NRC Publications Archive Archives des publications du CNRC

The occluding loop of Cathepsin B prevents its effective inhibition by human kininogens

Naudin, C.; Lecaille, F.; Chowdhury, S.; Krupa, J. C.; Purisima, E.; Mort, J.S.; Lalmanach, G.

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. / La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

For the publisher's version, please access the DOI link below. / Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

Publisher's version / Version de l'éditeur:

<https://doi.org/10.1016/j.jmb.2010.06.006>

Journal of Molecular Biology, 400, 5, pp. 1022-1035, 2010-06-09

NRC Publications Record / Notice d'Archives des publications de CNRC:

<https://nrc-publications.canada.ca/eng/view/object/?id=103e4b49-5312-4f3f-934e-0cac077ad726>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=103e4b49-5312-4f3f-934e-0cac077ad726>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

The Occluding Loop of Cathepsin B Prevents Its Effective Inhibition by Human Kininogens

C. Naudin¹†, F. Lecaille¹†, S. Chowdhury², J. C. Krupa³, E. Purisima⁴, J. S. Mort³ and G. Lalmanach¹*

¹Inserm U618, Protéases et Vectorisation Pulmonaires, Université François Rabelais, Tours, France

²Lady Davis Institute of Medical Research, McGill University Health Centre, Montreal, Quebec, Canada

³Shriners Hospital for Children and Department of Surgery, McGill University, Montreal, Quebec, Canada

⁴National Research Council Canada, Biotechnology Research Institute, Computational Chemistry, Montreal, Quebec, Canada

Received 15 March 2010;
received in revised form

26 May 2010;

accepted 1 June 2010

Available online

9 June 2010

Edited by R. Huber

Kininogens, the major plasma cystatin-like inhibitors of cysteine cathepsins, are degraded at sites of inflammation, and cathepsin B has been identified as a prominent mediator of this process. Cathepsin B, in contrast to cathepsins L and S, is poorly inhibited by kininogens. This led us to delineate the molecular interactions between this protease and kininogens (high molecular weight kininogen and low molecular weight kininogen) and to elucidate the dual role of the occluding loop in this weak inhibition. Cathepsin B cleaves high molecular weight kininogen within the N-terminal region of the D2 and D3 cystatin-like domains and close to the consensus QVVAG inhibitory pentapeptide of the D3 domain. The His110Ala mutant, unlike His111Ala cathepsin B, fails to hydrolyze kininogens, but rather forms a tight-binding complex as observed by gel-filtration analysis. K_i values (picomolar range) as well as association rate constants for the His110Ala cathepsin B variant compare to those reported for cathepsin L for both kininogens. Homology modeling of isolated inhibitory (D2 and D3) domains and molecular dynamics simulations of the D2 domain complexed with wild-type cathepsin B and its mutants indicate that additional weak interactions, due to the lack of the salt bridge (Asp22-His110) and the subsequent open position of the occluding loop, increase the inhibitory potential of kininogens on His110Ala cathepsin B.

© 2010 Elsevier Ltd. All rights reserved.

Keywords: cysteine protease; cystatin; kininogen; protease inhibitor; proteolysis

Introduction

*Corresponding author. INSERM U618 Protéases et Vectorisation Pulmonaires, Equipe Protéases et Pathologies Pulmonaires, Université François Rabelais—Faculté de Médecine, 10 Boulevard Tonnellé, F—37032 Tours cedex, France. E-mail address: gilles.lalmanach@univ-tours.fr.

† C.N. and F.L. contributed equally to this work.

Abbreviations used: AMC, 7-amino-4 methylcoumarin; CA-074, (L-3-*trans*-(propylcarbamoyl)oxirane-2-carbonyl)-L-isoleucyl-L-proline; CP, cysteine protease; CPI, cysteine protease inhibitor; E-64, L-3-carboxy-*trans*-2,3-epoxypropionyl-leucylamido-(4-guanidino) butane; EDTA, ethylenediaminetetraacetic acid; HMWK, high molecular weight kininogen; LMWK, low molecular weight kininogen; MD, molecular dynamics; PBS, phosphate-buffered saline; PDB, Protein Data Bank; SCR, structurally conserved region; Z, benzyloxycarbonyl.

Under physiological conditions, the proteolytic activities of cysteine cathepsins [cysteine proteases (CPs)] are specifically regulated by inhibitors of the cystatin superfamily [cysteine protease inhibitors (CPIs), MEROPS family I25¹], which includes stefins, cystatins, and kininogens.^{2,3} Globally, the inhibitory sites of these proteins consist of three regions that form a wedge-shaped binding surface, an N-terminal region with a conserved glycyl residue, a hairpin loop containing the consensus sequence Gln-X-Val-X-Gly, and a second C-terminal hairpin loop.^{4,5} Stefins (family 1) are single-domain proteins of about 100 amino acids devoid of disulfide bonds (~11 kDa), while family 2 cystatins (~12 kDa) have two internal disulfide bonds and may be glycosylated or phosphorylated.^{6,7} Two

multi-functional, multi-domain glycoproteins, termed high molecular weight kininogen (HMWK, 90–120 kDa) and low molecular weight kininogen (LMWK, 50–70 kDa),^{8,9} make up family 3. Besides being precursors of proinflammatory peptides (kinins), novel functions of kininogens or their cleaved forms (such as angiostatic or microbicidal properties) have been recently reviewed by Lalmach *et al.*¹⁰ Both kininogens consist of an identical N-terminal heavy chain, three tandemly repeated cystatin-like domains (D1, D2, and D3), a short domain (D4) corresponding to the kinin moiety, and distinct light chains (HMWK: domains D5 and D6; LMWK: domain D5) at their C-terminus. The first domain (D1) does not possess inhibitory properties, while domains D2 and D3 are tight-binding inhibitors of cysteine cathepsins.^{11,12}

Genes for 11 functional papain-like cysteine cathepsins (clan CA, family C1) are present in the human genome.^{13,14} In addition to protein degradation and turnover in the endosomal/lysosomal system, these proteases mediate extracellular matrix turnover, antigen presentation, and processing events^{15–17} and represent promising drug targets for osteoporosis, Alzheimer disease, asthma, arthritis, and cancer.^{18–20} Their catalytic site is highly conserved and is formed by three residues: Cys25 (nucleophile), His159, and Asn175 (papain numbering). Activity depends on the protonated His159/deprotonated Cys25 ion pair, which is stabilized by the side chain of Asn175.²¹ Modifications to two members of this family allow them to act as carboxypeptidases. A short loop is present in cathepsin X, which constrains its S' site, restricting its activity to the removal of single residues from the substrate C-terminus.²² In contrast, cathepsin B acts as a broad-spectrum endopeptidase, with an optimal activity in mildly acidic pH,²³ but it is also a peptidyl dipeptidase, sequentially removing C-terminal dipeptides.²⁴ X-ray structural analysis disclosed the molecular basis for this dual character, showing the presence of a 20-residue insertion (termed the "occluding loop"^{25–29}) that is held in place by salt bridges and blocks off the C-terminal side of the substrate-binding cleft.³⁰ Deletion of the cathepsin B occluding loop leads to a complete elimination of exopeptidase activity,³¹ and two critical residues (His110 and His111) have been identified to exert a key role in the positioning of the occluding loop.³²

Enzymes lacking a loop insertion such as cathepsins L, S, H, and K are efficiently inactivated by "three finger tip" inhibitors, that is, cystatins or chagasin (MEROPS family I42).^{9,33} Cathepsin X with its non-flexible "mini-loop"²² is therefore resistant to cystatin C inhibition.³⁴ Since the occluding loop in cathepsin B is flexible, it is inhibited by cystatins (low micromolar).³⁵ While general features of the binding and inhibition of cathepsin B by stefins (family 1) and cystatins (family 2) have been well established,³⁶ fewer studies have been reported for kininogens (family 3 cystatins). LMWK cystatin-like domains (D2 and D3 domains) bind each cathepsin

molecule independently and D3 has a larger k_{on} , resulting in greater inhibitor potential.³⁷ Similarly, HMWK obeys a 2:1 binding stoichiometry, with different rate constants for each binding site.³⁸ Numerous three-dimensional structures of cystatins and stefins have been described,^{39–42} whereas no structural or mechanistic data are presently available for kininogens.

