Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 1168–1172, February 1989
Biochemistry

v-*cbl*, an oncogene from a dual-recombinant murine retrovirus that induces early B-lineage lymphomas

(pre-B cell tumors/gag fusion protein/GCN4 homology/hemopoietic cells)

WALLACE Y. LANGDON*†, JANET W. HARTLEY†, S. PETER KLINKEN*†, SANDRA K. RUSCETTI‡, AND HERBERT C. MORSE III†

*Walter and Eliza Hall Institute of Medical Research, P.O. Royal Melbourne Hospital, Parkville, Victoria 3050, Australia; †Laboratory of Immunopathology, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892; and ‡Laboratory of Genetics, National Cancer Institute, Bethesda, MD 20892

Communicated by Michael Potter, November 14, 1988 (received for review September 1, 1988)

ABSTRACT **Cas NS-1 is an acutely transforming murine retrovirus that induces pre-B and pro-B cell lymphomas. Molecular cloning showed it was generated from the ecotropic Cas-Br-M virus by sequential recombinations with endogenous retroviral sequences and a cellular oncogene. The oncogene sequence shows no homology with known oncogenes but some similarity to the yeast transcriptional activator GCN4. A 100-kDa gag–cbl fusion protein, with no detectable kinase activity, is responsible for the cellular transformation. The cellular homologue of v-*cbl*, present in mouse and human DNA, is expressed in a range of hemopoietic lineages.**

The ability of retroviruses to recombine with host chromosomal DNA has enabled the identification of genes with cancer-causing potential, collectively known as oncogenes (for reviews, see refs. 1–3). From studies of avian and mammalian retroviruses numerous oncogenes have been isolated and this has led to dramatic progress in the definition and understanding of the molecular basis of cancer. We have been examining the tumors that develop in mice after infection with an ecotropic retrovirus called Cas-Br-M. This virus was originally isolated from a wild mouse of the Lake Casitas region of California (4) and induces T- and B-cell lymphomas, erythroleukemia, and myeloid leukemia with latent periods of 5–8 months (5). It remains unclear whether Cas-Br-M generates this range of tumors by infecting multipotential hemopoietic cells or by infecting cells committed to a lineage. Furthermore, the etiological agent in some of these tumors is a recombinant virus formed soon after inoculation (6). This class of murine retrovirus is termed mink cell focus-forming (MCF) virus and evolves by recombination between the infecting ecotropic virus and nonecotropic retroviral sequences present in the mouse genome (7, 8).

To determine the relative importance of Cas-Br-M and MCF viruses in the induction of hemopoietic neoplasms, cell-free extracts were prepared from a range of tumors arising in NFS/N mice inoculated at birth with Cas-Br-M (6). These extracts were inoculated into newborn NFS/N mice since it was reasoned if Cas-Br-M was the tumor-inducing virus, then each extract should produce a range of tumor phenotypes with a long latent period. Alternatively, if lineage-specific MCF viruses in the extracts were the dominant tumor-inducing agents, then the tumors should correspond to the primary tumor and the latency should be reduced. The results from 3 of 12 extracts tested were consistent with the latter possibility (6).

One of the three extracts producing a high proportion of tumors with the same phenotype as the primary tumor and a reduced latency was called Cas NS-1. This extract was prepared from the spleen of a mouse with a pre-B cell lymphoma that developed 28 weeks after Cas-Br-M infection. Injection of the Cas NS-1 extract into newborn NFS/N mice resulted in a high proportion of pre-B cell lymphomas with a mean latent period of 16 weeks. This result indicated that a tumorigenic virus with a lineage preference had been generated. To characterize this virus further, the extract was added to mouse fibroblasts. Unexpectedly, the extract transformed the cells raising the possibility that a fibroblast-transforming virus, and not an MCF virus, induced the pre-B cell lymphomas. The results presented here describe the molecular characterization of this transforming virus and show that it was formed from Cas-Br-M by two recombinational events that involved endogenous MCF sequences and a cellular oncogene. The DNA and amino acid sequences[§] of the oncogene show no significant similarities to any known oncogene, and we have called this oncogene v-*cbl* [pronounced Sybil] for *C*asitas *B*-lineage *l*ymphoma.

MATERIALS AND METHODS

Virus Assays and Biological Cloning. Ecotropic murine retroviruses were assayed by XC plaque tests (9) and MCF viruses were detected by direct assay in mink lung cells (7). Transforming virus was detected by morphological changes of infected NIH 3T3 or normal rat kidney (NRK) cells. Biological cloning of the transforming virus was performed by two limiting dilution titrations in NRK cells.

Tumor Diagnosis. Cas NS-1-induced tumors were examined by staining of viable cells with monoclonal antibodies labeled with fluorescein isothiocyanate or Texas Red and analyzed by flow cytometry (6). Histologic diagnoses were made on the basis of microscopic examinations of tissues fixed in Tellyesnickie's solution, sectioned, and stained with hematoxylin and eosin (5).

