Murine Type-C Virus Group-Specific Antigens: Interstrain Immunochemical, Biophysical, and Amino Acid Sequence Differences

STEPHEN OROSZLAN, MURRAY R. SUMMERS, CAROL FOREMAN, AND RAYMOND V. GILDEN

Flow Laboratories, Inc., Rockville, Maryland 20852

Received for publication 28 June 1974

The 30,000-molecular-weight internal protein, p30, was purified from seven strains of mouse type-C viruses. The individual p30's showed variation in isoelectric points and also intrastrain heterogeneity. The individual p30's could be distinguished by peptide map and quantitative complement fixation techniques with relatedness estimates of >95%. Amino terminal sequence analysis showed variability at position 4 for several p30's with complete homology otherwise through 24 residues. The intrastrain heterogeneity in p30 isoelectric points could not be explained by common contaminants, as shown by peptide mapping, and is more likely based on post-transcriptional modifications. These data provide a chemical basis for the recently described type-specific immunological properties of individual p30's.

The 30,000-molecular-weight internal protein, the major group-specific (gs) antigen or p30, of murine type-C viruses is an antigenic mosaic with determinants restricted to mouse viruses (species-specific) and determinants shared by other mammalian type-C viruses (interspecies) (13, 33, 35, 40; J. T. August, D. P. Bolognesi, E. Fleissner, R. V. Gilden, and R. C. Nowinski, Virology, in press). Current evidence indicates a multiplicity of determinants in each category (9; J. Davis, H. P. Charman, S. Oroszlan, and R. V. Gilden, Intervirology, in press). More recently, type-specific determinants have also been detected by use of radioimmunoassay (40) and quantitative C' fixation (R. V. Gilden, S. Oroszlan, and M. Hatanaka, In K. Maramorosch and E. Kurstak (ed.), Virus evolution and cancer, in press). These various determinants have value for viral genetic studies and studies of distribution of viral proteins in the absence of infectious virus (1, 17, 34, 37). In support of the immunological data showing type specificity of murine p30's, we now report differences in primary structure obtained using peptide map and amino acid sequencing techniques. Evidence for heterogeneity and strain differences in the isoelectric point (pI) of murine p30's is also described.

MATERIALS AND METHODS

Murine leukemia virus (MuLV). Several strains of mouse leukemia virus were used in these studies. Rauscher virus (R-MuLV) was grown in monolayer cultures of chronically infected mouse (BALB/c) bone marrow, JLS-V9 cells (50). The virus-shedding cell line was obtained from Electro-Nucleonics, Bethesda, Md. AKR virus was obtained from NIH Swiss embryo cells infected with an isolate from the spleen of an AKR mouse. These cells were kindly provided by Janet Hartley (National Cancer Institute). The La Puente isolate (1504E) of the wild mouse leukemia virus (WMLV) was obtained from Earle Officer (University of Southern California, Los Angeles, Calif.) (31; M. B. Gardner, J. E. Officer, R. W. Rongey, H. P. Charman, J. W. Hartley, J. D. Estes, and R. J. Huebner, Bibl. Haematol., in press) and was grown in monolayer wild mouse embryo cultures. New Zealand black (NZB) mouse type-C virus was grown in suspension cultures established from a fibrosarcoma, SCRF 60A, at Scripps Clinic, La Jolla, Calif. (18). AT-124 mouse virus grown in human rhabdomyosarcoma (RD) cells was propagated as described by Todaro et al. (44). Moloney mouse sarcoma virus (M-MSV) was produced in a rat tissue culture cell line, 78A1 (4), obtained through the courtesy of Maurice Green (Institute of Molecular Virology, St. Louis, Mo.). The wild mouse cell line shedding Kaplan radiation leukemia virus (Rad LV) derived from an X-irradiated C4Bl mouse (19) was obtained from Janet Hartley.

Rat type-C virus. Rat type-C virus, a pseudotype of murine sarcoma virus, M-MSV(RaLV), was obtained from the MSB-1 cell line (43) derived from a tumor induced by M-MSV in a female rat of the Brown-Norway strain, as previously described and characterized (28).

Feline type-C viruses. Feline leukemia virus (FeLV, Theilen strain) was obtained from a chronically infected cat lymphocytic cell suspension culture (42). RD 114, an endogenous cat virus, was the isolate.
from the RD rhabdomyosarcoma cell line of McAllister et al. (21).

All monolayer cultures were supplemented by 10% fetal bovine serum, and suspension cultures were supplemented by 20% fetal bovine serum.

**Virus purification.** Tissue culture fluids were clarified by filtration through a membrane filter (Millipore Corp., 1.2 μm pore size). The virus was concentrated by continuous-flow centrifugation with isopycnic banding in Tris-buffered (0.01 M, pH 7.4) sucrose gradients (20 to 50%, wt/wt). Fluid volumes of 20 to 30 liters were collected at flow rates of 4.0 to 4.5 liters/h with a CF-32 continuous-flow rotor operated by an L-350 ultracentrifuge (Spinco). Larger volumes (50 to 100 liters) were collected at flow rates of 13 liters/h on the K-6 continuous-flow rotor of the model K Mark II ultracentrifuge (Electro-Nucleonics). The concentrated virus was thereafter diluted with TNE buffer (0.01 M Tris-hydrochloride, 0.1 M NaCl, 0.001 M EDTA, pH 7.4) and reband a second time using either a Ti-14, Ti-15, or JCF-Z zonal rotor (Spinco) operating for 1, 2, or 3 h, respectively, at maximum speed.

