Avian Oncovirus MH2: Molecular Cloning of Proviral DNA and Structural Analysis of Viral RNA and Protein

HANS W. JANSEN, TILO PATSCHINSKY, AND KLAUS BISTER*
Max-Planck-Institut für Molekulare Genetik, D-1000 Berlin 33, Germany

Received 31 March 1983/Accepted 29 June 1983

Viral RNA, molecularly cloned proviral DNA, and virus-specific protein of avian retrovirus MH2 were analyzed. The complexity and sequence conservation of the transformation-specific v-myc sequences of MH2 RNA were compared with those of the other members of the MC29 subgroup of acute leukemia viruses, MC29, CMII, and OK10, and with chicken cellular c-myc sequences. All T1 oligonucleotides mapping within the 1.3-kilobase coding region of MC29 v-myc have homologous counterparts in the RNAs of all MC29 subgroup viruses and in c-myc. These counterparts are either identical in composition or altered by single point mutations. Hence, the 47,000-dalton carboxy-terminal sequences of the transforming proteins of these viruses and of the cellular gene product are probably highly conserved but may contain single amino acid substitutions. T1 oligonucleotide mapping of MH2 RNA indicated that the MH2 v-myc sequences map close to the 3' end of viral RNA. A genomic library of an MH2-transformed quail cell line was prepared by using the Charon 4A vector system. By screening with an myc-specific probe, a clone containing the entire MH2 provirus (λMH2-1) was isolated. Digestion of cloned DNA with KpnI yielded a 5.1-kilobase fragment hybridizing to both gag- and myc-specific probes. Further restriction mapping of λMH2-1 DNA showed that about 1.6 kilobases of the gag gene are present near the 5' end of proviral DNA, and the conserved part of v-myc, i.e., 1.3 kilobases, is present near the 3' end of proviral DNA. These two domains are separated by a segment of at least 1 kilobase of different genetic origin, including additional unique sequences unrelated to virion genes. Tryptic peptide analysis of the gag-related protein of MH2, p100, revealed gag-specific peptides and several unique methionine-containing peptides. One of the latter is possibly shared with the polymerase precursor protein Pr180gag-pol, but no myc-specific peptides, defined for the MC29 protein p110gag-myc, appear to be present in MH2 p100. The data on viral RNA, proviral DNA, and protein of MH2 reveal a unique genetic structure for this virus of the MC29 subgroup and suggest that its v-myc gene is not expressed as a gag-related protein.

Avian acute leukemia viruses of the MC29 subgroup are distinguished from other oncogenic avian retroviruses by the presence of a unique class of transformation-specific (onc) sequences, termed v-myc, in their genomic RNAs (8, 9, 23). These sequences apparently are derived from a normal cellular gene, since sequences closely related to v-myc, termed c-myc, are found in the chromosomally integrated avian leukemia virus. Their oncogenic spectra in vivo largely overlap, especially in the induction of myelocytomatosis, endotheiomas, and sarcomas in infected birds (3). Characteristic of MH2 is the high incidence of hepatocarcinomas induced by this agent (1, 3). All four viruses transform avian fibroblasts and bone marrow cells in vitro (22).

Analysis of the genetic structures of the replication-defective MC29 subgroup viruses revealed that the v-myc sequences of MC29 and CMII map directly adjacent to the 3' end of partial (Δ) gag sequences and are expressed via genome-sized mRNAs as gag-myc hybrid proteins with transforming function, MC29 p110gag-myc and CMII p90gag-myc, respectively (11, 13, 26). The Δgag-myc genetic units of the genomic RNAs are flanked at their 3' ends by partial env or pol and env sequences, respectively (8, 9, 13, 31). In OK10 a complete gag gene is present, and myc is inserted between partial pol and env sequences (8, 9, 14). The gag-Δpol-myc unit is expressed as a hybrid protein, OK10 p200gag-pol-myc (35). In addition, OK10-trans-
formed cells contain a subgenomic myc-containing mRNA species which would allow the expression of v-myc independent of structural genes (17, 41).

The genetic structure of MH2 had not been directly analyzed previously. Based on $T_1$ oligonucleotide analysis of viral RNA (20), cDNA hybridization (40, 44), and analysis of the gag-related MH2-specific protein p100 (25), it was believed that the genomic organization of MH2 was similar to that of MC29. However, the following recent observations were inconsistent with this presumed analogy: (i) MH2 variants that are oncogenic but that do not synthesize p100 in transformed cells were isolated (29), (ii) a subgenomic myc-containing mRNA species of 2.6 kilobases (kb) was detected in MH2-transformed cells (33), and (iii) phosphopeptide analysis of MH2 p100 did not reveal myc-specific phosphoserine- and phosphothreonine-containing peptides, typically found in MC29 p110 gag-myv and CMII p90 gag-myv (36).

In this communication, we present analyses of MH2 RNA, molecularly cloned MH2 proviral DNA, and the gag-related protein p100. All data consistently show that the genetic structure of MH2 is different from that of MC29 and that the MH2 v-myc gene is not expressed as a gag-related protein.

**MATERIALS AND METHODS**

**Cells and viruses.** A nonproducer line, clone Q8, of MC29-transformed quail embryo fibroblasts and quail embryo fibroblasts transformed by CMII (CM-associated virus) have been described previously (11, 13). Quail embryo fibroblasts transformed by MC29 variant HBI and superinfected with nondefective ring-necked pheasant virus (RPV) have been described before (12). Cell line MH2-A10 was derived from a colony of quail embryo fibroblasts transformed by MH2 (MH-associated virus [MHAV]). MH2-B4 and MH2-B9 are quail cells transformed by MH2 (MHAV + RPV).

