Proc. Natl. Acad. Sci. USA
Vol. 75, No. 3, pp. 1404–1408, March 1978
Cell Biology

Amino- and carboxyl-terminal amino acid sequences of proteins coded by *gag* gene of murine leukemia virus

(Rauscher leukemia virus/AKR leukemia virus/type C virus/structural proteins)

STEPHEN OROSZLAN*†, LOUIS E. HENDERSON*, JOHN R. STEPHENSON‡, TERRY D. COPELAND*, CEDRIC W. LONG*, JAMES N. IHLE§, AND RAYMOND V. GILDEN*

* Viral Oncology Program, Frederick Cancer Research Center, Frederick, Maryland 21701; ‡ Laboratory of RNA Tumor Viruses, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20014; and § Cancer Biology Program, Frederick Cancer Research Center, Frederick, Maryland 21701

Communicated by David Baltimore, December 21, 1977

ABSTRACT **The amino- and carboxyl-terminal amino acid sequences of proteins (p10, p12, p15, and p30) coded by the *gag* gene of Rauscher and AKR murine leukemia viruses were determined. Among these proteins, p15 from both viruses appears to have a blocked amino end. Proline was found to be the common NH_2 terminus of both p30s and both p12s, and alanine of both p10s. The amino-terminal sequences of p30s are identical, as are those of p10s, while the p12 sequences are clearly distinctive but also show substantial homology. The carboxyl-terminal amino acids of both viral p30s and p12s are leucine and phenylalanine, respectively. Rauscher leukemia virus p15 has tyrosine as the carboxyl terminus while AKR virus p15 has phenylalanine in this position. The compositional and sequence data provide definite chemical criteria for the identification of analogous *gag* gene products and for the comparison of viral proteins isolated in different laboratories. On the basis of amino acid sequences and the previously proposed H-p15-p12-p30-p10-COOH peptide sequence in the precursor polyprotein, a model for cleavage sites involved in the post-translational processing of the precursor coded for by the *gag* gene is proposed.**

The genome of type C RNA tumor viruses has a capacity of coding for approximately 3×10^5 daltons of protein sequence. Virion component translational products have been identified for three viral genes designated *gag* (group-specific antigen), *env* (envelope glycoprotein), and *pol* (RNA-dependent DNA polymerase) (1). The *gag* gene of a type C virus isolate of mouse origin codes for a precursor polyprotein which is subsequently cleaved by proteolytic enzymes, giving rise to four internal structural proteins (2–7). Based on approximate molecular weights (in thousands), these antigenically distinct proteins have been designated p30, p15, p12, and p10 (8). Based on analysis of *gag* gene expression in cells nonproductively infected by various replication-defective transforming viruses (4), the gene order has been shown to be p15-p12-p30-p10. Confirmation of this proposed sequence has been obtained by immunological methods using conditional lethal mutants defective in precursor cleavage (9). An enzymatic activity present in purified virus preparations and capable of *in vitro* cleavage of the precursor has also been recently identified (10). NH_2-terminal amino acid sequences of p30 proteins of several mouse and other mammalian type C viruses have been reported (11). As a first step towards the biochemical characterization of the proteolytic activities and possible cleavage sites involved in the post-translational processing of primary *gag* gene products of type C viruses, we have now subjected each of the structural proteins coded by *gag* gene of Rauscher (R) and AKR murine leukemia virus (MuLV) isolates to amino- and carboxyl-terminal amino acid sequence analyses. The results demonstrate extensive terminal sequence homology between analogous *gag* gene products of the above two related viruses and may have implications regarding the nature of proteolytic enzymes involved in precursor processing.

MATERIALS AND METHODS

Viral Proteins. R-MuLV and AKR-MuLV structural proteins coded by *gag* used in the present study were purified by published methods as summarized in Table 1.

Determination of NH_2-Terminal Amino Acids. NH_2-terminal amino acids were analyzed by the method of Weiner *et al.* (19) without major modification, with 0.5- to 1.0-nmol protein samples as described (11). Solvent systems of heptane/butanol/formic acid (10:10:1, vol/vol) and 0.15 M aqueous ammonium hydroxide were used (20).

