Proc. Natl. Acad. Sci. USA
Vol. 73, No. 12, pp. 4467–4470, December 1976
Biochemistry

Determination of molecular weight of the protein moiety in protein–detergent complexes without direct knowledge of detergent binding

(membrane proteins/lipoproteins/sedimentation equilibrium)

JACQUELINE A. REYNOLDS AND CHARLES TANFORD

Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710

Contributed by Charles Tanford, October 6, 1976

ABSTRACT **Sedimentation equilibrium measurements can be used to determine the molecular weight of the protein moiety of a protein–detergent complex without prior knowledge of detergent binding. The procedure is to adjust the solvent density by addition of D_2O so as to blank out the contribution of bound detergent to the sedimentation potential. An approximate measure of detergent binding can be obtained from the effect of solvent density on the sedimentation result. The procedure is also applicable to protein–lipid complexes. It can be used for complexes containing both lipid and detergent if the lipid content is known. The use of the method is demonstrated by experimental data for the AI polypeptide of serum high density lipoprotein, in separate complexes with nonionic detergents and with a phospholipid.**

Many membrane proteins can be solubilized only in the form of protein–detergent complexes. Nonionic detergents are being used for this purpose with increasing frequency because they are less likely than ionic detergents to cause alterations in the state of aggregation or other conformational properties of the protein (1, 2). For this reason one of the measurements one wants to make in the presence of nonionic detergents is that of molecular weight, because there is a good chance that it may reflect the state of aggregation of the protein in its native state in the membrane. The molecular weight that is sought is, however, that of the protein alone, exclusive of bound detergent (or lipid, if present). The desired molecular weight can be obtained by use of the analytical ultracentrifuge if the amount of detergent (and lipid, if present) associated with the protein in the solubilized complexes can be determined, as we have demonstrated in a previous paper (3). The determination of bound detergent is unfortunately not feasible for many nonionic detergents because they are not available in radioactive form and because suitable methods for analysis by chemical means do not exist.

The purpose of this paper is to show that the difficulty created by inability to determine the amount of bound detergent can be circumvented by using D_2O to adjust the density of the solvent to the point where it matches the buoyant density of the detergent. The detergent then makes no contribution to the sedimentation potential, no matter how much is bound. The method of course requires that the detergent's buoyant density be very different from that of the protein, a condition satisfied by all nonionic detergents in common use. It is also satisfied by biological lipids, so that the method can be used for soluble protein–lipid complexes as well.

The use of D_2O to alter solvent density in sedimentation analysis has been used previously by Edelstein and Schachman (4) in a procedure for the indirect determination of protein partial specific volume. Their purpose was different from ours, but the underlying physical principle is the same. D_2O has been used to blank out the detergent contribution in neutron scattering by protein–detergent complexes (5), and sucrose has been used for the same purpose in x-ray scattering (6).

Abbreviations: HDL, high density serum lipoprotein; cmc, critical micelle concentration. The symbol C_nE_x is used to represent detergents of the type $C_nH_{2n+1}(OCH_2CH_2)_xOH$.

THEORETICAL BASIS

Treatment of sedimentation equilibrium in terms of the thermodynamics of multicomponent solutions (7) shows that the experimentally observable protein concentration gradient in a solution, in which all manner of interactions may be taking place, is a measure of $M_P(1 - \phi'\rho)$, where M_P is the molecular weight of the protein moiety in the sedimenting particle, ρ is the solvent density, and ϕ' is the volume increment per g of protein, measured under conditions where the *chemical potentials* of all solvent components are kept constant. The measurement of ϕ' is difficult and time-consuming, and especially so in the presence of detergents. It is thus preferable to replace $M_P(1 - \phi'\rho)$ by an equivalent expression (7) in terms of the separate contribution of the individual constituents of the sedimenting particle. For a particle containing δ_D g of detergent and δ_L g of lipid, each per g of protein, the appropriate expression is

$$M_P(1 - \phi'\rho) = M_P[(1 - \bar{v}_P\rho) + \delta_D(1 - \bar{v}_D\rho) + \delta_L(1 - \bar{v}_L\rho)] \quad [1]$$

