SIZE AND CHEMICAL COMPOSITION OF INFLUENZA VIRUS HEMAGGLUTININ CHAINS

Colin W. WARD and Theo A. A. DOPHEIDE
Division of Protein Chemistry, CSIRO, 343 Royal Parade, Parkville (Melbourne), Victoria 3052, Australia

Received 26 January 1976
Revised version received 23 April 1976

1. Introduction

Influenza viruses are filamentous or spherical particles composed of a segmented viral genome, which together with several types of protein is enclosed within a lipoprotein envelope from which two types of glycoprotein project outwards [1]. Although three virus types (A, B and C) can infect man [2], successive epidemics are only caused by the A viruses. New strains arise as a result of major (antigenic shift) or minor (antigenic drift) changes in the structure of the major spike protein, the hemagglutinin [3–5].

This letter reports the molecular weight and the amino acid and carbohydrate compositions of the two chains of the hemagglutinin from a Hong Kong virus variant A_2/Memphis/102/72. The results obtained confirm that the heavy (HA_1) chain contains most of the carbohydrate, and show that the true mol. wt. of the heavy chain is considerably less than that estimated from standard SDS gels. We further show, that the apoprotein portion of the heavy chain is only 29% larger than the apoprotein part of the light chain (HA_2).

2. Materials and methods

The virus used was a recombinant, possessing the hemagglutinin of A_2/Memphis/102/72 and the neuraminidase of A_0/Bel/42, prepared as described by Webster [6]. The inoculum was obtained from Dr W. G. Laver, Department of Microbiology, J.C.S.M.R., Australian National University. Virus was grown and purified as described by Laver [7].

The hemagglutinin was purified by electrophoresis in cellulose acetate [8] and the reduced chains separated on guanidine hydrochloride gradients [9].

Amino acid analysis was performed on 3 nmol samples, hydrolysed in 4 M methylsulphonic acid (0.4 ml) containing 0.2% tryptamine [10] at 115°C, in vacuo, for 13, 24 and 48 h (table 1).

For carbohydrate analysis, samples (50–80 nmol) were hydrolysed in vacuo at 100°C for 5 h in 0.5 M HCl. Inositol was used as internal standard. The estimation was carried out essentially as described by Sloneker [11].

Polyacrylamide gel electrophoresis in SDS gels [12] was carried out on 100 μg samples dissolved by heating to 100°C for 2 min in 50 μl of 0.05 M sodium phosphate–8.0 M urea buffer, pH 7.0, containing 5% (w/v) SDS and 4% (v/v) mercaptoethanol. Mol. wt. markers used were bovine serum albumin monomer and dimer, ovalbumin, carbonic anhydrase and sperm whale myoglobin.

The molecular weight of HA_1 was also determined by meniscus depletion high speed equilibrium ultracentrifugation [13] in 6.0 M guanidine hydrochloride–0.001 M dithiothreitol, at 20°C, using a calculated partial specific volume.

3. Results and discussion

3.1. *Molecular weight*

Reported mol. wts. of influenza hemagglutinins have generally been based upon relative mobilities in SDS gels (cf. [4] for review). When compared with 'normal' proteins, however, glycoproteins bind less SDS per gram of protein, resulting in lower electro-

Table 1
Amino acid and carbohydrate compositions of influenza (strain A_2/Memphis/102/72) hemagglutinin heavy (HA_1) and light (HA_2) chains, expressed as residues per mole protein (A) and residues per 100 residues recovered (B)

Residue	HA_1		HA_2	
	A	B	A	B
Lys	15.1 (15)	4.56	15.6 (16)	6.38
His	5.7 (6)	1.72	5.6 (6)	2.27
Arg	13.9 (14)	4.22	11.3 (11)	4.62
Asp	43.7 (44)	13.24	32.0 (32)	13.06
Thr[a]	28.6 (29)	8.67	12.5 (13)	5.10
Ser[a]	30.4 (30)	9.22	13.1 (13)	5.36
Glu	25.4 (25)	7.68	33.1 (33)	13.50
Pro[b]	20.5 (21)	6.22	3.2 (3)	1.31
Gly	29.0 (29)	8.80	22.6 (23)	9.21
Ala	14.0 (14)	4.23	14.9 (15)	6.10
Val[b]	20.9 (21)	6.34	10.3 (10)	4.21
Met	4.0 (4)	1.23	3.9 (4)	1.58
Ile[b]	21.8 (22)	6.61	21.1 (21)	8.64
Leu[b]	22.3 (22)	6.75	18.8 (19)	7.66
Tyr	11.0 (11)	3.32	7.9 (8)	3.24
Phe	10.8 (11)	3.26	11.0 (11)	4.49
Trp[a]	4.8 (5)	1.46	3.7 (4)	1.49
Cys[c]	8.1 (8)	2.46	7.5 (8)	3.04
Met[c]	4.0 (4)	1.23	3.9 (4)	1.58
Glucosamine[ad]	19.1 (19)		6.6 (7)	
Galactose	9.9 (10		0.4	
Mannose	33.0 (33)		0.6	
Fucose	4.9 (5)		0.3	

