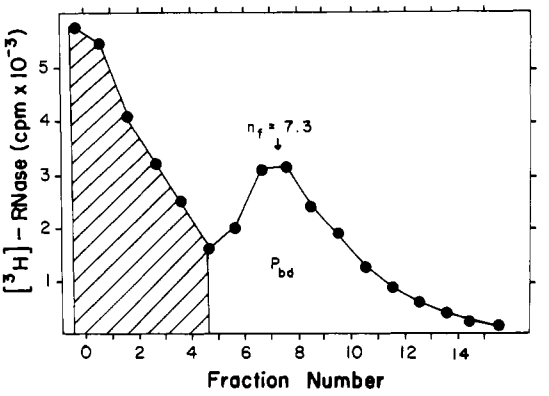

FIGURE 2: Sedimentation of [^{3}H]RNase with native λ DNA. Conditions of sedimentation are described under Materials and Methods and in Table I. The top of the gradient is at the left. The unhatched area is the fraction of total protein taken as still associated with the DNA peak and is 47% of the total protein.

in Figure 2. The concentration of DNA (expressed in units of base pairs) was 98 μM in the initial sample; the amount of [^{3}H]RNase used was enough to saturate less than 1% of the potential DNA binding sites.[1] The sedimentation profile shows that a significant portion of the [^{3}H]RNase has been left behind near the top of the gradient, while some [^{3}H]RNase is clearly associated with the faster migrating DNA (under these sedimentation conditions unbound [^{3}H]RNase does not move a significant distance). By estimation that 47% of the total recovered [^{3}H]RNase is associated with the DNA peak centered at n = 7.3 fractions,[2] and by use of the initial concentration of DNA (98 μM) as the concentration of DNA binding sites, eq 3 yields $K = 0.9 \times 10^5$ M^{-1}. This value is within a factor of two of the value of 1.7×10^5 M^{-1} measured by Jensen & von Hippel (1976) with native calf thymus DNA under identical conditions by a very different method (Jensen & von Hippel, 1977).

Table I summarizes binding constants determined by this sedimentation method for several interactions of known affinity. In all cases agreement with values determined by other methods is within approximately a factor of two.

A common (and generally incorrect) procedure in analyzing sedimentation patterns of the kind shown in Figure 2 has often

[1] In considering protein–nucleic acid interactions one must generally take into account the site size of the binding protein. This follows because the concentration of free nucleic acid binding sites will not be a linear function of bound protein, since one protein can cover more than one potential binding site (McGhee & von Hippel, 1974). However, as the level of saturation of the nucleic acid with protein becomes small, the effect of site size becomes insignificant. At the protein to nucleic acid ratios of these experiments, the concentration of total and of free binding sites is simply taken to be equal to the total concentration of nucleic acid bases or base pairs.

[2] In certain cases (see below) the sedimenting band may be appreciably skewed. In such situations the center-of-mass of the peak, rather than the maximum ordinate, should be used to define the center of the band.

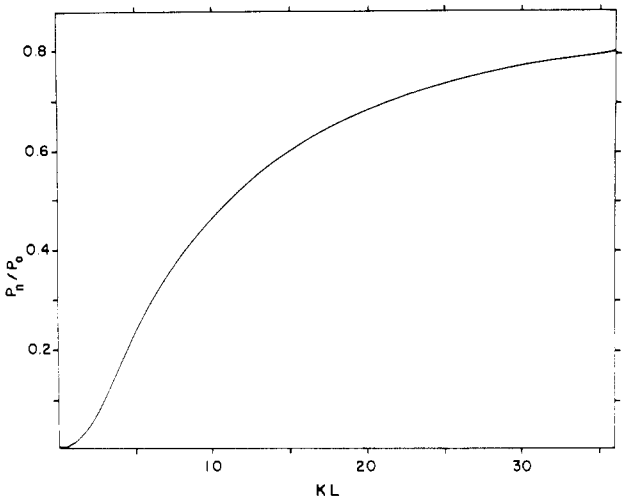

FIGURE 3: P_n/P_0 as a function of KL, calculated by use of eq 3 with $n = 8$.

been to define the protein at the top of the gradient and the protein moving with the faster sedimenting peak as "free" and "bound" protein, respectively, and then to apply the mass action law to this situation directly (e.g., see Laughrea & Moore, 1978). (In essence this represents the total nonequilibrium, or "frozen" original equilibrium, case.) This treatment of the data of Figure 2 yields $K = 0.9 \times 10^4$ M^{-1}, approximately 10-fold lower than the correct value.

