ISOLATION AND SEQUENCE ANALYSIS OF SERINE PROTEASE cDNAs FROM MOUSE CYTOLYTIC T LYMPHOCYTES

BY BYOUNG S. KWON,* DANIEL KESTLER,* ESTHER LEE,* MARK WAKULCHIK,* AND JOHN DING-E YOUNG

From the *Laboratory of Molecular Genetics, Guthrie Research Institute, Sayre, Pennsylvania 18840; and the 1Laboratory of Cellular Physiology and Immunology, The Rockefeller University, New York, New York 10021

Cytolytic T lymphocytes (CTL)1 and NK cells are well-characterized effector cells with potent cytolytic activity (1-4). How these killer cells damage their targets has not been resolved. Biochemical analysis performed in the last 3 yr with homogeneous populations of killer lymphocytes has already identified a pore-forming protein (PFP, perforin or cytolsin) (5-8) and several serine esterases (9-13) localized in the cytoplasmic granules of these cell types. More recent studies by Masson and Tschopp (14), and Tschopp and Jongeneel (15) described the identification and partial sequences of six different granule serine esterases (which they call granzymes) found in murine CTL. PFP lyses a variety of target cells, whereas the function of granule serine esterases is still unknown although it has been suggested that these enzymes and PFP may form a novel cytolytic cascade analogous to the humoral complement system (16, 17). Proteolytic enzymes have long been implicated in cytotoxic reactions, as implied from earlier studies with protease inhibitors (18-21). A more detailed structural analysis of serine esterases is clearly warranted to elucidate their functional role in cell-mediated killing.

An alternative strategy to direct purification and characterization of killer cell proteins, which has been used to study cell-mediated killing, has consisted of subtractive cDNA cloning. In this approach CTL-specific cDNA clones are generated first and then probed secondarily for their function. Notably, of the recently isolated CTL-specific cDNA clones, three code for distinct serine proteases, designated H factor (22), CCP1/CTLA1 (23, 24), and CCP2/granzyme C (23, 25). Recently, a modified differential screening procedure was developed in this laboratory to generate a broad representation of T lymphocyte subset-specific cDNA clones (26). Four CTL-specific
cDNA clones have been identified using this procedure and are shown here to code for serine proteases. One clone is identical to H factor, whereas the other three represent novel transcripts. One of these clones has been expressed as a fusion protein. Antibodies prepared against this fusion protein stain granules of CTL and react with two protein bands derived from granules.

Materials and Methods

Cells and Subcellular Fractionation. Methods for isolating and maintaining the cloned Th L2 and the cloned CTL L3 have been described before (27). L3 cells are Thy-1+, LFA-1+, Lyt-2-, L3T4+, and H-2Ld reactive. L2 cells are Mls reactive and are Thy-1+, LFA-1+, Lyt-2-, and L3T4+. To stimulate the cloned T cells, they were resuspended at 10^6-10^7 cells/ml and were cultured with Con A (Pharmacia Fine Chemicals, Piscataway, NJ) either at 10 μg/ml for L2 cells or 2 μg/ml for L3 cells. The specificities of CTLL-A11 and CTLL-R8 have been given elsewhere (28).

EL-4 (mouse thymoma), YAC-1 (mouse leukemia), A20.2j (mouse B cell line), and K46 (mouse B cell line) cells were maintained in RPMI 1640 medium containing 5% FCS. When indicated, EL-4 cells were treated with 12-0-tetradecanoylphorbol-13-acetate (10 ng/ml) for up to 20 h and the stimulation was monitored by IL-2 assay of EL-4 cell supernatant using IL-2-dependent CT6 cells. Murine melanoma cells, Cloudman S-91, were grown in Ham's F-10 medium containing 8% newborn calf serum, 8% NU-serum (Collaborative Research, Lexington, MA), 100 μg/ml penicillin, and 200 U/ml streptomycin.

Large granular lymphocytes (LGL) were obtained from murine spleen cells essentially as described (29). Briefly, splenic mononuclear cells, enriched by Ficoll/Hypaque gradient centrifugation, were subjected to adherence in plastic dishes and passed over nylon wool columns to remove monocytes/macrophages and B cells, respectively. Nylon wool-nonadherent cells were cultured in T flasks at 2 x 10^6 cells/ml in medium containing 1,000 U/ml of rIL-2 (a generous gift of Cetus Corp., Emeryville, CA). After 4 h of incubation at 37°C, adherent cells received separately fresh medium containing 1,000 U/ml of rIL-2. Cells were incubated for 3-5 d, after which they were harvested and used for immunofluorescence or immunoblotting analysis. In a parallel experiment, cells prepared this way were shown to be cytotoxic to YAC-1 targets and were stained positively (>90%) with anti-asialo GM1 antibodies. The rat LGL used here consisted of RNK cells, kindly provided by Dr. P. Henkart (National Institutes of Health, Bethesda, MD).

MLRs were established by using DBA/2 and C57BL/6 splenocytes as stimulating cells and C57BL/6 and CBA/J spleen cells as responders. Stimulating cells were irradiated with 2,000 rad before mixing with responder cells at a ratio of 1:4. MLRs were carried out for 5-7 d at 10^6 cells/ml, after which cells were harvested and used for immunofluorescence.

