agno, S. L. Dalton, R. K. Assoian, *ibid.* **122**, 461 (1993).
4. M. Philippe and J. M. Roberts, *Cell* **66**, 1217 (1991); F. Fang and J. W. Newport, *ibid.*, p. 731; C. J. Sherr, *ibid.* **73**, 1059 (1993); E. A. Nigg, *BioEssays* **17**, 471 (1995); J. Pines, *Biochem. J.* **308**, 697 (1995); A. Koff *et al.*, *J. Cell Biol.* **130**, 755 (1995); M. Ohtsubo, A. M. Theodoras, J. Schumacher, J. M. Roberts, M. Pagano, *Mol. Cell. Biol.* **15**, 2612 (1995).
5. Cells were cultured in 10-cm dishes until they were 70% confluent, synchronized by incubating in serum-free medium [Dulbecco's minimum essential medium (DMEM) supplemented with bovine serum albumin (1 mg/ml), bovine insulin (5 μg/ml), transferrin (5 μg/ml), and sodium selenite (5 ng/ml)] for 48 to 72 hours, and then stimulated by transfer into DMEM containing 20% fetal bovine serum (FBS). For the attached cultures, cells were allowed to reattach to tissue culture dishes. For the suspended cultures, cells were gently scraped from tissue culture dishes, separated by centrifugation, and resuspended in DMEM supplemented with 20% FBS and 1% Methocel [J. Rheinwald and H. Green, *Cell* **2**, 287 (1974)]. Hydroxyurea, an inhibitor of DNA synthesis, was added at 1 or 2 mM to the cultures to prevent any of the cells from progressing into the S phase of the cell cycle. The cell suspension was sealed in a 50-ml Corning tube, gassed with sterile 5% CO_2, and incubated in a rolling incubator at 37°C. Methocel (1%) had no effect on the protein kinase activity of CDK2 in attached cells.
6. F. Fang, G. Orend, N. Watanabe, T. Hunter, E. Ruoslahti, unpublished data.
7. Immunoprecipitations and in vitro kinase assays were performed as described [H. Matsushime *et al.*, *Mol. Cell. Biol.* **14**, 2066 (1994)], with the exception that 1 μg of glutathione-S-transferase–pRb(768–928) fusion protein was used in each kinase reaction. Antibodies were from the following sources: Antisera to CDK2 (SC-163), CDK4 (SC-260), CDK5 (SC-173), CDK6, and CDK7 (SC-529) were provided by M. Pagano or were from Santa Cruz Biotechnology; antiserum to cyclin D1 was from S. Reed, Santa Cruz Biotechnology (SC-450), or Upstate Biotechnology (06-137); antibodies to cyclin E (SC-248, SC-198), cyclin H (SC-855), cyclin D2 (SC-181), and cyclin D3 (SC-182) were from Santa Cruz Biotechnology; antiserum to $p21^{Cip1}$ was provided by S. Elledge or obtained from Oncogene Sciences (OP64, OP68) or Pharmingen (15091A); antiserum to $p27^{Kip1}$ was from Transduction Laboratories (K25020) or Santa Cruz Biotechnology (SC-776, SC-528); and antiserum to $p57^{Kip2}$ was provided by W. Harper.
8. T. Kakunaga, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1334 (1978).
9. Y. Gu, J. Rosenblatt, D. O. Morgan, *EMBO J.* **11**, 3995 (1992).
10. KD cells were synchronized in serum-free medium for 2 days, and then cultured attached or suspended for 20 hours in DMEM supplemented with 10% FBS and 1 mM hydroxyurea as described (*5*). The cells were transferred to phosphate-free DMEM supplemented with 20% dialyzed FBS and 1 mM hydroxyurea, and incubated for 1 hour. The cells in each 10-cm dish were then switched to 5 ml of fresh phosphate-free DMEM containing 20% dialyzed FBS, 1 mM hydroxyurea, and 5 mCi of [^{32}P]orthophosphate, and incubated for 6 hours. Suspended cells were incubated similarly except that the phosphate-free medium contained 1% Methocel. Cells were later collected and lysed with radioimmunoprecipitation assay buffer. Immunoprecipitation was performed as described (*7*).
11. W. J. Boyle, P. van der Geer, T. Hunter, *Methods Enzymol.* **201**, 110 (1991).
12. G. Draetta, *Trends Cell Biol.* **3**, 287 (1993).
13. J. Shuttleworth, R. Godfrey, A. Colman, *EMBO J.* **9**, 3233 (1990); R. Y. C. Poon, K. Yamashita, J. P. Adamczewski, T. Hunt, J. Shuttleworth, *ibid.* **12**, 3123 (1993); M. J. Solomon, J. W. Harper, J. Shuttleworth, *ibid.*, p. 3133; R. P. Fisher and D. O. Morgan, *Cell* **78**, 713 (1994); T. P. Mäkelä *et al.*, *Nature* **371**, 254 (1994); R. Y. C. Poon and T. Hunter, *Curr. Biol.* **11**, 1243 (1995).
14. J.-P. Tassan, S. J. Schultz, J. Bartek, E. A. Nigg, *J. Cell Biol.* **127**, 467 (1994).
15. R. Y. C. Poon and T. Hunter, *Science* **270**, 90 (1995).
16. T. Hunter and J. Pines, *Cell* **79**, 573 (1994); S. J. Elledge and J. W. Harper, *Curr. Opin. Cell Biol.* **6**, 847 (1994); C. J. Sherr and J. M. Roberts, *Genes Dev.* **9**, 1149 (1995).
17. K. Polyak *et al.*, *Cell* **78**, 59 (1994); O. Aprelikova, Y. Xiong, E. T. Liu, *J. Biol. Chem.* **270**, 18195 (1995).
18. M. Pagano *et al.*, *Science* **269**, 682 (1995).
19. H. Z. Zhang, G. J. Hannon, D. Beach, *Genes Dev.* **8**, 1750 (1994).
20. $p21^{CIP1}$ inhibited Tyr^{15} phosphorylation in cyclin A–CDK2 complexes by Wee1 in vitro (R. Y. C. Poon and T. Hunter, unpublished data).
21. I. Reynisdóttir, K. Polyak, A. Iavarone, J. Massagué, *Genes Dev.* **9**, 1831 (1995).
22. T. M. Guadagno, M. Ohtsubo, J. M. Roberts, R. K. Assoian, *Science* **262**, 1572 (1993).
23. J. Pines and T. Hunter, *Nature* **346**, 760 (1991); F. Girard, U. Strausfeld, A. Fernandez, N. J. C. Lamb, *Cell* **67**, 1169 (1991); M. Pagano, R. Pepperkok, F. Verde, W. Ansorge, G. Draetta, *EMBO J.* **11**, 961 (1992).
24. J. E. Meredith, B. Fazeli, M. A. Schwartz, *Mol. Biol. Cell* **4**, 953 (1993); S. M. Frisch and H. Francis, *J. Cell Biol.* **124**, 619 (1994); E. Ruoslahti and J. C. Reed, *Cell* **77**, 477 (1994).
25. M. D. Schaller and J. T. Parsons, *Trends Cell Biol.* **3**, 258 (1993); E. A. Clark and J. S. Brugge, *Science* **268**, 233 (1995).
26. We thank J. Wang for the glutatathione-S-transferase–pRb (768–928) fusion protein, and M. Pagano, S. Reed, S. Elledge, W. Harper, R. Brent, and R. Poon for antibodies and other help with this work. Supported by grants CA 60725, CA 67224, and CA 28896 and Cancer Center Support Grant CA 30199 (E.R.), as well as grant CA 37980 (T.H.), from the National Cancer Institute. T.H. is an American Cancer Society research professor. N.W. is a staff scientist of the Tsukuba Life Science Center at the Institute of Physical and Chemical Research (RIKEN) in Tsukuba, Japan, and is a recipient of a long-term fellowship from the Human Frontier Science Program. G.O. was supported by a fellowship from the Deutsche Forschungsgemeinschaft, Germany.

