SDS–PAGE strongly overestimates the molecular masses of the neurofilament proteins

Eckhard Kaufmann, Norbert Geisler and Klaus Weber

Max Planck Institute for Biophysical Chemistry, D-3400 Göttingen, FRG

Direct molecular mass determination of the three porcine neurofilament proteins (H, M and L) was performed in 6 M guanidine–HCl using analytical gel filtration and sedimentation equilibrium centrifugation. The results show that SDS–PAGE strongly overestimates the values of the 'higher molecular mass' components H and M. This discrepancy stems from the carboxyterminal extensions known to have unusual amino acid composition.

Neurofilament   Molecular mass   Gel electrophoresis   Dodecyl sulfate   Guanidine–HCl
Intermediate filament

1. INTRODUCTION

Mammalian neurofilaments consist of 3 distinct polypeptides whose molecular masses have only been analyzed by SDS–PAGE (see, for instance, [1–5]). In continuation of previous suggestions these triplet proteins can be called L(ow), M(iddle) and H(igh) and have for porcine neurofilaments been given values of 68–72 kDa (L), 140–160 kDa (M) and 200–220 kDa (H). Much emphasis has been put on the accuracy of these determinations (see, for instance, [1]) and species-specific changes of these apparent molecular masses have been reported particularly for M and H [4,5]. Whereas L was recognized early on as an intermediate filament protein by partial amino acid sequence data [3,6] and in vitro polymerization properties [7,8], the possible interrelation of H, M and L remained obscure until the properties of subdomains isolatable from mildly proteolyzed material were studied [1–3]. Our current view of H and M is that of polyproteins which harbor in their aminoterminal part the sequence information of non-epithelial intermediate filament proteins and display additional domains in the carboxyterminal part [3]. The latter domains increase in length upon transition from L to M or H. Since non-epithelial

intermediate filament proteins present outside neurones have molecular masses between 51 and 54 kDa the unusually high values of H and M have posed a special problem, which we have analyzed here by direct physical-chemical methods.

2. MATERIALS AND METHODS

Neurofilament triplet proteins were from porcine spinal cord [3,7]. The carboxyterminal tailpiece domains of H and M were obtained by limited chymotryptic digestion and purified by chromatography on DEAE–cellulose in the presence of urea (for documentation see [3,7]). Proteins concentrated by ethanol precipitation (70%) were dissolved in the guanidine–HCl buffers and extensively dialyzed against these solvents. Gel filtration was on Sephacryl S-400 and S-300 (Pharmacia). Protein elution was monitored by trichloroacetic acid precipitation or SDS–PAGE. Alternatively proteins were labelled either by ^{125}I or by the fluorescent reagent fluoresceine isothiocyanate [9]. Standards were from Boehringer with the exception of RNA-polymerase from *Escherichia coli*, which was kindly provided by Dr C. Biebricher. Sedimentation equilibrium centrifugation at 0.6–1.5 mg/ml protein was performed in a

Spinco Beckman centrifuge using the meniscus depletion method [10].

3. RESULTS AND DISCUSSION

Fig.1 shows the standardized elution profiles for Sephacryl S-400 and S-300 when 6 M guanidine–HCl was used at room temperature in 0.02 M Tris–HCl buffer (pH 7) in the presence of 0.1 M

2-mercaptoethanol. The particular molecular mass plot chosen was introduced in [9]. Elution profiles were highly reproducible. Apparent molecular masses deduced are 60 ± 4 kDa for L (two experiments), 106 ± 5 kDa for M (3 experiments) and 138 ± 5 kDa for H (5 experiments). In addition we characterized the carboxyterminal tailpiece of H and M with 107 and 63 kDa, respectively. The combined results are summarized in table 1.

