Proc. Natl. Acad. Sci. USA
Vol. 75, No. 6, pp. 2737–2740, June 1978
Biochemistry

Purification of the fusion protein of Sendai virus: Analysis of the NH_2-terminal sequence generated during precursor activation

(activation of infectivity/influenza hemagglutinin/hydrophobic amino terminus)

M. J. GETHING*, J. M. WHITE†, AND M. D. WATERFIELD*

* Department of Protein Chemistry, Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London, WC2A 3PX England; and † Gibbs Laboratory, Harvard University, Cambridge, Massachusetts 02138

Communicated by William N. Lipscomb, April 10, 1978

ABSTRACT **The two glycoproteins of Sendai virus, the hemagglutinin–neuraminidase and the fusion protein (F), were separated and purified by affinity chromatography on a *Lens culinaris* lectin–Sepharose column. F was shown to consist of two disulfide-bonded glycopolypeptide chains, F_1 and F_2, of molecular weights 51,000 and 11,000, each of which contained 15% carbohydrate by weight. Amino-terminal sequence analysis showed that F_2 was blocked and that the hydrophobic sequence NH_2-Phe-Phe-Gly-Ala-Val-Ile-Gly-Ile-Ile-Ala-Leu-Gly-Pro-Ala-Thr- was at the amino terminus of F_1. This sequence shows identity at six positions with the hydrophobic amino-terminal sequence of the smaller glycopolypeptide chain, HA_2, of the hemagglutinin of influenza virus. Both F_1 and HA_2 are formed by proteolytic cleavage of precursor glycoproteins (F_0, Sendai virus; HA_0, influenza virus). Since these cleavages confer infectivity upon both Sendai and influenza viruses and the ability to induce cell-to-cell fusion upon Sendai virus, the hydrophobic NH_2-terminal sequences on F_1 and HA_2 may play a role in fusion of viral and host-cell membranes.**

The results of recent studies of the biology and structure of Sendai virus grown in different types of cells suggest that one of the envelope glycoproteins is involved in the processes of virus-cell and cell-to-cell fusion. Virus grown in L cells or MDBK cells is noninfectious and lacks fusion activity, whereas virus grown in embryonated eggs is both infectious and capable of fusing cells. Comparative analysis of the virus proteins by sodium dodecyl sulfate ($NaDodSO_4$)/polyacrylamide gel electrophoresis has shown that biological activation is accompanied by proteolytic cleavage of a precursor envelope glycoprotein, F_0, and the production of the glycoprotein F (1, 2), which consists of two disulfide-linked polypeptide chains (3, 4). An analogous phenomenon has been observed for the activation of infectivity of influenza virus involving proteolytic cleavage of the precursor polypeptide HA_0 to two disulfide-bonded polypeptide chains, HA_1 and HA_2 (5, 6). Since the activated forms of the fusion protein of Sendai virus (F) and the hemagglutinin of influenza virus (HA) are required for infectivity (and cell fusion in the case of Sendai), we have compared the primary structures of the polypeptides F and HA.

MATERIALS AND METHODS

Iodo[^{14}C]acetamide (57 Ci/mol) and potassium boro[^{3}H]hydride (3.3 Ci/mmol) were obtained from the Radiochemical Centre, Amersham, Bucks, United Kingdom. Iodoacetamide (British Drug Houses, Poole, Dorset, United Kingdom), was recrystallized from ethyl acetate. α-Methyl-D-mannoside and galactose oxidase (type 1 from *Polyporus circinatus*) were from Sigma Chemical Co., Ltd. Empigen BB was a gift from Albright and Wilson Ltd., Marchon Division, Whitehaven, United Kingdom, and sodium deoxycholate was obtained as a specially pure grade from Fisons Scientific Apparatus, Loughborough, Leicestershire, United Kingdom. Ultrogel AcA34 was obtained from LKB-Produkter AB, Sweden. Reagents for sequence analysis were from Beckman, Glenrothies, Scotland.