**Analysis of viral RNA.** Labeling of virus with $H_2^3$PO$_4$, virus purification, RNA extraction and sedimentation, and fingerprinting of RNase $T_1$-resistant oligonucleotides by two-dimensional electrophoresis and homochromatography were carried out essentially as described previously (13, 14). Viral RNA was hybridized to denatured DNA of pBR322 plasmid $pmyc$-5' or $pmyc$-3' (see below) or to denatured DNA of Bc-myc-3, a phage clone containing the chicken c-myc locus (39), under conditions described previously (12, 15). Hybridized RNA was isolated and fingerprinted as reported before (12, 15). Mapping of $T_1$ oligonucleotides was done by fingerprinting size-selected, polyadenylate-containing viral RNA fragments (8, 13, 14).

**Molecular cloning and restriction enzyme analysis of proviral DNA.** High-molecular-weight DNA of MH2-A10 cells was prepared by the following procedure: 1.7 x 10$^8$ cells were trypsinized, washed with Tris-buffered saline, and suspended in 7 ml of 1 x SSC (SSC is 150 mM NaCl plus 15 mM sodium citrate, pH 7.0). Sodium dodecyl sulfate and pronase (Calbiochem, La Jolla, Calif.) were added to final concentrations of 0.5% (wt/vol) and 250 μg/ml, respectively. After incubation at 37°C for 30 min, the mixture was extracted three times with phenol-chloroform-isooamyl alcohol (50:24:1 [vol/vol]). DNA was precipitated with ethanol, spooled onto a glass rod, and dissolved in 3.5 ml of 0.1 x SSC. A 175-μl amount of 20 x SSC was added. The mixture was incubated first with RNase A at 100 μg/ml at 37°C for 30 min and then with pronase at 250 μg/ml at 37°C for 30 min. The mixture was extracted twice with phenol-chloroform-isooamyl alcohol (50:24:1 [vol/vol]) and once with chloroform-isooamyl alcohol (24:1 [vol/vol]). DNA was precipitated with ethanol, spooled onto a glass rod, and dissolved in 3.0 ml of 10 mM Tris-hydrochloride (pH 8.0)-1 mM EDTA. The total yield was 500 μg of DNA.

MH2-A10 DNA (200 μg) was fragmented by partial digestion with 70 U of EcoRI (Bethesda Research Laboratories, Gaithersburg, Md.) in the appropriate buffer (18) for 1 h at 37°C. Fragments of 10 to 22 kb were purified by sedimentation through 10 to 40% (wt/vol) sucrose in 100 mM NaCl-10 mM Tris-hydrochloride (pH 7.4)-1 mM EDTA (30). DNA (10 μg) of a similar size class was ligated to 25 μl of EcoRI DNA which had 4A arms joined at the cohesive end site (30), using 20 U of T4 DNA ligase (Bethesda Research Laboratories) in 100 μl of ligation buffer (66 mM Tris-hydrochloride [pH 7.5], 5 mM MgCl$_2$, 5 mM dithiotreitol, 1 mM ATP). Incubation was at 12°C for 16 h. The ligated, concatameric DNA was reacted with the in vitro packaging extract prepared from *Escherichia coli* BHB2688 and BHB2690 by the method of Hohn (24) without UV irradiation of the extract. Packaged phage particles (10$^{12}$) were amplified at subconfluent densities on plates of E. coli BHB2660. Approximately 4 x 10$^7$ recombinant phages from these plate lysates were screened on six 150-mm petri dishes using the in situ plaque hybridization technique of Benton and Davis (5). The plaque from each plate were adsorbed to two duplicate nitrocellulose filters (Schleicher & Schuell, Dassel, Germany). Filters were hybridized under conditions described below with 32P-labeled $pmyc$-3' DNA. One strong signal was obtained, and the corresponding plaque area was isolated. Phages were subjected to three subsequent cycles of plating at lowered densities and screening with the myc-specific probe until a phage stock which was derived from a single myc-containing phage, termed λMH2-1, was obtained.

Bacteriophage DNA was prepared by a modified plate lysis method. λMH2-1 was grown on plates of strain B1H2600 at confluent plaque density. Phages were eluted in phage buffer (20 mM NaCl, 10 mM MgCl$_2$, 10 mM Tris-hydrochloride [pH 7.4]). 1% (vol/vol) glycerol, 0.05% [wt/vol] gelatin) at 4°C for 20 h. Bacterial debris was removed by low-speed centrifugation, and an equal volume of 20% (wt/vol) polyethylene glycol-2 M NaCl in phage buffer was added. After 1 h at 0°C, precipitated phage particles were collected by centrifugation at 10,000 x g for 20 min at 4°C. Phages were suspended in CsCl solution in 10 mM Tris-hydrochloride (pH 7.5)-10 mM MgCl$_2$-0.1 mM EDTA, and the density was adjusted to 1.7 g/ml. This mixture was overlaid with CsCl solutions of 1.6 and 1.4 g/ml. Centrifugation in an SW41 rotor was set at 25,000 rpm at 20°C for 4 h. The phage band was
collected and dialyzed against 10 mM Tris-hydrochloride (pH 7.5)–10 mM MgCl₂–0.1 mM EDTA. Phage DNA was obtained by phenol extraction and ethanol precipitation.

Restriction endonuclease digestions of phage DNA were performed in the appropriate buffers (18). The agarose gel electrophoresis procedure used has been described elsewhere (15, 39). DNA was transferred to nitrocellulose by the method described by Southern (45). To facilitate hybridization to two different probes, DNA was transferred to two nitrocellulose filters. Blotting was for 12 min for the first filter and 30 min for the second filter.