5 October 1995; accepted 15 November 1995

Angiotensin II–Forming Activity in a Reconstructed Ancestral Chymase

Unnikrishnan M. Chandrasekharan, Subramaniam Sanker, Manuel J. Glynias, Sadashiva S. Karnik, Ahsan Husain*

The current model of serine protease diversity theorizes that the earliest protease molecules were simple digestive enzymes that gained complex regulatory functions and restricted substrate specificities through evolution. Among the chymase group of serine proteases are enzymes that convert angiotensin I to angiotensin II, as well as others that simply degrade angiotensins. An ancestral chymase reconstructed with the use of phylogenetic inference, total gene synthesis, and protein expression had efficient and specific angiotensin II–forming activity (turnover number, about 700 per second). Thus, angiotensin II–forming activity is the more primitive state for chymases, and the loss of such activity occurred later in the evolution of some of these serine proteases.

Chymases are a family of closely related mast cell serine proteases that are involved in diverse functions such as peptide hormone processing (*1–4*), the inflammatory response (*5*), and parasite expulsion (*6*). Although all chymases resemble chymotrypsin in that they hydrolyze peptide bonds at the COOH-termini of hydrophobic aromatic residues (such as Phe, Tyr, and Trp), the use of an extended substrate-binding site allows a remarkable degree of selectivity in the bonds that are ultimately hydrolyzed (*7*). For example, human and baboon chymases efficiently form the potent vasoconstrictor hormone angiotensin (Ang) II by cleaving the Phe^8–His^9 bond in Ang I (*2*). In the primate heart and its vessels, chymase is a major Ang II–forming enzyme (*8*, *9*). In contrast to chymotrypsin, which degrades Ang II, the high specificity of human chymase is illustrated by the fact that it does not cleave the Tyr^4–Ile^5 bond in Ang II, and thus the generated Ang II is not degraded (*2*). Rat chymase-1 readily degrades Ang I and Ang II (*3*).

Is the specificity observed in some chymases (that is, the formation of Ang II without its subsequent degradation) an early or late feature of evolution? Parsimony-based reconstruction of ancestral ribonuclease (*10*) suggested that a similar approach could be used to reconstruct the original state for chymase evolution. We created an alignment of chymase and related leukocyte serine protease sequences by using several conserved primary and secondary structural motifs as anchors (Fig. 1). The most parsimonious phylogenetic tree of leukocyte serine proteases was inferred by parsimony analysis, which has been shown to accurately predict known phylogeny (*11*). The bootstrap value (95%) supports the segregation of chymases as a distinct group of enzymes within the leukocyte serine protease family of enzymes. Within the chymase group of enzymes, human chymase, baboon chymase, dog chymase, and mouse chymase-5 form a distinct subgroup that we have classified as α-chymases (bootstrap value, 98%). Although we classified the remaining known mammalian chymases as β-chymases, the

Department of Molecular Cardiology, Mail Code FF3-30, Research Institute, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA.

*To whom correspondence should be addressed.

moderate bootstrap value (78%) assigned to this branch makes it less certain whether the β-chymases represent a homogeneous subgroup. However, we believe that the topology of evolutionary relations predicted by parsimony is the best estimate of the true phylogeny of the leukocyte serine protease sequences, because analysis of these sequences by the unweighted pair-group method of arithmetic averages yielded a tree with an identical topology.

