



HAL
open science

Rat cathepsin K: Enzymatic specificity and regulation of its collagenolytic activity

Fabien Lecaille, Thibault Chazeirat, Krzysztof Bojarski, Justine Renault, Ahlame Saidi, V. Gangadhara N.V. Prasad, Sergey Samsonov, Gilles Lalmanach

► To cite this version:

Fabien Lecaille, Thibault Chazeirat, Krzysztof Bojarski, Justine Renault, Ahlame Saidi, et al.. Rat cathepsin K: Enzymatic specificity and regulation of its collagenolytic activity. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2020, 1868 (2), pp.140318 -. 10.1016/j.bbapap.2019.140318 . hal-03677265

HAL Id: hal-03677265

<https://hal.science/hal-03677265v1>

Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Rat cathepsin K: enzymatic specificity and regulation of its collagenolytic activity

Fabien Lecaille^{1,2,¶}, Thibault Chazeirat^{1,2,*}, Krzysztof K. Bojarski^{3,*}, Justine Renault^{1,2,#}, Ahlame Saidi^{1,2}, V. Gangadhara N. V. Prasad⁴, Sergey Samsonov³ & Gilles Lalmanach^{1,2,¶}

¹ University of Tours, F-37032, Tours, France.

² INSERM, UMR 1100, Research Center for Respiratory Diseases (CEPR), Team: "Proteolytic Mechanisms in Inflammation", Tours, France.

³ Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, Gdańsk 80-308, Poland.

⁴ Department of Veterinary Pharmacology and Toxicology, College of Veterinary Science, Hyderabad, India.

¶ To whom correspondence should be addressed:

Fabien Lecaille, PhD (Tel : (+33) 247366047 ; e-mail : fabien.lecaille@univ-tours.fr) & Gilles Lalmanach, PhD (Tel : (+33) 247366151 ; e-mail : gilles.lalmanach@univ-tours.fr)
INSERM, UMR 1100 CEPR, University of Tours, Faculty of Medicine, 10 Boulevard Tonnellé, F-37032 Tours cedex, France.

* Authors contribute equally to the work

Present address: Biologie Fonctionnelle et Adaptative, CNRS UMR 8251, University of Paris Diderot-Paris7, Paris, France.

ABSTRACT (247 words)

Human cathepsin K (hCatK), which is highly expressed in osteoclasts, has the noteworthy ability to cleave type I and II collagens in their helical domain. Its collagenase potency depends strictly on the formation of an oligomeric complex with chondroitin 4-sulfate (C4-S). Accordingly, hCatK is a pivotal protease involved in bone resorption and is an attractive target for the treatment of osteoporosis. As rat is a common animal model for the evaluation of hCatK inhibitors, we conducted a comparative analysis of rat CatK (rCatK) and hCatK, which share a high degree of identity (88%) and similarity (93%). The pH activity profile of both enzymes displayed a similar bell-shaped curve (optimal pH: 6.4). Presence of Ser134 and Val160 in the S2 pocket of rCatK instead of Ala and Leu residues, respectively, in hCatK, led to a weaker peptidase activity, as observed for mouse CatK. Also, regardless of the presence of C4-S, rCatK cleaved in the nonhelical telopeptide regions of both type I (tail) and type II (articular joint) rat collagens. Structure-based computational analyses (electrostatic potential, molecular docking, molecular dynamics, free energy calculations) sustained that the C4-S mediated collagenolytic activity of rCatK obeys distinct molecular interactions from those of hCatK. Additionally, T-kininogen (a.k.a. thioestatin), a unique rat serum acute phase molecule, acted as a tight-binding inhibitor of hCatK ($K_i = 0.11 \pm 0.05$ nM). Taken into account the increase of T-Kininogen level in inflamed rat sera, this may raise the question of the appropriateness to evaluate pharmacological hCatK inhibitors in this peculiar animal model.

Keywords: cathepsin K, chondroitin 4-sulfate, collagen, cysteine protease, protein-glycosaminoglycan interactions, T-kininogen.

Abbreviations:

C4-S : chondroitin 4-sulfate ; CatK : cathepsin K; DMF : N,N-dimethylformamide; E-64: trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane ; dp : degree of polymerization ; DTT: dithiothreitol, GAG : glycosaminoglycan; MM-GBSA : Molecular Mechanics-Generalized Born Surface Area ; NRS : normal rat serum; IRS : inflamed rat serum; TK : T-kininogen; RIEP : rocket immunoelectrophoresis.

INTRODUCTION

Cathepsin K (CatK), a cysteine protease of the papain family that is highly expressed in osteoclasts, is one of the major enzymes involved in bone and articular cartilage resorption in humans [1–4]. Among mammalian peptidases CatK is a unique collagenase that cleaves efficiently type I and II collagen (the major organic matrix of bone and articular cartilage, respectively) at multiple sites in their helical domain [5,6]. It has been shown that cartilage-resident glycosaminoglycan chains such as chondroitin 4-sulfate (C4-S), enhances specifically the collagenolytic activity of human CatK (hCatK) by forming an active complex, via multiple contacts between the negatively charged C4-S sulfate and carboxyl groups and positive residues on the backside of the enzyme located away from its active site [7–9]. In the absence of C4-S, hCatK loses its ability to degrade fibrillar collagens but retains its overall peptidase activity. Genetic and pharmacological studies have demonstrated that CatK inhibition reduces bone resorption and increases overall bone mineral density and strength [10]. As a result, hCatK is considered as a valuable target for the development of novel anti-resorptive drugs in osteoporosis and arthritic diseases [11]. Conventional inhibitors target efficiently and selectively the active site of hCatK, thus causing a complete inhibition of its entire enzymatic activity, which may affect other physiological functions. As a matter of fact, the most advanced and promising inhibitor of hCatK, Odanacatib failed in human clinical trial phase III due to safety concerns in 2016 [12]. New strategies focusing on the blockade of the matrix-degrading function of hCatK without affecting its non-collagenolytic activity would be greatly advantageous. One promising approach consists in disrupting C4-S/hCatK interactions by an ectosteric inhibitor to block its collagenase activity without affecting the overall proteolytic activity [13–16].

Besides mouse (*Mus musculus*), rat (*Rattus norvegicus*) is a commonly used animal model for the study of human diseases including osteoporosis and rheumatoid arthritis. Interestingly, rats have the unique property among mammals to possess an additional third kininogen, T-kininogen (TK) (a.k.a. thiostatin, rat α 1-MAP, or rat α 1-cysteine protease inhibitor [17] besides high molecular weight kininogen (HK) and low molecular weight kininogen (LK). Kininogens are multifunctional glycosylated tight-binding inhibitors of cysteine cathepsins related to cystatins (clan IH, family I25) [18]. Conversely to HK and LK, TK is the only mammalian kininogen described to date whose plasma concentration increases markedly during inflammatory episodes

including arthritis [19]. Although analyses of substrate specificity of rat CatK (rCatK) have been surveyed earlier [20], whether collagenolytic activity of rCatK is mediated by C4-S has never been reported at our knowledge. Moreover, various CatK inhibitors display distinctly different potencies to block the active site of rat/mouse CatK versus hCatK regardless of their high amino acid sequence identity (rCatK *vs* hCatK: 88.4%) [20–22].

Herein, we conducted a comparative study of the substrate specificity of both rCatK and hCatK using AMC-derived peptides and internally quenched fluorescent peptidyl substrates, with a focus on the S2 binding pocket. Differences in the substrate specificity were supported by molecular modeling. Then the *in vitro* C4-S-mediated collagenolytic activity of rCatK was investigated and compared to hCatK. Conversely to hCatK, recombinant rCatK showed a poor collagenolytic activity. Structural differences in the GAG-binding site of rCatK were considered via *in silico* studies. Also, K_i values showed that TK is a potent serum inhibitor of both rCatK and hCatK. Taken together the ability of normal rat serum (NRS) and inflamed turpentine-treated rat serum (IRS) to inhibit hCatK and the increased level of TK in IRS, the biological relevance of using this experimental animal model for the pharmacological evaluation of hCatK inhibitors is therefore questionable.