Virus Growth, Purification, and Labeling. Sendai virus was grown in 10-day-old embryonated chicken eggs. The virus was concentrated from allantoic fluid by centrifugation at 120,000 × *g* for 1 hr at 4°, resuspended in phosphate-buffered saline (P_i/NaCl), and purified by density gradient centrifugation [20–55% sucrose (wt/vol) in P_i/NaCl] at 100,000 × *g* for 2 hr at 4°.

Virus particles (10 mg/ml in P_i/NaCl and 2 mM phenylmethylsulfonyl fluoride) were treated with galactose oxidase (40 units, 100 μl in P_i/NaCl) for 1 hr at 20°; after centrifugation through 20% sucrose (wt/vol) onto a shelf of 55% sucrose (wt/vol), they were reduced with potassium boro[^{3}H]hydride (50 μl in 10 mM NaOH; 5 mCi; 3.3 Ci/mmol) for 20 min at 20°.

Polyacrylamide Gel Electrophoresis. $NaDodSO_4$ gel electrophoresis was performed in gels of 10% acrylamide/0.27% *N,N'*-methylenebisacrylamide. The discontinuous buffer system was that described by Laemmli (7); electrophoresis was conducted at a constant current of 22 mA. Samples contained 1% $NaDodSO_4$ and 1% 2-mercaptoethanol. Gels of ^{3}H-labeled proteins were treated for fluorography by the method of Bonner and Laskey (8).

Separation of F and Hemagglutinin–Neuraminidase (HN) Glycoproteins on *Lens culinaris* Lectin (LcH)–Sepharose. Purified Sendai virus (70 mg) was dissociated in 50 ml of 1% Empigen BB/10 mM Tris·HCl, pH 6.8/1 mM phenylmethylsulfonyl fluoride, and sonicated for 30 sec at setting 3 on a Dawe Soniprobe (type 7530A). The suspension was dialyzed for 16 hr at 4° against the same buffer and then centrifuged at 120,000 × *g* for 1 hr at 4° to remove nucleocapsids and insoluble matrix protein. The supernatant containing the solubilized glycoproteins was applied to a column (2.6 × 45 cm) of LcH–Sepharose (9) equilibrated with 1% Empigen BB (pH 5.5) which was washed sequentially with 300 ml of the equilibrating buffer, 500 ml of 1% Empigen BB/0.1 M α-methyl-D-mannoside at pH 6.8, and 500 ml of 1% sodium deoxycholate/0.1 M α-methyl-D-mannoside at pH 8.3. The flow rate was 60 ml/hr and 10-ml fractions were collected. The two glycoprotein peaks were collected separately and concentrated by ultrafiltration.

The partially purified F glycoprotein (3 ml, 1.3 mg/ml) was applied to a column of Ultrogel AcA34 (1.6 cm × 2 m) equili-

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: F, fusion protein of Sendai virus; HA, hemagglutinin of influenza virus; $NaDodSO_4$, sodium dodecyl sulfate; P_i/NaCl, phosphate-buffered saline; LcH, *Lens culinaris* lectin; HN, hemagglutinin–neuraminidase.

brated with 1% sodium deoxycholate (pH 8.3) and the column was eluted at 10 ml/hr. Fractions containing pure F were pooled and concentrated by ultrafiltration. Deoxycholate was removed by dialysis for 4 days against 0.5% pyridine (pH 8.5).

Reduction and Alkylation of F. F glycoprotein (3 mg) in 6 M guanidine hydrochloride/0.1 M Tris at pH 8.3 was reduced with 1 mM dithiothreitol for 2 hr at 37°. Iodo[^{14}C]acetamide (10 μmol, 5.7 mCi/mol) was then added and the mixture was incubated for 1 hr at 37°. Excess iodoacetamide was reacted with 2-mercaptoethanol (150 μmol) and the solution was dialyzed for 16 hr against 6 M urea/0.2 M formic acid. The alkylated sample was then applied to a column of Sephadex G-100 equilibrated with 6 M urea/0.2 M formic acid and the column was eluted at 10 ml/hr. Fractions containing F_1 and F_2 were pooled separately, dialyzed against 0.5% pyridine (pH 8.3), and lyophilized. Hemagglutination and neuraminidase assays were made by standard procedures (10, 11).

