

- Clure, *Nucleic Acids Res.* **9**, 5493 (1981)], and the replacement of the BioGel A-50 m column with 1% to 2% agarose horizontal gel electrophoresis, both to remove the excess Eco RI linkers and to size-fractionate double-stranded cDNA. Spermine precipitation simplified the protocol and increased the yield, allowing all reactions to be performed in one siliconized tube. The agarose horizontal gel electrophoresis gave simple, reproducible separation, permitting visualization of the DNA with ethidium bromide. The double-stranded cDNA's were selected initially for lengths greater than 0.5 kb and subsequently for lengths greater than 0.95 kb. The selected agarose slices were electroeluted in dialysis bags [H. O. Smith *Methods Enzymol.* **65**, 371 (1980)] and precipitated with spermine. All RNA's for the cDNA libraries, Northern blots, and S1 analysis were prepared by guanidium thiocyanate extraction [J. M. Chirgwin, A. E. Przybyla, R. J. MacDonald, W. M. Rutter, *Biochemistry* **18**, 5294 (1979)] and poly(A) selected with oligo(dT)-cellulose. The 1E4 and AR1 killer cell lines were harvested and extracted 4 days after stimulation, at which time their cytotoxicity was confirmed.
11. The probe was prepared as described (32) and the cDNA libraries were plated at a density of ~3000 plaque-forming units per 150-cm plate as described by Huynh *et al.* (9).
 12. T. P. St. John and R. W. Davis, *Cell* **16**, 443 (1979).
 13. C. G. Lobe, B. B. Finlay, W. Paranchych, V. H. Paetkau, R. C. Bleackley, *Science* **232**, 858 (1986).
 14. R. B. Herberman and H. T. Holden, *Adv. Cancer Res.* **27**, 305 (1978).
 15. C. W. Reynolds, E. W. Bere, J. M. Ward, *J. Immunol.* **132**, 534 (1984); C. W. Reynolds *et al.*, *J. Exp. Med.* **161**, 1249 (1985).
 16. G. Spangrude and T. Mossman, unpublished data.
 17. J. Messing, *Methods Enzymol.* **101**, 20 (1983); S. Henikoff, *Gene* **28**, 351 (1984).
 18. B. S. Hartley, *Philos. Trans. R. Soc. London Ser. B* **257**, 77 (1970); J. Kraut, *Annu. Rev. Biochem.* **46**, 331 (1977).
 19. C. L. Young, W. C. Barker, C. M. Tomaselli, M. O. Dayhoff, in *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, DC, 1978), vol. 5 (Suppl. 3), p. 73. The four variant residues in HF, namely Ala²⁷, Phe⁵⁰, Thr¹⁴³, and Ser¹⁸⁵ are equivalent to chymotrypsin's Gly⁴³, Val⁶⁶, Pro¹⁶¹, and Gly¹⁹⁷, respectively.
 20. G. I. Bell *et al.*, *J. Biol. Chem.* **259**, 14265 (1984); G. H. Swift *et al.*, *ibid.*, p. 14271; C. S. Craik *et al.*, *ibid.*, p. 14255; A. J. Mason, B. A. Evans, D. R. Cox, J. Shine, R. I. Richards, *Nature (London)* **303**, 300 (1983).
 21. K. A. Walsh, in *Proteases and Biological Control*, E. Reich, D. B. Rifkin, E. Shaw, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1975), p. 1; K. B. M. Reid and R. R. Porter, *Annu. Rev. Biochem.* **50**, 433 (1981); E. W. Jones, *Annu. Rev. Genet.* **18**, 233 (1984); R. G. Woodbury and H. Neurath, *FEBS Lett.* **114**, 189 (1980).
 22. R. R. Porter, P. J. Lachman, K. B. M. Reid, *Philos. Trans. R. Soc. London Ser. B* **306**, 277 (1984).
 23. J. B. Graham, E. S. Barrow, H. M. Reisner, C. Edgell, *Adv. Hum. Genet.* **13**, 1 (1983).
 24. D. Redelman and D. Hudig, *J. Immunol.* **124**, 870 (1980); T. Chang and H. N. Eisen, *ibid.*, p. 1028; D. Hudig, T. Haverly, C. Fulcher, D. Redelman, J. Mendelsohn, *ibid.* **126**, 1569 (1981); M. S. Pasternack, M. V. Sitkovsky, H. N. Eisen, *ibid.* **131**, 1477 (1983); E. W. Ades *et al.*, *Immunology* **48**, 1 (1983); S. C. Wright and B. Bonavida, in *Natural Killer Activity and Its Regulation*, T. Hoshino, H. S. Koren, A. Uchida, Eds. (Excerpta Medica, Amsterdam, 1984), p. 145.
 25. V. B. Hatcher, M. S. Oberman, G. S. Lazarus, A. Grayzel, *J. Immunol.* **120**, 665 (1978).
 26. M. S. Pasternack and H. N. Eisen, *Nature (London)* **314**, 743 (1985).
 27. S. L. Swain, G. Dennert, S. Wormley, R. W. Dutton, *Eur. J. Immunol.* **11**, 175 (1981); D. Vidovic, A. Juretic, Z. A. Nagy, J. Klein, *ibid.*, p. 499; J. P. Tite and C. A. Janeway, *ibid.* **14**, 878 (1984).
 28. J. L. Sullivan, K. S. Byron, F. E. Brewster, D. T. Purtilo, *Science* **210**, 543 (1980); J. L. Sullivan, *Adv. Pediatr.* **30**, 365 (1983); D. T. Purtilo, L. Paquin, D. DeFlorio, F. Virzi, R. Sukhija, *Semin. Hematol.* **16**, 309 (1979).
 29. B. S. Hartley, *Proc. R. Soc. London Ser. B* **205**, 443 (1979); H. Neurath, *Science* **224**, 350 (1984); R. F. Doolittle, *Trends Biochem. Sci.* **10**, 233 (1985); L. Parthy, *Cell* **41**, 657 (1985).
 30. K. McGowan, C. F. Denek, G. M. Thorne, S. L. Gorbach, *J. Infect. Dis.* **146**, 616 (1982); J. D. Young, T. M. Young, L.-P. Lu, J. L. Unkeless, Z. A. Cohn, *J. Exp. Med.* **156**, 1677 (1982); C. Gitler, E. Calcif, I. Rosenberg, *Philos. Trans. R. Soc. London Ser. B* **307**, 73 (1984).
 31. G. K. McMaster and G. G. Carmichael, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 4835 (1977).
 32. J. Meinkoth and G. Wahl, *Anal. Biochem.* **138**, 267 (1984).
 33. J. Favaloro, R. Treisman, R. Kameh, *Methods Enzymol.* **65**, 718 (1980).
 34. C. G. Lobe, C. Havele, R. C. Bleackley, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
 35. T. Mossman, H. Cherwinski, M. W. Bond, M. Giedlin, R. Coffman, *J. Immunol.* **136**, 2348 (1986).
 36. We thank E. Pillemmer for the gift of 10⁹ active AR1 and 1E4 T killer cells; T. St. John for many helpful discussions; C. Reynolds for the rat NK cell leukemia RNA; R. Joho for provocative conversations in the larval stages of this work; T. Mossman for providing the T-helper cell lines; G. Spangrude for the Con A-stimulated cytotoxic T-cell lines and the lectin-dependent cytotoxicity studies; G. Griffiths, A. Schmitt-Verhulst, and C. Muller-Sieburg for critical reading of the manuscript; J. Boleslaw Carlson for innumerable edifying discussions; our Stanford colleagues for advice, DNA, and intellectual stimulation; and J. Mason for expeditious preparation of the manuscript. We also thank C. Lobe and R. C. Bleackley for the RNA blot in Fig. 2A and for generously sharing their results prior to publication. We acknowledge the National Biomedical Library for the alignment in Fig. 4. Supported by PHS grant AI 19512 and by a predoctoral fellowship (CA09032) in the Stanford Cancer Biology Program (to H.K.G.).