Cell rupture was done by nitrogen cavitation as described (28). Nucleus-free cell lysates (2 x 10^7 cell equivalents/ml) were subjected to centrifugation at 13,000 g for 20 min in a Sorvall SS-34 centrifuge. The pellet was resuspended to the same original volume with a high-phosphate extraction buffer (28). Both the supernatant and the resuspended pellet were centrifuged at 350,000 g for 1 h using a table-top ultracentrifuge (model TL100; Beckman Instruments, Inc., Fullerton, CA) to sediment membranes. The supernatants were collected and used for immunoprecipitation. In some experiments, cells were biosynthetically labeled with [35S]methionine at 1 mCi/10^7 cells/50 ml of methionine-free medium for 2 h, followed by incubation with 4% normal medium for an additional 12 h.

Construction of cDNA Libraries of CTLL-A11 and L3 Cells and Isolation of CTL-specific cDNA Clones. Cytoplasmic RNA of CTLL-A11 and Con A-stimulated L3 cells was extracted by a guanidine isothiocyanate-cesium chloride gradient centrifugation method (30). Poly(A)^+ mRNA of the L3 and CTLL-A11 cells was purified by chromatography on oligo-dT-cellulose column. dsDNA was synthesized from the poly(A)^+ mRNA. The cDNA was methylated at EcoRI sites to prevent cleavage, and Eco RI linkers were ligated to cDNA. After Eco RI cleavage, cDNA was enriched for molecules >250 kD by passage over 1 ml of Bio-Gel A-150m columns (Bio-Rad Laboratories, Richmond, CA). The cDNAs were inserted into the unique
Eco RI site of both λgt10 and λgt11 bacteriophage cloning vectors (31). We have previously isolated a group of cDNAs that are specific for T cells in contrast with B cells, using both positive and negative differential screening and RNA blot analysis of various lymphoid cells. The T cell-specific cDNAs were further studied to determine whether they were specific for cloned Th L2 or cloned CTL L3 cells and whether they were inducible by Con A or IL-2. A salient feature of the method was to analyze L2 or L3 cDNA species that failed to hybridize to ssDNA probes prepared from closely related cells, and thus allowed us to isolate T cell subset transcripts that exist at a low level (26).

RNA Blot Hybridization. Total cytoplasmic RNA (10 μg) or poly(A)+ mRNA (10 μg) were fractionated on 1.2% agarose-formaldehyde gels and transferred to Gene Screen Plus (New England Nuclear, Boston, MA). Gel-purified inserts of cDNA were 32P-labeled by nick translation and used as probes. Filters were prehybridized and hybridized at 42°C in 50% formamide, 5 × SSC (1 × SSC = 150 mM NaCl, 15 mM sodium citrate, pH 7), 0.1% SDS, 250 μg/ml of salmon sperm DNA, and 10% dextran sulfate. Filters were washed at room temperature for 15 min in 2 × SSC and 0.1% SDS, and at 42°C for 3 min in 0.1 × SSC and 0.1% SDS several times. When a Northern blot was used multiple times for hybridization, the previous probe was removed by treating the membrane in 10 mM Tris-HCl (pH 7.0) and 0.2% SDS at 85°C for 1 h.

Genomic DNA Blot Analysis. High molecular weight DNA of mouse spleens was prepared as described previously (32). Restriction endonuclease digests of DNA were subjected to electrophoresis in 0.8% agarose gel at 4°C. The gel was denatured with 0.5 M NaOH/1 M NaCl and neutralized with 1 M Tris-HCl, pH 8.0, 1 M NaCl. The DNA was transferred to Gene Screen Plus as described by Southern (33). The blot hybridized with 32P-labeled cDNA in a solution composed of 6 × SSC, 5 × Denhardt’s solution, 0.5% SDS, and 100 μg/ml denatured salmon sperm DNA at 65°C for 18 h. The filters were then washed three times at room temperature for 10 min each in 2 × SSC and 0.1% SDS, and two times at 65°C for 45 min each in 0.1 × SSC and 0.1% SDS.

DNA Sequencing DNA restriction fragments, subcloned in M13 vectors (34), were sequenced by the dideoxy chain termination technique (35), with modification made to accommodate 2'-deoxyadenosine 5'-[α-[35S]thio]triphosphate (36). A forward primer (New England Biolabs, Beverly, MA) complementary to the lacZ sequence adjacent to the 5' side of the Eco RI site in λgt11 was used for the direct sequencing of cDNA insert end point in λgt11 (37).

Antibody Production, Affinity Purification, Immunofluorescence, and Immunoblotting. The λgt11 cDNA bearing L3-5 was plated onto Escherichia coli Y1089. Lysogens, selected essentially as described (38), were induced with 20 mM isopropyl β-D-thiogalactopyranoside (IPTG) for 3–4 h. After three washes, bacteria were resuspended to OD600 of one in water. NP-40 was added to 1% and the bacterial suspension was lysed by freezing and thawing (3 ×) followed by extensive sonication using a probe. After sedimentation of the cell debris in a microfuge (10,000 g, 15 min), the supernatant was diluted threefold in PBS and applied to a Sepharose 4B column coupled with rabbit anti-β-galactosidase IgG (CooperBiomedical, Inc., Malvern, PA) at a load of 108 bacteria equivalents to 1 mg of coupled antibody. The column was washed with two cycles of alternate washes of PBS, PBS/0.4 M NaCl, and PBS/1% NP-40, and finally eluted with 4 M guanidine-Cl (pH 7.0). The enriched fusion protein was further purified on a G4000 column (TSK America Inc., North Bend, WA) and eluted at 0.5 ml/min with 0.1 M sodium phosphate buffer, pH 6.5, containing 0.1% SDS. The eluted protein was boiled and a portion of the antigen was crosslinked with 0.5% glutaraldehyde. These antigens were injected into a rabbit through multiple sites in 0.5-mg doses. Booster injections were given every 3 wk and antiserum was collected 4 d after each booster.