Nitrocellulose filters were prehybridized at 42°C for at least 3 h in a solution containing 50% (vol/vol) formamide, 5× SSC, 10× Denhardt solution (0.2% Ficoll, 0.2% bovine serum albumin, and 0.2% polyvinylpyrrolidone), 0.1% sodium dodecyl sulfate, and 70 μg of denatured total nucleic acids from strain BHB2600 per ml. Hybridization was for ca. 20 h at 42°C in fresh prehybridization solution plus 0.2 to 0.5 μg of denatured radioactive DNA labeled by nick translation with [α-32P]dCTP (Amersham Corp., Amersham, England) to a specific activity of 6 × 10⁶ cpm/μg (38). Filters were washed three times for 30 min at 65°C in 700 ml of 0.1× SSC–0.1% sodium dodecyl sulfate, air dried, and exposed to X-ray film XAR-5 (Eastman Kodak Co., Rochester, N.Y.) at −70°C with intensifying screens (DuPont Co., Wilmington, Del.).

The following probes were used as probes for screening and blot hybridizations. pMC29, originally provided by T. Papas, National Cancer Institute, Bethesda, Md., is a pBR322 plasmid containing a 3-kb BamHI fragment encompassing most of Δagg, all of v-myc, and part of Deuv of cloned MC29 proviral DNA (28) from the MC29-transformed nonproducer line Q5 (11). pmyc-5' and pmyc-3' are subclones in pBR322 of 1.9- (4, 34) and 1.1-kb BamHI-Sall fragments of the insert in pMC29, representing the 5' domain of v-myc and part of Δagg and the 3' domain of v-myc and part of Deuv, respectively. Their construction has been described previously (15). pgag and ppol are pAT153 plasmids containing a 1.4-kb BamHI fragment from the gag gene and a 1.4-kb EcoRI-BamHI fragment from the 5' half of the pol gene, respectively, of cloned proviral DNA of Schmidt-Ruppin A Rous sarcoma virus (19). They were prepared and kindly provided by M. Hayman, Imperial Cancer Research Fund, London, England.

Analysis of viral proteins. Q8 cells and MH2-A10 cells were seeded onto 100-mm dishes at densities of 6 × 10⁵ and 1 × 10⁶ cells per dish, respectively. After 24 h, the cells were washed once and then starved in labeling medium for 1 h at 37°C. Labeling was performed with 1.15 mCi of [35S]methionine (1.040 Ci/ mmol; Amersham Corp.) in 3 ml of labeling medium per dish for 4 h at 37°C. Labeling medium consisted of minimal essential medium containing 5% normal methionine concentration and 4% dialyzed (against saline) calf serum. Lysates from two dishes each of labeled Q8 and MH2-A10 cells were prepared by using RIPA buffer as described previously (34, 43). Immunoprecipitation was performed as described previously (34, 43). Complexes were precipitated by the addition of fixed Staphylococcus aureus, pelleted, and then centrifuged through a cushion of 10% sucrose in RIPA buffer, followed by three washes in RIPA buffer. The precipitates were then suspended in a solution containing 0.15 M NaCl and 0.01 M sodium phosphate, pH 7.2, transferred into a microcentrifuge tube, and pelleted. The gag-related proteins were isolated by using an anti-gag serum produced in rabbits against the purified, detergent-disrupted MC29 (MC-associated virus) complex. A sample of the MH2-A10 cell lysate was precipitated with a rabbit anti-reverse transcriptase serum, which was a gift from H. Oppermann, University of California, San Francisco, Calif. Immunoprecipitated proteins were separated through 1-mm-thick sodium dodecyl sulfate-polyacrylamide gels that contained 12.5% acrylamide and 0.1% bisacrylamide (34, 43). Electrophoresis was at 18 mA until the dye front entered the separation gel and was then continued at 25 mA for 4 h. Preparative gels were dried without fixing and exposed to XAR-5 film at room temperature. Analytical gels were processed for fluorography and exposed to prefogged film at −70°C (27).

Proteins to be analyzed by tryptic digestion were localized by autoradiography and extracted from homogenized gel strips in 0.05 M NH₄HCO₃–0.1% (wt/vol) sodium dodecyl sulfate–5% (vol/vol) mercaptoethanol as described previously (4, 34), except that two extractions were performed for 2 h, each at 37°C. The eluted proteins were precipitated with trichloroacetic acid in the presence of 20 μg of carrier protein, oxidized with performic acid, and digested with tolyl-sulfonl phenylalanyl chloromethyl ketone-treated trypsin (Worthington Diagnostics, Freehold, N.J.) based on a detailed description (4) with the following minor modification: samples were digested with 30 μg of trypsin for 4 h at 37°C; 20 μg of trypsin was then added, and the digestion was continued for another 4 h at 37°C. Tryptic digests were separated on cellulose thin-layer plates (E. Merck AG, Darmstadt, Germany), as described previously (4, 34), by electrophoresis for 27 min at 1.3 kV with a buffer of pH 4.7 containing n-butanol-pyridine-acetic acid-water (2:1:1:36 [vol/vol]), followed by ascending chromatography in n-butanol-pyridine-acetic acid-water (97:75:15:60 [vol/vol]). The plates were then processed for fluorography and exposed to prefogged film at −70°C (16).