Amino Acid Sequence Analysis. Amino-terminal sequences were determined with a Beckman 890C Sequencer and the 0.1 M Quadrol program of Brauer *et al.* (12). Phenylthiohydantoin-amino acids were identified by reverse-phase high-pressure liquid chromatography by using a Waters chromatograph with a Bondapak C18 column. Assignments of amino acids were confirmed by back hydrolysis of phenylthiohydantoin-amino acids in hydroiodic acid (13), followed by amino acid analysis on a Durrum D500 amino acid analyzer, and by gas chromatographic analysis by the techniques of Pisano and Bronzert (14). Amino acid analyses were performed in 0.5 ml of constant boiling HCl containing 3 μl of butane dithiol and 0.1% phenol according to Smyth *et al.* (15) with a Durrum D500 amino acid analyzer. Carbohydrates were analyzed as trimethyl silyl glucosides by the method of Clamp (16).

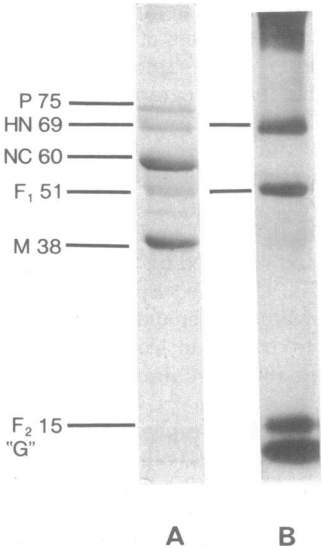

FIG. 1. Polyacrylamide gel electrophoresis of the polypeptides of Sendai virus grown in embryonated fowl eggs. Sizes of proteins are given in kilodaltons. P, Polymerase; HN, hemagglutinin-neuraminidase; NC, nucleocapsid protein; M, matrix protein. (*A*) Stained with Coomassie brilliant blue. The unlabeled band between F_1 and M may contain cellular actin (17). The three bands at the top of the gel may contain unreduced HN and P oligomers (3). (*B*) Autoradiogram of a gel of virus labeled with boro[^{3}H]hydride by the galactose oxidase method. Because this gel was intentionally overloaded, [^{3}H]aggregates of HN and F are seen at the top of the gel.

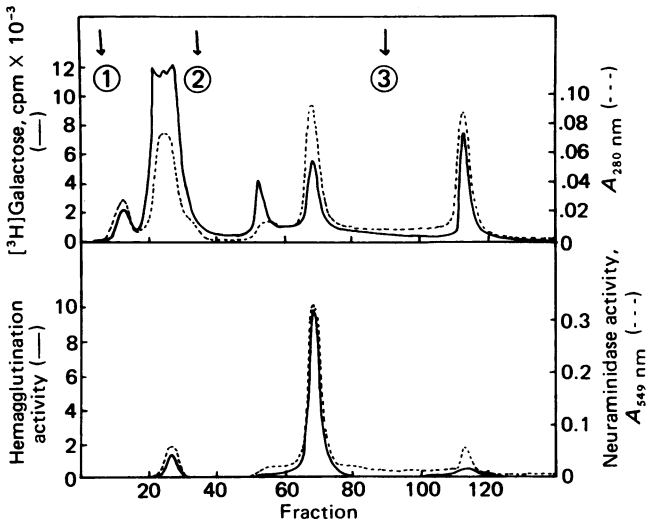

FIG. 2. Fractionation of Sendai virus glycoproteins by affinity chromatography on LcH–Sepharose. Aliquots of 50 μl were analyzed for radioactivity; 5-μl aliquots for hemagglutinin and neuraminidase activities. Arrow 1, 1% Empigen BB (pH 5.5); arrow 2, 1% Empigen BB/0.1 M α-methylmannoside at pH 6.8; arrow 3, 1% deoxycholate/0.1 M α-methylmannoside at pH 8.3.