Pulse-Labeling and Immune Precipitation. Cells were pulse-labeled for 30 min with [^{35}S]methionine (10) and extracts were immune precipitated with goat anti-Rauscher murine leukemia virus p15, p12, p30, or p10 antisera (from the Division of Cancer Cause and Prevention, National Cancer Institute, Bethesda, MD). Immune complexes were precipitated with a 10% (wt/vol) suspension of formalin-fixed *Staphylococcus aureus* and electrophoresed on 7% polyacrylamide gels, and the gels were fluorographed and exposed to x-ray film.

Molecular Cloning and DNA Sequencing. Genomic libraries were prepared from DNA of transformed NRK cells by partial digestion with *Mbo* I and complete digestion with *Eco*RI and packaged using the bacteriophages BF101 and Charon 4A (11). Recombinant bacteriophage were screened

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

Abbreviations: MCF virus, mink cell focus-forming virus; NRK cells, normal rat kidney cells.

[§]The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04169).

with murine retroviral long terminal repeat, gag, and env probes (12, 13). For sequencing, DNA fragments were subcloned into M13mp18 and -mp19 by using the restriction sites indicated in Fig. 3. The sequence was determined unambiguously on both strands and across all cloning sites using the chain-termination method (14).

Amino Acid Sequence Alignment. GCN4 (15, 16) and v-*cbl* sequences were first compared with the programs RELATE and SEARCH developed by Dayhoff (17). Optimal alignments of similar sequences were then obtained with the program ALIGN using the mutation data matrix, a matrix bias of +6, and a break penalty of 6 (17). The program calculates an alignment score by comparison of the score for the optimal alignments of the given sequences with the mean score for the optimal alignments of 100 random permutations of the given sequences. Alignment scores >3.0 SDs are usually taken as an indication of relatedness (17).

DNA and RNA Analysis. DNA samples (10 μg) were digested with restriction enzymes, separated on 0.8% agarose gels, denatured, and transferred to nitrocellulose membranes (18). Poly(A)$^+$ RNA (4 μg) was electrophoresed through a 1% agarose gel containing 6% (vol/vol) formaldehyde and 1× Mops buffer (11) and transferred to nitrocellulose (19). DNA and RNA membranes were hybridized overnight at 42°C in a solution containing 5× SSC, 50% formamide, 0.2% $NaDodSO_4$, 1× Denhardt's solution, denatured salmon sperm DNA at 50 μg/ml, and 10% (wt/vol) dextran sulfate (19) (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0; 1× Denhardt's = 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin). The probes were ^{32}P-labeled by using a random-primer kit (BRESA Biotechnology, Adelaide, South Australia). Stringent washes were at 60°C in 0.1× SSC plus 0.1% $NaDodSO_4$.

RESULTS

Transformation of Early B-Lineage Cells. The Cas NS-1 tumor extract showed the expected presence of Cas-Br-M (assayed by XC plaques) and MCF viruses (assayed by focus formation in mink lung cells). To separate the transforming virus, the extract was added to NRK cells and transformed clones that lacked reverse transcriptase activity were selected. A clonal, transformed nonproducer cell line, free of ecotropic and MCF viruses, was obtained and used for further biological and molecular analyses.

To ensure that the nonproducer cells contained viral sequences associated with transforming activity for hemopoietic cells, the line was superinfected with ecotropic viruses and the supernatants were inoculated into newborn and adult mice. Twenty newborn NFS/N mice infected with Cas-Br-M, Moloney, or Akv pseudotypes of Cas NS-1 and 13 newborn BALB/c mice infected with Moloney virus pseudotypes developed pre-B or pro-B cell lymphomas with a mean latent period of 15 weeks. These diagnoses were based on the expression of early B-lineage cell-surface markers (see *Discussion*) and immunoglobulin heavy chain gene rearrangements (data not shown). In contrast to newborns, no tumors were observed in adult-infected NFS/N, 129/J, SWR/J, C58/J, or DBA/2J mice.

Cas NS-1 Encodes a gag Fusion Protein. To determine whether the transforming Cas NS-1 virus encodes a unique protein, virus-infected NIH 3T3 cells were metabolically labeled with [^{35}S]methionine and the labeled extracts were precipitated with antiserum to the viral gag protein p12. Cells infected with the Cas-Br-M virus alone (Fig. 1*A*, lane 3) express the 65-kDa gag precursor protein Pr65^{gag}. Cells infected with the Cas NS-1 virus complex (Fig. 1*A*, lane 1) express the 65-kDa gag protein and a new p12-related protein with a molecular mass of 100 kDa. Extracts from nonproducer cells immunoprecipitated with the anti-p12 antiserum

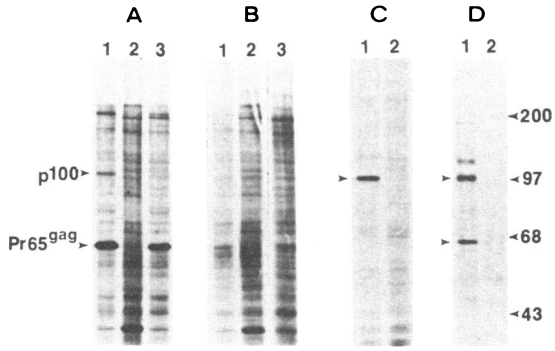

FIG. 1. Autoradiographs of metabolically labeled cell extracts precipitated with goat anti-gag p12 antiserum. (*A*) NIH 3T3 cells infected with the Cas NS-1 virus complex (lane 1), uninfected NIH 3T3 cells (lane 2), and NIH 3T3 cells infected with Cas-Br-M ecotropic virus (lane 3) precipitated with anti-p12 antiserum. (*B*) As in *A* except the extracts were precipitated with normal goat serum. (*C*) Nonproducer clone selected from NRK cells transformed with the Cas NS-1 virus complex and precipitated with anti-p12 antiserum (lane 1), or normal goat serum (lane 2). (*D*) Spleen cells from a Cas NS-1-induced pre-B cell lymphoma precipitated with anti-p12 antiserum (lane 1) or normal goat serum (lane 2). Molecular mass markers are in kDa.