EXPERIMENTAL PROCEDURES

Substrates

Z-Phe-Arg-AMC, Z-Leu-Arg-AMC, Z-Gly-Pro-Arg-AMC were purchased from Bachem (Weil am Rhein, Germany). Intramolecularly quenched fluorogenic substrates Abz-HPGGPQ-(3-NO₂-Tyr) and Abz-RPPGFSPFR-(3-NO₂-Tyr) were prepared as reported elsewhere [23,24]. Chondroitin 4-sulfate (C4-S) from bovine trachea (~20-30kDa) was purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). Rat tail type I collagen was purchased from Thermo Fisher Scientific (Illkirch, France) and rat type II collagen (articular joints) was from Chondrex (Redmond, WA, USA).

Rat Serum

Adult Wistar rats (300g of body weight) were provided by Mahaveer Enterprises (Hyderabad, India). Briefly turpentine (0.5ml/100g weight) was administered subcutaneously to induce inflammation; 48h later, blood was withdrawn from either untreated or turpentine-injected animals. Experiments were approved by Institutional Animal Ethics Committee, N.T.R College of Veterinary Science, Gannavaram, Andhra Pradesh, India [25]. Normal (NRS) and inflamed (IRS) rat sera were prepared as previously reported [26].

Enzymes and inhibitors

Rat recombinant cathepsin K (rCatK) was from BioVision (Milpitas, CA, USA). Human recombinant cathepsin K (hCatK) was produced as described elsewhere [27]. E-64 came from Sigma-Aldrich. The activity buffer for the enzyme assay was 100mM sodium acetate buffer, pH5.5, containing 2mM DTT (dithiothreitol, Bachem). The enzyme active site of both enzymes was titrated by E-64, using Z-Leu-Arg-AMC as substrate [28]. Rat T-kininogen (TK) was purified according to [29]. TK was titrated by E-64-titrated commercial papain [28].

T-kininogen immunoassay

Barbital buffer and polyethylene glycol (PEG) came from Sigma-Aldrich. Agarose was supplied by Amersham Biosciences (GE Healthcare, Buckinghamshire, UK). Rabbit polyclonal anti-TK antibody was obtained as reported elsewhere [26]. The concentration of TK in the serum of turpentine-treated rats was determined by rocket immunoelectrophoresis (RIEP) [30]. Briefly 10ml of 0.8% agarose gel was prepared in 60mM barbital buffer, pH8.6, containing 1% PEG

(v/v) by gentle heating. After cooling molten agarose to 45-50°C, 0.5ml of anti-TK antiserum was added to agarose solution and mixed smoothly to ensure uniform distribution of antiserum. Agarose plate was further placed on the template holder and 3mm wells were crunched then loaded by 5µl of different samples (IRS dilutions, 1/16-1/256; NRS dilutions, 1/2-1/64). Electrophoresis was run (60mM barbital buffer, pH8.6) for 4h at constant voltage (200V). The agarose gel was washed six times with NaCl 150mM, then washed overnight under gentle shaking. Colloidal Coomassie G-250 dye was used for staining (Sigma-Aldrich). According to NRS contains 0.5±0.05g/l of TK [26,31] the antigen (IRS TK) concentration was deduced from the standard NRS calibration graph [30].

Kinetic measurements

Assays using *peptidyl-AMC substrates* (Z-Phe-Arg-AMC, Z-Leu-Arg-AMC and Z-Gly-Pro-Arg-AMC) were monitored using a microplate reader fluorimeter (SpectraMax Gemini, Molecular Devices, Saint Grégoire, France) (λ_{ex} :350nm; λ_{em} :460nm). Alternatively, a Kontron SFM 25 spectrofluorimeter (Kontron AG, Zurich, Switzerland) was used to monitor the fluorescence release from intramolecularly quenched fluorogenic peptides Abz-HPGGPQ-(3-NO₂-Tyr) and Abz-RPPGFSPFR-(3-NO₂-Tyr) (λ_{ex} :350nm; λ_{em} :460nm). Cleavage sites of Abz-HPGGPQ-(3-NO₂-Tyr) or Abz-RPPGFSPFR-(3-NO₂-Tyr) by rCatK were identified as reported elsewhere [23]. The determination of *k_{cat}* and *K_m* was performed at 28°C at a constant enzyme concentration (1nM) and variable substrate concentrations (1-100µM) in 100mM sodium acetate buffer, pH5.5, containing 2mM DTT, and in the absence of 0.15% (w/v) C4-S. Data were plotted in Michaelis-Menten representation and analyzed by non-linear regression, running the Enzfitter software (Biosoft, Cambridge, UK). Alternatively, *K_m* was determined graphically from Hanes linear plot using various concentrations of substrate (0.1-10µM), while second order rate constants (*k_{cat}/K_m*) were measured under pseudo first-order conditions (i.e. using a substrate concentration far below the *K_m*). Kinetic data were reported as mean±SD (n=2). Assays were repeated in the presence of 0.15% (w/v) C4-S.

Inhibition assays

Rat CatK (2.5nM) was pre-incubated in the presence of increasing concentration of TK (0-20nM) in the activity buffer for 15min at room temperature 25°C. After addition of Z-Leu-Arg-AMC (5-

20 μ M), the residual enzymatic activity was followed by monitoring the fluorescence release. Data were plotted and inhibition constants (K_i) were determined using Easson-Steadman procedure as reported previously [32]. Alternatively, lyophilized normal rat serum (NRS) and inflamed rat serum (IRS) were solubilized in 10ml of 0.1M, sodium acetate buffer pH5.5 with 10% DMF. Total protein quantification was determined using BCA Protein assay kit (Thermo Fisher Scientific). Human CatK (0.5nM) was pre-incubated in the activity buffer with increasing doses of rat sera for 15min at 28°C before adding Z-Leu-Arg-AMC (20 μ M) as substrate. All measurements were at least repeated three times.

pH activity profile

The pH activity profile of both hCatK (1nM) and rCatK (1nM) were determined with Z-Leu-Arg-AMC (0.5 μ M) at a low concentration ($S < K_m$, where the initial rate, V_0 , is directly proportional to the k_{cat}/K_m value). The following buffers were used for the pH activity profile: 100mM sodium citrate (pH2.8-5.8), 100mM sodium phosphate (pH5.8-8), and 100mM sodium borate (pH8-10). All buffers contained 2mM DTT. A three-protonation model was used for least-squares regression analysis of the pH activity data, according to [33]. The data were fitted to the equation: $(k_{cat}/K_m)_{obs} = (k_{cat}/K_m)/([H^+]/K_1 + 1 + K_2/[H^+])$, running the Enzfitter software (Biosoft).

Collagenolytic activity

Type I rat collagen (0.4mg/ml) was incubated with rCatK or hCatK (20nM) in the activity buffer (pH5.5) at 28°C for 0-8h, in the absence or presence of 0.15% (w/v) C4-S. In parallel, assays were performed by pre-incubating samples with 10 μ M E-64 before adding collagen. Before stopping the reactions by addition of reducing SDS-PAGE sample buffer, an aliquot of each sample was withdrawn to monitor the residual cathepsin activity (0, 2, 4, and 8h) using Z-Leu-Arg-AMC as substrate (10 μ M) (triplicate assays). Collagen digests were subjected to SDS-PAGE (10 % or 4-20%), and degradation products were visualized by Coomassie Blue staining. Experiments were repeated with type II rat collagen (0.4mg/ml). Assays were repeated to confirm the reproducibility of the results (n=3).