Amino Acid Analysis. Amino acid analysis was carried out with a Beckman-Spinco automatic amino acid analyzer model 121 adapted for the high sensitivity of model 121 M. The reaction coil temperature was maintained at 150° (Henderson, unpublished).

NH_2-Terminal Sequence Analysis. Automated Edman degradations (21) were performed with the Beckman Sequencer model 890B. A microsequencing technique (22) using ^{14}C- or ^{35}S-labeled phenylisothiocyanate was modified as described (23). The single coupling, single cleavage dimethylallylamine program provided by Beckman Instruments, Inc., was used. The phenylthiohydantoin derivatives of amino acids were identified by analyzing the organic phase of each sequenator fraction by at least two independent methods: (*a*) thin-layer chromotography on polyamide sheets (24) combined with autoradiography using Kodak X-omat RP film, and (*b*) amino acid analysis after back conversion with HI to the parent amino acid (25). Early residues were also analyzed by gas chromatography (26). Phenylthiohydantoin-arginine present in the water phase was identified on the amino acid analyzer after HI hydrolysis.

Carboxypeptidase Digestion. Carboxyl-terminal amino acids sequences were determined by digestion of viral proteins with diisopropylfluorophosphate-carboxypeptidase A and phenylmethylsulfonyl fluoride-carboxypeptidase B (both obtained from Worthington Biochemical Corp., Freehold, NJ).

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: MuLV, murine leukemia virus; R-MuLV, Rauscher MuLV.

† To whom reprint requests should be addressed.

Table 1. Summary of AKR- and R-MuLV protein preparation analyzed and of purification methods

Proteins analyzed		Method of purification
p30	(A)	Isoelectric focusing, gel filtration (12, 13)
p15	(A)	Phosphocellulose–Ultrogel–GdmCl-agarose chromatography (9, 14)
	(B)	Lithium diiodosalicylate solubilization–Sephadex G-200–DEAE-Sephadex chromatography (refs. 15 and 16)
	(C)	$LiDodSO_4$–hydroxyapatite chromatography–$LiDodSO_4$/polyacrylamide gel electrophoresis (ref. 17; Olpin and Oroszlan, unpublished)
p12	(A)	Same as for p15, procedure A
	(B)	Same as for p15, procedure B
p10	(A)	GdmCl-DNA-cellulose chromatography (18)
	(B)	Same as for p15, procedure C

GdmCl, guanidinium chloride; $LiDodSO_4$, lithium dodecyl sulfate.

Carboxypeptidase A stock solution was prepared fresh daily by adding 0.1 μl of diisopropylfluorophosphate-carboxypeptidase A suspension (about 0.5 μg of protein) to 4.0 ml of 10 mM $NaHCO_3$, pH 8.3. The stock solution was kept at 4° until use on the same day. To each of four glass tubes 0.3–0.6 nmol of viral protein in 25 μl of 10 mM $NaHCO_3$, pH 8.3, containing 0.1% sodium dodecyl sulfate was added and the mixtures were incubated at 80° for 10 min. After the samples were cooled to room temperature, 2.5–25 μl of carboxypeptidase A stock solution was added to three of the sample tubes. Digestion was at room temperature. At various time intervals the reaction was terminated by adding 300 μl of 0.2 M sodium citrate, pH 2.2. The sample in the fourth tube was the zero time blank to which 300 μl of 0.2 M sodium citrate, pH 2.2, was added before addition of 25 μl of carboxypeptidase A stock solution. The amino acids released were analyzed by applying 250 μl of the digests to the amino acid analyzer. When digestion was with the mixture of carboxypeptidases A and −B, the procedure was the same except that 2 μl of phenylmethylsulfonyl fluoride-carboxypeptidase B (about 10 μg of protein) was added to the carboxypeptidase A stock solution.

