



DETERMINATION OF MOLECULAR WEIGHT OF GLYCOPROTEINS BY ANALYTICAL ULTRACENTRIFUGATION

S. J. Shire

Genentech, Inc.

South San Francisco, CA 94080

Introduction

Analytical ultracentrifugation is a classical technique that has played a critical role in laying the foundations for modern molecular biology. Among its achievements is the demonstration that proteins are macromolecules rather than complexes of smaller units, and direct support for the semiconservative replication mechanism of DNA as proposed by Watson and Crick. Unlike techniques such as SDS PAGE or gel permeation chromatography, analytical ultracentrifugation can be used to determine absolute molecular weights without the use of molecular weight standards or interference from the sieving matrix used for separation.

With the advent of recombinant DNA technology and the explosion of protein drugs being developed, there has been a renewed interest in analytical ultracentrifugation. Many of the proteins being developed as pharmaceuticals exist as glycosylated proteins. It is unclear how the composition or content of carbohydrate affects behavior on typical gel permeation chromatography media. Often the standards used to calibrate the chromatography are globular proteins and are inadequate to obtain an accurate molecular weight of the glycosylated protein in solution. The apparent molecular weights determined may be sufficiently in error to result in the conclusion that the protein exists in an associated state in solution. The experiments presented here show two examples using glycoproteins produced at Genentech, Inc.: the recombinant DNA-derived envelope glycoprotein, rgp120 IIIIB, of human immunodeficiency virus type 1, and sTNF-R1, the extracellular domain of human tumor necrosis factor (TNF- α) type 1 receptor. Gel permeation chromatography suggests that these molecules exist as dimers in solution. Molecular weights, determined by analytical sedimentation equilibrium experiments, showed these conclusions to be incorrect.

Materials and Methods

rgp120 IIIIB

The recombinant DNA-derived envelope glycoprotein of human immunodeficiency virus type 1 (rgp120 IIIIB) was expressed as a fusion protein in Chinese hamster ovary cells and purified by immunoaffinity chromatography (1). The fusion protein has a molecular weight of ≈ 53 kDa and consists of a short segment of the herpes gD1 protein fused to the gp120 protein at the amino terminus. However, SDS PAGE analysis reveals a broad band centered at approximately 120 kDa because it is glycosylated at 24 Asn residues. The broadness of the electrophoretic band is typical of glycosylated proteins and may reflect the heterogeneity of the carbohydrate moieties. Gel permeation chromatography yields apparent molecular weights that range from 200 to 280 kDa suggesting that the rgp120 IIIIB self associates.

Ultracentrifugation of rgp 120 IIIB

The rgp120 IIIB isolate at 0.4 mg/mL in 20 mM sodium phosphate, 117 mM NaCl, pH 6.5 was centrifuged at 10,000 rpm for ≈ 48 h at 20°C. The resulting absorbance gradient at 280 nm was fit to a single ideal species. The partial specific volume, 0.68 for this protein, was estimated from the amino acid and carbohydrate composition. The buoyant molecular weight determined from this analysis was converted to a molecular weight by dividing by $(1 - \bar{v} \rho)$, where the density, $\rho = 1.005$ g/mL.

sTNF-R1

The extracellular domain of human tumor necrosis factor (TNF- α) type 1 receptor (sTNF-R1) was expressed in 293S human embryonic kidney cells (2). The secreted soluble receptor was purified by chromatography on a TNF- α affinity column and reversed phase HPLC. The molecular mass of this protein is ≈ 19 kDa based on the theoretical amino acid composition derived from the protein sequence encoded by the cDNA. This protein has three potential N-linked glycosylation sites and this is reflected in its electrophoretic migration in SDS polyacrylamide gels. The purified protein migrates on SDS PAGE as two major bands with apparent molecular weights of 28 and 33 kDa and a faint minor band at ≈ 24 kDa. The presence of these bands are due to small differences in N-terminal and C-terminal processing as well as heterogeneity in the carbohydrate composition. Analysis by gel permeation chromatography yields an apparent molecular weight of 55-60 kDa, and suggests that this protein exists as a dimer in solution.

