CHROM. 4108

# ESTIMATION OF THE MOLECULAR WEIGHT OF FLEXIBLE DISORDERED PROTEINS BY EXCLUSION CHROMATOGRAPHY*

LI-PEN CHAO AND ELIZABETH ROBOZ EINSTEIN

*Departments of Neurology, and Biochemistry and Biophysics, University of California School of Medicine, San Francisco, Calif. 94122 (U.S.A.)*

(Received March 31st, 1969)

## SUMMARY

Estimation of the molecular weight of the encephalitogenic protein by exclusion chromatography, when globular proteins were used for comparison, resulted in a much higher value (39800) than that obtained by sedimentation equilibrium (17800 ± 600). This protein, unlike the reference compounds, is a flexible, disordered molecule. Therefore, globular proteins are not suitable for calibration.

In order to adapt the technique and render it suitable for general use, regardless of molecular shape, all proteins (references and unknowns used in the measurements) were denatured in concentrated guanidine hydrochloride and urea, reduced with $\beta$-mercaptoethanol to rupture the disulfide bonds, and then chromatographed on Sephadex G-100 columns. Under these conditions, the protein polypeptide chains are true random coils retaining no elements of their original native conformation. The molecular weight of the encephalitogenic protein estimated under these conditions is 19500. The method was also applied for the encephalitogenic active fragment and the lysine-rich histone. The results obtained are in agreement with the data obtained by physical methods.

## INTRODUCTION

For the estimation of the molecular weight of globular proteins, exclusion chromatography has been used frequently in recent years. However, anomalous results have been reported for proteins such as the glycoproteins[1], lysozyme[2], and hemoglobin[3]. Aggregation of some proteins under certain experimental conditions also affected the results[1]. When the technique was applied to the encephalitogenic protein from bovine spinal cord[4], to a basic protein from pig brain[5,6], and to histone from thymus gland[7,8], it resulted in considerably higher molecular weights than in the case of sedimentation equilibrium. It has been postulated that this is caused by the difference in molecular shape between the standard proteins and the flexible disordered proteins[5,8].

Because of the simplicity of exclusion chromatography, in which no special equipment is needed, unlike sedimentation equilibrium, it was desirable to find suitable conditions which whould eliminate the anomalous results so that the method could be used universally, regardless of the shape of the protein. According to TANFORD[9], protein polypeptide chains without cross-links are true random coils in concentrated guanidine hydrochloride (GuHCl) solution, and have dimensions that agree with the theoretical dimensions for randomly coiled chains. Therefore, estimation of the molecular weight in concentrated GuHCl in the presence of reducing and chelating agents should be free of anomalous results and independent of their molecular shape. The present study offers such a method for estimation of the molecular weights of the flexible, disordered proteins represented here by the encephalitogenic protein, its active fragment[10], and a histone. For the purpose of calibration, typical globular proteins were denatured, reduced, and then chromatographed on a Sephadex G-100 column.

EXPERIMENTAL PROCEDURE

*Materials*

Sephadex G-100 (Lot No. 6164) and Blue Dextran 2000 (a high molecular weight, $2 \times 10^6$, dextran covalently bound to a blue dye) were purchased from Pharmacia. Guanidine hydrochloride was obtained from Sigma and purified according to the method of SATAKE et al.[11] by passing the concentrated GuHCl solution through a column of activated charcoal (Norite) and Celite in a ratio of 1:2. This purified GuHCl solution (54 g/100 ml) had no or negligible absorption at 277 nm. Urea was purchased from Baker, and the concentrated urea solution was passed through a column of Rexyn 1-300 ($H^+$–$OH^-$ resin from Fisher) before use. All other solvents and reagents were reagent grade.

Bovine $\alpha$-chymotrypsinogen-A (6 × crystallized) and lysozyme (3 × crystallized) were obtained from Sigma. $\beta$-Lactoglobulin (B grade), insulin, and L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin (B grade) were purchased from Calbiochem. Pepsin (2 × crystallized) and horse heart myoglobin were obtained from Nutritional Biochem. Co. Bovine serum albumin and human $\gamma$-globulins (Fr-II) were purchased from Pentex. Ribonuclease and glucagon were obtained from Worthington and Eli Lilly & Co., respectively. Lysine-rich histone from calf thymus gland was kindly prepared and given to us by Dr. R. D. COLE of University of California at Berkeley.