RESULTS

Purification of F and Separation of Its Polypeptide Components. The proteins of egg-grown infectious Sendai virus can be resolved on NaDodSO$_4$/polyacrylamide gels stained with Coomassie brilliant blue into five distant bands (Fig. 1*A*): the polymerase, the hemagglutinin–neuraminidase, the nucleocapsid protein, the larger component of the fusion protein (F_1), and the matrix protein. When virus particles are treated with galactose oxidase and then reduced with boro[^{3}H]hydride (18), four labeled bands appear (Fig. 1*B*). One of these bands runs close to the dye front and is probably glycolipid. The other bands correspond to the HN and F_1 polypeptides and to the smaller component of the fusion protein, F_2, which does not stain with Coomassie brilliant blue (4).

Radioactively labeled Sendai virus was disrupted in detergent and the solubilized components were applied to a column of LcH–Sepharose. The column effluent was monitored for protein ($A_{280\ nm}$), ^{3}H label, and hemagglutinating and neuraminidase activities (Fig. 2). Analysis on NaDodSO$_4$/polyacrylamide gels showed that the material that did not bind to the column contained lipid and residual nonglycosylated proteins (polymerase and nucleocapsid and matrix proteins). Both glycoproteins bound to the column; elution with α-methyl-D-mannoside solutions containing Empigen BB resulted in the displacement of a small amount of lipid, followed by fractions which contained about 90% of the HN and about 15% of the F glycoproteins. The remainder of the F glycoprotein was eluted with α-methyl-D-mannoside solutions containing deoxycholate. This procedure resulted in the partial purification of both glycoproteins. The F component was further purified by chromatography on Ultrogel AcA34 in 1% deoxycholate (pH 8.3). As shown in Fig. 3, F was eluted as a single asymmetric peak, and NaDodSO$_4$/acrylamide gel electrophoresis of the peak fractions showed that only those of the trailing edge of the peak contained HN.

The F protein purified in this way was dialyzed to remove deoxycholate. Reduction and alkylation of the F protein with dithiothreitol and iodo[^{14}C]acetamide followed by chromatography on Sephadex G-100 in 6 M urea/0.2 M formic acid

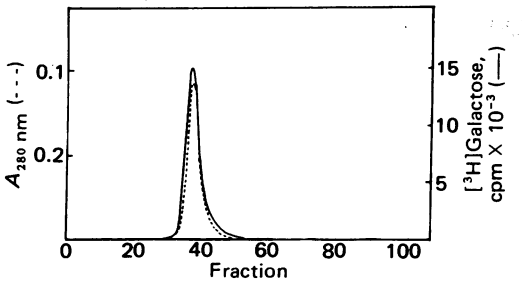

FIG. 3. Chromatography of F on Ultrogel AcA34. The fractions containing F from the LcH–Sepharose column (Fig. 2) were pooled, concentrated, and chromatographed on Ultrogel AcA34. Aliquots of 50 μl were analyzed for radioactivity.

showed that the F protein consists of two disulfide-bonded polypeptide chains, F_1 and F_2/ F_1 was excluded from the column and F_2 was eluted close to the included volume of the column (Fig. 4). Similar separations of F_1 and F_2 were also made by chromatography on Sepharose 4B in 6 M guanidine hydrochloride (pH 4.5). Under these conditions, F_1 was again excluded from the column while F_2 was eluted as a polypeptide of apparent molecular weight 11,000 by comparison with standard globular proteins (19).

