332, 174 (1988); H. R. MacDonald *et al.*, *ibid.*, p. 40; H. Pirchner, K. Burke, R. Lang, H. Hengartner, R. M. Zingernagel, *ibid.* **342**, 559 (1989); J. White *et al.*, *Cell* **56**, 27 (1989).
3. J. W. Kappler *et al.*, *Cell* **49**, 263 (1987); J. W. Kappler, N. Roehm, P. Marrack, *ibid.* **49**, 273 (1987); J. Bill, O. Kanagawa, D. L. Woodland, E. Palmer, *J. Exp. Med.* **169**, 405 (1989); D. L. Woodland, P. Happ, J. Bill, E. Palmer, *Science* **247**, 964 (1990).
4. W. Sha *et al.*, *Nature* **336**, 73 (1988); L. J. Berg, B. Fazekas de St. Groth, A. M. Pullen, M. M. Davis, *ibid.* **340**, 559 (1989).
5. H. Smith, I. M Chen, R. Kubo, K. S. K. Tung, *Science* **245**, 749 (1989); R. J. Hodes, S. Sharrow, A. Solomon, *ibid.* **246**, 1041 (1989).
6. A. M. Fry, L. A. Jones, A. M. Kruisbeek, L. A. Matis, *ibid.*, p. 1044; L. A. Jones, L. T. Chin, G. Merriam, L. A. Nelson, A. M. Kruisbeek, *J. Exp. Med.* **172**, 1277 (1990).
7. F. Ramsdell, T. Lantz, B. J. Fowlkes, *Science* **246**, 1038 (1989); J. L. Roberts, S. O. Sharrow, A. Singer, *J. Exp. Med.* **171**, 935 (1990).
8. G. Morohan, J. Allison, J. F. A. P. Miller, *Nature* **339**, 622 (1989); K. M. Murphy, C. T. Weaver, M. Elish, P. M. Allen, D. Loh, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 10034 (1989); D. Lo *et al.*, *Cell* **53**, 159 (1989); J. Markmann *et al.*, *Nature* **336**, 476 (1988).
9. H. von Boehmer and K. Schubiger, *Eur. J. Immunol.* **14**, 1048 (1984); H. von Boehmer and K. Hafen, *Nature* **320**, 626 (1986); A. R. Ready, E. J. Jenkinson, R. Kingston, J. J. T. Owen, *ibid.* **310**, 231 (1984); R. K. Jordan, J. H. Robinson, N. A. Hopkinson, K. C. House, A. L. Bentley, *ibid.* **314**, 454 (1985); J. Miller, L. Daitch, S. Rath, E. Seleing, *J. Immunol.* **144**, 334 (1990); J. Boehme *et al.*, *Science* **244**, 1179 (1989).
10. C. A. Smith, G. T. Williams, R. Kingston, E. J. Jenkinson, J. T. T. Owen, *Nature* **337**, 181 (1989); M. Mercep, A. M. Weissman, S. J. Frank, R. D. Klausner, J. A. Ashwell, *Science* **246**, 1162 (1989).
11. P. Marrack *et al.*, *Cell* **53**, 627 (1988).
12. D. E. Speiser, R. Schneider, H. Hengartner, H. R. MacDonald, R. M. Zingernagel, *J. Exp. Med.* **170**, 595 (1989); D. E. Speiser, R. K. Lees, H. Hengartner, R. M. Zinkernagel, H. R. MacDonald, *ibid.*, p. 2165.
13. S. R. Webb and J. Sprent, *ibid.* **171**, 953 (1990); H. R. MacDonald, R. K. Lees, Y. Chvatchko, *ibid.*, p. 1381.
14. S. Qin, S. Cobbold, R. Benjamin, H. Waldmann, *ibid.* **169**, 779 (1989); H. Rammensee, R. Kroschewski, B. Frangoulis, *Nature* **339**, 541 (1989).
15. C. S. David and G. Anderson, in *Transgeneic Mice and Mutants in MHC Research*, I. K. Egorov and C. S. David, Eds. (Springer-Verlag, Berlin, 1990), pp. 229–236.
16. C. A. Janeway, Jr., E. A. Lerner, J. M. Jason, B. Jones, *Immunogenetics* **10**, 481 (1980); C. A. Janeway, Jr., and M. E. Katz, *J. Immunol.* **134**, 2057, (1985); D. H. Lynch, R. E. Gress, B. W. Needleman, S. A. Rosenberg, R. J. Hodes, *ibid.*, p. 2071 (1985); A. L. Glasebrook and F. W. Fitch, *J. Exp. Med.* **151**, 876 (1980); R. Abe and R. J. Hodes, *J. Immunol.* **140**, 4132 (1988); S. MacPhail and O. Stutman, *Immunogenetics* **24**, 139 (1986); K. A. Wall, M. I. Lorber, M. Loken, S. McClatchey, F. W. Fitch, *J. Immunol.* **131**, 1056 (1983); S. MacPhail and O. Stutman, *Eur. J. Immunol.* **14**, 318 (1984).
17. A. M. Kruisbeek *et al.*, *J. Exp. Med.* **161**, 1029 (1985).
18. S. Marusic-Galesic, D. L. Longo, A. M. Kruisbeek, *Nature* **333**, 180 (1988).
19. L. A. Jones and A. M. Kruisbeek, unpublished data.
20. L. A. Jones, J. C. Zuniga-Pflucker, J. S. Fine, D. L. Longo, A. M. Kruisbeek, *Progress Immunol.* **VII**, 289 (1989) (blocking of deletion of $V_\beta 11$ cells); treatment with anti-E_α also prevents deletion of $V_\beta 5$-expressing T cells in the strains used here, as well as in BALB/c and C3H/HeJ mice. Both $V_\beta 11$ and $V_\beta 5$ T cells are deleted when recognizing self antigens (cotolerogens) in association with $E_\alpha E_\beta$ (*3*).
21. N. S. Liao, J. Maltzman, D. H. Raulet, *J. Exp. Med.* **170**, 135 (1989).
22. M. S. Vacchio and R. J. Hodes, *ibid.*, p. 1335.