also express the 100-kDa protein (Fig. 1*C*, lane 1), thus implicating it as a gag–onc fusion protein responsible for fibroblast and pre-B cell transformation. The 100-kDa protein could also be precipitated with antisera to two other gag proteins, p15 and p30, but was precipitated only weakly with antiserum to the p10 gag protein that is encoded by sequences at the 3′ end of the viral *gag* gene (data not shown). Spleen cells from pre-B cell lymphomas induced by the Cas NS-1 viral extract also express high levels of the 100-kDa protein (Fig. 1*D*, lane 1), along with the gag precursor of the helper virus. The 100-kDa Cas NS-1 protein can be labeled with [^{32}P]orthophosphate but does not contain detectable kinase activity, as it lacks the ability to autophosphorylate or phosphorylate immunoglobulin heavy chain proteins (data not shown). In addition, it does not appear to undergo N-linked glycosylation or myristoylation (data not shown).

A Transforming Virus with an MCF Envelope. DNA extracted from the transformed NRK cells and digested with restriction enzymes was examined by Southern blot hybridization by using probes spanning the murine retrovirus genome. These experiments identified a provirus with a 7.4-kilobase-pair (kbp) genome that contained a recombinant MCF envelope gene. This was determined by hybridization to a probe that recognizes the *env* sequences of xenotropic and MCF viruses but not ecotropic viruses (12). Consistent with this, the transforming virus did not hybridize to an ecotropic-specific env probe (13).

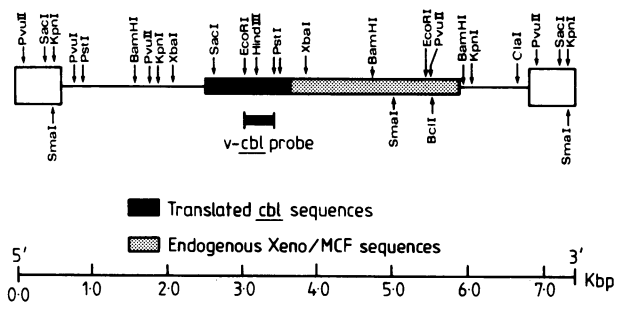

FIG. 2. Restriction map of the Cas NS-1 transforming virus. The *gag–cbl* and *cbl–pol* junctions of Cas NS-1 were determined from the DNA sequence shown in Fig. 3. The junction between endogenously derived and Cas-Br-M-derived *env* sequences was estimated from the restriction maps.

To molecularly clone the Cas NS-1 provirus, genomic DNA libraries were prepared from the transformed NRK cells by using the λ bacteriophages Charon 4A and BF101. Both libraries were screened with a long terminal repeat probe and positive clones were further characterized with gag and env probes. The largest λBF101 clone in combination with a Charon 4A clone gave the entire genome of the transforming virus. The two clones were ligated using an overlapping *Bcl* I site common to MCF virus envelopes (Fig. 2), and for ease of manipulation the viral DNA was subcloned into a pUC plasmid. DNA from this clone was transfected into NIH 3T3 fibroblasts and within 2 weeks foci of transformed cells were evident.