Structural Analysis of F Glycoprotein and Its Glycopolypeptide Components. The results of amino acid and carbohydrate analyses of the fusion protein are shown in Table 1. The calculated figures presented are based on apparent molecular weights of F and F_1 determined by NaDodSO$_4$/polyacrylamide gel electrophoresis and on the apparent molecular weight of F_2 determined by column chromatography on Sepharose 4B in guanidine hydrochloride.

The results of amino acid sequence analysis of the 15 amino-terminal residues of F showed a single sequence (Fig. 5). Since F contains F_1 and F_2, this result suggested that the amino terminus of one of these was blocked to Edman degradation. Separate analyses of F_1 and F_2 showed that F_2 was blocked whereas F_1 had the amino-terminal sequence found on intact F. This sequence shows identity at six positions with the hydrophobic NH_2-terminal sequence on the smaller polypeptide chain (HA_2) of the hemagglutinin of influenza virus.

DISCUSSION

Milligram quantities of the Sendai virus fusion protein have been purified by affinity and molecular sieve chromatography in detergents and a preliminary analysis of its primary structure has been made. The two viral glycoproteins were separated by sequential elution of HN and F components bound to a lectin affinity column. Final purification of F involved gel filtration in deoxycholate.

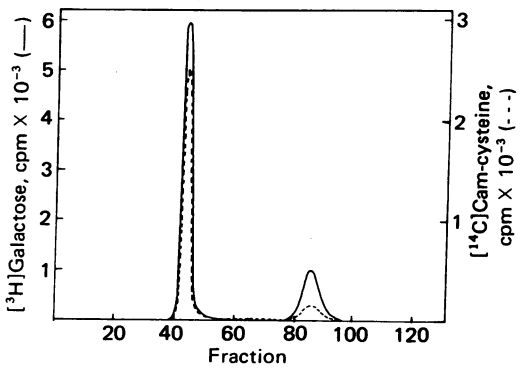

FIG. 4. Separation of F_1 and F_2 polypeptides on Sephadex G-100 after reduction and alkylation of F. Aliquots of 100 μl were analyzed for radioactivity. Cam-cysteine, carbamoylmethyl-cysteine.

Table 1. Amino acid and carbohydrate compositions of F, F_1, and F_2 glycopolypeptides

Residue	F, residues/ 500 amino acids	F_1, residues/ 410 amino acids	F_2, residues/ 90 amino acids	F_1 + F_2
Amino acid				
Aspartic acid	48.8	40.8	10.6	51.4
Threonine	36.8	31.4	5.4	36.8
Serine	35.8	26.9	6.8	33.7
Glutamic acid	50.7	41.2	10.7	51.9
Proline	25.9	20.0	4.7	24.7
Glycine	39.8	33.0	7.9	40.9
Alanine	39.3	33.8	5.7	39.5
Valine	33.8	28.9	7.1	36.0
Cysteine*	ND†	4.9	0.9	5.8
Methionine	6.0	3.7‡	0.0	3.7
Isoleucine	42.3	35.5	6.5	42.0
Leucine	54.7	40.8	11.8	52.6
Tyrosine	14.4	11.8	2.3	14.1
Phenylalanine	10.4	10.2	0.8	11.0
Histidine	8.0	6.1	1.1	7.2
Lysine	22.9	18.7	3.1	21.8
Arginine	25.4	22.4	4.5	26.9
Total amino acids	500	410	90	500
M_r of polypeptide	54,100	44,400	9,680	54,080
Carbohydrate				
N-Acetylglucosamine	21.4	11.7	3.8	15.5
N-Acetylgalactosamine	0.0	0.0	0.0	0.0
Fucose§	ND	ND	ND	ND
Mannose	17.9	11.9	2.8	14.7
Galactose	19.3	16.3	2.5	18.7
Sialic acid	0.0	0.0	0.0	0.0
Total % carbohydrate	16	14	15	14
M_r of glycopolypeptide	64,700	51,500	11,300	62,800

* Determined as carboxymethyl cysteine.
† Not determined.
‡ This determination is probably low due to partial oxidation of methionine.
§ Fucose was not determined due to interfering compounds.

