Proc. Nat. Acad. Sci. USA
Vol. 73, No. 3, pp. 915–918, March 1976
Immunology

Localization of the papain cleavage site of H-2 glycoproteins

(transplantation antigens/*H-2* gene products/cell membranes/membrane orientation/sequence analysis)

BRUCE M. EWENSTEIN*, JOHN H. FREED*, LAWRENCE E. MOLE†, AND STANLEY G. NATHENSON* ‡

* Departments of Microbiology and Immunology, and Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461; and † MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, Oxford, England

Communicated by Harry Eagle, December 11, 1975

ABSTRACT The antigenic products of the murine *H-2K* and *H-2D* genes are glycoproteins of about 45,000 molecular weight which are tightly integrated within the cell surface membrane. A glycoprotein fragment (F_{Ag}, antigenic fragment) of 37,000 daltons carrying the carbohydrate, antigenic sites, and the associated putative β_2-microglobulin of 12,000 daltons can be generated by papain cleavage either of the native molecules in the cell membrane or of immune precipitates made from the antigen solubilized by nonionic detergent. Partial NH_2-terminal sequence analyses of the native H-2 glycoprotein and of the papain-cleaved glycoprotein fragment establish that the fragment is, in fact, the NH_2-terminal portion of the native molecule. Thus, the cleavage by papain proteolysis is near the COOH-terminus, and removal of the COOH-terminal portion (F_m, membrane fragment) converts the glycoprotein to a water-soluble form. This observation suggests that the NH_2-terminus of the native glycoprotein extends out of the hydrophobic bilayer of the cell membrane, and that the COOH-terminus contains the membrane binding region and is buried within the bilayer.

The major histocompatibility complex (MHC) located on chromosome 17 of the mouse is a complex of genes which determines a diversity of immune and other functions (see reviews, refs. 1–4) associated with histocompatibility (1), resistance to leukemogenesis (5), immune responsiveness (6), mixed lymphocyte reactivity and cell mediated cytolysis (see reviews, refs. 2–4), cyclic-AMP levels (7), testosterone levels (8), and complement levels (9). The *H-2K* gene and the *H-2D* gene are located at the extreme ends of the MHC and are separated by about 0.5 centimorgan. The glycoprotein products of these two genes are the targets of graft rejection. Each is tightly integrated into the cell surface membrane matrix, has a basic molecular weight of about 45,000 (10), and carries carbohydrate chains of 3,300 daltons (11). In the cell membrane, the glycoproteins are noncovalently associated, probably in a 1:1 molar ratio, with a smaller protein of about 12,000 daltons, the H-2 associated protein, which recently has been tentatively identified as the murine β_2-microglobulin (12–15).

The antigenic reactivity of the H-2 glycoproteins is determined by their protein structure (see review, ref. 16). The extreme polymorphism of these genes and their products is documented by the large number of different serologically detectable determinants on the H-2 products carried by a variety of mouse strains, and by the remarkable and distinct differences between products of alleles of the *H-2K* and *H-2D* gene series revealed by comparative tryptic peptide mapping studies (17).

The H-2 glycoproteins can be extracted from their membrane matrix by dissolution of the membrane with nonionic detergents (18). The glycoprotein–detergent complexes can be purified by biochemical fractionation[§] or isolated by indirect immunoprecipitation (18). While the molecular weight of this native form of the H-2 glycoprotein is 45,000 (10), an antigenically active glycoprotein fragment of about 37,000 daltons (10) can also be isolated if the molecule is cleaved proteolytically from the surface of whole cells (19) or cell membranes with papain (20). This cleavage presumably also leaves an 8,000 dalton fragment in the membrane, although no such fragment has yet been recovered intact. For conciseness, we shall refer to the antigenically active 37,000 dalton fragment as F_{Ag} (antigenic fragment) and the putative membrane fragment left after proteolysis as the F_m (membrane fragment). The F_{Ag} is water soluble and carries the carbohydrate moiety, the private as well as some public H-2 antigenic sites (20), and, by analogy with findings for the homologous HLA glycoproteins, the site for binding of the H-2 associated 12,000 dalton β_2-microglobulin protein (ref. 13; Shimada, A., Yamane, K., and Nathenson, S. G. unpublished observation).