RESULTS

Comparison of v-myc sequences of MC29-subgroup viruses. The combination of hybridization and fingerprinting techniques has facilitated high-resolution analysis of v-myc specific T₁ oligonucleotides of MC29 RNA (15) and OK10 RNA (14). Here we analyzed MH2 and CMII RNA sequences related to cloned MC29 v-myc DNA or chicken c-myc DNA. ³²P-labeled 70S RNA of MH2 (MHAV plus RPV) was hybridized with pmyc-5’ (Fig. 1A), pmyc-3’ (Fig. 1B), or λc-myc-3 DNA (Fig. 1C), and the hybridized RNA was fingerprinted. A total of eight myc-specific T₁ oligonucleotides were detected, and they are indicated by underlined numbers in Fig. 1. The hybrid with pmyc-5’ also contained gag-specific T₁ oligonucleotides (oligonucleotides 13 and 20a) derived from both MH2 RNA and helper viral RNAs present in excess in the 70S
FIG. 1. Fingerprint analysis of v-myc RNA sequences of MH2, CMII, and HBI-MC29. [32P]RNA of MH2 (MHAV plus RPV) was hybridized to pmyc-5 DNA (A), pmyc-3 DNA (B), or λc-myc-3 DNA (C); [32P]RNA of CMII (CM-associated virus) was hybridized to pmyc-5 DNA (D) or pmyc-3′ DNA (E); and [32P]RNA of HBI-MC29 (RPV) was hybridized to λc-myc-3 DNA (F). The following conditions were used: 0.5 × 10⁶ to 2.0 × 10⁶ cpm of 32P-labeled 70S RNA (~0.1 to 0.4 μg) was hybridized with 15 to 30 μg of denatured plasmid or phage DNA in 20 to 30 μl of a solution containing 70% (vol/vol) formamide, 0.3 M sodium chloride, 0.03 M sodium citrate, and 0.02 M sodium phosphate (pH 7.0) at 40°C for 18 to 40 h. Unhybridized RNA was digested by incubation with RNase T₁, and the hybrid was isolated by exclusion chromatography on a Bio-Gel P100 column. The hybrid was melted, and the RNA was fingerprinted as described previously (12, 15). Numbers indicating inyc oligonucleotides, as determined by RNase A digestion (Table 1), are underlined. Oligonucleotides 13, 20a, and 20b were identified as gag-related sequences (12, 13, 15); they are derived from MH2, CMII, and helper viral RNAs.

[32P]RNA complex. The composition of T₁ oligonucleotides in terms of RNase A-resistant fragments was determined, and three myc-specific oligonucleotides (oligonucleotides 6a, 120a, and 43a) of MH2 RNA were found to contain single base changes compared with their MC29 homologues (Table 1). Fingerprint analysis of CMII (CM-associated virus) RNA sequences related to pmyc-5′ and pmyc-3′ is shown in Fig. 1D and 1E. A total of 12 myc-specific oligonucleotides, including the 6 described previously for CMII (13), were found. One of them (oligonucleotide 43a) is the same modified homologue of MC29 oligonucleotide 43 found in MH2 RNA. Figure 1F shows the complete set of 13 myc-specific T₁ oligonucleotides of MC29 RNA found in a hybrid with λc-myc-3 DNA. Hence, different subsets of v-myc-specific T₁ oligonucleotides, defined originally for MC29 RNA, are found in MH2 and CMII RNAs.
TABLE 1. T1 oligonucleotides of v-myc genes

<table>
<thead>
<tr>
<th>Oligonucleotide no.</th>
<th>RNase A digestion products</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2U, 4C, G, 3AC, AU, 3AAC</td>
</tr>
<tr>
<td>3</td>
<td>16U, 3C, G, 2AAC, A4U</td>
</tr>
<tr>
<td>6</td>
<td>7C, 3AU, AG, A3U</td>
</tr>
<tr>
<td>6a</td>
<td>7C, G, 3AU, A3U</td>
</tr>
<tr>
<td>7b</td>
<td>3U, 9C, AC, AAG</td>
</tr>
<tr>
<td>8b</td>
<td>U, 3C, AC, AG, AAU, A4C</td>
</tr>
<tr>
<td>15</td>
<td>5U, 3C, AC, 2AU, AAU</td>
</tr>
<tr>
<td>26</td>
<td>5C, 2AC, AG</td>
</tr>
<tr>
<td>31a</td>
<td>U, 2AU, AG, A4U</td>
</tr>
<tr>
<td>32a</td>
<td>6U, 3C, G, 2AC</td>
</tr>
<tr>
<td>43</td>
<td>2U, 3G, AC, AAC, A3C</td>
</tr>
<tr>
<td>43a</td>
<td>2U, 3G, AC, 2A4C</td>
</tr>
<tr>
<td>44</td>
<td>2U, 2C, G, AAC, A4C</td>
</tr>
<tr>
<td>45</td>
<td>3U, C, G, 3AC, A4C</td>
</tr>
<tr>
<td>120</td>
<td>U, 4C, 2AC, AAG</td>
</tr>
<tr>
<td>120a</td>
<td>3C, 2AC, A4G</td>
</tr>
</tbody>
</table>

* Underlined numbers refer to modified T1 oligonucleotides.
* Analysis of T1 oligonucleotides by RNase A digestion has been described previously (13, 15). Comparison of the published nucleotide sequences (2) for oligonucleotides 6, 43, and 120 of MC29 (data in parentheses) shows that they can be derived from each other by single base changes (underlined bases).