The presence of two of the major CPIs in plasma, cystatin C and kininogens, keeps under control CP activities under normal conditions (see, for review, Refs. 2 and 3). However, an imbalance may occur under pathophysiological conditions. For instance, cathepsin B levels are increased in sera from patients with breast, ovarian, bladder, and uterine cervical cancer and melanoma (see, for review, Ref. 43). Besides cancer, cathepsin B has been implicated in inflammatory diseases, and an upregulated level of cathepsin B was found in patients with rheumatoid arthritis⁴⁴ while circulating cathepsin B–kininogen complexes were detected in plasma from patients with inflammatory disorders, but not from healthy individuals.⁴⁵ Recently, we identified active cathepsin B as the most abundant CP in bronchoalveolar lavage fluids from patients with acute inflammatory diseases or silicosis,^{46–48} and interestingly, the kininogens were found to be extensively degraded (for review, see Ref. 49). Inhibitory kininogen fragments (*circa* 20–40 kDa) that may be produced under pathophysiological conditions, for example, in disseminated intravascular coagulation or polytrauma,⁵⁰ were found in body fluids including ascites, synovial fluid, and plasma and in placenta.^{12,44,51,52} Among proteases mediating such cleavages, cathepsin D, an aspartic protease, may process kininogens,⁵² and we demonstrated that kininogen degradation partly depends on CP activities in bronchoalveolar lavage fluids.⁴⁸

The importance of cathepsin B as drug target in human illness, including tumor metastasis, inflammation, or Alzheimer disease,¹⁹ requires an improved understanding of cathepsin B regulation, which can be exploited in the design of more selective and potent inhibitors.⁵³ The absence of robust interactions between cathepsin B and kininogens that favor the CP/CPI imbalance and the deleterious activity of cathepsin B is of significant interest. The rationale for the present work was to account for the ineffective inhibition of cathepsin B by kininogens. Molecular events arising during cathepsin B/kininogen interactions have been studied, and reasons for the weak inhibition were elucidated. Cathepsin B cleaved within the N-terminal region of both D2 and D3 cystatin-like domains and close to the QVVAG consensus region of the D3 domain, leading to subtle structural rearrangements of HMWK as suggested by circular dichroism (CD) analysis. In addition to the well-documented impairment of peptidyl dipeptidase activity, the lack of a functionally "closed" occluding loop allows the formation of a stable, tight-binding, and uncleaved complex between His110Ala cathepsin B mutant and kininogens, displaying K_i values similar to those for cathepsin L inhibition. Modeling

analysis and molecular dynamics (MD) calculations consistently support these experimental findings.

Results and Discussion

Processing of HMWK by cathepsin B

In contrast to cathepsins L and K, cathepsin B is poorly inhibited (K_i in the micromolar range³³) by kininogens that show a 2:1 binding stoichiometry with their target enzymes. Cathepsin B retained approximately 70% activity towards the synthetic substrate benzyloxycarbonyl (Z)-Phe-Arg-7-amino-4 methylcoumarin (AMC) under equimolar conditions (1:1), as well as in the presence of a stoichiometric excesses of HMWK (inhibitory sites/cathepsin B molar ratios of 4:1 and 2:1), confirming the inability of HMWK to effectively control cathepsin B activity. However, this residual and incomplete inhibition was stable for 60 min and remained unaffected by overnight incubation (data not shown), while cathepsin B was fully inhibited by addition of tight-binding cystatin C in a parallel experiment. Nevertheless, SDS-PAGE analysis (Coomassie Blue staining) indicated that cathepsin B generated two major cleavage products of ~80 and 60 kDa in a dose-dependent manner and a third fragment of ~35 kDa from HMWK (Fig. 1a) following 1 h incubation. This cleavage is abrogated in the presence of L-3-carboxy-*trans*-2,3-epoxypropionyl-leucylamido-(4-guanidino) butane (E-64), supporting the earlier proposal that HMWK is a substrate for cathepsin B under pathophysiological conditions.⁵⁴ In addition to its predominant native form, two lower bands may be observed for HMWK (Fig. 1b, lane 1) in the presence of a polyclonal anti-human HMWK, corresponding to faint amounts of degradation products that are not observed by Coomassie Blue staining (Fig. 1a, lane 1). Western blot analysis also revealed two additional minor hydrolysis products (M_r ~30 kDa and M_r ~20 kDa, respectively) that are hard to distinguish because of their low intensity and a band of M_r ~12 kDa in the presence of the anti-HMWK antibody. Cleavage positions within HMWK were identified by N-terminal sequencing (Fig. 1c). Two of these cleavage sites correspond to primary attack sites in the so-called "proteinase-sensitive regions" described by Vogel *et al.*¹² These regions that are located at the junctions of the three cystatin-like domains constituting the kininogen heavy chain were initially identified as two stretches of 11–12 residues susceptible to hydrolysis by bromelain, trypsin, chymotrypsin, and elastase. Indeed, cathepsin B cleaved the D2 cystatin-like domain of HMWK at the Cys124-Leu125 bond, two residues upstream of the N-terminal, highly conserved Gly126, and the D3 domain at the Lys244-Ile245 bond, four residues upstream of its conserved glycyl residue (Gly248), the latter cleavage occurring at the same position as that of trypsin.¹² The Ile283-Asp284 bond near the

consensus QVVAG inhibitory pentapeptide of the D3 domain is also vulnerable upon incubation with cathepsin B. These cleavages occurring in the vicinity of critical residues of the inhibitory machinery most probably participate in the weak inhibition of cathepsin B, corroborating that HMWK could be considered as a cathepsin B substrate under pathological conditions.⁵⁴ It could be also hypothesized that the conformational constraint imposed by the presence of the cathepsin B occluding loop may enhance the proteolytic susceptibility of HMWK. Interestingly, it was shown that, following the degradation of the HMWK heavy chain into smaller cystatin-like domains by the action of serine proteases, the released domains remain potent inhibitors of cathepsin L.¹² Furthermore, truncation of the N-terminal region of the third cystatin-like domain of rat T-kininogen (thiostatin) does not prevent it from inhibiting cathepsin L, which is in contrast to findings with family 2 cystatin.⁵⁵

Similarly, we assessed the ability of proteolysis products of HMWK, generated by the action of wild-type cathepsin B, to inhibit cathepsin L. Despite the fact that cathepsin B cleaves both D2 and D3 domains in the vicinity of highly conserved residues thought to participate in the reactive inhibitory site, after incubation of HMWK with cathepsin B for 1 h, followed by the addition of the selective inhibitor of cathepsin B (L-3-*trans*-(propyl-carbamoyl)oxirane-2-carbonyl)-L-isoleucyl-L-proline (CA-074), the inhibitory potential of HMWK towards cathepsin L remained unaffected and was comparable to uncleaved kininogen (data not shown). The *in vivo* formation of kininogen inhibitory fragments was proposed to be a relevant way of inhibiting cysteine cathepsins during inflammatory episodes that favor the local release of lysosomal proteinases from cells and tissues.^{56,57} In addition, four extra cleavage sites were identified for cathepsin B; one inside the non-inhibitory domain D1 (junction Gly49-Ser50), one between Gly366 and Phe367 within domain 4, corresponding to the insertion site of bradykinin, and two in the middle part of domain 5 (His421-Asn422 and Gly452-Leu453 bonds). Identification of a single cleavage at the Gly366-Phe367 bond within the bradykinin moiety confirms that cathepsin B is unable to generate pharmacologically active kinins from kininogens.⁵⁸

Further investigation by CD spectroscopy was performed as shown in Fig. 2. The far-UV CD spectrum of unbound HMWK corresponded to ~28% β -structure (SOMCD software⁵⁹), agreeing with its estimated secondary structures calculated by Jpred 3.⁶⁰ For the sake of clarity, the far-UV CD spectra of HMWK corresponding to HMWK/cathepsin B mixtures were reported after subtraction of the individual CD spectrum of wild-type cathepsin B. A decrease of β -structure (~16%) was observed in the presence of 2 mM dithiothreitol (DTT) in comparison to free HMWK. In contrast, the CD spectrum of the HMWK/cathepsin B mixture in the absence of DTT in the assay buffer did not differ

Fig. 1. Hydrolysis of HMWK by cathepsin B. (a) Dose-dependent degradation of HMWK by cathepsin B. HMWK (480 nM) was incubated with increasing concentrations of cathepsin B (inhibitory sites/cathepsin B molar ratios of 4:1, 2:1, and 1:1 respectively) in the assay buffer at 37 °C for 1 h prior to separation of the cleavage products on a (10–15%) gradient polyacrylamide gel under non-reducing conditions. Lane 1, HMWK; lane 2, HMWK/cathepsin B (4:1); lane 3, HMWK/cathepsin B (2:1); lane 4, HMWK/cathepsin B (1:1); lane 5, HMWK/cathepsin B (1:1) preincubated by E-64 (10 μM); lane 6, cathepsin B (indicated by an arrow). (b) Immunodetection of cleavage products of HMWK by cathepsin B. HMWK (two cystatin-like inhibitory sites/molecule) was incubated with cathepsin B (molar ratio of 4:1) at 37 °C for 1 h. Following their separation by 12% SDS-PAGE under reducing (lanes 1, 2, 5, 6, and 7) and non-reducing (lanes 3 and 4) conditions and transfer onto a nitrocellulose sheet, hydrolysis products were immunodetected by using a rabbit polyclonal anti-HMWK antibody⁵¹ (lanes 1 to 4) and a goat polyclonal anti-kininostatin (domain D5 of the light chain) antibody (lanes 5 to 7). Lanes 1, 3, and 5, HMWK; lanes 2, 4, and 6, HMWK/cathepsin B; lane 7, kininostatin. (c) Location of cathepsin B cleavage sites within HMWK. Schematic representation of the primary sequence of HMWK (Uniprot-Swissprot ID: P01042) that shows (delineated with vertical bars), starting from the amino-terminus, the three cystatin-like domains of the heavy chain (domains 1–3), the bradykinin (BK) moiety (D4), and both domains D5 (kininostatin) and D6 that compose the carboxy-terminal light chain. The QVVAG consensus sequence of inhibitory domains 2 and 3 is underlined. The two N-terminal invariant glycine residues of domains D2 and D3 are highlighted (*). Seven cleavage sites (▲) were identified by N-terminal peptide sequencing.