where $\bar{v}_P$ is the partial specific volume of the protein (i.e., volume increment per g of protein measured under conditions where the *masses* of other components are kept constant) and $\bar{v}_D$ and $\bar{v}_L$ are the partial specific volumes of detergent and lipid, respectively, in their protein-bound states. These latter quantities are no less difficult to measure than ϕ', and to make Eq. 1 usable one has to be willing to assume that partial specific volumes measured for the unbound components under appropriate circumstances can be substituted. For detergents and lipids it is in fact reasonable to assume that $\bar{v}$ values measured for the *self-associated* detergent or lipid (in micelles or bilayers) will not differ significantly from the $\bar{v}$ values of the protein-associated constituents (3). We have previously tested this assumption by applying Eq. 1 to sedimentation equilibrium results for protein–detergent complexes under conditions where δ_D could be measured directly. (No association with lipid was involved.) In one test (3) we used the AI apoprotein of human high density serum lipoprotein (HDL), the molecular weight of which is known because the amino acid sequence has been determined. Three different detergents were used ($\bar{v}_D$ ranging from 1.110 to 0.778 cm^3/g) and the measured values of $M_P(1 - \phi'\rho)$ varied over a 3-fold range. The values of M_P calculated by means of Eq. 1 were all within 5% of the correct value. In the second test (8), using the *M,N*-sialoglycoprotein of human

erythrocyte membranes, dissociated to the monomeric state by sodium dodecyl sulfate, Eq. 1 was used to obtain an M_P value far below the molecular weight suggested by previous workers. The result obtained was subsequently confirmed by the determination of the amino acid sequence of the protein (9).

To circumvent the problem created when δ_D cannot be measured, we have made use of the ability to change the density of the solvent by substitution of D_2O for H_2O. By interpolation or extrapolation we can obtain $M_P(1 - \phi'\rho)$ at the density $\rho = 1/\bar{v}_D$. At this point $1 - \bar{v}_D\rho = 0$ and the contribution of detergent to Eq. 1 vanishes, *regardless of the value of* δ_D. The same technique can be used to blank out the contribution of bound lipid. If both detergent and lipid are present, both cannot of course be blanked out simultaneously.

It is essential that D_2O be used to adjust the density, rather than sucrose, CsCl, or other added solute. A high concentration of additive will be necessary to achieve a reasonable range of density and it cannot then be assumed that preferential exclusion of solute from the incorporated solvent would be a negligible factor (10, 11). This would require the addition of another term in Eq. 1, the magnitude of which could not be readily determined. On the other hand, D_2O and H_2O should exchange freely whenever they occur, so that the isotopic composition of protein-bound water should remain the same as that of the solvent. Even exchangeable protons of the protein molecule are expected to exchange freely, which leads to a small increase in protein molecular weight as the D_2O content of the solvent is increased. If we make the reasonable assumption (4) that the molecular volume is unaffected by isotopic substitution, the change in M_P will lead to a reciprocal change in $\bar{v}_P$ which has to be taken into account when Eq. 1 is applied. If the detergent or lipid being used has exchangeable protons, a similar small decrease in $\bar{v}_D$ or $\bar{v}_L$ will occur.

An assumption inherent to this procedure, and to the conceptually similar methods mentioned in the introduction, is that the composition of the sedimenting particle is unaffected by the substitution of D_2O for H_2O. In the rare instances where this assumption might prove incorrect, marked curvature in a plot of $M_P(1 - \phi'\rho)$ against ρ should be observed. It may be noted that such a plot should normally be very nearly linear, and that the slope can be used as an approximate measure of detergent binding. When lipid is absent, for example,

$$-d[M_P(1 - \phi'\rho)]/d\rho = M_P(\bar{v}_P + \delta_D\bar{v}_D). \qquad [2]$$

The evaluation of δ_D can only be approximate, unless δ_D is large, because it depends on the difference between the measured slope and the term $M_P\bar{v}_P$ alone. (Eq. 2 neglects the expected small change in $\bar{v}$ values with density.)