Docking P2 residue of fluorogenic peptidyl substrates

Model of rCatK was obtained from SWISS-MODEL database (UniProtKB AC: O35186). Sequence alignment of rCatK and hCatK (UniProtKB AC: P43235) was performed using Pairwise Sequence Alignment tool available at EMBL-EBI [34] with BLOSUM62 similarity matrix. Z-Phe-Arg-AMC structure was obtained from previous study [35]. For docking simulation of substrate, Autodock 3 [36] was used first with hCatK with a grid box with dimensions of 60Å×60Å×60Å and a grid spacing of 0.375Å. One run of the Lamarckian genetic algorithm with default parameters was carried out. Among the 10 structures obtained, the structure closest to the position of an inhibitor in analogous complex with vinyl sulfone inhibitor (pdb ID: 1MEM) was selected for further MD calculations. Based on this position, docking (Autodock 3) was then performed with rCatK using Z-Phe-Arg-AMC, Z-Leu-Arg-AMC and Gly-Pro-Arg-AMC, respectively.

Docking C4-S to rat and human cathepsin K

First, electrostatic potential isosurfaces were calculated for rCatK and hCatK (pdb ID: 4N8W). A PBSA program from AmberTools [37] was used with a grid spacing of 1Å. Hexameric (dp6) and dodecameric (dp12) C4-S were modeled with AMBER16. Their charges were obtained from GLYCAM06 force field [38] and from literature for sulfate groups [39]. For docking simulations Autodock 3 was further used. A grid box with dimensions of 126Å×126Å×126Å and a grid spacing of 0.453Å containing the whole cathepsin K molecule was applied for GAGs docking. Independent runs (100) of the Lamarckian genetic algorithm with an initial population size of 300 and a termination condition of 10⁵ generations and 9995×10⁵ energy evaluations were carried out. The 50 top docking results were clustered using the DBSCAN algorithm [40]. Three representative poses from each cluster were selected for MD calculations for C4-S (dp6), whereas the further analyzed binding pose for C4-S (dp12) was chosen as it is described in the results section.

Molecular dynamics simulations

The structures of rCatK and hCatK:C4-S complex obtained from the docking calculations were refined by applying MD simulations carried out with AMBER16 and compared with the crystal

structures of hCatK with C4-S (pdb ID: 3C9E and 4N8W). Parameters from the ff14SB and GLYCAM-06 force fields were used for protein and GAGs, respectively. The partial ligand (GAG) charges used in the simulations were the same ones as in the docking calculations for consistency and were compatible with the applied force fields: they were either included already in the default libraries from GLYCAM06 or obtained in our previous study where the charges for the sulfated groups calculated by Huige and Altona [39] were introduced to the GAG units [41]. The complexes were solvated in a TIP3P octahedral periodic box with a minimal distance to the periodic box border of 6Å, and counter ions were used to neutralize the system. Two energy-minimization steps were carried out: first 0.5×10^3 steepest descent cycles and 10^3 conjugate gradient cycles with harmonic force restraints on solute atoms, and then 3×10^3 steepest descent cycles and 3×10^3 conjugate gradient cycles without constraints. Afterwards, the system was heated up to 300K for 10ps, equilibrated for 50ps at 300K and 10^6 Pa in isothermal isobaric ensemble (NPT). Finally, a 20ns of productive MD run was performed in an NTP ensemble. The SHAKE algorithm, 2fs time integration step, 8Å cutoff for non-bonded interactions and the Particle Mesh Ewald method were used.

Rat and human cathepsin K dimer/C4-S complexes simulations

Based on the experimental structure of hCatK with C4-S (pdb ID: 4N8W) we modeled the structures of hCatK and rCatK dimers with C4-S being bound in the same binding site using Chimera [42]. MD protocols for these complexes as well as for unbound CatK dimers were the same as described above. Ten independent runs for each rCatK dimer/C4-S and hCatK dimer/C4-S complexes were performed. Energetic post-processing of the trajectories and per residue energy decomposition were done in a continuous solvent model using MM-GBSA with $igb=2$ of AMBER16. Only frames of MD were taken for this analysis, where the complex was stable in terms of RMSD convergence and no dissociation of the ligand was observed.

Results and discussion

S2 substrate specificity of rat CatK and human CatK: a comparative analysis

Primary structure of mature rCatK exhibits a high degree of identity (88%) and similarity (93%) with hCatK, displaying 25 non-conserved residues over the 215-length amino acid chain (**Figure 1A**). The predicted 3D structure of rCatK obtained by homology modeling was superimposed with the structure of hCatK (RMSD= 0.2Å) (**Fig. 1B**). The catalytic triad (Cys25, His159, and Asn175) is conserved and located in a cleft between the L domain and the R domain. However, differences exist within the S2 pocket, which contributes substantially to the substrate specificity for papain-related cysteine cathepsins. The S2 subsite is defined by residues Tyr67, Met68, Ser134, Val160, Ala163 and Leu209. Ser134 and Val160 that are forming a side part of the S2 pocket of rCatK are replaced by Ala and Leu residues in hCatK, respectively (numbering of the mature protease used throughout this article). In particular, the hydroxyl group in the side chain of Ser134 may affect the binding of substrates by favoring hydrogen bonds with residues that occupy position P2. As previously reported, the S2 subsite pocket of hCatK exhibits a strong preference for aliphatic residues over aromatic amino acids (Leu>>Phe) and possesses an exclusive preference for Pro at P2 among mammalian cathepsins [6,23,33,43]. Therefore, we evaluated the S2 subsite specificity of rCatK using Z-Leu-Arg-AMC, Z-Phe-Arg-AMC, and Z-Gly-Pro-Arg-AMC as substrates and compared with that of hCatK. The kinetic parameters for the hydrolysis of these substrates are given in **Table 1**. rCatK displayed a stronger preference (20-fold increase of k_{cat}/K_m) for Leu over Phe compared to its orthologs (3-fold) (**Figure 2A**). Pro is also suitably accepted in the S2 pocket of rCatK with a K_m value identical to that of hCatK. Additionally, docking stimulations were performed within the S2 subsite of each enzyme (**Fig. 2B**). The binding energies of rCatK S2 pocket showed a stronger affinity for Leu at P2 ($\Delta G = -31.6 \pm 3.2$ kcal/mol) over Phe ($\Delta G = -23.6 \pm 3.6$ kcal/mol) and Pro ($\Delta G = -25.2 \pm 4.2$ kcal/mol). Nevertheless, the specificity constants (k_{cat}/K_m) of rCatK towards Z-Gly-Pro-Arg-AMC, Z-Leu-Arg-AMC and Z-Phe-Arg-AMC were, respectively, ~32-, ~34- and ~169-fold lower than that of hCatK, supporting that rCatK displays a weaker peptidase activity than hCatK in agreement with [21]. We expanded our analysis with two selective quenched fluorogenic substrates, Abz-HPGGPQ-(3-NO₂-Tyr) and Abz-RPPGFSPFR-(3-NO₂-Tyr) of hCatK, which encompass a Pro residue at P2 and span the entire active site of hCatK [23,24]. rCatK cleaved more efficiently the