significantly from that of free HMWK (Fig. 2). Observed structural changes of HMWK are DTT-dependent and a result of its hydrolysis by cathepsin

B rather than to a rearrangement due to weak HMWK–cathepsin B binding interactions. Despite the fact that cathepsin B cleaves HMWK in a time-

Fig. 2. Analysis by CD spectroscopy of HMWK–cathepsin B interactions. Far-UV CD spectra. Data acquisitions were made at 1-nm intervals (1 s dwell time) between 200 and 250 nm, at 37 °C, and averaged from repeated scans. HMWK (250 nM) was incubated with cathepsin B at molar ratio 1:1 for 5 min in 300 mM sodium phosphate, pH 5.5, containing 2 mM DTT at 37 °C, before recording (○). The same procedure was repeated in the absence of DTT (□). Individual CD spectrum of free HMWK was measured as control (▲). Spectra of HMWK corresponding to both HMWK/cathepsin B mixtures (in the presence of DTT, ○; in the absence of DTT: □) are presented after subtraction of the individual CD spectrum of cathepsin B.

dependent manner, the apparent lack of difference in the CD spectra observed after 5 and 60 min of incubation (data not shown) remains unexplained. It can be speculated that cleavage of HMWK within β -sheets of cystatin-like domains of its native heavy chain leads to a rapid and time-stable structural reorganization void of any further CD spectral changes.

His110Ala cathepsin B, but not wild-type cathepsin B, forms a tight-binding complex with kininogens

Both LMWK and HMWK serve as cathepsin B substrates in addition to weak inhibitors. This inability to strongly control cathepsin B activity appears to be due to the presence of the occluding loop, which is held in place by residues forming salt bridges.³⁰ Two residues in the occluding loop are of particular interest, His110 and His111. The His110Ala mutant exhibited increased occluding loop mobility as well as an improved endopeptidase activity, whereas the His111Ala cathepsin B lacked its peptidyl dipeptidase activity while the occluding loop remained in the closed position.³² HMWK was incubated with wild-type cathepsin B as well as both His110Ala and His111Ala mutants (inhibitory sites/enzyme molar ratio=4:1). The catalytically inactive recombinant form of cathepsin B (His110Ala/Cys29Ala), which is deprived of its nucleophilic residue, was used as a control in order to evaluate

the importance of the arrangement of the occluding loop in the absence of proteolytic cleavage. The His111Ala mutant processed HMWK similar to wild-type cathepsin B (Fig. 3). Conversely, no cleavage of HMWK by the His110Ala form was observed after 1 h of incubation. A similar pattern of proteolysis was observed for LMWK (data not shown). We further evaluated the ability of kininogens to form a stable, long-lasting complex with cathepsin B and its mutants by gel-filtration chromatography. However, experiments were conducted with LMWK rather than HMWK because of nonspecific electrostatic interactions between the light chain of HMWK and the column (Superdex 200). LMWK was eluted at the retention time expected from the calibration curve. When LMWK (120 nM) was incubated with either wild-type cathepsin B (30 nM) or the His111Ala mutant (30 nM) for 1 h prior to loading on Superdex 200, the enzymatic activity towards Z-Phe-Arg-AMC was fully recovered in fractions corresponding to the apparent molecular mass (M_r) of cathepsin B. In contrast, in a similar experiment with the His110Ala mutant, the fractions containing enzymatic activity eluted at a position corresponding to a higher M_r for cathepsin B (Fig. 4). Similarly, the maximal inhibitory activity of LMWK was observed in fractions corresponding to free LMWK ($M_r \sim 70$ kDa) for both wild-type cathepsin B/LMWK and His111Ala mutant/LMWK mixtures, but shifted to a lower elution volume (11–13 ml) for the His110Ala

Fig. 3. Cleavage of HMWK by cathepsin B and its occluding loop mutants. HMWK was incubated in the presence of human liver cathepsin B, recombinant wild-type form, and its site-directed mutants at 37 °C for 1 h before running on a 12% SDS-PAGE under reducing conditions as previously described. Following transfer, polyclonal anti-HMWK antibody was used for immunodetection.⁵¹ Lane 1, HMWK; lane 2, human liver cathepsin B/HMWK; lane 3, recombinant wild-type cathepsin B/HMWK; lane 4, His111Ala cathepsin B/HMWK; lane 5, His110Ala cathepsin B/HMWK; lane 6, His110Ala/C29A cathepsin B/HMWK.

Fig. 4. His110Ala cathepsin B mutant forms a stable complex with kininogen. LMWK (120 nM) and wild-type cathepsin B or its site-directed mutants His1110Ala and His111Ala (inhibitory sites/cathepsin B molar ratio of 4:1) were incubated for 1 h at 37 °C in 0.1 M sodium acetate buffer, pH 5.5, containing 2 mM DTT, 2 mM EDTA, and 0.01% Brij 35 before gel-filtration analysis of complex-forming capabilities of LMWK with wild-type cathepsin B and its mutants ($\lambda_{220\text{ nm}}$). Residual inhibitory potential towards exogenous papain (a) and residual enzymatic activity (b) were measured for each fraction (1 ml) collected using Z-Phe-Arg-AMC as substrate. Fractions 11–17 were slot-blotted, and immunodetection was performed using an anti-human cathepsin B antibody (c). (○): LMWK, (◆): wild-type cathepsin B/LMWK, (□): His111Ala cathepsin B/LMWK, (△): His110Ala cathepsin B/LMWK.

mutant/LMWK mixture. Subsequent apparent masses (M_r range = ~90–120 kDa) are consistent with the presence of stable complexes between His110Ala cathepsin B and LMWK and may correspond to both 1:1 and 2:1 binding enzyme/inhibitor stoichiometry. In addition, immunodetection of His110Ala mutant by slot-blot analysis was shifted to earlier fractions compared to its parent molecule or His111Ala mutant, again supporting the formation of a tight-binding complex between His110Ala mutant and LMWK (Fig. 4c). The pre-

sence of the occluding loop that restricts access to the primed region of the cathepsin B active-site cleft was initially proposed to be responsible for the overall lower affinities of cystatins and stefins, compared with those endopeptidases with open active sites (such as cathepsin L²⁷) since partial deletion of the occluding loop of cathepsin B increased the inhibitory activity.³¹ The present data with kininogens support the above hypothesis.

Michaelis constants for hydrolysis of Z-Phe-Arg-AMC by cathepsin B and its mutant forms were measured, prior to determining the K_i values. K_m values varied from 180 μM for wild-type cathepsin B to 5 μM for His110Ala cathepsin B and to 21 μM for His111Ala cathepsin B. Apparent inhibition constants $K_{i(\text{app})}$ were further determined as described in Materials and Methods and K_i values deduced from Eq. (1) are reported in Table 1. Slightly enhanced inhibition by HMWK was observed for the His111Ala cathepsin B variant ($K_i = 12\text{ nM}$), while an ~20,000-fold increase of inhibition occurred for the His110Ala cathepsin B mutant ($K_i = 0.005\text{ nM}$) when compared to its wild-type form ($K_i = 100\text{ nM}$). As expected, K_i values for cathepsin B mutants were similar for LMWK. Potent inhibition of His110Ala cathepsin B by both kininogens accounts for the stability of the His110Ala cathepsin B/LMWK complex on gel filtration, which is in contrast to that observed for His111Ala and wild-type cathepsin B. Furthermore, the K_i values for the His110Ala cathepsin B variant were comparable to those reported for cathepsin L for both kininogens (Table 1). Association rate constants for the inhibition of His110Ala cathepsin B ($k_{\text{ass}} = 1.4 \times 10^6\text{ M}^{-1}\text{ s}^{-1}$) and cathepsin L ($k_{\text{ass}} = 6.0 \times 10^6\text{ M}^{-1}\text{ s}^{-1}$) by HMWK are similar, suggesting that the improved affinity of the cathepsin B mutant compared to the wild-type enzyme for kininogens mostly depends on the increase in k_{ass} favoring the formation of the inhibitory complex. Alternatively, the dissociation rate constant for the inhibition of His110Ala cathepsin B ($k_{\text{diss}} = 7.0 \times 10^{-6}\text{ s}^{-1}$) differs from that of cathepsin L ($k_{\text{diss}} = 1.2 \times 10^{-4}\text{ s}^{-1}$), suggesting that the presence of the loop in its open conformation partly stabilizes the complex and delays dissociation between the two partners. Previous studies have demonstrated that binding of both stefin A and cystatin C to wild-type cathepsin B is a two-step reaction, with a rate-limiting step corresponding to the displacement of the occluding loop.^{61,62} This is in contrast to papain and cathepsins L and S, which bind family 1 and 2 cystatins according to a one-step mechanism.^{63–66} Likewise, the present experimental kinetics values, corresponding to the tight-binding inhibition of His110Ala cathepsin B by both kininogens, strongly suggest that the disruption of the salt bridge between His110 and Asp22 results in increased mobility of the occluding loop and a subsequent improved accessibility to the active site as previously discussed in detail by Pavlova *et al.*³⁶ and Redzyna *et al.*,³⁵ leading to a single, one-step reaction mechanism similar to that observed for