It may be noted that it is theoretically permissible to use sedimentation velocity to determine $M_P(1 - \phi'\rho)$ if the Stokes radius of the particle is known (Eq. 7 of ref. 3), with considerable saving in time as compared to equilibrium sedimentation. In the present procedure this means that equilibrium data need be obtained at only one density. Sedimentation velocity measurement on the same solution will yield a value for the Stokes radius, which, like other molecular parameters, is likely to be unaffected by the substitution of D_2O for H_2O, and should therefore remain constant as the density is varied.

EXPERIMENTAL PROCEDURE

The AI apoprotein of HDL was prepared as previously described (12). The exact molecular weight of the polypeptide chain, based on the amino acid sequence (13), is 28,342. The detergent $C_{12}E_8$ was a homogeneous compound from Nikko Chemicals Co., Tokyo, Japan. It was recrystallized from hexane before use. The commercial detergent Lubrol WX (Sigma Corp.) is stated by the manufacturer (ICI United States, Inc.) to be an approximately equimolar mixture of $C_{16}E_x$ and $C_{18}E_x$. The number of oxyethylene groups follows a Poisson distribution (14), with an average value of x of 16.4. Didecanoyllecithin was synthesized for us by Dr. W. L. Stone. The procedure and tests for purity will be described in a subsequent paper. D_2O (99.84 mol %) was obtained from Bio-Rad Laboratories, and its density was checked periodically to guard against possible contamination by atmospheric H_2O.

Sedimentation equilibrium measurements were made using a Beckman model E analytical ultracentrifuge, and protein concentration gradients were determined using the photoelectric scanner at a wavelength of 280 nm. The temperature at equilibrium was measured for each run, and ranged from 21 to 23°. Solutions for measurement were prepared from a protein stock solution in H_2O, containing 0.1 M NaCl, 0.02 M Tris· Cl (pH 8.3), 1.0 mM EDTA, and 1.0 mM NaN_3. Aliquots of this solution were diluted by weight with D_2O, and solvent densities were calculated with the aid of appropriate tabulated data (15, 16), assuming that the partial specific volumes of buffer and salts could be taken as independent of isotopic composition. Sedimentation velocities were used to measure $M_P(1 - \phi'\rho)$ in one determination: the solvent viscosities required (Eq. 7 of ref. 3) were obtained from the same sources as the densities.

Partial specific volumes for the detergents were determined by means of an Anton Paar precision densimeter, as previously described (3). We obtained $\bar{v} = 0.973$ cm^3/g for $C_{12}E_8$ and $\bar{v} = 0.929$ cm^3/g for Lubrol WX, both at 25° in H_2O. Each detergent contains one exchangeable proton, and the effect of full replacement by deuterium would be to diminish $\bar{v}$ by 0.002 and 0.001 cm^3/g, respectively. The value of $\bar{v}$ for didecanoyllecithin (0.927 cm^3/g) was based on the measurements for shorter chain homologs by Tausk *et al.* (17). The published molar volumes are for lecithin monohydrates and were corrected to the anhydrous state by subtracting the molar volume of H_2O. The increment due to the longer alkyl chain length was calculated by adding 16.1 cm^3 to the molar volume per added CH_2 group. Lecithins have no exchangeable protons. For the partial specific volume for the protein (in H_2O) we used the calculated value of 0.735 cm^3/g, based on the amino acid sequence, as was done previously (3). Such calculated values usually agree closely with measured values when enough material is available to make the measurement. An example is provided by the *M,N*-sialoglycoprotein of the erythrocyte membrane (8). It should be noted that an error of 0.01 cm^3/g in $\bar{v}_P$ would alter the value of M_P estimated from the data by only 5%. The AI polypeptide contains about 440 exchangeable protons per molecule, which would increase the molecular weight by 1.5% for complete isotopic substitution in pure D_2O, and proportionally less in D_2O–H_2O mixtures. There is a corresponding decrease in $\bar{v}_P$ (to 0.724 cm^3/g for complete substitution). The appropriate corrections were made in obtaining molecular weights from the data.

