THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 34, Issue of August 26, pp. 30367–30375, 2005
Printed in U.S.A.

Identification and Characterization of Human Archaemetzincin-1 and -2, Two Novel Members of a Family of Metalloproteases Widely Distributed in Archaea*

Received for publication, April 26, 2005, and in revised form, June 13, 2005
Published, JBC Papers in Press, June 22, 2005, DOI 10.1074/jbc.M504533200

Araceli Díaz-Perales‡, Víctor Quesada‡, Juan R. Peinado‡, Alejandro P. Ugalde‡, Jesús Álvarez§, María F. Suárez‡, F. Xavier Gomis-Rüth¶, and Carlos López-Otín‡‖

From the Departamentos de ‡Bioquímica y Biología Molecular y §Biología Celular, Facultad de Medicina, Instituto Universitario de Oncología, Universidad de Oviedo, 33006 Oviedo, Spain and the ¶Institut de Biología Molecular de Barcelona, Consejo Superior de Investigaciones Cientiíficas, 08034 Barcelona, Spain

Systematic analysis of degradomes, the complete protease repertoires of organisms, has demonstrated the large and growing complexity of proteolytic systems operating in all cells and tissues. We report here the identification of two new human metalloproteases that have been called archaemetzincin-1 (AMZ1) and archaemetzincin-2 (AMZ2) to empl… teases p… nomes. … with a … includes … residue … cysteine … position … genome … widely … tribute … that has … sequenc… from b… showed … evolutio… transfer… have sh… ern blo… distinct… mainly … domina… are also… human … the pur… substrates and bioactive peptides, demonstrating that they are functional proteases. Finally, these activities were abolished by inhibitors of metalloproteases, providing further evidence that AMZs belong to this catalytic class of proteolytic enzymes.

Proteases mediate many key physiological processes (1). These enzymes play essential roles in a variety of events that determine cell life and death in all living organisms. Thus, proteases participate in the control of cell cycle progression, tissue morphogenesis and remodeling, cell proliferation and migration, ovulation and fertilization, angiogenesis, host de… ecause of … ecessary … ic activi… ibutes to … ding ar… isorders, … man de… the de… steps to … Accord… rotease… vi.es/de… genes in … mplexity … nsion in … ncoding … se. Simi… r model … *melano*… dict that … ranging … r.ac.uk), … systems present in all organisms.

Recently, and as part of our studies aimed at identifying novel human proteases, we have evaluated the possibility that human tissues could produce proteases described previously in evolutionarily distant organisms but whose occurrence in mammals has yet not been reported. This approach led us to identify and characterize human and mouse ovastacin, a novel metalloprotease similar to hatching enzymes from arthropods, birds, amphibians, and fish (20). Following this strategy, we have also tried to explore the putative occurrence in the human genome of genes encoding proteases with sequence similarity to putative metalloproteases annotated in the course of genome sequencing projects of prokaryotic organisms (21). In this work, we report the identification, cloning, and characterization of two novel human metalloproteases called archaemetzincin-1 (AMZ1)[1] and arch-

* This work was supported by grants from Ministerio de Ciencia y Tecnología-Spain, Fundación La Caixa, the European Union (FP5 and FP6-Cancer Degradome), and the Daiichi Fine Chemical Company, Limited. (Toyama, Japan). The Instituto Universitario de Oncología is supported by Obra Social Cajastur-Asturias and Red de Centros de Cáncer-Instituto Carlos III of Spain. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AJ635357, AJ635358, AJ879912, AJ879913, AJ879914, and AJ879915.

‖ To whom correspondence should be addressed. Tel.: 34-985-104201; Fax: 34-985-103564; E-mail: clo@uniovi.es.

[1] The abbreviations used are: AMZ, archaemetzincin; AMC, 7-amino-4-methylcoumarin; contig, group of overlapping clones; GST, glutathione *S*-transferase; Mca, (7-methoxycoumarin-4-yl)-acetic acid.

aemetzincin-2 (AMZ2), which are closely related to proteins whose sequence has been predicted by bioinformatic analysis of archaeal genomes. We perform a detailed phylogenetic analysis of these enzymes to clarify the origin and complex evolutionary history of this new family of metalloproteases. Finally, we examine the tissue distribution of AMZ1 and AMZ2 in human tissues and analyze their enzymatic properties.

EXPERIMENTAL PROCEDURES

Materials—Restriction endonucleases and other reagents used for molecular cloning were from Roche Diagnostics. Double-stranded DNA probes were radiolabeled with [α-^{32}P]dCTP (3000 Ci/mmol) from Amersham Biosciences, using a commercial random priming kit purchased from the same company. Human cDNA libraries and Northern blots containing polyadenylated RNAs from different tissues were from Clontech. Fluorogenic substrates and biologically active peptides (neurogranin, angiotensin II, and angiotensin III) were purchased from Bachem, and protease inhibitors and AMC were from Sigma. Albumin, fibrillar collagens, gelatin, plasminogen, and aprotinin were also from Sigma. Antibodies against GST were developed in our laboratory as described previously (22).

Bioinformatic Analysis and cDNA Cloning—The BLAST program was used to screen public (www.ncbi.nlm.nih.gov) and private (www.celera.com) human genome databases, searching for regions with sequence similarity to prokaryotic metalloproteinase sequences (21). We found two partial sequences located in the human chromosomes 7p22.3 and 17q24.2 exhibiting similarity to putative metalloproteinase sequences identified during the course of large scale genome-sequencing projects involving Archaea (23–26). After the identification of these human sequences, we designed specific oligonucleotides to PCR amplify the cDNAs for these metalloproteases using a human brain cDNA library as a template. All PCR amplifications were performed [illegible] GeneAmp 2400 PCR system from PerkinElmer Life Scien[illegible] cloning the PCR products in pBluescript, their identity [illegible] by nucleotide sequencing

Nucleotide Sequence Analysis—Cloned cDNA [illegible] Oviedo University DNA analysis facility usi[illegible] sion 3.1) chemistry on an ABI PRISM [illegible] (Applied Biosystems). Computer [illegible] quences was performed with [illegible] sity of Wisconsin Geneti[illegible]

Phylogenetic Analysis—[illegible] archaeal, bacterial, and eu[illegible] genomic sequences with th[illegible] quences were aligned autom[illegible] (www.igbmc.u-strasbg.fr/BioInfo/Clus[illegible]neDoc version 2.6. (www.psc.edu/biomed/g[illegible]metalloprotease from the enterobacteria *Yersinia*[illegible] added to this alignment as an out group. The most p[illegible] as tree according to this alignment was calculated with the [illegible] pars program included in the Phylip package version 3.6 (evolution.genetics.washington.edu/phylip/getme.html). The resulting tree was plotted with TreePlot (www.bioinformatics.nl/tools/plottree.html). Additionally, a tree of the selected species was constructed based on a diverse array of phylogenetic resources (www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi).