23. Maximal responsiveness was reached with anti-$V_\beta 6$ at 50 μg/ml with spleen cells from other mouse strains: for BALB/c (H-2^d, Mls-1^b): 40,000 cpm; for DBA/1(H-2^q, Mls-1^a but nondeleting for $V_\beta 6$): 19,000 cpm; compare to 17,000 cpm for anti-E_α–treated mice. Note that DBA/1 mice are not only nondeleting, but apparently also nontolerizing for $V_\beta 6$-expressing T cells, presumably because $A_\alpha{}^q A_\beta{}^q$ molecules are unable to present Mls-1^a (*15*).
24. Supernatants from the cultures shown in Fig. 2, A and B, were diluted 1:2, 1:4 and 1:8 and tested for their ability to support proliferation of a subline of CTLL that responds to both IL-2 and IL-4; CTLL proliferation was observed in a dose dependent fashion only with supernatants from T cells from anti-E_α–treated mice, with cultures giving maximal proliferation in Fig. 2, A and B, yielding maximal proliferation of CTLL cells as well.
25. Control untreated mice, although unresponsive to anti-$V_\beta 6$ cross-linking, did respond by proliferation and lymphokine production to TCR cross-linking with control MAbs to $V_\beta 14$ and pan MAbs to αβ TCR in a dose dependent fashion.
26. Peripheral T cells from anti-E_α–treated (DBA/2 × C57B1/6)F_1 mice not only respond to spleen cells from DBA/2 (H-2^d, Mls-1^a) and C57B1/6 (H-2^b, Mls-1^b) mice, as shown in Fig. 2C, but also to spleen cells from BALB/c (H-2^d, Mls-1^b) and B10.D2 (H-2^d, Mls-1^b) mice. When stimulator cells express Mls-1^a, $V_\beta 6$ T cells constitute part of the responding population (*27*). Thus responses to Mls-1^b stimulator cells probably reflect representation of Mls-1^a gene products derived from the responder cells on class II MHC molecules of the stimulator cells; the degree of responsiveness is consistent with this hypothesis: the H-2^b haplotype is weakly stimulatory for Mls-1^a whereas H-2^d is highly stimulatory (*15*). In addition, the responding cells in these MLCs included T cells responding to cotolerogens presented in association with $E_\alpha E_\beta$ (*3*) that also were not deleted in anti-E_α–treated mice (*20*); the response to BALB/c and B10.D2 stimulator cells may include such T cells.
27. T cell enriched spleen cells from anti-E_α treated mice were cultured for 5 days at 3×10^6 cells per well of 24-well plates with 3×10^6 mitomycin C-treated syngeneic (DBA/2) or allogeneic (C3H/HeJ) stimulator cells. On days 5 and 8 of culture, viable cells were purified by centrifugation on Lympholyte M and analyzed for V_β expression by flow cytometry. Allogeneic stimulation did not enrich for $V_\beta 6^+$ T cells compared to analysis before culture. Stimulation with syngeneic Mls-1^a (DBA/2) expressing cells resulted in a three-fold enhancement of $V_\beta 6^+$ T cells by day 5 that further expanded to 25% of the αβ TCR expressing cells by day 8.
28. Mice were analyzed by two-color flow cytometry for IL-2 receptor (IL-2R) and $V_\beta 6$ or IL-2R and αβ TCR at each of the time points indicated in Fig. 4; control and treated mice did not differ. In addition, $V_\beta 6^+$ T cells did not increase in size (indicating blast formation) by forward scatter.
29. J. Sprent, *Cell. Immunol.* **21**, 278 (1976).
30. B. J. Fowlkes, R. H. Schwartz, D. M. Pardoll, *Nature* **334**, 620 (1988); H. R. MacDonald, H. Hengartner, T. Pedrazzini, *ibid.* **335**, 174 (1988).
31. B. L. Rellahan, L. A. Jones, A. M. Kruisbeek, A. M. Fry, L. A. Matis, *J. Exp. Med.* **172**, 1091 (1990).
32. R. H. Schwartz, *Science* **248**, 1349 (1990).
33. MAbs used: 14-4-4 (anti-E_α), K. Ozato, N. Mayer, D. H. Sachs, *J. Immunol.* **124**, 533 (1980); GK1.5 (anti-CD4), Dialynas *et al.*, *ibid.* **131**, 2445 (1983); H57-597 (anti-pan αβ TCR), R. T. Kubo *et al.*, *ibid.* **142**, 2736 (1989); J11D (anti-HSA), J. Bruce, F. Symington, T. McKern, J. Sprent, *ibid.* **127**, 2496 (1981); RR4-7 (anti-Vβ6), O. Kanagawa, E. Palmer, J. Bill, *Cell. Immunol.* **119**, 214 (1989); MKD6 (anti-A^d_β), J. W. Kappler, B. Skidmore, J. White, P. Marrack, *J. Exp. Med.* **153**, 1198 (1981).
34. We thank R. Kubo, O. Kanagawa, and D. H. Raulet for providing MAbs to TCRs, F. Hausman and D. Stephany for expert and cheerful flow cytometry analysis, and R. N. Germain for many stimulating discussions and critical reading of the manuscript. L. T. C. is an HHMI/NIH Research Scholar.