Recently, the complete nucleotide sequence of the v-myc region of a proviral clone of MC29 (46) was determined (2), and a partial sequence analysis around the myc-env junction in pmyc-3' was reported (10). Based on this, all 13 myc-specific oligonucleotides and adjacent gag and env oligonucleotides (oligonucleotides 119, 7a, 14a, and 2a) of MC29 RNA can be located precisely on the map of the MC29 genome (Fig. 2) in good agreement with previous oligonucleotide mapping (12, 15). The sequence analysis of Alioto et al. (2) also showed that the largest open reading frame in v-myc terminates 303 bases upstream from the myc-env junction. Figure 2 shows the v-myc-specific oligonucleotide sets of all myc-containing viruses analyzed to date, including a transformation-defective mutant, td10H-MC29 (15), and the MC29 variant HBI (12). All T1 oligonucleotides mapping within the coding region defined for MC29 v-myc (oligonucleotides 32a, 6, 120, 26, 7b, 1b, and 43) and one mapping immediately downstream from the translational termination codon (oligonucleotide 8b) are found in the v-myc sequences of CMII, OK10, and MH2 RNAs (Fig. 1 and 2). Allelic counterparts of MC29 oligonucleotides, modified by single base changes are indicated by underlined numbers in Table 1 and italicized numbers in Fig. 2. One of them (number 43a) is found in CMII, OK10, and MH2 RNAs, and two additional ones (numbers 6a and 120a) are found in MH2 RNA. T1 oligonucleotides from the noncoding region of MC29 v-myc (oligonucleotides 31a, 15, 45, 44 and 3) are not as highly conserved in the other members of the MC29 subgroup, and the flanking gag and env oligonucleotides (oligonucleotides 119, 7a-7c, 14a, and 2-2a) are not found at all in viruses other than those derived from MC29 (Fig. 2). The presence in the chicken c-myc locus of sequences homologous to all 13 v-myc specific oligonucleotides is deduced by their recovery from a hybrid between MC29 RNA and c-myc DNA (Fig. 1F). When such a hybrid was treated with RNase A in addition to RNase T1 before isolation and melting, some v-myc T1 oligonucleotides were missing from the subsequent fingerprint of hybridized RNA, indicating RNase A-sensitive mismatches in the RNA-DNA hybrid (data not shown; 12). Hence, the corresponding sequences in c-myc, indicated by numbers in parentheses (Fig. 2), probably contain single base changes when compared with those of MC29 v-myc. For a complete comparison of all v-myc genes, the v-myc region of td10H-MC29, lacking a central segment of the coding region (15, 21), and that of HBI-MC29 are also shown in Fig. 2. HBI probably was generated by recombination of td10H-MC29 with c-myc (12, 37). Since its v-myc sequences contain all T1 oligonucleotides defined previously for wild-type MC29 (12), HBI (RPV) RNA was used as a reference v-myc source for the hybridization experiments described here.

Location of v-myc on MH2 RNA. Purification and complete T1 oligonucleotide analysis of MH2 RNA, for which a size of 5.4 to 5.7 kb has been reported (20, 33), were not observed. The large excess of 8.5-kb helper viral RNAs found in the virus complexes released from several MH2-transformed quail cells analyzed (data not shown). Nevertheless, several characteristic v-myc T1 oligonucleotides are well resolved from helper viral oligonucleotides on two-dimensional fingerprints. Hence, total viral RNA of MH2 (MHAV plus RPV) was fragmented, and size-selected, polyadenylated fragments were fingerprinted (Fig. 3). Figure 3A shows that in the 34S pool only MHAV plus RPV oligonucleotides are present, and MH2-specific oligonucleotides are first observed in the 28S pool (Fig. 3B), corresponding to the smaller MH2 genome size. MH2 oligonucleotides are numbered, and the v-myc-specific ones are underlined. Numbers in parentheses refer to env oligonucleotides of RPV RNA (12). The resolved v-myc oligonucleotides are present in the 20S pool, and some are present in the 14S pool (Fig. 3D and E). This is not observed when MC29 or CMII RNAs are analyzed in a similar way (data not shown; 8).
FIG. 2. Partial sequence comparison of v-myc genes of MC29 subgroup viruses and of the c-myc locus of chicken DNA. The locations of v-myc-specific oligonucleotides, numbered as in Fig. 1 and Table 1, on the map of MC29 RNA are shown at the top (boldface arrows). The precise map position of these oligonucleotides is based on a recent sequence analysis of MC29 v-myc (2). The relative positions of the gag translational initiation codon (position 380) and of the gag-myc junction (position 1728) are based on sequence comparison of MC29 and Rous sarcoma virus proviral DNAs (2, 42), and v-myc and env nucleotides are numbered consecutively. The oligonucleotide patterns of the v-myc genes of the other MC29 subgroup viruses are shown in the lower portion of the figure; italic numbers and numbers in parentheses indicate modified alleles of MC29 sequences (see text and Table 1).

Hence, the v-myc sequences of MH2 RNA map close to the 3' end of the genome, within ca. 2.5 kb from the 3' terminus.