Northern Blot Analysis—Nylon membranes containing 2 μg of poly(A$^+$) RNA from diverse human tissues were prehybridized at 42 °C for 3 h in 50% formamide, 5× SSPE (1× SSPE is 150 mM NaCl, 10 mM NaH_2PO_4, 1 mM EDTA, pH 7.4) 10× Denhardt's solution, 2% SDS, and 100 μg/ml denatured herring sperm DNA. Membranes were then hybridized with specific radiolabeled probes containing nucleotides from positions 180 to 580 of AMZ1 cDNA and from 340 to 750 of AMZ2 cDNA. Hybridization was performed for 20 h under the same conditions used for prehybridization. Finally, blots were washed once with 2× SSC, 0.05% SDS for 30 min and three times in 0.1× SSC and 0.1% SDS for 30 min at 50 °C and exposed to autoradiography. RNA integrity and loading was assessed by hybridization with an actin probe.

Production and Purification of Recombinant Proteins—cDNAs for the predicted catalytic domains of AMZ1 (positions 1–320) and AMZ2 (positions 1–300) were obtained by PCR amplification using specific oligonucleotide pairs containing defined restriction sites. The AMZ1 catalytic domain oligonucleotides were 5′-GGGGATCCCATGCTGCAGTGTAGACCCGCACAGGA-3′ and 5′-GGGTCGACGATTGAGAGAAGGGGTAGGGTCCCTG-3′, and the AMZ2 catalytic domain oligonucleotides were 5′-GGGGATCCCATGCAAATAATACGGCACTCCG-3′ and 5′-CAGGAATTCAGTAAAAACCTCTTGACGGTCCG-3′ (where the restriction sites are underlined). PCR amplifications were performed with 30 cycles of denaturation (95 °C for 30 s), annealing (60 °C for 30 s), and extension (68 °C for 1 min) using the Expand™ long template, high fidelity PCR system. PCR products were then digested with the corresponding restriction enzymes and cloned in the appropriate sites of the pGEX-5x-2 expression vector (Amersham Biosciences). The resulting constructs were transformed into BL21(DE3)-pLysE-competent *Escherichia coli* cells, and expression was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (final concentration 1 mM), followed by 3 h of incubation at 28 °C. The cells were then harvested by centrifugation, washed with phosphate-buffered saline, and lysed by incubation in phosphate-buffered saline with 100 μg/ml lysozyme, 10 μg/ml DNase, and 0.1% Triton X-100 overnight at 4 °C. The recombinant catalytic domain proteins contained in the corresponding supernatants were purified by affinity chromatography using a glutathione-Sepharose column. The identity of the recombinant proteins was verified by Western blot and trypsin digestion followed by mass spectrometry analysis.

Trypsin Digestion—Gel bands were manually excised and placed into 0.5-ml tubes. Then, gel pieces were washed three times with 180 μl of 25 mM ammonium bicarbonate/acetonitrile (70:30) (v/v), dried at 90 °C for 15 min, and incubated with 12 μg/ml trypsin (Promega) in 25 mM ammonium bicarbonate at 60 °C for 1 h. Likewise, soluble proteins were incubated with trypsin (12 μg/ml) in 25 mM ammonium bicarbonate for 1 h at 60 °C. The resulting peptide mixtures were placed into ice for 2 min, and 2 μl of 10% trifluoroacetic acid were ad[illegible] to each sample. Samples were then desalted by C18 reverse ph[illegible]tography (ZipTip; Millipore). Peptides were eluted with [illegible]4-hydroxycinnamic acid in acetonitrile and 0.1% tri[illegible]0) (v/v). In a typical experiment, 1 μl of this sol[illegible] spectrometry.

Mass S[illegible]sisted laser desorption ionizati[illegible]ss spectrometer equipped [illegible]R; Applied Biosystems). [illegible]ed to produce a mass

[illegible] the purified recombinant [illegible]sing AMC-coupled amino acids [illegible]lu-AMC, His-AMC, Val-AMC, Asn-[illegible]-AMC, Phe-AMC, Ala-AMC, Gln-AMC, [illegible]MC, Pro-AMC, Met-AMC, and Arg-AMC) or [illegible]ides QF35 (Mca-Pro-Leu-Ala-Nva-Dpa-Ala-Arg-[illegible]Mca-Pro-Cha-Gly-Nva-His-Ala-Dpa-NH_2), where Mca [illegible]ycoumarin-4yl)-acetic acid, Nva is norvaline, Dpa is L-[illegible]henyl-diamino propionic acid, and Cha is cyclohexyl alanine. [illegible]ays were carried out at 37 °C at a substrate concentration of 5 μM in a buffer containing 50 mM Tris-HCl, 150 mM NaCl, and 0.05% Brij-35, pH 7.5. The fluorometric measurements were made in an LS55 spectrofluorometer from PerkinElmer Life Sciences (λ_{ex} = 360 nm and λ_{em} = 460 nm for AMC-coupled amino acids and λ_{ex} = 328 nm and λ_{em} = 393 nm for Mca-containing peptides). The fluorescent signal was calibrated using known concentrations of AMC and Mca. For inhibition experiments, the recombinant proteins were preincubated for 30 min at 37 °C with *o*-phenantroline, E-64, 4-(2-aminoethyl)-benzenesulfonyl fluoride, batimastat, tissue inhibitor of metalloproteinase-1, -2, -3, and -4, arphamenine A, and amastatin, and then the hydrolyzing activity against Ala-AMC for AMZ1 or against Arg-AMC for AMZ2 was determined by fluorometric measurements as described above. Kinetic studies were performed using different concentrations of the fluorogenic compounds (0.5–100 μM) in 100 μl of assay buffer containing recombinant enzymes (5 nM), and peptide hydrolysis was measured from the increase in fluorescence at 37 °C over time. Initial velocities were calculated using the analysis package FL WinLab 2.01 (PerkinElmer Life Sciences), and data were fitted to the Michaelis-Menten equation (27) using GraFit version 4.0 (Erithacus). Assays with purified proteins (albumin, fibrillar collagens, gelatin, plasminogen, or aprotinin) and bioactive peptides (neurogranin, angiotensin II, and angiotensin III) were performed by incubation of 2.5 nM recombinant AMZ1 or AMZ2 with each substrate (10 μM). Reactions were carried out at 37 °C in a buffer containing 50 mM Tris-HCl and 150 mM NaCl, pH 7.5, overnight for purified proteins or in the course of 2 h for bioactive peptides. The digestions of purified proteins were analyzed by SDS-PAGE. Peptide digestions were purified using a ZipTip and analyzed by mass spectrometry.