It is reasonable to suggest that along the protein backbone of each native H-2 glycoprotein there are various regions organized for different special functional purposes. Thus, present knowledge (16) predicts that there should be a region for membrane integration, one for β_2-microglobulin binding, one where the carbohydrate chain is attached, and one or more regions that express the highly polymorphic antigenic site(s).

The present study approaches the problem of the intramolecular arrangement of these sites by analysis of the partial NH_2-terminal sequence of the native H-2 glycoprotein, and its derivative F_{Ag}. In these analyses automatic sequencing techniques were used on radiolabeled material. Only isotopically labeled residues are detected and, thus, only minute quantities of antigen are required. However, only a few amino acids can be determined at a time during each sequencer run. The method, therefore, provides only a partial sequence or profile for a particular amino acid.

The profiles for leucine and arginine were determined for the 30 NH_2-terminal residues of both the Nonidet P-40(NP-40)-solubilized native H-2 glycoprotein, and the F_{Ag}. Comparison of these profiles would establish whether the F_{Ag} arose from the COOH- or NH_2-terminal part of the native molecule. Since cleavage by papain converts the water-insoluble native glycoprotein into the water-soluble F_{Ag}, identification of this fragment permits localization of the site of papain cleavage, and by deduction, can also position the region of membrane binding.

MATERIALS AND METHODS

Antisera. Antisera to detect H-2 specificity 2 were raised

Abbreviations: MHC, major histocompatibility complex; F_{Ag} and F_m, antigenic and membrane fragment, respectively; $NaDodSO_4$, sodium dodecyl sulfate; NP-40, Nonidet P-40.

‡ To whom reprint requests should be addressed.

§ J. H. Freed, D. W. Sears, D. L. Mann, and S. G. Nathenson, in preparation.

in (HTI × BALB/c)F_1 ($H\text{-}2^i \times H\text{-}2^d$) mice that were injected with EL-4 ($H\text{-}2^b$) cells; antisera to detect H-2 specificity 33 were raised in (B10.D2 × A)F_1 ($H\text{-}2^d \times H\text{-}2^a$) mice injected with HTI ($H\text{-}2^i$) cells.

Preparation and Purification of Antigen Fractions. H-2K and H-2D alloantigens were isolated from EL-4, a continuous lymphoblastoid cell culture line ($H\text{-}2^b$). Radiolabeled NP-40 extracts were prepared as described previously (17); however, before isolation by specific anti-H-2 immunoprecipitation, the extracts were partially purified by lentil lectin affinity chromatography.§ The native H-2 antigen prepared by immunoprecipitation (10) contained the H-2 glycoprotein and its associated protein, the putative β_2-microglobulin. The glycoprotein and the β_2-microglobulin were separated by column chromatography (0.9 × 55 cm) on Sephadex G-150 Superfine in a sodium dodecyl sulfate ($NaDodSO_4$)-containing buffer (0.1% $NaDodSO_4$, 0.01 M Tris·HCl; pH 7.4); the elution rate was 1–2 ml/hr. The separated components each migrated as a single band upon $NaDodSO_4$-discontinuous alkaline polyacrylamide gel electrophoresis (21).

The 37,000 dalton F_{Ag} was prepared either from immune precipitates (10) or directly from membrane fractions (20). As shown in Fig. 1, separation of this fragment from other radioactive products after cleavage of immune precipitates was obtained by column chromatography (0.9 cm × 115 cm) on Sephadex G-75 in 1 M formic acid (90% of applied radioactivity recovered from pools). The F_{Ag} (Fig. 1, pool II) ran as a single peak on $NaDodSO_4$/polyacrylamide gel electrophoresis with an apparent molecular weight of 37,000. Material from pool III also appeared to be homogeneous by this criterion and to have an apparent molecular weight of 12,000 daltons.

Sequence Determination. A model 890C Beckman Sequencer was used for the NH_2-terminal sequence analysis using the 0.1 M Quadrol program (22). Immune precipitates, containing from approximately 25,000 to 50,000 cpm of radiolabeled antigen and 10 mg of carrier immunoglobulin (goat anti-mouse IgG plus mouse anti-H-2 IgG), plus an additional 2 mg of myoglobin, were applied to the sequencer cup in 50% acetic acid. The radioactivity in the thiazolinone derivatives from each step was determined directly by measuring the radioactivity of aliquots in an LKB–Wallach Ultro Beta 1210 liquid scintillation counter (LKB Instruments, Ltd.) in a solubilizer-scintillation fluid (NE320: NE233, 1:5; Nuclear Enterprises, Ltd.). The remainder of each sample was converted to the phenylthiohydantoin (23) and analyzed by gas chromatography and by thin-layer chromatography (24). The dominant residue at each step, as determined by the latter methods, corresponded to that of the known sequence of the myoglobin carrier.