Virus bands were localized by monitoring the absorbance at 280 nm and by the microcomplement-fixation (CF) tests for the major gs antigen, p30 (31). Sucrose gradient-purified virus used for the purification of viral proteins was diluted with TNE buffer and pelleted to the bottom of the centrifuge tubes.

**Disruption of virus and purification of the major gs antigen (p30) by isoelectric focusing:** procedure A. Purified virus was disrupted with Tween 80-ether and exhaustively extracted with the organic solvent by previously described procedures (31–33). Isoelectric focusing was carried out in 1 to 2% ampholine solution having a pH range of 5 to 8, in a sucrose gradient of 0 to 40% as described previously. When protein inputs were higher than 10 mg per column (LKB 8101, 110 ml), 2% ampholine was used.

**Procedure B.** Purified virus pellets were resuspended in 0.01 M Tris-acetate buffer, pH 7.8, containing 0.1 M NaCl. The protein concentration was determined, and sodium dodecyl sulfate (SDS, Seqeuental grade, Pierce Chemical Co., Rockford, Ill.) was added at a fivefold weight excess. After addition of 2-mercaptoethanol (2-ME, Bio-Rad Labs, Richmond, Calif.) to a final concentration of 0.01 M, the mixture was kept at room temperature for 1 h. The solution was then made 6 M in deionized and recrystallized urea and incubated at room temperature for 30 min. SDS was removed on a Dowex AG-1X2 (Bio-Rad) anion exchange column by the method of Weber and Kuter (46). In agreement with these authors, we found that S-labeled SDS was completely removed by this procedure. The SDS- and RNA-free (the viral RNA is also strongly bound by the anion exchange resin) viral protein solution was analyzed chemically for total protein content and serologically for gs antigen after the removal of urea by dialysis. Essentially complete recovery of the viral proteins was obtained. SDS-polyacrylamide gel electrophoresis patterns of the disrupted viral proteins before and after removal of detergent by the Dowex resin were identical after staining with Coomassie brilliant blue (Fig. 1). The fact that the strong anionic detergent could be completely removed by the above ion exchange procedure made possible the use of the isoelectric focusing technique for the purification of gs antigen after complete denaturation of viral proteins. After these denatured RNA-free viral protein preparations were electrofocused in the presence of 2% ampholine, with a pH range 5 to 8, and a 0 to 40% sucrose gradient usually for 64 h at 700 V and 6 C. In these electrofocusing runs a strong precipitate was formed, usually appearing in three distinct precipitation rings in the acidic region (below pH 4.5) of the pH gradient. Therefore, instead of emptying the column through the bottom outlet tubing, another fractionation method was used after the completion of the run. A piece of narrow Tygon tubing was inserted carefully via the upper nipple of the column-filling compartment with its end as close to the upper precipitin ring as possible, with care taken not to disturb the precipitate. It was secured in this position, and the gradient was pumped out with a peristaltic pump with a slow, constant flow rate. The effluent was monitored at 280 nm with an ISCO UA-2 ultraviolet analyzer, and 1.5- to 2-ml fractions were collected as usual.

**Procedure C.** Virus was disrupted as in procedure B but electrofocused was carried out in 1 to 6 M urea gradient instead of sucrose. Other conditions for the electrofocusing run and method of fractionation were the same as in procedure B.

All gs antigens purified by electrofocusing were chromatographed on Bio-Gel P-10 or Bio-Gel P-100 (Bio-Rad) columns equilibrated with 0.33 M ammonium acetate (30) or 0.25 M ammonium bicarbonate, and protein was recovered by lyophilization.

**Guandine-hydrochloride-agarose gel chromatography.** The previously described procedure of Fish et al. (11), as adapted by Fleisner (12) and Nowinski et al. (25), for the separation of type-C virus proteins was used. Purified virus pellets were disrupted with 8 M guandine-hydrochloride (GuHCl) (Heico, Inc., Delaware Water Gap, Pa.) and 2% 2-ME, in 0.05 M Tris-hydrochloride buffer, pH 8.5, containing 0.01 M EDTA, at 56 C for 45 min. Agarose (Bio-Gel A-5M, Bio-Rad) columns were developed with 0.02 M sodium phosphate buffer, pH 6.5, containing 6 M GuHCl and 0.01 M dithiothreitol (Calbiochem, La Jolla, Calif.). To recover proteins for chemical analysis, fractions were dialyzed exhaustively against 0.1 M NH4HCO3 and lyophilized. When necessary, p30-containing fractions were rechromatographed to obtain high purity protein preparations.

**Quantitative complement fixation.** Quantitative CF tests were performed as described (R. V. Gilden, K. Frank, M. Hanson, S. Bladen, R. Toni, and S. Oroszlan, Intervirology, in press) with adjustment of test volumes of 3.0 ml. These procedures used limited C', ~1.1 C'H4 units, and hemoglobin released by lysis is estimated at 413 nm using a spectrophotometer. Controls of antigen alone, antigen and complement, and serum and complement were employed for each test system to control for nonspecific contributions to A413 and anti-C' activity. With the dilutions of serum employed (≥1:1,000) and the low concentrations of
purified antigen (≤ 1 µg/ml) used, no corrections were necessary. A series of eight twofold dilutions of each purified p30 starting at 1 µg/ml was mixed with increasing dilutions of guinea pig antiserum prepared against each protein. At least four increasing dilutions of antiserum were tested in each case one the approximate titer (~75% maximum binding) was established. Each block of tests was performed on the same day (one serum, all antigens), and end points were repeated at least three times to insure relative reproducibility. The mixtures of antigen and antibody (0.5 ml each) were incubated with C' (1.0 ml usually of 1:300 guinea pig serum) for 18 h at 5 C. A 1-ml volume of sensitized red blood cells was added for 60 min at 37 C before centrifugation and determination of absorbance at 413 nm.