23 August 1990; accepted 6 November 1990

Hemolin: An Insect-Immune Protein Belonging to the Immunoglobulin Superfamily

SHAO-CONG SUN, INGRID LINDSTRÖM, HANS G. BOMAN, INGRID FAYE, OTTO SCHMIDT*

Insects have an efficient defense system against infections. Their antibacterial immune proteins have been well characterized. However, the molecular mechanisms by which insects recognize foreignness are not yet known. Data are presented showing that hemolin (previously named P4), a bacteria-inducible hemolymph protein of the giant silk moth *Hyalophora cecropia*, belongs to the immunoglobulin superfamily. Functional analyses indicate that hemolin is one of the first hemolymph components to bind to the bacterial surface, taking part in a protein complex formation that is likely to initiate the immune response.

INSECTS ARE WELL ADAPTED TO DIverse environments and defend themselves against many parasites and microorganisms by highly effective cellular (*1*) and humoral immune mechanisms (*2*). It has been generally accepted that these reactions do not depend on immunoglobulins (Igs) (*3*) although attempts have been made to identify antibodies in insects (*4*). Some insect proteins (*5*–*7*) have been reported to contain Ig-like domains (*8*). However, these proteins do not seem to be involved in the immune response, but rather they play a role in neural cell interaction.

Antibacterial proteins have been identified in the giant silk moth, *Hyalophora cecropia*, and these proteins constitute an important part of the humoral immune defense. A set of proteins induced by bacterial infection

Department of Microbiology, University of Stockholm, S-106 91 Stockholm, Sweden.

*On leave from Biologie III, Universität Freiburg, D-78 Freiburg, FRG.

has been isolated from pupal hemolymph. This set includes three groups of antibacterial proteins—lysozyme, attacins, and cecropins (*9*)—and hemolin, which was previously called P4 (*10*). Experiments performed on other insect species have confirmed that induction of these immune proteins is a common response to bacterial infection (*2*). These proteins are also induced by wounding or by injection of components of the bacterial cell wall, such as lipopolysaccharides (LPS) or degradation products from peptidoglycans (*2, 11*).

```
                    ___B___      ___C___                 _D_    __E__         ___F___
Hemolin I      NNPTVLECIIEGN-DQGVKYSWKKDGKSYNWQEH--NAALRKDE------GSLVFLRP---QASDEGHYQCFAETPAG
Hemolin II     GRPFQLDCVLPNA-YPKPLITWKKRLSGADPNAD-VTDFDRRNHSW--TDGNLYFTIVTKEDVSDIYKYVCTAKNAAV
Hemolin III    GDVTMIYCMYGSN-PMGYPNYFKNGKDVNGNPED---RITRHNRTS---GKRLLFKTT---LPEDEGVYTCEVDNGVG
Hemolin IV     GQDVTIPCKVTGL--PAPNVVWSHNAKPL--------SGGRATVTD----SGLVIKGV---KNGDKGYYGCRATNEHG
                     *  *        *                  *         *             *  *  *  *  *
Neuroglian I   DNPFIIECEADGQ--PEPEYSWIKNGIIFDWQ----AYDNRMLRQP--GRGTLVITIP---KDEDRGHYQCFASNEFG
NCAM V         GNQVNITCEVFAY--PSATISWFRDGQLLPSS---NYSNIKIYNTPS--ASYLEVTPD---SENDFGNYNCTAVNRIG
Ig Vκ          GDRVTITCRARQG--ISSWLAWYQQKPEKAPK(-13-)PSRFSGSGSGTDFTLTISSL---QPEDFATYYCQQYNSYP
Ig Cκ          SGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKD---EYERHNSYTCEATHKTS
```

Fig. 2. Alignment of the Ig-like domains of hemolin and four other representative members from the Ig superfamily. The sequences were aligned with the GAP program and visually optimized. The solid bars above the alignment indicate the approximate positions of the potential β strands (*8*). The β strands A and G and the β strands C′ and C″ from IgVκ are omitted. Conserved amino acids, present in at least four of the eight domains, are typed in bold letters, and those that are highly conserved among the Ig superfamily members are indicated by asterisks below the last domain of hemolin.