RESULTS

The purity of all the viral proteins studied was established by several independent methods, such as isoelectric focusing, sodium dodecyl sulfate/polyacrylamide gel electrophoresis, and/or competition radioimmunoassay (9, 13–15). Moreover, NH_2-terminal amino acid analyses revealed either none or a single amino acid residue for each of the purified viral proteins studied, except p10. We found that alanine was the major NH_2-terminal amino acid for both AKR and R-MuLV p10s independent of the purification method used. In addition, threonine (second residue in the sequence) and valine (third residue) were detected in decreasing amounts. The average relative yields were 70, 25, and 5%, respectively. This was due to an apparent molecular heterogeneity of p10 at the NH_2 terminus and not to contamination by extraneous proteins (unpublished data). Proline alone was found for both p12s. The NH_2-terminal end of both AKR- and R-MuLV p15 was blocked since neither dansylation nor the combined dansyl-Edman procedure revealed any free α-amino acid end group. Furthermore, Edman degradation through several steps yielded negative results. Only purified proteins satisfying the above criteria for purity were subjected to further chemical analysis.

Table 2. Amino acid composition of the structural proteins coded by the *gag* gene of R-MuLV and AKR-MuLV*

Amino acid	p10†‡		p12†‡		p15†§		p30¶	
	R-MuLV	AKR	R-MuLV	AKR	R-MuLV	AKR	R-MuLV	AKR
Lys	8	9	2	1	7	7	15	16
His	2	2	0	0	6	4	4	3
Arg	13	15	8	7	6	7	29	27
Asp	7	7	15	13	12	11	31	28
Thr	2	3	9	7	13	12	14	16
Ser	3	3	12	18	12	13	12	12
Glu	12	12	9	11	13	10	44	48
Pro	11	11	25	18	25	21	19	18
Gly	10	10	12	13	9	9	17	17
Ala	6	5	5	7	6	9	13	14
Val	3	3	6	5	14	10	8	9
Met	0	0	1	1	0	0	2	2
Ile	0	0	1	6	4	4	7	5
Leu	5	5	10	11	12	10	32	33
Tyr	1	1	1	2	2	2	7	7
Phe	0	0	1	1	5	5	6	6

* Number of residues per mole is given. The respective approximate molecular weights (in thousands) were used for calculation.
† Average values of analyses performed on individual protein preparations as listed in Table 1.
‡ Hydrolyzed for varying length of time to correct for hydrolytic losses of threonine and serine.
§ Hydrolyzed for 24 hr only.
¶ Previously published data (11).

Amino acid composition

The amino acid composition of AKR- and R-MuLV *gag* gene products is given in Table 2. R-MuLV and AKR-MuLV p10s, which represent the RNA-associated nucleoproteins of the respective viruses (27), have exceptionally high lysine and arginine content and lack both methionine and phenylalanine. While the overall composition of AKR- and R-MuLV p12s is different, neither contains histidine. The main differences appear in serine, proline, and isoleucine content. The fact that the p15s of both AKR- and R-MuLV appear to be free of methionine is consistent with previously published data showing the lack of [^{35}S]methionine incorporation into this protein (28).

In Table 3 the amino acids are divided into three classes: polar, intermediate, and nonpolar (29). The p10s have the highest percentage of polar amino acids and the p30s, which

Table 3. Polarities of AKR- and R-MuLV structural proteins

Protein	% of total amino acids*		
	Polar†	Intermediate‡	Nonpolar§
R-MuLV p10	48.2	21.7	30.1
AKR p10	50.0	22.1	27.9
R-MuLV p12	29.1	29.1	41.9
AKR p12	26.4	33.1	40.5
R-MuLV p15	26.0	28.8	45.2
AKR p15	26.1	29.9	44.0
R-MuLV p30	45.8	20.7	33.5
AKR p30	45.6	21.1	33.3

* Amino acids were divided into three classes according to Capaldi and Vanderkooi (29).
† Polar: Lys, Arg, Asp, Asn, Glu, and Gln (the number of amides was not separately determined).
‡ Intermediate: His, Ser, Thr, Gly, and Tyr.
§ Nonpolar: Ala, Val, Leu, Ile, Cys (not determined), Met, Pro, and Phe.