**Determination of N-terminal amino acids.** The procedure described by Weiner et al. (49) was used without major modifications for the determination of N-terminal amino acids. Purified proteins (25 to 50 µg) dissolved in 100 aliters of 0.1 M NaHCO₃ were dansylated, hydrolyzed, and prepared from chromatography according to Weiner et al. (49). Two-dimensional thin-layer chromatography was performed on double-layered Cheng-Chin polyamide sheets (Gal- lard-Schlesinger, Carle Place, N.Y.) using a solvent system of 1.5% formic acid and benzene-acetic acid (9:1).

**Determination of carboxyl terminal amino acids.** C-terminal amino acids of murine type-C virus p30’s were determined using carboxypeptidase A (2), or a combined hydrazinolysis-dansylation micro method (Summers, unpublished data), analogous to the previously described hydrazinolysis-dinitrophenol technique of Akabori et al. (23).

**Amino acid analysis.** Samples to be analyzed were placed in an ignition tube and dissolved in constant-boiling HCl containing 0.05% mercaptethanol and 0.5% phenol. The tubes were then evacuated and flushed with ultra-pure N₂ gas three times before being sealed at <500 µm pressure. Hydrolysis was effected by heating at 110 C for varying periods of time. Samples were quenched in ice to terminate hydrolysis and then taken to dryness in vacuo. For application to the Beckman 121 H amino acid analyzer, these dried samples were redissolved in pH 2.2 sodium citrate buffer. The component amino acids were eluted using the Beckman single-column methodology, and peak areas were determined electronically with an Autolabs System AA computing integrator. For each of the proteins analyzed, the reproducibility of the values for the stable amino acids obtained with this technique was 96% or better based on standard deviation of the mean. The number of residues per mole was calculated by dividing the total recovery from the analyzer in nanograms by the known gram molecular weight of the protein x 10⁻³ to obtain the nanomolar input. The nanomolar value for each residue is then divided by the total protein input in nanomoles to give the number of residues per mole.

**Peptide mapping.** Between 100 and 500 µg of protein was digested overnight with 2% by weight of TPCK trypsin (Worthington Biochemical Corp., Freehold, N.J.). The digestion was stopped by drying the mixture in vacuo and then redissolving it in a small volume of chromatographic solvent n-butanol: pyridine:acetic acid: H₂O (90:60:18:72, vol/vol/vol/vol). No less than 100 µg of protein was then applied in a compact spot (by repeated application and cold air drying) to a 100-µm layer of micro-crystalline (E. Merck, Darmstadt, W. Germany) or Macherey-Nagel (Cel 300-10 Macherey-Nagel, Duren, W. Germany) cellulose on a glass plate (20 by 20 cm). The plate was then allowed to equilibrate with chromatographic solvent overnight in a closed chamber. After a 5-h ascending development at constant temperature, the plate was air-dried for 1 h and then subjected to electrophoresis at right angles to the direction of chromatographic separation in a pH 3.6 pyridine-acetic acid buffer at 900 V. After oven drying, the peptide spots were visualized with a collidine-buffered ninhydrin spray (M. R. Summers, R. V. Gilden, and S. Oroszlan, Fed. Proc. 33:1564, 1974).

**SDS-polyacrylamide gel electrophoresis (SDS-PAGE).** Viral protein mixtures or purified proteins were subjected to electrophoresis in 10% acrylamide gels as previously described (47). The gels were stained with Coomassie blue by the method of Fairbanks et al. (10).

**Immunodiffusion.** Double immunodiffusion was done in plates that contained 0.8% agarose, pH 7.4, ionic strength 0.15. In preparing plates, 8 g of agarose, 9.3 g of Tris 2-amino-2-(hydroxymethyl) 1,3-propanediol, 74 ml of 1 N HCl, and 7.0 g of NaCl were made up to 1 liter with distilled water (8). Merthiolate (1:10,000) was added as a preservative. Plates were kept at room temperature, observed for 72 h, and photographed when optimum precipitation lines developed.

**Protein determination.** Protein was determined by the modified Lowry method (20) using crystalline bovine serum solution as standard or by the recently described method of Schaffner and Weissman (39), based on colorimetric determination of bound dye Amidoschwarz 10B (E. Merck), using lypophilized p30 protein dried to constant weight as standard.

**Antiserum.** The preparation of all antiserum used in these studies was described in previous publications (13, 29, 31–33).

### RESULTS

**Isoelectric heterogeneity of the major gs antigen (p30).** Isoelectric focusing (49), with its superb resolving power, has proved to be a valuable method for the purification of gs antigens. Our initial experiments carried out with a relatively low amount of input mouse virus (R-MuLV and AKR) proteins (31, 32) showed no overt heterogeneity of the gs protein; however, isoelectric heterogeneity has been consistently detected over a period of several years with larger viral protein input.

