The Role of Basic Amino Acid Surface Clusters on the Collagenase Activity of Cathepsin K

ARTICLE *in* BIOCHEMISTRY · OCTOBER 2013

Impact Factor: 3.02

4 AUTHORS, INCLUDING:

Ferez Nallaseth
Life Sciences Institute of New Jersey
151 PUBLICATIONS **710** CITATIONS
SEE PROFILE

Fabien Lecaille
University of Tours
56 PUBLICATIONS **1,455** CITATIONS
SEE PROFILE

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

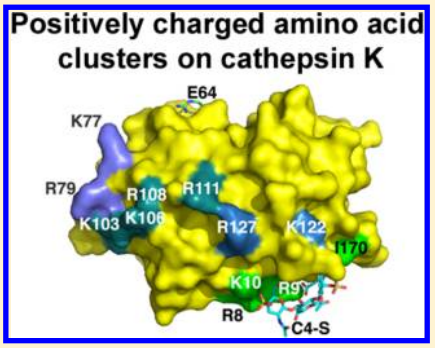
The Role of Basic Amino Acid Surface Clusters on the Collagenase Activity of Cathepsin K

Ferez S. Nallaseth,[†,‡] Fabien Lecaille,[†,§] Zhenqiang Li,[†,∥] and Dieter Brömme*[,†,⊥]

[†]Department of Human Genetics, Mount Sinai School of Medicine, New York, New York 10029, United States

[⊥]Department of Oral Biological and Medical Sciences, The University of British Columbia, Vancouver, British Columbia V6T1Z3, Canada

ABSTRACT: Cathepsin K is a highly potent collagenase in osteoclasts and is responsible for bone degradation. We have previously demonstrated that its unique collagenolytic activity is modulated by glycosaminoglycans that form high molecular weight complexes with the protease. However, mutational analysis of a specific glycosaminoglycan−cathepsin K binding site only led to a 60% reduction of the collagenolytic activity suggesting additional glycosaminoglycan binding sites or other determinants controlling this activity. We identified eight cathepsin K specific arginine/lysine residues that form three positively charged clusters at the bottom part of the protease opposing the active site. These residues are highly conserved among mammalian, avian, and reptilian cathepsin K orthologues and to a lesser degree in amphibian and fish specimens. Mutational analysis of these residues revealed an approximately 50% reduction of the collagenolytic activity when the basic amino acids in cluster 2 (K103, K106, R108, R111) were mutated into alanine residues and resulted in a 100% loss of this activity when the mutations were expanded into cluster 3 (K122, R127). Cluster 1 mutations (K77, R79) had no effect. A partial rescue effect was observed when the hexamutant variant was combined with three mutations in the previously identified glycosaminoglycan binding site (N190, K101, L195K) indicating the relevance of at least two independent interaction sites. Amino acid substitutions in all sites had no effect on the catalytic efficacy of the protease variants as reflected in their unaltered peptidolytic and gelatinolytic activities and their overall protein stabilities. This study suggests that the basic amino acid clusters in cathepsin K are involved in alternative glycoasaminoglycan binding sites, play other roles in the formation of collagenolytically active protease complexes, or contribute in a yet unknown manner to the specific binding to collagen.



Collagen degradation is crucial for normal bone remodeling, skin development, wound healing, and organ development. Triple helical collagens constitute the major organic components in bone (90% type collagen), cartilage (50% type II collagen), and skin (80% type I and III collagens). Various pathologies including arthritis, osteoporosis, skin disorders, and some vascular disorders are thought to result from unregulated and excessive collagen degradation. For example, menopause and aging related collagen loss not only affects bones but also decreases the skin collagen content by 50−70%.[1−3] Besides the well-known collagenases of the matrix metalloprotease family, the papain-like cysteine protease cathepsin K (CatK) has been recognized as a major collagenase in bone, cartilage, and other organs as well.[4] CatK is the predominant peptidase in bone degrading osteoclasts[5,6] and has the unique ability to cleave triple helical collagen at multiple sites.[7−9] It has been shown that the collagenase activity of this enzyme depends on its oligomerization with glycosaminoglycans (GAG).[10,11] In the absence of GAGs, CatK only exerts non-collagenase activities such as the hydrolysis of gelatin or synthetic peptide substrates. The formation of complexes with GAGs also distinguishes CatK from other papain-like proteases.[11] With the exception of the non-collagenase cathepsin V, none of the other cathepsins exhibits the formation of high-molecular-weight complexes with GAGs.[11,12]

We have previously solved the X-ray structure of a CatK/GAG complex and performed a mutation analysis of the identified GAG binding sites.[13,14] The structure revealed a "beads-on-a-strand"-like organization of CatK molecules on a single strand of a chondroitin 4-sulfate molecule with six sugar residues interacting with the surface of CatK on its L-domain and remote from its active site. The replacement of these interaction sites in CatK with appropriate amino acid residues of collagenolytically inactive cathepsin L revealed a 60% reduction in the collagenase activity without affecting its gelatinase and peptidase activities.[14] On one side, these experiments underlined the relevance of the GAG-binding to CatK for its collagenase activity but also indicated that additional residues are probably contributing to its collagenase activity. A unique feature of the CatK structure is a set of positively charged lysine and arginine residues at the bottom of the molecule (on the opposing site of the active site cleft). We identified eight lysine and arginine residues that are unique to CatK compared with other cathepsin L-like proteases (cathepsins L, S, and V), which lack a collagenase activity,

Received: August 2, 2013
Revised: September 28, 2013
Published: October 2, 2013

and that are highly conserved among CatK orthologues in the animal kingdom.

In order to identify additional or alternative putative interaction sites involved in complex formation, we devised a set of single and multiple mutations at these sites. Our results indicate that besides the GAG binding sites identified in the crystal structure, additional CatK-specific lysine and arginine residues specifically contribute to the collagenase activity. This may include alternative GAG binding sites, specific CatK/collagen interaction sites, or protein−protein interface sites within the complexes.

EXPERIMENTAL PROCEDURES

Materials. Z-LR-MCA and Z-GPR-MCA, as well as chondroitin 4-sulfate (C4-S), dithiothreitol (DTT), and EDTA, were purchased from Sigma-Aldrich (St. Louis, MO).

Identification of Unique Basic Surface Residues on CatK. The Merops database (http://merops.sanger.ac.uk) was used to identify CatK specific lysine and arginine residues. First, all 11 human cathepsin sequences were searched for CatK specific lysine and arginine residues, and then the identified residues were evaluated for their conservation among vertebrate species.

Construction of CatK Mutants and Their Expression in *Pichia pastoris*. CatK conserved positively charged lysine and arginine residues were replaced with alanine residues. Cloned human pro-CatK cDNA[15] was used as a template to generate CatK mutants by site-directed mutagenesis. Multiple coding sites were sequentially mutated, with appropriate mutant cDNAs used as PCR templates. Each single nucleotide substitution was introduced into the cDNA using the PCR ligation method with *Pfu* polymerase (New England Biolabs Inc., Beverly, MA). The CatK mutant nomenclature is an extension of the nomenclature of previously published mutants[14] and is summarized below. Composite mutants of the lysine/arginine to alanine mutants and the previously identified C4-S binding site mutants were generated. The C4-S binding site mutants contained analogous cathepsin L residues at the mutation sites. The mutants are as follows: (M6), K77A, R79A; (M7), K103A, K106A, R108A, R111A; (M7.1), K103A; (M7.2), K106A; (M7.3), R108A; (M7.4), R111A; (M8), K77A, R79A, K103A, K106A, R108A, R111A; (M9), K77A, R79A, K103A, K106A, R108A, R111A, K122A, K127A; (M10), K122A, R127A; (M11), K77A, R79A, K103A, K106A, R108A, R111A, K122A, R127A, K9E; (M12), K77A, R79A, K103A, K106A, R108A, R111A, K122A, R127A, K9E, N190M, K191G, L195K; (M13), K77A, R79A, K103A, K106A, R108A, R111A, K122A, R127A, K9E, I171E, Q172S.