Table 4. NH_2-terminal amino acid sequences of AKR- and R-MuLV proteins coded by the *gag* gene

R-MuLV	AKR-MuLV	5	10	15	20
p30	p30	Pro-Leu-Arg-Leu-Gly-Gly-Asn-Gly-Gln-Leu-Gln-Tyr-Trp-Pro-Phe-Ser-Ser-Ser-Asp-Leu-			
p15	p15	Blocked			
p12		*Pro*-Thr-*Leu-Thr*-Ser-Pro-*Leu*-Asn-Thr-*Lys-Pro-Arg-Pro*-Gln-Val-Leu-*Pro*-Asp- X -Gly-			
	p12	*Pro*-Ala-*Leu-Thr*-Pro-Ser-*Leu*- * - * -*Lys-Pro-Arg-Pro*-Ser-Leu-X -*Pro*-Phe-			
p10		Ala-Thr-Val-Val- X - Gly-Gln-Arg-Gln-Asp-Arg-Gln-Gly-Gly-Glu-			
	p10	Ala-Thr-Val-Val- X - Gly-Gln-Arg-Gln-Asp-Arg-Gln-Gly-Gly-Glu-			

Italicized residues are positionally identical.

also show very similar percentage values independent of origin (AKR- or R-MuLV), are the next most polar proteins among the *gag* gene products. The calculated values for AKR- and R-MuLV p12s and p15s are very similar. Both these *gag* gene products are much less polar and contain relatively more amino acids with hydrophobic side chains than p10 or p30. The p12s are phosphorylated in various mouse type C viruses (30) and thus their polarity is greatly altered by post-translational modification by the addition of negatively charged phosphate groups. Together with this, the amino acid compositional data indicate that p15 is the least hydrophilic protein derived from the *gag* precursor.

NH_2-terminal amino acid sequence

The amino acid sequences of AKR- and R-MuLV *gag* gene products are shown in Table 4. Because of the blocked end-group, p15 is not amendable to stepwise Edman degradation. The finding that p15 is the only protein among the MuLV *gag* gene products that does not have a free amino terminus is consistent with the results of previous studies that indicate that p15 is located at the amino end of the precursor polyprotein Pr65-70 (4, 9). Each of the other *gag* gene products, p30, p12, and p10, shows unique sequences. The data for p30s taken from previous studies (11) show that the NH_2-terminal sequences of AKR- and R-MuLV p30 are identical. The p10s of these viruses also show identical initial sequences. The NH_2-terminal amino acid sequences of AKR- and R-MuLV p12s, while distinctly different, also show substantial regions of homology. It should be noted that p12 is the structural protein of mouse type C viruses that has been identified as having the strongest type-specific immunological properties among the *gag* gene products (31). In order to maintain positional homology within the p12 sequences (Table 4), it was necessary to introduce a gap in the sequence of AKR-MuLV p12 at positions 8 and 9 in front of a major homologous region consisting of Lys-Pro-Arg-Pro. In the initial sequence, 9 out of 18 amino acids appear to be positionally identical (*italicized*), representing an overall 50% homology of this region.

For those amino acids that are positionally different in the two p12s the number of base changes is as follows: a change from threonine to alanine requires two changes; from serine to proline and vice versa, one; from glutamine to serine, two; from valine to leucine, one; and from aspartic acid to phenylalanine, two. Including the two-amino-acid gap, it would require minimal change of 33% in base sequence to give complete homology in nucleic acid sequences.

Carboxyl-terminal sequence

Carboxyl-terminal amino acid sequence analysis of *gag* gene coded proteins were performed as described in *Materials and Methods*. The number of moles of amino acids released per mole of protein determined at different time intervals is given in Table 5 for both R-MuLV and AKR-MuLV p15 and p12. Although kinetic data for p30 are not available, digestion with carboxypeptidase A for 1 hr liberated 1 mol of leucine, 0.5 mol of lysine, and 0.3 mol of serine or amide per mole of protein. Digestion with a mixture of carboxypeptidases A and B in addition released arginine. The most probable carboxyl-terminal

Table 5. Amino acids liberated* by carboxypeptidase digestion of R-MuLV and AKR-MuLV p12 and p15