RESULTS

Cloning and Characterization of Two Human cDNAs Encoding Novel Metalloproteases Similar to Archaeal Metzincins—A bioinformatic search of the human genome to look for sequences similar to those of archaeal or bacterial metallopro-

teases led us to identify two DNA contigs located in chromosomes 7p22.3 and 17q24.2 and encoding two uncharacterized proteins with sequence similarity to putative archaeal metalloprotease sequences (21, 23–26). The full-length cDNAs for both human enzymes were PCR-amplified using specific oligonucleotides and a brain cDNA library. These experiments led us to the amplification of 1.5- and 1-kb cDNAs, both containing in-frame initiator and stop codons. After cloning and sequencing of the PCR-amplified products, we confirmed by conceptual translation that the generated sequences encoded two novel proteins of 498 and 360 amino acids, respectively (Fig. 1*A*, and GenBank™ accession numbers AJ635357 and AJ635358). Domain analysis with the InterPro (www.ebi.ac.uk/interpro) and SMART (smart.embl-heidelberg.de) programs confirmed the presence in both human protein sequences of a catalytic domain related to neutral zinc metalloproteases. A search for orthologous sequences using the TBLASTN algorithm showed that both human sequences are closely related to members of a family of predicted metalloproteases originally identified during the analysis of archaeal genomes and tentatively called archaemetzincins (21). Accordingly, we propose to call the newly identified proteins human archaemetzincin-1 and -2. The maximum percentages of identities between the catalytic domains of human and archaeal enzymes were 27% between human AMZ1 and the predicted archaemetzincin from the genome sequence of *Thermococcus kodakaraensis* and 39% between human AMZ2 and the corresponding enzyme from *Pyrococcus abyssi* (Fig. 1*B*). Likewise, the percentage of identity between the catalytic domains of human AMZ1 and [illegible] ~40%. Further bioinformatic analysis of available [illegible] quences revealed that archaemetzincins are [illegible] in Archaea as well as in vertebrates [illegible] ans, and fish (Fig. 1*B*). However, [illegible] were found in a large numb[illegible] exception of *Aquifex aeo*[illegible] wise, archaemetzinci[illegible] non-vertebrate Metazo[illegible] and *Anopheles gambiae*.

Amino acid sequence a[illegible] metzincins with all related seq[illegible]nt species allowed us to identify a high[illegible] catalytic motif, HE*XX*H*XX*GX_3CX_4C*X*MX_{17}C[illegible] the putative metalloprotease zinc-binding site is u[illegible]ined. The conserved methionine would be part of the "Met-turn" described in the metzincin clan of metalloproteases (21), although it should be noted that the sequences deduced for mouse and rat AMZ1 contain a Leu residue at this position (Fig. 1*B*, and GenBank™ accession numbers AJ879912 and AJ879913, respectively). By contrast, mouse and rat AMZ2 contain the archetypal Met at this position (Fig. 1*B*, and GenBank™ accession numbers AJ879914 and AJ879915, respectively). Notably, human AMZ1 and rodent AMZ1 lack a conserved His residue that is present in human and rodent AMZ2, as well as in archaeal and fungal archaemetzincins (Fig. 1*B*). Accordingly, this His residue can be used as a distinctive structural feature between AMZ1 and AMZ2. It is also noteworthy that the core catalytic motif of these enzymes contains four Cys residues that are absolutely conserved in all archaemetzincins but are absent at the equivalent positions of other metalloproteases. Accordingly, we propose that these four Cys residues can be used as a specific signature to distinguish this family of metalloproteases within the metzincin clan.

Evolutionary Analysis of AMZs—An extensive search of the publicly available genome sequences allowed us to identify AMZ genes in multiple eukaryotic and prokaryotic organisms. However, AMZs are absent in several eukaryotic model organisms, such as *Saccharomyces cerevisiae, A. thaliana*, *D. melanogaster*, and *C. elegans*, as well as in most bacterial organisms. All of the predicted prokaryotic AMZs were classified as AMZ2 because of the presence of a third histidine residue in their catalytic site. Additionally, the archaeal organisms belonging to the Thermococcaceae family presented a second AMZ2 with a highly divergent N-terminal extension that we called AMZ2b. After alignment of these sequences, a phylogenetic tree was calculated and rooted with an unrelated metalloprotease (Fig. 2). The obtained tree shows eukaryotic and prokaryotic AMZs separated into two large groups. Only AMZ2b genes and AMZ2 from bacterial *M. xanthus* stand outside these groups. Interestingly, AMZ2 from *A. aeolicus* groups with AMZs from archaeal organisms, suggesting a relatively late lateral gene transfer event from Archaea to bacteria as was previously proposed for other *A. aeolicus* genes (28, 29). Finally, data from this analysis were fitted to a taxonomic tree to construct a model that could explain the evolution of AMZ genes (Fig. 3). According to this model, the primordial AMZ arose in a common ancestor of Archaea and Eukaryota. Some bacterial species acquir[illegible] this gene through lateral gene transfer from archaeal or[illegible] On the other hand, two duplication events would [illegible]esence of AMZ2b in Thermococcaceae and A[illegible] The lack of AMZ genes in several eu[illegible]nisms would be likely explaine[illegible] different times (Fig. 3).

[illegible]*MZ2 in Human Tis-*

[illegible]archaemetzincins in [illegible] poly(A$^+$) RNAs pre[illegible] tissues were hybridized [illegible]MZ1 or AMZ2. As can be seen [illegible]scripts are detected predominantly [illegible]ough there are significant mRNA levels [illegible]y, and testis. On the other hand, AMZ2 [illegible]ainly present in heart and testis, although there [illegible]etectable transcripts in pancreas, kidney, liver, lung, [illegible]nta, brain, and prostate. Notably, both human AMZs display several mRNA transcripts, possibly derived from alternative splicing events. On the other hand, both AMZ mRNAs are present in all fetal tissues analyzed (Fig. 4), being mainly detected in kidney and liver in the case of AMZ1 and in kidney and brain in the case of AMZ2.

Enzymatic Properties of Human AMZ1 and AMZ2 Produced in E. coli—To analyze the enzymatic properties of both human AMZs, we produced in *E. coli* two fusion proteins containing the putative catalytic domains of these enzymes linked to GST at their N termini. The catalytic domains were defined based on the alignments of human AMZs with the related archaeal proteins, which showed the maximum degree of conservation in the N-terminal region of these proteins. Then, these constructs (encoding amino acids 1–320 of AMZ1 and 1–300 of AMZ2) were transformed in *E. coli* BL21, and, after isopropyl-1-thio-β-D-galactopyranoside induction, bands of the expected size (55 kDa) were detected by SDS-PAGE and Western blot analysis of protein extracts using antibodies against GST (Fig. 5*A*). These recombinant GST-proteases were then purified by glutathione-Sepharose chromatography. To assess the identity of the proteins present in these bands, they were digested with trypsin and analyzed by mass spectrometry. The obtained spectra confirmed that the 55-kDa bands corresponded to GST-AMZ1 and GST-AMZ2 fusion proteins.