Ultracentrifugation of sTNF-R1

sTNF-R1 at 75 μ g/mL in PBS was sedimented to equilibrium at 15,000 rpm in a Proteomelab XL-A analytical ultracentrifuge over a period of 18 to 24 h at 20°C. The absorbance as a function of radial position was determined at 232 nm. Buoyant molecular weights, $M(1 - \bar{v} \rho)$, were determined by analyzing the absorbance gradient as one ideal species as described below. Molecular weights were computed from the buoyant molecular weights using estimated values for the partial specific volume of the glycosylated protein. The partial specific volume for sTNF-R1 was calculated assuming that the mass of sTNF-R1 is between 10 and 30% carbohydrate and that the contribution from a carbohydrate chain is 0.63.

Data Analysis

The sedimentation equilibrium data for sTNF-R1 and rgp120 IIIB were analyzed as a single ideal species. The absorbance, A_r at any radial position, r , is related to the molecular weight, M , by

$$A_r = A_{r_0} e^{[M(1 - \bar{v} \rho)\omega^2/2RT] (r^2 - r_0^2)}$$

where A_{r_0} is the absorbance at a radial reference distance, r_0 , ω is the angular velocity, R is the gas constant, T is the temperature in Kelvin, \bar{v} is the partial specific volume of the molecule, and ρ is the solution density. The reference radial distance, r_0 , for the analysis was set to a radial position 2/3 of the column height. The data were fitted to this equation either using a Simplex algorithm implemented in Turbo Pascal, version 5.5 (3), or with the general curve fitting routines in the commercially available graphics software package, KaleidaGraph™.

Results and Discussion

Analysis as a single ideal species of the sedimentation equilibrium concentration gradient in the cell yields buoyant molecular weights. In order to convert these values to absolute molecular weights, it is necessary to determine the partial specific volume for the glycoprotein. Experimentally this is done with high-precision density measurements as a function of protein concentration. Standard physicochemical methods such as pycnometry require an inordinate amount of material. It is possible to obtain \bar{v} values using density meters that make use of the mechanical oscillator technique (4), but this technique also requires a fair amount of protein that may not always be available. Alteration of solvent density using D₂O or D₂O¹⁸ in sedimentation velocity or sedimentation equilibrium experiments can also be used to determine \bar{v} (5). An alternative method is to compute the partial specific volume from a weight average of the partial specific volumes of the component amino acid residues. These calculated and observed partial specific volumes of many proteins are in good agreement (6, 7).

In the case of a glycoprotein, a similar calculation can be used to estimate partial specific volume if the carbohydrate content is known, and again good agreement between experimental and calculated values has been shown for glycoproteins (8). Obviously, if the carbohydrate content is unknown, it is difficult to make estimates for partial specific volume for a glycoprotein. However, the types of carbohydrate structures found in glycoproteins are generally understood, and it is possible to estimate the contribution from these oligosaccharide chains. Table 1 shows the typical type of N-linked carbohydrates found in glycoproteins along with an estimate of partial specific volume for each type of structure. The estimated values range from 0.62 to 0.64 mL/g, and thus it would appear that an estimate of 0.63 mL/g for the average contribution is not unreasonable. What remains, then, is to compute the partial specific volume for a glycoprotein with unknown carbohydrate composition, and this is discussed further below.

The concentration gradients for both glycoproteins after attaining sedimentation equilibrium are shown in Figures 1 and 2 along with the fit to a single ideal species model. The residuals are also given in the figures, and aside from some obvious variations of absorbance (probably due to dirt or imperfections in the quartz windows of the cells), these data suggest that a single ideal species model is adequate in both analyses.

Table 2 summarizes the buoyant molecular weights and resulting absolute molecular weights determined from the estimated partial specific volumes.