Direct molecular mass determinations using sedimentation equilibrium centrifugation under meniscus depletion conditions were performed at 20°C in 6 M guanidine–HCl (20 mM Tris–HCl, pH 7.8, 1 mM dithiothreitol, 1 mM phenylmethylsulfonylfluoride and 1 mM NaN$_3$). Fig.2 gives examples for the molecular mass of the triplet proteins from 3 experiments performed at different protein concentrations. Apparent partial specific volumes were calculated from the known amino acid composition [3] correcting for guanidine–HCl binding as in [11]. The resulting values in ml/g
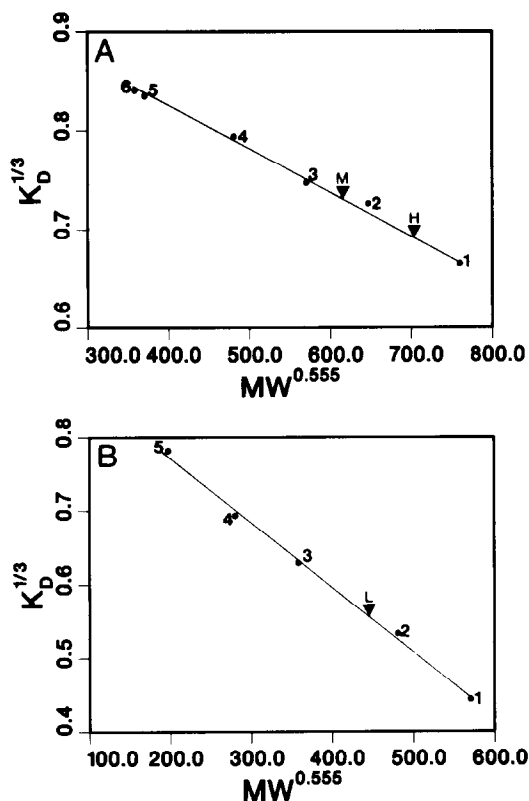
Fig.1. Gel filtration on Sephacryl S-400 (A) and Sephacryl S-300 (B) in 6 M guanidine–HCl, 0.02 M Tris–HCl (pH 7) and 0.1 M 2-mercaptoethanol. The plot shows the cubic root of the distribution coefficient K_d as a function of the molecular mass (MW) to the 0.555 power as in [9]. Triplet proteins are H and M in (A) and L in (B). Standards in (A) are $\beta\beta'$-chains of *Escherichia coli* RNA polymerase (1; 155 kDa), β-galactosidase from *E. coli* (2; 116 kDa), phosphorylase b (3; 93 kDa), bovine serum albumin (4; 68 kDa), actin (5; 43 kDa) and aldolase (6; 40 kDa). Standards 1–3 of (B) are 3, 4 and 6 in (A). Additional standards in (B) are chymotrypsinogen (4; 26 kDa) and ribonuclease (5; 14 kDa).

Table 1

Molecular masses in kDa of the three mammalian neurofilament proteins and their carboxyterminal domain

Protein	SDS–PAGE[a]	Gel filtration	Sedimentation	Proposed value
H	200–205	138	112	~110–140
H tailpiece	150–170	107	73	
M	145–160	106	108	~107
M tailpiece	100–120	63		
L	68–72	60	63	62

[a] Values are from [1–8]

were 0.727, 0.746 and 0.749 for L, M and H, respectively. Using these values apparent molecular masses were 63 kDa for L (3 determinations

62; 63.5; 65.5), 108 kDa for M (4 determinations 102; 107; 112; 112) and 112 kDa for H (4 determinations 103; 110; 110; 113). The carboxyterminal domain of H yielded a value of 73 kDa in 6 M guanidine–HCl and 63 kDa in 0.05 M Mes buffer (pH 6.5), 0.1 M NaCl, 1 mM dithiothreitol, in the absence of the denaturant.