Oligonucleotides used for mutagenesis were as follows (sites of mutations are underlined): variant M6, 5′-CAA TAT GTG CAG GCG AAC GCG GGT ATT GAC TCT-3′ (mutant primer 1) and 5′-AGA GTC AAT ACC CGC GTT CGC CTG CAC ATA TTG-3′ (mutant primer 2; template, CatK wt plasmid); variant M7, 5′-TAC AAC CCA ACA GGC GCG GCA GCT GCA TGC GCA GGG TAC GCA GAG ATC CCC GAG GGG-3′ (mutant primer 3) and 5′-CCC CTC GGG GAT CTC TGC GTA CCC TGC GCA TGC AGC TGC CGC GCC TGT TGG GTT GTA-3′ (mutant primer 4; template, CatK wt plasmid); variant M8, template M7 reamplified with mutant primers 1 and 2; variant M9, template M8 reamplified with mutant primer 5, 5′-GGG AAT GAG AAA GCC CTG GCG AGG GCA GTG GCC GCA GTG GGA CCT GTC-3′, and mutant primer 6, 5′-GAC AGG TCC CAC TGC GGC CAC TGC CCT CGC CAG GGC TTT CTC ATT CCC-3′; variant M10, template CatK wt plasmid with mutant primers 5 and 6; variant M11, M9 template reamplified with 5′-GAC TCT GTC GAC TAT CGA GAG AAA GGA TAT GTT ACT CCT-3′ (mutant primer 7) and 5′-TAT CCA GTG CTT GTT TCC CTT CTG GAT TCC ATA TCC CAC TGC CAA AAC CGC ATG GTT-3′ (mutant primer 8); variant 11, M11 template reamplified with 5′-GGA GAA AAC TGG GGA ATG GGA AGG ATA TAT ATC AAA ATG GCT CGA AAT AAG AAC AAC GC-3′ (mutant primer 9) and 5′-GC GTT GTT CTT ATT TCG AGC CAT TTT GAT ATA TCC TCC CAT TCC CCA GTT TTC TCC-3′ (mutant primer 10); variant M13, K77A, R79A, K103A, K106A, R108A, R111A, K122A, R127A, K9E, I171E, Q172S; variant 13, M11 template reamplified with 5′-AAC CAT GCG GTT TTG GCA GTG GGA TAT GGA GAA TCG AAG GGA AAC AAG CAC TGG ATA-3′ (mutant primer 11) and 5′-TAT CCA GTG CTT GTT TCC CTT CTA TCC ATA TCC CAC TGC CAA AAC CGC ATG GTT-3′ (mutant primer 12).

For the individual variants K103, K106, R108, and R111, the following mutant primers were used: variant M7.1, 5′-TAC AAC CCA ACA GGC GCG GCA GCT AAA TGC AGA GGG TAC AGA GAG ATC CCC GAG GGG-3′ (mutant primer 13) and 5′-CCC CTC GGG GAT CTC TCT GTA CCC TCT GCA TTT AGC TGC CGC GCC TGT TGG GTT GTA-3′ (mutant primer 14; template CatK wt plasmid); variant M7.2, 5′-TAC AAC CCA ACA GGC AAG GCA GCT GCA TGC AGA GGG TAC AGA GAG ATC CCC GAG GGG-3′ (mutant primer 15) and 5′-CCC CTC GGG GAT CTC TCT GTA CCC TCT GCA TGC AGC TGC CGC GCC TGT TGG GTT GTA-3′ (mutant primer 16; template CatK wt plasmid); variant M7.3, 5′-TAC AAC CCA ACA GGC AAG GCA GCT AAA TGC GCA GGG TAC AGA GAG ATC CCC GAG GGG-3′ (mutant primer 17) and 5′-CCC CTC GGG GAT CTC TCT GTA CCC TGC GCA TTT AGC TGC CTT GCC TGT TGG GTT GTA-3′ (mutant primer 18, template CatK wt plasmid); variant M7.4, 5′-TAC AAC CCA ACA GGC AAG GCA GCT AAA TGC AGA GGG TAC GCA GAG ATC CCC GAG GGG-3′ (mutant primer 19) and CCC CTC GGG GAT CTC TGC GTA CCC TCT GCA TTT AGC TGC CTT GCC TGT TGG GTT GTA-3′ (mutant primer 20; template CatK wt plasmid). All mutant primers were synthesized in the forward and reverse orientations. The flanking anchoring vector primers were 5′-CGG TCG TAA CGA CGA TTT CTT CTT CCC CAT A-3′ (forward) and 5′-GGC AAA TGG CAT TCT GAC ATC CTC TTG ATT-3′ (reverse). Each cDNA amplicon, encoding a mutation, was generated with either cloned wild-type CatK or an intermediate PCR product as template, a forward vector primer/reverse mutant primer, and forward mutant primer/reverse vector primer. These two segments of amplicons were annealed and reamplified with the forward/reverse vector primers thus inserting the nucleotide substitutions encoding the desired mutant substitutions in one or more steps. Amplified PCR products and the pPIC9K expression vector were cleaved with *Eco*RI and *Not*I, gel purified, and ligated to yield cDNAs, which were therefore under the regulation of the methanol inducible *Paox1* promoter (Invitrogen Corp., Carlsbad, CA). The insertions of substitutions into the plasmid encoding mutant coding sequences were confirmed by sequencing of both strands of DNA templates with primers annealing to either the vector or coding sequences. Sequences were analyzed with an Applied Biosystems model 3777 automated sequencer. The 8.0

Table 1. Conservation of Cathepsin K/Chondroitin Sulfate Interaction Sites in Vertebrates[a]

species, Merops no.	amino acid residues								
	D6	R8	K9	K10	I171	Q172	N190	K191	L195
Mammals									
human 000644	D	R	K	K	I	Q	N	K	L
chimpanzee 071739	D	R	K	K	I	Q	N	K	L
bonobo 344477	D	R	K	K	I	Q	N	K	L
orangutan 292869	D	R	K	K	I	Q	N	K	L
white cheeked gibbon 292870	D	R	K	K	I	Q	N	K	L
rhesus macaque 005622	D	R	K	K	I	Q	N	K	L
crab-eating macaque 004489	D	R	K	K	I	Q	N	K	L
galago (lemur) 344482	D	R	K	K	I	Q	N	K	L
European rabbit 000645	D	R	K	K	I	Q	N	K	L
Norwegian rat 003747	D	R	K	K	*T*	Q	N	K	L
house mouse 002412	D	R	K	K	*T*	Q	N	K	L
naked mole rat 292868	D	R	K	K	I	Q	N	K	L
guinea pig 271203	D	R	K	K	**M**	Q	N	K	L
common marmoset 344476	D	R	K	K	I	*L*	N	K	L
pig 14559	D	R	K	K	I	Q	N	K	L
cow 048532	D	R	K	K	I	Q	N	K	L
alpaca 260489	D	R	K	K	I	Q	N	K	L
horse 98351	D	R	K	K	I	Q	N	K	L
African elephant 292876	D	R	K	K	I	Q	N	K	L
panda bear 233540	D	R	K	K	I	Q	N	K	L
dog 053414	D	R	K	K	I	Q	N	K	L
European hedgehog 104162	D	R	K	K	**V**	Q	N	K	L
little brown bat 102107	D	R	K	K	I	Q	N	K	L
large flying fox 263759	D	R	K	K	I	Q	N	K	L
European polecat 292874	D	R	K	K	**V**	Q	N	K	L
bottlenose dolphin 261510	D	R	K	K	I	Q	N	K	L
gray opossum 165395	D	R	K	K	I	Q	N	K	L
duck billed platypus 097967	D	R	K	K	**A**	Q	N	K	L
Birds									
chicken 001933	D	R	**R**	K	**A**	Q	N	K	L
Reptiles									
green anole lizard 292875	D	R	K	K	**S**	Q	*D*	K	L
Amphibians									
western clawed frog 052064	D	R	K	K	**T**	Q	*D*	K	L
African clawed frog 117331	D	R	K	K	**N**	*E*	*K*	K	L
Bony Fish									
Atlantic salmon 144119	D	R	K	K	*Q*	*T*	*K*	*Q*	**M**
rainbow smelt 183924	D	R	R	K	**V**	*T*	N	K	L
olive flounder 142304	D	R	K	*E*	I	*S*	*K*	K	L
mangrove red snapper 183938	D	R	K	K	**V**	*T*	*K*	*G*	L
stickleback fish 112083	D	R	K	K	**V**	*T*	**Q**	K	**M**
rice fish 111226	D	R	K	K	**V**	*N*	N	K	L
goldfish 118327	D	R	K	*L*	**A**	*T*	*K*	K	L
zebrafish 056529	D	R	K	*L*	**A**	*T*	*K*	K	L
catfish 232420	D	R	K	*L*	**A**	*T*	*M*	K	L
Takifugu pufferfish 265080	D	R	K	K	**V**	**N**	N	*G*	L
pufferfish 126698	D	R	K	K	**V**	*T*	*T*	*E*	L