Table 1. Estimated Partial Specific Volume for N-Linked Oligosaccharides

Carbohydrate*	Calculated Partial Specific Volume
High Mannose	
	0.63
Hybrid	
	0.64
Complex	
	0.62–0.64

*All three classes have identical core structure to the right of the dashed line. Variations may occur due to increased branching or the presence of other sugars. The values for partial specific volume are calculated on a weight basis using calculated partial specific volumes for individual sugar residues (7). The range computed for the complex form is based on a range of typical structures found in proteins

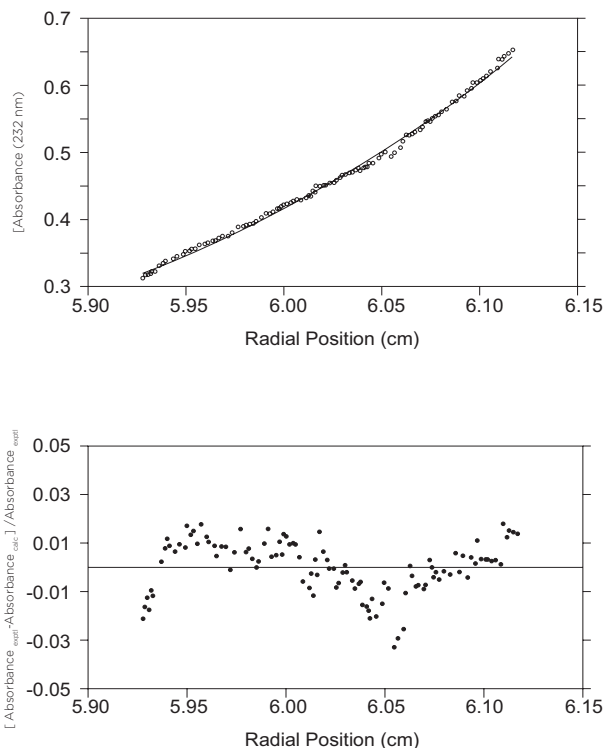


Figure 1. Absorbance gradient in centrifuge cell for sTNF-R1, after attaining sedimentation equilibrium. Solid line is the result of a fit to a single ideal sedimenting species. Bottom panel shows the difference in the fitted and experimental values as a function of radial position.

It is clear from these data that both glycoproteins are essentially monomeric in solution rather than dimers as suggested by the gel permeation chromatography. The determined molecular weight for rgp120 IIIIB is ≈ 104 kDa, which is in very good agreement with the expected molecular weight of 102 kDa based on amino acid and average carbohydrate composition. This agreement is quite good considering that the partial specific volume was also estimated from amino acid and average carbohydrate composition. In the case of sTNF-R1, the carbohydrate composition is unknown. Calculated molecular weights assuming 10, 20 and 30% carbohydrate with an average partial specific volume of 0.63 mL/g are shown in Table 2. As can be seen, the final values are relatively insensitive to the carbohydrate content and show that this protein is a monomer rather than the dimer suggested by the gel permeation chromatography experiments.

In conclusion, the sedimentation equilibrium technique can be used to obtain molecular weights of glycoproteins in solution. Estimation of partial specific volume using amino acid and carbohydrate composition may be sufficient for these analyses. Moreover, it may even be possible to estimate the partial specific volume for a glycoprotein with unknown carbohydrate composition as outlined in this report. The molecular weights obtained by gel permeation chromatography of the two glycoproteins discussed in this report are unreliable, but the sedimentation equilibrium technique, which does not rely on calibration with standard proteins, is an appropriate method to determine molecular weights of glycoproteins in solution.