The antiserum obtained this way was extensively preadsorbed on induced λgt11 lysogens (without inserts) immobilized onto nitrocellulose membranes (3 ml antiserum/50,000 plaques per 10-cm filter) to remove nonspecific anti-E. coli and anti-β-galactosidase antibodies. Before this adsorption, nitrocellulose membranes were fixed with 10% acetic acid/15% isopropanol and subsequently blocked with 1% dry milk to saturate nonspecific sites. The preadsorbed antiserum was then further enriched by affinity chromatography by incubating this antiserum with blocked nitrocellulose membranes carrying replicas of induced λgt11 L3-5 lysogens prepared as before. Membranes were washed extensively with alternate washes of PBS, PBS/0.4
MOUSE CYTOLYTIC T LYMPHOCYTE SERINE PROTEASE cDNAs

M NaCl, PBS/1% NP-40, and the membrane-bound antibodies were eluted with 0.1 M glycine, pH 2.2, and neutralized immediately with 0.1 vol of 1 M Tris-HCl, pH 8.0. After dialysis against PBS, these antibodies were used for immunofluorescence and immunoprecipitation studies.

Indirect immunofluorescence was carried out with cells sedimented onto slides in the cytocentrifuge that were then fixed with 1% paraformaldehyde for 10 min and permeabilized with acetone (−20°C for 3 min). After blocking with 1% goat serum in PBS, slides were stained with specific IgG (50 μg/ml) as the first antibody and visualized with FITC-conjugated goat anti-rabbit IgG (Fab′)2 (Boehringer Mannheim Biochemicals, Indianapolis, IN; 1:100 dilution). Slides were mounted with glycerol and observed under a Nikon fluorescence microscope.

For immunoprecipitation, subcellular fractions of cells were reacted with affinity-purified anti-L3-5 IgG (at 1 mg/10⁸ cells) for 2 h, followed by incubation with protein A-agarose (Boehringer Mannheim Biochemicals) for an additional 2 h and sedimentation in the microfuge. The immunoprecipitate was resuspended in gel sample buffer containing 1% SDS and boiled for 5 min before its use in gel electrophoresis and autoradiography. Gel electrophoresis was performed on 10–20% gradient gels according to Laemmli (39). After being developed at constant current, the gels were soaked in ENHANCE (New England Nuclear) for 1 h, dried under vacuum, and subjected to autoradiography for 5–7 d.

[^3]H[DFP] Labeling of Proteins. [^3]H[diisopropyl fluorophosphate ([^3]H[DFP]) labeling was performed essentially as described before (13). Briefly, to 100–200-μl protein samples, a reaction mixture of 1 M Tris-HCl buffer (pH 7.5), containing 1 mM[^3]H[DFP] (5 μCi/μmol; New England Nuclear) was added to 10% final volume. After 60 min at 37°C, the reaction was terminated by addition of boiling buffer (2x) containing 2% SDS and 50 mM DTT, followed by additional boiling in Neville’s buffer for 5 min. Samples were then applied directly to gel slabs and processed as before.

Results

Isolation of CTL-specific Serine Esterase cDNAs. We have recently isolated T lymphocyte subset-specific cDNAs from Th and CTL by means of a modified differential screening procedure (26). The specific expression of 21 potential CTL-specific cDNAs was further studied by Northern blot analysis of RNA from K46 B cells, unstimulated or TPA-stimulated EL-4 thymoma, unstimulated or Con A-stimulated L2 Th, unstimulated or Con A-stimulated L3 CTL, and rat LGL cells. Of these, eight inserts produced specific hybridization signals with L3 RNA while 13 inserts failed to show hybridization signal with any RNA used. We reasoned that the failure of the 13 inserts to hybridize with any RNA tested was probably due to the low expression of corresponding transcripts. Subsequently, nucleotide sequence analysis revealed that one clone among the eight inserts, L3G10 no. 6, was identical to H factor (22), while one clone among the latter 13 inserts, designated L3-1, represented a novel transcript with ~64% sequence homology to CCP1 (23).

L3G10 no. 6 is expressed in unstimulated and Con A-stimulated L3 cloned CTLs, but not in L2 Th, K46 B cells, EL-4 thymoma cells, or NK-like rat LGL (Fig. 1 A). L3G10 no. 6 is also expressed at a high level in another CTL line, CTLL-All (Fig. 1 B). On the other hand, L3-1 fails to detect mRNA in any of the RNA used in Fig. 1 A but detects 900 base mRNAs in CTLL-All cells (Fig. 1 C). L3-7 probe (Fig. 1 D) is an anonymous cDNA from L3 cDNA library and was used to demonstrate that each lane of the RNA blot used for Fig. 1 B and C contains similar amounts of RNA. For Fig. 1, B–D, the same filter was used successively in that order. These results appear to confirm the CTL specificity of L3-1 cDNA. Its level of expression is undetectable in L3 cells although the clone is originated from these same cells,
an observation that illustrates the usefulness of our differential screening procedure in detecting gene transcripts present in cells at low levels.