[a]Chondroitin sulfate binding sites on cathepsin K were taken from ref 13. Bold, conservative substitutions; italic, nonconservative substitutions.

kb pPIC9K plasmids with the CatK variants, *HIS4* 5′ and 3′ *AOX1* sequences, secretory coding, and transcriptional terminator sequences were released by digestion with *Bgl*II, electrophoretically resolved from ampicillin/pBR322 sequences, and gel purified with Qiagen gel extraction kits. This purified fragment was then electroporated into *P. pastoris* GS115 host cells (Invitrogen Corp), which were seeded on MDA plates, and colonies were sequentially purified. Multiple clones expressing putative CatK mutant protein were screened for Mut^S, His^+ phenotypes on selective MDA, MMA, and rich BMGYA replica plates. This phenotype is consistent with the integration of the fragment at the *AOX1* locus of the host strain that is necessary for high-level secretion of the CatK mutant protein regulated by the plasmid borne *Paox1* promoter.

Protein Production in *P. pastoris*. Recombinant wild-type human CatK and variants were expressed in *Pichia pastoris* as described previously.[16] Briefly, frozen stocks of transformed yeast cells at −80 °C were thawed, streaked out onto rich

dx.doi.org/10.1021/bi401051j | *Biochemistry* 2013, 52, 7742−7752

Table 2. Conservation of Cathepsin K-Specific Positively Charged Residues in Vertebrates[a]

	amino acid residues								
species, Merops no.	K9	K77	R79	K103	K106	R108	R111	K122	R127
Mammals									
human 000644	K	K	R	K	K	R	R	K	R
chimpanzee 071739	K	K	R	K	K	R	R	K	R
bonobo 344477	K	K	R	K	K	R	R	K	R
orangutan 292869	K	K	R	K	K	R	R	K	R
white cheeked gibbon 292870	K	K	R	K	K	R	R	K	R
rhesus macaque 005622	K	K	R	K	K	R	R	K	R
crab-eating macaque 004489	K	K	R	K	K	R	R	K	R
galago (lemur) 344482	K	R	R	K	K	R	R	K	R
European rabbit 000645	K	K	R	K	K	R	R	K	R
Norwegian rat 003747	K	*Q*	*G*	K	K	R	R	K	R
house mouse 002412	K	*Q*	*G*	K	K	R	R	K	R
naked mole rat 292868	K	*Q*	R	K	K	R	R	K	R
guinea pig 271203	K	*E*	R	K	K	R	R	K	R
common marmoset 344476	K	K	R	K	K	R	R	K	R
pig 14559	K	K	R	K	K	R	R	K	R
cow 048532	K	K	R	K	K	R	R	K	R
alpaca 260489	K	K	R	K	K	R	R	K	R
horse 98351	K	K	R	K	K	R	R	K	R
African elephant 292876	K	K	R	K	K	R	R	K	R
panda bear 233540	K	K	R	K	K	R	R	K	R
dog 053414	K	K	R	K	K	R	R	K	R
European hedgehog 104162	K	K	R	K	K	R	R	K	R
little brown bat 102107	K	**R**	R	K	K	R	**K**	K	R
large flying fox 263759	K	**R**	R	K	K	R	**K**	K	R
European polecat 292874	K	K	R	K	K	R	R	K	R
bottlenose dolphin 261510	K	K	R	K	K	R	R	K	R
gray opossum 165395	K	K	R	K	K	R	R	K	R
duck billed platypus 097967	K	K	R	K	K	R	R	K	R
Birds									
chicken 001933	**R**	K	R	K	K	R	R	K	R
Reptiles									
green anole lizard 292875	K	K	R	K	K	R	**K**	K	R
Amphibian									
western clawed frog 052064	K	*D*	K	K	*T*	**K**	**K**	K	*N*
African clawed frog 33	K	*D*	R	**R**	*A*	**K**	**K**	K	*L*
Bony fish									
Atlantic salmon 144119	K	*E*	*G*	*M*	Q	R	**K**	K	**K**
rainbow smelt 183924	**R**	*D*	Q	*M*	*S*	R	**K**	*T*	**K**
olive flounder 142304	K	*E*	*G*	*M*	Q	**K**	**K**	*A*	**K**
mangrove red snapper 183938	K	*D*	*G*	*M*	Q	**K**	**K**	*A*	**K**
stickleback fish 112083	K	*D*	*G*	*M*	Q	**K**	**K**	*A*	**K**
rice fish 111226	K	*E*	*G*	*L*	Q	**K**	**K**	*A*	**K**
goldfish 118327	K	*D*	Q	*R*	*T*	**K**	**K**	*T*	**K**
zebrafish 056529	K	*N*	Q	*V*	*S*	R	**K**	*T*	*N*
catfish 232420	K	*D*	*G*	K	*E*	R	**K**	*A*	**K**
Takifugu pufferfish 265080	K	*T*	R	K	*E*	*S*	*E*	*A*	**K**
pufferfish 126698	K	*T*	**K**	K	*E*	R	*E*	*A*	**K**

[a]Bold, conservative substitutions; italic, nonconservative substitutions.

media (BMGYA) plates, and incubated for 24−48 h at 30 °C to allow recovery. Colonies were restreaked onto minimal/selective media (MDA) plates and incubated for 24−48 h at 30 °C to reselect for *His*$^+$ colonies and against the potential loss or rearrangements of integrants at the *AOX1* loci. Single colonies from this plate were inoculated into 50 mL of MD broth and incubated on shakers at 250 rpm and 30 °C but restricted to A600 of ≤140 to preempt possible sporulation and clumping. Each MD culture was inoculated into 500 mL of prewarmed BMGY broth per 2 L flask, in two to four flasks per enzyme preparation and incubated in shakers at 250 rpm and 30 °C. Cell growth was monitored with aliquots periodically withdrawn at 12 h intervals until cell growth plateaued due to the exhaustion of glycerol as a source of carbon. Cells were pooled, centrifuged at 1000 rpm and 25 °C, washed once with 1× PBS, and resuspended in 250−300 mL of BMMY inductive medium. Secretion of CatK was monitored by taking aliquots of cultures at 12 h intervals and testing their protease activity

dx.doi.org/10.1021/bi401051j | *Biochemistry* 2013, 52, 7742−7752

using the fluorogenic cathepsin substrate Z-LR-MCA after pepsin activation as described.[16,17] Cultures were grown in shaker flasks for approximately 4 days where they reached maximal cathepsin expression, and the liquid culture medium was concentrated using a YM10 ultrafiltration membrane with a cutoff size of 10 kDa (Amicon Inc., Beverly, MA). The enzyme was activated with pepsin purified by *N*-butyl-sepharose 4 fast flow chromatography (Amersham Pharmacia Biotech, Piscataway, NJ) as previously described.[7,17] The homogeneity of the proteases was demonstrated by SDS-PAGE. Active site concentrations of CatK peptidases were determined in duplicate titrations with LHVS as previously described using E64.[18] Yields for the mutant and wild-type proteins were between 1 and 2 mL of 5.7−570 μM stocks.

Substrate Assays using Z-Peptidyl-MCA Peptides. Steady-state kinetics of each enzyme were determined in duplicate with the fluorogenic substrates Z-LR-MCA and Z-GPR-MCA, as previously described.[15] The assays were based on excitation and emission wavelengths set at 380 and 460 nm, respectively, using the spectrofluorometer, PerkinElmer LB50. Nonlinear regression analysis was applied to determine Michaelis−Menten constants (K_m) and rate constants (k_{cat}). All enzymes were assayed at fixed enzyme concentrations (1−5 nM) but variable substrate concentrations (1−200 μM) in 100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM DTT and 2.5 mM EDTA at room temperature. The stability of cathepsin proteins in the presence of C4-S was determined by incubating 600 nM wild-type or variant proteins for 8 h in 100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM DTT, 2.5 mM EDTA, and 0.4 mg/mL of soluble calf skin type I collagen at 28 °C. Residual activities in cleaving the synthetic substrate, Z-FR-MCA (2 μM), were determined in duplicate at time zero and after 8 h.