Figure 2 illustrates an isoelectric focusing profile of R-MuLV proteins disrupted with Tween-80 ether (Twee 80/protein = 2.5 wt/ wt). After removal of all the ether-extractable
lipids and detergent, the isoelectric focusing was carried out in a pH gradient of 5 to 8 with 2% ampholine. The CF data show clearly a major peak at pH 6.7 and a minor peak at pH 6.2 superimposable on the optical density peaks in those pH ranges. Similar data were obtained by screening the isoelectric focusing fractions in immunodiffusion (Fig. 3). The strongest precipitin lines were obtained with fractions 32, 33, and 26, respectively. In this figure, the antigenic identity of the two distinct isoelectric molecular species (pI 6.7, fraction 33, and pI 6.2, fraction 26) is also shown with a monospecific guinea pig serum and a polyspecific rat serum (reacts with other viral components, p30, and p15 in immunodiffusion) prepared against M-MSV-induced rat tumor homogenates. In this particular experiment, a total of 18.8 mg of viral protein was used and approximately 25% of that was recovered in the pH 6.7 and 6.2 peak fractions. In addition, the quantitative data obtained by determination of protein after trichloroacetic acid precipitation from the ampholine-sucrose solution, or by estimation from the absorbance at 280 nm and also by CF titers, indicated that approximately 10 to 12% of the total p30 recovered in the CF-positive fractions was focused at the lower pH value. Similar isoelectric heterogeneity could be demonstrated with other MuLV p30's. These results were also obtained when completely denatured (SDS-urea-2-ME, see Materials and Methods, procedures B and C) virus protein mixtures were electrofocused.

A comparison of the isoelectric focusing patterns of R-MuLV and WMLV can be seen in Fig. 4. In the case of R-MuLV, the major isoelectric protein species focused again at pH 6.7, and the minor focused at pH 6.2 (upper graph). With WMLV, one of the two gs (p30) reactive peaks was found again at pH 6.2, whereas the larger peak focused at pH 7.1. Serological tests with specific anti-p30 sera, both in CF and immunodiffusion, indicated a similar resolution superimposable on the absorbance tracing. With a total protein input of 11.4 mg of R-MuLV and 13.8 mg of WMLV, the quantitative determinations data and CF tests indicated a recovery of about 25 to 30% of total protein and CF units in these two peaks in each case. However, the p30 protein in the pH 6.2 minor peak was found to be approximately 10% of the total recovered R-MuLV and 30 to 33% for WMLV p30. Subsequent analysis in SDS-PAGE confirmed the presence of a main protein component of identical size, with a molecular weight of about 30,000 (p30) in the peak fractions and having different isoelectric points both for R-MuLV and WMLV. Although R-MuLV pl 6.7 and 6.2 and WMLV pl 7.1 fractions were homogeneous to the extent that p30 was the only detectable protein in the gel, WMLV pl 6.2 fraction contained, besides the main p30, three other minor components (Fig. 5). To ascertain whether the altered I was due to complex formation between p30 and the other protein components, WMLV pl 6.2 peak was re-electrofocused in the presence of deionized urea utilizing a 1 to 6 M urea gradient instead of sucrose to prevent reassociation of the protein components. The serologically active gs p30 antigen re-electrofocused to the same pH, thus excluding the possibility of the involvement in charge heterogeneity of protein-protein interactions through binding forces sensitive to urea. The apparent pH of ampholine solutions in the presence of urea is higher than the actual value (38), and therefore appropriate corrections were made when pH of urea-containing fractions was determined. The urea concentration at this peak (pH 6.2) was 4.0 to 4.3 M.

It was of interest to serologically characterize gs antigens with different isoelectric points. The results of immunodiffusion tests of WMLV pl 6.2 and pl 7.1 components are shown in Fig. 6. In these tests both species-specific and interspecies antisera were included. From the results,
Immunoprecipitin reactions of electrofocused fractions (Fig. 2) with a multispecific (MSV I-7, see text) rat serum and a monospecific gs guinea pig (GP anti-MuLV) serum. The identity of pI 6.2 (Fr 26 and 27) and pI 6.7 (Fr 33) components is shown in the right-hand patterns.

the following conclusions can be made. As shown previously for R-MuLV pI 6.2 and 6.7 components, WMLV pI 6.2 and 7.1 p30's are also serologically similar as demonstrated by the fusion of precipitin lines obtained with all antiserum tested (see legend to Fig. 6). It is also evident that both protein molecules carry interspecies determinants as well (see reaction with goat serum IS-8).

Interstrain isoelectric charge differences of mouse p30's. Table 1 summarizes the results obtained by determining the pI value of gs antigens (p30) of seven different strains of the murine type-C virus group. For several viral p30's, a comparison was made in three different systems, where the method of disruption or the condition for isoelectric focusing or both were varied as follows: (i) Tween-ether disruption and electrofocusing in sucrose gradient (Materials and Methods, procedure A); (ii) disruption of virus with SDS-urea-2-ME, followed by electrofocusing in sucrose gradient (procedure B), and (iii) disruption of virus as per procedure B but electrofocusing in 1 to 6 M urea gradient to maintain dissociating conditions (procedure C). Other strains were not as extensively studied and were compared primarily only in a single experimental system (i.e., procedure B).