Hemolin is a 48-kD protein present in low but significant amounts in the hemolymph of cecropia pupae (*10*). After injection of live bacteria into pupae, the concentration of hemolin is 18 times greater than normal. However, the hemolin does not exhibit direct bactericidal effects (*12*). We have characterized the cDNA of hemolin (Fig. 1). Analysis of the deduced 413 amino acid sequence revealed unexpectedly that the protein contains four internal repeats characteristic of immunoglobulin (Ig)-like domains. Each domain is 90 to 110 amino acids in length and contains two cysteine residues at homologous positions. A number of other amino acids are conserved among the repeated domains, particularly around the Cys residues. Most of these amino acids are also conserved among the members of the Ig superfamily (*8*). Representative types of Ig domains are aligned in Fig. 2 such as insect neuroglian (*7*), mouse neural cell adhesion molecule (NCAM) (*13*) (C2 type), the variable region of the human κ light chain (Ig V_κ) (*14*), and the constant region of the mouse κ light chain (Ig C_κ) (*15*). The four Ig-like domains of hemolin resemble the C2 type. Chou-Fasman analysis (*16*) of the hemolin sequence predicts large stretches of β sheets. Folding of hemolin primarily as β structure, which is typical of Ig-like domains, has also been predicted by analysis of circular dichroism spectra of the protein (*12*).

Although glycosylation of purified hemolin has not been detected (*12*), the deduced protein sequence reveals three potential N-glycosylation sites. Hemolin contains a typical eukaryotic signal peptide (*17*) and lacks a transmembrane sequence. These properties are consistent with the finding that hemolin is a soluble protein released into the hemolymph of insects. However, immunocytochemical localization of hemolin on the surfaces of the fat body and certain hemocytes (*12*) indicates that hemolin may also attach to membranes.

Comparison of the hemolin with molecules from the Ig superfamily by the GAP program (*18*) revealed a close relation to the insect molecules neuroglian, amalgam (*6*), and fasciclin II (*5*), and to vertebrate proteins such as mouse neural adhesion molecule L1 (L1) (*19*) and mouse NCAM (*13*). The relative sequence identity is strongest to neuroglian (38%), followed by amalgam (28%), fasciclin II (24%), L1 (23%), and NCAM (20%). A particularly strong overall sequence homology was found to the first four Ig-like domains of neuroglian. The fifth domain of neuroglian shows homology mainly to the last domain of hemolin. There is no significant homology between hemolin and the fibronectin-like domains of neuroglian (Fig. 3A). Because mouse L1 is closely related to neuroglian (Fig. 3B), we compared the hemolin sequence to that of L1. Although hemolin is less similar to L1 than to neuroglian, conserved regions are still found in the five Ig-like domains located at the amino end of the molecule (Fig. 3C). Homology to NCAM, amalgam, and fasciclin II is particularly high in the last Ig-like domain of hemolin (Fig. 3D).

```
CTGGAGTATAAGGCACGGGTACACAATGGCGTTCAAGAGTATAGCAGTTTTAAGCGCCTGCATAATTGTGGGTTCAGCGCTTCCCGTGGATAAATACCCG   100
                          M  A  F  K  S  I  A  V  L  S  A  C  I  I  V  G  S  A  L  P  V  D  K  Y  P   25

GTGCTGAAAGACCAGCCGGCCGAGGTTCTCTTCAGGGAGAATAACCCAACGGTCCTCGAATGTATCATCGAAGGGAACGATCAGGGAGTCAAGTACTCCT   200
 V  L  K  D  Q  P  A  E  V  L  F  R  E  N  N  P  T  V  L  E  C  I  I  E  G  N  D  Q  G  V  K  Y  S  W   59
                                            ▲
GGAAAAAGGATGGGAAATCCTACAACTGGCAGGAGCATAACGCCGCTCTTCGCAAGGATGAAGGATCTTTGGTATTCCTGAGACCGCAAGCCTCAGACGA   300
  K  K  D  G  K  S  Y  N  W  Q  E  H  N  A  A  L  R  K  D  E  G  S  L  V  F  L  R  P  Q  A  S  D  E   92

GGGTCACTACCAGTGCTTCGCTGAAACCCCAGCCGGTGTTGCCAGCTCGAGGGTGATCAGCTTCAGGAAGACTTACCTAATCGCGTCGCCAGCAAAGACA   400
  G  H  Y  Q  C  F  A  E  T  P  A  G  V  A  S  S  R  V  I  S  F  R  K  T  Y  L  I  A  S  P  A  K  T   125

CACGAGAAAACGCCAATCGAAGGCAGGCCTTTCCAATTGGACTGCGTCCTCCCTAACGCTTACCCTAAACCCTTGATTACTTGGAAGAAACGTTTGTCCG   500
 H  E  K  T  P  I  E  G  R  P  F  Q  L  D  C  V  L  P  N  A  Y  P  K  P  L  I  T  W  K  K  R  L  S  G   159

GAGCCGATCCTAACGCTGACGTGACTGACTTTGATCGCCGCAATCACAGCTGGACTGACGGAAACCTCTACTTTACAATCGTCACTAAAGAGGACGTCAG   600
  A  D  P  N  A  D  V  T  D  F  D  R  R  N  H  S  W  T  D  G  N  L  Y  F  T  I  V  T  K  E  D  V  S   192
                                         ▲
TGACATTTATAAATACGTATGCACCGCCAAGAACGCAGCGGTTGACGAAGAGGTAGTTTTGGTGGAGTATGAAATCAAGGGAGTGACAAAAGACAACTCT   700
  D  I  Y  K  Y  V  C  T  A  K  N  A  A  V  D  E  E  V  V  L  V  E  Y  E  I  K  G  V  T  K  D  N  S   225

GGGTACAAGGGTGAGCCGGTCCCTCAATACGTAAGCAAGGATATGATGGCTAAAGCTGGTGACGTTACCATGATATACTGTATGTATGGAAGCAATCCTA   800
 G  Y  K  G  E  P  V  P  Q  Y  V  S  K  D  M  M  A  K  A  G  D  V  T  M  I  Y  C  M  Y  G  S  N  P  M   259

TGGGTTATCCTAACTACTTCAAGAATGGTAAGGACGTGAATGGAAACCCTGAAGACCGTATCACCCGCCACAATAGAACCTCAGGCAAACGTCTCCTCTT   900
  G  Y  P  N  Y  F  K  N  G  K  D  V  N  G  N  P  E  D  R  I  T  R  H  N  R  T  S  G  K  R  L  L  F   292
                                                                      ▲
CAAGACAACACTGCCAGAAGACGAGGGCGTATACACTTGTGAAGTCGACAATGGAGTCGGCAAACCCCAGAAACACAGTTTGAAATTGACTGTAGTCAGT   1000
  K  T  T  L  P  E  D  E  G  V  Y  T  C  E  V  D  N  G  V  G  K  P  Q  K  H  S  L  K  L  T  V  V  S   325

GCACCGAAGTACGAACAGAAACCGGAAAAGGTGATCGTCGTCAAACAAGGACAGGATGTCACGATCCCTTGCAAGGTGACCGGTCTGCCAGCGCCCAACG   1100
 A  P  K  Y  E  Q  K  P  E  K  V  I  V  V  K  Q  G  Q  D  V  T  I  P  C  K  V  T  G  L  P  A  P  N  V   359

TCGTCTGGAGCCATAACGCGAAGCCTCTAAGCGGTGGTAGAGCTACGGTCACCGACAGTGGTCTGGTCATCAAAGGCGTAAAAAATGGTGACAAGGGATA   1200
  V  W  S  H  N  A  K  P  L  S  G  G  R  A  T  V  T  D  S  G  L  V  I  K  G  V  K  N  G  D  K  G  Y   392

CTACGGCTGCAGGGCTACTAACGAGCATGGAGATAAATACTTCGAGACCCTTGTACAAGTTAACTAAACAGGTTAACTTGAATGTGGACTGTCAAAATAC   1300
  Y  G  C  R  A  T  N  E  H  G  D  K  Y  F  E  T  L  V  Q  V  N  STOP                                    413

GTAAAAAAGACATAAATGACAGTTGTGCGTCAACATCAAACTATAAGAACTTTTTAGGTTATATTATAATCGAAATAAGTAATAAAATAAAATTATTATA   1400

AAATAT                                                                                                 1406
```