25 October 1985; accepted 25 February 1986

Novel Serine Proteases Encoded by Two Cytotoxic T Lymphocyte-Specific Genes

CORRINNE G. LOBE, B. BRETT FINLAY, WILLIAM PARANCHYCH, VERNER H. PAETKAU, R. CHRIS BLEACKLEY

Genes that are expressed exclusively in cytotoxic T cells should encode proteins that are essential for target cell lysis in cell-mediated immune responses. The sequences of two cytotoxic T lymphocyte-specific complementary DNA's (cDNA's) suggest that the two genes encode serine proteases. A full-length cDNA corresponding to one of the genes was isolated and sequenced. The predicted protein resembles serine proteases in that it includes all the residues that form the catalytic triad of the active site of serine proteases. Moreover, it has sequence characteristics thought to occur only in rat mast cell protease type II. These results are in accord with the view that a protease cascade plays a key role in cytotoxic T-cell activation.

CYTOTOXIC T LYMPHOCYTES (CTL's), also referred to as T killer cells, are effector cells in cell-mediated immune reactions. They specifically recognize foreign antigens on the surface of target cells, bind to them, and cause the target cells to lyse. Although the various steps in this process have been analyzed in considerable detail (1, 2), most studies have not provided insight into the mechanism by which the killer cell effects the lysis of a target cell. A means of identifying relevant molecules, based on cloning CTL-specific genes, has been described (3). The transcripts corresponding to two of these genes (B10 and C11) were detected exclusively in activated CTL's (4). Moreover, the kinetics of messenger RNA (mRNA) expression, as detected by these two cloned probes, closely paralleled but preceded cytotoxicity throughout cytotoxic responses *in vitro* (3).