bradykinin-mimicking substrate Abz-RPPGFSPFR-(3-NO₂-Tyr) than Abz-HPGGPQ-(3-NO₂-Tyr) ($k_{cat}/K_m=115,148M^{-1}.s^{-1}$ vs $k_{cat}/K_m=9,687M^{-1}.s^{-1}$). Comparably to AMC-derived substrates, specificity constants were dramatically lowered for rCatK compared to hCatK, mainly due to a ~5 to 360-fold reduction in k_{cat} values (Abz-RPPGFSPFR-(3-NO₂-Tyr): $k_{cat}=0.2s^{-1}$ vs $k_{cat}=1.1s^{-1}$; Abz-HPGGPQ-(3-NO₂-Tyr): $k_{cat}=0.005s^{-1}$ vs $k_{cat}=1.8s^{-1}$). Since the kinetic data revealed critical differences between the S2 subsite of rCatK and hCatK, the binding free energy of both Ser134 and Val160 was estimated using the MM/GBSA per residue decomposition method to gain information on their potential interactions with Leu, Phe, or Pro residues at P2. Detailed results are listed in **Table 2**. Overall, Ser134 and Val160 residues in rCatK contribute in a lesser extent to the binding of Leu, Phe, or Pro than their S2 subsite counterparts (Ala134 and Leu160) in hCatK. Accordingly, site-directed mutagenesis studies, using Ser134Ala and Val160Leu rCatK mutants, supported that the weaker activity of rCatK compared to hCatK partly depends on these two S2 residues [20]. These results suggest that rCatK S2 subsite specificity is more restricted than that of hCatK. A similar finding was obtained for mouse CatK S2 subsite, which differs from a single residue (Met209) with rCatK (Leu209) [22]. Additionally, the pH activity profile of both enzymes displayed a similar bell-shaped curve with an optimal activity at pH 6.4 (pK values for rCatK: $pK_1=3.9\pm 0.1$ and $pK_2=8.6\pm 0.1$; pK values for hCatK: $pK_1=4.07\pm 0.1$ and $pK_2=7.8\pm 0.1$) (**Fig. 2C**), supporting that the overall structural integrity of the catalytic site is somewhat well preserved between rat CatK and human CatK.

Inhibition of rat and human CatK by T-kininogen

T-kininogen (TK, a.k.a. thiostatin), which is an additional kininogen only found in rat (for review: [18]), is both a serum acute phase reactant and a tight-binding inhibitor of rat cathepsins B, H and L [26,44]. Here we established that TK is also a potent inhibitor of rCatK ($K_i=2.8\pm 0.8nM$). Importantly, TK inhibits also hCatK with a ~25-fold lower K_i value ($K_i=0.11\pm 0.05nM$). Two days after subcutaneous turpentine injection, a circa 3-fold increase of immunoreactive TK was observed by immunoblotting in IRS compared to NRS used as control (data not shown). Alternatively, dosage of TK in inflamed serum was achieved by rocket immunoelectrophoresis (RIEP) (**Figure 3A**), using NRS for calibration (TK concentration: $0.5\pm 0.05\mu g/\mu l$, [25]). Although this immunochemical technique is not very sensitive, RIEP is

very robust and reproducible [30], making possible to overcome the lack of other available quantitative immunoassays such as ELISA. Here we found that the concentration of immunoreactive TK in IRS samples from turpentine-treated rats (Wistar strain) was $\sim 1.7\mu\text{g}/\mu\text{l}$, thus confirming the ~ 3 -fold increase of TK compared to NRS [26]. It should be noted that substantially variable assay results were reported, TK levels in inflamed sera (from 2- to 20-fold TK increase compared to NRS) depending on rat strains, experimental protocols of inflammation and most probably also depending on age and sex [31,45,46]. Thereafter, hCatK-inhibiting properties of rat serum were evaluated. A dose-dependent decrease of hCatK activity was observed upon addition of NRS, while IRS led to a ~ 3 -fold greater inhibition of hCatK, consistently with the ~ 3 -fold rise of immunoreactive TK (**Fig. 3B**). Taken together the unique presence of TK in rat, its tight-binding (nM range) inhibitory properties and its increased level in IRS after an inflammatory challenge, present data raise the accuracy issue of the therapeutic evaluation of pharmacological inhibitors of hCatK using rat as an experimental animal model.

Structural analysis of rat CatK:C4-S and human CatK:C4-S interactions

Comparison of C4-S mediated collagenolytic activity between rat and human CatK

While collagenases of the matrix metalloprotease family (MMP-1, -2, -8, -13, and -14, [47]) cleave triple helical collagen (type I and II) at a single site toward the C-terminal end, resulting in the release of $\frac{3}{4}$ N-terminal and $\frac{1}{4}$ C-terminal fragments, hCatK is a unique and potent collagenase that cleaves at various site in the helical region of both type I and II collagens [6]. This cleavage event takes place at the site of bone resorption, within the so-called resorption lacuna, in which acidic pH conditions prevail. Its collagenolytic activity requires specifically the presence of the cartilage-resident glycosaminoglycan C4-S, a major component of extracellular matrix with which it forms an active complex. Indeed, monomeric hCatK has no significant collagenase activity [8]. To determine whether C4-S may enhance the ability of rCatK to cleave interstitial collagens, soluble rat type I and II triple helical collagens were incubated with rCatK in the absence and presence of C4-S (~ 20 - 30kDa), and its activity was compared with those of hCatK at equal enzyme concentration (**Figure 4**). After 8h of incubation, in the presence or absence of C4-S (0.15%), rCatK activity was limited to a cleavage in the nonhelical telopeptide regions leading to the disappearance of the high-molecular-weight β and γ bands of type I collagen and the accumulation of $\alpha 1/2$ chains (**Fig. 4A**). In contrast, hCatK degraded efficiently

rat triple-helical type I collagen only in the presence of C4-S. Moreover, C4-S has a negligible effect on rCatK collagenase activity against type II rat collagen, contrary to hCatK (**Fig. 4B**). The degradation of both type I and II rat collagen was blocked by E-64. Other studies reported that higher concentrations of hCatK ($\geq 125\text{nM}$) are required *in vitro* to degrade efficiently both soluble calf skin type I collagen and soluble calf articular joints type II collagen with C4-S (0.15%) [7,48–50]. A plausible explanation of this apparently discrepant result is that the ultrastructural organization of fibrillar collagens may depend both on the animal species considered and on the tissue source. In a convergent way with present results, rCatK displayed comparable kinetics constants for the hydrolysis of the peptidyl substrate Z-Leu-Arg-AMC in the presence of C4-S ($k_{\text{cat}}/K_{\text{m}}=183,000\text{M}^{-1}\cdot\text{s}^{-1}$, $K_{\text{m}}=21.7\pm 4.3$, $k_{\text{cat}}=3.9\pm 0.7\text{s}^{-1}$) and in the absence of C4-S ($k_{\text{cat}}/K_{\text{m}}=144,569\text{M}^{-1}\cdot\text{s}^{-1}$; $K_{\text{m}}=13.3\pm 2.5\mu\text{M}$; $k_{\text{cat}}=1.9\pm 0.4\text{s}^{-1}$).

Next, in order to delineate differences in collagenolytic activity, we evaluated the conservation of hCatK residues that are involved in the binding of hexameric C4-S (dp6) in rCatK, using two hCatK-C4-S complex X-ray structures [51] (pdb ID: 3C9E) [52] (pdb ID: 4N8W). A previous study demonstrated that mutation of Ile171, which is substituted by Thr171 in rCatK (Fig. 1), contributes in combination with other residues (Lys9, Lys77, Arg79, Lys103, Lys106, Arg108, Arg111, Lys122, Arg127, and Gln172) in the loss of the collagenase activity of hCatK, due to its close vicinity to a C4-S sulfate group [49]. Moreover, five substitutions were observed between rCatK and hCatK: i.e. Ile5 vs Val5, Gln77 vs Lys77, Gly79 vs Arg79, Tyr177 vs His177, and Val194 vs Ile194, respectively. These residues are located at the boundary of GAG-binding sites, but do not interact directly with C4-S as revealed by the two structures of hCatK-C4-S complexes (3C9E and 4N8W). The basic side chains of Arg79 and to a lesser extent His177 in hCatK are more likely favorable to bind to the highly acid sulphates groups of C4-S, via stable hydrogen bonds and electrostatic interactions, than aliphatic Gly79 and aromatic Tyr177 in rCatK. In the aforementioned “beads-on-a-strain”-like organization (pdb ID: 3C9E), a single C4-S (dp6) binds to GAG-binding residues (Asp6, Arg8, Lys9, Lys10, Ile171, Gln172, Asn190, Lys191, and Leu195) via hydrogen-bonded ion pairs in the groove of hCatK surface, located in the R-domain, opposite to that of the active site. Alternatively, two distinct long-stretched GAG binding sites, which include basic residues Lys119, Lys122, Arg123, and Arg127 for the first one and Lys40, Lys41, Arg108, Arg111, Arg127, and Lys214 for the second binding site were identified (pdb ID: 4N8W). Beside this second binding site, an additional or alternative GAG binding site (i.e.