Fig. 1. Sequence of cDNA encoding hemolin. The cDNA clone pCP434 (*23*) coding for a partial sequence of hemolin was sequenced with the dideoxy chain termination method (Sequenase, U.S. Biochemical). Based on the 5′ sequence of pCP434, an oligonucleotide (21 nucleotides) primer was synthesized and used to construct a hemolin-specific cDNA library, from which a clone containing the missing information was isolated. The first strand synthesis of cDNA was accomplished by primer extension (*24*). The second strand synthesis and the following steps including the ligation of the cDNA into pTZ19R were done with a cDNA synthesis kit (Pharmacia). This library was screened with a genomic subclone containing the 5′ coding region of the hemolin gene (*25*). The isolated clones were sequenced. The DNA sequence was analyzed with MacGene Plus software (Applied Genetic Technology). The sequence is numbered from the first base at the 5′ end. The polyadenylation signal, AATAAA, is underlined. The deduced amino acid sequence is numbered from the initiating methionine. The amino acid sequence matching the NH_2-terminal sequence determined from the purified hemolin (*12*) is underlined. The putative N-linked glycosylation sites are indicated by arrowheads. The Cys residues in the mature part of the sequence are circled.

To investigate the function of hemolin, we used an in vitro attachment assay to analyze binding of hemolymph components to the surface of bacteria. After 1 minute of incubation, only hemolin and a 125-kD protein can be extracted from the bacteria (Fig. 4, lane 6). After a 2-min incubation period, a band of about 170 kD is detected (lane 5) that becomes the most abundant protein associated with the bacterial surface after 3 minutes of incubation (lane 4, arrowhead). After 3 minutes several other proteins are also detected (lane 4).

To determine whether the protein with apparent molecular size of 170 kD is a complex of hemolin and the 125-kD protein, we analyzed protein extracts under native conditions. Proteins eluted from bacteria migrated in the upper region of the gel (Fig. 5A), whereas purified hemolin migrates to the lower half of the gel (*20*). However, when the most heavily stained band was excised from the gel and separated further by SDS–polyacrylamide gel electrophoresis (PAGE), it generated a protein of 125 kD and two others of about 48 kD (Fig. 5B). The two smaller proteins are related as revealed by analytical papain digests (*20*), but only one of them is clearly stained by antiserum against hemolin on a protein immunoblot (Fig. 5C). These results suggest that hemolin, together with a related protein, may form a protein complex on the bacterial surface with a 125-kD protein. The complex can be dissociated with SDS shortly after its formation but becomes stable to SDS treatment after longer incubation. The complex formation appears to be dependent on the presence of sugar residues in the LPS core of the bacterial cell wall since no protein of 170 kD was detected in attachment assays performed with LPS mutants of *Escherichia coli* K12 (*21*), whereas the binding of hemolin was not affected (*20*).