Using gel filtration and ultracentrifugation in 6 M guanidine–HCl we have shown that SDS–PAGE overestimates the molecular masses of the 3 neurofilament triplet proteins (table 1). This is already noticeable with the L polypeptide but becomes extreme for the two high molecular mass components M and H. This discrepancy is located to the carboxyterminally situated non-α-helical domains (tailpieces) known to have an unusual amino acid composition [3] and in the case of H an extremely high content of serine phosphate [2,12]. In addition these domains are rather acidic [2,3]. There are several acidic proteins which, probably due to restricted SDS-binding, give too high apparent molecular masses in SDS–PAGE (see, e.g., [13]), although the discrepancies seem not as extreme as observed for H and M. That the higher degree of phosphorylation in the tailpiece of H contributes to an overestimate of the molecular mass can also be seen by inspections of fig.4 in [2]. Enzymatic removal of at least some phosphate leads to an increased mobility in SDS–PAGE.

For L we estimate a polypeptide molecular mass around 62 kDa. This value is in line with current sequence data covering the aminoterminal 82 and carboxyterminal 271 residues [3]. If along the remaining sequence gap L is built like other non-epithelial intermediate filament proteins [3,14] the total length of L should be around 545 residues corresponding to a molecular mass of 62 kDa. This value is in excellent agreement with our
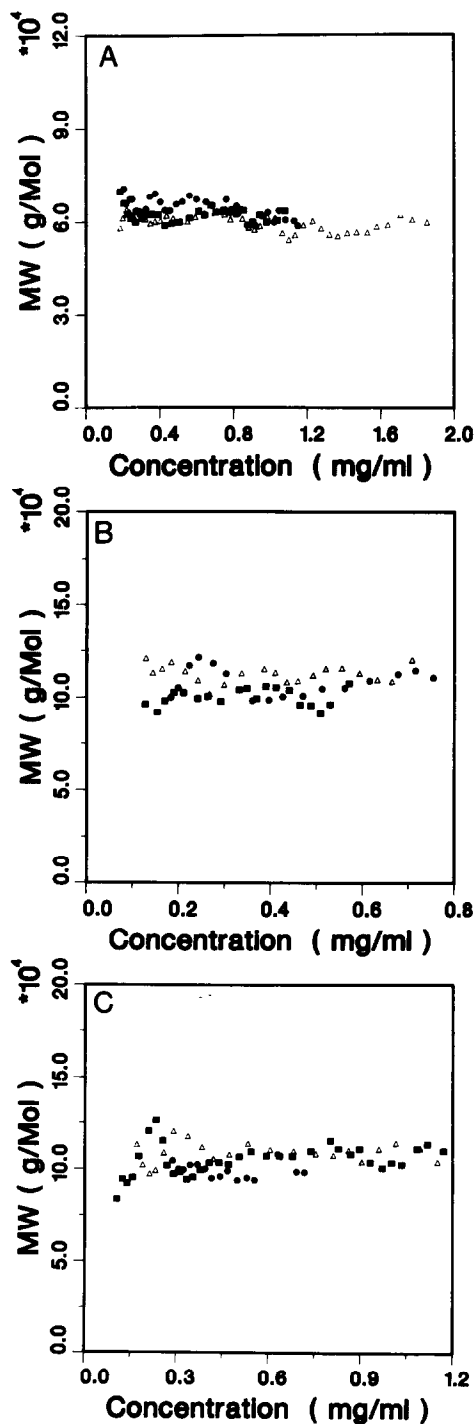

Fig.2. High-speed equilibrium centrifugation of the triplet proteins in 6 M guanidine–HCl, 0.02 M Tris–HCl (pH 7.8), 1 mM phenylmethylsulfonylfluoride, 1 mM dithiothreitol using different original protein concentrations [15]. L polypeptide at 0.9 ($\triangle$), 1.2 ($\bullet$) and 1.5 mg/ml ($\blacksquare$) see (A); M polypeptide at 0.6 ($\triangle$), 0.9 ($\bullet$) and 1.2 mg/ml ($\blacksquare$) see (B); H polypeptide at 0.9 ($\blacktriangle$), 1.3 ($\bullet$) and 0.6 mg/ml ($\blacksquare$) see (C). The column height was about 4 mm and all experiments were performed at 20°C.

measurements of 60–63 kDa but significantly lower than the 68–72 kDa numbers consistently derived from SDS–PAGE [1–5].