Sequence analysis (5) of B10 and C11 (Fig. 1A) revealed that they were related to each other and that the hypothetical proteins they encode contain a short region characteristic of serine proteases, Asp-Ser-Gly-Gly (a sequence homologous to that surrounding Ser¹⁹⁵ of chymotrypsin). With B10 and C11 as probes, another CTL com-

plementary DNA (cDNA) library was screened, in which inserts greater than 1000 base pairs were cloned in λ gt10. Forty thousand recombinants were screened and 39 plaques corresponding to C11 were isolated, but no evidence for a B10 recombinant could be found.

A cDNA insert of 1400 base pairs, which hybridized with C11, was selected for sequence analysis (5). The predicted protein sequence encoded, of molecular weight 25,319, is shown in Fig. 1B. The putative start codon is preceded by a potential ribosome binding site CCUCCG (6), and a polyadenylation signal sequence AAUAAA (7) occurs just upstream from the poly(A) tract. Of the first 12 amino acids predicted, ten are hydrophobic, and the amino acid in position 2 (Lys) is basic, suggesting that this sequence may act as a signal to direct secretion or intracellular organelle location (8).

A search of the National Biomedical Research Foundation (NBRF) protein sequence data bank revealed that the protein encoded by C11 resembles a number of serine proteases (Table 1). When the se-

Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7.

quences were optimally aligned according to the Dayhoff algorithm (9), the homologies generally varied between 30 and 40 percent. The greatest homology was found with rat mast cell protease type II (RMCPII), which had amino acids identical to 109 of 215 amino acids encoded by C11, giving a match per length of 51 percent. The amino acid residues known to form the catalytic triad of the active site in serine proteases (His⁵⁷, Asp¹⁰², and Ser¹⁹⁵) (10) were all found in the protein encoded by C11. The sequences around these residues, which are highly conserved among serine proteases, are also conserved in the C11 gene product. Indeed, largely because of conservation around this region, the protein encoded by C11 appears to be somewhat homologous (about 30 percent of 209 residues) even to the prokaryotic proteases trypsin and type B from *Streptomyces griseus*.

The cytotoxic T cell-specific proteins (CCP's) encoded by C11 and B10 will be referred to as CCPI and CCPII, respectively. In Fig. 2 the optimal protein alignment with CCPI is presented for RMCPII, bovine chymotrypsin, bovine trypsin, and CCPII. RMCPII is an intracellular serine protease found in the granules of atypical mast cells (11). The high level of homology of CCPI with RMCPII is particularly intriguing as RMCPII has a number of structural features that make it exceptional in the serine protease superfamily (12). Protein CCPI contains cysteines in precisely the same positions as RMCPII which, by analogy with RMCPII, form three disulfide bonds. These occur in the same positions in chymotrypsin, trypsin, and elastase. Both CCPI and RMCPII lack a disulfide bond that is present in all other known serine proteases, including several from prokaryotes (13), and that links Cys¹⁹¹ with Cys²²⁰ in chymotrypsin. In both CCPI and RMCPII the first of these two half-cysteines is replaced by a phenylalanine, while the second half-cysteine has been deleted along with other residues. Linkage of Cys¹⁹¹ to Cys²²⁰ is thought to be important in stabilizing the conformation of the substrate binding site (12). Its absence in CCPI and RMCPII may lead to significant changes in that site and, hence, in substrate specificity.

Two other primary structure changes previously seen only in RMCPII and thought to alter substrate binding are also present in the predicted CCPI protein. In RMCPII and CCPI the amino acid six residues before the active-site serine is alanine. In chymotrypsin-like proteases it is serine and in trypsin-like proteases, aspartic acid. The residue in this position lies at the bottom of the S₁ binding site (14), so the change to a less polar residue would indicate a preference for a hydrophobic amino acid at the P₁ position

in the substrate (13). Furthermore, the sequence Ser-Trp-Gly²¹⁶ in chymotrypsin, which forms hydrogen bonds with the P₁ and P₃ residues of the substrate (14), is replaced by Ser-Tyr-Gly in CCPI and RMCPII, again suggesting altered substrate specificity. Both of these changes are also seen with CCPII.