RESULTS

Measurements were made at detergent or lipid concentrations that were deemed sufficient to assure saturation of the protein. All individual runs yielded linear plots of logarithm of the concentration against the square of radial displacement, to as close to the bottom of the cell as one could read, indicating that one is dealing with a single sedimenting particle whose properties are not affected by the slight gradient of detergent or lipid

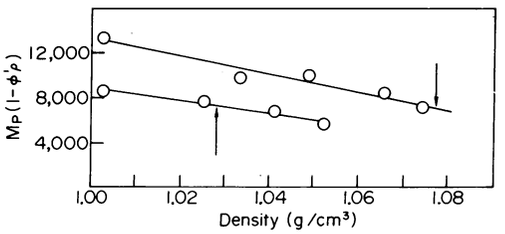

FIG. 1. Results for the AI apoprotein of HDL in Lubrol WX (upper line) and in $C_{12}E_8$ (lower line). Total detergent concentrations were close to 5 mM, which is a factor of more than 50 above the cmc for $C_{12}E_8$ and a factor of about 500 for Lubrol WX. The protein concentration was about 5 μM. Arrows indicate the value of $\rho = 1/\bar{v}_D$ for each detergent.

concentration that must exist at any density other than $\rho = 1/\bar{v}_D$, nor by the increase in protein concentration that occurs in the cell as the bottom of the cell is approached.

Results at different densities in $C_{12}E_8$ and Lubrol WX are shown in Fig. 1. The densities at which $1 - \bar{v}_D\rho$ is equal to zero are shown by arrows, and the protein molecular weights obtained from the values of $M_P(1 - \phi'\rho)$ at those points are 29,000 for $C_{12}E_8$ and 32,000 for Lubrol WX. The former is in excellent agreement with the polypeptide chain molecular weight (28,500 after correction for the probable level of substitution of D for H), indicating that AI is a monomer in its complexes with $C_{12}E_8$, in agreement with what has been found for complexes of AI with other detergents (18, 19).

The result obtained in Lubrol WX, however, is higher than expected (corrected molecular weight 28,600) by about 12%, a discrepancy that is probably outside the purely experimental uncertainty. The linearity of the semilogarithmic plots of the data precludes partial self-association as an explanation, and one is forced to conclude that the protein is again in the monomeric state, but that an incorrect value for the molecular weight is obtained. This error can almost certainly be attributed to the heterogeneity of the detergent and the resulting likelihood that the average composition of bound detergent (in terms of individual species of the mixture) is not the same as that of the detergent as supplied. A difference between these compositions is in fact theoretically predictable, as the following argument shows. (*i*) All previous studies indicate that the free energy of association between detergents and AI is primarily hydrophobic in origin but only slightly dependent on alkyl chain length above 10 carbons (19). The binding constants for all species in the mixture of $C_{16}E_x$ and $C_{18}E_x$ should therefore be about the same, and the amount of each component bound should be proportional to its thermodynamic activity (a_i) in solution. (*ii*) The values of a_i are, however, not expected to be in the same proportion as the mole fractions (X_i) in the micellar state, i.e., the state to which the $\bar{v}$ measurement refers. If the micellar phase can be considered to be thermodynamically ideal, we expect to have $a_i = (\text{cmc})_i X_i$, where $(\text{cmc})_i$ is the critical micelle concentration (cmc) of micelles formed from a single component of the mixture (20). (*iii*) Protein-bound detergent should thus be enriched, relative to detergent micelles, in species with higher $(\text{cmc})_i$, which means species with the shorter alkyl chain and with the larger value of x, as can be seen from the cmc values determined for pure detergents of the C_nE_x type (21). Both shortening the alkyl chain and increasing the number of oxyethylene units decrease $\bar{v}$, so that the protein-bound detergent should have a smaller $\bar{v}$ than the bulk sample. A value of $\bar{v}_D = 0.919$ cm^3/g instead of the experimental value of 0.929 cm^3/g for the micellar state would in fact yield the correct monomer molecular weight from the experimental data. It may be noted that a single additional oxyethylene group decreases $\bar{v}_D$ by about 0.005 cm^3/g, and the effect of removing one CH_2 group is about half as great. The observed result is thus entirely reasonable.