Lys10, Lys39, and Lys77) was pinpointed by *in silico* modeling [53]. According to these data, presence of both Gln77 and Thr171 (instead of Lys77 and Ile171 for hCatK), in concert with other dissimilar residues, may support the lower binding of C4-S to rCatK.

Likewise, analysis of PBSA electrostatic potential of rCatK model and hCatK X-ray structure (pdb ID: 4N8W) depicted a noticeable difference in the vicinity of position 5 (**Fig. 5A**). Indeed, a positively charged electrostatic patch was found on rCatK surface conversely to hCatK that displayed a negative electrostatic patch. This difference suggests that rCatK may interact preferentially with C4-S (dp6) in the binding pose reported by [51].

Furthermore, molecular docking studies were conducted to scrutinize potential alternative binding poses of C4-S (dp6) with rCatK and to relate to the two X-ray structures of hCatK-C4-S complexes (i.e. pdb ID: 3C9E and pdb ID: 4N8W). Although, three clusters of GAG binding sites were comparable for both proteases, they displayed different free binding energies (ΔG) (**Fig. 5B**, **Table 3**). Among the top 50 binding poses obtained, also experimentally observed GAG binding sites, which predominantly correspond to the binding site from 4N8W structure, were identified. The next docking step was completed using a longer linear chain of C4-S (dp12) in order to better mimic *in vitro* collagen-degrading experiments in the presence of ~20-30kDa C4-S. The dodecameric saccharide (dp12) model bound to hCatK experimental structure in two binding sites, either in the binding site corresponding to 3C9E or 4N8W structure (**Fig. 5C**). Although similar results were found for rCatK model, supplementary rCatK/C4-S (dp12) complexes were predicted at the interface of both GAG binding sites connecting them. Finally, comparative MD-based analysis of the modeled structure with both hCatK:C4-S complexes (3C9E and 4N8W), which allowed us to identify additional hydrogen bonds between C4-S (dp12) and 6 residues of rCatK (e.g. Arg8, Tyr12, Lys40, Gln172, Asn190, and Tyr193), also reflected by the computed binding free energies (**Fig. 5D**).

Molecular dynamics analysis and free energy calculations

To gain further insights into the deficiency of C4-S to enhance collagenolytic activity of rCatK, MD simulations were carried out for nine monomeric rCatK and hCatK:C4-S (dp6) complexes previously obtained by docking experiments. From each cluster of structures obtained from molecular docking (**Fig. 5B**), we chose three random structures to increase the sampling within a

cluster. Differences of binding free energy (ΔG) of the C4-S binding pose were found for both enzymes, the 4N8W structure being more favorable than the 3C9E structure (**Table 3**). Interestingly, MD simulations (20 ns) performed on monomeric rCatK in complex with C4-S (dp12) resulted in matching C4-S binding poses to both 3C9E and 4N8W binding sites. The corresponding computed free energy value (-56.7 ± 16.1 kcal/mol) was among the lowest energy calculated for all rCatK:C4-S complexes.

Besides monomeric forms of rCatK and hCatK, analysis of hCatK:C4-S complex (4N8W) revealed a dimeric organization of hCatK held together by protein-protein and protein-C4-S interactions [52]. Based on this structure, we built a model of rCatK dimer (in complex with C4-S (dp6)) that was further analyzed by MD and compared to hCatK dimer. Again, we did not observe ligand dissociation during MD simulations, providing evidences that complexes were stable. The free energy analysis results also showed that the free energy values of rCatK dimer complexes with C4-S (**Table 4**) were in most case more favorable in comparison to the free energy values obtained for rCatK and hCatK:C4-S complexes from the experimental structures (**Table 3**), furthermore suggesting the stability of rCatK dimer complexes with C4-S calculated by applying the same methodology. At the same time, we are aware that the long MD simulations may provide significantly different results [54] and potentially result in the dissociation of a ligand. MM-GBSA free energy analysis showed that - in eight out of ten repeated MD simulations - complex formation of hCatK dimer with C4-S was more stable than rCatK dimer with C4-S (**Table 4**). Also, C4-S interacted preferentially with the site corresponding to the 4N8W structure.

Conclusion

Despite rCatK and hCatK share a close substrate specificity as well as a high degree of identity (88%) and similarity (93%), some distinctive features, which are mostly related to non-conserved residues, were disclosed in terms of peptidase activity, S2 subsite specificity and C4-S-mediated collagen degradation. Based on our present *in silico* data, we propose a hypothetical molecular model, which could sustain the difference between the mechanisms of collagen degradation by rCatK and hCatK (**Figure 6**). In this model, C4-S first binds to both to rCatK and hCatK by its short hexameric fragment, which according to the experimental structural data represents a

sufficient binding unit, in the binding site corresponding to the crystal structure from pdb ID: 4N8W. Upon the CatK dimer formation stimulated by C4-S binding, depending on the protein, the long chains of C4-S could either further combine with the CatK dimer in a binding pose as found in pdb ID:4N8W for hCatK or as in pdb ID:3C9E for rCatK, respectively, which may correspond to different levels of collagenolytic activities of these two proteases. Otherwise, following an inflammatory challenge, the increased level of serum TK, a unique tight-binding inhibitor of cysteine cathepsins, raised the question of the appropriateness to evaluate pharmacological hCatK inhibitors in rat. Also, present data support limitations of this rodent model to delineate molecular mechanisms underlying C4-S-mediated collagenolytic activity of CatK during inflammatory diseases like rheumatoid arthritis or osteoporosis. Besides the absence of T-kininogen, humanized CatK mouse models would be a beneficial alternative in order to assess the potential efficiency of hCatK inhibitors.

Acknowledgements

Authors acknowledge Prof. Dieter Brömme (Department of Oral Biological and Medical Sciences, The University of British Columbia, Vancouver, British Columbia, Canada) for providing human recombinant cathepsin K. This work was supported by institutional fundings from the Institut National de la Santé et de la Recherche Médicale (INSERM), the University of Tours and GagoSciences (Structure, function and regulation of glycosaminoglycans; GDR 3739, Centre National de la Recherche Scientifique). We acknowledge the RTR (Thematic Research Network) MotivHealth (Région Centre-Val de Loire, France) for support. TC holds a doctoral fellowship from MESRI (Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, France). KKB holds a BMN (Badania Młodych Naukowców) grant from the Faculty of Chemistry, University of Gdańsk (BMN-538-8370-B249-18) and a grant (UMO-2018/31/N/ST4/01677) from the National Science Center of Poland (Narodowe Centrum Nauki). SS received funding (grant UMO-2018/30/E/ST4/00037) from the National Science Center of Poland (Narodowe Centrum Nauki). A friendly and warmly grateful thought to Dr Francis Gauthier, Professor Emeritus at the University of Tours.

Author Contributions

Conceived and designed the experiments: FL, KKB, SS, and GL. Performed the experiments: FL, TC, JR, AS, KKB, VGNVP, and GL. Analyzed the data: FL, TC, KKB, SS, and GL. Contributed reagents/materials/analysis tools: PSNDDB. Prepared the tables and figures: FL, KKB, and TC. Wrote the paper: FL and GL. All authors read and approved the final version of the manuscript.

Competing Interests: The authors declare no competing interest.

Table 1: Kinetic parameters for hydrolysis of peptidyl-AMC substrates by rCatK and hCatK.