**Molecular cloning and restriction mapping of MH2 proviral DNA.** A recombinant DNA library of partially EcoRI-digested DNA from the MH2-A10 cell line was prepared by using the Charon 4A vector system. A single clone (λMH2-1) which hybridized to both gag- and myc-specific probes was selected. A restriction enzyme analysis of that clone is shown in Fig. 4. After agarose gel electrophoresis (Fig. 4A), DNA was transferred to two nitrocellulose filters and hybridized to either ϕgag (Fig. 4B) or pMC29 DNA (Fig. 4C). Since ϕgag contains all the gag sequences present in pMC29, additional bands observed in Fig. 4C in comparison with those observed in Fig. 4B are derived from DNA fragments containing myc sequences. Filters were also hybridized to pmyc-3' (data not shown), and all results are summarized and interpreted in Fig. 5. Hybridization with gag- and myc-specific probes is confined to a 5.1-kb KpnI fragment. The restriction pattern to the right of the left-hand KpnI site up to the second BamHI site is identical with that reported for the Rous sarcoma virus gag gene (19, 42) and is consistent with the pattern of hybridization to the gag probe (Fig. 4B). The restriction pattern around the internal SalI site of the 5.1-kb KpnI fragment is identical to that of the v-myc region of MC29 proviral DNA (2, 21, 28, 32, 46) and is consistent with the pattern of hybridization to pMC29 (Fig. 4C) and pmyc-3' (data not shown). The location of the 3' border of the gag region, to the right of the BamHI site, is somewhat uncertain and shown as a broken line in Fig. 5, but it is clear from the restriction map that not all of gag is present. The 5' border of myc, also shown as a broken line in Fig. 5, is defined here by comparison with the restriction pattern and the sequence data of MC29 v-myc (Fig. 2). A segment of at least 1 kb (depending on the precise location of the gag border) in the center of the 5.1-kb fragment is unrelated to gag or MC29 v-myc (Fig. 5). Blot hybridization (data not shown) with ppol DNA, containing the 1.4-kb fragment from the EcoRI site at the 3' end of the gag gene to the BamHI site in the center of the pol gene of Rous sarcoma virus, showed that sequences from the 3' end of gag and the 5' half of pol are not present in λMH2-1. In summary, the restriction mapping data suggest that λMH2-1 contains the entire MH2 provirus terminating...
FIG. 3. Fingerprint analysis of size-selected, polyadenylated RNA fragments of MH2 (MHAV plus RPV) obtained by oligodeoxymethylidylic acid chromatography and sucrose gradient sedimentation of heat-denatured 70S [32P]RNA (8, 13). The approximate pool sizes are as follows: 34S (A), 28S (B), 24S (C), 20S (D), 14S (E), and 7S (F). Numbers indicate oligonucleotides of MH2 RNA; v-myc-specific oligonucleotides, identified by RNase A digestion, are underlined (Table 1). Numbers in parentheses indicate env oligonucleotides of RPV RNA.

in long terminal repeats containing KpnI sites similar to those in the long terminal repeats of MC29 proviral DNA (21, 28, 46). The observed complexity of the KpnI fragment (5.1 kb) agrees well with the size of polyadenylated MH2 RNA (5.4 to 5.7 kb), estimated from electrophoretic mobility in agarose or polyacrylamide gels (20, 33). The solid line at the bottom of Fig. 5 indicates the coding region needed for the synthesis of a 100,000-dalton protein, starting at the normal gag translational initiation codon. For further proof of the identity of cloned MH2 provirus, 32P-labeled RNA of HBI-MC29 (RPV) was hybridized to denatured λMH2-1 DNA, and the RNA present in the hybrid was fingerprinted (Fig. 6). T1 oligonucleotides were identified by their location on the fingerprint and by RNase A digestion. They are numbered accordingly, and myc-specific oligonucleotides underlined. The set of MC29 v-myc oligonucleotides hybridized by λMH2-1 DNA (Fig. 6, oligonucleotides 1, 6, 7b, 8b, 26, 32a, 43, and
FIG. 4. Restriction enzyme analysis of clone λMH2-1 DNA. (A) Phage DNA (1.0 μg) was digested with BamHI (1), BamHI-EcoRI (2), EcoRI (3), EcoRI-KpnI (4), KpnI (5), BamHI-BglII (6), BamHI-ClaI (7), BamHI-PstI (8), BamHI-PvuII (9), BamHI-SaiI (10), BamHI-SstI (11), BamHI-SstII (12), EcoRI-BglII (13), EcoRI-ClaI (14), EcoRI-PstI (15), EcoRI-PvuII (16), EcoRI-SaiI (17), EcoRI-SaiI (18), and EcoRI-SstII (19). Molecular weight markers were DNA from XcI857 S7 digested with HindIII (m1) or HindIII-EcoRI (m2). DNA digests were analyzed on an 0.8% agarose gel.

DNA was transferred to two subsequent nitrocellulose sheets and hybridized to 32P-labeled DNA from p100 (B) or pMC29 (C) (see text).

120) corresponds exactly to the homologous set of v-myc oligonucleotides found in MH2 RNA (Fig. 1 and 2). This is direct confirmation that the cloned DNA contains the MH2 v-myc gene. This experiment also shows that the 3' noncoding myc sequences of MC29 represented by T1 oligonucleotides 31a, 15, 45, 44, and 3 have no homologous counterparts in MH2 proviral DNA, since these oligonucleotides are not found in the hybrid. This extends the analysis of MH2 v-myc RNA sequences (Fig. 1 and 2) and shows that the lack of these T1 oligonucleotides from MH2 RNA is not due to point mutations but to absence of most of the 3' noncoding myc sequences. Oligonucleotides specific for the gag gene (oligonucleotides 9a, 13, 20a, 25, and 52)

FIG. 5. Restriction map of λMH2-1 DNA. The genetic map of MH2 proviral DNA is deduced from comparison of the restriction site pattern with those of Rous sarcoma and MC29 proviral DNAs and from the hybridization data shown in Fig. 4 and 6. Dotted lines in the genetic map indicate that the precise location of gene borders is not known yet (see text). The solid line at the bottom indicates the coding region needed for the synthesis of a 100,000-dalton protein from genomic mRNA, starting at the normal gag translational initiation site. LTR, Long terminal repeat.
as described in the text. Methionine-containing peptides specific to MH2 p100, MC29 p110^{ag-myc}, and Pr180^{ag-pol} were determined by comparing the map of each of the three proteins to that of Pr76^{ag} of MHAV (Fig. 8A) and by analyzing appropriate mixtures of tryptic peptides (data not shown). Only the non-gag peptides are marked in the maps shown in Fig. 8. The map of p100 of MH2 contained 17 tryptic peptides (Fig. 8C). Of these, nine peptides (indicated by numbers) were not detected in the map of Pr76^{ag} (Fig. 8A) and therefore were specific to p100. Peptides 5 and 6 appeared only as minor spots but were reproducibly observed, whereas peptide 8 was not always detectable. Analysis of MC29 p110^{ag-myc} (Fig. 8E) revealed four methionine-containing tryptic peptides (arrows) that