Collagen and Gelatin Digests. Collagenase activities of wild-type CatK (600 nM) and each of the variant proteins (600 nM) were determined by incubating the proteases with 0.6 mg/mL of soluble calf skin type I collagen (U.S. Biochemical Corp., Cleveland, OH) in 100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM DTT and 2.5 mM EDTA. Collagen was digested at 28 °C for 8 h in the presence or absence of 0.1% (w/v) C4-S. The digestion reaction was stopped by the addition of 5 μM of E64. Gelatinase activity of each protease mutant was determined with heat denatured type I collagen (30 min at 70 °C). In the presence of wild-type CatK (20 nM) or mutant (20 nM), the reaction mix was incubated for 30 min at 28 °C. Before addition of LHVS (10 μM), residual activity of each enzyme was monitored spectrofluorometrically using the fluorogenic substrate Z-FR-MCA (2 μM). Peptides from collagen and gelatin digests were resolved by SDS-polyacrylamide gel electrophoresis using 4−20% Tris/glycine gradient gels (Novex, San Diego, CA) and visualized by Coomassie blue staining. Stained cleavage products were quantified by NIH ImageJ densitometry. Densities of pixels of signals bracketing the α1 and α2 chains of collagen in presence or absence of C4-S were measured. Signal densities of peptides resulting from the digestions in the absence of C4-S for each variant protein (no or only minimal intrahelical degradation) were set as 100% and normalized to signal densities of α1 and α2 chains resulting from digestion with the same protease in the presence of C4-S. The ratio of α1/α2 chains in the presence or absence of C4-S was multiplied by 100 to yield the value of percent (%) undigested α1/α2 chains. Densitometric measurements were taken from at least three independent experiments, except for M10 (two readings), averaged, and expressed in percentage of α-chain cleavage.

Electromobility Gel Shift Assay. Gel shift assays demonstrating the formation of a complex between the M6−M9 variants and C4-S were performed as described for the wild-type protein and the M5 variant of the GAG binding site.[11,14]

RESULTS

Conservation of Arginine and Lysine Residues on the Surface of Mammalian CatK Orthologues. Using the Merops protease database, we analyzed 42 vertebrate CatK sequences (27 mammals, 1 bird, 1 reptile, 2 amphibians, and 11 bony fish). We first evaluated the conservation of nine amino acid residues involved in the binding of chondroitin sulfate as deduced from the previously published crystal structure of human CatK[13] (Table 1). From the nine positions, only I171, Q172, and N190 (human CatK sequence) showed some variation among the mammalian, avian, reptilian, and amphibian CatK orthologues. Observed amino acid changes were mostly conservative (I171 → M, V, A, S, T). Residue Q172 was only altered into a leucine residue in the marmoset and a glutamate residue in the African clawed frog. Residue N190 showed changes in the reptilian and amphibian specimens with replacements by aspartate and lysine residues. The same was true for all fish species regarding this position. Fish species showed altogether a higher variability in positions K10, I171, Q172, N190, and K191. With exception of position 171, various nonconservative substitutions were observed, such as the replacement of positively charged lysine residues by neutral or negatively charged residues. Residues Q172 or N190 were frequently either replaced by threonine and serine residues or lysine residues, respectively. Altogether, all positions were highly conserved in mammals with rodents representing a noticeable exception for their nonconservative substitutions in position 171.

Next, we analyzed the positively charged residues K9, K77, R79, K103, K106, R108, R111, K122, and R127 that were specific for CatK among human cathepsins (Table 2). All residues were highly conserved in mammals and avian and reptilian specimens with occasional lysine to arginine and vice versa substitutions. The only noticeable exception was observed in rodent sequences in positions 77 and 79. Here, nonconservative substitutions were frequent such as lysine or arginine to glutamine, glutamate, or glycine. Amphibian CatK orthologues revealed nonconservative amino acid exchanges in positions 77, 106, and 127. The lowest degree of conservation was observed in fish, where with the exception of positions K9, K108, K111, and 127 all other positions showed high variability (77, 79, 103, 106, and 122).

Expression and Characterization of CatK Variants. To extend the identification of positively charged surface amino acid residues of CatK putatively interacting with C4-S or contributing to the formation of oligomeric CatK complexes, we generated eight CatK variants. The identified residues were located in three clusters (Figure 1). Cluster 1 contains K77 and K79, cluster 2 is formed by K103, K106, R108, and R111, while K122 and R127 compose cluster 3. CatK variants were generated containing the individual clusters and combinations thereof. Furthermore, mutants containing all eight substitutions were combined with variants previously characterized for their interactions with the known C4-S site.[13] The eight arginine and lysine residues were replaced with alanine residues, and the C4-

dx.doi.org/10.1021/bi401051j | *Biochemistry* 2013, 52, 7742−7752

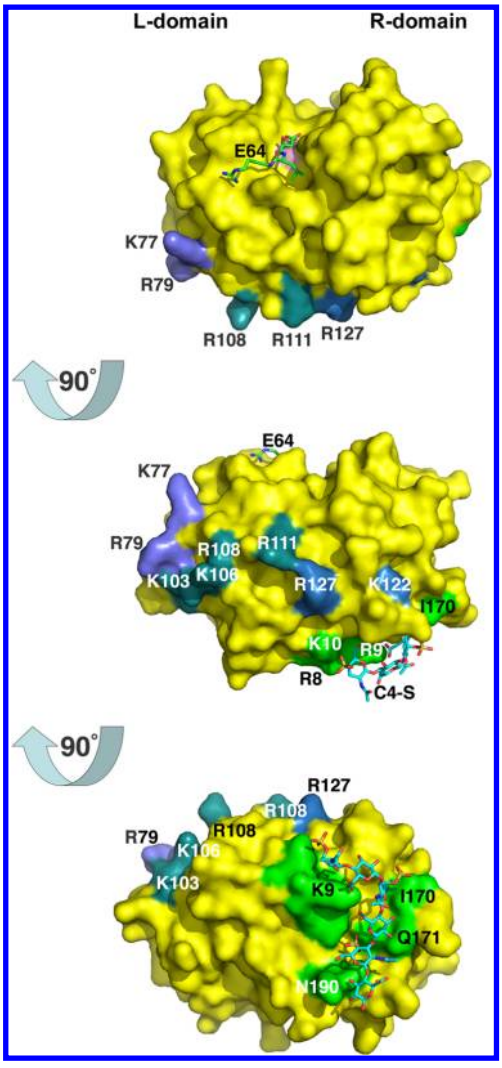


Figure 1. Surface model of human CatK with E64 bound in the active site and chondroitin 4-sulfate (C4-S) bound on the backside of the R-domain (Pymol presentation of 3C9E). The molecule is turned stepwise backward by 90° to visualize the positively charged clusters of arginine and lysine residues at the bottom of the protease. Cluster 1 (K77, R79 in blue), cluster 2 (K103, K106, R108, R111 in light teal), and cluster 3 (K122, R127 in marine). Residues interacting with C4-S (K9, I170, Q171, N190, K191, L195) are in green. The active site cysteine residue (C25) is in purple.

S binding site was replaced with residues present in the cathepsin L sequence. Proteases encoded by wild-type and mutant full-length cDNAs were expressed in *Pichia pastoris*, activated with pepsin, and purified as previously described.[16] All CatK variants were both homogeneous and completely processed into catalytically active mature forms as determined by SDS-PAGE analysis. Variant and wild-type proteins were titrated with LHVS yielding stock concentrations of 293.6 μM (wild-type), 2.4 μM (variant M6), 533 μM (variant M7), 1.1 μM (variant M8), 132 μM (variant M9), 228.7 μM (variant M10), 570 μM (variant M11), 167.8 μM (variant M12), and 58.2 μM (variant M13), as well as 399 μM (variant M7.1), 16.7 μM (variant M7.2), 11.7 μM (variant M7.3), and 74.6 μM (variant M7.4). Enzyme stocks were stable for months when stored at −80 °C.