Figure 3, in which we plot P_n/P_0 as a function of KL (using a typical value of $n = 8$), illustrates this point further. The curve is not a normal binding isotherm; the exponent (n) in eq 2 makes the curve sigmoid, even though no cooperativity is present. For a normal binding isotherm the product KL is 1.0 when the fraction of bound ligand is 0.5 (eq 1); here $KL \simeq 10$ when $P_n/P_0 = 0.5$ (eq 2 with $n = 8$). Thus experiments in the literature in which binding constants have been estimated from band sedimentation experiments (under binding equilibrium conditions) by application of the mass action law directly to the final distribution yield association constants which are seriously underestimated, though corrections are easily possible if the experimental details of the measurement are available (e.g., see Draper & von Hippel, 1979).

Treatment of Band Spreading. To justify the application of eq 3 to sedimentation runs, we show that an approximate consideration of band spreading will change the calculated value of KL by less than 30% under most sedimentation conditions. Sedimentation of a real band soon results in the generation of a peak which is generally roughly Gaussian in shape (though see footnote 2) and broadens with time. An exact analytical solution of this problem is impossible (though numerical solutions have been carried out; see Cann, 1970), so, for present purposes, we approximate the sedimenting peak as a rectangle with width equal to the width at half-height of the actual Gaussian peak.

Equation 2 must be modified in two ways to take into account the broadening of the sedimentation peak. First, n transpositions of the peak are no longer required to reach fraction number n; rather, a smaller number of band transpositions, x, are needed (see Figure 1b). In addition, the concentration of nucleic acid binding sites (L) must decrease as the peak broadens; this will be taken into account by calculating a dilution factor, D.

To calculate x, we assume that the width of the (rectangular) peak varies linearly with its position in the gradient, from one fraction wide at the 0th fraction to a final width, W_f, at $n = n_f$.[3] Thus

$$W_n = (W_f - 1)\left(\frac{n}{n_f}\right) + 1 \tag{4}$$

where W_n is the width of the peak (in units of initial fractions; see Figure 1b) when the peak reaches fraction n. The fraction number of the xth peak transposition, n_x, is related to that of the previous peak transposition (n_{x-1}) by

$$n_x = n_{x-1} + \frac{W_{x-1}}{2} + \frac{W_x}{2} \tag{5}$$

Substituting eq 4 into eq 5 and solving for n_x gives

$$n_x = \frac{(n_{x-1})(2n_f + W_f - 1) + 2n_f}{2n_f - W_f + 1} \tag{6}$$

This equation is easily solved repetitively, starting with $n_0 = 0$, to give the fraction numbers of the successive peak transpositions and thus x. Assuming that the peak is diluted by a constant factor between each transposed zone, we define a dilution factor, D, as the ratio of successive peak widths

$$D = \frac{W_{x-1}}{W_x} \tag{7}$$

Equation 2 can now be rewritten to take into account the broadening and dilution of the peak during sedimentation

$$\frac{P_x}{P_0} = \left(\frac{KL}{1 + KL}\right)\left(\frac{KLD}{1 + KLD}\right)\cdots\left(\frac{KLD^{x-1}}{1 + KLD^{x-1}}\right)$$

or

$$\frac{P_x}{P_0} = \prod_{n=0}^{n=x-1}\left(\frac{KLD^n}{1 + KLD^n}\right) \tag{8}$$

This equation predicts the fraction of initial protein still migrating with the nucleic acid after x peak transpositions. In the case of $W_f = 1$ (i.e., no broadening of the sedimenting peak), eq 8 reduces to eq 2.

To demonstrate the effect of band spreading on the calculated value of KL, we consider a hypothetical band sedimenting to fraction $n = 8$ with $P_x/P_0 = 0.5$ and assume a series of values for the final peak width from $W_f = 1$ to $W_f = 11$ fractions. For each value of W_f, x and D are calculated by use of eq 6 and 7 and then KL is determined from a plot of P_x/P_0 vs. KL, calculated by use of eq 8. The outcome, for KL as a function of W_f, is presented in Figure 4 (upper curve). It can be seen that KL calculated ignoring band spreading (the $W_f = 1$ case) is not more than 30% larger (upper dashed line) than the value calculated taking peak spreading into account over the range $1 < W_f < 5$. The error involved in using the "non-band-spreading" approximation varies slightly with the distance of sedimentation, becoming somewhat larger with increasing n.