The molecular cloning of the Cas NS-1 virus enabled detailed restriction mapping (Fig. 2). Cas NS-1 has identical long terminal repeat and gag restriction sites to Cas-Br-M (20), but on the 3′ side of the *gag Xba* I site the restriction sites are unique to Cas NS-1. This divergence extends through the endogenously derived *Bam*HI and *Xba* I sites that are also present in the *pol* gene of some MCF viruses (21). Restriction mapping in this region, therefore, indicates the generation of Cas NS-1 involved the loss of the entire Cas-Br-M *pol* gene and substitution with an oncogene and endogenous *pol* sequences. The Cas NS-1 envelope contains restriction sites common to MCF viruses, notably the *Eco*RI, *Pvu* II, and *Bcl* I sites.

```
                                                ←gag p30→
  Leu Glu Arg Leu Lys Glu Leu Ser Gln Val His Pro Tyr Asp Pro Glu Asp Pro Gly Gln Glu Thr Asn Val Ser Met Ser Phe Ile
T CTA GAG AGA CTT AAG GAG CTA TCG CAG GTA CAC CCT TAT GAT CCT GAA GAT CCT GGG CAA GAA ACT AAT GTG TCT ATG TCT TTT ATT      88
 XbaI
 Trp Gln Ser Ala Pro Asp Ile Gly Arg Lys Leu Glu Arg Leu Glu Asp Leu Lys Ser Lys Thr Leu Gly Asp Leu Val Arg Glu Ala Glu
 TGG CAG TCT GCC CCA GAC ATT GGG AGA AAG TTA GAG AGA TTG GAA GAT TTA AAA AGT AAG ACG CTT GGA GAT CTA GTA AGA GAG GCA GAA  178

 Lys Ile Phe Asn Lys Arg Glu Thr Pro Glu Glu Arg Glu Glu Arg Ile Lys Arg Glu Thr Lys Glu Lys Glu Glu Arg Arg Arg Ala Glu
 AAG ATA TTT AAT AAG CGA GAA ACC CCG GAA GAA AGA GAG GAA CGT ATC AAG AGA GAA ACA AAG GAA AAA GAA GAA CGC CGT AGG GCA GAG  268
                                                ←gag p10
 Asp Glu Gln Arg Glu Lys Glu Arg Asp Arg Arg Arg His Arg Glu Met Ser Lys Leu Leu Ala Thr Val Val Ile Gly Gln Arg Gln Asp
 GAT GAG CAG AGA GAG AAA GAA AGG GAC CGC AGA AGA CAT AGA GAG ATG AGC AAG CTC TTG GCC ACT GTA GTT ATT GGT CAG AGA CAG GAT  358
                                                                                                  ←cbl
 Arg Gln Gly Gly Glu Arg Arg Arg Pro Gln Leu Asp Lys Asp Gln Cys Ala Tyr Cys Lys Glu Thr Gly His Trp Ala Ala Ser Ala Gly
 AGA CAG GGG GGA GAG CGG AGG AGG CCC CAA CTT GAT AAG GAC CAA TGC GCC TAC TGC AAA GAA ACG GGA CAC TGG GCT GCG AGC GCC GGC  448

 Gly Gly Cys Arg Arg Gly Pro Ser Phe Ser Pro Gly Ser Ile Pro Ser Leu Ala Ala Glu Arg Ala Pro Asp Pro Pro Leu Ala Met Ala
 GGT GGT TGC CGG AGA GGC CCC TCC TTC TCG CCC GGC TCC ATT CCC TCG CTC GCG GCC GAG CGG GCT CCC GAC CCT CCG CTG GCC ATG GCC  538

 Gly Asn Val Lys Lys Ser Ser Gly Ala Gly Gly Gly Gly Ser Gly Gly Ser Gly Ala Gly Gly Leu Ile Gly Leu Met Lys Asp Ala Phe
 GGC AAC GTG AAG AAG AGC TCG GGC GCC GGC GGC GGC GGC TCT GGG GGC TCG GGA GCG GGC GGC CTG ATC GGG CTC ATG AAG GAC GCC TTC  628
    SacI
 Gln Pro His His His His His His Leu Ser Pro His Pro Pro Cys Thr Val Asp Lys Lys Met Val Glu Lys Cys Trp Lys Leu Met Asp
 CAG CCG CAC CAC CAC CAC CAC CAC CTC AGC CCG CAC CCT CCC TGC ACG GTG GAC AAG AAG ATG GTG GAG AAG TGC TGG AAG CTC ATG GAC  718

 Lys Val Val Arg Leu Cys Gln Asn Pro Asn Val Ala Leu Lys Asn Ser Pro Pro Tyr Ile Leu Asp Leu Leu Pro Asp Thr Tyr Gln His
 AAG GTG GTG CGG TTG TGT CAA AAC CCA AAC GTG GCG CTC AAG AAC AGC CCG CCT TAT ATC TTA GAC CTG CTG CCT GAC ACC TAC CAG CAC  808

 Leu Arg Thr Val Leu Ser Arg Tyr Glu Gly Lys Met Glu Thr Leu Gly Glu Asn Glu Tyr Phe Arg Val Phe Met Glu Asn Leu Met Lys
 CTC CGC ACT GTC TTG TCA AGA TAT GAG GGG AAG ATG GAG ACG CTT GGA GAA AAT GAG TAT TTC AGG GTG TTC ATG GAA AAT TTG ATG AAG  898

 Lys Thr Lys Gln Thr Ile Ser Leu Phe Lys Glu Gly Lys Glu Arg Met Tyr Glu Glu Asn Ser Gln Pro Arg Arg Asn Leu Thr Lys Leu