One of the few RMCPII-specific differ-

ences that is not present in CCPI is the substitution of isoleucine at position 99 in chymotrypsin for asparagine. In most mammalian serine proteases this residue is hydrophobic, and indeed in CCPI it appears to be phenylalanine. However, most of the RMCPII-specific changes are present in CCPI protein, suggesting that the substrate binding site of CCPI will resemble that of

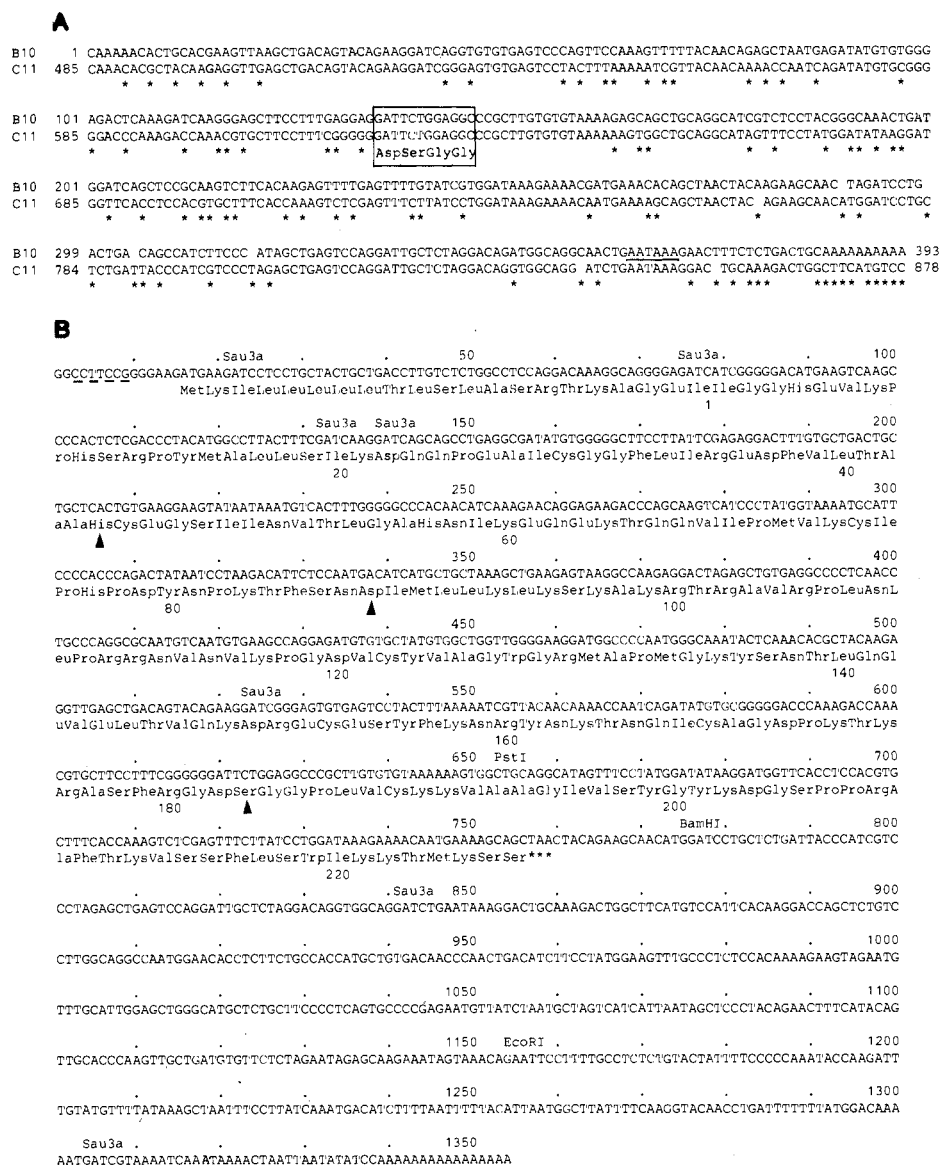


Fig. 1. (A) Nucleotide sequence comparison of clones B10 and C11. Insert DNA was purified from the CTL-specific clones designated B10 and C11 (3), recloned in M13 or pUC vectors and sequenced by the dideoxy method (5). The two sequences were maximally aligned; * designates nucleotide mismatches. The sequence of B10 is numbered from the first nucleotide of the insert, whereas that of C11 has been renumbered to correspond with the full-length sequence shown in (B). The region encoding the characteristic serine protease sequence Asp-Ser-Gly-Gly is boxed, and the polyadenylation signal sequence of B10 is underlined. (B) Nucleotide and predicted protein sequence of C11 insert. A size selected cDNA library (>1000 base pairs) was screened with a C11 insert isolated from the original library (3). A plaque that contained a cross-hybridizing sequence, whose apparent length was 1400 base pairs was selected for further analysis. Restriction fragments were subcloned in M13 or pUC13 and sequenced as above. A potential initiator methionine codon is present at nucleotide 16 followed by an open reading frame to nucleotide 756. A putative ribosome binding site (-----) and a polyadenylation signal (—) are underlined. By analogy with RMCPII (see text and Fig. 2), the numbering of the C11 protein starts at the isoleucine 21 residues downstream from the proposed translational initiation site. The amino acids which form the catalytic triad of the serine proteases (His⁵⁷, Asp¹⁰², and Ser¹⁹⁵ in chymotrypsin) are marked (▲).