The recombinant human AMZ1 and AMZ2 proteins were then used in enzymatic assays with the fluorescent substrates commonly employed for assaying other proteases. These assays showed that recombinant AMZ1 exhibits a significant hydrolytic activity against Ala-AMC, whereas recombinant AMZ2 preferentially cleaves Arg-AMC (Fig. 5*B*). By contrast, we did

FIG. 1. **Sequences of human AMZs and comparison to proteases from other organisms.** *A*, sequences of human AMZ1 and AMZ2. The nucleotide and amino acid sequences of AMZ1 and AMZ2 are shown. The characteristic core catalytic motif of archaemetzincins, including the zinc-binding site, is in *black*. *B*, amino acid sequence alignment around the zinc-binding site of the catalytic domains of AMZ enzymes. Residues common to all of them are *shaded*. The alignment was performed using ClustalX (version 1.81). *Hsa*, *Homo sapiens*; *Mmu*, *Mus musculus*; *Rno*, *Rattus novergicus*; *Gga*, *Gallus gallus*; *Xtr*, *Xenopus tropicalis*; *Tni*, *Tetraodon nigroviridis*; *Pfu*, *Pyrococcus furiosus*; *Mja*, *Methanoccocus jannaschii*; *Neq, Nanoarchaeum equitans*; *Sso*, *Sulfolobus solfataricus*; *Aae*, *A. aeolicus*.

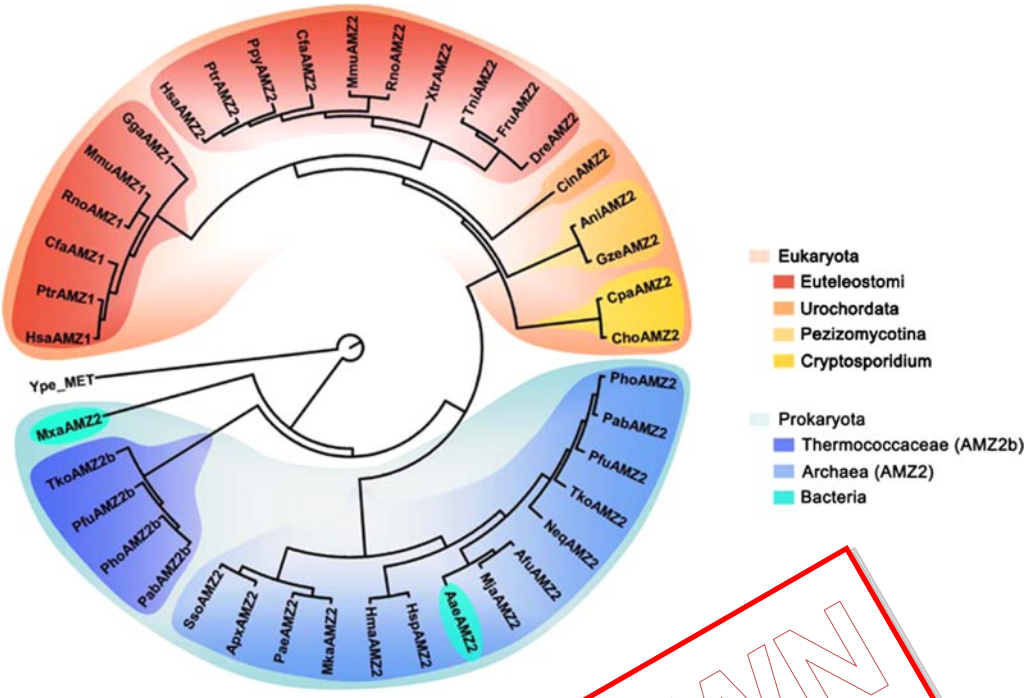

FIG. 2. **Phylogenetic relationships of archaemetzincins.** The tree was … Phylip program package (version 3.6). *Hsa*, *Homo sapiens*; *Ptr*, *Pan troglodytes; Ppy*, *Pongo pygm… Rno*, *Rattus novergicus*; *Gga*, *Gallus gallus*; *Xtr*, *Xenopus tropicalis*; *Dre*, *Danio rerio*; *Fru*, *…iona intestinalis*; *Ani*, *Aspergillus nidulans*; *Gze*, *Giberella zeae*; *Cho*, *Cryptosporidium… abyssi*; *Tko*, *T. kodakaraensis*; *Afu*, *Archaeoglobus fulgidus*; *Apx*, *Aeropyrum pern…*; *Hsp*, *Halobacterium* sp.; *Hma*, *Haloarcula marismortui*; *Neq*, *Nanoarchaeum … fataricus*; *Mka*, *Methanopyrus kandleri; Pae*, *Pyrobaculum aerophilum*; Aae, *A. ae…*

not detect any significant activity of … its for substrate specificity and sensitivity against QF35 or QF41, two pep… obtained by using the catalytic domains of both metalloendopeptidases. Lik… ta not shown).

dence of endoproteolyti… rther characterize the enzymatic activity of the identified AMZ metalloproteases, several commercially available bioactive peptides were incubated in the presence of purified catalytic domains of AMZ1 or AMZ2, and the resulting samples were analyzed by mass spectrometry. As shown in Fig. 6, these experiments demonstrated that human AMZ1 exhibited aminopeptidase activity against neurogranin, whereas human AMZ2 was active against angiotensin III. Thus, as can be seen in Fig. 6*A*, neurogranin is detected as a 1800.1-Da peak consistent with its amino acid sequence (AAKIQASFRGHMARKK), whereas incubation of neurogranin with AMZ1 produced a single additional peak with a mass of 1657.9 Da, corresponding to the processed peptide KIQASFRGHMARKK (Fig. 6*B*). Notably, AMZ2 did not process neurogranin under the same experimental conditions (data not shown). Similarly, neither AMZ1 nor AMZ2 hydrolyzed angiotensin II (data not shown). However, AMZ2 cleaved the N-terminal Arg residue of angiotensin III (RVYIHPF), albeit with low efficiency, to produce angiotensin IV (VYIHPF) (Fig. 6, *C* and *D*).