*Preparation of the encephalitogenic protein and its active fragment*

The encephalitogenic protein and its active fragment were prepared according to the method of CHAO AND EINSTEIN[10].

*Preparation of the buffer systems*

Buffer system 1. The buffer system was prepared according to ANDREWS[3,12], and consisted of 0.05 $M$ Tris–HCl buffer (pH 7.5) containing 0.1 $M$ KCl.

Buffer system 2. The buffer system consisted of 0.05 $M$ Tris–HCl, containing 0.1 $M$ KCl, 5 $M$ GuHCl, 0.2 $M$ $\beta$-mercaptoethanol, and 0.1 $M$ ethylenediaminetetraacetate disodium (EDTA). The pH of this solution was adjusted to 7.5 with 6 $N$ HCl.

Buffer system 3. The composition of this buffer was identical with buffer system 2 except that the denaturing agent was 8 $M$ urea instead of GuHCl.

## Preparation of the columns

The Sephadex G-100 was well soaked in 0.02 $M$ HCl in the cold room. The solvent was removed by suction and the Sephadex gel was resuspended in the corresponding eluting buffer system overnight at room temperature with several changes of the suspending buffer system. The glass tube (1.0 × 185 cm, coated with Siliclad), with a plug of glass wool at the bottom, was filled up to approximately one third with the buffer. The Sephadex gel suspension was carefully poured to bring the liquid level to the top of the tube. After a 5- to 6-cm layer of the gel had settled to the bottom, the outlet was opened and the addition of the gel was continued until the bed height reached 180 cm. A buffer reservoir (1-l separatory funnel) was connected to the top of the column, and the flow of buffer was maintained by gravity. After the bed height was stabilized, a filter paper disc was placed on top of the gel to prevent disturbing the gel when the sample was applied. The column was equilibrated with the corresponding buffer system, usually for 48 h.

## Use of the columns

All experiments with buffer system 1 were performed in the cold room (4°). The work with buffer systems 2 and 3 was carried out at room temperature. The flow rate was maintained by the application of hydrostatic pressure to the top of the columns by adjusting the buffer reservoir. The approximate flow rate was 2.4 ml/h per tube for buffer system 1, and 2.0 ml/h per tube for buffer systems 2 and 3. Three milligrams of protein, dissolved in 1.0 ml of the equilibrating buffer, were added immediately to the top of the gel bed for buffer system 1. Under flow, the protein was washed into the gel with approximately 0.5 ml of additional buffer and the buffer was then added above the gel. Fractions were collected with an LKB fraction collector from the moment the samples touched the gel bed. The collector was equipped with a drop-counting device. Three milligrams of a protein sample were dissolved in 0.5 ml of either buffer system 2 or 3 and left for 24 h before use. (The sample solution was stored in the cold for buffer system 2 and at room temperature for buffer system 3.) Otherwise, the procedure was the same as for buffer system 1. The samples in most cases were chromatographed in succession. In a few cases, the second protein sample was applied to the column as soon as there was no danger of overlapping of the two components. No differences in the ratio between elution volume ($V_e$) and void volume ($V_0$) were found under these two conditions. The elution volume was taken from the initial addition of protein sample to the peak of the eluted protein. The void volume was determined with blue dextran before each run (unless it became constant).

## Examination of the effluents

Blue dextran in the column eluates was read at the optical density at 625 m$\mu$ on a Zeiss spectrophotometer. The protein content of the fractions from buffer system 1 was determined by optical density, either at 280 m$\mu$ or 230 m$\mu$. For determination of the protein content from buffer systems 2 and 3, the turbidimetric method of LUCK et al.[13] was used. Eluates (0.25 ml from each fraction) were diluted to a total volume of 2.0 ml with water in a Bausch and Lomb colorimetric tube, then 1 ml of 3 $M$ TCA

was added. After the solution had remained at room temperature for 15 min, the optical density at 400 m$\mu$ was read on a Bausch and Lomb Spectronic 20 colorimeter.