FIG. 6. Fingerprint analysis of HBI-MC29 (RPV) RNA sequences related to λMH2-1 DNA. A total of 5.0 × 10^6 cpm (1.0 μg) of 32P-labeled 70S RNA from HBI-MC29 (RPV) was hybridized to 30 μg of denatured λMH2-1 DNA under conditions described in the legend to Fig. 1. The RNA in the hybrid was isolated and fingerprinted as described in the legend to Fig. 1. T1 oligonucleotides were identified by RNase A digestion and numbered accordingly. Underlined numbers indicate v-myc oligonucleotides (Table 1). Oligonucleotides specific for the gag region of HBI-MC29 and of RPV are numbered as in previous analyses (12).

(12, 15) are also found in the fingerprint of HBI-MC29 (RPV) RNA hybridized with λMH2-1 DNA (Fig. 6), confirming the presence of 5' gag sequences in λMH2-1 DNA. Interestingly, no characteristic T1 oligonucleotides from the pol and env genes of RPV or from the partial env sequences of HBI-MC29 (12) are detected in the hybrid between HBI-MC29 (RPV) RNA and λMH2-1 DNA.

Analysis of gag-related viral proteins. The structural relationship of MH2 p100 to p110^{ag-myc} of MC29 and to both Pr76^{ag} and Pr180^{ag-pol} of MHAV was analyzed by tryptic peptide mapping. The proteins were isolated by immunoprecipitation from [35S]methionine-labeled Q8 and MH2-A10 cells, and analytical gel electrophoresis of immunoprecipitates is shown in Fig. 7. MH2 p100, like MC29 p110^{ag-myc}, is precipitated by anti-gag serum but not by anti-reverse transcriptase serum. Tryptic digests of proteins eluted from preparative gels were prepared and separated by electrophoresis at pH 4.7, followed by ascending chromatography,
FIG. 8. Comparison of methionine-containing tryptic peptides of gag-related proteins encoded by MH2 and its helper virus MHAV and by MC29. Proteins were labeled biosynthetically with [35S]methionine, isolated by immunoprecipitation and preparative gel electrophoresis, and digested with trypsin as described in the text. The digests were separated into two dimensions by electrophoresis at pH 4.7 from the left to the cathode at the right, followed by ascending chromatography from bottom to top. (A) MHAV Pr76<sup>gag</sup>, 9,900 cpm, 6 days; (B) MHAV Pr180<sup>gag-pol</sup>, 4,500 cpm, 13 days; (C) MH2 p100, 9,700 cpm, 6 days; (D) MH2 p100, 4,800 cpm, and MHAV Pr180<sup>gag-pol</sup>, 4,500 cpm, 13 days; (E) MC29 p110<sup>gag-myc</sup>, 3,300 cpm, 16 days; (F) MC29 p110<sup>gag-myc</sup>, 3,300 cpm, and MH2 p100, 3,300 cpm, 16 days. Arrowheads in panels B and D indicate pol-specific peptides of Pr180<sup>gag-pol</sup>, numbers in panels C, D, and F indicate unique peptides of MH2 p100, and arrows in panels E and F indicate myc-specific peptides of MC29 p110<sup>gag-myc</sup>. Unmarked peptides in panels B through F represent gag-related protein sequences identified by comparison with the map of Pr76<sup>gag</sup> (A) and by direct mixing experiments (data not shown). They include p19-specific peptides.

did not comigrate with any of the peptides present in Pr76<sup>gag</sup> (Fig. 8A), confirming previous analysis of p110<sup>gag-myc</sup> (26). One of the myc-specific peptides appeared as a doublet, probably as a result of incomplete oxidation. The comparison of the map of MC29 p110<sup>gag-myc</sup> (Fig. 8E) with that of MH2 p100 (Fig. 8C) and direct analysis of a mixture of peptides derived from the two proteins (Fig. 8F) demonstrated that none of the myc-specific tryptic peptides of p110<sup>gag-myc</sup> comigrated with methionine-containing peptides specific to MH2 p100. The map of Pr180<sup>gag-pol</sup> of MHAV (Fig. 8B) contained most of the methionine-containing tryptic peptides of MHAV Pr76<sup>gag</sup> (Fig. 8A), as expected. In addition, nine peptides not shared with Pr76<sup>gag</sup> (arrowheads) were detected. One of these presumably pol-specific peptides may be related to the
unique peptide 1 of MH2 p100 (Fig. 8C). This was suggested by comigration of the two peptides in a map containing a mixture of tryptic peptides of the two proteins. (Fig. 8D).