against a number of … fibrillar collagens, gelat… shown). It is also notewo… ited different optimal pH v… against AMC derivatives. W… of activity at pH 8.0, the opti… (Fig. 5*C*). We then tested the … hibitors to block the enzymatic a… As can be seen in Fig. 5*D*, the activity of both peptidases was inhibited by the general metalloprotease inhibitors *o*-phenantroline and batimastat, but not by 4-(2-aminoethyl)-benzenesulfonyl fluoride, E-64, and tissue inhibitors of metalloproteinases (TIMPS), which are inhibitors of serine, cysteine, and matrix metalloproteases, respectively. Interestingly, AMZ1 and AMZ2 activities were significantly inhibited by amastatin, which is an inhibitor of aminopeptidases (Fig. 5*D*). We next performed a kinetic analysis of the proteolytic reaction catalyzed by the catalytic domains of AMZ1 and AMZ2 with their preferred substrates (Ala-AMC and Arg-AMC, respectively). The fitting of the resulting data to the Michaelis-Menten equation yielded k_{cat}/K_m values of 46 M^{-1} s^{-1} and 22 M^{-1} s^{-1} for catalytic domain proteins of AMZ1 and AMZ2, respectively, which are similar to the value reported for recombinant aminopeptidase O produced in the same expression system (30). We have also tried to perform similar experiments with the full-length proteins produced in bacterial systems. To date, these experiments have been hampered by the low amounts of full-length AMZs, which can be recovered in active form by using different expression systems. Nevertheless, preliminary experiments performed with full-length AMZ1 and AMZ2 produced as His tail fusion proteins in *E. coli* BL21 pLysS have con-

DISCUSSION

In this work we describe two new human proteases that have been tentatively called archaemetzincin-1 and -2. According to a series of structural and enzymatic features, these proteins belong to a new family of metalloproteases characterized by a conserved motif ($HEXXHXXGX_3CX_4CXMX_{17}CXXC$) that contains an archetypal zinc-binding site and four Cys residues that contribute to defining the specific signature of this novel metalloprotease family. Furthermore, enzymatic assays performed with human recombinant AMZs have provided the first evidence that these proteins are catalytically active metalloproteases that exhibit substrate specificity and sensitivity to in-

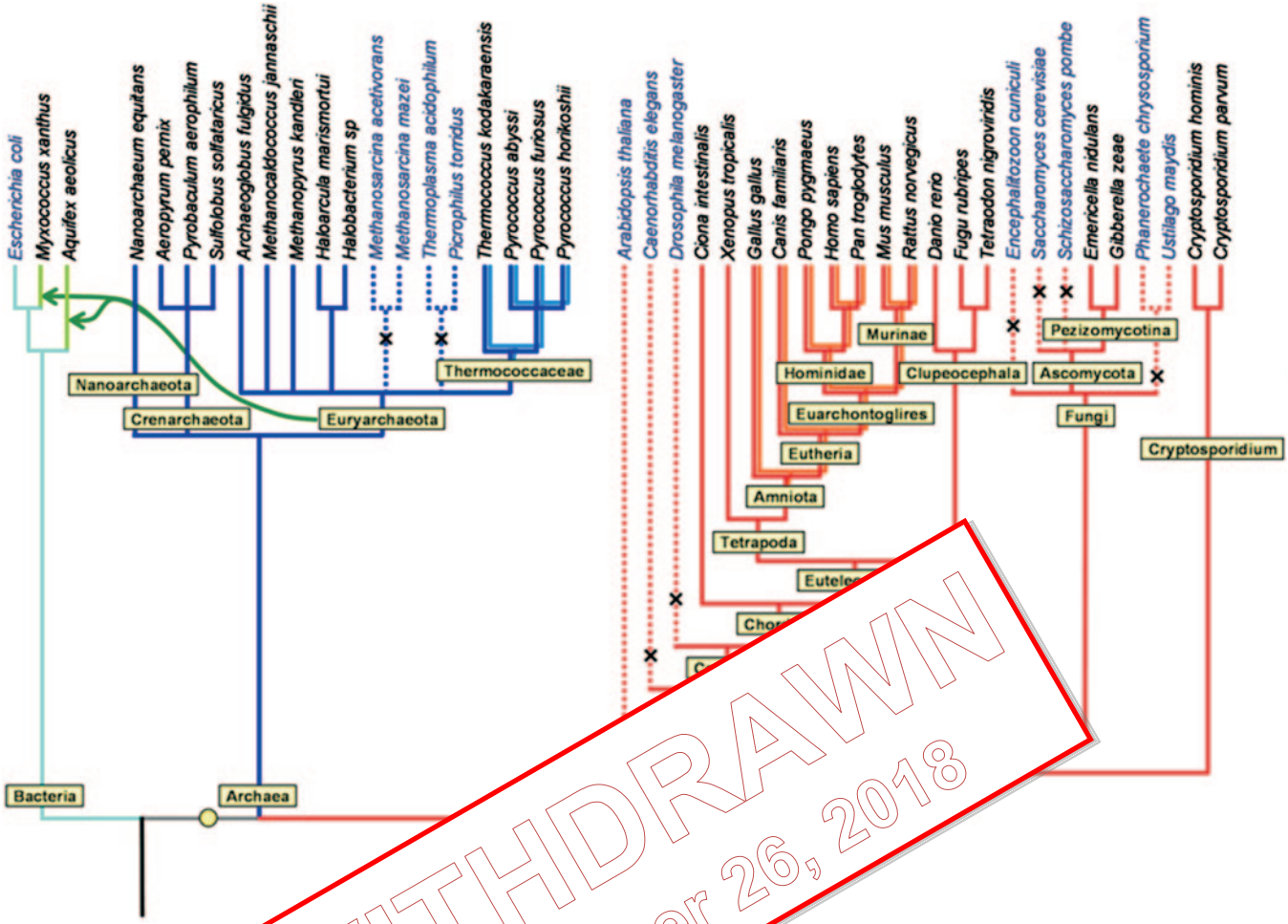

FIG. 3. **Model of the ev** a diverse array of phylogenetic resources was constructed using the NCBI taxonomy bro . Putative sites of origin (*yellow circle*), duplication (*double line*), lateral transfer (*green ar* MZ genes are shown.

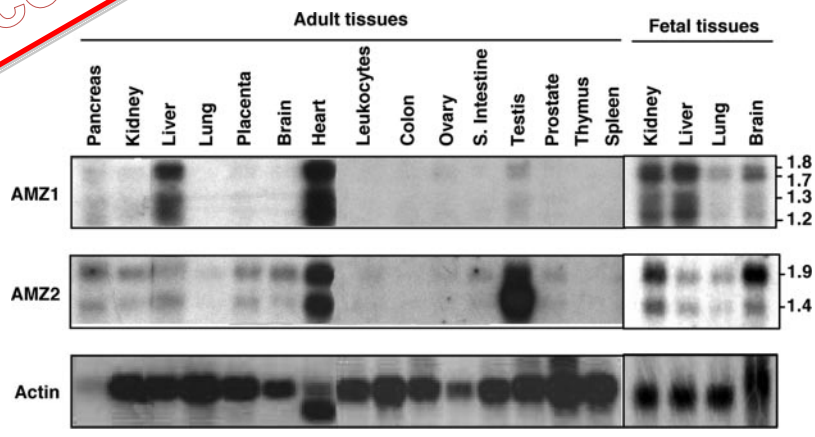

FIG. 4. **Analysis of AMZ1 and A expression in adult and fetal human tissues.** Filters containing ~2 μg of polyadenylated RNAs were hybridized with human AMZ1 or AMZ2 specific probes. RNA sizes are indicated in kilobases. *S. Intestine*, small intestine.

hibitors, which appears to indicate that both proteases may act predominantly as aminopeptidases.