RESULTS

Table I contains the proteins, used for the standard to calibrate the Sephadex G-100 column, together with their molecular weights and intrinsic viscosities under various conditions. Table II shows the ratio of elution volume to void volume ($V_e/V_0$)

TABLE I

THE MOLECULAR WEIGHTS AND INTRINSIC VISCOSITIES OF THE STANDARD PROTEINS IN THEIR NATIVE AND DENATURED STATES

| Protein | Native state | | Denatured state | | | References |
|---|---|---|---|---|---|---|
| | Mol. wt. | $[\eta]$ (cc/g) | Mol. wt. | $[\eta]$ (cc/g) | | |
| | | | | In conc. GuHCl | In conc. urea | |
| Glucagon | 3 500 | — | — | — | — | 12 |
| Insulin | 5 700 | — | 2 970 | 6.1 | — | 9 |
| Ribonuclease | 13 700 | 3.3 | 13 700 | 16.6 | 15.6 | 9 |
| Lysozyme | 14 300 | 2.7 | 14 300 | 17.1 | 16.0 | 9 |
| Myoglobin | 17 800 | 3.1 | 17 200 | 20.9 | — | 9 |
| $\beta$-Lactoglobulin | 36 800 | 3.4 | 18 400 | 22.8 | 21.6 | 9 |
| Trypsin | 23 800 | — | 23 800 | — | — | 24 |
| Chymotrypsinogen | 25 700 | 2.5 | 25 700 | 26.8 | 22.6 | 9 |
| Pepsin | 35 000 | 3.1 | 35 000 | —[a] | — | 25 |
| Serum albumin | 69 000 | 3.7 | 69 000 | 52.2 | 43.2 | 9 |
| $\gamma$-Globulin | 150 000 | 6.0 | — | — | — | 12 |

[a] The intrinsic viscosity of pepsinogen is 31.5 (cc/g) in 6 $M$ GuHCl (ref. 9).

TABLE II

THE RATIO OF ELUTION VOLUME TO VOID VOLUME ($V_e/V_0$) OF PROTEINS IN THEIR NATIVE AND DENATURED STATES ON A SEPHADEX G-100 COLUMN

| Protein | $V_e/V_0$ | | |
|---|---|---|---|
| | Native state | Denatured state | |
| | | In conc. GuHCl | In conc. urea |
| Glucagon | 2.82 | — | — |
| Insulin | — | 2.61 | 2.62 |
| Ribonuclease | 2.23 | 1.71 | 1.72 |
| Lysozyme | 2.54 | 1.66 | 1.66 |
| Myoglobin | 2.14 | 1.59 | 1.60 |
| $\beta$-Lactoglobulin | 1.81 | 1.47 | 1.47 |
| Trypsin | — | 1.32 | 1.36 |
| Chymotrypsinogen | 1.95 | 1.30 | 1.31 |
| Pepsin | 1.47 | 1.14 | 1.15 |
| Serum albumin | 1.43 | 1.07 | — |
| $\gamma$-Globulin | 1.19 | — | — |
| Encephalitogenic protein | 1.71 | 1.46 | 1.46 |
| Active fragment | 1.84 | 1.74 | 1.73 |
| Lysine-rich histone | 1.23 | 1.23 | 1.22 |

under various experimental conditions. The linear relationship between $V_e/V_0$ and log (mol. wt.) was plotted (Fig. 1). The results show all the typical globular proteins when chromatographed in their native state under the conditions used by ANDREWS[1,12] (buffer system 1). The molecular weights of the encephalitogenic protein, its active fragment, and the lysine-rich histone were computed from this curve. Molecular weights of 39800, 30200, and 114800 resulted for the encephalitogenic protein, its active fragment, and the lysine-rich histone, respectively. Similar results with this technique were also reported for the encephalitogenic protein[4], a basic protein from pig brain[5,6], and histone[7,8]. It was obvious that these results compared with those obtained from sedimentation equilibrium were too high (Table III). It has been postulated that this discrepancy arises from the difference in molecular shape between the calibration standards and the flexible, disordered proteins[5,8].