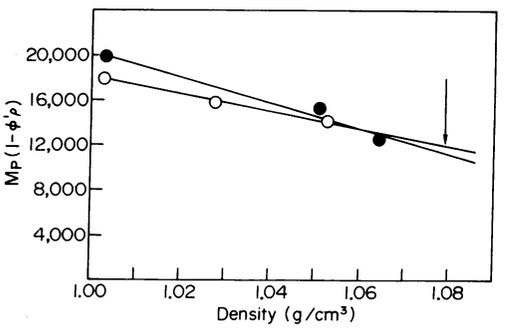

FIG. 2. Results for the AI apoprotein of HDL in didecanoyllecithin, at total lipid concentrations of 1.29 mM (open circles) and 1.94 mM (filled circles). The cmc is 4.9 μM. The arrow indicates $\rho = 1/\bar{v}_D$.

Fig. 2 shows the results obtained for complexes between AI and didecanoyllecithin, at two different concentrations of lipid. In spite of the large excess of lipid, the two solutions proved to correspond to somewhat different levels of binding, both short of saturation, but there was no curvature of the sedimentation plots, and the failure to achieve saturation thus did not interfere with the molecular weight determination. The molecular weights of the protein moiety that were obtained from the two sets of experiments are 55,500 and 52,500. Both are close to twice the polypeptide chain molecular weight, i.e., the data indicate that the AI polypeptide is in a dimeric state in these complexes with a diacyl lipid, in contrast to the monomeric state observed with detergents in this and in previous studies. This is an interesting result because *native* HDL, in association with lipid, invariably contains two AI chains per molecule regardless of the density form of the lipoprotein that is examined (22). This aspect of the interaction of AI with didecanoyllecithin, as well as the high lipid concentration required to achieve saturation, will be discussed in detail in a forthcoming paper.

The approximate level of detergent binding calculated from these data by Eq. 2 is 70 mol/mol of AI for both $C_{12}E_8$ and Lubrol WX. No attempt was made to correct for preferential association of individual detergent complexes in the case of Lubrol WX. The two sets of data in Fig. 2 correspond to 35 and 78 mol of bound lipid per mol of AI.

DISCUSSION

The method described here can be used for any pure protein in a complex with detergent or lipid. If both are present, the binding of one of them (normally the lipid) would have to be measured independently. The partial specific volume of the detergent must be significantly different from that of the protein: a value of $\bar{v}_D > 0.90$ cm^3/g is desirable for optimal accuracy. Only microgram quantities of protein are needed.

The most important applications of the method are likely to be in the category illustrated by the data of Fig. 2, i.e., the determination of the state of aggregation under conditions approaching the native state, where the molecular weights of completely dissociated polypeptide chains are already known. We have made one such application previously, to determine the state of aggregation of Ca^{++}-stimulated ATPase from sarcoplasmic reticulum, in complexes that retained enzymatic activity and contained both bound lipid and detergent (23). The experiments were preliminary in nature, and the results not as

accurate as those in this paper. Nevertheless, they showed unambiguously that the active enzyme was an oligomer, containing three or four ATPase polypeptide chains.

A further illustration of how powerful this method may be is provided by a crude calculation of the protein contribution to the molecular weight of native human low density lipoprotein, based on experimental data of Fisher *et al.* (24). These workers determined $M_P(1 - \phi'\rho)$ from sedimentation velocity and diffusion data, the precision of the latter being relatively poor. They obtained data in H_2O and D_2O, to evaluate $\bar{v}$ for the whole lipoprotein particle, and used it to obtain the molecular weight of the whole particle. Analytical data for the lipid/protein ratio are then needed to determine M_P. We have used only the two values of $M_P(1 - \phi'\rho)$ from this study, and a $\bar{v}_L$ value of 1.030 cm^3/g for the lipid, which is appropriate for the average lipid composition of this class of lipoproteins (25). The two available data points were extrapolated to $\rho = 1/\bar{v}_L$, and yielded (with a calculated $\bar{v}_P$ value) a protein molecular weight of 570,000. Considering the fact that the data were not designed for this purpose, this result is in satisfactory agreement with the presumably true M_P of 510,000 ± 25,000, corresponding to a content of two polypeptide chains per particle (25, 26). The remarkable aspect of this calculation is that the native lipoprotein particle contains 80% lipid and only 20% protein. The result suggests that our method can be extended to other lipid-rich particles, such as membrane proteins embedded in phospholipid vesicles.