The predicted sequence of the 226-residue ancestral chymase (Fig. 2) was reconstructed by parsimony analysis (node ψ, Fig. 1B); 93.4% of these residues were determined unambiguously by parsimony analysis. Assignments for the remaining 15 residues were made from chymase sequences (Fig. 2A). The ancestral chymase differs from known chymases by 52 to 100 residues (23 to 34% difference) (Fig. 2B). Because of the number and wide distribution of these changes (Fig. 2A), we found it necessary to chemically synthesize the entire gene for the ancestral chymase. This gene was constructed with the human chymase signal and activation peptide sequences, and re-

A

```
                                                                                                                                 1111111111111
                  11111111112222222222333333333344444444    4455555555556666666666777777777788888888           8899999999990000000000111
        1234567890123456789012345678901234567890123456789abcd8901234567890123456789012345678901234567abcdefghijk8901234567890123456789012
        * **      * *                * * *     * *****            **       *   *                               *  *  *         *
m-GrG   IIGGHEVKPHSRPYMAFIKSVDIEGKKKYCGGFLVQDDFVLTAAHCR----NRSMTVTLGAHNIKAKEETQQIIPVAKAIPHPAFNRKHGT-----------NDIMLLKLESKAKRTKAVRPLKLPR
m-GrF   IIGGHEVKPHSRPYMARVRFVKDNGKRHSCGGFLVQDYFVLTAAHCT----GSSMRVILGAHNIRAKEETQQIIPVAKAIPHPAYDDKDNT-----------SDIMLLKLESKAKRTKAVRPLKLPR
m-GrE   IIGGHVVKPHSRPYMAFVKSVDIEGNRRYCGGFLVQDDFVLTAAHCR----NRTMTVTLGAHNIKAKEETQQIIPVAKAIPHPDYNATAFF-----------SDIMLLKLESKAKRTKAVRPLKLPR
m-GrD   IIGGHVVKPHSRPYMAFVMSVDIKGNRIYCGGFLIQDDFVLTAAHCKNSSVQSSMTVTLGAHNITAKEETQQIIPVAKDIPHPDYNATIFY-----------SDIMLLKLESKAKRTKAVRPLKLPR
m-GrB   IIGGHEVKPHSRPYMALLSIKDQQPEA-ICGGFLIREDFVLTAAHCE----GSIINVTLGAHNIKEQEKTQQVIPMVKCIPHPDYNPKTFS-----------NDIMLLKLKSKAKRTRAVRPLNLPR
r-KCP1  IIGGHEAKPHSRPYMAYLQIMDEYSGSKKCGGFLIREDFVLTAAHCS----GSKINVTLGAHNIKEQEKMQQIIPVVKIIPHPAYNSKTIS-----------NDIMLLKLKSKAKRSSAVKPLNLPR
m-GrC   IIGGNEISPHSRPYMAYYEFLKVGGKKMFCGGFLVRDKFVLTAAHCK----GSSMTVTLGAHNIKAKEETQQIIPVAKAIPHPDYNPDDRS-----------NDIMLLKLVRNAKRTRAVRPLNLPR
h-GrB   IIGGHEAKPHSRPYMAYLMIWDQKSLK-RCGGFLIQDDFVLTAAHCW----GSSINVTLGAHNIKEQEPTQQFIPVKRPIPHPAYNPKNFS-----------NDIMLLQLERKAKRTRAVQPLRLPS
m-Chy2  IIGGVEAKPHSRPYMAYLKFTTKNGSKERCGGFLIAPQFVMTAAHCR----GSEISVILGAHNINKNEPTQQIIKTEKTFVHPKFQYLSGF-----------NDIMLLKLQKKAELNSDVDVISLPS
r-Chy2  IIGGVESIPHSRPYMAHLDIVTEKGLRVICGGFLISRQFVLTAAHCK----GREITVILGAHDVRKRESTQQKIKVEKQIIHESYNSVPNL-----------HDIMLLKLEKKVELTPAVNVVPLPS
m-Chy1  IIGGVEARPHSRPYMAHLKIITDRGSEDRCGGFLIAPQFVLTAAHCK----GREITVTLGAHDVSKSESTQQRIKVEKQIIHKNYNVSFNL-----------YDIMLLKLEEKAELTPTVDVIPLPG