All proteins were then treated and chromatographed in 5 $M$ GuHCl, $\beta$-mercaptoethanol, and EDTA solution (buffer system 2). In order to check the validity of this method, the same technique was applied with 8 $M$ urea in the eluting buffer (buffer system 3). Fig. 2 shows the plot of $V_e/V_0$ versus log (mol. wt.). The results in Fig. 2 and Table II show an almost indistinguishable difference between the buffer systems containing concentrated GuHCl and urea. The molecular weights of the encephalitogenic protein, its active fragment, and the lysine-rich histone, computed on the basis of this curve, were approximately 19000, 12000, and 29000, in this order (Table III).
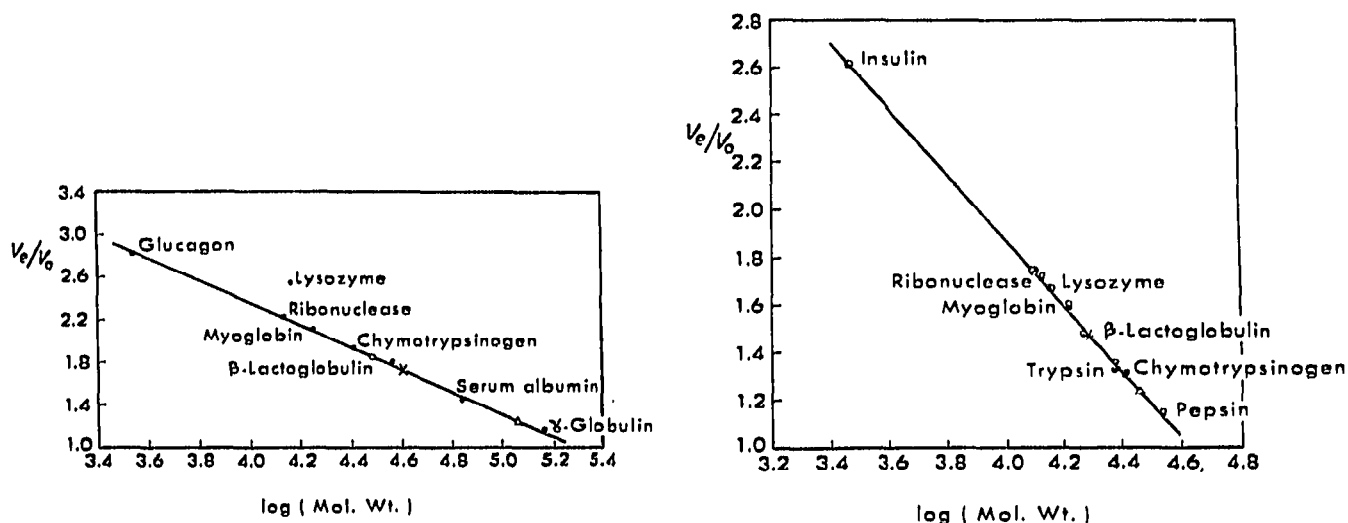


Fig. 1. Plot of $V_e/V_0$ versus log (mol. wt.) for proteins in their native states on a Sephadex G-100 column (1.0 × 180 cm) at pH 7.5. Experimental details are described in the text. The symbols ○, × and △ represent the active fragment, the encephalitogenic protein, and the lysine-rich histone, respectively.

Fig. 2. Plot of $V_e/V_0$ versus log (mol. wt.) for proteins in their denatured and reduced states on a Sephadex G-100 column (1.0 × 180 cm). Experimental details are described in the text. Values for $V_e/V_0$ were determined in guanidine hydrochloride (●) and in urea (○). The symbols ⊗, × and △ represent the active fragment, the encephalitogenic protein, and the lysine-rich histone, respectively.