An additional distinctive feature of this family of metalloproteases is the complex series of evolutionary events that have contributed to its creation and diversification in different organisms. In fact, our bioinformatic analysis revealed that these enzymes are widely distributed in vertebrate and archaeal organisms but are absent in the genomes of a number of model organisms such as *E. coli*, *S. cerevisiae*, *A. thaliana*, *D. melanogaster*, and *C. elegans*. The occurrence of genes shared by prokaryotes and vertebrates but absent in other eukaryotes has been widely considered as an indication of lateral gene transfer events from prokaryotes to vertebrates (31). Accordingly, AMZs could represent novel and interesting examples of these rare evolutionary events. However, the recent accumulation of data questioning many cases of lateral gene transfer to the vertebrate lineage (32–34) prompted us to perform an exhaustive bioinformatic search for AMZ genes in all available genome sequences. This analysis led us to identify additional AMZ-related sequences in other non-vertebrate eukaryotes and in two bacterial species, as well as to uncover a series of complex evolutionary events underlying the formation of this metalloprotease family (Fig. 3). According to this phylogenetic analysis, the evolutionary history of AMZs is best described by

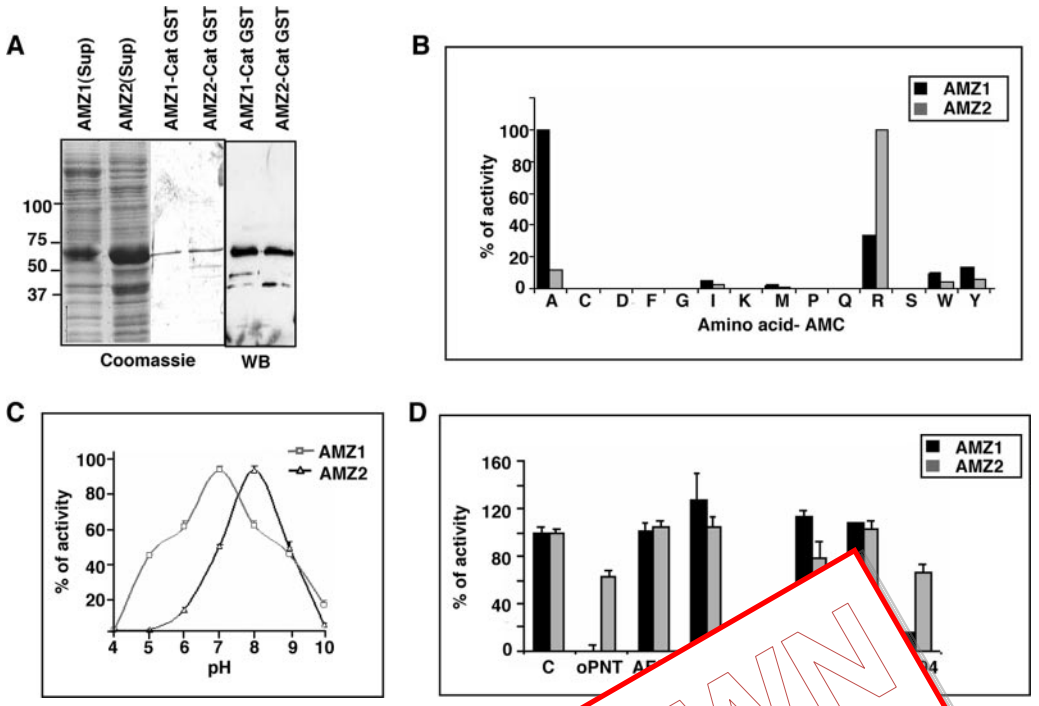

FIG. 5. **Production and enzymatic analysis of recombinant human AM** SDS-PAGE and Western blot (*WB*) showing the results of the expression and purification of the AMZ1 and AMZ2. 5-μl aliquots of soluble fraction (*AMZ1(Sup)* and *AMZ2(Sup)*) and puri he two *far right lanes* correspond to the Western blot analysis of purified AMZ1 and A re shown on the left. *B*, analysis of the hydrolyzing activity of AMZ1 and AMZ2 us nalysis using Ala-AMC and Arg-AMC as substrates for AMZ1 and AMZ2, respectivel GST-AMZ1 or GST-AMZ2 were incubated alone (*C*) or in the presence of 10 mM *o*-p nyl fluoride (*AEB*), 20 μM E-64, 20 μM amastatin (*Amas*), 20 μM arphamenine A (*A* d 5 μM batimastat (*BB94*); the remaining activity was measured using Ala-AMC or *A*

a scenario in which the pri e two human metalloproteases show different subrent AMZ2 enzymes, preferences. Whereas AMZ1 preferentially targets suband eukaryotic organi ates that contain Ala at their N termini, AMZ2 mainly all the analyzed bacter hydrolyzes substrates with Arg at that position. Consistent with the proposed origin the with this finding, AMZ1 could hydrolyze the N-terminal Ala of primordial bacterial orga AMZ in neurogranin, whereas AMZ2 processed the N-terminal Arg of *A. aeolicus*, a hyperthermop pies an eco- angiotensin III. Furthermore, AMZ1 and AMZ2 failed to hylogical niche dominated by be explained by drolyze synthetic peptides such as QF35 and QF41, which are lateral gene transfer from som archaeal organisms. used for the analysis of metalloendopeptidases. The activities Furthermore, the clear phylog tic relationship between of both enzymes were abolished by general metalloprotease *A. aeolicus* AMZ2 and archaeal AMZs provides additional sup- inhibitors such as *o*-phenantroline and batimastat and by the port to the occurrence of the proposed lateral transmission. specific aminopeptidase inhibitor amastatin, providing further Similarly, the finding of AMZ-related sequences in *M. xanthus* evidence that AMZs could be classified as metalloaminopeptishould likely be the result of lateral transfer from Archaea, but dases. Also in this regard, it is remarkable that the activity of in this case this event was followed by a rapid accumulation of the recombinant catalytic domains of AMZs is relatively low, mutations, which would explain its location as an out group in which is somewhat reminiscent of the limited proteolytic activthe phylogenetic tree. The evolutionary history of AMZs in ity observed with the catalytic domains of ADAMTSs (a disineukaryotic organisms has also involved a series of diverse tegrin and metalloprotease with thrombospondin motifs) but is events since their separation from their common ancestor with in contrast to the case of matrix metalloproteinase catalytic Archaea. First, the absence of AMZ genes in plants, nematodes, domains, which readily hydrolyze linear peptides and gelatin or insects is remarkable, suggesting the occurrence of multiple (35–37). On this basis, it is tempting to speculate that the gene loss events in these organisms. Consistent with this pro- catalytic domains of some metalloproteinases including AMZs posal, codon usage or nucleotide composition analysis of AMZ and ADAMTSs may have very strict requirements in terms of genes failed to provide any evidence of lateral transmission the presence of additional motifs or domains to exhibit their from Archaea to vertebrates. Finally, our phylogenetic analysis full potential as proteolytic enzymes.