CatK Variant Activities toward Synthetic Fluorogenic Substrates, Gelatin, and Peptidase Stability. The kinetic parameters, K_m and k_{cat}, were determined to exclude impairment of the activity of the variant proteases due to nonspecific conformational effects of the introduction of substitutions of residues. The proteolytic activity of each mutant and wild-type CatK was measured with the general cathepsin substrate, Z-LR-MCA, and a CatK specific substrate, Z-GPR-MCA. Michaelis–Menten constants (K_m) and second order rate constants (k_{cat}/K_m) were comparable for the wild-type protein and the variants M6–M13 for each of the substrates tested (Table 3). As for the wild-type enzyme, Z-GPR-MCA was hydrolyzed approximately 30-times less efficiently than Z-LR-MCA by all eight CatK variants. With exception of variant M11, the K_m values for each of the substrates varied by equal or less than a factor of 2, the k_{cat} values by less than 2.1, and the k_{cat}/K_m values by less than a factor of 2.2 among all CatK variants. Variant M11 was characterized by a 5-fold reduction of the k_{cat} value resulting compared with the wild-type protease (Table 2). This suggested that the introduced substitutions did not dramatically altered the integrity of the protease variants and thus allowed them to maintain the catalytic and substrate binding efficacies of the mutant enzymes relative to that of the wild-type CatK enzyme.

Moreover, all wild-type and variant CatK proteins displayed after 8 h of incubation in the presence of type I collagen and C4-S at 28 °C residual peptidase activities of approximately 25–45% (Table 4). This indicates comparable enzyme stabilities among all variants.

Heat-denatured collagen (gelatin) was used as a protein substrate to compare the general proteolytic activities of CatK variant proteins with that of the wild-type enzyme. The gelatinase activities of all variants were potent at low enzyme concentrations (20 nM) and hydrolyzed gelatin similarly to wild-type CatK at pH 5.5 (Table 4). Figure 2A displays the gelatinase activity of M9, which is comparable to that of wild-type CatK but does not reveal any activity toward triple helical collagen (Figure 2C). The comparable gelatinase activities between wild-type CatK and its variants suggest that the introduced substitutions do not impair the active site and thus the catalytic activity of the variants.

Degradation of Triple Helical Type I Collagen by CatK Variants and Gel-Shift Mobility. In the presence of C4-S, wild-type CatK degrades triple-helical type I collagen efficiently into small peptides that are not detectable by SDS-polyacrylamide gel electrophoresis. In the absence of C4-S, CatK activity is limited to a cleavage in the nonhelical telopeptide regions leading to the disappearance or reduction of the high-molecular-weight β and γ bands of collagen and an accumulation of α chains (Figure 2B–D). The degree of undigested α1 and α2 chains is considered as a measure of the inhibition of the collagenase activity of a mutant variant. Wild-type CatK and all its variants lacked significant collagenase activities in the absence of C4-S, and the generation of collagen α chains was comparable among all CatK variants. In the presence of C4-S, however, significant differences in the ability of the CatK variants to cleave collagen were observed (Figure 2B–D). The CatK-specific arginine and lysine residues selected for mutational analysis formed three clusters: cluster 1 (K77, R79), cluster 2 (K103, K106, R108, R111), and cluster 3 (K122, R127). Whereas the double mutant M6 (K77A, R79A) had no effect on the collagenolytic activity of this variant, the substitution of K103, K106, R108, and R111 with alanine residues in tetramutant M7 led to a reduction of its collagenolytic activity by about 50% (Figure 2B,E). Combina-

dx.doi.org/10.1021/bi401051j | *Biochemistry* 2013, 52, 7742–7752

Table 3. Kinetic Parameters for the Hydrolysis of ZR-MCA and Z-GPR-MCA by Recombinant Cathepsin K Mutants[a]

	Z-LR-MCA			Z-GPR-MCA		
CatK variant	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\times 10^6$ M^{-1} s^{-1})	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\times 10^6$ M^{-1} s^{-1})
WT	6.2 ± 2.3	3.3 ± 0.5	1.9	2.9 ± 1.4	48 ± 9	6
M6	4.7 ± 0.3	2.7 ± 0.2	1.7	2.3 ± 0.2	44 ± 7	5.2
M7	5.3 ± 0.7	4.4 ± 0.1	1.2	2.1 ± 0.4	49 ± 2	4.3
M7.1	7.3 ± 0.8	5.4 ± 0.7	1.4	3.1 ± 0.2	46 ± 5	6.7
M7.2	6.7 ± 0.5	3.8 ± 0.2	1.8	4.2 ± 0.2	63 ± 0.2	6.7
M7.3	7.3 ± 0.6	3.5 ± 0.3	2.1	3.0 ± 0.2	52 ± 2.0	5.8
M7.4	9.5 ± 1.5	5.4 ± 0.8	1.8	3.4 ± 0.1	54 ± 9.0	6.3
M8	7.0 ± 1.3	3.4 ± 0.4	2.1	4.3 ± 0.2	47 ± 4	9.2
M9	4.6 ± 0.7	2.9 ± 0.3	1.6	2.4 ± 0.1	55 ± 8	4.4
M10	4.7 ± 0.4	3.4 ± 0.3	1.4	1.0 ± 0.03	41 ± 1.9	2.5
M11	0.7 ± 0.2	1.9 ± 0.2	0.4	0.5 ± 0.07	37.6	1.4
M12	7.3 ± 1.0	2.8 ± 0.5	2.6	2.8 ± 0.4	41 ± 7	6.8
M13	4.6 ± 1.3	2.8 ± 0.4	1.6	3.6 ± 0.2	54 ± 12	6.7

[a]Steady state kinetics of each enzyme were determined in duplicate with the fluorogenic substrates Z-LR-MCA and Z-GPR-MCA, as previously described.[15] The rates were monitored at excitation and emission wavelengths set at 380 and 460 nm, respectively, using the spectrofluorometer PerkinElmer LB50. Nonlinear regression analysis was applied to determine the Michaelis–Menten (K_m) and the catalytic constants (k_{cat}). All enzymes were assayed at fixed enzyme concentrations (1–5 nM) but variable substrate concentrations (1–200 μM) in 100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM DTT and 2.5 mM EDTA at room temperature. Kinetic constants are expressed in SEM ± SD derived from at least two independent assays.

Table 4. Gelatinase and Residual Activities of Wild-Type and Cathepsin K Variants[a]

CatK variant	% gelatinase activity	% residual Z-FR-MCA activity (SEM ± SD)
wild-type	100	36.7 ± 10.9
M6	100	38.0 ± 2.1
M7	100	45.2 ± 12.3
M7.1	100	55.7 ± 4.3
M7.2	100	59.4 ± 6.1
M7.3	100	36.9 ± 3.6
M7.4	100	25.9 ± 1.1
M8	100	24.7 ± 5.4
M9	100	42.1 ± 16.0
M10	100	34.2 ± 9.7
M11	100	27.4 ± 1.2
M12	100	45.1 ± 16.7
M13	100	20.3 ± 7.4

[a]To evaluate gelatinase activity, wild-type CatK and its variants (20 nM) were incubated with 0.4 mg/mL of heat-denatured calf skin type I collagen (gelatin) for 30 min at 28 °C in 100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM DTT and 2.5 mM EDTA. The degree of hydrolysis was monitored by SDS-PAGE analysis, which resulted in the complete loss of the γ, β, and α bands of gelatin (see, for an example, Figure 2A). For CatK stability measurements, active wild-type CatK and CatK variants (each 600 nM) were incubated at 28 °C with 0.4 mg/mL of soluble calf skin type I collagen in 100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM DTT and 2.5 mM EDTA. The residual activities of each enzyme were measured by withdrawing aliquots of 5 μL of each reaction into a 2 mL standard Z-FR-MCA assay (2 μM) at time zero and 8 h. Residual activities are expressed in SEM ± SD derived from two to five independent assays.