For separation of the H-2 glycoproteins from their associated β_2-microglobulin, a whole precipitate was dissolved in 2% $NaDodSO_4$, boiled for 2 min under nitrogen, and chromatographed on a Sephadex G-150 Superfine column (0.9 × 45 cm) by eluting with a 0.1% $NaDodSO_4$-containing buffer. The peaks containing the separated molecules were pooled and dialyzed twice against a 500-fold excess of 0.001% $NaDodSO_4$ and an appropriate aliquot was added directly to the sequencer cup.

RESULTS

Antigenic material was prepared from $H\text{-}2^b$ cells that had been radiolabeled in culture with either [^{3}H]leucine or [^{3}H]arginine. The purified NP-40 extracted glycoprotein

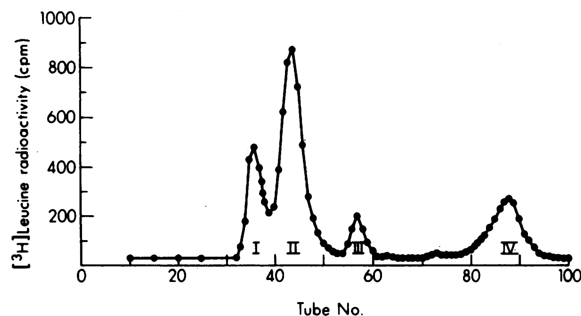

FIG. 1. Sephadex G-75 column chromatographic separation of the products of papain cleavage of an H-2 glycoprotein. The mixture produced by papain cleavage of a [^{3}H]leucine-labeled H-2.33 (350,000 cpm) immune precipitate was separated as described in *Materials and Methods*.

and the papain-solubilized F_{Ag}, prepared as described in *Materials and Methods* (Fig. 1), each gave a single peak, as judged by $NaDodSO_4$/polyacrylamide gel analysis. The β_2-microglobulin also ran as a single band.

The native NP-40-solubilized glycoproteins were analyzed first. Both the $H\text{-}2K^b$ (H-2.33) and the $H\text{-}2D^b$ (H-2.2) gene products were examined. Fig. 2 shows the results from the automatic sequencer run for the glycoprotein bearing the H-2.33 specificity. Panel A, which gives results obtained for the whole, unfractionated immune precipitate, was carried out as an internal control to verify the specificity of the method since mouse β_2-microglobulin has an unblocked NH_2-terminus, and leucine occurs at position 23 in the β_2-microglobulin of all four mammalian species reported, including rat (Poulik, Shinnick, and Smithies, personal communication). A plot of the radioactivity for the thiazolinone derivatives (Fig. 2A) shows the major peaks of leucine at positions 5 and 17, as well as a smaller peak, roughly half as large, at residue 23. These results thus are consistent with the

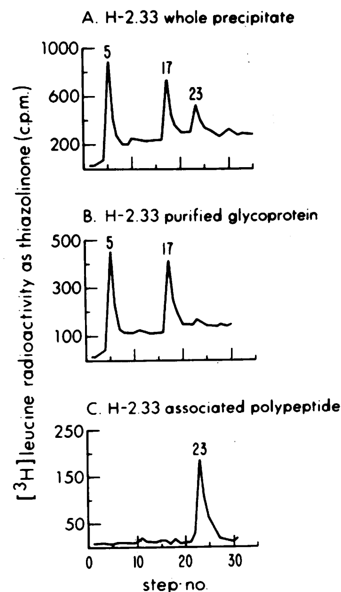

FIG. 2. Radioactive leucine amino-acid profile analysis for the 30 NH_2-terminal residues (H-2.33) of: (A) an immune precipitate (75,000 cpm) containing both the native glycoprotein and the associated 12,000 dalton β_2-microglobulin; (B) the native H-2 glycoprotein (35,000 cpm); and (C) the 12,000 dalton β_2-microglobulin (15,000 cpm). Preparations and sequencing methods are described in *Materials and Methods*.