If peak broadening is entirely due to diffusion and not to heterogeneity of the sedimenting molecules, then peak width will increase with the square root of time (or peak position) and not linearly with time as assumed above. The case can be treated the same way, by use of eq 9 in place of eq 4, to

$$W_n = (W_f - 1)\left(\frac{n}{n_f}\right)^{1/2} + 1 \tag{9}$$

determine the variation of peak width with distance of sedimentation. An equation similar to eq 6 is obtained (which

[3] This assumes that the gradient is such as to make sedimentation approximately linear in time and that most of the band spreading is due to sedimentational heterogeneity, rather than to simple diffusion. If other situations are known to prevail, eq 4 and 5 must be modified (see below; eq 8 and 9).

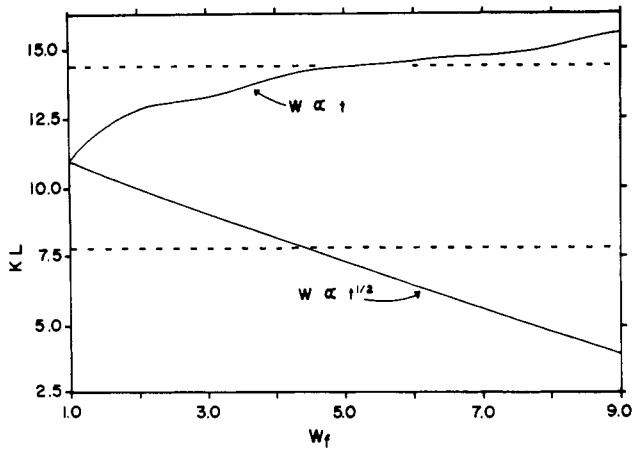

FIGURE 4: KL as a function of W_f, calculated by use of eqs 6–9 with $n_f = 8$ and $P_x/P_0 = 0.5$. We assume either that peak broadening is entirely due to heterogeneity (W proportional to t; upper curve) or to diffusion (W proportional to $t^{1/2}$; lower curve). The upper and lower dashed lines represent values of KL which are, respectively, 30% greater or 30% smaller than the value of KL for $W_f = 1$.

is, however, a quadratic in n_x) and is used to calculate successive peak positions and x. The dilution factor is now not a constant but must be calculated separately for each successive peak transposition.

If we now consider the same hypothetical band sedimenting to $n = 8$ with $P_x/P_0 = 0.5$ and plot KL as a function of W_f, the lower line of Figure 4 is obtained. The error in KL again becomes steadily larger with increasing width of the sedimenting peak but here in the opposite direction from that observed for the peak broadening linearly with time. Thus here the value calculated by use of the non-band-spreading approximation is too *large*, rather than too small. The error is still less than 30% in the range $1 < W_f < 5$. Since in many cases both heterogeneity and diffusion contribute to peak broadening, the actual value of KL will lie somewhere between the two extreme cases shown, and the error made in KL as a result of using eq 3 will be acceptable over a larger range of W_f. Therefore we conclude that eq 3, which neglects the effects of band spreading, is suitable for analyzing sucrose gradients under most sedimentation conditions.

As indicated above, the reason that band spreading has such a small effect on the value of KL calculated with eq 3 is that the two modifications introduced into the equation have opposite effects. Peak broadening, which reduces the exponent n to a smaller value, tends to make KL larger. However the concomitant reduction in the concentration of binding sites (L), due to dilution, decreases KL. The two effects nearly cancel one another over a wide range of values of W_f.

Limitations of the Method. For accurate results, concentrations and sedimentation conditions should fall within certain limits which we summarize here.

(a) The upper limit of association constants which can be measured by this technique is determined by the half-life of the interaction complex; when this becomes too large, equilibrium is no longer maintained during the run. As noted above, this limit is reached at $K \simeq 10^{12}$ M^{-1} for proteins of $\sim$50 000 daltons binding nonspecifically to a large nucleic acid. It will be lower for very large proteins, for proteins which have an association rate much slower than diffusion controlled, and for protein–nucleic acid (or protein–ribosome) interactions involving direct (i.e., no nonspecific binding intermediates) specific site binding.