**DISCUSSION**

**Structure of v-myc genes.** The cell-derived oncogene v-myc has so far been detected in four completely independent isolates of avian acute leukemia viruses: MC29, CMII, OK10, and MH2. The oncogenic spectra of these viruses are largely the same, and the correlation between oncogenicity and presence of v-myc in different viral isolates provided the first genetic evidence for involvement of v-myc in tumor induction. Analyses of these four viruses showed that their genomes contain overlapping but possibly not identical complements of v-myc (8, 9, 13, 14, 20, 40, 44). We analyzed the sequence conservation and complexity of the different v-myc genes in more detail to determine which genetic elements are essential for oncogenicity. From the data summarized in Fig. 2, it was evident that the v-myc sequences coding for the transforming protein of MC29 appear highly conserved in all other myc-containing viruses with transforming potential. Within the sequences corresponding to the open reading frame defined for the MC29 v-myc gene (2), only single base changes were detected in the other viral genomes by the techniques used. Most of these point mutations were found in MH2 v-myc. They lead to single amino acid substitutions, as a comparison with the nucleotide sequence of MC29 v-myc shows (Table 1; reference 2). The coding sequences of v-myc also have almost completely homologous counterparts in c-myc. One possible divergence between the MC29 v-myc coding sequences and c-myc was detected, and the same region is also represented by an allelic oligonucleotide (oligonucleotide 43a) in CMII, OK10, and MH2 RNAs. More extensive modifications, such as the deletion of sequences from the coding region of v-myc, lead to partial deafective for transformation, as in td10H-MC29 (15, 21). The sequences from the 3'-noncoding region of MC29 v-myc are far less conserved between the different v-myc genes and the c-myc gene (Fig. 2), and we showed (see below) that they are actually not even present in the MH2 genome. This strongly suggests that all v-myc genes and probably also the cellular c-myc gene use the same translation-termination codon as defined for the MC29 oncogene and that the 3'-noncoding sequences are not essential for v-myc function. Hence, the transforming proteins of these viruses and the normal cellular gene product appear to be strongly conserved in their ca. 47,000-dalton carboxy-terminal sequences but may contain single amino acid substitutions. The amino-terminal sequences of the viral proteins vary depending on the mode of gene expression for v-myc in the different viruses, and in the case of the hybrid proteins, they contain partial sequence elements from structural genes (8, 9). The conservation of the coding v-myc sequences strongly argues that the transforming function of the viral onc gene products is basically due to the myc domain of these proteins. It remains to be determined whether the different strategies of v-myc expression or single amino acid substitutions within v-myc or both could possibly modulate the function of the transforming proteins, leading to some of the observed oncogenic specificities, such as the high incidence of carcinoma induction in MH2-infected animals.

**Genetic structure of MH2.** The structural analyses of proviral DNA, viral RNA, and the gag-related protein of MH2 consistently demonstrate that the genetic design of MH2 is different from that of the other members of the MC29 subgroup (Fig. 5). The 1.3-kb v-myc sequences homologous to MC29 v-myc are located near the 3' terminus of the MH2 genome, separated from partial gag sequences at the 5' terminus by at least 1 kb of other sequences. The topography of the cloned proviral DNA implies that the 100,000-dalton gag-related protein of MH2 is encoded by the partial gag gene and the central sequences of the genome unrelated to v-myc. This is based on assumptions that synthesis of this protein is directed by genome-sized mRNA and that the normal gag initiation codon is used. Both assumptions appear correct, since cell-free translation of MH2 genomic RNA yields p100 (33), and p100 does contain typical tryptic peptides of p19, the amino-terminal protein encoded in the gag gene (Fig. 8). Tryptic peptide analysis directly shows that p100 does not contain myc-specific sequences, as defined by p110\(^{\text{myc}}\)-MC29. This is in agreement with a previous report on the lack of myc-specific phosphopeptides in p100 (36). The genetic structure of MH2 reported here is also consistent with the presence of a 2.6-kb subgenomic myc-related mRNA in MH2-transformed cells (33) and suggests that MH2 v-myc is expressed independent from major coding regions of structural genes. A putative MH2 v-myc gene product, with a molecular weight of 57,000, unrelated to the gag gene has recently been identified by cell-free translation of mRNAs from MH2-transformed cells hybrid selected with an myc-specific probe (33). The genetic origin of the central segment of the MH2 genome is of particular interest. Based on hybridization with ppol, the possibility that sequences from the 3' end of gag or the 5' half of pol are present can be ruled out. The hybridization of LMK2-1 DNA to HB1-MC29 (RPV) RNA
(Fig. 6) also gave no clues to the presence in MH2 of pol or env sequences shared with MC29 or RPV. Furthermore, pol-specific antigenic determinants recognized on Pr180\textsuperscript{pol} by the anti-reverse transcriptase serum used in these experiments are not present on p100 (Fig. 7). Hence, the restriction mapping, the hybridization data, and the protein analyses suggest that the central region of the MH2 genome contains additional sequences unrelated to virion genes, a conclusion also supported by further analysis of λMH2-I DNA by hybridization to Rous sarcoma virus proviral probes (H. W. Jansen and K. Bister, unpublished data). However, the presence of small residual complements of pol or env sequences in the center of the MH2 genome or at the 3' site of the v-myc region cannot be ruled out. A detailed view of the genetic arrangement of virion sequences, v-myc, and the additional specific sequences in the MH2 genome needs to be obtained by nucleotide sequence analysis of cloned proviral DNA. This would also give more insight into the mechanism of the v-myc expression in MH2, which may involve the use of viral splice acceptor and translational initiation sites or additional cell-derived, possibly c-myc-related sequences not present in the other viruses of the MC29 subgroup.

ACKNOWLEDGMENTS

We thank Ellen Dietzsch and Birgit Schroer for expert technical assistance.

ADDITIONUM IN PROOF

We have recently shown that the additional MH2-specific sequences, termed v-mil, are closely related to chicken DNA sequences, termed c-mil, which apparently represent a single-copy locus with a complex split gene structure (Jansen et al., EMBO J, in press).

LITERATURE CITED

29. Linial, M. 1982. Two retroviruses with similar transform-