TABLE III

THE MOLECULAR WEIGHTS OF THE ENCEPHALITOGENIC PROTEIN, THE ACTIVE FRAGMENT, AND HIS-
TONE OBTAINED BY EXCLUSION CHROMATOGRAPHY AND SEDIMENTATION EQUILIBRIUM

| Methods | Encephalitogenic protein | Active fragment | Histone |
|---|---|---|---|
| Exclusion chromatography | | | |
| in Tris–HCl buffer | 39 800 | 30 200 | 114 800 |
| in GuHCl | 19 500 | 12 300 | 28 800 |
| in urea | 19 500 | 12 500 | 29 200 |
| Sedimentation equilibrium | 17 800 ± 600 (ref. 26) | — | 22 100 ± 1 400 (ref. 27) |

DISCUSSION

The present study clearly demonstrates that the methods of WHITAKER[2] and
ANDREWS[3,12] are only suitable for the estimation of the molecular weights of certain
globular proteins. Anomalous results have been reported with this technique where
the molecular shape is globular, such as lysozyme[2], hemoglobin[3], and $\beta$-lactoglobulin[3].
Aggregation of some proteins can also occur under certain experimental conditions[1].
Results presented in Fig. 1 and Table II show that lysozyme was eluted more slowly
than expected from its molecular weight. Whitaker[2] reported the same observation.
Pepsin was eluted much faster than expected from its molecular weight of 35 000
(Table I). It gave a molecular weight of approximately 70 000. No such results can be
observed using buffer systems 2 and 3 (Table II and Fig. 2). It should be added that
GuHCl is a good solvent for most of the proteins[14].

The relationship between $V_e/V_0$ and log (mol. wt.), shown in Fig. 2, is linear only
in the molecular weight range of 3000 to 35 000. From Table II, we see that the de-
natured globular proteins are eluted much faster than in their native states. For the
higher molecular weight proteins (>35 000), agarose with low cross-linkage will be
the choice[15,16].

JIRGENSONS and co-workers[17,18] reported that the lysine-rich histone is a flexible,
disordered protein in its native state. Recently, we demonstrated by various physico-
chemical methods[19] that the encephalitogenic protein from bovine spinal cord is
similar in molecular shape to histone. The active fragment is derived from the en-
cephalitogenic protein and should have the same conformation as the parent material.
The intrinsic viscosities of these proteins in their native and denatured states are
presented in Table IV. The data indicated that there are no significant differences in
in the intrinsic viscosities in both states. These proteins are not globular (they have
low intrinsic viscosities[20], between 3 and 4 cc/g; see Table I). Consequently, the globular
proteins are not suitable for calibration. Table II definitely shows that the elution
behavior of the disordered protein is quite different from the globular one on the Se-
phadex G-100 column. The latter proteins were eluted much faster in the denatured
state than in their native conformation. In contrast there is not much difference be-
tween the flexible, disordered proteins before and after treatment.

TANFORD[9] and his co-workers concluded, on the basis of the measurements of the

TABLE IV

THE INTRINSIC VISCOSITIES OF THE ENCEPHALITOGENIC PROTEIN AND HISTONE IN THEIR NATIVE AND DENATURED STATES[a]

| Proteins | $[\eta]$ (cc/g) | | References |
|---|---|---|---|
| | Native state | Denatured state | |
| Encephalitogenic protein | 16.9 | 20.8 | 26 |
| Lysine-rich histone | 20.5 | 22.0 | 9, 17 |

[a] The intrinsic viscosity of the active fragment was not measured. The molecular shape should be the same as that of the encephalitogenic protein, the parent compound.