A useful test for the validity of the equilibrium assumption in any particular case is to take advantage of the "tracer" nature of the method at very low ratios of (e.g.) protein to nucleic acid binding sites. If equilibrium indeed prevails, the protein wil spend such a small fraction of its time associated with any particular nucleic acid molecule that the sedimentation rate of the latter should be totally unaffected by the binding process. Thus a demonstration that the same sedimentation rate is observed whether the progress of the band is monitored by the sedimentation of (e.g.) the radioactive protein or by the sedimentation of the nucleic acid or ribosomal component itself often comprises an excellent way of assuring that equilibrium prevails (see also Draper & von Hippel, 1979).

Whether or not equilibrium is maintained during a run can also be checked by determining K using different concentrations of binding sites or different distances of sedimentation (i.e., different values of n_f). The fraction of protein remaining bound to nucleic acid should vary as predicted by eq 3 and not by the mass action law.

The lower limit of measurable association constants is determined only by the lowest experimentally resolvable concentration of protein cosedimenting with the nucleic acid component. For reasonable accuracy at least 1/20th of the total protein should still be associated with the nucleic acid at the end of the run. The association constant to which this limit applies can be decreased somewhat further by increasing the nucleic acid concentration or by stopping the run earlier (i.e., when the peak is at a smaller value of n).

(b) The concentration of binding protein must be much less than the concentration of binding sites for the protein to satisfy the condition $L_{free} \simeq L_{total}$. Initial saturation levels of less than 10% are sufficient for this. Cases in which the saturation level is much higher can be treated, though a simple equation for KL is not obtained (see discussion of stoichiometry determinations, below).

(c) The optimum range for P_n/P_0 is roughly 0.2–0.6. Outside of this range the error in the determination of P_n/P_0 rises, since the uncertainty in deciding which fractions to include with "shed" protein and which with "bound" protein increases substantially. As Figure 3 shows, small errors of P_n/P_0 ratios outside the 0.2–0.6 range may also be amplified in calculating KL.

(d) The breadth of the peak at half-height should not exceed the distance of sedimentation, unless eq 8 is used to take into account the effects of band spreading. For very broad peaks the possibility that some of the initially shed protein will have "caught up" with the trailing part of the peak (due to either diffusion or sedimentation of this protein from its "shed" position) should also be taken into account.

Possible Sources of Error. The experimenter should be aware of several problems which may give rise to erroneous determinations of K and controls which may be done to avoid them.

(a) Many binding proteins have a significant affinity for nitrocellulose centrifuge tubes, which can present a severe problem at low protein concentrations. Calculation of the total recovery of protein after a run is important to avoid this artifact. Several approaches can be used to alleviate the problem: (1) polyallomer tubes, which generally have a lower affinity for protein than nitrocellulose, can be used; (2) high salt concentrations, which generally decrease the binding of proteins to the tubes and other irrelevant sites, can be employed; and (3) bovine serum albumin or other proteins which do not bind to the nucleic acid but which do coat competing binding sites on the centrifuge tubes can be added to the system.

(b) Since the walls of sucrose gradient centrifuge tubes are parallel rather than sector shaped, sedimenting material will tend to accumulate at the walls of the tubes. This will lead to dilution of protein in the middle of the tube and concentration at the perimeter. A high concentration of protein at the walls may encourage binding of protein to the tube walls and, in severe cases, also density instability and convection. To avoid this kind of artifact, we generally limit the distance of sedimentation to about one-third of the 5-mL gradient (about 1.5 cm).

(c) For any binding reaction involving a substantial change in the partial specific volume of the reactants on complex formation, the hydrostatic pressure developed in the gradient during sedimentation can shift the equilibrium constants by an enormous amount (Harrington & Kegles, 1973). This possibility can be checked by simply determining K at different rotor speeds or by pouring a truncated gradient, layering on the sample band, and layering additional buffer on top. The increased pressure on the sample during the run should show whether pressure has any effect on K. For the several protein–nucleic acid interactions in which pressure dependence of binding has been tested [e.g., the RNase–DNA interaction (Jensen & von Hippel, 1977), the *lac* repressor–DNA interaction (Revzin & von Hippel, 1977), and the S1 protein–ribosome interaction (Draper & von Hippel, 1979)], no rotor speed effects have been detected.