The major source of error in the data presented here comes from uncertainty in the value of $\bar{v}_D$ or $\bar{v}_L$, i.e., uncertainty in the exact density required to blank out the contribution of detergent or lipid. The possible error increases when heterogeneous commercial detergents are used because the composition of bound detergent may differ from that of micellar detergent, as illustrated by the results obtained for AI in Lubrol WX. Even here, however, the result obtained differs from the correct molecular weight by only 12%. Much larger errors are likely when semi-empirical methods, such as the procedure of Siegel and Monty (27), are applied to protein-detergent complexes, as has been done in a number of recent studies.

This work was supported by Research Grants HL-14882 and AM-04576 from the National Institutes of Health and by a grant from the National Science Foundation.

1. Helenius, A. & Simons, K. (1972) *J. Biol. Chem.* **247**, 3656–3661.
2. Makino, S., Reynolds, J. A. & Tanford, C. (1973) *J. Biol. Chem.* **248**, 4926–4932.
3. Tanford, C., Nozaki, Y., Reynolds, J. A. & Makino, S. (1974) *Biochemistry* **13**, 2369–2375.
4. Edelstein, S. J. & Schachman, H. K. (1967) *J. Biol. Chem.* **242**, 306–311.
5. Schoenborn, B. P. (1976) *Biochim. Biophys. Acta* **457**, 41–55.
6. Sardet, C., Tardieu, A. & Luzzati, V. (1976) *J. Mol. Biol.* **105**, 383–407.
7. Casassa, E. F. & Eisenberg, H. (1964) *Adv. Protein Chem.* **19**, 287–395.
8. Grefrath, S. P. & Reynolds, J. A. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3913–3916.
9. Tomita, M. & Marchesi, V. T. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2964–2968.
10. Vinograd, J. & Hearst, J. E. (1962) *Fortschr. Chem. Org. Naturst.* **20**, 372–422.
11. Timasheff, S. N., Lee, J. C., Pittz, E. P. & Tweedy, N. (1976) *J. Colloid Interface Sci.* **55**, 658–663.
12. Reynolds, J. A. & Simon, R. H. (1974) *J. Biol. Chem.* **249**, 3937–3940.
13. Baker, H. N., Gotto, A. M., Jr. & Jackson, R. L. (1975) *J. Biol. Chem.* **250**, 2725–2738.
14. Shachat, N. & Greenwald, H. L. (1967) in *Nonionic Surfactants*, ed. Schick, M. J. (Marcel Dekker, New York), pp. 8–43.
15. *Landolt-Börnstein Tabellen* (1969–1971) (Springer-Verlag, Berlin), 6th ed.
16. Jones, G. & Fornwalt, H. J. (1936) *J. Chem. Phys.* **4**, 30–33.
17. Tausk, R. J. M., Van Esch, J., Karmiggelt, J., Voordouw, G. & Overbeek, J. Th. G. (1974) *Biophys. Chem.* **1**, 184–203.
18. Makino, S., Tanford, C. & Reynolds, J. A. (1974) *J. Biol. Chem.* **249**, 7379–7382.
19. Haberland, M. E. & Reynolds, J. A. (1975) *J. Biol. Chem.* **250**, 6636–6639.
20. Tanford, C. (1973) *The Hydrophobic Effect* (John Wiley & Sons, New York), pp. 81–85.
21. Becher, R. (1967) in *Nonionic Surfactants*, ed. Schick, M. J. (Marcel Dekker, New York), pp. 478–515.
22. Friedberg, S. J. & Reynolds, J. A. (1976) *J. Biol. Chem.* **251**, 4005–4009.
23. le Maire, M., Møller, J. V. & Tanford, C. (1976) *Biochemistry* **15**, 2336–2342.
24. Fisher, W. R., Granade, M. E. & Mauldin, J. L. (1971) *Biochemistry* **10**, 1622–1629.
25. Schumaker, V. N. (1973) *Acc. Chem. Res.* **6**, 398–403.
26. Smith, R., Dawson, J. R. & Tanford, C. (1972) *J. Biol. Chem.* **247**, 3376–3381.
27. Siegel, L. M. & Monty, K. J. (1966) *Biochim. Biophys. Acta* **112**, 346–362.