The selectivity factor corresponds to the ratio of the second order rate constants (k_{cat}/K_m) of hCatK vs rCatK. The k_{cat}/K_m values are reported as the average of three independent experiments.

	rCatK			hCatK			Selectivity factor
	K_m (μM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (M ⁻¹ .s ⁻¹)	K_m (μM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (M ⁻¹ .s ⁻¹)	
Z-LR-AMC	13.3±2.5	1.9±0.4	144 569	2.6±0.3	12.7±0.1	4 900 000	33.9
Z-FR-AMC	19.2±1.2	0.13±0.01	7 823	9.7±1.0	12.8±0.2	1 321 000	168.8
Z-GPR-AMC	32.8±5.0	0.07±0.01	2 073	36±3.0	2.4±0.02	65 319	31.5

Table 2: MM-GBSA free energy decomposition of rat (model) and human CatK (experimental structure) S2 pocket.

P2 residue	rCatK		hCatK	
	Ser134	Val160	Ala134	Leu160
	ΔG (kcal/mol)			
Phe	-0.31	-0.35	-0.96	-0.44
Leu	-0.34	-0.18	-0.31	-0.44
Pro	-0.02	-0.02	-0.30	-0.11

Table 3: Free binding energy in rat (model) and human (experimental structure) CatK monomer/C4-S (dp6) complexes. Mean values and standard deviations are provided.

Run No.	ΔG_{rCatK} (kcal/mol)	ΔG_{hCatK} (kcal/mol)
cluster 1		
1	-43.2±8.8	-48.8±7.8
2	-49.9±8.1	-58.6±11.4
3	-52.2±16.5	-68.7±15.0
cluster 2		
1	-48.5±8.9	-62.5±10.9
2	-52.4±8.0	-39.8±7.6
3	-45.3±11.2	-44.9±9.4
cluster 3		
1	-63.5±13.5	-47.8±13.3
2	-60.1±8.9	-42.4±9.1
3	-52.5±8.5	-38.4±14.7
reference experimental structure		
3C9E	-42.2±7.8	-36.2±9.6
4N8W	-49.4±14.5	-55.1±17.3

Table 4: Free binding energy in human (experimental structure) and rat (model) CatK dimer/C4-S (dp6) complexes. Mean values and standard deviations are provided.

Run No.	ΔG_{hCatK} (kcal/mol)	ΔG_{rCatK} (kcal/mol)	$\Delta G_{hCatK} - \Delta G_{rCatK}$ (kcal/mol)
1	-32.5±11.1	-40.7±10.9	8.2
2	-52.2±10.2	-27.9±9.0	-24.3
3	-60.6±14.3	-44.9±8.8	-15.7
4	-49.1±9.2	-44.8±9.6	-4.3
5	-64.6±14.3	-61.1±11.6	-3.5
6	-46.7 ± 7.8	-59.5±17.5	12.8
7	-64.2±16.2	-44.0±9.0	-20.2
8	-51.6±11.2	-41.2±7.9	-10.4
9	-52.7±12.6	-35.2±7.9	-17.5
10	-53.4±10.5	-50.8±10.8	-2.6

Legend to Figures

Figure 1: Structural comparison between rat and human CatK.

A) Primary sequence alignment of rCatK (*Rattus norvegicus*, UniProtKB accession: 035186) and hCatK (*Homo sapiens*, UniProtKB accession: P43235). The nonconserved residues in both proteins are highlighted in grey. The residues of the catalytic triad (Cys25, His159 and Asn175) are marked with an asterisk. The residues of the S3, S2, S1 and S1' subsites are indicated. B) Superposition of rCatK (homology model, blue) and hCatK (pdb ID: 4N8W, grey) structures (cartoon representation). Cys25 is in yellow, and the nonconserved residues of rCatK are shown as stick representation.

Figure 2: S2 specificity and pH profile comparison between rat and human CatK.

A) k_{cat}/K_m values for the hydrolysis of Z-Leu-Arg-AMC, Z-Phe-Arg-AMC and Z-Gly-Pro-Arg-AMC by rCatK (black bar) and hCatK (grey bar) (normalized to the best substrate: Z-Leu-Arg-AMC). B) 3D representation of the S2 pocket of rCatK (model) and hCatK (pdb ID: 1MEM) with Z-Leu-Arg-AMC (panels 1, 2), Z-Phe-Arg-AMC (panels 3, 4) and Z-Gly-Pro-Arg-AMC (panels 5, 6). C) pH activity profile for rCatK and hCatK. Residues of the S2 pocket are indicated. Relative k_{cat}/K_m values for the hydrolysis of Z-Leu-Arg-AMC were plotted against the pH values.

Figure 3: Dosage of TK and inhibition of human CatK by rat serum.

A) TK immunoassay. The concentration of TK in the serum of inflamed turpentine-treated rats (IRS) was determined by rocket immunoelectrophoresis (RIEP, a.k.a. Laurell method) as in-depth described in the experimental section. Lanes a-e: IRS dilutions (1/16-1/512); Lanes f-k: non-inflamed rat serum (NRS) dilutions (1/2-1/64). According to NRS contains $0.5 \pm 0.05 \mu\text{g}/\mu\text{l}$ of TK [26],[31], NRS dilutions (lanes f-k) correspond to 0.25-0.008 $\mu\text{g}/\mu\text{l}$ of TK. The concentration of TK in IRS was deduced from the calibration graph [30]. B) Inhibition of hCatK activity by rat serum. hCatK (0.5nM) was incubated with increasing doses of non-inflamed (NRS, grey) and inflamed (IRS, black) rat serum proteins in sodium acetate buffer 0.1M pH5.5, DTT 2mM. After 15 minutes, residual peptidase activity of hCatK was measured using Z-Phe-Arg-AMC (20 μM) (n=3).

Figure 4: Comparison of collagenolytic activity between rat and human CatK in the absence or presence of C4-S.

A) Soluble type I or type II (B) rat collagen (0.5mg/mL) was incubated in the absence or presence of C4-S (0.15%, m/v) at 28°C with rCatK and hCatK (20nM), respectively for 0-8h. Representative gels of the collagenolytic activities of rat and human CatK are shown. Control experiments were performed with rCatK and hCatK preincubated with E-64 (10µM) before adding collagen. Samples were analyzed by SDS-PAGE and subsequently stained by Coomassie blue. Residual activity (%) of each enzyme in the absence or in the presence of 0.15% C4-S was monitored spectrofluorometrically using the fluorogenic substrate, Z-Leu-Arg-AMC (10µM) at different incubation times. The data represent the mean±SD values of residual activity (%) from independent experiments (n=3) performed in triplicate.

Figure 5: Comparison between rat and human CatK:C4-S interactions by docking and MD simulations.

A) Electrostatic potential isosurfaces for rCatK (model) and hCatK (pdb ID: 4N8W) in surface representation (the isovalues are: -3kcal/mole in red, 3kcal/mole, in blue, respectively). The C4-S (dp6) binding poses of X-ray structures are in sticks: orange and cyan correspond to pdb ID:3C9E and pdb ID:4N8W, respectively. B) Structural clusters of C4-S (dp6) obtained by molecular docking (Autodock 3) for rCatK and hCatK receptors. Proteins are shown in white surface and clustered C4-S structures are in blue, red and green sticks, respectively. C) Top 50 structures of C4-S (dp12, black sticks) docked into rCatK (model) and hCatK (pdb ID: 4N8W) (proteins are in white surface). D) Clusters of rCatK:C4-S (dp12) interactions are shown in blue, red and green sticks, respectively. The structural pose obtained by docking that connects the structures corresponding to pdb ID: 3C9E (orange) and pdb ID: 4N8W (cyan) structures is shown in green. Hydrogen bonds formed between C4-S (dp12) and rCatK residues are depicted in dashed lines in the selected pose.

Figure 6: Proposed mechanism of the collagen degradation by rat CatK in the presence of C4-S.

1) C4-S (dp6) binds preferentially both rCatK and hCatK in the binding site corresponding to pdb ID: 4N8W structure, 2) rCatK and hCatK dimer formation is mediated by the prebound C4-S,

and 3) C4-S elongation leads to diverse structures for rCatK and hCatK:GAG complexes, which could explain the differences in their collagenolytic activity.