Table 1. Amino-acid profiles for arginine- and leucine-labeled H-2 alloantigens

		Position																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Native (NP-40-solubilized)	H-2.33 ($H\text{-}2K^b$)	—	—	—	—	L	R	—	—	—	—	—	—	—	R	—	—	L	—	—	—	R	—	—	—	—	—	—	—	—	—
	H-2.2 ($H\text{-}2D^b$)	—	—	—	—	—	R	—	—	—	—	—	—	—	R	—	—	L	—	—	—	R	—	—	—	—	—	—	—	—	—
Papain-cleaved (F_{Ag})	H-2.33	—	—	—	—	L	R	—	—	—	—	—	—	—	R	—	—	L	—	—	—	R	—	—	—	—	—	—	—	—	—

Amino-acid code: L = leucine; R = arginine; — = neither leucine nor arginine is present at this residue. These data were obtained from experiments discussed in the *text*.

conclusion that the sequence technique is correctly detecting the leucine 23 of β_2-microglobulin, and that the leucines appearing at 5 and 17 are from the H-2.33 molecule.

This conclusion was verified further by sequence analysis of the separate H-2.33 glycoprotein (Fig. 2B) and the 12,000 dalton β_2-microglobulin (Fig. 2C). Clearly, the H-2.33 glycoprotein has leucine at positions 5 and 17 (Fig. 2B). Panel C shows that the profile for the H-2 associated 12,000 dalton protein has a leucine at residue 23. These data come from a protein isolated from a papain-digested preparation, although identical results were found with material separated from the native glycoprotein by $NaDodSO_4$/gel filtration chromatography (see *Materials and Methods*).

The arginine and leucine profiles for the first 30 residues were determined for both H-2.33 and H-2.2. Table 1 summarizes data obtained from sequencer runs similar to those shown in Fig. 2. The H-2.33 glycoprotein has leucine at residue 5 and residue 17 and shows arginine at residues 6, 14, and 21. The H-2.2 glycoprotein lacks leucine at position 5 but has leucine at residue 17, as well as arginine at residues 6, 14, and 21.

Analysis of the papain-cleaved F_{Ag} was determined next. The sequencer runs of a leucine-labeled (Fig. 3A) and arginine-labeled (Fig. 3B) H-2.33 F_{Ag} showed, as was found for the native glycoprotein, that leucine was present at residues 5 and 17, and arginine at residues 6, 14, and 21. The samples for these studies were prepared from immune precipitates. The same results were obtained with samples obtained from papain digests of cell membranes carried out as previously described (20). Thus, we can conclude that the F_{Ag} either of the cell membrane or of the NP-40-solubilized molecule is in fact from the NH_2-terminus of the native glycoprotein.

DISCUSSION

It is apparent that micromethods are prerequisites for the study of the protein structure of biologically important molecules that are available only in small quantities. In previous studies of the peptide structure of the H-2 glycoproteins (17), materials isolated by specific immunochemical precipitation from different H-2 genotypes could be analyzed by comparative tryptic peptide methods because they were internally tagged with isotopically labeled amino-acid residues. Because of the extreme sensitivity of the methods we were able to complete a comparative study of three genotypes that showed a significant degree of peptide differences between products of haplotypes both within a gene series, e.g., K^b versus K^d, and between the two H-2 genes K^b versus D^b.

We have refined our structural analysis to approach the problems of the primary structure of the H-2 glycoproteins. Recent results with immunoglobulins and viral proteins have shown the feasibility of partial sequence analysis through the use of radioisotopically incorporated amino acids (25, 26). We used this methodology, as have others [Silver and Hood (27); J. D. Capra, personal communication; O. Smithies, personal communication], in the study of the structure of the cell surface antigens.