Time, min	p15		p12	
	R-MuLV	AKR	R-MuLV	AKR
0	0	0	0	0
1	Tyr 0.27	—	—	Phe 0.30
2	—	Phe 0.46	Phe 0.67	—
		Leu 0.22	Ser, Gln, or Asn 0.16†	
3	Tyr 0.67	—	—	—
	Leu 0.09			
5	Tyr 0.79	Phe 0.87	—	Phe 0.69
	Leu 0.20	Leu 0.42	—	Ala 0.47
				Ser, Gln, or Asn 0.29†
10	—	Phe 0.96	—	—
		Leu 0.55		
15	Tyr 1.07	—	Phe 0.88	—
	Leu 0.72		Ser, Gln, or Asn 0.55†	
			Ala 0.10	
60	—	—	Phe 1.10	
			Ser, Gln, or Asn 0.72†	—
			Ala 0.63	

† Although the amino acid analyzer program used in these experiments could not distinguish between Ser, Gln, and Asn, the color constants for these amino acids were essentially identical.

* Moles of amino acids per mole of protein.

Table 6. Carboxyl-terminal amino acid sequences of R-MuLV and AKR-MuLV *gag* gene products

Protein		
R-MuLV	AKR-MuLV	COOH-terminal sequence*
p30		-(Arg,X)-Lys-Leu
	p30†	-Leu
p15		-Leu-Tyr
	p15	-Leu-Phe
p12		-Ala- X -Phe
	p12	-(Ala, X)-Phe

* In the sequences, X indicates the presence of one of three possible amino acid residues, i.e., Ser, Asn, or Gln. See footnote to Table 5.
† Previously published result (13)

amino acid sequences of p30, p15, and p12 of both R-MuLV and AKR-MuLV are given in Table 6.

DISCUSSION

The determination of amino acid sequences of oncornavirus structural proteins is of major importance to assure adequate definition for each constituent viral polypeptide. Their differentiation and unambiguous characterization must ultimately be based upon chemical analyses in addition to immunological and functional properties. Immunologically homologous and, presumably, functionally similar proteins of different mammalian type C viruses have been shown to have different molecular weights (32), probably due to shifts in cleavage sites on the precursor polyprotein. To date, immunological methods providing an indirect measure of the extent of amino acid sequence homology have proven useful in determining inter- and intraviral relationships among the viral components (33, 34). Primary structure analyses of p30s (11) from diverse mammalian type C viruses supported the immunological findings and provided definitive evidence for extensive amino acid sequence homology, indicating close phylogenetic relatedness among these viruses.

The amino-terminal sequence data derived in the present study for the remaining *gag*-coded proteins are also in general agreement with the relatedness estimate based on immunologic analysis. In addition, these findings provide definitive criteria for comparison of viral proteins isolated in different laboratories. For example, regardless of nomenclature, it is clear that R-MuLV and AKR-MuLV contain an homologous protein, designated p12, which exhibits type-specificity in immunoassays. A recent report from one laboratory (35) claiming the homologue of Gross p12 to be p15 in R-MuLV is difficult to reconcile in light of the present findings. Differences in acrylamide gel systems, presumably accounting for such discrepancies, are eliminated by definitions of proteins in terms of their sequence. Short sequences at the amino terminus may suffice for this purpose. Previous studies with p12 have demonstrated a high specificity in ability to bind to homologous viral RNA (36). The sequence differences shown here between two strains of MuLV could account for this specificity with the assumption that the nucleic acid sequence(s) involved in binding p12 show the same degree of individuality as the amino acid sequences of p12. The apparently complete identity of the NH_2-terminal sequence of the p10s of AKR- and R-MuLV is consistent with previous results indicating a lack of type-specific immunologic reactivity (14, 18).

Carboxypeptidase analyses indicate that the individual *gag* gene products, p15, p12, and p30, have unique carboxyl-terminal amino acid sequences. These findings combined with the amino-terminal sequence data and the reported peptide order

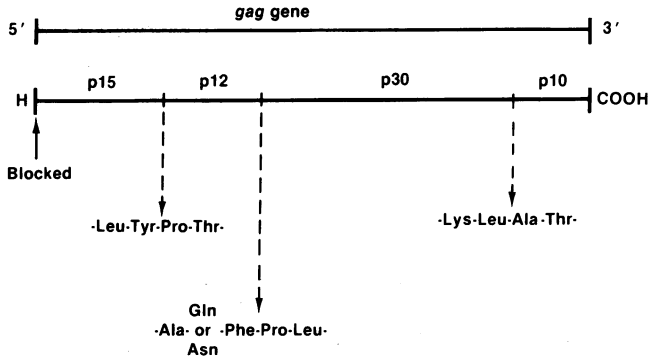

FIG. 1. Proposed order of *gag* gene products in R-MuLV precursor polyprotein and possible cleavage sites.