Bibliography

- [1] D. Brömme, K. Okamoto, Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution, *Biol. Chem. Hoppe. Seyler.* 376 (1995) 379–384.
- [2] D. Brömme, K. Okamoto, B.B. Wang, S. Biroc, Human Cathepsin O2, a Matrix Protein-degrading Cysteine Protease Expressed in Osteoclasts: Functionnal expression of Human Cathepsin O2 in *Spodoptera Frugiperda* and Characterization of the Enzyme, *J. Biol. Chem.* 271 (1996) 2126–2132.
- [3] F. Lecaille, D. Brömme, G. Lalmanach, Biochemical properties and regulation of cathepsin K activity, *Biochimie.* 90 (2008) 208–226.
- [4] V. Turk, V. Stoka, O. Vasiljeva, M. Renko, T. Sun, B. Turk, D. Turk, Cysteine cathepsins: From structure, function and regulation to new frontiers, *Biochim. Biophys. Acta BBA - Proteins Proteomics.* 1824 (2012) 68–88.
- [5] W. Kafienah, D. Brömme, D.J. Buttle, L.J. Croucher, A.P. Hollander, Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix, *Biochem. J.* 331 (1998) 727–732.
- [6] P. Garnero, O. Borel, I. Byrjalsen, M. Ferreras, F.H. Drake, M.S. McQueney, N.T. Foged, P.D. Delmas, J.-M. Delaissé, The Collagenolytic Activity of Cathepsin K Is Unique among Mammalian Proteinases, *J. Biol. Chem.* 273 (1998) 32347–32352.
- [7] Z. Li, W.-S. Hou, D. Brömme, Collagenolytic Activity of Cathepsin K Is Specifically Modulated by Cartilage-Resident Chondroitin Sulfates, *Biochemistry.* 39 (2000) 529–536.
- [8] Z. Li, W.-S. Hou, C.R. Escalante-Torres, B.D. Gelb, D. Brömme, Collagenase Activity of Cathepsin K Depends on Complex Formation with Chondroitin Sulfate, *J. Biol. Chem.* 277 (2002) 28669–28676.
- [9] Z. Li, Y. Yasuda, W. Li, M. Bogyo, N. Katz, R.E. Gordon, G.B. Fields, D. Brömme, Regulation of Collagenase Activities of Human Cathepsins by Glycosaminoglycans, *J. Biol. Chem.* 279 (2004) 5470–5479.
- [10] D. Brömme, F. Lecaille, Cathepsin K inhibitors for osteoporosis and potential off-target effects, *Expert Opin. Investig. Drugs.* 18 (2009) 585–600.
- [11] M. Vizovišek, M. Fonović, B. Turk, Cysteine cathepsins in extracellular matrix remodeling: Extracellular matrix degradation and beyond, *Matrix Biol.* 75–76 (2019) 141–159.
- [12] S. Khosla, L.C. Hofbauer, Osteoporosis treatment: recent developments and ongoing challenges, *Lancet Diabetes Endocrinol.* 5 (2017) 898–907.
- [13] J. Selent, J. Kaleta, Z. Li, G. Lalmanach, D. Brömme, Selective Inhibition of the Collagenase Activity of Cathepsin K, *J. Biol. Chem.* 282 (2007) 16492–16501.
- [14] P. Panwar, L. Xue, K. Sjøe, K. Srivastava, S. Law, J.-M. Delaisse, D. Brömme, An Ectosteric Inhibitor of Cathepsin K Inhibits Bone Resorption in Ovariectomized Mice, *J. Bone Miner. Res.* 32 (2017) 2415–2430.
- [15] P. Panwar, S. Law, A. Jamroz, P. Azizi, D. Zhang, M. Ciufolini, D. Brömme, Tanshinones that selectively block the collagenase activity of cathepsin K provide a novel class of ectosteric antiresorptive agents for bone: Ectosteric inhibitors of cathepsin K, *Br. J. Pharmacol.* 175 (2018) 902–923.

- [16] S. Law, X. Du, P. Panwar, N.S. Honson, T. Pfeifer, M. Roberge, D. Brömme, Identification of substrate-specific inhibitors of cathepsin K through high-throughput screening, *Biochem. J.* 476 (2019) 499–512.
- [17] F. Gauthier, N. Gutman, T. Moreau, A. el Moujahed, Possible relationship between the restricted biological function of rat T kininogen (thiostatin) and its behaviour as an acute phase reactant, *Biol. Chem. Hoppe. Seyler.* 369 Suppl (1988) 251–255.
- [18] G. Lalmanach, C. Naudin, F. Lecaille, H. Fritz, Kininogens: More than cysteine protease inhibitors and kinin precursors, *Biochimie.* 92 (2010) 1568–1579.
- [19] A. Barlas, H. Okamoto, L.M. Greenbaum, T-kininogen - the major plasma kininogen in rat adjuvant arthritis, *Biochem. Biophys. Res. Commun.* 129 (1985) 280–286.
- [20] S. Tada, K. Tsutsumi, H. Ishihara, K. Suzuki, K. Gohda, N. Teno, Species Differences Between Human and Rat in the Substrate Specificity of Cathepsin K, *J. Biochem. (Tokyo).* 144 (2008) 499–506.
- [21] R.W. Marquis, Y. Ru, S.M. LoCastro, J. Zeng, D.S. Yamashita, H.J. Oh, K.F. Erhard, L.D. Davis, T.A. Tomaszek, D. Tew, K. Salyers, J. Proksch, K. Ward, B. Smith, M. Levy, M.D. Cummings, R.C. Haltiwanger, G. Trescher, B. Wang, M.E. Hemling, C.J. Quinn, H.Y. Cheng, F. Lin, W.W. Smith, C.A. Janson, B. Zhao, M.S. McQueney, K. D'Alessio, C.P. Lee, A. Marzulli, R.A. Dodds, S. Blake, S.M. Hwang, I.E. James, C.J. Gress, B.R. Bradley, M.W. Lark, M. Gowen, D.F. Veber, Azepanone-based inhibitors of human and rat cathepsin K, *J. Med. Chem.* 44 (2001) 1380–1395.
- [22] S. Law, P.-M. Andrault, A.H. Aguda, N.T. Nguyen, N. Kruglyak, G.D. Brayer, D. Brömme, Identification of mouse cathepsin K structural elements that regulate the potency of odanacatib, *Biochem. J.* 474 (2017) 851–864.
- [23] F. Lecaille, E. Weidauer, M.A. Juliano, D. Brömme, G. Lalmanach, Probing cathepsin K activity with a selective substrate spanning its active site, *Biochem. J.* 375 (2003) 307–312.
- [24] E. Godat, F. Lecaille, C. Desmazes, S. Duchêne, E. Weidauer, P. Saftig, D. Brömme, C. Vandier, G. Lalmanach, Cathepsin K: a cysteine protease with unique kinin-degrading properties, *Biochem. J.* 383 (2004) 501–506.
- [25] V.G.N.V. Prasad, Ch. Vivek, P. Anand Kumar, P. Ravi Kumar, G.S. Rao, Turpentine oil induced inflammation decreases absorption and increases distribution of phenacetin without altering its elimination process in rats, *Eur. J. Drug Metab. Pharmacokinet.* 40 (2015) 23–28.
- [26] F. Esnard, F. Gauthier, Rat alpha 1-cysteine proteinase inhibitor. An acute phase reactant identical with alpha 1 acute phase globulin, *J. Biol. Chem.* 258 (1983) 12443–12447.
- [27] C.J. Linnevers, M.E. Mcgrath, A. Armstrong, F.R. Mistry, M.G. Barnes, J.L. Klaus, J.T. Palmer, B.A. Katz, D. Brömme, Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex, *Protein Sci.* 6 (1997) 919–921.
- [28] A.J. Barrett, A.A. Kembhavi, M.A. Brown, H. Kirschke, C.G. Knight, M. Tamai, K. Hanada, L- *trans* -Epoxy succinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L, *Biochem. J.* 201 (1982) 189–198.
- [29] G. Lalmanach, A. Adam, T. Moreau, N. Gutman, F. Gauthier, Discrimination between rat thiostatin (T-kininogen) and one of its cystatin-like inhibitory fragments by a monoclonal antibody, and localization of the epitope, *Eur. J. Biochem.* 196 (1991) 73–78.
- [30] C.B. Laurell, Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies, *Anal. Biochem.* 15 (1966) 45–52.
- [31] J. Urban, D. Chan, G. Schreiber, A rat serum glycoprotein whose synthesis rate increases greatly during inflammation, *J. Biol. Chem.* 254 (1979) 10565–10568.