The data permit the following main conclusions to be made. Each virus strain studied shows a certain degree of heterogeneity similar to that described above in detail for R-MuLV, AKR, and WMLV strains (e.g., a major and at least one minor component). While the pI of both major and minor components of R-MuLV and AKR strains is the same as previously reported (31), and as discussed in this report, a substantial variation occurs in the isoelectric pH of p30's according to the virus strain, with pI values ranging from 5.4 to 7.1 for the major and from 5.6 to 6.4 for the minor components. In several strains, the major isoelectric molecular species represents the more basic protein, but just the opposite is true for others where the major component is the more acidic species. The results for those systems which have been studied in the greatest detail reveal that the pI values obtained for both the major and minor
components probably represent stable physicochemical parameters which appear to be constant under the various experimental conditions and independent of variation in disruption and electrofocusing procedures. This suggests that heterogeneity and the strain differences in pI of p30's are probably not due to some kind of ligand binding, but appear to be well-defined intrinsic properties of the individual molecular species.

A major role of polynucleotide binding in the

---

**FIG. 4.** Isoelectric focusing of R-MuLV and WMLV after disruption and denaturation with SDS-urea-2ME (Materials and Methods, procedure B). Electrofocusing was carried out in 2% ampholine, pH 5 to 8, at 700 V and 6 C for 64 h. The absorbance at 280 nm was monitored as described in the legend to Fig. 2. The total protein inputs were 11.4 mg for R-MuLV and 13.8 mg for WMLV. In the acidic region (below pH 5), a heavy precipitate was formed and fractions were collected from above the precipitation zone as described in the text. The peak fractions at highly alkaline pH (between pH 9 and 10) contained a protein with a molecular weight of about 10,000 (p10).

**FIG. 5.** SDS-PAGE analysis of pI 6.2 (left) and 7.1 (right) electrofocused p30 components of WMLV (see Fig. 4). Migration from top to bottom.

TABLE 1. *Isoelectric points (pI) of MuLV gs Antigen* 

<table>
<thead>
<tr>
<th>Virus strain</th>
<th>Component</th>
<th>Tween-ether EF in sucrose</th>
<th>SDS-urea-2ME</th>
<th>EF in sucrose</th>
<th>EF in urea</th>
</tr>
</thead>
<tbody>
<tr>
<td>RLV</td>
<td>Major</td>
<td>6.7</td>
<td>6.7</td>
<td>6.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td>6.2</td>
<td>6.2</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>AKR</td>
<td>Major</td>
<td>6.7</td>
<td>6.7</td>
<td>6.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td>6.2</td>
<td>6.2</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>WMLV</td>
<td>Major</td>
<td>7.1</td>
<td>7.1</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td>6.2</td>
<td>6.2</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>NZB</td>
<td>Major</td>
<td>5.9</td>
<td>5.6, 6.9</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-MSV</td>
<td>Major</td>
<td>6.2</td>
<td>6.2</td>
<td>6.2</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td>6.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT 124a</td>
<td>Major</td>
<td>5.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K-Rad</td>
<td>Major</td>
<td>5.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minor</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* EF, Isoelectric focusing.
* Heterogeneity was also seen with this virus but the accurate pI of the minor peak(s) could not be determined due to scarcity of input material.

Interstrain pI differences and also in the intrastrain heterogeneity of p30's has been ruled out by experiments utilizing [°H]uridine-labeled virus as previously reported (31, 32). This is corroborated by the results presented here, which were obtained by electrofocusing SDS-urea-2-ME-disrupted viral protein preparations from which acidic polynucleotides were completely removed with Dowex 1 X-2 anion exchange resin (see Materials and Methods).

Quantitative C' fixation. Reciprocal tests were made with guinea pig antisera to three purified MuLV p30's, namely from the AKR, R-MuLV, and NZB-SCRF 60A strains (major isoelectric components). The p30 isolated from WMLV strain 1504E was also tested versus the three antisera; however, as yet a guinea pig antiserum to this p30 has not been prepared. In each set of assays a serum dilution could be found at which only the homologous p30 gave significant C' fixation. With each serum dilution maximum fixation was obtained with about 0.02 ng of p30/ml, with no significant difference between homologous or heterologous reactions. When the percent of fixation at the peak of the symmetrical curves was plotted versus the log of serum dilution, linearity was obtained with no difference in slope of homologous and heterologous reactions. This enables the calculation of the important parameters of the CF comparisons, namely the index of dissimilarity, and from this the immunological distance (I.D.) as previously described (36; Gilden et al., Intervirology, in press). The I.D. was converted to percentage sequence difference, with the assumption that the approximation of 5 I.D. units equals 1% sequence difference as determined for a variety of other proteins (15, 36) was valid in these comparisons. The results of these comparisons are shown in Table 2. Although each of the p30's could be distinguished, the estimated degree of sequence relatedness was >95% in each comparison. For comparative purposes, Table 2 includes data obtained with the above sera and non-MuLV p30's. No C' fixation was obtained in these assays at I.D. values of >130, thus attesting to the close relationship of the MuLV p30's.