It has previously been shown by NaDodSO$_4$/polyacrylamide gel electrophoresis that F consists of two polypeptide chains (3, 4) of molecular weights approximately 50,000 and 13,000. In the present study, these two chains, F_1 and F_2, were isolated by gel filtration chromatography after complete reduction and alkylation of disulfide bonds. The smaller component, F_2, eluted from Sepharose 4B in guanidine hydrochloride columns with an apparent molecular weight of 11,000, a somewhat smaller estimate than that obtained by NaDodSO$_4$/polyacrylamide gel electrophoresis. The size of F_1 could not be estimated by chromatography because this component aggregates in both urea and guanidine hydrochloride solutions. An apparent molecular weight of 51,000 was, however, determined by NaDodSO$_4$/polyacrylamide gel electrophoresis. The result of analyses of the amino acid and carbohydrate compositions of F, F_1, and F_2 indicate that F contains about 500 amino acid residues, F_1, 410, and F_2, 90. All three components contain

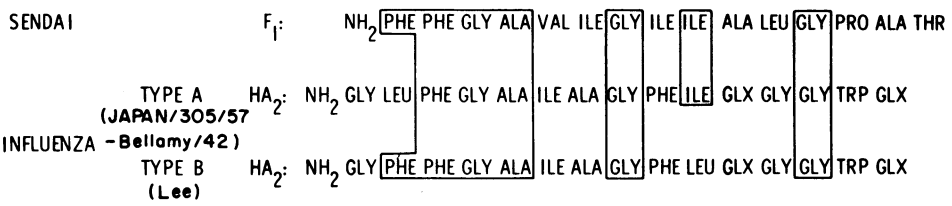

FIG. 5. Amino-terminal sequences of the F_1 component of the Sendai virus fusion protein and the HA_2 component of the hemagglutinin of influenza virus generated during precursor activation.

N-acetylglucosamine, mannose, and galactose, which together account for approximately 15% of their weight. These composition estimates allow molecular weights of 64,700, 51,500, and 11,300 to be calculated for F, F_1, and F_2, respectively. The determination of more accurate molecular weights must await complete sequence analysis of these glycopolypeptides.

It has previously been shown that F is generated from the precursor protein F_0 by proteolytic cleavage and that this cleavage is essential for the production of infectious Sendai virus that is also capable of fusing cells (1, 2). The amino acid sequence analysis presented here has shown that the amino terminus of F_2 is blocked to Edman degradation while the amino terminus of F_1 is unblocked. These results, together with the observation that F_0 cannot be dansylated, suggest that the amino terminus of F_1 is generated by proteolysis and that the order of biosynthesis of the glycoprotein components is NH_2-F_2-F_1-COOH (see ref. 4).

Details of the association of F with the lipid bilayer are not known. Studies of the hemagglutinin (HA) of influenza virus have shown that the carboxy-terminal region of the HA_2 polypeptide is associated with the bilayer (20), and in certain strains of virus (the Hong Kong viruses) the amino terminus of HA_1 is blocked to Edman degradation (unpublished observation). Studies of the major glycoproteins of erythrocytes (21, 22) and of vesicular stomatitis virus (23) have shown that these proteins are anchored into their membranes by a carboxy-terminal region that spans the bilayer. We therefore suggest that the carboxy-terminal region of F_1 is associated with the viral membrane. Further evidence in support of this orientation is the observation that F_1 but not F_2 aggregates in solutions of urea and guanidine hydrochloride. [The HA_2 polypeptide of influenza virus also aggregates in these solvents (20).]

Biological activation of Sendai virus involves proteolytic cleavage of F_0, thereby generating a new carboxyl terminus on F_2 and a new amino terminus on F_1. The amino terminus of F_1 is very hydrophobic and may participate in the interaction of the fusion protein with the host-cell plasma membrane during infection and fusion. However, there is no direct experimental evidence to support this speculation.