 AAA ACT AAG CAG ACT ATC AGC CTC TTC AAG GAG GGA AAA GAA AGG ATG TAT GAG GAG AAT TCC CAG CCT AGG CGA AAC CTG ACC AAA TTA  988
                                          EcoRI
 Ser Leu Ile Phe Ser His Met Leu Ala Glu Leu Lys Gly Ile Phe Pro Ser Gly Leu Phe Gln Gly Asp Thr Phe Arg Ile Thr Lys Ala
 TCC CTG ATC TTC AGC CAC ATG CTG GCA GAA CTG AAA GGC ATC TTT CCG AGC GGA CTC TTC CAA GGA GAC ACT TTC CGG ATT ACT AAA GCT 1078

 Asp Ala Ala Glu Phe Trp Arg Lys Ala Phe Gly Glu Lys Thr Ile Val Pro Trp Lys Ser Phe Arg Gln Ala Leu His Glu Val His Pro
 GAT GCT GCC GAA TTT TGG AGA AAA GCT TTT GGT GAA AAG ACG ATA GTC CCG TGG AAG AGC TTT CGA CAG GCC CTG CAT GAA GTG CAT CCC 1168
                            HindIII
 Ile Ser Ser Gly Leu Glu Ala Met Ala Leu Lys Ser Thr Ile Asp Leu Thr Cys Asn Asp Tyr Ile Ser Val Phe Glu Phe Asp Ile Phe
 ATC AGT TCT GGG CTG GAG GCC ATG GCT CTG AAG TCC ACT ATT GAT CTG ACC TGC AAT GAT TAT ATT TCT GTC TTT GAA TTT GAT ATT TTT 1258

 Thr Arg Leu Phe Gln Pro Trp Ser Ser Leu Leu Arg Asn Trp Asn Ser Leu Ala Val Thr His Pro Gly Tyr Met Ala Phe Leu Thr Tyr
 ACA CGG CTG TTT CAG CCC TGG TCC TCT TTG CTC AGA AAT TGG AAC AGC CTT GCT GTA ACT CAC CCT GGT TAC ATG GCT TTC CTG ACA TAC 1348

 Asp Glu Val Lys Ala Arg Leu Gln Lys Phe Ile His Lys Pro Gly Ser Tyr Ile Phe Arg Leu Ser Cys Thr Arg Leu Gly Gln Trp Ala
 GAT GAA GTG AAA GCG CGC CTG CAG AAG TTC ATC CAC AAA CCT GGC AGT TAC ATC TTT CGG CTG AGC TGT ACT CGT TTG GGT CAG TGG GCT 1438
                          PstI
 Ile Gly Tyr Val Thr Ala Asp Gly Asn Ile Leu Gln Thr Ile Pro His Asn Lys Pro Leu Phe Gln Ala Leu Ile Asp Gly Phe Arg Glu
 ATT GGG TAT GTT ACT GCC GAT GGG AAC ATT CTG CAG ACA ATC CCA CAC AAT AAA CCG CTC TTC CAA GCA CTG ATT GAT GGC TTC AGG GAA 1528
                              PstI
 Gly Phe Tyr Leu Phe Pro Asp Gly Arg Asn Gln Asn Pro Asp Leu Thr Gly Leu Cys Glu Pro Thr Pro His Phe Ser ***
 GGC TTC TAT TTG TTT CCT GAT GGA CGA AAT CAA AAT CCT GAC CTG ACA GGT TTA TGT GAA CCA ACT CCT CAC TTC TCA TAG AGACCTCAACC 1620
                                                                                                    pol→
 CCGTACACTCCAGACTTCTTCCATTATACTGAGACAGATATAAAGAACCTACAAGAGTTGGGAGCCACATATGATAGAGAGAAAAAATATTGGGTCCTGCAAGGTAAACCTGTGATGCCT 1740

 GACCAGTTCACCTTTGAATTATTAGACTTCCTTCACCAGCTCACCCACCTTAGCTATCAGAAGATGAGGGCACTTCTAGACCGGAAAGAAAGCCCCTATTACATGCTAAATAAAGATAAG 1860
                                                                  XbaI
 ATCCTCCACGAGGTGGCGGAATCATGCCAAGCCTGTGTCCAAGTAAATGCCAGTAAGACTAAGATCAGGGCCGGAACACGAGTAAGAGGACATCGACCAGGCACCCATTGGAAATTGACA 1980

 TTTACTGAAGTGAAGCCCGGACTGTATGGGTACACGTATCTCCTGGTATTCGTGGACACGTTCTCTGGCTGGGTTGAAGCCTTCCCAACCAAACATGAGACTGCCAAATAGTGACCAAGA 2100

 AACTTTGGAAGAAATATTTCCAAGGTTTAGAATGCCCCAAGTATTGGGGACTGATAATGGGCCTGCCTTCGTCTCCCAGGTAAGTCAGTCGGTGGCCAAGCTACTGGGGATTGATTGGAA 2220

 ACTACATTGTGCTTACAGACCCCAGAGTTCAGGTCAGGTAAAAAGAATGAATAGGACAATCAAGGACACTTTGACCAAATTAACGCTTCGAACTGGCACTAGAGACTGGGTACTCCTACT 2340

 TCCCTTGGCCCTCTACCGAGCCCGCAACACTCCGGGCCCCCATGGACTCATTCCGTATAAAATCCTGTATGGCGCCCCCGCCCCTGTTAATTTCCATGATCCTGAAATGTCAAAGTTTAC 2460

 TAATAGCCCCTCTCTCCAAGCTCACTTACAGGCCCTCCAACCAGTACAACGAGAGGTCTGGAAGCCACTGGCCGCTGCCATCTAGGACCAGCAGGACCAGCCAGTGATACCACACCCCTT 2580

 CCGTGTAGGCGACACCGTGTGGGTACGCCGGCACCAGACTAAGAACTTGGAACCTCGCTGGAAAGGACCCTACACCGTCCTGCTGACCACCCCCACCGCTCTCAAAGTAGACGGCATCGC 2700

 TGCGTGGATCC                                                                                                          2711
      BamHI
```