The structural homology between hemolin and neuroglian suggests that they have a common ancestor molecule. Hemolin might have diverged from neuroglian by gene rearrangements that also result in distinct functions of the two types of molecules; the membrane bound neuroglian is involved in neural cell-cell interaction (*7*), whereas the soluble hemolymph protein hemolin attaches to the surface of bacteria and possibly other foreign objects.

Nonclonal forms of recognition, termed "pattern recognition" (*22*), may have been important in the evolution of the immune system. Several of our observations indicate that hemolin is a recognition molecule that meets certain requirements for "pattern recognition": hemolin belongs to the Ig superfamily, it is present in normal hemolymph,

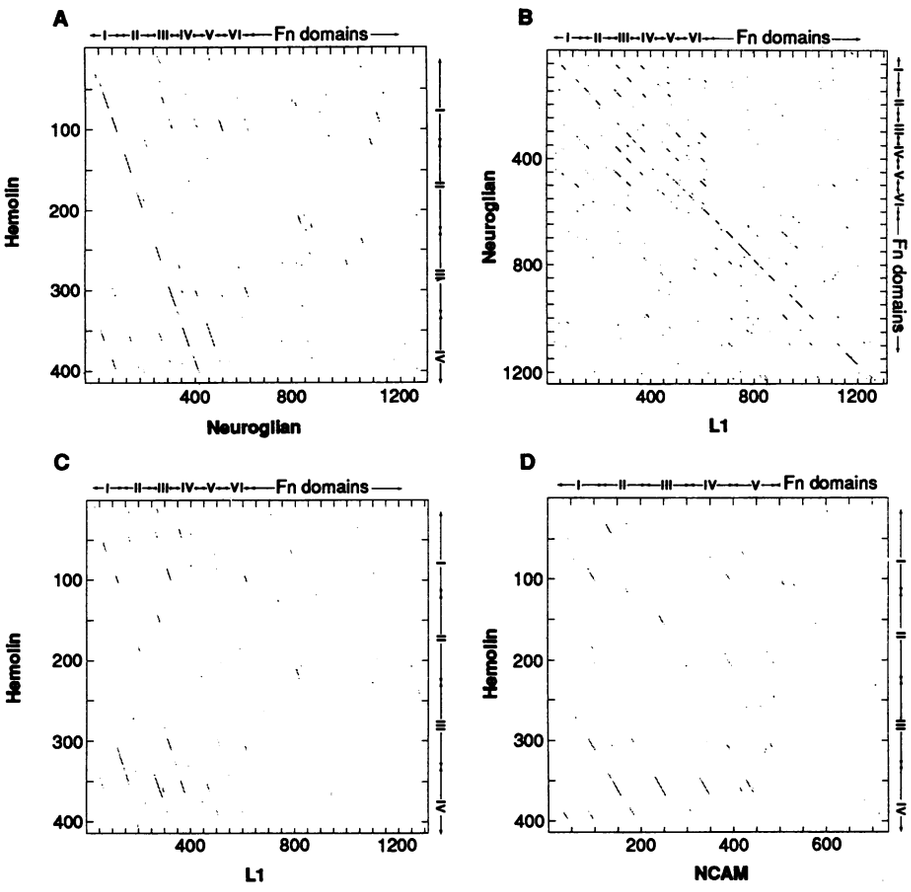

Fig. 3. Dot-matrix sequence comparison. The comparisons were made with the HOMOLOGY option of the MacGene program (Applied Genetic Technology) with a window of 17 residues (range 8) and cutoff of 64%. The complete amino acid sequence of hemolin was compared with *Drosophila* neuroglian (**A**), mouse L1 (**C**), and NCAM (**D**). Comparisons of hemolin to amalgam and fasciclin II gave similar results to those shown for NCAM (D). As a reference, neuroglian was compared with L1 (**B**). The amino acid numbers, starting from the initiating methionines, and the structural domains of the molecules are indicated on the axes. The Ig-like domains are indicated by roman numerals, and the regions covering the fibronectin-like domains are labeled (Fn).