(d) An important control, especially at high final binding densities ($P_n/P_0 > 0.6$–0.8), is to run a sample under very tight and stoichiometric binding conditions (for example, at low salt concentrations for many protein–nucleic acid interactions) to ascertain that all the protein is indeed active in binding. A small contaminant of nonbinding (i.e., defective) protein can result in serious underestimations of large binding constants if not detected and taken into account.

(e) At high concentrations the sedimentation coefficient of DNA becomes concentration dependent, and an asymmetric peak is often obtained (e.g., see Burgi & Hershey, 1963). These effects become noticeable at DNA concentrations greater than 1 μg/mL and severe at concentrations exceeding $\sim$12 μg/mL. Extreme asymmetry of the sedimenting peak may make it necessary to modify the above procedure; specifically, under these conditions the center-of-mass of the sedimenting peak should be used, instead of the maximum ordinate, to define n. In practice, for DNA binding proteins characterized by K values greater than 10^5 M^{-1}, this matter presents no problem. The concentrations of DNA which must be used to measure weaker binding constants are sufficiently large so that the band shape can be determined optically.

Proteins with Multiple Binding Sites. Many binding proteins have two or more binding sites. Ribosomal protein S1, for instance, has two (nonidentical) sites for polyribonucleotides (Draper et al., 1977). For a protein with two independent binding sites, the fraction of protein not bound via *either* site (i.e., the fraction which is left behind during sedimentation) is the product of the probabilities of each site being free

$$\frac{P_f}{P_0} = \frac{1}{(K_1L + 1)(K_2L + 1)}$$

where K_1 and K_2 are the binding affinities of the two different sites, P_f is the concentration of protein not bound in either site, and P_0 is the total protein concentration. The fraction of bound protein is thus

$$\frac{P_{bd}}{P_0} = 1 - \frac{1}{[K_1K_2L^2 + (K_1 + K_2)L + 1]}$$

This expression can be developed in the same way as eq 1 to determine the fraction of protein migrating with the nucleic acid peak after some distance of sedimentation. Qualitatively, the outcome is that at low concentrations of nucleic acid the middle term of the denominator, $(K_1 + K_2)L$, will dominate, and the apparent binding constant measured will be related to the *sum* of the two binding constants. At much higher concentrations the K_1K_2 term will dominate, and a value related to the *product* of the binding constants will be measured. For a protein with two binding sites of identical affinity, under typical experimental conditions (e.g., at $KL \simeq 10$), the apparent binding constant will be three to five times greater than the individual binding constants.

Cooperatively Binding Proteins. A number of nucleic acid binding proteins have the property that binding of two proteins at contiguous sites on the polynucleotide lattice gives rise to an extra free energy of binding (i.e., binding is cooperative in protein concentration). One can define an intrinsic binding constant, K_{int}, which represents the affinity of the protein for an isolated polynucleotide site, and a cooperativity parameter, ω, which represents the increased binding affinity, relative to an isolated site, for a site next to an already bound protein (McGhee & von Hippel, 1974).

Since band sedimentation experiments are generally performed with a vast excess of polynucleotide binding sites over protein molecules, cooperativity of binding will tend to be overwhelmed by the excess of free binding sites available. For example, if the total protein concentration is 1% of the concentration of binding sites, only when $\omega \simeq 100$ will enough protein be bound at contiguous sites for there to be a significant increase in the apparent binding constant relative to the intrinsic binding constant.

If the concentration of binding protein in an experiment is increased so that the amount of protein becomes significant relative to the number of binding sites, and thus the approximation that $L_{total} \simeq L_{free}$ no longer holds, the fraction of total protein bound (in the absence of cooperativity) decreases (see eq 1), and the apparent binding constant calculated will be too small. If cooperativity is present, however, increasing the concentration of binding proteins will increase the number of proteins binding contiguously to the nucleic acid and hence the number of cooperative interactions; the apparent binding constant approaches $K_{int}\omega$ as the lattice approaches saturation (McGhee & von Hippel, 1974). In an appropriate range of nucleic acid concentrations, a small increase in apparent binding affinity will produce a large change in the fraction of protein bound after sedimentation (see Figure 3). Thus a sensitive test for cooperativity is to determine the binding constant with increasing concentrations of binding protein; any increase in apparent binding affinity is diagnostic of cooperativity. Cooperativity in the binding of *E. coli* ribosomal protein S1 to inactive 30S ribosomal subunits has been detected in this way (Draper & von Hippel, 1979).