also revealed that eukaryotic AMZ1 diverged from AMZ2 re- In this work, we have also analyzed the distribution of AMZ1 cently, probably by gene duplication, again illustrating the and AMZ2 in human tissues. These studies allowed us to detect genomic plasticity of this family of metalloproteases. the predominant expression of AMZ1 in liver and heart,

To further explore the functional relevance of AMZ1 and whereas AMZ2 mRNA was mainly detected in heart and testis. AMZ2, we performed an enzymatic analysis of both recombi- This pattern of expression suggests that both AMZs are implinant enzymes produced in *E. coli*. This analysis revealed that cated in the development or physiology of these organs. Nevthe recombinant proteins are catalytically active and that their ertheless, further studies will be required to validate at the activities seem to correspond to those of aminopeptidases. protein level the observed distribution of AMZ RNAs in human

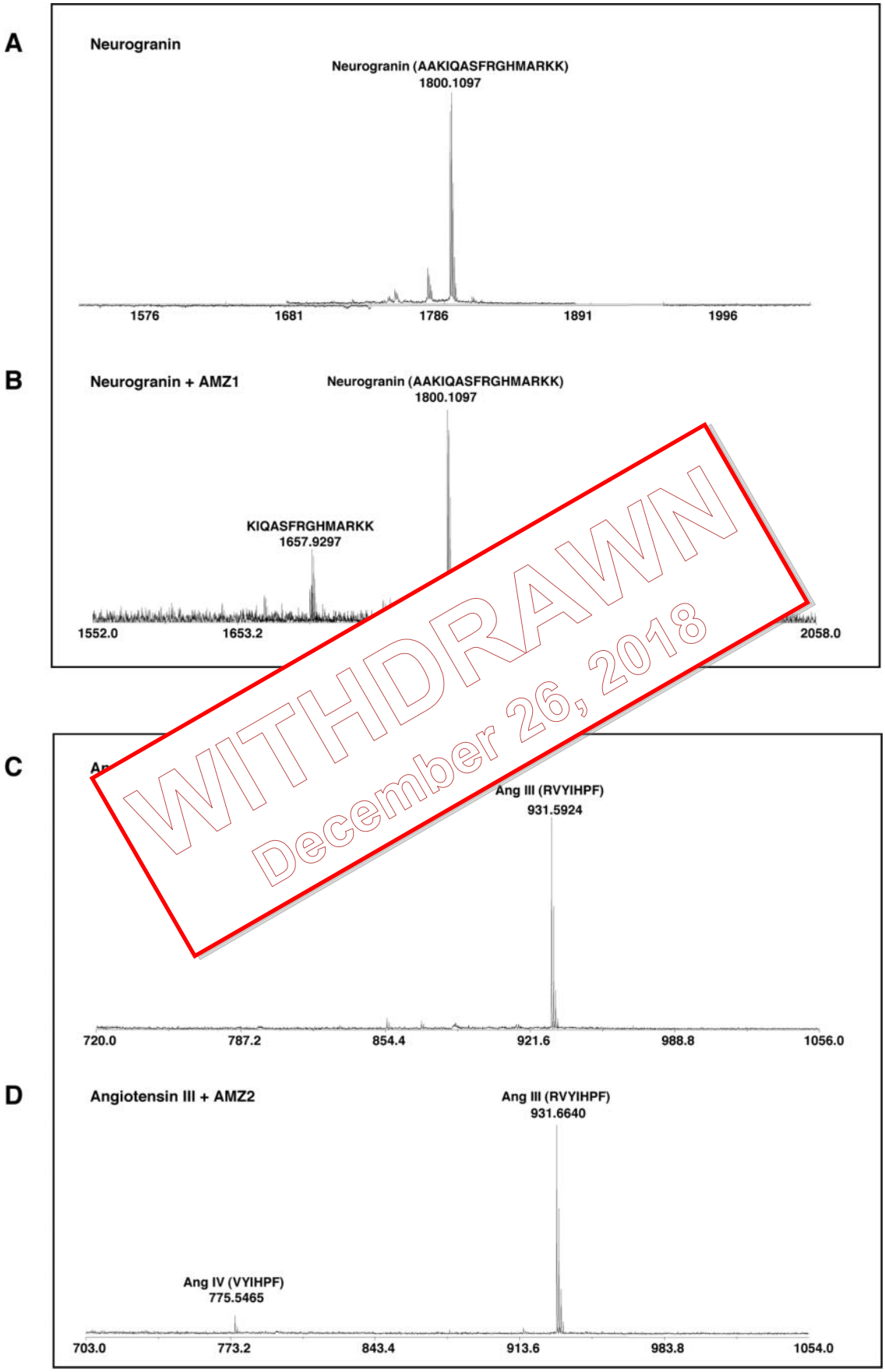

FIG. 6. **Mass spectrometry analysis of the biological peptide proteolysis catalyzed by AMZs.** Human neurogranin or angiotensin III were incubated alone (*A* and *C*, respectively) or in the presence of recombinant GST-AMZ1 (*B*) or GST-AMZ2 (*D*) for 2 h. The resulting peptide mixture was analyzed by mass spectrometry. The peaks corresponding to neurogranin (1800.1 Da), processed neurogranin (1657.9 Da), angiotensin III (931.6 Da), and angiotensin IV (775.4 Da) are shown.

tissues. Additional clues about the physiological and pathological roles of these enzymes may be derived from their chromosome locations at 7p22 and 17q24, respectively. Alterations in these regions have been frequently associated with cancer and other diseases such as hypertension or multiple sclerosis (38–41). Further studies will be required to ascertain whether AMZs could be a direct target of any of these genetic abnormalities resulting in cancer or other pathological conditions. Likewise, further experimental work, including the three-dimensional structural analysis of these enzymes and the generation of mutant organisms deficient in these proteases, will be necessary to clarify their functional roles and to define their precise relevance in the context of the growing complexity of proteolytic systems operating in all living organisms.

Acknowledgments—We thank G. Velasco, J. P. Freije, A. M. Pendás, S. Cal, L. M. Sánchez, and X. S. Puente for helpful comments, and S. Alvarez for excellent technical assistance.