r-Chy1  IIGGVESRPHSRPYMAHLEITTERGYKATCGGFLVTRQFVMTAAHCK----GRETTVTLGVHDVSKTESTQQKIKVEKQIVHPNYNFYSNL-----------HDIMLLKLQKKAKVTPAVDVIPLPQ
m-Chy4  IIGGVESRPHSRPYMAHLEITTERGFTATCGGFLITRQFVMTAAHCS----GREITVTLGAHDVSKTESTQQKIKVEKQIVHPKYNFYSNL-----------HDIMLLKLQKKAKETPSVNVIPLPR
m-ChyL  IIGGVESEPHSRPYMAYVNTFRRKGYVAICGGFLITPQFVMTAAHCR----GRRMTVTLGAHNVRKRECTQQKIKVEKYILPPNYNVSSKF-----------NDIVLLKLKKQANLTSAVDVVPLPG
h-Chy   IIGGTECKPHSRPYMAYLEIVTSNGPSKFCGGFLIRRNFVLTAAHCA----GRSITVTLGAHNITEEEDTWQKLEVIKQFRHPKYNTSTLH-----------HDIMLLKLKEKASLTLAVGTLPFPS
b-Chy   IIGGTECKPHSRPYMAYLEIVTSNGPSKSCGGFLIRRNFVLTAAHCA----GRSITVTLGAHNITEKEDTWQELEVIKQFRHPKYNTSTLH-----------HDIMLLKLKEKASLTLAVGTLPFPS
d-Chy   IIGGTESKPHSRPYMAHLEILTLRNHLASCGGFLIRRNFVLTAAHCA----GRFIMVTLGAHNIQKKEDTWQKLEVIKQFPHPKYDDLTLR-----------HDIMLLKLKEKANLTLAVGTLPLSP
m-Chy5  IIGGTECIPHSRPYMAYLEIVTSENYLSACSGFLIRRNFVLTAAHCA----GRSITVLLGAHNKTSKEDTWQKLEVEKQFLHPKYDENLVV-----------HDIMLLKLKEKAKLTLGVGTLPLSA
h-CathG IIGGRESRPHSRPYMAYLQIQSPAGQS-RCGGFLVREDFVLTAAHCW----GSNINVTLGAHNIQRRENTQQHITARRAIRHPQYNQRTIQ-----------NDIMLLQLSRRVRRNRNVNPVALPR
m-GrA   IIGGDTVVPHSRPYMALLKLSSNT----ICAGALIEKNWVLTAAHCN---VGKRSKFILGAHSINK-EPEQQILTVKKAFPYPCYDEYTRE-----------GDLQLVRLKKKATVNRNVAILHLPK
h-GrA   IIGGNEVTPHSRPYMVLLSLDRKT----ICAGALIAKDWVLTAAHCN---LNKRSQVILGAHSITREEPTKQIMLVKKEFPYPCYDPATRE-----------GDLKLLQLTEKAKINKYVTILHLPK
m-Kal5  IFGGFNCEKNSQPWQVAVYRFTKY----QCGGVLLNANWVLTAAHCH----NDKYQVWLGKNNFFEDEPSAQHRLVSKAIPHPDFNMSLLNEHTPQEDDYSNDLMLLRLKKPADITDVVKPIDLPT
m-Kalr  IVGGFNCEKNSQPWQVAVYRFTKY----QCGGILLNANWVLTAAHCH----NDKYQVWLGKNNFLEDEPSAQHRLVSKAIPHPDFNMSELNEHTPQEDDYSNDLMLLRLKKPADITDVVKPIDLPT

        11111111111111111 111111111111111111111111111     1 111111111111111111111111111111111111111 111  222222  22222222 2222222222222
        11111112222222222 333333333344444444445555555     5 556666666666777777777788888888889999999 999  000000  00001111 1111112222222
        34567890123456789a012345678901234567890123456abcde7a890123456789012345678901234567890123456a789ab012345ab67890123a4567890123456
               *  *   ***            *            *                  * *            ***** *     *                             **
m-GrG   PNARVKPGDVCSVAGWGKTSINATKASARLREAQLIIQEDEECKKL----W-YTYSKTTQICAGDPKKVQAPYEGESGGPLVCDNLAYGVV-SYG--INRTIT--PGVFTKVV-HFLPWISTNMKLL
m-GrF   PNARVKPGHVCSVAGWGRTSINATQRSSCLREAQLIIQKDKECKKY----F-YKYFKTMQICAGDPKKIQSTYSGDSGGPLVCNNKAYGVL-TYG--LNRTIG--PGVFTKVV-HYLPWISRNMKLL
m-GrE   PNARVKPGDVCSVAGWGSRSINDTKASARLREAQLVIQEDEECKKR----F-RHYTETTEICAGDLKKIKTPFKGDSGGPLVCDNKAYGLL-AYA--KNRTIS--SGVFTKIV-HFLPWISRNMKLL