[a] Corrected by extrapolation to zero time.
[b] Values after 48 h hydrolysis.
[c] Determined after duplicate 24 h hydrolysis of performic acid oxidized proteins.
[d] Determined by amino acid analyzer.

phoretic mobilities and hence higher apparent mol. wts. [14,15]. Since the degree of overestimation decreases as the acrylamide gel concentration increases, the true molecular weight of glycoproteins can be obtained from a plot of their apparent mol. wt. at a range of gel concentrations against gel concentration [16]. By this method, we obtained the results shown in fig.1. The apparent mol. wt. of the light chain fell only from a value of 34 000 in 3% gels through 31 000 in 5% gels to a constant 29 700 in 7.5–15% gels. In contrast, the apparent molecular weight of the heavy chain fell from 91 800 in 3% gels to 51 000 in 15% gels, suggesting an asymptotic minimum mol. wt. of approx. 50 000. Since the estimation of these minimal mol. wts. appears to be analogous to the determination of maximum velocities in enzyme kinetics, a double reciprocal plot should give a straight line whose intercept should be the true mol. wt.

As shown in the inset to fig.1 the points do fall on a straight line and the minimal mol. wt. as determined by the intercept is 46 000. This value agrees well with that of 46 800 obtained by ultracentrifugal analysis in 6.0 M guanidine hydrochloride. The partial specific volume for HA_1 was 0.689 cc/g, calculated by the method of Lee and Timasheff [17] using the data of Cohn and Edsall [18] for amino acids and that of Gibbons [19] for sugar residues.

The considerable variation in the mol. wts. reported in the literature for the HA_1 from other strains of influenza (see [4] for review) may be due to differ-

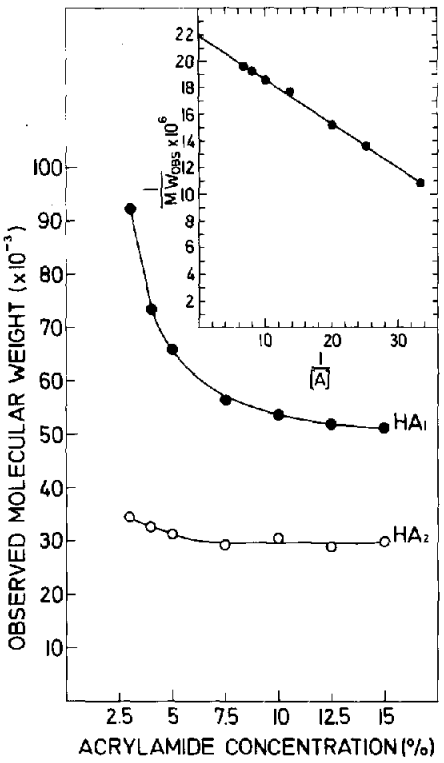

Fig.1. Observed mol. wts. of hemagglutinin chains in SDS gels of different acrylamide concentrations. (•) Heavy chain, HA_1; (○) light chain HA_2. Insert. Double reciprocal plot of (HA_1 observed mol. wt.)$^{-1}$ versus (acrylamide concentration)$^{-1}$.

ences in the acrylamide gel concentrations used by the different workers. For example, the mol. wt. of the HA_1 from A_0/Bel/42 has been reported as 65 000 when 5% gels were used [20], 58 000 when 7.5% gels were used [21] but only 47 000 when the mol. wt. was determined by gel filtration in guanidine hydrochloride [22]. This compares with our values of 66 000 in 5% gels, 56 000 in 7.5% gels and 46 800 from ultracentrifugal analysis.