Several arguments may be made to support our contention that the sequences obtained by our techniques are in fact the sequences of H-2 gene products. First, all materials used for the sequence studies were prepared by immunoprecipitation techniques which utilize strict immunogenetic specificity controls, as described previously (17). Second, since each sample after reduction showed essentially single bands upon discontinuous $NaDodSO_4$/polyacrylamide gel analysis, any contaminating protein must have been present in only very small relative molar quantities. Third, the se-

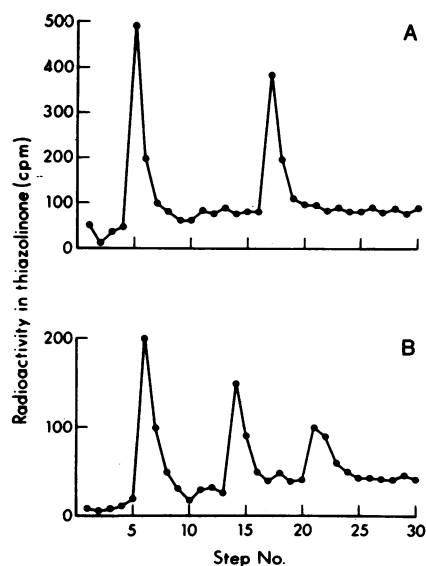

FIG. 3. Radioactive amino-acid profile for the H-2.33 glycoprotein F_{Ag} (37,000 daltons) derived by papain cleavage as described in *Materials and Methods* and Fig. 1. (A) Data from a sample containing 25,000 cpm of [^{3}H]leucine. (B) The results from a sample containing 14,000 cpm of [^{3}H]arginine.

quences for H-2.2 and H-2.33 were obtained using material prepared from the same starting extract by precipitating with different alloantisera. Thus, differences in these sequences must represent differences in molecules with which the antisera reacted, and must not be from contaminants in the preparation.

A final argument involves the yield of radioactive material in the sequencer. For example, in the experiments shown in Fig. 3, on the basis of previously published results (20), the number of leucine residues in the F_{Ag} would be estimated to be around 20 (7 moles percent). If one assumes a constant specific activity of [^{3}H]leucine throughout the molecule, an initial yield on the sequencer of 60% (typical run), and a repetitive yield of 94–96% at each successive step, then between 80 and 100% of the expected radioactivity was found in the thiazolinone derivatives of leucine at residue 5 (see Fig. 3). It is also known that under the conditions used, radiolabeled leucine or arginine is incorporated only as the labeled amino acid added to the cell culture (28).

Comparison of the native H-2 glycoprotein and the H-2 F_{Ag} amino-acid profiles has allowed us to establish that the NH_2-terminal segment of the F_{Ag} is identical with the NH_2-terminal segment of its native molecule. Thus, we can conclude that the site of cleavage by papain is at a point approximately 1/6 from the COOH-terminal end. It is not clear whether or not there are other cleavage sites. Previous data (20) have suggested that an antigenically active 28,000 dalton fragment can also be isolated from papain digests, hence suggesting that secondary cleavage points might exist on the NH_2-terminal fragment. In addition, we have never been able to recover an intact piece of approximately 8000 daltons which would represent the COOH-terminus. It is possible that once the cleavage is made this piece is further degraded.

Since the loss of the small COOH-terminal portion of the native molecule (F_m) correlates with a loss of water insolubility (20), the COOH-terminus must be the region responsible for the presumably hydrophobic interaction either with the membrane or with other membrane proteins.

This fact bears on the problem of membrane orientation. Each membrane integrated glycoprotein is thought to be specifically oriented with its longer axis perpendicular to the plane of the lipid bilayer. It is generally assumed that several features dictate this precise and specific orientation, including a "membrane binding" region, which is hydrophobic and anchors the molecule in the membrane, and the carbohydrate moiety, which is hydrophilic and is positioned on a portion of the protein extending out of the membrane bilayer and in contact with the aqueous external medium. Since our studies suggest that the COOH-terminal 1/6 portion of the H-2 glycoprotein has the membrane binding region (F_m), it follows that the NH_2-terminal portion (F_{Ag}), bearing the carbohydrate, extends out of the hydrophobic membrane environment and thus is in a position to be released by papain cleavage.

Examination of our preliminary data on the NH_2-terminal amino-acid profile reveals that the H-2.2 and H-2.33 molecules share four of the five positions identified as containing arginine or leucine. While such a sample is clearly too small to generalize to the whole molecule, these data support our previous findings of consistent differences in tryptic peptide profiles between these two gene products (17). If borne out by further studies now in progress, it appears that the sequence data support the present concept that each H-2 product is indeed strikingly unique. It is of interest that other studies using similar methods report a leucine at position 5 but not 17 of H-2D^d (H-2.4) and leucines at both positions 5 and 17 of H-2K^d (H-2.31) (D. W. Sears, unpublished observation).