Table 1. Inhibition of wild type and recombinant variants of cathepsin B by stefin B and kininogens

Cathepsin B	Stefin B	LMWK	HMWK		
	K_i (nM)	K_i (nM)	K_i (nM)	k_{ass} ($\text{M}^{-1} \text{s}^{-1}$)	k_{diss} (s^{-1})
Wild type (human liver)	4 ± 0.1	70 ± 6.5	100 ± 20	n.d.	n.d.
Recombinant wild type (<i>P. pastoris</i>)	2 ± 0.2	62 ± 10	107 ± 8	n.d.	n.d.
Mutant H110A	0.008 ± 0.002	0.009 ± 0.005	0.005 ± 0.002	$1.4 \pm 0.2 \times 10^6$	7×10^{-6a}
Mutant H111A	n.d.	27 ± 5	12 ± 2	n.d.	n.d.
Cathepsin L	0.0095^b	$0.017^c - 0.014^b$	$0.013^b - 0.019^d$	6.0×10^{6d}	1.2×10^{-4d}

n.d., not determined.
^a Deduced from Eq. (3).
^b Taken from Ref. 33.
^c Taken from Ref. 11.
^d Taken from Ref. 9.

occluding loop-lacking cathepsins interacting with stefins and cystatins. Human stefin B, similarly to kininogens and their isolated inhibitory domains, binds tightly to papain and cathepsin L and more weakly to cathepsin B. Stefin B also has a lower affinity to cathepsin B than ovocystatin or stefin A.^{67,68} A dramatic increase in inhibition for the His110Ala cathepsin B mutant ($K_i = 8$ pM) was observed, compared to its wild-type form (K_i in the nanomolar range) (Table 1). Furthermore, the inhibitory potency for the His110Ala mutant is identical with that for cathepsin L, supporting the hypothesis that the occluding loop in the wild-type enzyme limits access to the active site and thus results in a greater K_i . Nevertheless K_i variations for stefin B between His110Ala cathepsin B *versus* its wild-type form (~200- to 500-fold difference) are less dramatic than those measured for both kininogens (~7000- to 20,000-fold difference) and correspond to the K_i values reported in the literature for cystatins (families 1 and 2) with cathepsin L and cathepsin B.

Fig. 5. Analysis of interactions between HMWK and cathepsin B and its mutants by CD. Far-UV CD spectra were obtained using 250 nM for both HMWK and enzyme in 300 mM sodium phosphate, pH 5.5, 2 mM DTT, at 37 °C, as described in Fig. 3. CD spectra of HMWK—corresponding to HMWK/cathepsin B (or its mutants) complexes—are reported after subtraction of the individual CD spectrum of cathepsin B (or its mutants). (◇): HMWK, (▲): wild-type cathepsin B/HMWK, (□): His111Ala cathepsin B/HMWK, (■): His110Ala cathepsin B/HMWK.

Complementary far-UV CD analyses were conducted. As described above, CD spectra of HMWK corresponding to HMWK/cathepsin B mixtures were reported after subtraction of the individual CD spectrum of wild-type cathepsin B. CD spectra of HMWK corresponding to both wild-type cathepsin B/HMWK and His111Ala cathepsin B/HMWK complexes were closely superimposable (Fig. 5). On the other hand, a notable increase attributable to β -structure (~39%) was observed for the His110Ala cathepsin B/HMWK complex. This most probably reflects rearrangements associated with the formation of a tight-binding complex, in contrast to His111Ala cathepsin B/HMWK and wild-type cathepsin B/HMWK that form weak inhibitory complexes.

Molecular modeling of kininogen inhibitory domains and of the D2 domain complexed with wild-type cathepsin B and occluding loop mutants

A molecular analysis was conducted in order to account in a qualitative manner for the experimentally observed inhibition profile of kininogens with wild-type cathepsin B and its two mutant forms. Since the three-dimensional structures of kininogens are not available, we built molecular models of the inhibitory cystatin-like D2 and D3 domains. The 3-D modeled D2 domain was then docked into cathepsin B and its His111Ala and His110Ala mutants for a comparative analysis.

Molecular modeling of both domains was conducted using human cystatin F [Protein Data Bank (PDB) ID: 2CH9⁶⁹] as template. The well-conserved protease binding segments, the flexible N-terminal part, and two hairpin loops, one in the central region (L1) (QXVXG) and another towards the C-terminal region L2 (PW), were aligned at the same position as cystatin F. Both HMWK and LMWK (family 3) harbor type 2 cystatin-like domains (D2 and D3) containing two intramolecular disulfide bonds at their C-termini. Alignment with the template structure places the Cys residues in close proximity to form a disulfide bond. Stepwise energy refinement of the initial complex generates a convincing model of the D2 domain of human kininogen (Fig. 6a).

Fig. 6. Binding mode of the D2 cystatin-like domain to cathepsin B. The inhibitory cystatin-like domain of the heavy chain of human kininogens is shown in green. (a) Molecular model of the D2 domain. (b) Wild-type cathepsin B (white). (i) H-bond: backbone carbonyl of Cys119 and side chain of Asn58, (ii) conserved ionic interaction between Asp122 and Arg48. Unliganded wild-type cathepsin B is overlaid (brown) for reference. Bonds are depicted as broken lines. Distances are labeled in angstrom. (c) His111Ala mutant (violet). (i) H-bond: backbone amide of Gly121 and side chain of Gln52, (ii) H-bond: backbone carbonyl of Cys119 and side chain of Asn96, (iii) H-bond: backbone carbonyl of Ala111 and side chain of Tyr98, (iv) conserved ionic interaction between Asp122 and Arg48. (d) His110Ala mutant (magenta). (i) H-bond: backbone carbonyl of Cys119 and side chain of Arg60, (ii) H-bond: backbone carbonyl of Val112 and side chain of Asn111, (iii) H-bond: backbone amide of Gly114 and backbone carbonyl of Gln110, (iv) H-bond: side chain of Asn113 and side chain of Asn34, (v) conserved ionic interaction between Asp122 and Arg48. The D2 cystatin-like domain is shown in green.

From the Ramachandran plot, 89% of the residues are located within the most favored regions and 11% are located in additionally allowed regions. The modeled D2 domain retains the overall fold of template structures; a five-stranded antiparallel β -sheet wrapped around a five-turn α -helix, while the N-terminal β 1 structure was less preserved. The three conserved domains that form the typical tripartite 'wedge' shown in other cystatin crystal structures important for protease inhibition are also

preserved, consistent with its role as a functional CPI. A homology model of the D3 domain was also obtained (data not shown). In this model, three structurally conserved regions (SCRs) were built using the template. Separating the three SCRs are two structurally variable regions. L2 residues of the D3 domain showed excellent sequence identity with the template. Despite both the N- and C-terminal regions of the D3 kininogen domain showing better homology with the human cystatin F compared to

D2, the superposition of the energy-refined 3-D models of both kininogen domains and the template structure revealed similar protease binding sites.