L3-1 contained a cDNA insert of 281 bp and was used to screen λgt11 L3 cDNA library to obtain a full-length version of L3-1. We obtained a cDNA clone L3-5 containing a 635-bp cDNA insert excluding the poly(A) tail. L3-1 and L3-5 shared 221 bp.

Fig. 2 A shows the nucleotide sequence and deduced amino acid sequence of the longest open reading frame derived from L3-1 and L3-5 cDNA inserts. The sequence was designated a mouse CTL serine protease 1 (MCSP-1). Comparison between MCSP-1 and the CCPl/CTLA1 sequences revealed that MCSP-1 potentially did not contain ~30 amino acids equivalent in the 5' coding sequence of the transcript. Therefore, a cDNA library prepared from CTLL-A11 mRNA was screened using L3-1 probe to obtain a full-length version of L3-1. In this screening, we isolated many cDNA clones that hybridized to L3-1 probe. Partial nucleotide sequencing of the 14 new cDNA inserts revealed that none of the 14 clones was identical to MCSP-1. The data suggest that MCSP-1 expression may also be very low in CTLL-A11 cells and that the hybridizing band seen in CTLL-A11 RNA (Fig. 1 C) is due to a cross-hybridization to L3-1-related sequences. Among the 14 clones, five cDNA clones whose insert sizes were >700 bases were analyzed further. The five clones represented two other related sequences, designated as MCSP-2 and MCSP-3.

**Nucleotide and Deduced Amino Acid Sequences of the Three New Serine Esterase cDNAs.** Fig. 2, A–C shows the nucleotide and deduced amino acid sequences of MCSP-1, MCSP-2, and MCSP-3. The homology of nucleotide sequence among the three cDNAs is ~64% and the nucleotide changes are dispersed along the cDNA sequence. Each open reading frame of MCSP-2 and MCSP-3 code for a potential mature polypeptide of 228 amino acids.
MOUSE CYTOLYTIC T LYMPHOCYTE SERINE PROTEASE cDNAs

A

1  GCA GGC TTC TTC TGT GTA CAA GAT TTT CCC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA
2  Gly Gly Phe Leu Val Glu Asp Phe Gly Leu Thr Ala His Cys Arg Asn Arg Ser

3  AAT ACA CTC ACA CTC GGG CCC AAG TTC AAC CCC GTT GCA
4  Met Thr Val Leu Gly Asp Ala Arg Val Arg Cys Arg Thr Ser Ala Gly Val

5  CCT GTG CCA AAA GGC ATT CCC CAT CCA GCT TTT ATT AAG ACA ACT GCC ACC AAT GAC ATT
6  Pro Val Ala Leu Ala Pro His Pro Ala Phe Asn Arg Lys His Thr Asn Arg Ile

7  181  ATG TTA TTA AAA CTC GAG AGT AAG AGA ACT AAA GCT GTC AGC CCT AGC TCA
8  Met Leu Leu Leu Leu Ser Ala Tyr Ala Arg Asp Asa Ala Gln Thr Val Thr

9  502  CAC CTG GCA GAC TGC GCC GAG CTG CAC ACC TCA GAG GCT
10  Arg Val Ala Glu Glu Val Arg Asp Asn Ser Ala Glu Gly Val

11  140  Met Pro Pro Ile Leu Leu Leu Met Leu Leu Leu Pro Leu Arg Ala Gly Ala Gly

12  321  GAA GGC TTC TTC TGT GTA CAA GAT TTT CCC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA
13  Glu Gly Phe Leu Val Glu Asp Phe Gly Leu Thr Ala His Cys Arg Asn Arg Ser

14  601  Met Thr Val Leu Gly Asp Ala Arg Val Arg Cys Arg Thr Ser Ala Gly Val

15  120  CCT GTG CCA AAA GGC ATT CCC CAT CCA GCT TTT ATT AAG ACA ACT GCC ACC AAT GAC ATT
16  Pro Val Ala Leu Ala Pro His Pro Ala Phe Asn Arg Lys His Thr Asn Arg Ile

17  240  ATG TTA TTA AAA CTC GAG AGT AAG AGA ACT AAA GCT GTC AGC CCT AGC TCA
18  Met Leu Leu Leu Leu Ser Ala Tyr Ala Arg Asp Asa Ala Gln Thr Val Thr

19  481  GAC AAC CTA CTA GTT GGA GCT TCT AGC TCT GCA GAT GCT GCT TCA
20  Asp Asp Ala Ser Gly Val Gly Ser Ser Ser Ala Glu Val Gly

21  640  Pro Arg Pro Ala Arg Val Lys Pro Gly Asp Ser Val Ala Gly Trp

22  811  ATC ATC AAC CCC CAC TCA TCA GCA GCC TAC GCA TCC TCA ATC
23  Ile Ile Asn Cys Cys Ser Ser Ala Ala Ser Ala Glu Ser Ser Ala Glu