intrinsic viscosities[21], hydrogen ion titration[22], and optical rotatory dispersion[23] of proteins in 6 $M$ GuHCl, that protein polypeptide chains in concentrated (5–8 $M$) GuHCl are true random coils retaining no elements of their original native conformation. Proteins which have been denatured by urea probably exist as random coils, similar to proteins denatured by GuHCl. For this reason, the globular proteins and the disordered proteins were denatured by both 5 $M$ GuHCl and 8 $M$ urea in the presence of excess $\beta$-mercaptoethanol to rupture the disulfide bonds and in EDTA to remove the metal from any metalloprotein to prevent folding of the polypeptide chains. Under these conditions, the chromatographic behavior of either the globular or the flexible, disordered proteins on the Sephadex G-100 column should be independent of their shape, since now they both exist as random coils. The molecular weights measured under these conditions are summarized in Table III. According to AN-DREWS[3,12] the values obtained by exclusion chromatography are accurate within 10 %. The values are in agreement with the results obtained from the sedimentation equilibrium method. From the amino acid analyses, the approximate minimum molecular weight for the encephalitogenic protein is about 10 000 and 6 000 for its active fragment[10]. The results of the present study are consistent with these findings when doubling the approximate minimum molecular weights.

REFERENCES

1 P. ANDREWS, in H. PEETERS (Editor), Protides of the Biological Fluids, 14th Colloquium, Elsevier, Amsterdam, 1966, p. 573.
2 J. R. WHITAKER, Anal. Chem., 35 (1963) 1950.
3 P. ANDREWS, Biochem. J., 91 (1964) 222.
4 A. NAKAO, W. J. DAVIS AND E. R. EINSTEIN, Biochim. Biophys. Acta, 130 (1966) 163.
5 L. G. TOMASI AND S. E. KORNGUTH, J. Biol. Chem., 242 (1967) 4933.
6 L. G. TOMASI AND S. E. KORNGUTH, J. Biol. Chem., 243 (1968) 2507.
7 H. J. CRUFT, Biochim. Biophys. Acta, 54 (1961) 611.
8 K. MURRAY, Ann. Rev. Biochem., 34 (1965) 209.
9 C. TANFORD, Advan. Protein Chem., 23 (1968) 121.
10 L.-P. CHAO AND E. R. EINSTEIN, J. Biol. Chem., 243 (1968) 6050.
11 K. SATAKE, P. S. RASMUSSEN AND J. M. LUCK, J. Biol. Chem., 235 (1960) 2801.
12 P. ANDREWS, Biochem. J., 96 (1965) 595.
13 J. M. LUCK, P. S. RASMUSSEN, K. SATAKE AND A. N. TSVETIKOV, J. Biol. Chem., 233 (1958) 1407.
14 C. TANFORD, K. KAWAHARA AND S. LAPANJE, J. Biol. Chem., 241 (1966) 1921.
15 P. F. DAVISON, Science, 161 (1968) 906.
16 K. A. PIEZ, Anal. Biochem., 26 (1968) 305.
17 B. JIRGENSONS, L. S. HNILICA AND S. C. CAPETILLO, Makromol. Chem., 97 (1966) 216.

18 B. JIRGENSONS, *J. Biol. Chem.*, 242 (1967) 912.
19 L.-P. CHAO AND E. R. EINSTEIN, *Federation Proc.*, 28 (1969) 842.
20 J. T. YANG, *Advan. Protein Chem.*, 16 (1961) 323.
21 C. TANFORD, K. KAWAHARA AND S. LAPANJE, *J. Am. Chem. Soc.*, 89 (1967) 729.
22 Y. NOZAKI AND C. TANFORD, *J. Am. Chem. Soc.*, 89 (1967) 742.
23 C. TANFORD, K. KAWAHARA, S. LAPANJE, T. M. HOOKER, Jr., M. H. ZARLENGO, A. SALAHUD-
DIN, K. AUNE AND T. TAKAGI, *J. Am. Chem. Soc.*, 89 (1967) 5023.
24 P. DENUELLE, in P. D. BOYER, H. LARDY AND K. MYRBÄCK (Editors), *The Enzymes*, Vol. 4,
Academic Press, New York, 1960, p. 119.
25 F. A. BOVEY AND S. S. YANARI, in P. D. BOYER, H. LARDY AND K. MYRBÄCK (Editors), *The
Enzymes*, Vol. 4, Academic Press, New York, 1960, p. 63.
26 L.-P. CHAO AND E. R. EINSTEIN, unpublished results.
27 D. C. TELLER, J. M. KINKADE AND R. D. COLE, *Biochem. Biophys. Res. Commun.*, 20 (1965) 739