In contrast to stefins and cystatins, cathepsin B cleaves HMWK close to the central consensus region (L1) of the D3 domain, but not the D2 domain. Thus, despite the D3 domain showing higher sequence identity with the template, we chose to further focus on the D2 domain for modeling complexes. Refinement of the complexes between the D2 domain and the three different forms of cathepsin B was carried out by MD simulations. Differences between the three complexes are illustrated with snapshots representing the MD average structure of the last 4 ns of each trajectory. Superimposition of the D2 domain complexed with wild-type, His111Ala, and His110Ala forms of cathepsin B, respectively, showed an overall similar global fold and exhibited significant differences in the loop regions (Supplementary Fig. 1a). The overall architecture and mode of binding of isolated D2 kininogen domain to cathepsin B, in agreement with the chicken cystatin structure determined by X-ray diffraction and NMR spectroscopy,⁷⁰ are characterized by the three well-conserved segments basically involved in the inhibitory mechanism with their target proteases: the flexible N-terminal portion, a hairpin loop (L1) at the central region that is conserved throughout the cystatin superfamily, and partly by another loop (L2) towards the C-terminus. The L1 (QVVAG) segment forms a wedge that enters the active site of cathepsin B near the catalytic cysteine residue. The occluding loop in the wild-type structure is only partially displaced, restricting complete access to the active-site cleft (Supplementary Fig. 1b). The L2 loop and adjacent β -strand of kininogen D2 are unable to bind to the active-site cleft and ultimately embrace the enzyme active site laterally. The N-terminal flexible region is partially solvent exposed and interacts with residues outside the active-site cavity. The incomplete entry of the binding regions of the D2 domain into the active-site cavity may explain why it inhibits wild-type cathepsin B with much lower affinity than other CPs. The MD simulation of the His111Ala mutant shows some differences with respect to the wild-type cathepsin B/D2 complex (Fig. 6b and c). The occluding loop is differently placed in the two structures. A number of electrostatic interactions occur between wild-type cathepsin B and the kininogen D2 domain; however, no H-bonds are observed between the residues of the occluding loop and residues of the D2 domain. On the other hand, an H-bond between the backbone carbonyl of Ala111 (cathepsin B) and the side-chain hydroxyl of Tyr98 (kininogen D2 domain) is formed in the His111Ala mutant/D2 complex. This strengthened interaction of the kininogen domain with His111Ala cathepsin B might participate to the fairly improved inhibition of His111Ala mutant by kininogens as compared to wild-type cathepsin B. His110 is responsible for anchoring the loop in occluding position in wild-type cathepsin B.³² The mutation His110Ala allows this occluding loop to

adopt an open conformation, allowing intimate interactions between the inhibitory domain and the active-site pocket of cathepsin B. In an MD snapshot representing the average structure of the complex, with respect to the wild type protein, residues of the occluding loop of His110Ala cathepsin B interact with the D2 domain through three supplementary H-bonds in addition to the conserved extended electrostatic interactions (see Fig. 6d).

Taken together, the weak kininogen-dependent inhibition of cathepsin B and its His111Ala cathepsin B variant, the lack of detectable complexed forms under mild chromatographic conditions (gel permeation), and their susceptibility to proteolysis by their target enzyme support the fact that kininogens primarily act as cathepsin B substrates. On the contrary, the disruption of a salt bridge through a single mutation (His110Ala cathepsin B variant) and the establishment of supplementary close contacts allow the recovery of tight-binding inhibition by both kininogens. The present data suggest that the occluding loop of cathepsin B (in particular the Asp22-His110 salt bridge), apart from its established role in discriminating between exopeptidase *versus* endopeptidase activities, is used to distinguish members of the three cystatin families and define its inhibitory profile.

Materials and Methods

Reagents

Z-Phe-Arg-AMC, AMC, and DTT were purchased from Bachem (Weil am Rhein, Germany). Ethylenediaminetetraacetic acid (EDTA) was from Merck KGaA (Darmstadt, Germany). E-64 and CA-074 were obtained from Sigma-Aldrich (St. Quentin le Fallavier, France). Molecular weight calibration kits were from Bio-Rad (Ivry-sur-Seine, France).

Enzymes and inhibitors

Human wild-type recombinant procathepsin B and site-directed mutants (His110Ala, His111Ala, and His110Ala/C29A) were produced in the yeast *Pichia pastoris* and processed to the mature forms as described elsewhere.^{31,32,36} Cathepsin B purified from human liver was supplied by Calbiochem (VWR International, Libourne, France). Papain was obtained from Boehringer (Roche Molecular Biochemicals, Mannheim, Germany). Cathepsins were activated in assay buffer, 100 mM sodium acetate, pH 5.5, containing 2 mM DTT and 2 mM EDTA for 5 min at 37 °C prior to kinetic assays [Kontron SFM25 spectrofluorimeter ($\lambda_{\text{ex}}=350$ nm; $\lambda_{\text{em}}=460$ nm)]. Cathepsins were active site titrated using E-64⁷¹ and Z-Phe-Arg-AMC as the substrate. Human LMWK and HMWK and human cystatin C were purchased from Calbiochem. Both kininogens and cystatin C were titrated with E-64-titrated commercial papain. Recombinant kininostatin, expressed in *Escherichia coli* and corresponding to domain 5 of HMWK (Lys438-Ser531), was purchased from R&D Systems Europe (Lille, France).

Proteolysis of kininogens by cathepsin B

HMWK (two cystatin-like inhibitory sites/molecule) was incubated with increasing concentrations of cathepsin B (120, 240, and 480 nM, corresponding to inhibitory sites/cathepsin B molar ratios of 4:1, 2:1, and 1:1 respectively), in 0.1 M sodium acetate, pH 5.5, containing 2 mM DTT, 2 mM EDTA, and 0.01% Brij 35 for 60 min at 37 °C (final volume: 10 µl). Reactions were stopped by addition of reducing SDS-PAGE sample buffer [40 mM Tris/HCl buffer, pH 6.8, containing 4% SDS (w/v), 33% glycerol (w/v), 10% 2-mercaptoethanol (w/v), and 0.002% Bromophenol Blue (w/v)]. Samples were boiled for 5 min and submitted to electrophoresis, using a (10–15%) gradient polyacrylamide gel system (PhastSystem Pharmacia; GE Healthcare Europe, Paris, France) and visualized by Coomassie blue staining. Preincubation of cathepsin B with E-64 (10 µM) was used as a control. Simultaneously, HMWK (480 nM) was incubated with cathepsin B (inhibitory sites/enzyme molar ratios of 4:1, 2:1, and 1:1, respectively) for 0–60 min at 37 °C in activity buffer (20 µl). Aliquots (2.5 µl) were removed, and the residual enzymatic activity was recorded in the presence of Z-Phe-Arg-AMC (20 µM) at 37 °C. Assays were carried out in triplicate. Controls were made using human cystatin C (10 µM) or E-64 (10 µM). Trials were also repeated under non-reducing conditions. Apart from wild-type cathepsin B, HMWK was incubated in the presence of site-directed mutants of cathepsin B (His111Ala, His110Ala, and His110Ala/Cys29Ala) using the same procedure. Assays were performed with LMWK (480 nM) and wild-type cathepsin B and its mutants under the same conditions. Alternatively, in order to assess the inhibitory potency of proteolysis products generated by wild-type cathepsin B, HMWK was incubated with cathepsin B (inhibitory sites/cathepsin B molar ratios of 4:1, 2:1, and 1:1 respectively) in the assay buffer at 37 °C for 1 h. After incubating cathepsin B with CA-074 (10 µM), we measured the inhibitory potential of HMWK towards cathepsin L (0.5 nM) using Z-Phe-Arg-AMC as described above. Preincubation of cathepsin B with CA074 was used as a control.

Detection and identification of cleavage sites

Rabbit anti-human HMWK was prepared as previously described.⁵¹ Goat anti-human kininostatin was supplied by R&D Systems Europe. Goat anti-rabbit IgG–peroxidase conjugate and rabbit anti-goat IgG–peroxidase conjugate were from Sigma-Aldrich. Wild-type cathepsin B (120 nM) was incubated with HMWK (480 nM) for 0–60 min at 37 °C in assay buffer. The reaction was quenched by the addition of reducing buffer, followed by incubation at 100 °C for 5 min. The cleavage products were separated by 12% SDS-PAGE under reducing conditions. After transfer onto a nitrocellulose sheet by semi-dry electrotransfer (2 h, 30 mA), the membrane was blocked with phosphate-buffered saline (PBS) containing 0.1% Tween 20 and 5% (w/v) dried milk for 1 h at room temperature. The membrane was in contact with anti-human HMWK (1:1000, in PBS, 0.1% Tween 20, and 5% dry milk) overnight at 4 °C, washed three times, and then incubated with goat anti-rabbit IgG–peroxidase conjugate (1:5000 in PBS, 0.1% Tween 20, and 5% dry milk) for 1 h at room temperature. Bands were visualized using the Renaissance 4 CN plus (4-chloro-1-naphthol) system (Perkin Elmer, Courtaboeuf, France). Immunodetection of kininostatin was performed by using goat anti-human kininostatin (1:100) and rabbit anti-goat IgG–peroxidase conjugate

(1:2000). For N-terminal sequencing of cleavage products, after incubating the kininogen/cathepsin B mixture (molar ratio 4:1) for 2 h at 37 °C, and separation by SDS-PAGE 12%, the products were immobilized onto a polyvinylidene fluoride membrane (Applied Biosystems Prosorb Inserts, Foster City, CA, USA), and analyzed with a Procise 491 automated amino acid sequencer (Applied Biosystems), as previously described.⁷²

Gel-filtration studies of the complex-forming capacity of HMWK with wild-type cathepsin B and its mutants