in Pr65-70 (4, 9) permit us to postulate possible proteolytic cleavage sites for the post-translational processing of this precursor (Fig. 1). This model is based on the assumption that proteolytic cleavage of the polyprotein between the carboxyl terminus of p15 and and the NH_2 terminus of p10 occurs only at three sites to yield the four final *gag* gene products, p15, p12, p30, and p10, and that excision of any amino acid residues does not occur during processing. It must be further assumed that the cleaved gene products are not altered by proteolytic enzymes before or during their purification. However, our data for at least one protein (p10) indicated the presence of molecules with one or two of the amino acid residues missing at the NH_2 terminus, suggesting, among other possibilities, either degradation by amino peptidases or a degenerate cleavage site between p30 and p10.

The peptide bond hydrolyzed between R-MuLV p15 and p12 is tyrosylproline, and between p12 and p30 is phenylalanylproline, according to the scheme proposed in Fig. 1. In AKR-MuLV, both cleavage sites appear to be phenylalanylproline bonds. Bachrach (37) has recently postulated that post-translational cleavages of picornavirus polyproteins occur most frequently at protease-sensitive β-bonds involving the carbonyl function of an amino acid to the left of a helix-breaker amino acid. While this may not necessarily be a generality, we do note that our results show that proline is the common NH_2 terminus of p12 and p30. Thus, one could assume that a single protease with specificity for the carbonyl function of aromatic amino acids adjacent to proline could be responsible for cleavages at the above two sites. However, tyrosylproline and phenylalanylproline bonds are resistant to most of the known proteases (38). This could suggest the involvement of a novel enzyme for generation of this cleavage. Another type of cleavage site, the leucylalanine bond, appears between p30 and p10. It should be noted that α-chymotrypsin has a high specificity for peptide bonds involving the carbonyl function of leucine (38). A similar cleavage site yielding carboxyl-terminal leucine has been postulated in the processing of picornavirus polyproteins (39).

Recently, Yoshinaka and Luftig reported the presence in purified R-MuLV preparations of a proteolytic factor capable of cleaving Pr65-70 to the constituent polypeptides in two steps (10, 40). Their results are consistent with our suggestion, based on sequence data, that proteases with at least two different specificities for cleavage sites may be involved in the post-translational processing of the precursor polyprotein coded by the *gag* gene.

This work was supported by the Virus Cancer Program, Contract N01-CO-25423, National Cancer Institute, National Institutes of Health.