tion of the mutations present in M6 and M7 to yield the hexamutant, M8, had a negligible effect indicating that residues K77 and R79 had no role in enhancing the collagenase activity. However, further addition of the mutations K122A and R127A to the hexamutant to yield the octamutant, M9, led to a collagenolytically inactive CatK variant (Figure 2C,E). In contrast, the double mutant M10 (K122A, R127A) revealed only an approximately 20% inhibition of the collagenase activity. This implies that residues K103, K106, R108, R111, K122, and R127 contribute to the collagenolytic activity of CatK, whereas K77 and R79 play either no role or a minimal role. An individual mutational analysis of residues K103 (M7.1), K106 (M7.2), R108 (M7.3), and R111 (M7.4) only showed a weak contribution of the single residues, which excludes an additive effect to the enhancement of the collagenolytic activity (Figure 3A,B). These mutations did not reveal any effect on the peptidolytic (Table 3) and gelatinase activities (Table 4) of the CatK variants. The octamutant, M9, resulted in a total loss of the collagenase activity. However, using a gel-shift electromobility assay, we could demonstrate that these mutations did not eliminate the binding between CatK and C4-S. The M9 variant revealed the same electromobility as wild-type CatK in the presence of C4-S (Figure 3C). This is not surprising because the amino acid residues involved in the structure-verified interaction site[13] remain intact. It should be noted that this assay does not allow differentiation between the formation of collagenolytically active and nonactive GAG complexes as previously shown for the M5 variant, which despite mutation in the GAG-binding site retained a slightly shifted GAG binding mode but lost 50% of its collagenolytic activity.[14]

Based on the CatK/C4-S crystal structure, we have shown that the C4-S binding site is scattered around three CatK sequence areas (K9, I171, Q172, and N190, K191, L195).[14] Variants M1 (K9E) and M2 (K9E, I171E, Q172S) lost about 25–35% of their collagenase activities compared with the wild-type enzyme. In contrast, a third variant, M4 (K9E, N190M, K191G, L195K), only displayed an insignificant level of collagenase activity loss.[14] This was interpreted as a stabilization of the C4-S binding in this variant, which thus led to a restoration of its collagenase activity. Therefore, we evaluated this effect on variant M9, by generating variants M11, M12, and M13, which in addition to the octamutations (K77–R127) bear the mutations present in former variants M1, M2, and M4. Similar to the previous observation, M12, revealed a partial recovery of its collagenase activity whereas M11 and M13 lacked the collagenase activities as did variant M9 (Figure

dx.doi.org/10.1021/bi401051j | Biochemistry 2013, 52, 7742–7752

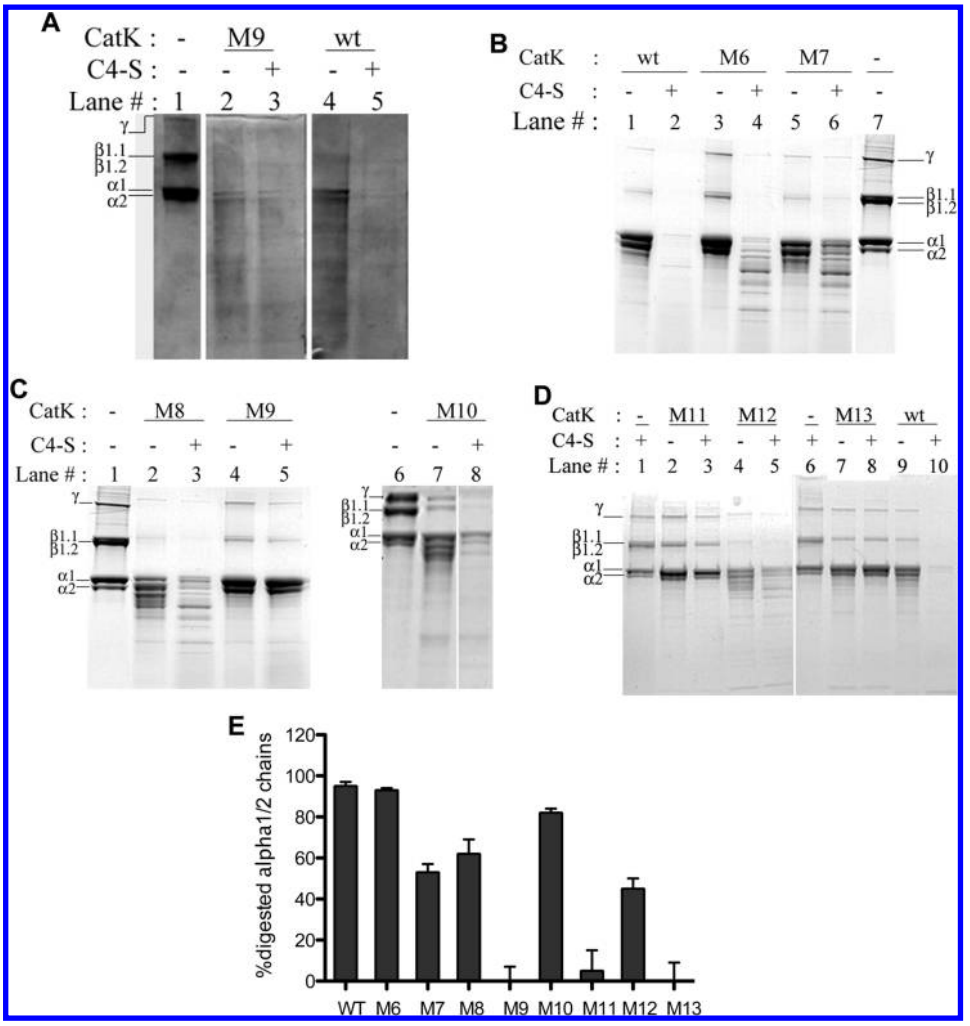


Figure 2. (A) Gelatin degradation by wild-type CatK and M9. (B) Type I collagen degradation by wild-type CatK and variants M6 and M7. (C) Type I collagen degradation by variants M8, M9, and M10. (D) Type I collagen degradation by wild-type CatK and variants M11, M12, and M13. All degradations were performed either in the absence or in the presence of C4-S. The degradation was monitored by SDS-PAGE analysis of degradation mixtures of either heat-denatured type I collagen or triple helical type I collagen in 100 mM acetate buffer, pH 5.5, containing 2.5 mM DTT and 2.5 mM EDTA at 28 °C. The collagen degradation assays contained 600 nM CatK, and the substrate was incubated for 8 h. The gelatinase assay contained 20 nM CatK, and the incubation time was 30 min. Coomassie stained gels were scanned and quantified with Image J (NIH) densitometry. (E) Bar representation of the % digested α-bands of type I collagen after incubation with wild-type CatK and its variants.

2D,E). This result again confirms the contribution of the previously identified C4-S binding site to the collagenase activity of CatK.

■ DISCUSSION

We have previously demonstrated that the collagenase activity of CatK requires specific interactions with GAG components present in bone and cartilage.[10,11] The crystal structure of the complex between wild-type CatK and chondroitin 4-sulfate (C4-S) revealed nine amino acid residues interacting directly or indirectly via water molecules with the GAG.[13] Sequence alignment analyses of 42 CatK sequences from mammalian, avian, reptilian, amphibian, and fish species showed a high degree of conservation in these residues. Besides occasional conservative substitutions, all nine residues were completely conserved among the 27 mammalian CatK sequences. The degree of conservation gradually decreased from reptiles through amphibian to fish specimens. The replacement of the specific C4-S interacting residues in CatK with appropriate amino acid residues present in cathepsin L only led to a 60% reduction of the collagenolytic efficacy suggesting additional factors contributing to this unique activity. We hypothesized that there exist additional or alternative GAG binding sites on the highly positively charged lower area of the molecule. Three arginine/lysine clusters containing a total of eight CatK-specific residues were identified compared with the other members of the cathepsin L subfamily (cathepsins L, V, and S). An analysis of all vertebrate CatK sequences presently available revealed that basic arginine and lysine residues are fully conserved in positions 9, 103, 106, 108, 111, 122, and 127 among mammalian, avian, and reptilian CatK orthologues. The only noticeable exception was observed in rodents, which displayed nonconservative substitutions in residues 77 and 79. This finding was supported by our mutational analysis of the three arginine/lysine clusters. Cluster 1 (K77, R79) had no role in enhancing the collagenase activity of CatK. Of more importance is cluster 2, which contains residues K103, K106, R108, and R111. All four basic residues were completely conserved among mammalian, avian, and reptilian species. Not surprisingly, cluster 2 contributed about 50% of the collagenase