Table 1. A selection of proteins that are homologous to the predicted C11 protein, CCPI. The protein sequence predicted from the longest open reading frame encoded by CCPI was compared with the National Biomedical Research Foundation protein sequence data bank. The numbering for CCPI (see text) is given in Fig. 1B. The data bank numbering system has been used for homologous proteins. All of the proteins that were significantly homologous (>30%) with CCPI over a large portion of the molecule (>150 residues) were serine proteases.

Protein	E.C. number	Species source	Residues compared		Percent homology
			CCPI	Bank protein	
7S nerve growth factor	3.4.21	Murine	29-224	26-229	40
Chymotrypsin A	3.4.21.1	Bovine	1-200	16-216	35
Chymotrypsin B	3.4.21.1	Bovine	1-200	16-216	36
Complement C1r	3.4.21.41	Human	52-224	56-238	35
Elastase	3.4.21.11	Porcine	3-220	3-233	33
Factor X	3.4.21.6	Bovine	1-225	192-421	33
RMCPH	3.4.21	Rat	1-214	1-213	51
Kallikrein	3.4.21.8	Rat	26-225	51-262	36
Plasminogen	3.4.21.7	Human	3-224	563-787	37
Plasminogen activator	3.4.21.31	Human	72-224	389-560	35
Trypsin	3.4.21.4	<i>S. griseus</i>	29-220	22-214	33
Trypsin	3.4.21.4	Rat	29-226	31-228	39