Molecular size. We have previously noted variability (+15%) in the molecular weight of the p30 homologues in viruses isolated from different species. The resolving power of the SDS gel system employed is such that these differences are clearly resolved in co-electrophoresis experiments. Figure 7A, for example, shows an example of resolution of MuLV, FeLV (Theilen), and RD 114 homologues. That these are homologues is clearly indicated by sequence homology (~80%) as previously reported (30), and by the presence of a cross-reactive determinant detected in several immunoassays (28, 29, 33). In contrast to the ability to distinguish between species, all MuLV p30's in spite of differences in isoelectric points co-migrate in such gels; thus significant variations in size (>5%) are unlikely. This is illustrated in Fig. 7B where co-electrophoresis of R-MuLV p30 (pl 6.7) and WMLV p30 (pl 7.1) is shown. Similar results were obtained for all mouse viruses studied when p30's exhibiting different isoelectric points of the same virus strain were subjected to co-electrophoresis. Therefore, the observed isoelectric heterogeneity is also not due to substantial variation in molecular size.

End-group analysis. All the mouse virus p30's purified by isoelectric focusing (major component), or by GuHCl-agarose chromatography have proline as the N-terminal residue and leucine as the carboxyl terminal residue. An intrastrain comparison of p30's with different isoelectric points also showed proline at the N-terminal, and preliminary experiments indicated leucine as carboxyl terminal; thus, in spite of the charge heterogeneity, p30 protein molecules appear to be homogeneous by terminal amino acid analysis.

Amino acid composition. Amino acid com-
The positional data of p30's of several mouse virus strains are given in Table 3. A gross examination of the number of individual residues per molecule indicates a very similar amino acid composition. All are rich in residues with polar side chains, e.g., arginine, lysine, aspartic acid, and glutamic acid (or their amides), and poor in cysteine, methionine, histidine, and isoleucine.

**Amino acid sequence analyses.** We previously reported that the amino terminal 15 residues of AKR and R-MuLV were identical (30). We have now extended this analysis to WMuLV and NZB SCRF60, and M-MSV p30's. The results (Table 4) indicate that whereas R-MuLV and AKR contain leucine in position 4, WMuLV contains serine, the other two strains contain alanine at this position, and complete homology is maintained at other residues up to position 20 or higher. Position 4 appears to be hypervariable since, in p30's from FeLV, RD 114, RaLV, and Gibbon ape type-C viruses, dissimilar residues are found (S. Orozslan, T. Copeland, M. R. Summers, G. W. Smythers, and R. V. Gilden, J. Biol. Chem., submitted for publication). In considering minimal number of base changes involved in changes at MuLV p30 residue 4, we note that Leu-Ala and Ala-Ser require only a single change, whereas Leu-Ser requires two base changes. This may have implications for the origin of the various strains although conclusions based on one position in one protein are unwarranted.

**Peptide maps.** Tryptic digests and peptide maps of several p30's have been prepared. Without exception, the maps show from 35 to 41 ninhydrin-positive spots, a number in close agreement with tryptic cleavage of approximately 87% of the susceptible bonds based on amino acid analysis, indicating the high purity of the input material. Of these 35 to 41 spots, about three-fifths consistently give much deeper ninhydrin coloration than the rest, which are well defined but faint. Apparently the color yield of certain peptides (even present in equivalent amounts) is much less on the chromatographic plate. The range of coloration of the spots varies from blue violet through green to brown and yellow. The yellow peptide migrates identically in all of the maps and has been identified as the amino terminal tripeptide by co-migration with synthetic prolylleucylarginine. Unfortunately, these lightly shaded peptides do not photograph well and are only faintly visible in the figures shown here. In a few instances, the photographic reproductions do not disclose the fine shading differences between closely migrating peptides, and thus give the appearance of a single spot. The staining intensity of obviously identical peptides occasionally shows some variation from map to map in both duplicate and homologous protein runs.

An examination of these maps discloses the striking homology among the p30's isolated from different strains of murine leukemia-sar-
Fig. 7. Co-electrophoresis in SDS polyacrylamide gels of electrofocus-purified murine and feline gs antigens. (A) 1. R-MuLV + RD 114 + FeLV (Theilen). 2. R-MuLV (mol wt 31,000). 3. RD 114 (mol wt 29,000). 4. FeLV (mol wt 27,000). (B) 1. WMLV. 2. R-MuLV. 3. WMLV + R-MuLV. We originally reported (29) a molecular weight of 33,500 for the major gs antigen of RD 114, thus having a larger size than the mouse protein. This has been confirmed by co-electrophoresis with mouse protein. After many passages of the cell line for a period of more than a year, we now consistently find a molecular weight of 29,000 for RD 114, without apparent change in the isoelectric point. The significance of this is unknown at present.
coma viruses. In fact, there are only four or five differences per map at most. Assuming that these peptides differ by a single amino acid substitution, this indicates a sequence homology of approximately 98% (5 out of 250 total), in close agreement with the homology calculated from quantitative serological analysis.

Peptide maps of the p30's obtained from AKR, WMLV, and M-MSV, purified by gel filtration through a guanidine chloride-agarose column, are shown in Fig. 8. Each of the purified proteins can be seen to generate an easily identifiable and unique pattern of peptides in spite of the predominant homology. For example, peptide A appears in both AKR and WMLV (also in R-MuLV, Fig. 9), and never in M-MSV. It is always more intensely stained in the AKR maps than in those from WMLV. On the other hand, peptide B is never present in AKR but is always seen in the maps of other strains shown in Fig. 8, as well as Fig. 9, including the map for R-MuLV. Peptide C is present on maps from AKR, WMLV, and R-MuLV (Fig. 9), but is either missing in M-MSV or altered to give peptide D with significantly slower migration in the chromatographic direction. This type of alteration could result from substitution of a more polar for a less polar group, i.e., serine for alanine (this type of replacement can be seen at position 4 of N-terminal amino acid sequence for example), or valine for isoleucine. Peptide E is seen only in WMLV maps. M-MSV maps invariably contain peptide F which was not seen elsewhere. The pattern of the group of peptides in the lower left region of the M-MSV and WMLV maps (Fig. 8) and R-MuLV (Fig. 9) is quite similar except for the presence of peptide G in M-MSV and its absence from the others. This region is different for AKR. In all of these maps, a faintly staining yellow spot was invariably seen. Knowing that ninhydrin gives a yellow colored reaction with proline and with peptides containing an N-terminal proline, it was assumed that this spot might be the p30 N-terminal tripeptide, prolyl-leucylarginine. To verify this, synthetic prolyl-leucylarginylglutamylglycine (prepared by S. Sallay of Purdue University) was cleaved with