dx.doi.org/10.1021/bi401051j | Biochemistry 2013, 52, 7742−7752

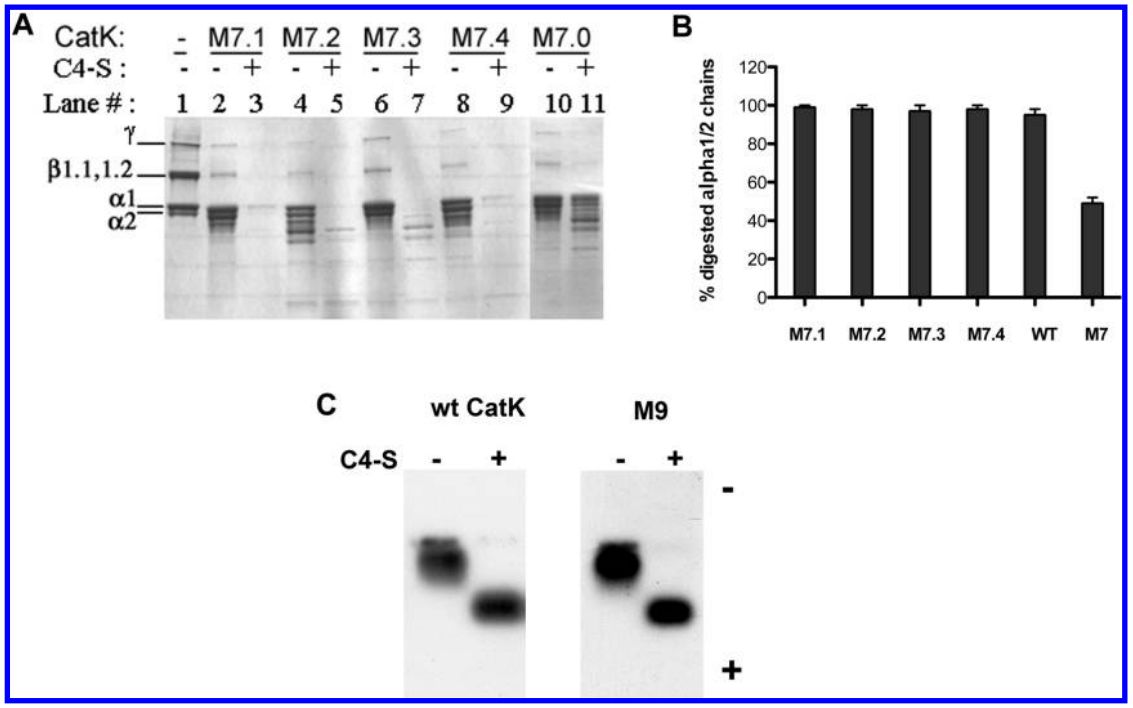


Figure 3. (A) Type I collagen degradation by single mutation variants M7.1, M7.2, M7.3, and M7.4 in comparison to the tetramutant variant M7. Degradations were performed either in the absence or presence of C4-S. The degradation was monitored by SDS-PAGE analysis of degradation mixtures of either heat-denatured type I collagen or triple helical type I collagen in 100 mM acetate buffer, pH 5.5, containing 2.5 mM DTT and 2.5 mM EDTA at 28 °C. The collagen degradation assays contained 600 nM CatK, and the substrate was incubated for 8 h. Coomassie stained gels were scanned and quantified with Image J (NIH) densitometry. (B) Bar representation of the % digested α-bands of type I collagen after incubation with wild-type CatK and its variants. (C) Electromobility assay of wild-type CatK and the M9 variant. In the presence of C4-S, the CatK protein migrates further to the anode.

activity suggesting (i) an additional GAG binding site, (ii) a specific site needed for triple helical collagen binding, or (iii) protein−protein interaction sites in the oligomeric CatK/GAG complexes. Residues K106 and R108 are in the vicinity of the closest contact between two alternating CatK molecules bridged by a C4-S chain as observed by crystallography.[13] However, none of the four individual residues appear decisive because no collagenase-inhibiting effects were observed from the single mutants. The third cluster, K122, R127, reduced the collagenase activity by about 20% but in combination with cluster 2 led to a complete loss of the collagenolytic activity. Both sites do not seem to act in an additive manner but rather in a cooperative mode. K122 and R127 are completely conserved among mammalian, avian, and reptilian CatK sequences. Amphibia appear to have residue 122 conserved and fish mostly residue 127.

It should be noted that all three clusters are remote from the active site of CatK and its associated substrate binding area (S3−S2′), and are located on the lower side of the R (cluster 1, K77, R79, and cluster 2, K103, K106, R108, R111) and L domains (cluster 3, K122, R127) (Figure 1). Therefore, it was not surprising that the specificity constants k_{cat}/K_m for variants M6−M13 were comparable to that of the wild-type enzyme and thus did not affect the integrity of the catalytic site (Table 3). The same was true for the unabridged gelatinase activities in all variants and their overall proteolytic stabilities.

Recently, Novinec et al.[19] proposed two additional or alternative GAG binding sites on CatK by *in silico* modeling. Interestingly, energetically, the most favorable second docking site included residues R108, R111, and R127. Besides residue K9, these three residues have the highest degree of conservation among all species analyzed (Table 2). It should be noted that variant, M9, and the composite variants, M11 and M13, all included alanine replacements of these residues that resulted in the total loss of the collagenase activity. When either R111A in variants M7 and M8 or R127A in variant M10 are individually present, the effect on the inhibition of the collagenolytic activity was only between 20% and 50%, whereas when R111A and K127A both are present as in variants M9, M11, and M13, a total inhibition was observed. Although R111 and R127 are on different amino acid sequence clusters, both residues are in close vicinity on the bottom surface of CatK and may jointly interact with a putative GAG binding side as proposed by Novinec et al.[19]

The mutational analysis of the collective arginine/lysine cluster variants and those of the C4-S binding site variants based on previous structural analysis[14] also revealed the activating effect of residues present in the C4-S binding site. The M12 variant retained about 50% of its collagenase activity despite all the arginine/lysine mutations that are present in M9 (resulting in the total loss of the collagenolytic activity). This may reflect the presence of an additional domain of mutations (N190M, K191E, L195K) that has been previously shown to partially rescue the inhibitory effect of a K9E mutation alone.[14] The introduction of additional mutations such as I171E and Q172S in variant 13 led to a complete blocking of collagen degradation and thus reflects the maximal anticollagenase effect previously observed in the M5 variant.[14] It is likely that the I171E mutation has the largest effect in this set of CatK variants because it is in close vicinity to a GAG sulfate group as revealed in the X-ray structure.[13,14] The glutamate mutation repositions

dx.doi.org/10.1021/bi401051j | *Biochemistry* 2013, 52, 7742−7752

the GAG chain that contributes to the reduction of the collagenase activity.[14]

The complete analysis of all variants revealed that there must be at least two sites of interaction required for the unique collagenase activity of CatK: (i) the previously identified GAG binding site on the left side of the R domain and (ii) additional lysine/arginine-mediated interactions on the bottom part of the protease. These could be additional GAG binding sites required for the formation of an active CatK/GAG complex or sites involved in protease−collagen or CatK−CatK interactions within the collagenase/GAG complex. It is of interest that wild-type CatK as well as its M5 variant form significant protein−protein interfaces that may play a role in the formation of the active collagenase complex of CatK with C4-S.[14] Crystallographic data suggest a dimerization of two CatK molecules along the Y98 loop extending from Y87 to G101.[14] Residue K102 is in the direct neighborhood to this loop. A Y98C mutation was identified to cause a CatK activity deficiency based disorder, pycnodysostosis, which is responsible for a bone-sclerosing phenotype.[20,21] Interestingly, the Y98C mutation led to a proteolytictically active enzyme that specifically lost its collagenase activity[22] and was unable to form complexes with GAGs as suggested by the lack of the electromobility shift observed in the wild-type protease.[10] On the other hand, residues K106, R108, and to a lesser extent R111 are in the vicinity of the boundary of two alternate neighboring CatK molecules held together by the GAG bridge as seen in the crystal structure.[13] This may suggest that at least four CatK and two GAG molecules are involved in a collagenolytically active complex. The exact knowledge of the architectures of this complex may help to design exosite inhibitors, which will selectively inhibit the disease-relevant collagenase activity of cathepsin K but leave the other proteolytic activities intact. Recent studies have shown that cathepsin K also degrades regulatory proteins such as TGF-β and thyroglobulin, which have been implicated in homoestasis of lung tissue and in brain function.[23−25]

AUTHOR INFORMATION

Corresponding Author

*Mailing address: University of British Columbia, 2350 Health Sciences Mall, Life Sciences Institute, Rm 4558, Vancouver, BC V6T 1Z3, Canada. Phone: (604) 822-1787. Fax: (604) 822-3562. E-mail: dbromme@dentistry.ubc.ca.