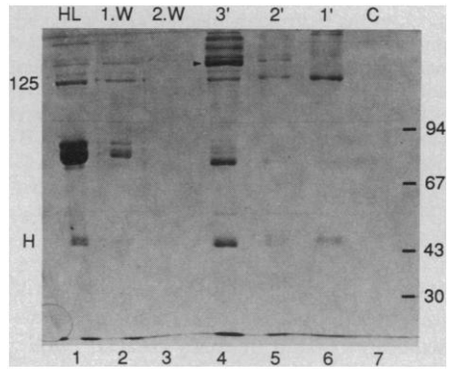

Fig. 4. In vitro binding assay. Hemolymph of diapausing pupae of *H. cecropia* were bled into ice-cold saturated ammonium sulfate solution containing 0.025% phenylthiourea. Water was added to adjust the concentration of ammonium sulfate to 40% of saturation, and the solution was incubated for 30 minutes on ice. The supernatant was recovered by centrifugation at 17,400*g* for 10 minutes at 0°C and stored at 4°C. Immediately before the incubation experiment, the supernatant (hemolymph fraction) was desalted over a PD10 column (Pharmacia) and diluted 1:10 with 10 mM tris-HCl (pH 8.2). The *E. coli* strain D21 was grown to late log phase, collected by centrifugation, and suspended in 1/10 of the original volume with 0.9% NaCl. An equal volume of 20% acetic acid was added, and the bacteria were left at room temperature for 5 minutes. Five volumes of 1 M tris-HCl (pH 8.2) was added, and the bacteria were collected by centrifugation and resuspended in one-tenth of the original culture volume in 10 mM tris-HCl (pH 8.2). The bacterial suspension was added to the hemolymph fraction (0.5 ml of bacteria to 2.5 ml of fraction) and incubated at room temperature for the time indicated. The suspension (1 ml) was centrifuged for 2 min at 10,000*g*, the pellet was washed twice in 1 ml of water and then suspended in 200 μl of 0.5 M ammonium formate (pH 6.4). The ammonium formate eluates were immediately adjusted to 0.1% SDS and 10 mM DDT, heated at 70°C for 5 minutes, and subjected to electrophoresis (*26*). The gels were stained with Coomassie brilliant blue after separation of the following: hemolymph fraction, 0.01 μl (lane 1); first washing, 10 μl (lane 2); second washing, 10 μl (lane 3); ammonium formate eluate after 3 min, 2 min, and 1 min of incubation, 10 μl (lanes 4, 5, and 6, respectively); and ammonium formate eluate from untreated bacteria, 10 μl (lane 7). Molecular mass markers indicated on the right side are in kilodaltons. The positions of hemolin (H) and the 125-kD protein are indicated. The 170-kD band is marked by an arrowhead.

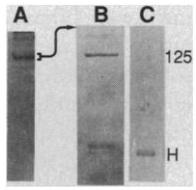

Fig. 5. Protein complex formation of hemolin with a 125-kD protein. The proteins were separated on a 10% polyacrylamide gel under nondenaturing conditions at pH 4, and visualized by Coomassie brilliant blue. Protein immunoblots were prepared (*27*) and tested with antiserum to hemolin that has been purified on a hemolin affinity column (*12*). Blots were incubated at 4°C overnight in an antibody solution (120 μg protein/ml), diluted 1:1000 in a buffer containing 0.5% bovine serum albumin, 150 mM NaCl, and 10 mM Tris-HCl (pH 8). After being washed with 0.02% Tween 20 in 0.9% NaCl the blots were incubated with alkaline phosphatase conjugated to the $F(ab')_2$ fragment of goat antibodies directed against rabbit Ig (sigma), and developed with a solution containing nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate. (**A**) Native gel electrophoresis of protein extracts eluted from the bacterial surface after 2 min. (**B**) Separation of the major band (A) by SDS-PAGE, stained with Coomassie brilliant blue. (**C**) Protein immunoblot of (B) with antiserum against hemolin.

and it is strongly induced after bacterial infection. Finally, hemolin binds to surface structures shared by many bacteria, where it forms a complex with another hemolymph protein, which might constitute an important part of the primary immune response.

REFERENCES AND NOTES

1. A. F. Rowley and N. A. Ratcliffe, in *Invertebrate Blood Cells*, A. F. Rowley and N. A. Ratcliffe, Eds. (Academic Press, New York, 1981), pp. 421–488.
2. P. E. Dunn, *Annu. Rev. Entomol.* **31**, 321 (1986); H. G. Boman and D. Hultmark, *Annu. Rev. Microbiol.* **41**, 103 (1987); S. Natori, "Hemolymph proteins participating in the defense system of sarcophaga peregrina," in *Molecular Entomology*, J. H. Law, Ed. (Liss, New York, 1987), vol. 49 of UCLA Symposia on Molecular Biology, new series, pp. 369–378.
3. J. Klein, *Scand. J. Immunol.* **29**, 499 (1989).
4. A. W. Bernheimer, E. Caspari, A. D. Kaiser, *J. Exp. Zool.* **119**, 23 (1952); J. D. Briggs, *ibid.* **138**, 155 (1958); G. F. George, R. D. Karp, B. L. Rellahan, J. L. Ressard, *Immunology* **62**, 505 (1987).
5. A. L. Harrelson and C. S. Goodman, *Science* **242**, 700 (1988).
6. M. A. Seeger, L. Haffley, T. C. Kaufman, *Cell* **55**, 589 (1988).
7. A. J. Bieber *et al.*, *ibid.* **59**, 447 (1989).
8. A. F. Williams and A. N. Barclay, *Annu. Rev. Immunol.* **6**, 381 (1988).
9. D. Hultmark, H. Steiner, T. Rasmuson, H. G. Boman, *Eur. J. Biochem.* **106**, 7 (1980); D. Hultmark *et al.*, *EMBO J.* **2**, 571 (1983); H. Steiner, D. Hultmark, A. Engström, H. Bennich, H. G. Boman, *Nature* **292**, 246 (1981).
10. T. Rasmuson and H. G. Boman, *Insect Biochem.* **17**, 133 (1979).
11. T. Trenczek and I. Faye, *ibid.* **18**, 299 (1988); M. R. Kanost, W. Dai, P. E. Dunn, *Arch. Insect Biochem. Physiol.* **8**, 147 (1988).
12. K. Andersson and H. Steiner, *Insect Biochem.* **17**, 133 (1987).
13. D. Barthels *et al.*, *EMBO J.* **6**, 907 (1987).
14. D. L. Bentley and T. H. Rabbitts, *Nature* **288**, 730 (1980).
15. P. H. Hamlyn, M. J. Gait, C. Milstein, *Nucleic Acid Res.* **9**, 4485 (1981).
16. P. Y. Chou and G. D. Fasman, *Adv. Enzymol. Relat. Areas Mol. Biol.* **47**, 45 (1978).
17. G. von Heijne, *J. Mol. Biol.* **184**, 99 (1985).
18. S. A. Needleman and C. D. Wunsch, *ibid.* **48**, 443 (1970).
19. M. Moos *et al.*, *Nature* **334**, 701 (1988).
20. O. Schmidt and I. Lindström, unpublished data.
21. P. Prehm, S. Stirm, B. Jann, K. Jann, H. G. Boman, *Eur. J. Biochem.* **66**, 369 (1976).
22. C. A. Janeway, Jr., *Cold Spring Harbor Symp. Quant. Biol.* **54**, 1 (1989).
23. J. Y. Lee, T. Edlund, T. Ny, I. Faye, H. G. Boman, *EMBO J.* **2**, 577 (1983).
24. R. E. Kingston, in *Current Protocols in Molecular Biology*. F. M. Ausubel *et al.*, Eds. (Wiley, New York, 1987), chapt. 4, sect. 8, pp. 1–3.
25. I. Lindström *et al.*, unpublished data.
26. U. K. Laemmli, *Nature* **227**, 680 (1970).
27. V. N. Burnette, *Anal. Biochem.* **112**, 195 (1981).