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>R-MuLV</th>
<th>AKR</th>
<th>WMLV</th>
<th>M-MSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Histidine</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Arginine</td>
<td>29</td>
<td>27</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>31</td>
<td>28</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Throneine</td>
<td>14</td>
<td>16</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Serine</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>44</td>
<td>48</td>
<td>47</td>
<td>48</td>
</tr>
<tr>
<td>Proline</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>17</td>
</tr>
<tr>
<td>Glycine</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Alanine</td>
<td>13</td>
<td>14</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Half cystine</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Valine</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Methionine</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Leucine</td>
<td>32</td>
<td>33</td>
<td>33</td>
<td>30</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

* Purified by GuHCl-agarose chromatography.
* A molecular weight of 30,000 was used for calculation.
* Based on a single complete analysis. A complete analysis included samples of 24, 48, and 72 h of hydrolysis, and a performic acid oxidized sample hydrolyzed for 24 h.
* Based on a duplicate complete analysis.
* Corrected for hydrolytic losses by extrapolation to zero time.
* Determined as cysteic acid. Preliminary experiments by cleavage with SH specific reagent TNB-CN indicate that there may be two cysteines per mole (Oroszlán and Copeland, unpublished observations).
* Determined as methionine-sulfone.

Table 4. Amino acid sequence differences in the N-terminal region of murine p30's.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Position</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-MuLV</td>
<td></td>
<td>P</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>G</td>
</tr>
<tr>
<td>AKR</td>
<td></td>
<td>P</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>G</td>
</tr>
<tr>
<td>WMLV</td>
<td></td>
<td>P</td>
<td>L</td>
<td>R</td>
<td>S</td>
<td>G</td>
</tr>
<tr>
<td>NZB</td>
<td></td>
<td>P</td>
<td>L</td>
<td>R</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>M-MSV</td>
<td></td>
<td>P</td>
<td>L</td>
<td>R</td>
<td>A</td>
<td>G</td>
</tr>
</tbody>
</table>

* The detailed quantitative data of amino acid sequence analyses will be published in a separate publication (Oroszlán et al., Biochemistry, submitted for publication). The one letter code used for amino acids is: A, alanine; D, aspartic acid; F, phenylalanine; G, glycine; K, lysine; L, leucine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; W, tryptophan; Y, tyrosine.
not been overspotted with the pentapeptide digest, showing that peptide I must be prolyl-leucylarginine.

Figure 9 shows a comparison of three mouse virus (AKR, WMLV, and R-MuLV) p30 antigens purified by isoelectric focusing. The map for AKR is nearly identical to that shown in Fig. 8. With the exception of a single spot (J) on the AKR map (appearing approximately in the center of the pattern), all of the well resolved peptides seen here are also seen in Fig. 8. The major difference among the maps on isoelectric-

trypsin, and the digest was then applied at the origin on plates which had already been charged with an AKR or M-MSV tryptic digest. On development, these two plates were found to contain only one new spot, peptide H. This was identified as the dipeptide glutamylglycine. The yellow spot on these two plates was considerably more intense than on plates which had

Fig. 8. Comparative peptide maps (Machery-Nagel Cel 300-10 layer) of the p30's from AKR, WMLV, and M-MSV. For each map, chromatography is the horizontal dimension and electrophoresis is the vertical dimension (with the cathode at the top). The origin of each map is in the lower left corner. Protein input varied from 4.2 nmol for AKR and 3.4 nmol for WMLV to 2.1 nmol for M-MSV. Peptides specifically discussed in the text are marked with an arrow and a letter.

Fig. 9. Comparative peptide maps (Machery-Nagel Cel 300-10 layer) of isoelectric-focused p30's from AKR, WMLV, and R-MuLV. Electrophoresis and chromatography are identical to Fig. 8. Protein input is 3.2 nmol (AKR), 3.4 nmol (WMLV), and 4.0 nmol (R-MuLV). RLV = R-MuLV.
focused protein and guanidine chloride-agarose-purified protein of AKR is the spot appearing in the lower right hand corner of the maps. The position and intensity of these peptides was highly variable in duplicate runs and so we chose not to include them in the discussion. For WMLV the p30 gave identical maps regardless of the method of preparation.

The uniformity of the overall pattern of peptides given by the p30 isolated from different mouse strains is further emphasized by a comparison of the R-MuLV map shown in Fig. 9 to those already discussed. Peptide K appears to be specific for this strain.