Present Addresses

‡F.S.N.: Cancer Institute of New Jersey, The State University of New Jersey, 679 Hoes Lane West, Piscataway, NJ 08854, United States.

§F.L.: INSERM, UMR 1100, Pathologies Respiratoires: protéolyse et aérosolthérapie, Centre d'Etude des Pathologies Respiratoires, Université François Rabelais F-37032 Tours cedex, France.

∥Z.I.: Imclone Systems, A wholly owned subsidiary of Eli Lilly and Company, 59/61 Imclone Dr, Branchburg, NJ 08876.

Author Contributions

F.S.N. and F.L. contributed equally to this work.

Funding

This work was supported, in part, by a Canada Research Chair Award: Protease in Diseases (to DB), a National Institutes of Health grant AR 48669 and the Canadian Institutes of Health Research grant MOP89974.

Notes

The authors declare no competing financial interest.

ABBREVIATIONS USED

CatK, cathepsin K; C4-S, chondroitin 4-sulfate; DTT, dithiothreitol; E64, L-3-carboxy-*trans*-2,3-epoxypropionyl-leucylamido-(4-guanidino)butane; EDTA, ethylenediaminetetraacetic acid; GAG, glycosaminoglycan; LHVS, morpholino-leucine-homophenylalanine-vinylsulfone; MCA, 4-methyl-7-coumarylamide; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; Z, benzyloxycarbonyl

REFERENCES

(1) Riis, B. J., Hansen, M. A., Jensen, A. M., Overgaard, K., and Christiansen, C. (1996) Low bone mass and fast rate of bone loss at menopause: Equal risk factors for future fracture: A 15-year follow-up study. *Bone 19*, 9−12.

(2) Affinito, P., Palomba, S., Sorrentino, C., Di Carlo, C., Bifulco, G., Arienzo, M. P., and Nappi, C. (1999) Effects of postmenopausal hypoestrogenism on skin collagen. *Maturitas 33*, 239−247.

(3) Baroni Edo, R., Biondo-Simoes Mde, L., Auersvald, A., Auersvald, L. A., Montemor Netto, M. R., Ortolan, M. C., and Kohler, J. N. (2012) Influence of aging on the quality of the skin of white women: The role of collagen. *Acta Cir. Bras. 27*, 736−740.

(4) Yasuda, Y., Kaleta, J., and Bromme, D. (2005) The role of cathepsins in osteoporosis and arthritis rationale for the design of new therapeutics. *Adv. Drug Delivery Rev. 57*, 973−993.

(5) Bromme, D., and Okamoto, K. (1995) Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution. *Biol. Chem. Hoppe-Seyler 376*, 379−384.

(6) Drake, F. H., Dodds, R. A., James, I. E., Connor, J. R., Debouck, C., Richardson, S., Lee-Rykaczewski, E., Coleman, L., Rieman, D., Barthlow, R., Hastings, G., and Gowen, M. (1996) Cathepsin K, but not cathepsins B, L, or S, is abundantly expressed in human osteoclasts. *J. Biol. Chem. 271*, 12511−12516.

(7) Brömme, D., Okamoto, K., Wang, B. B., and Biroc, S. (1996) Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O2 in Spodoptera frugiperda and characterization of the enzyme. *J. Biol. Chem. 271*, 2126−2132.

(8) Kafienah, W., Bromme, D., Buttle, D. J., Croucher, L. J., and Hollander, A. P. (1998) Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix. *Biochem. J. 331*, 727−732.

(9) Garnero, P., Borel, O., Byrjalsen, I., Ferreras, M., Drake, F. H., McQueney, M. S., Foged, N. T., Delmas, P. D., and Delaisse, J. M. (1998) The collagenolytic activity of cathepsin K is unique among mammalian proteinases. *J. Biol. Chem. 273*, 32347−32352.

(10) Li, Z., Hou, W. S., Escalante-Torres, C. R., Gelb, B. D., and Bromme, D. (2002) Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate. *J. Biol. Chem. 277*, 28669−28676.

(11) Li, Z., Yasuda, Y., Li, W., Bogyo, M., Katz, N., Gordon, R. E., Fields, G. B., and Bromme, D. (2004) Regulation of collagenase activities of human cathepsins by glycosaminoglycans. *J. Biol. Chem. 279*, 5470−5479.

(12) Yasuda, Y., Li, Z., Greenbaum, D., Bogyo, M., Weber, E., and Bromme, D. (2004) Cathepsin V, a novel and potent elastolytic activity expressed in activated macrophages. *J. Biol. Chem. 279*, 36761−36770.

(13) Li, Z., Kienetz, M., Cherney, M. M., James, M. N., and Bromme, D. (2008) The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex. *J. Mol. Biol. 383*, 78−91.

(14) Cherney, M. M., Lecaille, F., Kienitz, M., Nallaseth, F. S., Li, Z., James, M. N., and Bromme, D. (2011) Structure-activity analysis of cathepsin K/chondroitin 4-sulfate interactions. *J. Biol. Chem. 286*, 8988−8998.

(15) Lecaille, F., Choe, Y., Brandt, W., Li, Z., Craik, C. S., and Bromme, D. (2002) Selective inhibition of the collagenolytic activity of human cathepsin K by altering its S2 subsite specificity. *Biochemistry 41*, 8447−8454.
(16) Linnevers, C. J., McGrath, M. E., Armstrong, R., Mistry, F. R., Barnes, M., Klaus, J. L., Palmer, J. T., Katz, B. A., and Brömme, D. (1997) Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex. *Protein Sci. 6*, 919−921.
(17) Bromme, D., Nallaseth, F. S., and Turk, B. (2004) Production and activation of recombinant papain-like cysteine proteases. *Methods 32*, 199−206.
(18) Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M., and Hanada, K. (1982) L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochem. J. 201*, 189−198.
(19) Novinec, M., Kovacic, L., Lenarcic, B., and Baici, A. (2010) Conformational flexibility and allosteric regulation of cathepsin K. *Biochem. J. 429*, 379−389.
(20) Gelb, B. D., Brömme, D., and Desnick, R. J. (2001) Pycnodysostosis: Cathepsin K deficiency, in *The Metabolic and Molecular Bases of Inherited Diseases* (Sriver, C. R., Beaudet, A. L., Valle, D., and Sly, W. C. S., Eds.), pp 3453−3468, McGraw-Hill. Inc., New York, St. Louis, San Francisco.
(21) Gelb, B. D., Shi, G. P., Chapman, H. A., and Desnick, R. J. (1996) Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science 273*, 1236−1238.
(22) Hou, W.-S., Brömme, D., Zhao, Y., Mehler, E., Dushey, C., Weinstein, H., Miranda, C. S., Fraga, C., Greig, F., Carey, J., Rimoin, D. L., Desnick, R. J., and Gelb, B. D. (1999) Cathepsin K: Characterization of novel mutations in the pro and mature polypeptide regions causing pycnodysostosis. *J. Clin. Invest. 103*, 731−738.
(23) Tepel, C., Bromme, D., Herzog, V., and Brix, K. (2000) Cathepsin K in thyroid epithelial cells: Sequence, localization and possible function in extracellular proteolysis of thyroglobulin. *J. Cell Sci. 113*, 4487−4498.
(24) Dauth, S., Sirbulescu, R. F., Jordans, S., Rehders, M., Avena, L., Oswald, J., Lerchl, A., Saftig, P., and Brix, K. (2011) Cathepsin K deficiency in mice induces structural and metabolic changes in the central nervous system that are associated with learning and memory deficits. *BMC Neurosci. 12*, No. 74.
(25) Zhang, D., Leung, N., Weber, E., Saftig, P., and Bromme, D. (2011) The effect of cathepsin K deficiency on airway development and TGF-beta1 degradation. *Respir. Res. 12*, No. 72.