m-GrD   SNARVKPGDVCSVAGWGSRSINDTKASARLREVQLVIQEDEECKKR----F-RYYTETTEICAGDLKKIKTPFKGDSGGPLVCHNQAYGLF-AYA--KNGTIS--SGIFTKVV-HFLPWISWNMKLL
m-GrB   RNVNVKPGDVCYVAGWG-RMAPMGKYSNTLQEVELTVQKDRECESY----FKNRYNKTNQICAGDPKTKRASFRGDSGGPLVCKKVAAGIV-SYG--YKDGSP--PRAFTKVS-SFLSWIKKTMKSS
r-KCP1  RNVKVKPGDVCYVAGWG-KLGPMGKYSDTLQEVELTVQEDQKCESY----LKNYFDKANEICAGDPKIKRASFRGDSGGPLVCKKVAAGIV-SYG--QNDGST--PRAFTKVS-TFLSWIKKTMKKS
m-GrC   RNAHVKPGDECYVAGWG-KVTPDGEFPKTLHEVKLTVQKDQVCESQ----FQSSYNRANEICVGDSKIKGASFEEDSGGPLVCKRAAAGIV-SYG--QTDGSA--PQVFTRVL-SFVSWIKKTMKHS
h-GrB   NKAQVKPGQTCSVAGWG-QTAPLGKHSHTLQEVKMTVQEDRKCESD----LRHYYDSTIELCVGDPEIKKTSFKGDSGGPLVCNKVAQGIV-SYG--RNNGMP--PRACTKVS-SFVHWIKKTMKRY
m-Chy2  SSDFIKPGKMCWTAGWG-KTGKNNPLSVTLREVELRIMDQEACKD-----H-NDYDYQLQVCAGSPTTLKSIGQGDSGGPLVCDGVAHGIASSY-----EAKA--PAVFTRIS-YYLPWIYKVLKSK
r-Chy2  PSDFIHPGAMCWAAGWG-KTGVRDPTSYTLREVELRIMDEKACVD-----Y-RYYEYKFQVCVGSPTTLRAAFMGDSGGPLLCAGVAHGIV-SYG--HPDAKP--PAIFTRVS-TYVPWINAVINTS
m-Chy1  PSDFIDPGKMCWTAGWG-KTGEKEPTSETLREVELRIMDKEACKM-----Y-KHYDYNFQVCVGSSTKLKTAYMGDSGGPLLCAGVAHGIV-SYG--DSHGKP--PAVFTRIS-AYVPWIKTVINGK
r-Chy1  PSDFLKPGKMCRAAGWG-QTGVTKPTSNTLREVKQRIMDKEACKN-----Y-FHYNYNFQVCVGSPRKIRSAYKGDSGGPLVCAGVAHGIV-SYG--RGDAKP--PAVFTRIS-PYVPWINKVIKGK
m-Chy4  PSDFIKPGKMCRAAGWG-RTGVTEPTSDTLREVKLRIMDKEACKN-----Y-WHYDYNLQVCVGSPRKKRSAYKGDSGGPLLCAGVAHGIV-SYG--RGDAKP--PAVFTRIS-SYVPWINRVIKGE
m-ChyL  PSDFAKPGTMCWAAGWG-RTGVKKIISHTLREVELKIVGEKACKI-----F-RHYKDSLQICVGSSTKVASVYMGDSGGPLLCAGVAHGIV-SSG--RGNAKP--PAIFTRIS-PHVPWINRVIKGK
h-Chy   QFNFVPPGRMCRVAGWG-RTGVLKPGSDTLQEVKLRLMDPQACSH-----F-RDFDHNLQLCVGNPRKTKSAFKGDSGGPLLCAGVAQGIV-SYG--RSDAKP--PAVFTRIS-HYRPWINQILQAN
b-Chy   QFNFVPPGRMCRVAGWG-RTGVLKPGSDTLQEVKLRLMDPQACSH-----F-RYFDHNLQLCVGNPRKTKSAFKGDSGGPLLCAGVAQGIV-SYG--RLDAKP--PAVFTRIS-HYRPWINKILQAN
d-Chy   QFNFVPPGRMCRVAGWG-KRQVNGSGSDTLQEVKLRLMDPQACRH-----Y-MAFDHNLQLCVGNPRKTKSAFKGDSGGPLLCAGVAQGIV-SYG--QNDAKP--PAVFTRIS-HYRPWINKVLKQN
m-Chy5  NFNFIPPGRMCRAVGWG-RTNVNEPASDTLQEVKMRLQEPQACKH-----F-TSFRHNSQLCVGNPKKMQNVYKGDSGGPLLCAGIAQGIA-SYV--HRQAKP--PAVFTRIS-HYRPWINKILREN
h-CathG AQEGLRPGTLCTVAGWG-RVSMRR-GTDTLREVQLRVQRDRQCLRI----F-GSYDPRRQICVGDRRERKAAFKGDSGGPLLCNNVAHGIV-SYG--KSSGVP--PEVFTRVS-SFLPWIRTTMRSF
m-GrA   KGDDVKPGTRCRVAGWG-RFGNKSAPSETLREVNITVIDRKICNDEKHYNF-HPVIGLNMICAGDLRGGKDSCNGDSGSPLLCDGILRGIT-SFGGE-KCGDRRWPGVYTFLSDKHLNWIKKIMKGS
h-GrA   KGDDVKPGTMCQVAGWG-RTHNSASWSDTLREVNITIIDRKVCNDRNHYNF-NPVIGMNMVCAGSLRGGRDSCNGDSGSPLLCEGVFRGVT-SFGLENKCGDPRGPGVYILLSKKHLNWIIMTIKGA
m-Kal5  EEP--KLGSTCLASGWGSITPVIYEPADDLQCVNFKLLPNEDCVKA----H-IEKVTDVMLCAGDMDGGKDTCMGDSGGPLICDGVLHGIT-SWG-PSPCGKPNVPGIYTKLI-KFNSWIKDTIAKN
m-Kalr  EEP--KLGSTCLASGWGSITPVKYEYPDELQCVNLKLLPNEDCAKA----H-IEKVTDDMLCAGDMDGGKDTCAGDSGGPLICDGVLQGIT-SWG-PRPCGKPNVPGIYTRVL-NFNTWIRETMAEN
```