We thank Selma Hadjolian, Diane McGovern, and Jane Addis for their expert technical assistance, and Prof. R. R. Porter, Prof. M. D. Scharff, and Dr. B. Birshtein for illuminating discussions. This work was supported by USPHS Grants 07289 and 10702 and ACS Grant IM-77 (S.G.N.), and The Medical Research Council (L.E.M.). J.H.F. is a fellow of the Arthritis Foundation and S.G.N. is a faculty scholar of the Josiah Macy, Jr. Foundation. B.M.E. was supported by USPHS Training Grant 5T5 GM-1674.

1. Snell, G. D. & Stimpfling, J. H. (1966) in *Biology of the Laboratory Mouse*, ed., Green, E. (McGraw-Hill, Inc., New York), pp. 457–491.
2. Demant, P. (1973) *Transplant. Rev.* **15**, 162–200.
3. Klein, J. (1975) in *Biology of the Mouse Histocompatibility H-2 Complex* (Springer-Verlag, New York, Heidelberg, Berlin).
4. Shreffler, D. C. & David, C. S. (1975) *Adv. Immunol.* **20**, 125–195.
5. Lilly, F. & Pincus, T. (1973) *Adv. Cancer Res.* **17**, 231–277.
6. Benacerraf, B. & McDevitt, H. (1973) *Science* **175**, 273–279.
7. Meruelo, D. & Edidin, M. (1975) *Proc. Nat. Acad. Sci. USA* **72**, 2644–2648.
8. Iványi, P., Hámpl, R., Stárka, L. & Micková, M. (1972) *Nature New Biol.* **238**, 280–281.
9. Hizova, E., Demant, P. & Iványi, P. (1972) *Folia Biol. (Prague)* **18**, 237–243.
10. Schwartz, B. D., Kato, K. Cullen, S. E. & Nathenson, S. G. (1973) *Biochemistry* **12**, 2157–2164.
11. Nathenson, S. G. & Muramatsu, T. (1971) in *Glycoproteins of Blood Cells and Plasma*, eds., Jamieson, G. A. & Greenwalt, T. J. (Lippincott Co., Philadelphia), pp. 254–262.
12. Silver, J. & Hood, L. (1974) *Nature* **249**, 764–765.
13. Rask, L., Lindblom, J. B. & Peterson, P. A. (1974) *Nature* **249**, 833–834.
14. Vitetta, E. S. Uhr, J. W. & Boyse, E. A. (1975) *J. Immunol.* **114**, 252–254.
15. Natori, T., Tanagaki, N., Nahamuno, K., Apella, E. & Pressman, D. (1975) *Fed. Proc.* **34**, 553 (bs. 1881).
16. Nathenson, S. G. & Cullen, S. E. (1974) *Biochim. Biophys. Acta* **344**, 1–25.
17. Brown, J. L., Kato, K. Silver, J. & Nathenson, S. G. (1974) *Biochemistry* **13**, 3174–3178.
18. Schwartz, B. D. & Nathenson, S. G. (1971) *J. Immunol.* **107**, 1363–1367.
19. Schwartz, B. D. & Nathenson, S. G. (1971) *Transplant. Proc.* **3**, 180–182.
20. Shimada, A. & Nathenson, S. G. (1969) *Biochemistry* **8**, 4048–4062.
21. Maizel, J. V., Jr. (1971) *Methods Virol.* **5**, 179–246.
22. Brauer, A. W., Margolies, M. N. & Haber, E. (1975) *Biochemistry* **14**, 3029–3035.
23. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91.
24. Summers, M. R., Smythers, G. W. & Oroszlan, S. (1973) *Anal. Biochem.* **53**, 624–628.
25. McKean, D. J., Peters, E. H., Waldby, J. I. & Smithies, O. (1974) *Biochemistry* **13**, 3048–3051.
26. Schechter, I., McKean, D. J., Guyer, R., Terry, W. (1975) *Science* **188**, 160–162.
27. Silver, J. & Hood, L. (1975) *Nature* **256**, 63–64.
28. Yamane, K., Shimada, A. & Nathenson, S. G. (1972) *Biochemistry* **11**, 2398–2402.