FIG. 3. DNA and amino acid sequences of the Cas NS-1 virus from the *gag* p30 *Xba* I site to the endogenously derived *Bam*HI site at the 3′ end of *pol*. The p30–p10, p10–*cbl*, and *cbl*–*pol* junctions were determined by comparisons with the sequences of Moloney and Akv murine leukemia viruses (22, 23). Computer searches of DNA and protein data bases failed to find any oncogene with similarity to v-*cbl*.

Cas NS-1 Contains the Oncogene v-*cbl*. To define precisely the extent of oncogene and endogenous retroviral recombination involved in the generation of the Cas NS-1 virus and to determine whether the oncogene is related to known oncogenes, the viral DNA was sequenced from the *gag* p30 *Xba* I site to the endogenous *pol Bam*HI site (Fig. 3). The position of the oncogene junction was determined to be in the *gag* gene for p10 at a position that removed the terminal 24 *gag*-encoded amino acids. Although the Cas-Br-M *gag* region has not been sequenced, the *gag* sequence of Cas NS-1 shows sufficient similarity to Moloney (22) and Akv (23) viruses (90% and 83% identity, respectively) that the *gag–onc* junction could be determined readily. The sequence on the 3′ side of the *gag–onc* junction showed that the open reading frame continued for a further 1170 nucleotides to a TAG stop codon at nucleotide 1515 (Fig. 3). This results in the translation of 390 amino acids of nonviral origin that contributes 43 kDa to the 100-kDa fusion protein.

Extensive computer searches (GenBank Release 55) of the nucleic acid and protein data bases failed to find any oncogene with significant sequence similarity to the transforming gene of Cas NS-1. This indicates the virus contains an oncogene, which we named v-*cbl*. The only amino acid sequence to show any significant similarity to v-cbl is the yeast regulatory protein GCN4, which binds to DNA and functions as a transcriptional activator (24–26). Interestingly, the two regions of similarity occur in those parts of GCN4 that are required for transcriptional activation and DNA binding. Transcriptional activation by GCN4 requires a 19-amino acid segment in the middle of the molecule (27). In Fig. 4*a* the alignment of 15 of these GCN4 amino acids with 15 amino acids from the middle of v-cbl (residues 150–164 counted from the amino terminus) shows 7 identities with no breaks. The importance of this similarity is unclear since "activating regions" are characterized by a functionally important acidic domain but show marked variations in their amino acid sequences (27, 28). Possibly more relevant is the homology of v-cbl to the GCN4 DNA-binding domain of 60 amino acids in the carboxyl-terminal region (27). Within this region, v-cbl shows 18 identities out of 61 possible matches (i.e., 29% identity), which is equivalent to the similarity between v-fos and GCN4 (29). The alignment and statistical significance of this similarity are shown in Fig. 4*b*. The v-*cbl* oncogene does not show a significant similarity to the avian v-*jun* oncogene, which has a 44% identity to GCN4 within the carboxyl terminus (29). Although the two alignments shown in Fig. 4 are highly suggestive of homology between v-cbl and GCN4, it remains to be determined whether v-cbl functions as a transcriptional activator in the same manner as GCN4 and the v-*jun* and v-*fos* oncogenes (30–32).

FIG. 4. Amino acid sequence alignments of v-cbl and GCN4. (*a*) v-cbl amino acids 150–164 (Fig. 3) and GCN4 amino acids 111–125 (15, 16). (*b*) v-cbl amino acids 238–313 and GCN4 amino acids 216–276. Circles denote identical amino acid. The alignment scores for *a*, 4.65 standard deviations, and *b*, 3.43 standard deviations, are greater than the 3.0 SDs usually taken as an indication of relatedness (17).

Additional features of interest in the v-cbl sequence include a region containing 10 glycines over a stretch of 14 amino acids immediately after the *Sac* I site at the 5′ end of the v-*cbl* gene, and a repeat of 6 histidines 11 amino acids upstream from the last glycine repeat. The chicken c-*fps* oncogene also contains a repeat of 6 histidines (33), but the amino acids on either side of the histidines in c-fps show no similarity to v-cbl. The sequence data do not predict a protein kinase domain or a zinc-finger structure.

The nontranslated sequences on the 3′ side of the v-*cbl* open reading frame are similar to Moloney and Akv *pol* sequences (78% and 83% identity, respectively). This finding and the presence of the *Xba* I and *Bam*HI sites found in endogenous *pol* genes (34) proved this region to be retrovirus-related but derived from the NFS/N genome.

c-*cbl* in Mouse, Rat, and Human DNA. A 418-base pair (bp) *Eco*RI–*Pst* I fragment from v-*cbl* (Fig. 2) was used to probe mouse, rat, and human DNA and detected sequences in all species (Fig. 5). The degree of sequence similarity of v-*cbl* to the human c-*cbl* gene appears high since the hybridization was carried out under stringent conditions.

To study the expression of c-*cbl*, RNA was prepared from a range of hemopoietic tumor cell lines (35) and examined with the v-cbl probe. A large 10.7-kilobase (kb) mRNA was detected in cells of the B, T, erythroid, myeloid, and mast cell lineages. The B-cell tumor WEHI-231 showed two additional c-cbl mRNA species of 12.2 kb and 8.2 kb that appeared to be expressed at levels equivalent to the 10.7-kb species. Smaller mRNAs were also detected in a number of other lineages, notably a 6.2-kb mRNA in the plasmacytoma MPC-11, but they were less abundant than the 10.7-kb mRNA.