24  981  GAG ATC ATG CTG TTA AAG CTG TTA
25  Glu Ile Met Leu Lys Ala Ala Ala

26  1062  CCA GCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA

28  1162  GCA GGC TTC TTC TGT GTA CAA GAT TTT CCC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA
29  Glu Gly Phe Leu Val Glu Asp Phe Gly Leu Thr Ala His Cys Arg Asn Arg Ser

B

1  | CTG CCA GCT GGA GCA GAG GAG |
2  | Leu Gly Ala Gly Ala Gly Ala |

3  181  ATC ATC AAC CCC CAC TCA TCA GCA GCC TAC GCA TCC TCA ATC
4  Ile Ile Asn Cys Cys Ser Ser Ala Ala Ser Ala Glu Ser Ser Ala Glu

5  502  CAC CTG GCA GAC TGC GCC GAG CTG CAC ACC TCA GAG GCT
6  Arg Val Ala Glu Glu Val Arg Asp Asn Ser Ala Glu Gly Val

7  140  Met Pro Pro Ile Leu Leu Leu Met Leu Leu Leu Pro Leu Arg Ala Gly Ala Gly

8  321  GAA GGC TTC TTC TGT GTA CAA GAT TTT CCC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA
9  Glu Gly Phe Leu Val Glu Asp Phe Gly Leu Thr Ala His Cys Arg Asn Arg Ser

10  601  Met Thr Val Leu Gly Asp Ala Arg Val Arg Cys Arg Thr Ser Ala Gly Val

11  120  CCT GTG CCA AAA GGC ATT CCC CAT CCA GCT TTT ATT AAG ACA ACT GCC ACC AAT GAC ATT
12  Pro Val Ala Leu Ala Pro His Pro Ala Phe Asn Arg Lys His Thr Asn Arg Ile

13  240  ATG TTA TTA AAA CTC GAG AGT AAG AGA ACT AAA GCT GTC AGC CCT AGC TCA
14  Met Leu Leu Leu Leu Ser Ala Tyr Ala Arg Asp Asa Ala Gln Thr Val Thr

15  481  GAC AAC CTA CTA GTT GGA GCT TCT AGC TCT GCA GAT GCT GCT TCA
16  Asp Asp Ala Ser Gly Val Gly Ser Ser Ser Ala Glu Val Gly

17  640  Pro Arg Pro Ala Arg Val Lys Pro Gly Asp Ser Val Ala Gly Trp

18  811  ATC ATC AAC CCC CAC TCA TCA GCA GCC TAC GCA TCC TCA ATC
19  Ile Ile Asn Cys Cys Ser Ser Ala Ala Ser Ala Glu Ser Ser Ala Glu

20  981  GAG ATC ATG CTG TTA AAG CTG TTA
21  Glu Ile Met Leu Lys Ala Ala Ala

22  1062  CCA GCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA

24  1162  GCA GGC TTC TTC TGT GTA CAA GAT TTT CCC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA
25  Glu Gly Phe Leu Val Glu Asp Phe Gly Leu Thr Ala His Cys Arg Asn Arg Ser

C

-150  COTTON PEGGY MARGARET DIANNE ANNE CHADWICK SIMON KIRKLEY GERTRUDE AGNES ROGER CHANDLER PATRICIA CYNTHIA GEORGE CINNAMON

1  ATC CCA CCA ATC ATT CTC GAG ATT ACC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA |
2  Met Thr Val Leu Gly Asp Ala Arg Val Arg Cys Arg Thr Ser Ala Gly Val

3  601  Met Thr Val Leu Gly Asp Ala Arg Val Arg Cys Arg Thr Ser Ala Gly Val

4  120  CCT GTG CCA AAA GGC ATT CCC CAT CCA GCT TTT ATT AAG ACA ACT GCC ACC AAT GAC ATT
5  Pro Val Ala Leu Ala Pro His Pro Ala Phe Asn Arg Lys His Thr Asn Arg Ile

6  240  ATG TTA TTA AAA CTC GAG AGT AAG AGA ACT AAA GCT GTC AGC CCT AGC TCA
7  Met Leu Leu Leu Leu Ser Ala Tyr Ala Arg Asp Asa Ala Gln Thr Val Thr

8  481  GAC AAC CTA CTA GTT GGA GCT TCT AGC TCT GCA GAT GCT GCT TCA
9  Asp Asp Ala Ser Gly Val Gly Ser Ser Ser Ala Glu Val Gly

10  640  Pro Arg Pro Ala Arg Val Lys Pro Gly Asp Ser Val Ala Gly Trp

11  811  ATC ATC AAC CCC CAC TCA TCA GCA GCC TAC GCA TCC TCA ATC
12  Ile Ile Asn Cys Cys Ser Ser Ala Ala Ser Ala Glu Ser Ser Ala Glu

13  981  GAG ATC ATG CTG TTA AAG CTG TTA
14  Glu Ile Met Leu Lys Ala Ala Ala

15  1162  GCA GCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA CCA

17  12345  ATC CCA CCA ATC ATT CTC GAG ATT ACC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA |
18  Met Thr Val Leu Gly Asp Ala Arg Val Arg Cys Arg Thr Ser Ala Gly Val

19  321  GAA GGC TTC TTC TGT GTA CAA GAT TTT CCC CTG AGC OCT GCT CAC TGC AGA AAC AGC TCA |
20  Glu Gly Phe Leu Val Glu Asp Phe Gly Leu Thr Ala His Cys Arg Asn Arg Ser