Finally, it is of particular interest that certain features of the activation of Sendai virus infectivity and fusion capability are analogous to the proteolytic activation of the infectivity of influenza virus. In both cases, proteolytic cleavage generates two disulfide-bonded polypeptide chains. The newly generated amino-terminal sequences on F_1 and HA_2 are very hydrophobic and have 6 of the first 15 residues in common. These similarities may be significant for the processes of virus–cell and cell–cell fusion.

We thank Don Wiley and John Skehel for many helpful discussions and John Seaver for excellent technical assistance. J.W. acknowledges support from National Institutes of Health Training Grant in Biophysics (Harvard University) 5T01GM00782-19 and Grants GB-43573X from the Program in Human Cell Biology of the National Science Foundation and AI-13654 from the Institute for Allergy and Infectious Diseases of the National Institutes of Health to Don Wiley.

1. Homma, M. & Ohuchi, M. (1973) *J. Virol.* **12**, 1457–1465.
2. Scheid, A. & Choppin, P. W. (1974) *Virology* **57**, 475–490.
3. Ozawa, M., Asano, A. & Okada, Y. (1976) *FEBS Lett.*, **70**, 145–149.
4. Scheid, A. & Choppin, P. W. (1977) *Virology* **80**, 54–66.
5. Klenk, H.-D., Rott, R., Orlich, M. & Blodorn, J. (1975) *Virology*, **68**, 426–439.
6. Lazarowitz, S. & Choppin, P. (1975) *Virology* **68**, 440–454.
7. Laemmli, U. K. (1970) *Nature* **227**, 680–685.
8. Bonner, W. M. & Laskey, R. A. (1974) *Eur. J. Biochem.* **46**, 83–88.
9. Hayman, M. J., Skehel, J. & Crumpton, M. J. (1973) *FEBS. Lett.* **29**, 185–188.
10. Palmer, D. E., Coleman, M. T., Dowdle, W. R. & Schild, G. C. (1975) *Advanced Laboratory Techniques for Influenza Diagnosis*. U.S. Department of Health, Education and Welfare, Public Health Service Publication (U.S. Public Health Service, Washington, DC).
11. Aminoff, D. (1961) *Biochem. J.* **81**, 384–392.
12. Brauer, A. W., Margolies, M. N. & Haber, E. (1975) *Biochemistry* **14**, 3029–3035.
13. Smithies, O., Gilson, D., Fanning, E. M., Goodfleish, R. M., Gilman, J. G. & Ballantyne, D. C. (1971) *Biochemistry* **10**, 4912–4921.
14. Pisano, J. J. & Bronzert, T. J. (1969) *J. Biol. Chem.* **244**, 5597–5607.
15. Smyth, D. G., Stein, W. H. & Moore, S. (1963) *J. Biol. Chem.* **238**, 227–234.
16. Clamp, J. R. (1974) *Biochem. Soc. Symp.* **40**, 3–16.
17. Wang, E., Wolf, B., Lamb, R., Choppin, P. & Goldberg, A. (1976) in *Cell Motility*, eds. Goldman, R., Pollard, T. & Rosenbaum, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), pp. 589–599.
18. Critchley, D. R. (1974) *Cell* **3**, 121–125.
19. Mann, K. G. & Fish, W. W. (1972) in *Methods in Enzymology*, ed. Hirs, C. H. W. (Academic, New York), Vol. 26, pp. 28–42.
20. Skehel, J. J. & Waterfield, M. D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 93–97.
21. Bretscher, M. S. (1971) *Nature New Biol.* **231**, 229–232.
22. Tomita, M. & Marchesi, V. T. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2964–2968.
23. Katz, F., Rothman, J., Lingappa, V., Blobel, G. & Lodish, H. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3278–3282.