LMWK (120 nM) and wild-type cathepsin B (30 nM) were incubated together or separately for 1 h at 37 °C in 0.1 M sodium acetate, pH 5.5, containing 2 mM DTT, 2 mM EDTA, and 0.01% Brij 35. The mixture was then loaded on a size-exclusion chromatography Superdex 200 column (AKTA purifier 900 HPLC system, Amersham, GE Healthcare Bio-Sciences AB, Uppsala, Sweden), and the products eluted in the same buffer (0.4 ml/min, 1 ml fractions, $\lambda_{220\text{ nm}}$). The calibration curve was obtained using the standard proteins, ferritin (450 kDa), albumin (68 kDa), chymotrypsinogen (25 kDa), and cytochrome *c* (12.5 kDa). The same protocol was repeated with His111Ala and His110Ala mutants of cathepsin B. Residual enzymatic activity and residual inhibitory potential towards papain (1 nM) were recorded in the presence of Z-Phe-Arg-AMC. In addition, sample fractions were blotted onto nitrocellulose sheets (Cross-Blot system, Sebia, Issy-Les-Moulineaux, France) for 1 h at room temperature under gentle agitation. After blocking [PBS, 0.1% Tween 20, and 5% (w/v) dried milk for 1 h at room temperature], we performed visualization of wild-type cathepsin B and mutants using a rabbit anti-human cathepsin B antibody (dilution: 1/300) (Fitzgerald, Concord, MA, USA), while HMWK was detected by a rabbit anti-human HMWK (dilution: 1/1000).⁵¹ Enhanced chemiluminescence was used for detection (ECL Plus™ Western Blotting Detection, Amersham Biosciences).

Kinetic measurements

Assays were performed in 0.1 M sodium acetate, pH 5.5, containing 2 mM DTT, 2 mM EDTA, and 0.01% Brij 35 at 37 °C, using Z-Phe-Arg-AMC as substrate ($\lambda_{\text{ex}}=350\text{ nm}$; $\lambda_{\text{em}}=460\text{ nm}$). Michaelis constant (K_m) for wild-type cathepsin B as well as for His110Ala and His111Ala mutants was determined as previously reported.⁷³ Inhibition constants $K_{i(\text{app})}$ were determined under experimental conditions that gave nonlinear dose–response curves as described elsewhere.⁷⁴ Data were plotted using the Easson and Stedman procedure as described by Bieth⁷⁵ and fitted with Enzfitter software (Biosoft, Cambridge). Enzyme concentration $[E_0]$ was at least 1- to 10-fold higher than the K_i value, and less than 5% of substrate (Z-Phe-Arg-AMC) was hydrolyzed during these experiments. Data are reported as means $\pm$ SD ($n=3$). K_i values were calculated from $K_{i(\text{app})}$ values according to Eq. (1):

$$K_i = K_{i(\text{app})} / (1 + S_0 / K_m) \quad (1)$$

The rate constant for association (k_{ass}) between His110Ala cathepsin B mutant and HMWK was determined by monitoring the time dependence of association of equimolar amounts of enzyme and inhibitor active sites (final concentration: 1 nM).⁷⁵ Aliquots were incubated

at 37 °C, and the residual activity was measured as a function of time after addition of Z-Phe-Arg-AMC (3.3 μM). Fractional activity was plotted against time according to Eq. (2):⁷⁵

$$1/[E] = k_{\text{ass}}t + 1/[E_0] \quad (2)$$

with $[E_0]$ being the initial enzyme concentration and $[E]$ being the enzyme concentration at time t . The rate constant for dissociation (k_{diss}) was calculated from the experimental values of K_i and k_{ass} , which are related to k_{diss} as follows:

$$k_{\text{diss}} = K_i \cdot k_{\text{ass}} \quad (3)$$

CD spectroscopy

CD spectra were obtained using a Jasco J-810 spectropolarimeter (Jasco France, Bouguenais, France) equipped with a Peltier thermostat system. The path length of the cell was 1 mm (Hellma, VWR). Wild-type HMWK was incubated with cathepsin B at molar ratio 1:1 for 5 and 60 min in 300 mM sodium phosphate, pH 5.5, containing 2 mM DTT at 37 °C. Data acquisitions were made at 1-nm intervals with a dwell time of 1 s between 200 and 250 nm, at 37 °C in the same buffer, and averaged from repeated scans. The same procedure was repeated in the absence of DTT. Individual CD spectra of wild-type cathepsin B and HMWK were measured as controls. Complexes between HMWK and site-directed cathepsin B mutants (His111Ala and His110Ala) were analyzed similarly. Secondary-structure content was assessed from CD measurements by computational analysis based on Kohonen's self-organizing maps: SOMCD.⁵⁹ The amino acid sequence of human HMWK was obtained from UniProtKB/Swiss-Prot (Uniprot-Swissprot ID: P01042), and its predicted secondary-structure content was estimated by using Jpred 3.⁶⁰

Homology modeling

3-D models were built by using MODELLER9v9 software.⁷⁶ The template structure was selected using the BLAST⁷⁷ server against the PDB. Fifty initial models of D2 and D3 kininogens were generated using the cystatin F crystal structure (PDB ID: 2CH9) as a template. Coordinates of four SCRs, four structurally variable regions (three deletion and one insertion), the N-terminus, and the C-terminus from the template were assigned to the target sequence based on the satisfaction of spatial restraints. All side-chain conformations of the model protein were set using a rotamer library. Finally, the top-ranked model of D2 and D3 was selected (MODELLER GA341 scoring method)⁷⁸ and then subjected to conjugate gradient energy minimization using a distance-dependent dielectric function ($\epsilon = 4r$) and an 8-Å non-bonded cutoff up to an rms gradient of 0.01 kcal/mol-Å using the AMBER force field.⁷⁹ The final structures obtained were analyzed by Ramachandran map using PROCHECK.⁸⁰ These structures were used for docking into wild-type cathepsin B and the His110Ala and His111Ala mutants.

Protein-protein docking

The complex of D2 with the His110Ala mutant was modeled with the occluding loop in the open conformation, with the structure of cathepsin B taken from the

crystal structure of chagasin with His110Ala cathepsin B (PDB ID: 3CBJ).³⁵ ClusPro⁸¹ was used to dock D2 onto cathepsin B. ClusPro applies a series of filters to identify complexes with good surface complementarity and low desolvation energies and selects docked complexes with the broadest energy wells. Briefly, the top 20,000 complexes with the best shape complementarity are trimmed down to 2000 complexes based on electrostatic and desolvation free energies. These are then clustered and ranked according to their cluster size (default value of 9 Å as a radius of clustering) and subjected to a short energy minimization to relieve possible side-chain clashes. We used the top-ranked complex from ClusPro as the starting point for further refinement using MD (see below).

The complexes of D2 with wild-type cathepsin B and the His111Ala mutant were modeled with the occluding loop closed, with the structure of cathepsin B taken from a crystal structure with the small-molecule inhibitor CA030 (PDB ID: 1CSB).³³ The initial complexes were generated by manual docking using Sybyl (Tripos Inc., St. Louis, MO). The orientation of chagasin in its complex with His110Ala cathepsin B (PDB ID: 3CBJ) was used as a guide to position D2. Operationally, this was carried out by superposing the closed and open forms of cathepsin B and manually aligning the D2 structure with chagasin.

MD simulations

Refinement of the complexes between the D2 domain and the three different forms (wild type, His111Ala, and His110Ala, respectively) of cathepsin B was carried out by MD simulations. A corresponding energy-minimized model for each of the three complexes was used as a starting point for MD simulations using the AMBER ff03 force field in the AMBER 9 suite of programs.⁸² The protein was solvated in a truncated octahedron TIP3P water box. The solvated system was energy-minimized with harmonic restraints of 10 kcal mol⁻¹ Å⁻² on all solute atoms, followed by heating from 100 to 300 K over 25 ps in the canonical ensemble (NVT). Then, the solvent density was adjusted by running a 25-ps isothermal isobaric ensemble (NPT) simulation under 1 atm pressure. The harmonic restraints were then gradually reduced to zero with four rounds of 25-ps NPT simulations. After an additional 25-ps simulation, a 12-ns production NPT run was carried out with snapshots collected every 1 ps. For all simulations, a 2-fs time step and a 9-Å non-bonded cutoff were used. The particle mesh Ewald method was used to treat long-range electrostatic interactions,⁸³ and bond lengths involving bonds to hydrogen atoms were constrained by SHAKE.⁸⁴ After 7.5 ns, all three complexes seem to have equilibrated based on the RMSD fluctuations from the initial structures.