3.2. *Carbohydrate and amino acid compositions*

The carbohydrate and amino acid compositions for both HA_1 and HA_2 are shown in table 1. Both chains are glycosylated (see also [21,23]) but most of the carbohydrate is on the heavy chain. For a mol. wt. of 47 000, the carbohydrate content of HA_1 totals 11 450 daltons or 24.4% by weight and puts the estimate of the apoprotein mol. wt. at 35 500. Only four sugars were present, *N*-acetylglucosamine, galactose, mannose and fucose with molar ratios 3.8/2/6.6/1 respectively. The presence of *N*-acetylglucosamine and the absence of *N*-acetylgalactosamine and xylose suggests the carbohydrate is attached via *N*-glycosidic linkage to asparagine [24]. We find that the light chain carries about one-third as much *N*-acetylglucosamine as the heavy chain [9,20] but only traces of the other sugars. For a mol. wt. of 29 700 the *N*-acetylglucosamine content of HA_2 totals 1400 daltons, or 4.7% by weight, giving an apoprotein mol. wt. of 28 300.

The amino acid analyses for the two chains are also shown in table 1. Both chains contain four methionines and eight half-cystines per mole of protein. When compared with the amino acid compositions for A_0/Bel/42 [25,26] there are several quite significant differences. Peptide maps of tryptic digests indicate that there are also marked differences in amino acid sequence in the hemagglutinin chains of A_0/Bel/47 and the A_2 viruses [9,27].

Acknowledgements

We wish to thank Dr A. E. Clarke, School of Botany, University of Melbourne, for help with the carbohydrate analyses, Dr E. F. Woods of this laboratory for the ultracentrifugal analyses and Dr W. G. Laver for advice regarding cultivation and purification of virus.

References

[1] Fenner, F., McAuslan, B. R., Mims, C. A., Sambrook, J. and White, D. O. (1974) The Biology of Animal Viruses, 2nd edn. pp. 111–117, Academic Press, New York.
[2] Pereira, H. G. (1969) Progr. Med. Virol. 11, 46–79.
[3] Laver, W. G. and Webster, R. G. (1966) Virology 48, 445–455.
[4] White, D. O. (1964) Curr. Top. Microbiol. Immunol. 63, 1–48.
[5] Skehel, J. J. (1974) Symp. Soc. Gen. Microbiol. No. 24, 321–343.
[6] Webster, R. G. (1970) Virology 42, 633–642.
[7] Laver, W. G. (1969) in: Fundamental Techniques in Virology (Haber, K. and Salzmann, N. P., eds.) pp. 66–82, Academic Press, New York.
[8] Laver, W. G. (1964) J. Mol. Biol. 9, 109–124.

[9] Laver, W. G. (1971) Virology 45, 275–288.
[10] Liu, T.-Y., as quoted to S. Moore, (1972) Chemistry and Biology of Peptides, Proc. 3rd American Peptide Symposium 1972, (Meyenhofer, J., ed.), pp. 629–653, Science, Ann Arbor.
[11] Sloneker, J. H. (1973) in: Methods of Carbohydrate Chemistry IV (Whistler, R. and BeMiller, J. N., eds.), pp. 20–24, Academic Press, New York.
[12] Weber, K. and Osborn, M. (1969) J. Biol. Chem. 244, 4406–4412.
[13] Yphantis, D. A. (1964) Biochemistry 3, 297–317.
[14] Bretscher, M. S. (1971) Nature London New Biol. 231, 229–232.
[15] Segrest, J. P., Jackson, R. L., Andrews, E. P. and Marchesi, V. T. (1971) Biochem. Biophys. Res. Commun. 44, 390–395.
[16] Segrest, J. P. and Jackson, R. L. (1972) Methods Enzymol. 28, 54–63.
[17] Lee, J. C. and Timasheff, S. N. (1974) Arch. Biochem. Biophys. 165, 268–273.
[18] Cohn, E. J. and Edsall, J. T. (1943) Proteins, Amino Acids and Peptides, p. 372, Reinhold, New York.
[19] Gibbons, R. A. (1972) in: Glycoproteins (Gottschalk, A., ed.) 2nd edn., Part A, p. 78, Elsevier, Amsterdam.
[20] Stanley, P. M. and Haslam, E. A. (1971) Virology 46, 764–773.
[21] Skehel, J. J. and Schild, G. C. (1971) Virology 44, 396–408.
[22] Webster, R. C. (1970) Virology 40, 643–654.
[23] Skehel, J. J. (1972) Virology 49, 23–36.
[24] Neuberger, A. Gottschalk, A., Marshall, R. D. and Spiro, R. G. (1972) in: Glycoproteins (Gottschalk, A., ed.) 2 nd edn., Part A, pp. 450–490, Elsevier, Amsterdam.
[25] Skehel, J. J. and Waterfield, M. D. (1974) Proc. Nat. Acad. Sci. USA 72, 93–97.
[26] Laver, W. G. and Baker, N. (1972) J. Gen. Virol. 17, 61–67.
[27] Laver, W. G. and Webster, R. G. (1972) Virology 48, 445–455.