21  601  Met Thr Val Leu Gly Asp Ala Arg Val Arg Cys Arg Thr Ser Ala Gly Val

22  120  CCT GTG CCA AAA GGC ATT CCC CAT CCA GCT TTT ATT AAG ACA ACT GCC ACC AAT GAC ATT |
23  Pro Val Ala Leu Ala Pro His Pro Ala Phe Asn Arg Lys His Thr Asn Arg Ile

24  240  ATG TTA TTA AAA CTC GAG AGT AAG AGA ACT AAA GCT GTC AGC CCT AGC TCA |
25  Met Leu Leu Leu Leu Ser Ala Tyr Ala Arg Asp Asa Ala Gln Thr Val Thr

26  481  GAC AAC CTA CTA GTT GGA GCT TCT AGC TCT GCA GAT GCT GCT TCA |
27  Asp Asp Ala Ser Gly Val Gly Ser Ser Ser Ala Glu Val Gly

28  640  Pro Arg Pro Ala Arg Val Lys Pro Gly Asp Ser Val Ala Gly Trp

29  811  ATC ATC AAC CCC CAC TCA TCA GCA GCC TAC GCA TCC TCA ATC |
30  Ile Ile Asn Cys Cys Ser Ser Ala Ala Ser Ala Glu Ser Ser Ala Glu

Figure 2. Legend appears on opposing page.
FIGURE 2. Nucleotide sequence and the deduced amino acid sequence of MCSP-1 (A), MCSP-2 (B), and MCSP-3 (C). The nucleotide sequences of message strands are numbered in 5' to 3' direction. The amino acids forming a serine esterase active site (Δ) are His45, Asp9, and Ser184. The potential asparagine-linked glycosylation sites are underlined. The putative active site pocket (*) is at residue Ala in MCSP-1, Thr in MCSP-2, and Ser in MCSP-3. The AATAAA polyadenylation consensus sequence is boxed in the 3' noncoding region in MCSP-1 and MCSP-3 sequences.

Stop codon (---). An arrow indicates a putative cleavage site for active enzyme.

The molecular weights corresponding to the mature protein backbones of MCSP-2 and MCSP-3 are 25,477 and 25,360, respectively. There are three potential glycosylation signals in MCSP-1 and MCSP-3, and four in MCSP-2 as underlined. The putative active site pocket (Δ) is at residue Ala in MCSP-1, Thr in MCSP-2, and Ser in MCSP-3. The AATAAA polyadenylation consensus sequence is boxed in the 3' noncoding region in MCSP-1 and MCSP-3 sequences.

Stop codon (---). An arrow indicates a putative cleavage site for active enzyme.

Since the predicted amino acid sequence for MCSP-1 is substantially different from other reported T cell serine proteases, the amino acid sequence of MCSP-1 was compared with other serine protease-like proteins whose active site pocket residue is alanine. Fig. 3 B shows a comparison of predicted amino acids of MCSP-1 with those of CCPI (23), granzyme C (25), RMCPII (43, 44) and cathepsin G (45). Approximately 33% of the amino acids are identical in all five proteins (Fig.3 B, boxed amino acids). The amino acid identity was 56% between MCSP-1 and CCPI, 58% between MCSP-1 and granzyme C, 42% between MCSP-1 and RMCPII, and 46% between MCSP-1 and cathepsin G. The serine protease triads are conserved
Figure 3. Legend appears on opposing page.
Figure 3. Alignment of predicted amino acid sequences of T-cell serine proteases. Comparisons were made among MCSP-1, MCSP-2, and MCSP-3 deduced amino acid sequences (A), and MCSP-1 amino acids were aligned with four other predicted serine proteases whose active site pocket residue is alanine (B). An arrow indicates a putative cleavage site to generate an active enzyme. (A) The "charge-relay" system of the active site of serine proteases; (*) the potential amino acid residue that participates in the primary substrate binding site in serine proteases. The numbering begins at the NH$_2$ terminus of the predicted active enzyme. Gaps were introduced to optimize the alignment. Amino acids that are identical among the sequences are boxed. Numbers in parentheses indicate the positions of amino acids. (A) The underlined NH$_2$-terminal amino acids of MCSP-2 and MCSP-3 are identical to those of granzymes E and F, respectively. (B) Positions that the amino acids of MCSP-1 are different from those of other four proteins (+).
in similar space in all five proteins (Fig. 3 B, indicated by Δ). However, ∼30% of amino acids of MCSP-1 (60 positions indicated by + in Fig. 3 B) are different from those of the other four proteins. These data may indicate that MCSP-1 is a new member of T cell serine protease family.

Southern Blot Analysis. Southern blot analysis of mouse (C57BL/6) genomic DNA was performed with MCSP-1, MCSP-2, and MCSP-3 clones (Fig. 4, A–C, respectively). All three cDNAs detect the same strongly hybridizing bands but detect weakly hybridizing bands differentially. For example, when the mouse DNA is digested with Bam H1, all three cDNAs detect three strong bands whose sizes are 10.9, 7.4, and 2.5 kb. In addition, there are three weakly hybridizing bands whose sizes are 6.0, 5.0, and 2.6 kb. Of these, three weak bands are detected by MCSP-3 but only two bands each hybridize to MCSP-1 and MCSP-2 with different intensity. A similar phenomenon is observed in Eco RI and Hind III–digested genomic DNA (Fig. 4). These hybridization data and sequence information suggest that there is a family of genetic loci that contains separate genes representing these three cDNAs. Their exon and intron composition must be similar.