CCPI	Ile Ile Gly Gly His Glu Val Lys Pro His Ser Arg Pro Tyr Met Ala Leu Leu Ser Ile Lys Asp	22
RMCPH	Ile Ile Gly Gly Val Glu Ser Ile Pro His Ser Arg Pro Tyr Met Ala His Leu Asp Ile Val Thr	22
CA COW	Ile Val Asn Gly Glu Glu Ala Val Pro Gly Ser Trp Pro Trp Gln Val Ser Leu Gln Asp	37
TR COW	Ile Val Gly Gly Tyr Thr Cys Gly Ala Asn Thr Val Pro Tyr Gln Val Ser Leu Asn Ser	26
CCPI	Gln Gln Pro Glu Ala Ile Cys Gly Gly Phe Leu Ile Arg Glu Asp Phe Val Leu Thr Ala	42
RMCPH	Glu Lys Gly Leu Arg Val Ile Cys Gly Gly Phe Leu Ile Ser Arg Glu Asp Phe Val Leu Thr Ala	43
CA COW	Lys Thr Gly Phe His Phe Cys Gly Gly Ser Leu Ile Asn Glu Asn Trp Val Val Thr Ala	55
TR COW	Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Ser Gln Trp Val Val Ser Ala	44
CCPI	Ala His Cys Glu Gly Ser Ile Ile Asn Val Thr Leu Gly Ala His Asn Ile Lys Glu Gln	62
RMCPH	Ala His Cys Lys Gly Arg Glu Ile Thr Val Ile Leu Gly Ala His Asp Val Arg Lys Arg	63
CA COW	Ala His Cys Gly Val Thr Thr Ser Asp Val Val Val Ala Gly Glu Phe Asp Gln Gly Ser Ser	76
TR COW	Ala His Cys Tyr Lys Ser Gly Ile Gln Val Arg Leu Gly Gln Asn Ile Asn Val Val	64
CCPI	Glu Lys Thr Gln Gln Val Ile Pro Met Val Lys Cys Ile Pro His Pro Asp Tyr Asn Pro Lys Thr	84
RMCPH	Glu Ser Thr Gln Gln Lys Ile Lys Val Glu Lys Gln Ile Ile His Glu Ser Tyr Asn Ser Val Pro	85
CA COW	Ser Glu Lys Ile Gln Lys Leu Ile Ala Lys Val Phe Lys Asn Ser Lys Tyr Asn Ser Leu Thr	98
TR COW	Glu Gly Asn Gln Gln Phe Ile Ser Ala Ser Lys Ser Ile Val His Pro Ser Tyr Asn Ser Ala Thr	86
CCPI	Phe Ser Asn Asp Ile Met Leu Leu Lys Leu Lys Ser Lys Ala Lys Arg Thr Arg Ala Val Arg	105
RMCPH	Asn Leu His Asp Ile Met Leu Leu Lys Leu Glu Lys Lys Val Glu Leu Thr Pro Ala Val Asn	106
CA COW	Ile Asn Asn Asp Ile Thr Leu Leu Lys Leu Ser Thr Ala Ala Ser Phe Ser Gln Thr Val Ser	119
TR COW	Leu Asn Asn Asp Ile Met Leu Ile Lys Leu Lys Ser Ala Ala Ser Leu Asn Ser Arg Val Ala	107
CCPI	Pro Leu Asn Leu Pro Arg Arg Asn Val Asn Val Lys Pro Gly Asp Val Cys Tyr Val Ala Gly Trp	127
RMCPH	Val Val Pro Leu Pro Ser Pro Ser Asp Phe Ile His Pro Gly Ala Met Cys Trp Ala Ala Gly Trp	128
CA COW	Ala Val Cys Leu Pro Ser Ala Ser Asp Asp Phe Ala Ala Gly Thr Thr Cys Val Thr Thr Gly Trp	141
TR COW	Ser Ile Ser Leu Pro Thr Ser Cys Ala Ser Ala Gly Thr Gln Cys Leu Ile Ser Gly Trp	127
CCPI	Gly Arg Met Ala Pro Met Gly Lys Tyr Ser Asn Thr Leu Gln Glu Val Glu Leu	145
RMCPH	Gly Lys Thr Gly Val Arg Asp Pro Thr Ser Tyr Thr Leu His Glu Val Lys Leu	146
CA COW	Gly Leu Thr Arg Tyr Thr Asn Ala Asn Thr Pro Asp Arg Leu Gln Glu Val Glu Leu	160
TR COW	Gly Asn Thr Lys Ser Ser Gly Thr Ser Tyr Pro Asp Val Leu Lys Cys Leu Lys Ala	146
CCPI	Thr Val Gln Lys Asp Arg Glu Cys Glu Ser Tyr Phe Lys Asn Arg Tyr Asn Lys Thr Asn Gln	166
RMCPH	Arg Val Gln Lys Asp Gln Val Cys Glu Ser Gln Phe Gln Ser Phe Tyr Asn Arg Ala Asn Glu	165
CA COW	Arg Ile Met Asp Glu Lys Ala Cys Val Asp Tyr Arg Tyr Tyr Glu Tyr Lys Phe Gln	179
TR COW	Pro Leu Leu Ser Asn Thr Asn Cys Lys Lys Tyr Tyr Gly Thr Lys Ile Lys Asp Ala Met	166
CCPI	Ile Cys Ala Gly Asp Pro Lys Thr Lys Arg Ala Ser Phe Arg Gly Asp Ser Gly Gly Pro Leu Val	188
RMCPH	Ile Cys Val Gly Asp Ser Lys Ile Lys Gly Ala Ser Phe Glu Glu Asp Ser Gly Gly Pro Leu Val	187
CA COW	Val Cys Val Gly Ser Pro Thr Thr Leu Arg Ala Phe Met Gly Asp Ser Gly Gly Pro Leu Val	200
TR COW	Ile Cys Ala Gly Ala Ser Gly Val Ser Ser Cys Met Gly Asp Ser Gly Gly Pro Leu Val	188
CCPI	Phe Cys Ala Gly Tyr Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Val Val	
CCPI	Cys Lys Lys Val Ala Ala Gly Ile Val Ser Tyr Gly Tyr Lys Asp Gly Gly	204
RMCPH	Cys Lys Arg Ala Ala Ala Gly Ile Val Ser Tyr Gly Gln Thr Asp Gly Gly	203
CA COW	Cys Ala Gly Val Ala His Gly Ile Val Ser Tyr Gly His Pro Asp Ala	222
TR COW	Cys Lys Lys Asn Gly Ala Trp Thr Leu Val Gly Ile Val Ser Trp Gly Ser Ser Thr Cys Ser Thr	205
CCPI	Ser Pro Pro Arg Ala Phe Thr Lys Val Ser Ser Phe Leu Ser Trp Ile Lys Lys Thr Met Lys	225
RMCPH	Ser Ala Pro Gln Val Phe Thr Arg Val Leu Ser Phe Val Ser Trp Ile Lys Lys Thr Met Lys	223
CA COW	Lys Pro Pro Ala Ile Phe Thr Arg Val Ser Thr Tyr Val Pro Thr Ile Asn Ala Val Ile	243
TR COW	Ser Thr Pro Gly Val Tyr Ala Arg Val Thr Ala Leu Val Asn Trp Val Gln Gln Thr Leu Ala	227
CCPI	Lys Asn Lys Pro Gly Val Tyr Thr Lys Val Cys Asn Tyr Val Ser Trp Ile Lys Gln Thr Ile Ala	
CCPI	Ser Ser 227	
RMCPH	His Ser 224	
CA COW	Asn 224	
TR COW	Ala Asn 245	
	Ser Asn 229	

Fig. 2. Alignment of the predicted CCPI protein sequence with other serine proteases. The protein sequence predicted from C11 is aligned, according to the method of Dayhoff (9), with rat mast cell protease (RMCPH), bovine chymotrypsin (CA COW), bovine trypsin (TR COW), and the partial protein sequence predicted from insert B10 (CCPI). Amino acids that are identical to those of CCPI are boxed. Residue numbers are given at the end of each line; CCPI and RMCPH start at the Ile residue, whereas the numbering of CA and TR is based on the inactive zymogens. No numbering is presented for CCPI because the full sequence has yet to be determined.