Acknowledgements

We acknowledge for their financial support the French cystic fibrosis association "Vaincre la Mucoviscidose" (VLM, France) (G.L.: Grant IC0810). C.N. holds a PhD grant from VLM. J.C.K. and J.S.M. are supported by the Shriners of North America. We are grateful to Daniel Bourry for his technical assistance (iconography facilities, Faculté de Médecine,

Université François Rabelais). We thank M. Brillard-Bourdet (INSERM U618) for N-terminal peptide sequencing and Dr P.K. Nandi (Infectiologie Animale et Santé Publique, Centre INRA de Tours, Nouzilly, France) for circular dichroism facilities and helpful comments.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2010.06.006](https://doi.org/10.1016/j.jmb.2010.06.006)

References

- Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. & Barrett, A. J. (2008). MEROPS: the peptidase database. *Nucleic Acids Res.* **36**, D320–D325.
- Abrahamson, M., Alvarez-Fernandez, M. & Nathanson, C. M. (2003). Cystatins. *Biochem. Soc. Symp.* 179–199.
- Turk, V. & Bode, W. (1991). The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett.* **285**, 213–219.
- Abrahamson, M., Ritonja, A., Brown, M. A., Grubb, A., Machleidt, W. & Barrett, A. J. (1987). Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J. Biol. Chem.* **262**, 9688–9694.
- Machleidt, W., Thiele, U., Laber, B., Assfalg-Machleidt, I., Esterl, A., Wiegand, G. *et al.* (1989). Mechanism of inhibition of papain by chicken egg white cystatin. Inhibition constants of N-terminally truncated forms and cyanogen bromide fragments of the inhibitor. *FEBS Lett.* **243**, 234–238.
- Esnard, F., Esnard, A., Faucher, D., Capony, J. P., Derancourt, J., Brillard, M. & Gauthier, F. (1990). Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylation. *Biol. Chem. Hoppe-Seyler*, **371**, 161–166.
- Laber, B., Krieglstein, K., Henschen, A., Kos, J., Turk, V., Huber, R. & Bode, W. (1989). The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS Lett.* **248**, 162–168.
- Bhoola, K. D., Figueroa, C. D. & Worthy, K. (1992). Bioregulation of kinins: kallikreins, kininogens, and kininases. *Pharmacol. Rev.* **44**, 1–80.
- Muller-Esterl, W. (1987). Novel functions of the kininogens. *Semin. Thromb. Hemostasis*, **13**, 115–126.
- Lalmanach, G., Naudin, C., Lecaille, F. & Fritz, H. (2010). Kininogens: more than cysteine protease inhibitors and kinin precursors. *Biochimie*, in press (DOI information: [10.1016/j.biochi.2010.03.011](https://doi.org/10.1016/j.biochi.2010.03.011)).
- Salvesen, G., Parkes, C., Abrahamson, M., Grubb, A. & Barrett, A. J. (1986). Human low-M_r kininogen contains three copies of a cystatin sequence that are divergent in structure and in inhibitory activity for cysteine proteinases. *Biochem. J.* **234**, 429–434.
- Vogel, R., Assfalg-Machleidt, I., Esterl, A., Machleidt, W. & Muller-Esterl, W. (1988). Proteinase-sensitive regions in the heavy chain of low molecular weight kininogen map to the inter-domain junctions. *J. Biol. Chem.* **263**, 12661–12668.
- Lopez-Otin, C. & Bond, J. S. (2008). Proteases: multifunctional enzymes in life and disease. *J. Biol. Chem.* **283**, 30433–30437.
- Turk, V., Turk, B. & Turk, D. (2001). Lysosomal cysteine proteases: facts and opportunities. *EMBO J.* **20**, 4629–4633.
- Brix, K., Dunkhorst, A., Mayer, K. & Jordans, S. (2008). Cysteine cathepsins: cellular roadmap to different functions. *Biochimie*, **90**, 194–207.
- Chapman, H. A., Riese, R. J. & Shi, G. P. (1997). Emerging roles for cysteine proteases in human biology. *Annu. Rev. Physiol.* **59**, 63–88.
- Lecaille, F., Kaleta, J. & Bromme, D. (2002). Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. *Chem. Rev.* **102**, 4459–4488.
- Mohamed, M. M. & Sloane, B. F. (2006). Cysteine cathepsins: multifunctional enzymes in cancer. *Nat. Rev., Cancer*, **6**, 764–775.
- Turk, B. (2006). Targeting proteases: successes, failures and future prospects. *Nat. Rev., Drug Discov.* **5**, 785–799.
- Vasiljeva, O., Reinheckel, T., Peters, C., Turk, D., Turk, V. & Turk, B. (2007). Emerging roles of cysteine cathepsins in disease and their potential as drug targets. *Curr. Pharm. Des.* **13**, 387–403.
- Mladenovic, M., Junold, K., Fink, R. F., Thiel, W., Schirmeister, T. & Engels, B. (2008). Atomistic insights into the inhibition of cysteine proteases: first QM/MM calculations clarifying the regiospecificity and the inhibition potency of epoxide- and aziridine-based inhibitors. *J. Phys. Chem. B*, **112**, 5458–5469.
- Nagler, D. K. & Menard, R. (1998). Human cathepsin X: a novel cysteine protease of the papain family with a very short proregion and unique insertions. *FEBS Lett.* **434**, 135–139.
- Barrett, A. J. & Kirschke, H. (1981). Cathepsin B, cathepsin H, and cathepsin L. *Methods Enzymol.* **80**, 535–561.
- Mort, J. S. & Buttle, D. J. (1997). Cathepsin B. *Int. J. Biochem. Cell Biol.* **29**, 715–720.
- Cygler, M., Sivaraman, J., Grochulski, P., Coulombe, R., Storer, A. C. & Mort, J. S. (1996). Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion. *Structure*, **4**, 405–416.
- Jia, Z., Hasnain, S., Hiram, T., Lee, X., Mort, J. S., To, R. & Huber, C. P. (1995). Crystal structures of recombinant rat cathepsin B and a cathepsin B–inhibitor complex. Implications for structure-based inhibitor design. *J. Biol. Chem.* **270**, 5527–5533.
- Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber, R. *et al.* (1991). The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J.* **10**, 2321–2330.
- Turk, D., Podobnik, M., Kuhelj, R., Dolinar, M. & Turk, V. (1996). Crystal structures of human procathepsin B at 3.2 and 3.3 angstroms resolution reveal an interaction motif between a papain-like cysteine protease and its propeptide. *FEBS Lett.* **384**, 211–214.
- Turk, D., Podobnik, M., Popovic, T., Katunuma, N., Bode, W., Huber, R. & Turk, V. (1995). Crystal structure of cathepsin B inhibited with CA030 at 2.0-Å resolution: a basis for the design of specific epoxysuccinyl inhibitors. *Biochemistry*, **34**, 4791–4797.
- Nagler, D. K., Storer, A. C., Portaro, F. C., Carmona, E., Juliano, L. & Menard, R. (1997). Major increase in endopeptidase activity of human cathepsin B upon removal of occluding loop contacts. *Biochemistry*, **36**, 12608–12615.

31. Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T. & Mort, J. S. (1997). Role of the occluding loop in cathepsin B activity. *J. Biol. Chem.* **272**, 1197–1202.
32. Krupa, J. C., Hasnain, S., Nagler, D. K., Menard, R. & Mort, J. S. (2002). S2' substrate specificity and the role of His110 and His111 in the exopeptidase activity of human cathepsin B. *Biochem. J.* **361**, 613–619.
33. Lecaille, F., Bromme, D. & Lalmanach, G. (2008). Biochemical properties and regulation of cathepsin K activity. *Biochimie*, **90**, 208–226.
34. Nagler, D. K., Zhang, R., Tam, W., Sulea, T., Purisima, E. O. & Menard, R. (1999). Human cathepsin X: a cysteine protease with unique carboxypeptidase activity. *Biochemistry*, **38**, 12648–12654.
35. Redzynia, I., Ljunggren, A., Abrahamson, M., Mort, J. S., Krupa, J. C., Jaskolski, M. & Bujacz, G. (2008). Displacement of the occluding loop by the parasite protein, chagasin, results in efficient inhibition of human cathepsin B. *J. Biol. Chem.* **283**, 22815–22825.
36. Pavlova, A., Krupa, J. C., Mort, J. S., Abrahamson, M. & Bjork, I. (2000). Cystatin inhibition of cathepsin B requires dislocation of the proteinase occluding loop. Demonstration by release of loop anchoring through mutation of his110. *FEBS Lett.* **487**, 156–160.
37. Turk, B., Stoka, V., Bjork, I., Boudier, C., Johansson, G., Dolenc, I. *et al.* (1995). High-affinity binding of two molecules of cysteine proteinases to low-molecular-weight kininogen. *Protein Sci.* **4**, 1874–1880.
38. Turk, B., Stoka, V., Turk, V., Johansson, G., Cazzulo, J. J. & Bjork, I. (1996). High-molecular-weight kininogen binds two molecules of cysteine proteinases with different rate constants. *FEBS Lett.* **391**, 109–112.
39. Alvarez-Fernandez, M., Liang, Y. H., Abrahamson, M. & Su, X. D. (2005). Crystal structure of human cystatin D, a cysteine peptidase inhibitor with restricted inhibition profile. *J. Biol. Chem.* **280**, 18221–18228.
40. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A. *et al.* (1988). The 2.0 Å x-ray crystal-structure of chicken egg-white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **7**, 2593–2599.
41. Jenko, S., Dolenc, I., Guncar, G., Dobersek, A., Podobnik, M. & Turk, D. (2003). Crystal structure of stefin A in complex with cathepsin H: N-terminal residues of inhibitors can adapt to the active sites of endo- and exopeptidases. *J. Mol. Biol.* **326**, 875–885.
42. Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B. & Turk, V. (1990). The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J.* **9**, 1939–1947.
43. Berdowska, I. (2004). Cysteine proteases as disease markers. *Clin. Chim. Acta*, **342**, 41–69.
44. Lenarcic, B., Gabrijelcic, D., Rozman, B., Drobnic-Kosorok, M. & Turk, V. (1988). Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory and metabolic joint diseases. *Biol. Chem. Hoppe-Seyler*, **369**, 257–261.
45. Takeda, A., Nakamura, Y. & Aoki, Y. (1992). Enzyme-linked immunosorbent assay for the detection of cathepsin-kininogen complexes in human plasma. *J. Immunol. Methods*, **147**, 217–223.
46. Lalmanach, G., Diot, E., Godat, E., Lecaille, F. & Herve-Grepinet, V. (2006). Cysteine cathepsins and caspases in silicosis. *Biol. Chem.* **387**, 863–870.
47. Perdureau, C., Godat, E., Maurel, M. C., Hazouard, E., Diot, E. & Lalmanach, G. (2006). Cysteine cathepsins in human silicotic bronchoalveolar lavage fluids. *Biochim. Biophys. Acta*, **1762**, 351–356.
48. Serveau-Avesque, C., Martino, M. F., Herve-Grepinet, V., Hazouard, E., Gauthier, F., Diot, E. & Lalmanach, G. (2006). Active cathepsins B, H, K, L and S in human inflammatory bronchoalveolar lavage fluids. *Biol. Cell*, **98**, 15–22.
49. Veillard, F., Lecaille, F. & Lalmanach, G. (2008). Lung cysteine cathepsins: intruders or unorthodox contributors to the kallikrein-kinin system? *Int. J. Biochem. Cell Biol.* **40**, 1079–1094.
50. Turk, V., Stoka, V. & Turk, D. (2008). Cystatins: biochemical and structural properties, and medical relevance. *Front. Biosci.* **13**, 5406–5420.
51. Lalmanach, G., Hoebeke, J., Moreau, T., Ferrer-Di Martino, M. & Gauthier, F. (1992). An immunochemical approach to investigating the mechanism of inhibition of cysteine proteinases by members of the cystatin superfamily. *J. Immunol. Methods*, **149**, 197–205.
52. Lenarcic, B., Krasovec, M., Ritonja, A., Olafsson, I. & Turk, V. (1991). Inactivation of human cystatin C and kininogen by human cathepsin D. *FEBS Lett.* **280**, 211–215.
53. Lim, I. T., Meroueh, S. O., Lee, M., Heeg, M. J. & Mobashery, S. (2004). Strategy in inhibition of cathepsin B, a target in tumor invasion and metastasis. *J. Am. Chem. Soc.* **126**, 10271–10277.
54. Barros, N. M., Tersariol, I. L., Oliva, M. L., Araujo, M. S., Sampaio, C. A., Juliano, L. & Motta, G. (2004). High molecular weight kininogen as substrate for cathepsin B. *Biol. Chem.* **385**, 551–555.
55. Moreau, T., Gutman, N., Faucher, D. & Gauthier, F. (1989). Limited proteolysis of T-kininogen (thiostatin). Release of comparable fragments by different endopeptidases. *J. Biol. Chem.* **264**, 4298–4303.
56. Katunuma, N. & Kominami, E. (1986). Distributions and localizations of lysosomal cysteine proteinases and cystatins. In *Cysteine Proteinases and Their Inhibitors* (Turk, V., ed.), pp. 219–227, Walter de Gruyter & Co, Berlin, Germany.
57. Gauthier, F., Gutman, N., Moreau, T. & el Moujahed, A. (1988). 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**, 251–255.
58. Desmazes, C., Gauthier, F. & Lalmanach, G. (2001). Cathepsin L, but not cathepsin B, is a potential kininogenase. *Biol. Chem.* **382**, 811–815.
59. Unneberg, P., Merelo, J. J., Chacon, P. & Moran, F. (2001). SOMCD: method for evaluating protein secondary structure from UV circular dichroism spectra. *Proteins*, **42**, 460–470.
60. Cole, C., Barber, J. D. & Barton, G. J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* **36**, W197–201.
61. Estrada, S., Pavlova, A. & Bjork, I. (1999). The contribution of N-terminal region residues of cystatin A (stefin A) to the affinity and kinetics of inhibition of papain, cathepsin B, and cathepsin L. *Biochemistry*, **38**, 7339–7345.
62. Nycander, M., Estrada, S., Mort, J. S., Abrahamson, M. & Bjork, I. (1998). Two-step mechanism of inhibition of cathepsin B by cystatin C due to displacement of the proteinase occluding loop. *FEBS Lett.* **422**, 61–64.
63. Bjork, I., Brieditis, I., Raub-Segall, E., Pol, E., Hakansson, K. & Abrahamson, M. (1996). The importance of the second hairpin loop of cystatin C

- for proteinase binding. Characterization of the interaction of Trp-106 variants of the inhibitor with cysteine proteinases. *Biochemistry*, **35**, 10720–10726.
64. Estrada, S., Olson, S. T., Raub-Segall, E. & Bjork, I. (2000). The N-terminal region of cystatin A (stefin A) binds to papain subsequent to the two hairpin loops of the inhibitor. Demonstration of two-step binding by rapid-kinetic studies of cystatin A labeled at the N-terminus with a fluorescent reporter group. *Protein Sci.* **9**, 2218–2224.
65. Lindahl, P., Abrahamson, M. & Bjork, I. (1992). Interaction of recombinant human cystatin C with the cysteine proteinases papain and actinidin. *Biochem. J.* **281**, 49–55.
66. Turk, B., Colic, A., Stoka, V. & Turk, V. (1994). Kinetics of inhibition of bovine cathepsin S by bovine stefin B. *FEBS Lett.* **339**, 155–159.
67. Abrahamson, M., Barrett, A. J., Salvesen, G. & Grubb, A. (1986). Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. *J. Biol. Chem.* **261**, 11282–11289.
68. Barrett, A. J., Rawlings, N. D., Davies, M. E., Machleidt, W., Salvesen, G. & Turk, V. (1986). Cysteine proteinase inhibitors of the cystatin superfamily. In *Proteinase Inhibitors* (Barrett, A. J. & Salvesen, G., eds), vol. 12, pp. 515–569. Elsevier, Amsterdam, The Netherlands.
69. Schuttelkopf, A. W., Hamilton, G., Watts, C. & van Aalten, D. M. (2006). Structural basis of reduction-dependent activation of human cystatin F. *J. Biol. Chem.* **281**, 16570–16575.
70. Engh, R. A., Dieckmann, T., Bode, W., Auerswald, E. A., Turk, V., Huber, R. & Oschkinat, H. (1993). Conformational variability of chicken cystatin. Comparison of structures determined by X-ray diffraction and NMR spectroscopy. *J. Mol. Biol.* **234**, 1060–1069.
71. Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M. & Hanada, K. (1982). 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**, 189–198.
72. Linke, T., Ross, A. C. & Harrison, E. H. (2006). Proteomic analysis of rat plasma by two-dimensional liquid chromatography and matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *J. Chromatogr., A*, **1123**, 160–169.
73. Lecaille, F., Weidauer, E., Juliano, M. A., Bromme, D. & Lalmanach, G. (2003). Probing cathepsin K activity with a selective substrate spanning its active site. *Biochem. J.* **375**, 307–312.
74. Moreau, T., Gutman, N., el Moujahed, A., Esnard, F. & Gauthier, F. (1986). Relationship between the cysteine-proteinase-inhibitory function of rat T kininogen and the release of immunoreactive kinin upon trypsin treatment. *Eur. J. Biochem.* **159**, 341–346.
75. Bieth, J. G. (1980). Pathophysiological interpretation of kinetic constants of protease inhibitors. *Bull. Eur. Physiopathol. Respir.* **16**, 183–197.
76. Sali, A. & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.
77. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
78. John, B. & Sali, A. (2003). Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Res.* **31**, 3982–3992.
79. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M. et al. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
80. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
81. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, **20**, 45–50.
82. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr et al. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688.
83. Darden, T., York, D. & Pedersen, L. (1993). Particle mesh Ewald—an $n \log(n)$ method for ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.
84. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of cartesian equations of motion of a system with constraints—molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341.