Antibodies Derived Against Fusion Protein Carrying MCSP1 React with Granule Proteins. Direct sequence analysis with a λgt11 forward primer revealed that L3-5 cDNA was in frame with the lacZ gene of the λgt11 vector. Thus, the L3-5 clone was used to prepare antibodies against the fusion protein.

Lysogens harboring the L3-5 clone were induced with IPTG. The fusion protein was purified by immunoaffinity chromatography using agarose-bound β-galactosidase–specific mAbs as the immunoadsorbent. A rabbit antiserum raised against this fusion protein was further enriched for MCSP-1-specific antibodies through a differential adsorption method detailed in Materials and Methods. Affinity-purified

**Figure 4.** Southern blot analysis of mouse DNA. Genomic DNA from C57BL/6 liver was digested with Hind III (H), Bam HI (B), or Eco RI (E), electrophoresed on a 0.8% agarose gel, transferred to Gene Screen Plus, and hybridized to MCSP-1 (A), MCSP-2 (B), and MCSP-3 (C). (M) Molecular size marker in kilobase pairs.
antibodies were then used to stain various cell lines by indirect immunofluorescence. CTLL-A11, CTLL-1, and CTLL-R8, but not EL-4 thymoma or YAC-1 leukemia cells, were stained with these antibodies (not shown). Granulated staining patterns were always observed, suggesting the predominant granular localization of the MCSP-1 protease (Fig. 5). In parallel experiments no reactivity was given by the preimmune serum control. To assess the distribution of this antigen in vivo, anti-MCSP-1 antibodies were next used to stain bulk cell populations from murine spleen, liver, and peripheral blood. Staining was persistently negative for all the cell populations examined. However, lymphocyte populations that had been stimulated with IL-2 for 3–7 d displayed weakly positive stain in their granules. Positive staining was associated mainly with a subpopulation of LGL that had been enriched by their selective adherence to plastic dishes (not shown; see Materials and Methods). Nonadherent cells did not show any detectable reactivity. In parallel experiments murine primary CTL (3–5 d old), obtained from bulk spleen lymphocyte populations by MLRs, also displayed faint reactivity to the antibody as determined by immunofluorescence. These results suggest that MCSP-1-related antigens are induced to measurable levels in both alloimmune-stimulated CTL and IL-2-driven LGL/NK cells. Since L3-1 (a murine MCSP-1-related transcript) did not hybridize with the mRNA of rat tumor LGL (Fig. 1A), these results would argue that either there are some significant differences between the mouse and the rat transcripts that would restrict crosshybridiza-

**Figure 5.** Indirect immunofluorescence using anti-L3-5 fusion protein antibodies. CTLL-A11 cells, sedimented in the cyt centrifuge, were stained with affinity-purified antibodies prepared against L3-5 β-galactosidase fusion protein. (A) Phase-contrast; (B) same field, fluorescence; x600. The preimmune control serum produced undetectable reactivity (not shown).
MOUSE CYTOLYTIC T LYMPHOCYTE SERINE PROTEASE cDNAs

**Figure 6.** Immunoprecipitation of CTL proteins with anti-L3-5 fusion protein antibodies and [3H]DFP labeling. (Lanes 1-3) Immunoprecipitation of [35S]methionine-labeled CTLL-All (lanes 1 and 2) and CTLL-R8 cells (lane 3). Lanes 2 and 3 contain proteins from granule-enriched pellet that have been reacted with affinity-purified anti-L3-5 fusion protein IgG (see Materials and Methods). The precipitation profile corresponds to material derived from 10^8 cells. (Lane 1) Precipitation was carried out with granule-depleted cytosol under exact conditions as described above; material also corresponds to 10^8 cell equivalents. The autoradiogram was exposed for 5 d. (Lane 4) [3H]DFP labeling of immunoprecipitated proteins. Immunoprecipitation of CTLL-All granule-enriched proteins was carried out exactly as in lane 2 except that cells were not labeled. The immunoprecipitate was reacted with [3H]DFP as described in Materials and Methods and loaded onto the gel. Note that only the 35-40-kD protein band labels strongly with [3H]DFP. Autoradiography was performed for 3 d.

**Discussion**

To isolate mRNA species, including rare transcripts specific to a cell type, we developed a protocol for differential screening without prior selection and applied the approach to the analysis of cDNA libraries from Con A-stimulated helper (L2) and CTL (L3) cells. In the above analysis, we could not detect transcripts representing 30-40% of cDNA inserts from L2 and L3 cells. We speculated that those inserts could still be subset specific, but that the L2 and L3 RNA loads used in the RNA blot analysis were not high enough to detect a basal level of these RNA species. In fact, one of the cDNA inserts, L3-1, corresponded to a CTL-specific serine protease.
that was originated from L3 cells but whose RNA expression in these cells was too low to detect. On the other hand, the expression of the MCSP-1-related cDNAs, MCSP-2 and MCSP-3, was very high in another CTL clone, CTLL-All. Since we could not clone MCSP-1 cDNA from the CTLL-All cDNA library, this sequence must not be expressed at high level in the CTLL-All either. A 23-mer oligo nucleotide matching the 5' end of L3-1 sequence was used to prime cDNA synthesis on poly (A)+ mRNA from CTLL-All cells (data not shown). The primer extension products were undetectable in many repeated experiments, which was another indication that L3-1 mRNA level was low in CTLL-All cells. Notably, CTLL-All expressed H factor/CTLA 3, MCSP-2, and MCSP-3 at high levels, whereas L3 cells expressed H factor/CTLA 3 only at high level, while MCSP-1, MCSP-2, and MCSP-3 transcripts were expressed at undetectable levels, suggesting marked heterogeneity in the expression of serine esterase messages even among different CTL clones.