B

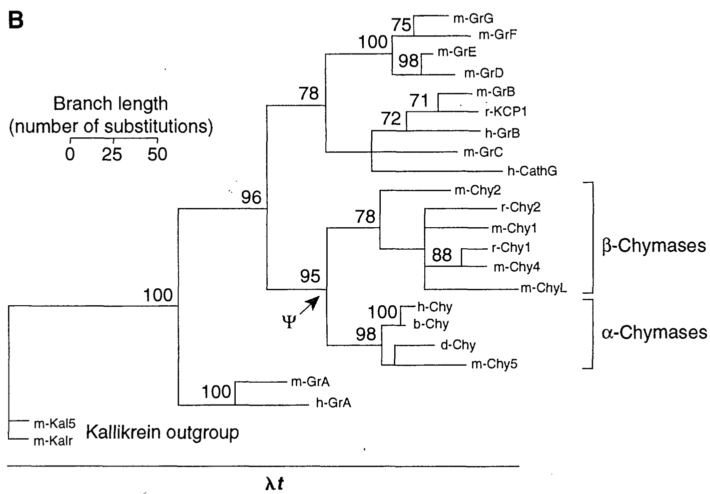

Fig. 1. Alignment and phylogenetic relations of several leukocyte serine proteases (*17*). (**A**) Alignment of chymases (Chy), granzymes [Gr and KCP (killer cell protease)], cathepsin G (CathG), and kallikreins (Kal). The prefixes h, b, d, r, and m refer to sequences from human, baboon, dog, rat, and mouse, respectively [see (*13*) for sources]. The numbering system used here is based on the sequence of the mature human chymase. Asterisks denote residues that are conserved in all serine proteases described here; dashes indicate alignment gaps. Enzyme sequences were trimmed to remove signal and activation peptide sequences. Where necessary, COOH-terminal sequences of some enzymes beyond residue 226 (chymase numbering) were also deleted. (**B**) Phylogenetic relations within leukocyte serine proteases as derived from parsimony analysis of the 23 aligned enzyme sequences shown in (A). Parsimony trees were inferred with PAUP (*18*). The trees were rooted with two closely related mouse kallikrein sequences as an outgroup. A heuristic method was used to find the shortest trees. However, because heuristic parsimony algorithms can get trapped in local optima (*18*), multiple starting trees were generated through random addition. One most parsimonious tree of 1171 steps was found by this method. The bootstrap resampling method (*19*) was used to evaluate the support for the internal branches of this tree. Bootstrap percentages from 525 replications are shown on each supported branch; values below 60 are omitted for clarity. ψ indicates the ancestral chymase node. The bar labeled λ*t* is the distance from the root to the most divergent tip.

A

```
....‡.**.. ......**‡* **‡***‡*‡. *...*†*†.. *.....‡.*†
IIGGVESKPH SRPYMAYLKI VTENGYKARC GGFLIRRNFV LTAAHCRGRS  50

**.*..*.** ‡*†.‡.‡.*‡ ‡**.**†*** **†*†††‡.. *....*****
ITVTLGAHNI NKKEPTQQKI KVEKQFPHPK YNDSTLFNDI MLLKLKEKAK 100

**‡*.‡††** **‡††.†‡.. ‡..*†*...* ******†.*. .‡..‡**‡**
LTSAVDTLPL PSPFDFVKPG KMCRVAGWGR TGVNEPGSDT LREVELRIMD 150

‡‡..*‡***‡ *‡**.*.*.‡ ***‡*****. ......*.*. *.‡..*.***
EKACKDYRDY DYNLQLCVGS PRKLKSAFKG DSGGPLLCAG VAHGIVSYGR 200

†****..*.. .*.†*‡...* ****‡‡
NDAKPPAVFT RISHYLPWIN KVLKSK                            226
```

B

	an-Chy	d-Chy	h-Chy	b-Chy	m-Chy5	r-Chy1	r-Chy2	m-Chy1	m-Chy2	m-ChyL
d-Chy	75	–								
h-Chy	77	82	–							
b-Chy	77	82	97	–						
m-Chy5	66	73	75	76	–					
r-Chy1	72	61	61	61	56	–				
r-Chy2	69	58	59	59	54	72	–			
m-Chy1	69	57	58	58	53	74	74	–		
m-Chy2	71	54	54	54	49	63	60	65	–	
m-ChyL	68	53	56	55	51	68	65	65	59	–
m-Chy4	75	64	65	64	59	89	74	76	64	68

Fig. 2. (**A**) Primary structure of the ancestral chymase at node ψ (Fig. 1B) predicted by parsimony analysis. Except for eight residues indicated below, the ancestral chymase sequence was inferred through PAUP (*17, 18*). The numbering system used is based on the sequence of the mature human chymase. Residues conserved in all chymases are indicated by a dot. Nonconserved residues in chymases that are found in one or more α- and β-chymases are indicated by an asterisk; those found in one or more α- but not β-chymases are indicated by the symbol †; and those found in one or more β- but not α-chymases are indicated by the symbol ‡. For 15 residues (underlined), parsimony-based assignments were ambiguous. Assignment for seven of these residues was arbitrarily determined by PAUP. The remaining eight residues were assigned to adjust the net positive charge of the ancestral chymase to +18. The net positive charge on a typical chymase is between +12 and +22 and is organized in two clusters. These positively charged clusters may play an important role in binding to heparin within the mast cell granule and in heparin-dependent zymogen activation (*13*). The substrate-binding cleft is located centrally but is far from these two charged clusters (*20*). In the final construction, 3 of the 15 ambiguously assigned residues were found exclusively in α-chymases; the rest were found exclusively in β-chymases. (**B**) Similarity of the ancestral chymase to known chymases. Values show percentage identity among known chymases and the ancestral chymase (an-Chy, ancestral chymase; other abbreviations as in Fig. 1).

combinant protein was generated (*12, 13*).

Kinetic studies revealed two properties of the ancestral chymase. First, the ancestral chymase specifically cleaves the Phe[8]–His[9] bond in Ang I to generate Ang II, but does not cleave the Tyr[4]–Ile[5] bond in Ang I and II that is sensitive to rat chymase-1 and chymotrypsin (Fig. 3 and Table 1). Ancestral chymase–dependent cleavage of the Phe[8]–His[9] bond is at least 1000 times as fast as cleavage of the Tyr[4]–Ile[5] bond (Table 1). Second, the turnover number for the cleavage of the Phe[8]–His[9] bond by the ancestral chymase ($k_{cat} \approx 700\ s^{-1}$) is about five times that for human chymase and about 80 times that for Ang I–converting enzyme (*14*); indeed, it is one of the highest turnover numbers reported for any protease.