Some of the peptide maps were run on glass plates coated with microcrystalline cellulose instead of a Macherey-Nagel layer. In general, the separation and resolution of peptides on the latter is superior to that on the microcrystalline plates. Figure 10 shows two AKR p30 maps which were run on microcrystalline plates. In spite of the poorer resolution, the pattern is still recognizable as one belonging to AKR based on the absence of peptide B. In Fig. 10, maps of AKR p30 components with different isoelectric points are compared. That the observed isoelectric heterogeneity is not due to a gross alteration in the primary sequence is evident from these two maps.

DISCUSSION

The data presented in this paper clearly indicate that the p30 of several MuLV strains can be distinguished in a type-specific manner by appropriate immunochemical and biophysical procedures. In addition, the primary structure analyses, i.e., peptide mapping and amino acid sequence determination, fully support these findings and provide unequivocal evidence for the existence of type-specific chemical differences.

It should be emphasized that the biophysical data alone are insufficient to determine the strain classification of a particular mouse type-C virus. The p30's of several MuLV strains appear to be indistinguishable by molecular size, and no evidence of significant change in the number of residues was found by amino acid analysis. Some of the p30's have identical isoelectric points (major forms, e.g., R-MuLV and AKR, NZB, AT 124, and RadLV).

The major p30 components of the various strains, although appearing identical by immunodiffusion, can be distinguished by radioimmuno-competition assays as previously described (40), by quantitative C′F tests as shown here, and also by fine structural analysis. Literature data suggest that identity reactions in gel diffusion may be obtained with proteins showing >80% sequence homology (36). Immunological changes induced by small variations in the primary structure of related proteins, including single residue substitutions, have been previously observed in several other cases (7, 24) using sensitive serological techniques. A single amino acid substitution may contribute to serological specificity by directly reacting with the antibody (thus being an intrinsic part of an antigenic determinant) or by influencing the structure of an antigenic determinant not including this amino acid.

Considering the intrasubstrate heterogeneity, it should be noted that the pl 6.2 form appears in many different strains; however, this clearly does not represent contribution from a virus common to all preparations, since peptide maps show virtual identity between pl 6.2 and 6.7 proteins of AKR, both quite distinct from the p30 protein of M-MSV where the main component has a pl of 6.2. No serological difference between the various molecular forms within a strain was found by immunodiffusion, but the assumption of complete identity made from gel diffusion analyses alone would not be appropriate for the same reasons as stated above. More thorough immunochemical and primary structure analyses, as done for the major isoelectric forms, will be needed to determine the exact reason for the observed heterogeneity; however, this could result from degradation or modification of the proteins during isolation, or the multiple forms could occur naturally. It is unlikely that proteolytic digestion during isolation causes the heterogeneity since the various molecular forms have identical size and both N-terminal and carboxyl-terminal amino acids remain the same.

Isoelectric heterogeneity of mouse gs proteins have been reported previously (6, 14); in these cases it was attributed to RNA or nucleotide binding. The possibility that the binding of RNA or its acidic degradation products to the gs antigen molecules was a major factor causing the interstrain pl differences and the intrasubstrate charge heterogeneity under our experimental conditions has been ruled out by studies utilizing [3H]uridine-labeled viruses and, more effectively, by electofocusing viral protein preparation from which acidic polynucleotides have been completely removed. It appears that lipid binding does not have a role either, since in separate experiments employing hydroxyapatite chromatography in the presence of SDS (22), a similar degree of heterogeneity was
Fig. 10. Comparative peptide maps (microcrystalline cellulose layer) of pI 6.7 and pI 6.2 p30 from AKR. Electrophoresis and chromatography are identical to Fig. 8. Protein inputs are 4.6 nmol for pI 6.7 and 5.0 nmol for pI 6.2.

observed with R-MuLV and WMLV proteins (Oroszlan, unpublished observations). The persistence of micro-heterogeneity of p30 molecules in the presence of urea suggests that there may be a primary structural basis (either synthetic or post-synthetic) for the charge differences. In addition to the changes in the amino acid sequence itself, examples of other chemical modifications, not genetically determined, can be mentioned: phosphorylation of serine residues by the viral enzyme protein kinase (16, 41); dephosphorylation by phosphatases, the presence of glycosidic linkages or deaminidation processes, etc.

The results of our immunochemical and primary structure analyses on the major p30's show minimal variation—perhaps 1 to 2% maximum sequence difference—with no evidence of significant change in the number of residues. This is a much closer relationship than indicated for the entire viral genome inferred from molecular hybridization experiments (3, 17). The present results indicate the need and possible reward of similar analyses of the p30 from multiple viruses obtained from individual inbred mouse strains, some of which now appear to be separable by genetic and host range manipulation (1, 37). Comparison of gs antigen detected in embryonic tissues with virion gs antigens, especially that of the xenotropic viruses, also now seem necessary to specify the origin of the embryonic antigens. This may have
relevance to the concept of tolerance based on presence of gs antigen in chicken and mouse embryos. Breaking of tolerance by related antigens is a well-known phenomenon (5, 48), and thus the gs antibody induced in chickens by hyperimmunization may result from differences in the exogenous virus gs compared to the gs of the endogenous inherited viral genome.

ACKNOWLEDGMENTS

This work was supported by Public Health Service Contract No1-CF-3-3247 with the Virus Cancer Program of the National Cancer Institute.

We wish to thank D. Bova, T. Copeland, M. Hansen, L. Masters, G. Symthers, and R. Toni for excellent technical assistance.

LITERATURE CITED

1574 OROSZLAN ET AL. J. VIROL.