Based upon the identity of NH2-terminal amino acids of MCSP-2 and MCSP-3 with the reported partial sequence of the so-called granzymes E and F (15), we assigned MCSP-2 as granzyme E and MCSP-3 as granzyme F. MCSP-1 differs from the published sequences of granzymes A-F (15). Masson and Tschopp (14) observed that granzymes D-F are closely related in that they all react with antibodies to granzyme D and are similar in their degree of glycosylation and the position of at least one tryptophan residue in their polypeptide chains. As shown in Fig. 3 A, there are indeed two tryptophans that appear at the same positions in all three deduced protein sequences presented here. Antibodies raised against L3-5 β-galactosidase fusion proteins react strongly with a 35-40-kD protein, which corresponds to the presumed molecular mass of granzyme D. It should be noted, however, that the number of potential N-glycosylation sites are three for MCSP-1 instead of five for granzyme D and that the active site pocket residue is alanine for MCSP-1 instead of threonine for granzyme D. Therefore, MCSP-1 protein may be a seventh member of murine T cell serine protease family.

With the completion of this work, six closely related serine esterase transcripts (granzymes A, B, C, E, F, and MCSP-1) will have been cloned and sequenced, fully supporting the notion that a family of serine esterases is found in lymphocyte granules. However, the function of these enzymes still remains to be determined (46). The three cDNAs described here were expressed only in CTL in our limited survey (Fig. 1, A and B). These transcripts may not be unique to CTL since at least one of them (MCSP-1 or a closely related product) was also expressed weakly by LGL/NK cells established in primary cultures. Studies with a large number of cell lines and more specific probes will be required to address vigorously the cell and tissue distribution of these cloned transcripts.

Serine esterases may be used as sensitive markers for lymphocytes that have been committed to become killer cells (10, 46). It is intriguing that many of the transcripts that have been cloned so far from subtractive CTL-specific cDNA libraries should turn out to code for serine proteases. T cell subsets may acquire such transcripts only upon activation via IL-2 or allo-immune stimulation. Serine esterase-specific probes could in principle be used to study the role of killer lymphocytes in health and in disease, particularly in the assessment of killer lymphocyte distribution in tissues afflicted with auto-immune disorders. The development of more specific nucleic acid probes and peptide-specific antisera for the various serine esterases should facilitate further studies on their function and expression.
MOUSE CYTOLYTIC T LYMPHOCYTE SERINE PROTEASE cDNAs

Summary

Three new cDNA clones (designated MCSP-1, MCSP-2, and MCSP-3) encoding mouse serine proteases were isolated from cloned cytolytic T lymphocytes (CTL) by a modified differential screening procedure. The putative mature proteins of MCSP-2 and MCSP-3 are each composed of 228 amino acids with molecular weights of 25,477 and 25,360, respectively. NH₂-terminal amino acids of MCSP-2 and MCSP-3-predicted proteins were identical to those reported for granzyme E and F, respectively. The third species, MCSP-1, was closely related to the two other cDNA species but ~30 amino acids equivalents of the NH₂-terminal portion of the cDNA were not cloned. The amino acids forming the active sites of serine proteases were well conserved among the three predicted proteins. The active site pocket residue positioned six residues before the active-site Ser¹⁸⁴ is alanine in MCSP-1, threonine in MCSP-2, and serine in MCSP-3, indicating that both MCSP-2 and MCSP-3 may have chymotrypsin-like specificity. There are three potential asparagine-linked glycosylation sites in MCSP-1 and MCSP-3, and four in MCSP-2-deduced amino acid sequences. Amino acid comparison of MCSP-1 with four other reported serine proteases whose active site pocket residue is alanine revealed that MCSP-1 was substantially different from the other molecules, indicating that MCSP-1 may be a new member of mouse T cell serine protease family. Antibodies made against a MCSP-1 lacZ gene fusion protein stain granules of CTL and react on immunoblots with [³H]DFP. Since a protease cascade may play a key role in cytolytic lymphocyte activation, our isolation of cDNAs representative of unique serine esterases should help to investigate such a cascade process.

Received for publication 20 July 1988.

Note added in proof: Jenne et al. (47) reported cDNA sequence for granzyme D, E, and F. Amino acid comparison revealed that MCSP-2 was identical to granzyme E except for four amino acid changes at positions 85 (Ala-Ile), 87 (Phe-Tyr), 112 (Lys-Arg), and 130 (Ser-Pro), and that MCSP-3 was identical to granzyme F at all positions. MCSP-1 was different from granzyme D sequence.

References

MOUSE CYTOLYTIC T LYMPHOCYTE SERINE PROTEASE cDNAs