We previously showed that multiple determinants, both in the acyl and leaving groups of the substrate, are necessary for the high specificity of human chymase (*7*), which implies that the extended substrate-binding site of human chymase is involved in determining substrate specificity. Because molecular modeling showed a mosaic of α- and β-chymase residues in the extended substrate-binding site of the ancestral chymase, it is unclear which residues in this site are responsible for the observed specificity. We believe that the 15 residues in the ancestral chymase sequence that could not be assigned unambiguously (Fig. 2) did not bias the kinetic results in favor of Ang II–forming activity (α-chymase phenotype), because (i) only 3 of these 15 residues were from α-chymases, and (ii) molecular modeling indicated that none of these 15 residues occur in the substrate binding site, and hence they are unlikely to alter specificity. These results imply that before α- and β-chymases diverged, the ancestral chymase was an efficient Ang II–forming enzyme. Thus, chymases with lower substrate specificity [for example, rat chymase-1 (*3*)] must have evolved later for less discrete digestive functions; however, more reconstructions

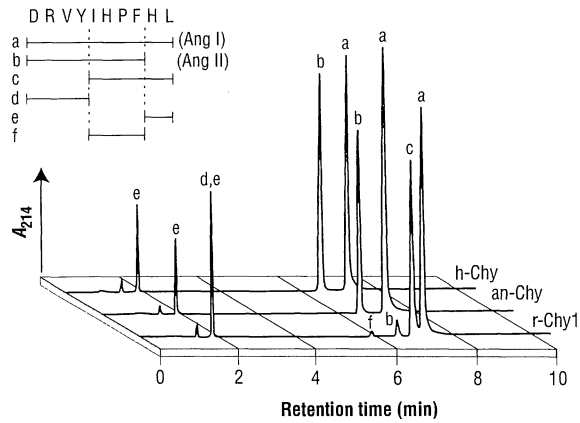

Fig. 3. Differences in specificities of rat chymase-1 (r-Chy1), human chymase (h-Chy), and ancestral chymase (an-Chy) toward angiotensins as shown by high-performance liquid chromatography (HPLC) profiles of degradation of Ang I. Time-dependent incubations (20 to 60 min) were performed with each protease (2 to 100 fmol) in assay buffer (20 mM tris-HCl buffer, pH 8.0, containing 0.5 M KCl and 0.1% Triton X-100) containing 500 μM Ang I; the products were resolved by reversed-phase HPLC as described (*7*). Only one time point, representing a ~50% degradation of Ang I, is shown for each chymase. HPLC peaks occurring at the retention time of Ang I, Ang II, Ang I-(5–10), Ang I-(1–4), Ang I-(9–10), and Ang I-(5–8) synthetic standards are labeled a, b, c, d, e, and f, respectively [the residues contained in each standard are shown (*17*)]. Note that rat chymase-1 cleaves the Tyr[4]–Ile[5] bond in Ang I more rapidly than it cleaves the Phe[8]–His[9] bond, thereby degrading Ang I at a much faster rate than it can form Ang II. Human chymase and ancestral chymase only cleave the Phe[8]–His[9] bond in Ang I, thereby leading to a stable formation of Ang II. A_{214}, absorbance at 214 nm.

Table 1. Kinetic constants for Ang II formation from Ang I by human and ancestral chymases, and for Ang II degradation by rat chymase-1 ($n = 3$; K_m, Michaelis constant). Kinetic constants were determined as described (*7*). A k_{cat} value for the degradation of Ang I by rat chymase-1 could not be determined (ND) because two peptide bonds were split simultaneously.

Enzyme	Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\ s^{-1}$)
Human chymase	Ang I	40 ± 1.3	146 ± 1.9	3.6
Ancestral chymase	Ang I	166 ± 16	720 ± 65	4.3
Rat chymase-1	Ang I	–	ND	–
Human chymase	Ang II	–	<0.05*	–
Ancestral chymase	Ang II	–	<0.05*	–
Rat chymase-1	Ang II	55 ± 1.6	4.7 ± 0.04	0.085

*Minimum k_{cat} detectable was 0.05 s^{-1}.

are required to determine at what point in evolution this loss of specificity occurred.

Only a single α-chymase gene is present in humans and baboons (*8*) and only a single α-chymase has been described in dogs. In contrast, five chymase isoenzymes have been identified in mice and two in rats. Four of the five mouse chymases are of the β subtype, and one is of the α subtype; both rat chymases are of the β subtype (*13*). One possibility that could explain this distribution of chymase isoenzymes is that the α- and β-chymases split apart when rodents branched off from other mammals. We believe, however, that the division into α- and β-isoenzymes occurred long before mammals branched off from therapsids because mouse chymase-5 segregated from other rodent chymases as an α-chymase (bootstrap value, 98%; Fig. 1B). This hypothesis implies that humans and baboons have lost their β-chymase genes and that rats have lost their α-chymase gene. Ang II contributes to cardiovascular regulation in several nonmammalian species (*15*), but pathways for its synthesis have been explored largely in mammals. Because mast cells occur in frogs, birds, and lizards (*16*) and because the reconstructed ancestral chymase is an efficient Ang II–forming enzyme, we speculate that a chymase-dependent pathway of Ang II formation occurred early in vertebrate evolution.