Figure 2

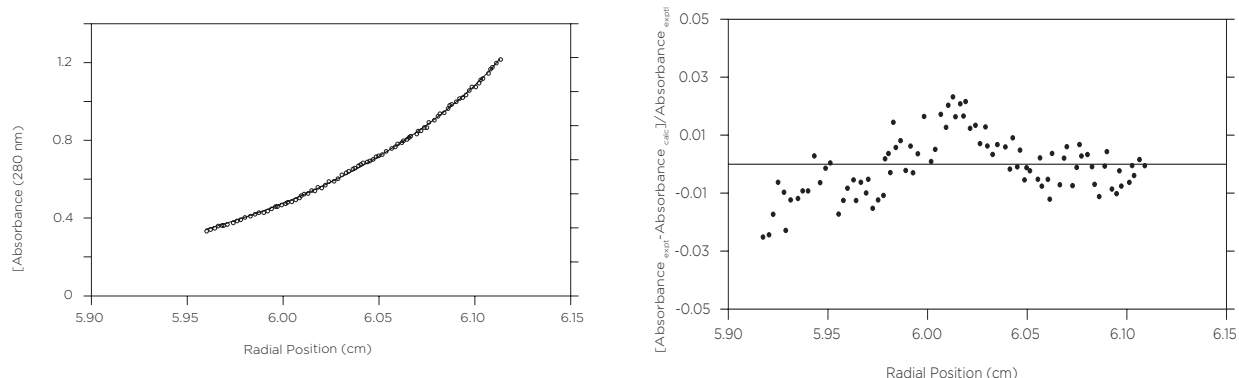


Figure 2. Absorbance gradient in centrifuge cell for rgp120 IIIIB, after attaining sedimentation equilibrium. Solid line is the result of a fit to a single ideal sedimenting species. Bottom panel shows the difference in the fitted and experimental values as a function of radial position.

Table 2. Molecular Weight of Two Glycoproteins Determined by Sedimentation Equilibrium

Glycoprotein	$M(1 - \bar{v}\rho)$ (kDa)	\bar{v} (mL/g)	M (kDa)
rgp120 IIIIB sTNF-R1	33 7.970 ± 0.870	0.68 0.68-0.70	104 26 ± 2.8
Effect of Assumed %Carbohydrate on Molecular Weight of sTNF-R1			
% carbohydrate of sTNF-R1	10	20	30
Estimated $\bar{v}_{\text{sTNF-R1}}$	0.70	0.69	0.68
$M_{\text{sTNF-R1}}$	26.9 ± 2.8	26 ± 2.7	25.2 ± 2.7

References

1. Leonard, C. K., Spellman, M., Riddle, L., Harris, R. J. Thomas, J. N., Gregory, T. J. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J. Biol. Chem.* 265, 10373-10382 (1990)
2. Pennica, D., Kohr, W. J., Fendly, B. M., Shire, S. J., Raab, H. E., Borchardt, P. E., Lewis, M., Goeddel, D.V. Characterization of a recombinant extracellular domain of the type 1 tumor necrosis factor receptor. Evidence for tumor necrosis factor- α induced receptor aggregation. *Biochemistry* 31, 1134-1141 (1992)
3. Shire, S. J., Holladay, L. A., Rinderknecht, E. Self association of human and porcine relaxin as assessed by analytical ultracentrifugation and circular dichroism. *Biochemistry* 30, 7703-7711 (1991)
4. Edelstein, S. J., Schachman, H. K. The simultaneous determination of partial specific volumes and molecular weights with microgram quantities. *J. Biol. Chem.* 242, 306-311 (1967)
5. Kratky, O., Leopold, H., Stabinger, H. The determination of the partial specific volume of proteins by the mechanical oscillator technique. *Methods in Enzymology*, Vol. 27, pp. 98-110. New York, Academic Press, 1973.
6. McMeekin, T. L., Marshall, K. Specific volumes of proteins and the relationship to their amino acid contents. *Science* 116, 142-143 (1952)
7. Charlwood, P. A. Partial specific volumes of proteins in relation to composition and environment. *J. Am. Chem. Soc.* 79, 776-781 (1957)
8. Gibbons, R. A. Physico-chemical methods for the determination of the purity, molecular size and shape of glycoproteins. *Glycoproteins, Part A*, pp. 128-140. Edited by A. Gottschalk. 2nd ed. Amsterdam, Elsevier Publ. Co., 1972.