Determination of Binding Stoichiometries. A common method for determining the number of specific, high-affinity binding sites for a protein on a nucleic acid or other rapidly sedimenting species (e.g., ribosomes) is to mix the nucleic acid with increasing excesses of protein, sediment, and determine the ratio of protein to nucleic acid in the rapidly sedimenting peak. The maximum ratio attainable is taken to be the maximum number of binding sites. The interaction of ribosomal proteins with the different ribosomal RNAs, for instance, has been examined by this approach (Spierer & Zimmermann, 1978; Newberry et al., 1978), and recently this procedure has been used to demonstrate a change in the

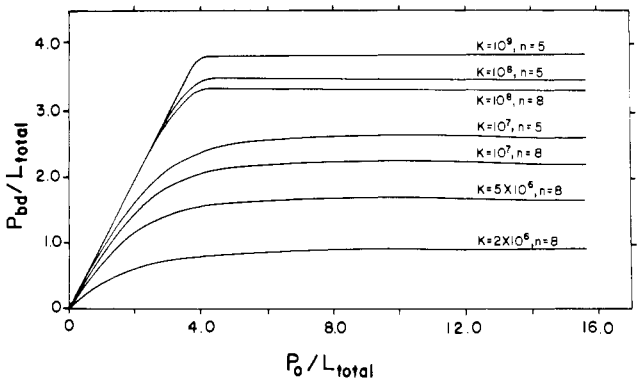

FIGURE 5: Amount of protein cosedimenting with a nucleic acid band as a function of initial protein to nucleic acid binding site ratios. A hypothetical nucleic acid at a concentration of $L_{total} = 10^{-6}$ M is sedimented with a binding protein at concentrations (P_0) indicated in the abscissa. The nucleic acid is assumed to have four identical (nonoverlapping) binding sites for the protein, with intrinsic association constants (K) as indicated. The ratio of protein to nucleic acids (P_{bd}/L_{total}) remaining with the nucleic acid peak after sedimentation over n fractions is plotted against P_0/L_{total}. Details of the calculation are given in the text.

number of sites for binding ribosomal protein S1 to different conformations of the 30S ribosomal subunit (Laughrea & Moore, 1977, 1978). A simple extension of the analysis given in the first section of this paper shows that stoichiometries determined in this way can easily be in error by appreciable factors.

To simulate the sedimentation of a polynucleotide with an excess of binding protein, we apply eq 1 (the mass action law) without the approximation that $L_{free} \simeq L_{total}$

$$\frac{P_{bd}}{P_0} = \frac{K(L_{total} - P_{bd})}{1 + K(L_{total} - P_{bd})}$$

The solution for P_{bd} in terms of K and the total concentration of binding sites (L_{total}) is a quadratic. As before, the concentration of protein which is carried from fraction zero into fraction one is simply the concentration of protein bound in fraction zero; thus to find the concentration of protein carried into fraction n we simply solve n times for P_{bd}, using for P_0 each time the value of P_{bd} from the previous round of calculation.[4]

Figure 5 shows the result for a hypothetical nucleic acid sedimented at a concentration of 10^{-6} M; four nonoverlapping protein binding sites of identical affinity are assumed. Curves calculated for varying binding site affinities and distances of sedimentation (i.e., different values of n) are shown. Several conclusions can be drawn. First, all the curves, regardless of binding affinity or distance of sedimentation, reach apparent maximum binding stoichiometries; however the apparent value may be low relative to the actual value by as much as a factor of four. The plateau level depends on the binding affinity and the distance of sedimentation; thus any treatment of a polynucleotide which appears to alter the binding stoichiometry of the protein might, in fact, be affecting binding site affinities or sedimentation behavior instead. Only in the case of extraordinarily tight binding ($KL > 10^3$) and relatively short sedimentation distances do the apparent binding stoichiometries come close to the actual values. The plateau value, therefore, should be taken as a *minimum* value for the binding stoichiometry, unless the association constant for the binding interaction is known to be sufficiently high.