RMCPH and could be significantly different from those of other mammalian serine proteases (15).

In addition to the interesting aspects of the structure of CCPI, the fact that both CCPI and CCPII appear to be serine proteases has significant implications for the mechanism by which CTL's are activated to become capable of lysing target cells (1, 2). Protease inhibitors have been shown to block T-cell effector function (16), and a trypsin-like esterase activity was shown to be induced during CTL activation (17). We have demonstrated that two genes that encode protease-like proteins are expressed in CTL's. Previous evidence that these genes are activated specifically in CTL's and that their expression correlates with cytolytic activity (3) suggests that they play a key role in the development of a cytotoxic response.

The two genes we have so far characterized appear to encode serine esterases, and another CTL-specific gene, AR10, cloned by Gershenfeld and Weissman (18) encodes a trypsin-like enzyme. These findings suggest that a protease cascade mechanism may be involved in T-cell cytotoxicity. Such a mechanism has been shown in a number of biological systems, including blood coagulation and complement-mediated lysis, and is advantageous in control, specificity, and amplification of enzymatic reactions (19).

REFERENCES AND NOTES

1. G. Berke, *Immunol. Rev.* 72, 5 (1983).
2. M. Nabholz and H. R. MacDonald, *Annu. Rev. Immunol.* 1, 273 (1983).
3. C. G. Lobe, C. Havelle, R. C. Bleackley, *Proc. Natl. Acad. Sci. U.S.A.* 83, 1448 (1986).
4. To date, B10 and C11 expression has been detected in seven of seven cloned CTL lines and in none of three cloned T-helper cell lines. Other cell types tested, including murine brain, liver, thymus and lipopolysaccharide-activated B cells and human CTL's, gave no signal (C. Lobe, unpublished data).
5. F. Sanger, A. R. Coulson, B. G. Barrell, A. J. H. Smith, B. A. Roe, *J. Mol. Biol.* 143, 161 (1980).
6. O. Hagenbuchle, M. Santer, J. A. Streitz, *Cell* 13, 551 (1978).

7. N. J. Proudfoot and G. G. Brownlee, *Nature (London)* **263**, 211 (1976).
8. G. von Heijne, *J. Mol. Biol.* **184**, 99 (1985).
9. M. O. Dayhoff, in *Atlas of Protein Sequence and Structure*, M. O. Dayhoff, Ed. (National Biomedical Research Foundation, Washington, DC, 1979), vol. 5 (suppl. 3), p. 1.
10. H. Neurath, *Science* **224**, 350 (1984).
11. R. G. Woodbury et al., *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5311 (1978); R. G. Woodbury and H. Neurath, *FEBS Lett.* **114**, 189 (1980).
12. R. G. Woodbury, N. Katunuma, K. Kobayashi, K. Titani, H. Neurath, *Biochemistry* **17**, 811 (1978).
13. P. Johnson and L. B. Smillic, *FEBS Lett.* **47**, 1 (1974); L. Jurásek et al., *Biochem. Biophys. Res. Commun.* **61**, 1095 (1974).
14. The peptide bond cleaved by proteases is between residues P₁ and P₁'. The corresponding amino acids in the protease are designated S₁ and S₁'. Other residues are numbered relative to the peptide bond cleaved. D. M. Segal et al., *Biochemistry* **10**, 3728 (1971).
15. J. C. Powers et al., *ibid.* **24**, 2048 (1985).
16. D. Redelman and D. Hudig, *J. Immunol.* **124**, 870 (1980); T. W. Chang and H. N. Eisen, *ibid.*, p. 1028.
17. M. S. Pasternack and H. N. Eisen, *Nature (London)* **314**, 743 (1985).
18. H. K. Gershenfeld and I. L. Weissman, *Science* **232**, 854 (1986).
19. H. Neurath and K. A. Walsh, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3825 (1976).
20. We thank M. James and W. Anderson for helpful discussion, I. Weissman and H. Gershenfeld for sharing material (including the λ gt10) and results prior to publication, and D. Oare for preparing the manuscript. Supported by the Medical Research Council of Canada, the National Cancer Institute of Canada, and the Alberta Heritage Foundation for Medical Research provided salary support.