- [32] C. Naudin, F. Lecaille, S. Chowdhury, J.C. Krupa, E. Purisima, J.S. Mort, G. Lalmanach, The Occluding Loop of Cathepsin B Prevents Its Effective Inhibition by Human Kininogens, *J. Mol. Biol.* 400 (2010) 1022–1035.
- [33] F. Lecaille, Y. Choe, W. Brandt, Z. Li, C.S. Craik, D. Brömme, Selective Inhibition of the Collagenolytic Activity of Human Cathepsin K by Altering Its S2 Subsite Specificity †, *Biochemistry.* 41 (2002) 8447–8454.
- [34] W. Li, A. Cowley, M. Uludag, T. Gur, H. McWilliam, S. Squizzato, Y.M. Park, N. Buso, R. Lopez, The EMBL-EBI bioinformatics web and programmatic tools framework, *Nucleic Acids Res.* 43 (2015) W580–W584.
- [35] J. Sage, F. Mallèvre, F. Barbarin-Costes, S.A. Samsonov, J.-P. Gehrcke, M.T. Pisabarro, E. Perrier, S. Schnebert, A. Roget, T. Livache, C. Nizard, G. Lalmanach, F. Lecaille, Binding of Chondroitin 4-Sulfate to Cathepsin S Regulates Its Enzymatic Activity, *Biochemistry.* 52 (2013) 6487–6498.
- [36] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.* 19 (1998) 1639–1662.
- [37] D.A. Case, D.S. Cerutti, T.E. Cheatham, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York, P.A. Kollman, Amber 2017, University of California, San Francisco, Univ. Calif. San Franc. (2017).
- [38] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. González-Outeiriño, C.R. Daniels, B.L. Foley, R.J. Woods, GLYCAM06: A generalizable biomolecular force field. *Carbohydrates, J. Comput. Chem.* 29 (2008) 622–655.
- [39] C.J.M. Huige, C. Altona, Force field parameters for sulfates and sulfamates based on ab initio calculations: Extensions of AMBER and CHARMM fields, *J. Comput. Chem.* 16 (1995) 56–79.
- [40] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, in: *Proc 2nd Int. Conf. Knowl. Discov.*, 1996: pp. 226–231.
- [41] A. Pichert, S.A. Samsonov, S. Theisgen, L. Thomas, L. Baumann, J. Schiller, A.G. Beck-Sickinger, D. Huster, M.T. Pisabarro, Characterization of the interaction of interleukin-8 with hyaluronan, chondroitin sulfate, dermatan sulfate and their sulfated derivatives by spectroscopy and molecular modeling, *Glycobiology.* 22 (2012) 134–145.
- [42] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera? A visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–1612.
- [43] Y. Choe, F. Leonetti, D.C. Greenbaum, F. Lecaille, M. Bogyo, D. Brömme, J.A. Ellman, C.S. Craik, Substrate Profiling of Cysteine Proteases Using a Combinatorial Peptide Library Identifies Functionally Unique Specificities, *J. Biol. Chem.* 281 (2006) 12824–12832.
- [44] T. Moreau, F. Esnard, N. Gutman, P. Degand, F. Gauthier, Cysteine-proteinase-inhibiting function of T kininogen and of its proteolytic fragments, *Eur. J. Biochem.* 173 (1988) 185–190.
- [45] Y.A.H. Gordon, L.N. Louis, al Acute-Phase Globulins of Rats, *Biochem. J.* (1969) 113, 481.
- [46] R. Walter, D.M. Murasko, F. Sierra, T-Kininogen is a biomarker of senescence in rats, *Mech. Ageing Dev.* 106 (1998) 129–144.

- [47] J.L. Lauer-Fields, D. Juska, G.B. Fields, Matrix metalloproteinases and collagen catabolism, *Biopolymers*. 66 (2002) 19–32.
- [48] W.-S. Hou, D. Brömme, Y. Zhao, E. Mehler, C. Dushey, H. Weinstein, C.S. Miranda, C. Fraga, F. Greig, J. Carey, D.L. Rimoin, R.J. Desnick, B.D. Gelb, Characterization of novel cathepsin K mutations in the pro and mature polypeptide regions causing pycnodysostosis, *J. Clin. Invest.* 103 (1999) 731–738.
- [49] F.S. Nallaseth, F. Lecaille, Z. Li, D. Brömme, The Role of Basic Amino Acid Surface Clusters on the Collagenase Activity of Cathepsin K, *Biochemistry*. 52 (2013) 7742–7752.
- [50] M. Novinec, M. Rebernik, B. Lenarčič, An allosteric site enables fine-tuning of cathepsin K by diverse effectors, *FEBS Lett.* 590 (2016) 4507–4518.
- [51] Z. Li, M. Kienetz, M.M. Cherney, M.N.G. James, D. Brömme, The Crystal and Molecular Structures of a Cathepsin K: Chondroitin Sulfate Complex, *J. Mol. Biol.* 383 (2008) 78–91.
- [52] A.H. Aguda, P. Panwar, X. Du, N.T. Nguyen, G.D. Brayer, D. Brömme, Structural basis of collagen fiber degradation by cathepsin K, *Proc. Natl. Acad. Sci.* 111 (2014) 17474–17479.
- [53] M. Novinec, L. Kovačič, B. Lenarčič, A. Baici, Conformational flexibility and allosteric regulation of cathepsin K, *Biochem. J.* 429 (2010) 379–389.
- [54] K.K. Bojarski, A.K. Sieradzan, S.A. Samsonov, Molecular dynamics insights into protein-glycosaminoglycan systems from microsecond-scale simulations, *Biopolymers*. 110 (2019).

Figure 1

A)



B)

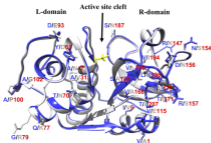


Figure 3

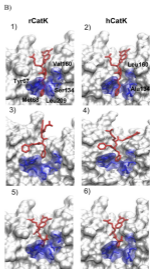
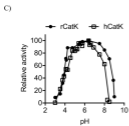
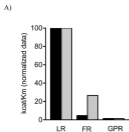


Figure 3

A)



B)

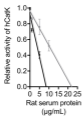


Figure 4

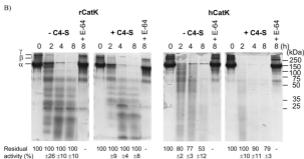
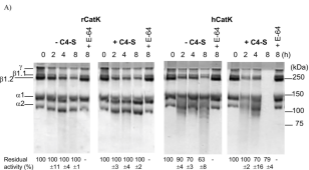


Figure 5

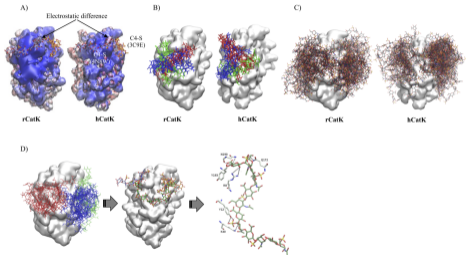


Figure 6