The reason that the apparent maximum binding stoichiometry observed in sedimentation experiments of this sort is substantially less than the actual value can be understood qualitatively by considering sedimentation of the hypothetical nucleic acid of Figure 5 with an *infinite* binding protein concentration. Initially, of course, all four binding sites will be saturated with protein; thus after sedimentation over one fraction the amount of protein remaining with the nucleic acid will be equal to the concentration of nucleic acid binding sites (i.e., $P_{bd}/L_{total} = 4.0$). However, sedimentation through each additional fraction will reduce this saturation level by an amount dependent on KL, regardless of the initial saturation ratio. Thus the net result will be a significant loss of protein from the complex after some distance of sedimentation (unless $KL > 10^3$), even when the *initial* ratio of protein to binding sites is essentially infinite.

Conclusions

The method described here for analyzing band sedimentation data can provide accurate estimates of macromolecular association constants ranging from $\sim 10^3$ to 10^{11} M^{-1}. The ease with which a determination of K can be carried out makes this method particularly applicable to surveys of binding conditions (salt concentrations, temperature, pH) or binding specificities (RNA vs. DNA, native vs. denatured DNA). We should emphasize that this method is applicable to any interaction in which the two components differ sufficiently in sedimentation coefficients; in principle the technique can also be applied when complex and free protein is separated by electrophoresis or by gel filtration rather than by sedimentation. Not only protein–nucleic acid interactions but also protein–macromolecular complex interactions (as illustrated by the cited measurement of the affinity of S1 protein for 30S ribosomal subunits) and small molecule–protein interactions (where the protein becomes the fast-sedimenting component) can, in principle, be analyzed by various modifications of this technique.

Added in Proof

Since this manuscript has been in press, the method of "sedimentation partition chromatography", devised by Yamamoto & Alberts (1974) for measuring binding constants of estradiol receptor protein to nonspecific DNA, has come to our attention. While quite different in detail, this method is based on the same multiple equilibrium principles underlying the present approach and represents an alternative procedure for use under certain experimental conditions. We are grateful to Bruce Alberts for pointing this out to us.

References

Burgi, E., & Hershey, A. D. (1963) *Biophys. J. 3*, 309–321.
Cann, J. R. (1970) *Interacting Macromolecules: The Theory and Practice of Their Electrophoresis, Ultracentrifugation and Chromatography*, Academic Press, New York.
deHaseth, P. L., Gross, C. A., Burgess, R. R., & Record, M. T., Jr. (1977) *Biochemistry 16*, 4777–4782.
Draper, D. E., & von Hippel, P. H. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, (in press).
Draper, D. E., Pratt, C., & von Hippel, P. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4786–4790.
Harrington, W. F., & Kegles, G. (1973) *Methods Enzymol. 27*, 306–345.

[4] The calculation can also be carried out taking into account band spreading as before. The effect of this is to lower the plateau level of the curves in Figure 5 by 5–20%, the largest effect being associated with the smallest K.

Hinkle, D. C., & Chamberlin, M. J. (1970) *Cold Spring Harbor Symp. Quant. Biol. 35*, 65–72.
Hinkle, D. C., & Chamberlin, M. J. (1972) *J. Mol. Biol. 70*, 157–185.
Jensen, D. E., & von Hippel, P. H. (1976) *J. Biol. Chem. 251*, 7198–7214.
Jensen, D. E., & von Hippel, P. H. (1977) *Anal. Biochem. 80*, 267–281.
Laughrea, M., & Moore, P. B. (1977) *J. Mol. Biol. 112*, 399–421.
Laughrea, M., & Moore, P. B. (1978) *J. Mol. Biol. 121*, 411–430.
McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469–489.
Molineaux, I. J., Pauli, A., & Gefter, M. L. (1975) *Nucleic Acids Res. 2*, 1821–1837.
Newberry, V., Brosius, J., & Garrett, R. (1978) *Nucleic Acids Res. 5*, 1753–1766.
Reuben, R. C., & Gefter, M. L. (1974) *J. Biol. Chem. 249*, 3843–3850.
Revzin, A., & von Hippel, P. H. (1977) *Biochemistry 16*, 4769–4776.
Riggs, A. D., Suzuki, H., & Bourgeois, S. (1970) *J. Mol. Biol. 48*, 67–83.
Silhavy, T. J., Szmelcman, S., Boos, W., & Schwartz, M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2120–2124.
Spierer, P., & Zimmermann, R. A. (1978) *Biochemistry 17*, 2474–2479.
Tal, M., Aviram, M., Kanarek, A., & Weiss, A. (1972) *Biochim. Biophys. Acta 281*, 381–392.
von Hippel, P. H. (1978) in *Biological Regulation and Development* (Goldberger, R., Ed.) Plenum Press, New York (in press).
Yamamoto, K. R., & Alberts, B. (1974) *J. Biol. Chem. 249*, 7076–7086.
Yarus, M., & Berg, P. (1970) *Anal. Biochem. 35*, 450–465.