28. We thank K. Andersson for providing the antiserum to hemolin; J.-Y. Lee for the plasmid pCP434; K. Andersson, D. Hultmark, S. D. Pereira, and U. Theopold for suggestions; and C. Fernandez, T. Matsunaga, and H. Steiner for critical reading of the manuscript. Supported by NFR grants (H.G.B. and I.F.; and H.G.B. and O.S.) and by a DFG grant O.S.

20 July 1990; accepted 24 October 1990

Dominant Negative Regulation of the Mouse α-Fetoprotein Gene in Adult Liver

JEAN VACHER AND SHIRLEY M. TILGHMAN

Transcription of the mouse α-fetoprotein gene is activated in the developing fetal liver and gut and repressed in both tissues shortly after birth. With germline transformation in mice, a cis-acting element was identified upstream of the transcription initiation site of the α-fetoprotein gene that was responsible for repression of the gene in adult liver. This negative element acts as a repressor in a position-dependent manner.

THE APPROPRIATE DEVELOPMENT OF a multicellular organism requires the expression of genes at the correct time and in the correct cell type. Expression of certain genes must also be repressed at a later stage of development, once their transcripts are no longer required by the organism. The α-fetoprotein (AFP) is activated in the embryonic yolk sac, the fetal liver, and, to a lesser extent, the fetal gut, but is subsequently repressed in both neonatal liver and gut (*1–3*). The decrease in liver AFP mRNA occurs primarily at the level of transcription (*3*).

Two opposing models can be proposed to explain postnatal repression. In one instance, the positive factors necessary for AFP transcription are inactivated after birth, thereby leading to a decrease in AFP mRNA. Alternatively, these factors remain unchanged in concentration, but are no longer able to activate transcription of the gene, due to the presence of a dominant repressor. This is the case for the two silent mating type loci in yeast (*4–6*). Genetic experiments in mice have identified at least one locus (termed *raf*, for regulation of AFP) that might encode such a repressor (*2, 7, 8*).

One way to discriminate between these two models is to ask whether the sequences that are necessary for both positive and negative regulation of the AFP gene coincide, which must be the case in the first model but need not be the case in the second. At least four elements that participate in high-level transcription of the AFP gene have been identified by transfection into tissue culture cells (*9–14*). In the mouse gene, these include a proximal promoter in the first 250 base pairs (bp) upstream of the AFP cap site, as well as three enhancers located upstream of the site of transcription (+1 bp) at −2.5 kilobases (kb), −5.0 kb, and −6.5 kb (*9, 10, 14*). In transgenic mice, these four elements were sufficient to direct high-level expression of a reporter gene in all chromosomal locations into which the gene was integrated (*15*). In addition, the reporter minigene was transcriptionally repressed after birth in a manner indistinguishable from that of the endogenous gene.

As a first test of the two models to explain repression, we established that the AFP enhancers play no obligatory role in the postnatal repression of the gene. The enhancers were used to direct transcription of the evolutionarily related albumin gene, which normally does not undergo repression. No repression in adult livers of transgenic mice was observed. On the other hand, the reciprocal construct that consisted of the albumin gene enhancer upstream of the AFP promoter and minigene exhibited appropriate repression after birth (*16*). These results argued that the enhancers were not responsible for repression, and we focused our attention on the region proximal to the AFP gene itself.

Three constructs bearing the AFP enhancers, an internally deleted AFP minigene (*15*), and different amounts of DNA between −838 and −118 bp were generated

Howard Hughes Medical Institute, Department of Biology, Princeton University, Princeton, NJ 08544.