2 October 1985; accepted 6 March 1986

The Locale Map of Honey Bees: Do Insects Have Cognitive Maps?

JAMES L. GOULD

Whereas higher vertebrates are able to construct a mental "map" of their home area and so use their knowledge of the spatial relations between landmarks to navigate along novel routes, invertebrates have been thought able to use landmarks in their navigation only as a familiar, route-specific series. Experiments with honey bees show that these insects have and use landmark maps thus invalidating this presumed invertebrate-vertebrate dichotomy.

ALTHOUGH HONEY BEES USE CELESTIAL cues in their navigation to unfamiliar food sources (1), they frequently begin to rely on prominent landmarks, at least on the outward journey, as the route becomes familiar (2, 3). Wehner (4) has proposed that the landmark memory of invertebrates is stored as a series of route-specific photographs; hence, bees and other insects could use landmarks only in serial order and along familiar routes. This model stands in sharp contrast to the way higher vertebrates are thought to use landmarks as part of a map; in this system, the relative location of familiar landmarks is understood—presumably stored in the brain as a map—so that novel routes based on new combinations of landmarks may be used, freeing the animal from dependence on route-specific combinations.

To put this presumed vertebrate-invertebrate dichotomy into concrete terms, consider the area around an animal's home (Fig. 1). If the hive is its home and site A is a food source, an insect would be presumed to know site A in terms of a set of landmarks regularly encountered there; if displaced to site B, the insect would, at best, recognize site B as part of some other route leading from home to another food source, and perhaps be able to follow that route back home. A higher vertebrate familiar with the same area, on the other hand, would be able to use the landmarks visible at site B to determine the direction of site A, and set off directly toward it even though it had never traveled from site B to site A before; the

vertebrate, it is supposed, has integrated its landmark knowledge into a map.

Two observations suggest that the map alternative might be available to bees. First, older foragers captured while leaving the hive, transported in darkness to a location hundreds of meters away and out of sight of the hive, fed a highly concentrated sugar

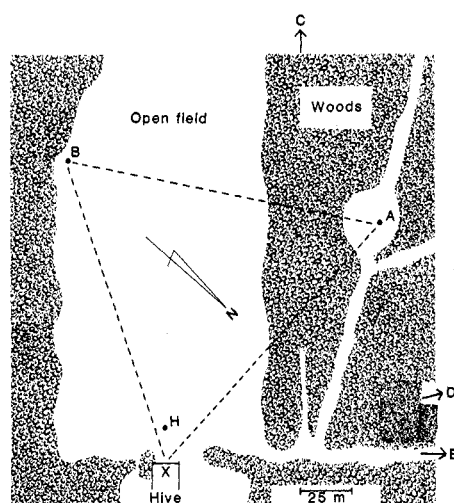


Fig. 1. Bees from a hive (bottom center) were trained to either site A, site B (both 160 m from the hive), or (off the top of the map) site C (350 m from the hive); on subsequent days these foragers and others were captured and transported to a different site and released—site B for those trained to A, site A for those trained to B, and site D (350 m from the hive) or site E (4425 m from the hive) for those trained to site C. As a control, some trained foragers were released at site H as well. Arrow N marks north.

solution, and released can fly directly home (5, 6); that young foragers cannot suggests that some familiarity with the locale may be important. This result is not conclusive since the foragers might have recognized the release site as part of their route-specific memory.

The second suggestion comes from experiments in which forager dances were manipulated to indicate a location either in the middle of a lake or on the far shore (7, 8). Recruitment was effective only to the latter, more plausible site, suggesting that the dance coordinates enabled recruits to make some sort of judgment about the suitability of the advertised site before leaving the hive; recruits make economic judgments about the distance versus the quality of the food being advertised before leaving (9). But the lake results could be explained on the basis of environmental factors (8). The purpose of the experiments reported here was to determine whether bees are limited to route-specific navigation or, instead, have a true locale map available to them.

Individually marked honey bees (*Apis mellifera ligustica*) were first trained to site A 160 m west of the hive (Fig. 1) along a path through the woods at Stony Ford, the Princeton University field station in Princeton, New Jersey. On the basis of the training procedure and the observed departure bearings from both the hive and the food source, foragers flew either directly to and from the hive to reach the feeding station or along a two-legged route that included the last portion of the training route (the path). All bees visiting the station were either marked or captured.

On subsequent days, foragers regularly visiting site A were captured at the hive entrance in a 600-ml beaker placed in their takeoff path. Each trapped forager was then carried in the dark across an open field to site B, 160 m south-southwest of the hive, and released. Two foragers were sometimes captured and transported together, but bees

Department of Biology, Princeton University, Princeton, NJ 08544.