REFERENCES

1. Lopez-Otin, C., and Overall, C. M. (2002) *Nat. Rev. Mol. Cell Biol.* **3,** 509–519
2. Glickman, M. H., and Ciechanover, A. (2002) *Physiol. Rev.* **82,** 373–428
3. Inada, M., Wang, Y., Byrne, M. H., Rahman, M. U., Miyaura, C., Lopez-Otin, C., and Krane, S. M. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17192–17197
4. Blobel, C. P. (2005) *Nat. Rev. Mol. Cell Biol.* **6,** 32–43
5. Curry, T. E., Jr., and Osteen, K. G. (2003) *Endocr. Rev.* **24,** 428–465
6. Noel, A., Maillard, C., Rocks, N., Jost, M., Chabottaux, V., Sounni, N. E., Maquoi, E., Cataldo, D., and Foidart, J. M. (2004) *J. Clin. Pathol.* **57,** 577–584
7. Di Cera, E. (2003) *Chest* **124,** 11S–17S
8. Riedl, S. J., and Shi, Y. (2004) *Nat. Rev. Mol. Cell Biol.* **5,** 897–907
9. Marino, G., Uria, J. A., Puente, X. S., Quesada, V., Bordallo, J., and Lopez-Otin, C. (2003) *J. Biol. Chem.* **278,** 3671–3678
10. Baker, A. H., Edwards, D. R., and Murphy, G. (2002) *J. Cell* 3719–3727
11. Apte, S. S. (2004) *Int. J. Biochem. Cell Biol.* **36,** 981–985
12. Liu, J., Sukhova, G. K., Sun, J. S., Xu, W. H., Libby, P. *Arterioscler. Thromb. Vasc. Biol.* **24,** 1359–1366
13. Marjaux, E., Hartmann, D., and De Strooper
14. Overall, C. M., and Lopez-Otin, C. (2002)
15. Borgono, C. A., and Diamandis, E. P.
16. Puente, X. S., Sanchez, L. M., Lopez-Otin, C. (2005) *Biochi*
17. Puente, X. S., Sanchez, *Rev. Genet.* **4,** 544–5
18. Puente, X. S., and Lopez-
19. Rawlings, N. D., Tolle, D. D160–D164
20. Quesada, V., Sanchez, L. M., *Chem.* **279,** 26627–26634
21. Gomis-Ruth, F. X. (2003) *Mol. B*
22. Llamazares, M., Cal, S., Quesada, **278,** 13382–13389
23. Bult, C. J., White, O., Olsen, G. J., Zh Blake, J. A., FitzGerald, L. M., Cla A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M., and Venter, J. C. (1996) *Science* **273,** 1058–1073
24. Slesarev, A. I., Mezhevaya, K. V., Makarova, K. S., Polushin, N. N., Shcherbinina, O. V., Shakhova, V. V., Belova, G. I., Aravind, L., Natale, D. A., Rogozin, I. B., Tatusov, T. L., Wolf, Y. I., Stetter, K. O., Malykh, A. G., Koonin, E. V., and Kozyavkin, S. A. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 4644–4649
25. Maeder, D. L., Weiss, R. B., Dunn, D. M., Cherry, J. L., Gonzalez, J. M., DiRuggiero, J., and Robb, F. T. (1999) *Genetics* **152,** 1299–1305
26. Fukui, T., Atomi, H., Kanai, T., Matsumi, R., Fujiwara, S., and Imanaka, T. (2005) *Genome Res.* **15,** 352–363
27. Northrop, D. B. (1999) *Adv. Enzymol Relat. Areas Mol. Biol.* **73,** 25–55
28. Forterre, P., Bouthier De La Tour, C., Philippe, H., and Duguet, M. (2000) *Trends Genet.* **16,** 152–154
29. Nesbo, C. L., L'Haridon, S., Stetter, K. O., and Doolittle, W. F. (2001) *Mol. Biol. Evol.* **18,** 362–375
30. Diaz-Perales, A., Quesada, V., Sanchez, L. M., Ugalde, A. P., Suarez, M. F., Fueyo, A., and Lopez-Otin, C. (2005) *J. Biol. Chem.* **280,** 14310–14317
31. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., W stock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Tov Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., W J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Br Robert, C., Wincker, P., Smith, D. R., Doucette-Stam einstock, K., Lee, H. M., Dubois, J., Rosenthal Taudien, S., Rump, A., Yang, H., Yu, J L., Rowen, L., Madan, A., Qin, S., D P., Proctor, M. J., Myers, R. M., D. R., Olson, M. V., Kaul, R., ans, G. A., Athanasiou, M., ser, J., Lehrach, H., Rein- Dedhia, N., Blocker, H., Aravind, L., Bailey, J. A., K, P., Brown, D. G., Burge, C. B., Clamp, M., Copley, R. R., Doerks, T., T. S., Galagan, J., Gilbert, J. G., Harmon, D., Hermjakob, H., Hokamp, K., Jang, W., A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, nn, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., kowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., eeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., Szustakowki, J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., and Chen, Y. J. (2001) *Nature* **409,** 860–921
32. Stanhope, M. J., Lupas, A., Italia, M. J., Koretke, K. K., Volker, C., and Brown, J. R. (2001) *Nature* **411,** 940–944
33. Salzberg, S. L., White, O., Peterson, J., and Eisen, J. A. (2001) *Science* **292,** 1903–1906
34. Roelofs, J., and Van Haastert, P. J. (2001) *Nature* **411,** 1013–1014
35. Kashiwagi, M., Enghild, J. J., Gendron, C., Hughes, C., Caterson, B., Itoh, Y., and Nagase, H. (2004) *J. Biol. Chem.* **279,** 10109–10119
36. Murphy, G., Allan, J. A., Willenbrock, F., Cockett, M. I., O'Connell, J. P., and Docherty, A. J. (1992) *J. Biol. Chem.* **267,** 9612–9618
37. Clark, I. M., and Cawston, T. E. (1989) *Biochem. J.* **263,** 201–206
38. Lafferty, A. R., Torpy, D. J., Stowasser, M., Taymans, S. E., Lin, J. P., Huggard, P., Gordon, R. D., and Stratakis, C. A. (2000) *J. Med. Genet.* **37,** 831–835
39. Kalikin, L. M., George, R. A., Keller, M. P., Bort, S., Bowler, N. S., Law, D. J., Chance, P. F., and Petty, E. M. (1999) *Genomics* **57,** 36–42
40. Chen, D. C., Saarela, J., Clark, R. A., Miettinen, T., Chi, A., Eichler, E. E., Peltonen, L., and Palotie, A. (2004) *Genome Res.* **14,** 1483–1492
41. Meuleman, J., Timmerman, V., Van Broeckhoven, C., and De Jonghe, P. (2001) *Neurogenetics* **3,** 115–118