DISCUSSION

The molecular isolation and DNA sequencing of the murine Cas NS-1 retrovirus have shown that this virus arose from the Cas-Br-M ecotropic virus by sequential recombinations with

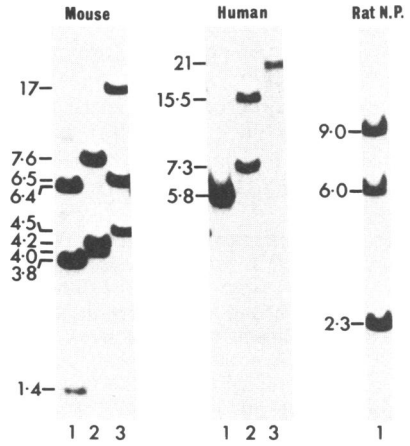

FIG. 5. *cbl* sequences in mouse, human, and rat nonproducer (N.P.) DNA. Genomic DNAs from C57BL/6 liver, human spleen, and NRK nonproducer cells were digested with *Eco*RI (lanes 1), *Xba* I (lanes 2), or *Bam*HI (lanes 3). Membranes were hybridized with the 418-bp *Eco*RI–*Pst* I v-cbl probe shown in Fig. 2. The 2.3-kbp *Eco*RI fragment in the rat nonproducer DNA is from the Cas NS-1 virus. Molecular mass markers are in kbp.

host cellular genes. The first recombination most likely involved the acquisition of endogenous *pol* and *env* sequences to generate an MCF intermediate that, in turn, transduced a cellular oncogene. The phenomenon of retroviral recombination with two distinct cellular genes has occurred in the generation of the Harvey and Kirsten murine sarcoma viruses, avian erythroblastosis virus, avian MH2 and E26 viruses, and the Gardner–Rasheed feline sarcoma virus (37–41). This, however, demonstrates that a murine MCF virus can transduce a cellular oncogene. It is clear that the recombinant MCF envelope is not involved in the transformation process since a construct lacking the envelope region readily transforms fibroblasts (data not shown). The transforming potential of Cas NS-1, therefore, correlates with the expression of the 100-kDa gag–cbl protein.

The induction of B-lineage lymphomas *in vivo* and the absence of protein kinase activity distinguish Cas NS-1 from other fibroblast-transforming retroviruses. The majority of fibroblast-transforming retroviruses isolated to date induce sarcomas in animals and most of these function as protein kinases (42). The only other naturally occurring transforming retroviruses that induce tumors in the B-lymphoid lineage are the Abelson murine leukemia virus and the avian reticuloendotheliosis virus (43, 44). The oncogenes of these viruses, v-*abl* and v-*rel*, respectively, show no significant nucleotide or amino acid similarities to v-*cbl*. Furthermore, the v-*abl* gene product is a tyrosine kinase (45). A comparison between v-*cbl* and v-*abl* is worth noting since both induce tumors that are histologically and phenotypically similar. The tumors are classified as lymphoblastic lymphomas and in the majority of cases they exhibit the same pre-B cell phenotype—i.e., they are negative for surface immunoglobin and Thy-1 but positive for the early B-lineage markers Ly-5 (B220) and Lyb-2. However, 10% of the v-*cbl*-induced lymphomas coexpress early B-lymphoid antigens and the macrophage marker Mac-1. Tumors of this phenotype have been designated pro-B cell lymphomas (46), and their normal counterparts are found in fetal liver and newborn marrow and spleen (47). Lymphomas induced by v-*cbl* also differ from those induced by v-*abl* in their longer latencies and the resistance exhibited by adult mice. This is consistent with the suggestion that a very early B-lineage cell, represented at higher levels in newborns than adults, is the target for transformation by v-*cbl*.

The isolation of this oncogenic retrovirus is a further example of Cas-Br-M generating an interesting recombinant. During our investigation of the tumors of Cas-Br-M infected mice, we have isolated a spleen focus-forming virus (36), an MCF virus that rapidly induces T-cell lymphomas (6), and the v-*cbl* oncogene reported here. These viruses were detected by inoculating newborn mice with cell-free extracts of primary tumor tissues and it is possible that further experiments using this approach would detect more recombinants.

We thank Drs. Iswar Hariharan, Graham Baldwin, and Charles Buckler for searches of the sequence data bases; Joan Austin and Craig Hyland for excellent assistance; and Drs. Jerry Adams, Wendy Davidson, Iswar Hariharan, and Alan Harris for helpful suggestions in the preparation of this manuscript. This work was supported in part by grants from the Australian National Health and Medical Research Council, the National Cancer Institute (CA 12421), and Contract N01-A1-22673 to Microbiological Associates (Bethesda, MD).