REFERENCES AND NOTES

1. C. F. Reilly, D. A. Tewksbury, N. M. Schechter, J. Travis, *J. Biol. Chem.* **257**, 8619 (1982).
2. H. Urata, A. Kinoshita, K. S. Misono, F. M. Bumpus, A. Husain, *ibid.* **265**, 22348 (1990).
3. H. Le Trong, H. Neurath, R. G. Woodbury, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 364 (1987).
4. G. H. Caughey, F. Leidig, N. F. Viro, J. A. Nadel, *J. Pharmacol. Exp. Ther.* **244**, 133 (1988).
5. S. D. Brain and T. J. Williams, *Nature* **335**, 73 (1988).
6. R. G. Woodbury *et al.*, *ibid.* **312**, 450 (1984).
7. A. Kinoshita, H. Urata, F. M. Bumpus, A. Husain, *J. Biol. Chem.* **266**, 19192 (1991).
8. H. Urata *et al.*, *J. Clin. Invest.* **91**, 1269 (1993).
9. B. D. Hoit *et al.*, *ibid.* **95**, 1519 (1995).
10. T. M. Jermann, J. G. Opitz, J. Stackhouse, S. A. Benner, *Nature* **374**, 57 (1995).
11. D. M. Hillis, J. J. Bull, M. E. White, M. R. Badgett, I. J. Molineux, *Science* **255**, 589 (1992).
12. H. G. Khorana, *ibid.* **203**, 614 (1979).
13. M. Murakami, S. S. Karnik, A. Husain, *J. Biol. Chem.* **270**, 2218 (1995) and references therein.
14. E. G. Erdös and R. A. Skidgel, *Biochem. Soc. Trans.* **13**, 42 (1985).
15. I. W. Henderson, J. A. Oliver, A. McKeever, N. Hazon, in *Advances in Animal and Comparative Physiology*, G. Pethes and V. L. Frenyo, Eds. (Pergamon, New York, 1981), pp. 355–363.
16. L. M. Pastor *et al.*, *J. Submicrosc. Cytol. Pathol.* **20**, 25 (1988); R. Hakanson *et al.*, *Histochemistry* **86**, 5 (1986).
17. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
18. D. L. Swofford, *PAUP: Phylogenetic Analysis Using Parsimony*, version 3.0 (Center for Biodiversity, Illinois Natural History Survey, Champaign, IL, 1992).
19. J. Felsenstein, *Evolution* **39**, 783 (1985).
20. S. J. Remmington, R. G. Woodbury, R. A. Reynolds, B. W. Matthews, H. Neurath, *Biochemistry* **27**, 8097 (1988).

21. Supported by NIH grants HL44201 and HL33713.

We thank W. Gilbert for helpful criticism, D. Wilk for excellent technical assistance, and C. Kassuba for editorial assistance.

21 August 1995; accepted 21 November 1995

Molecular Cloning and Disease Association of Hepatitis G Virus: A Transfusion-Transmissible Agent

Jeff Linnen, John Wages Jr., Zhen-Yong Zhang-Keck, Kirk E. Fry, Krzysztof Z. Krawczynski, Harvey Alter, Eugene Koonin, Margaret Gallagher, Miriam Alter, Stephanos Hadziyannis, Peter Karayiannis, Kevin Fung, Yoshiyuki Nakatsuji, J. Wai-Kuo Shih, Lavonne Young, Michael Piatak Jr., Cameron Hoover, John Fernandez, Stacie Chen, Jian-Chao Zou, Timothy Morris, Kenneth C. Hyams, Susan Ismay, Jeffrey D. Lifson, Georg Hess, Steven K. H. Foung,* Howard Thomas, Daniel Bradley, Harold Margolis, Jungsuh P. Kim†

An RNA virus, designated hepatitis G virus (HGV), was identified from the plasma of a patient with chronic hepatitis. Extension from an immunoreactive complementary DNA clone yielded the entire genome (9392 nucleotides) encoding a polyprotein of 2873 amino acids. The virus is closely related to GB virus C (GBV-C) and distantly related to hepatitis C virus, GBV-A, and GBV-B. HGV was associated with acute and chronic hepatitis. Persistent viremia was detected for up to 9 years in patients with hepatitis. The virus is transfusion-transmissible. It has a global distribution and is present within the volunteer blood donor population in the United States.

Although sensitive and specific tests for detection of the known hepatitis viruses are available (*1*), the etiology of a substantial fraction of post-transfusion (*2*) and community-acquired hepatitis (*3*) cases has remained undefined, suggesting the existence of additional causative agents. To identify such an agent, designated hepatitis G virus (HGV) (*4*), molecular cloning was initially performed with plasma from a patient designated PNF2161, who was originally identified as having non-A, non-B viral hepatitis through the Centers for Disease Control and Prevention (CDC) Sentinel Counties Study of Viral Hepatitis (*3*). Patient PNF2161 was initially believed not to be infected with hepatitis C virus (HCV), on the basis of consistently negative results with a first-generation immunoassay (the Ortho HCV ELISA Test System; Ortho Diagnostics, Raritan, New Jersey). However, subsequent testing with a second-generation HCV immunoassay (also from Ortho) and a polymerase chain reaction (PCR) assay based on HCV 5′ untranslated region primers (*5*) demonstrated that PNF2161 was infected with HCV.

Library construction and immunoscreening with plasma from PNF2161 were performed as described (*6*). Sequence analysis of immunoreactive clones isolated from the PNF2161 λ gt11 library revealed HCV-related sequences as well as several sequences that did not match any in the GenBank database (*7*). PCR primers designed from these nonmatching sequences were used to determine that they were exogenous to

J. Linnen, J. Wages Jr, Z.-Y. Zhang-Keck, K. E. Fry, K. Fung, L. Young, M. Piatak Jr, C. Hoover, J. Fernandez, S. Chen, J.-C. Zou, J. D. Lifson, S. K. H. Foung, J. P. Kim, Genelabs Technologies, Redwood City, CA 94063, USA.
K. Z. Krawczynski, M. Gallagher, M. Alter, T. Morris, H. Margolis, Hepatitis Branch, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA.
H. Alter, Y. Nakatsuji, J. W.-K. Shih, Department of Transfusion Medicine, National Institutes of Health, Bethesda, MD 20205, USA.
E. Koonin, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA.
S. Hadziyannis, Academic Department of Medicine, Hippokration Hospital, Athens, Greece.
P. Karayiannis and H. Thomas, Department of Medicine, St. Mary's Hospital Medical School, London, UK.
K. C. Hyams, Naval Medical Research Institute, Rockville, MD 20852, USA.
S. Ismay, New South Wales Red Cross Blood Transfusion Service, Sydney, Australia.
G. Hess, Boehringer Mannheim GmbM, D68305 Mannheim, Germany.
D. Bradley, 2938 Kelly Court, Lawrenceville, GA 30244, USA.

*Present address: Department of Pathology, Stanford University, Stanford, CA 94304, USA.
†To whom correspondence should be addressed.