An Octamer of Core Histones in Solution: Central Role of the H3·H4 Tetramer in the Self-Assembly†

Adolfo Ruiz-Carrillo* and José Luis Jorcano

ABSTRACT: The association of histones H2A, H2B, H3, and H4 in solution has been studied. In 2 M NaCl and at neutral pH they can assemble in a complex in which each histone is present in equimolar amounts. The complex has a weight average molecular weight of 98 000 (±3700) and a sedimentation coefficient ($s^0_{20,w}$) of 4.8. The value of the weight average molecular weight and the histone stoichiometry indicate that the complex is an octamer. The pairs of histones H2A,H2B and H3,H4 studied separately under identical conditions only associate as equimolar complexes consistent with dimeric and tetrameric structures, respectively. The stability of the core histone octamer is a function of the ionic strength, pH, and concentration of protein. The octamer dissociates by losing dimers of H2A,H2B until the main complexes existing in solution are the H3·H4 tetramer and the H2A·H2B dimer. This process is reversible upon reestablishing the original conditions.

The nucleosome core particle has been shown to consist of approximately 140 base pairs of DNA complexed with the core histones (H2A, H2B, H3, and H4) (Sollner-Webb & Felsenfeld, 1975; Axel, 1975). In recent years evidence has accumulated in support of the view that the four core histones occur in the vast majority of the core particles (Kornberg & Thomas, 1974; Boseley et al., 1976; Camerini-Otero et al., 1976; Simpson & Bustin, 1976; Bustin et al., 1977; Ruiz-Carrillo & Jorcano, 1977; Oudet et al., 1977; Jorcano & Ruiz-Carrillo, 1979). In addition, hydrodynamic (Shaw et al., 1976) and cross-linking (Thomas & Kornberg, 1975a,b) studies have indicated that an octamer containing two of each core histone makes up the protein moiety of the core particle. Whether the core histone octamer can exist in the absence of DNA is at present a matter of controversy.

In solutions of low ionic strength, histones interact in pairs H2A,H2B (dimer) and H3,H4 (tetramer) (Kelley, 1973; Roark et al., 1974; Kornberg & Thomas, 1974), although a relatively strong cross-interacting dimer H4·H2B has also been characterized (D'Anna & Isenberg, 1974). The way in which the core histones associate in solution is, however, not well established. Weintraub et al. (1975) reported that in solutions of high ionic strength (i.e., 2 M NaCl) and neutral pH, they associate in "heterotypic" tetramers (H2A·H2B·H3·H4), a view later supported by laser light scattering studies (Campbell & Cotter, 1976). On the other hand, cross-linking studies have suggested that an octamer may exist under similar conditions (Thomas & Kornberg, 1975a,b). During the course of this work, Thomas & Butler (1977) have also reported hydrodynamic studies which are consistent with the existence of a core histone octamer in solution. A clarification of this controversy will contribute not only to our understanding of the structure of the nucleosome but also to the way in which chromatin is assembled during DNA replication. We present evidence demonstrating the occurrence of a core histone octamer in solutions of high ionic strength and describe some of its properties.

Experimental Procedures

Preparation of Cells and Nuclei. Mature hen erythrocytes were obtained as described previously (Ruiz-Carrillo et al., 1975). Cells were lysed in buffer A[1] (Hewish & Burgoyne,

† From the Max-Planck-Institut für Molekulare Genetik, 1 Berlin 33 (Dahlem), West Germany. *Received July 10, 1978.* J.L.J. was supported by an EMBO long-term fellowship.

0006-2960/79/0418-0760$01.00/0