1. Bishop, J. M. (1983) *Annu. Rev. Biochem.* **52**, 301–354.
2. Varmus, H. E. (1984) *Annu. Rev. Genet.* **18**, 553–612.
3. Bister, K. & Jansen, H. W. (1986) *Adv. Cancer Res.* **47**, 99–188.
4. Hartley, J. W. & Rowe, W. P. (1976) *J. Virol.* **19**, 19–25.
5. Fredrickson, T. N., Langdon, W. Y., Hoffman, P. M., Hartley, J. W. & Morse, H. C., III (1984) *J. Natl. Cancer Inst.* **72**, 447–454.
6. Holmes, K. L., Langdon, W. Y., Fredrickson, T. N., Coffman, R. L., Hoffman, P. M., Hartley, J. W. & Morse, H. C., III (1986) *J. Immunol.* **137**, 679–688.
7. Hartley, J. W., Wolford, N. K., Old, L. J. & Rowe, W. P. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 789–792.
8. Rommelaere, J., Faller, D. V. & Hopkins, N. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 495–499.
9. Rowe, W. P., Pugh, W. E. & Hartley, J. W. (1970) *Virology* **42**, 1136–1139.
10. Ruscetti, S. K., Linemeyer, D., Field, J., Troxler, D. & Scolnick, E. M. (1979) *J. Virol.* **30**, 787–798.
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
12. Buckler, C. E., Hoggan, M. D., Chan, H. W., Sears, J. F., Khan, A. S., Moore, J. L., Hartley, J. W., Rowe, W. P. & Martin, M. A. (1982) *J. Virol.* **41**, 228–236.
13. Chan, H. W., Bryan, T., Moore, J. L., Staal, S. P., Rowe, W. P. & Martin, M. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5779–5783.
14. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
15. Thireos, G., Penn, M. D. & Greer, H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5096–5100.
16. Hinnebusch, A. G. & Fink, G. R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5374–5378.
17. Dayhoff, M. O. (1979) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), pp. 1–8.
18. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
19. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
20. Jolicoeur, P., Nicolaiew, N., DesGroseillers, L. & Rassart, E. (1983) *J. Virol.* **45**, 1159–1163.
21. Chattopadhyay, S. K., Cloyd, M. W., Linemeyer, D. L., Lander, M. R., Rands, E. & Lowy, D. R. (1982) *Nature (London)* **295**, 25–31.
22. Shinnick, T. M., Lerner, R. A. & Sutcliffe, J. G. (1981) *Nature (London)* **293**, 543–548.
23. Herr, W. (1984) *J. Virol.* **49**, 471–478.
24. Hinnebusch, A. G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6442–6446.
25. Penn, M. D., Galgoci, B. & Greer, H. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2704–2708.
26. Lucchini, G., Hinnebusch, A. G., Chen, C. & Fink, G. R. (1984) *Mol. Cell. Biol.* **4**, 1326–1333.
27. Hope, I. A. & Struhl, K. (1986) *Cell* **46**, 885–894.
28. Ma, J. & Ptashne, M. (1987) *Cell* **48**, 847–853.
29. Vogt, P. K., Bos, T. J. & Doolittle, R. F. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3316–3319.
30. Bohmann, D., Bos, T. J., Admon, A., Nishimura, T., Vogt, P. K. & Tjian, R. (1987) *Science* **238**, 1386–1392.
31. Struhl, K. (1988) *Nature (London)* **332**, 649–650.
32. Lech, K., Anderson, K. & Brent, R. (1980) *Cell* **52**, 179–184.
33. Huang, C.-C., Hammond, C. & Bishop, J. M. (1985) *J. Mol. Biol.* **181**, 175–186.
34. Khan, A. S., Rowe, W. P. & Martin, M. A. (1982) *J. Virol.* **44**, 625–636.
35. Kemp, D. J., Harris, A. W., Cory S. & Adams, J. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2876–2880.
36. Langdon, W. Y., Hoffman, P. M., Silver, J. E., Buckler, C. E., Hartley, J. W., Ruscetti, S. K. & Morse, H. C., III (1983) *J. Virol.* **46**, 230–238.
37. Chien, Y.-H., Lai, M., Shih, T. Y., Verma, I. M., Scolnick, E. M., Roy-Burman, P. & Davidson, N. (1979) *J. Virol.* **31**, 752–760.
38. Vennström, B. & Bishop, J. M. (1982) *Cell* **28**, 135–143.
39. Kan, N. C., Flordellis, C. S., Garon, C. F., Duesberg, P. H. & Papas, T. S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6566–6570.
40. Nunn, M. F., Seeburg, P. H., Moscovici, C. & Duesberg, P. H. (1983) *Nature (London)* **306**, 391–395.
41. Naharro, G., Robbins, K. C. & Reddy, E. P. (1984) *Science* **223**, 63–66.
42. Bishop, J. M. (1985) *Cell* **42**, 23–38.
43. Abelson, H. T. & Rabstein, L. S. (1970) *Cancer Res.* **30**, 2213–2222.
44. Chen, L., Lim, M. Y., Bose, H. & Bishop, J. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 549–553.
45. Witte, O. N., Rosenberg, N. & Baltimore, D. (1980) *Nature (London)* **283**, 826–831.
46. Davidson, W. F., Fredrickson, T. N., Rudikoff, E. K., Coffman, R. L., Hartley, J. W. & Morse, H. C., III (1984) *J. Immunol.* **133**, 744–753.
47. Holmes, K. L. & Morse, H. C., III (1984) *Curr. Top. Microbiol. Immunol.* **113**, 224–230.