1. Baltimore, D. (1974) *Cold Spring Harbor Symp. Quant. Biol.* **39,** 1187–1200.
2. Jamjoon, G., Karshin, W. L., Naso, R. B., Arcement, L. T. & Arlinghaus, R. B. (1975) *Virology* **68,** 135–145.
3. Naso, R. B., Arcement, L. J. & Arlinghaus, R. B. (1975) *Cell* **4,** 31–35.
4. Barbacid, M., Stephenson, T. R. & Aaronson, S. A. (1976) *Nature* **262,** 554–559.
5. Oroszlan, S., Long, C. W. & Gilden, R. V. (1976) *Virology* **72,** 523–526.
6. Shapiro, S. Z., Strand, M. & August, J. T. (1976) *J. Mol. Biol.* **107,** 459–477.
7. Van Zaane, D., Dekker-Michielsen, M. J. A. & Bloemers, H. P. J. (1976) *Virology* **75,** 113–129.
8. August, J. T., Bolognesi, D. P., Fleissner, F., Gilden, R. V. & Nowinski, R. C. (1974) *Virology* **60,** 595–600.
9. Reynolds, R. K. & Stephenson, J. R. (1977) *Virology* **81,** 328–340.
10. Yoshinaka, Y. & Luftig, R. B. (1977) *Biochem. Biophys. Res Commun.* **76,** 65–63.
11. Oroszlan, S., Copeland, T., Summers, M. R., Smythers, G. & Gilden, R. V. (1975) *J. Biol. Chem.* **250,** 6232–6239.
12. Oroszlan, S., Fisher, C. L., Stanley, T. B. & Gilden, R. V. (1970) *J. Gen. Virol.* **8,** 1–10.
13. Oroszlan, S., Summers, M. R., Foreman, C. & Gilden, R. V. (1974) *J. Virol.* **14,** 1559–1574.
14. Barbacid, M., Stephenson, J. R. & Aaronson, S. A. (1976) *J. Biol. Chem.* **251,** 4859–4866.
15. Ihle, J. N., Denny, T. & Bolognesi, D. P. (1976) *J. Virol.* **17,** 727.
16. Ihle, J. N., Joseph, D. R. & Pazmino, N. H. (1976) *J. Exp. Med.* **144,** 1406–1423.
17. Moss, B. & Rosenblum, E. N. (1972) *J. Biol. Chem.* **247,** 5194–5198.
18. Long, C. W., Berzinski, R. & Gilden, R. V. (1977) *Int. J. Cancer* **19,** 843–850.
19. Weiner, A. M., Platt, T. & Weber, K. (1972) *J. Biol. Chem.* **247,** 3242–3251.
20. Mihalyi, E. & Albert, A. (1971) *Biochemistry* **10,** 237–242.
21. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1,** 80–91.
22. Oroszlan, S., Copeland, T., Summers, M. & Smythers, G. (1975) in *Solid Phase Methods in Protein Sequence Analysis,* ed. Laursen, R. A. (Pierce Chemical Co., Rockford, IL), pp. 179–192.
23. Oroszlan, S., Copeland, G., Smythers, G., Summers, M. R. & Gilden, R. V. (1977) *Virology* **77,** 413–417.
24. Summers, M. R., Smythers, G. & Oroszlan, S. (1973) *Anal. Biochem.* **53,** 624–628.
25. Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G. & Ballantyne, D. L. (1971) *Biochemistry* **10,** 4912–2921.
26. Pisano, J. J., Bonzert, T. J. & Brewer, H. B., Jr. (1972) *Anal. Biochem.* **45,** 43–59.
27. Fleissner, E. & Tress, E. (1973) *J. Virol.* **12,** 1612–1615.
28. Arlinghaus, R. B., Naso, R. B., Jamjoon, G. A., Arcement, L. J. & Karshin, W. L. (1976) in *Animal Virology,* eds. Baltimore, D., Huang, A. J. & Fox, C. F. (Academic, New York), pp. 689–716.
29. Capaldi, R. A. & Vanderkooi, G. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 930–932.
30. Pal, B. K., McAllister, R. M., Gardner, M. B. & Roy-Burman, P. (1975) *J. Virol.* **16,** 123–131.
31. Stephenson, J. R., Anderson, G. R., Tronick, S. R. & Aaronson, S. A. (1974) *Cell* **2,** 87–94.
32. Stephenson, J. R., Reynolds, R. K. & Aaronson, S. A. (1976) *J. Virol.* **17,** 374–384.
33. Gilden, R. V. (1975) *Adv. Cancer Res.* **22,** 157–202.
34. Aaronson, S. A. & Stephenson, J. R. (1976) *Biochim. Biophys. Acta* **458,** 323–354.
35. Strand, M. & August, J. T. (1977) *Virology* **79,** 129–143.
36. Sen, A., Sherr, C. J. & Todaro, G. J. (1976) *Cell* **7,** 21–32.
37. Bachrach, H. L. (1977) in *Beltsville Symposia in Agricultural Research. I. Virology in Agriculture,* eds. Romberger, J. A. (Allanheld, Osum & Co., Montclair, NJ), pp. 3–32.
38. Hill, R. L. (1965) *Adv. Protein Chem.* **20,** 37–107.
39. Ziola, B. R. & Scraba, D. G. (1976) *Virology* **71,** 111–121.
40. Yoshinaka, Y. & Luftig, R. B. (1977) *Biochem. Biophys. Res. Commun.* **79,** 319–325.