For the M component there is very good agreement between the results obtained by gel filtration and sedimentation equilibrium centrifugation (table 1). Both methods point to a molecular mass of about 107 kDa, a value some 30–40% lower than the ones derived by SDS–PAGE. M protein is thought to resemble a normal non-epithelial intermediate filament protein with a further carboxyterminal tail region [3]. Since such filament proteins usually have molecular masses of 50–54 kDa [15] this hypothetical value plus the newly determined molecular mass of the tail region account for 113–118 kDa in fair agreement with the 107 kDa found experimentally for M. Thus further domains are not to be expected and emerging partial amino acid sequence data centering on CNBr fragments (unpublished) also favor this prediction.

The molecular mass of H cannot be predicted as clearly. Although both gel filtration and centrifugation point to values some 30–40% lower than SDS–PAGE (table 1) the two procedures involving 6 M guanidine–HCl reveal some disagreement. Gel filtration indicates a value around 138 kDa whereas sedimentation equilibrium centrifugation points to only 112 kDa. Currently we must favor the latter value since gel filtration in guanidine–HCl requires the presence of random coils for both standard and unknown proteins. Given the unusual amino acid composition of the carboxyterminal domain of H [3] and its extreme phosphate content [2,12] we do not know if the random coil requirement is really fulfilled. The latter two properties may, however, also influence the value for the partial specific volume entering the equation to calculate the molecular mass by sedimentation equilibrium centrifugation. A value of 0.749 ml/g was calculated from the known amino acid composition with correction for guanidine–HCl binding but deviations from such calculations can occur [11]. Thus the H component has to be characterized by either a direct determination of its partial specific volume or by detailed amino acid sequence data. Regardless of this remaining difficulty all data on H point to a much lower molecular mass than indicated by SDS–PAGE and pin the source of the disagreement to the carboxyterminal domain known to be very rich in glutamic acid and lysine [13] and harbor many serine phosphates [2,12].

Several studies have noticed changes in the apparent molecular masses particularly of H and M when neurofilaments of different vertebrates were analyzed by SDS–PAGE [4,5]. Given the results described above one wonders whether such changes indeed reflect any molecular mass properties or whether they arise from certain amino acid sequence changes of protein domains not reliably analyzable by SDS–PAGE.

REFERENCES

[1] Chin, T.K., Eagles, P.A.M. and Maggs, A. (1983) Biochem. J. 215, 227–237.

[2] Julien, J.P. and Mushynski, W.E. (1983) J. Biol. Chem. 258, 4019–4025.

[3] Geisler, N., Kaufmann, E., Fischer, S., Plessmann, U. and Weber, K. (1983) EMBO J. 2, 1295–1302.

[4] Liem, R.H.S., Yen, S.H., Solomon, G.D. and Shelanski, M.L. (1978) J. Cell Biol. 79, 637–645.

[5] Shaw, G. and Weber, K. (1984) Eur. J. Cell Biol., in press.

[6] Geisler, N., Plessmann, U. and Weber, K. (1982) Nature 296, 448–450.

[7] Geisler, N. and Weber, K. (1981) J. Mol. Biol. 151, 565–571.

[8] Liem, R.K.H. and Hutchinson, S.B. (1982) Biochemistry 21, 3221–3226.

[9] Fish, W.W., Mann, K.G. and Tanford, C. (1969) J. Biol. Chem. 244, 4989–4994.

[10] Yphantis, D.A. (1964) Biochemistry 3, 297–317.

[11] Lee, J.C. and Timasheff, S.N. (1979) Methods Enzymol. 61, 49–57.

[12] Julien, J.P. and Mushynski, W.E. (1982) J. Biol. Chem. 257, 10467–10470.

[13] Burton, Z., Burgess, R.R., Lin, J., Moore, D., Holder, S. and Gros, C.A. (1981) Nucleic Acids Res. 9, 2889–2903.

[14] Geisler, N. and Weber, K. (1983) EMBO J. 2, 2059–2063.

[15] Kim, H., Deonier, R.C. and Williams, J.C. (1977) Chem. Rev. 77, 659–690.