A Review of Small Molecule Inhibitors and Functional Probes of Human Cathepsin L

Dibyendu Dana [1,2,*,†] and **Sanjai K. Pathak** [1,2,*]

1 Chemistry and Biochemistry Department, Queens College of The City University of New York, 65-30 Kissena Blvd, Flushing, NY 11367, USA

2 Ph.D. Program in Biochemistry, The Graduate Center of the City University of New York (CUNY), 365 5th Ave, New York, NY 10016, USA

* Correspondence: dibyendu.iitm@gmail.com (D.D.); Sanjai.Kumar@qc.cuny.edu (S.K.P.); Tel.: +1-817-370-0490 (D.D.); +1-718-997-4120 (S.K.P.); Fax: +1-718-997-5531 (S.K.P.)

† Current Address: KemPharm Inc., 2200 Kraft Drive, Blacksburg, VA 24060, USA.

check for updates

Abstract: Human cathepsin L belongs to the cathepsin family of proteolytic enzymes with primarily an endopeptidase activity. Although its primary functions were originally thought to be only of a housekeeping enzyme that degraded intracellular and endocytosed proteins in lysosome, numerous recent studies suggest that it plays many critical and specific roles in diverse cellular settings. Not surprisingly, the dysregulated function of cathepsin L has manifested itself in several human diseases, making it an attractive target for drug development. Unfortunately, several redundant and isoform-specific functions have recently emerged, adding complexities to the drug discovery process. To address this, a series of chemical biology tools have been developed that helped define cathepsin L biology with exquisite precision in specific cellular contexts. This review elaborates on the recently developed small molecule inhibitors and probes of human cathepsin L, outlining their mechanisms of action, and describing their potential utilities in dissecting unknown function.

Keywords: human cathepsin L; cathepsin L probe; cathepsin L inhibitor; activity-based probes; clickable ABP; E-64; CA-074; KDP-1; cathepsin L ABP

1. Introduction

Lysosomes play critical roles in human biology receiving, trafficking, processing, and degrading biological molecules from seminal cellular processes, such as endocytosis, phagocytosis, autophagy and secretion. Discovered by the ground-breaking work of de Duve, these single-membrane enclosed cytosolic organelle maintain an acidic (~4.5–5) pH environment, and house close to sixty proteolytic enzymes [1,2]. Among these are the eleven members of the cysteine cathepsin enzymes with a versatile expression and functional profile: Cathepsin L (L1), B, C, F, H, K, O, V (L2), X, S, and W [3]. These enzymes closely mimic the CA1 clan of the papain structure and catalytic cycle and mediate numerous crucial cellular events. For example, they participate in processes involving cell death, protein degradation, post-translational modifications of proteins, extracellular matrix (ECM) remodeling, autophagy, and immune signaling. Given that their functions are aberrantly dysregulated in several human diseases, many are considered prime targets for therapeutic development [4]. Several elegant reviews have recently emerged describing the importance of cysteine cathepsins in both normal physiology and human diseases [5–11]. The focus of this review is specifically on human cathepsin L, a ubiquitously expressed endopeptidase whose involvement in several human diseases has emerged in recent years. These include liver fibrosis, Type I and II diabetes, cardiac and bone, immune and kidney disorders [12–24] In addition, its role in a wide variety of highly invasive forms of cancer is

also now well established and has been a subject of elegant reviews elsewhere [25–27]. Unfortunately, several overlapping and redundant functions of cathepsin L have also emerged [5,28–30]. It is therefore critically important that its functional biology in both normal and disease-specific cell types be clearly annotated before significant resources are directed in drug discovery endeavors. In this review, we will briefly outline the biogenesis of cathepsin L, describe its post-translational processing and trafficking, and highlight key structural features required for formation of an active and mature cathepsin L. A brief summary of cathepsin L biology and its role in human diseases is provided next. Finally, a detailed and up-to-date report on existing cathepsin L-targeting small molecule inhibitors and functional probes with their mechanistic details is described.

Human cathepsin L gene, CTHL (Uniprot primary accession number: P07711), located at the 9q21-q22 position of chromosome 9, encodes for a total of 333 amino acid peptide sequence (M.W. = 37,564 kDa) [31]. The coding space includes the regions of a N-terminal signal peptide, two pro-peptides, and two mature peptides comprising of a heavy (H) chain and a light (L) chain (Figure 1). After transcription, the signal peptide is co-translationally removed in polysome and the resulting 41 kDa pro-cathepsin L is translocated to endoplasmic reticulum where it undergoes N-linked glycosylation with mannose rich sugars. The mannose sugars on the pro-cathepsin L are then phosphorylated in *cis* Golgi by UDP-N-acetylglucosamine:N-acetylglucosaminephosphotransferase enzyme [32]. The mannose-6-phosphate (M6P) receptors located on the surface of the *trans* Golgi network recognize the M6P-pro-cathepsin peptide and deliver the pro-cathepsin L peptide to the lysosome via the endolysosomal pathway. The weakly acidic environment of endosome/lysosome releases M6P receptors and the phosphate group from mannose sugars is removed by a lysosomal acid phosphatase [33,34]. Activation to mature cathepsin L form then occurs by removal of propeptides either by autocatalysis [35] or by aspartyl cathepsin D in the acidic environment of lysosome [36]. This leads to the double chain form of mature and active cathepsin L, comprising of H and L domains, connected by disulfide bridges, (Figure 2). It is to be noted here that several isoforms of cathepsin L have also been observed in specific cell types due to alternative splicing of mRNA transcripts and alternative translation [4,37–41].


MNPTLILAAF CLGIASATLT FDHSLEAQWT KWKAMHNRLY GMNEEGWRRA

VWEKNMKMIE LHNQEYREGK HSFTMAMNAF GDMTSEEFRQ VMNGFQNRKP

RKGKVFQEPL FYEAPRSVDW REKGYVTPVK NQGQCGSCWA FSATGALEGQ

MFRKTGRLIS LSEQNLVDCS GPQGNEGCNG GLMDYAFQYV QDNGGLDSEE

SYPYEATEES CKYNPKYSVA NDTGFVDIPK QEKALMKAVA TVGPISVAID

AGHESFLFYK EGIYFEPDCS SEDMDHGVLV VGYGFESTES DNNKYWLVKN

SWGEEWGMGG YVKMAKDRRN HCGIASAASY PTV

Figure 1. The primary sequence of the full length human inactive human cathepsin L: Blue = 17 Amino acid signal (prepro) peptide, Purple = Propeptides/activation peptides: 96 amino acid (Thr18–Glu113) and 3 amino acid (Glu289–Asp291), Red = Heavy chain peptide, Pink = Light chain peptide; * Disulfide bond pair residues, Cys135–Cys178, $^{\theta}$ Disulfide bond pair Cys169–Cys211, $^{\Omega}$ Disulfide bond pair Cys268–Cys322, ¥ The site of N-linked glycosylation. The two key catalytic residues of the active site, Cys25 and His163 (numbering of mature cathepsin L) residing in the heavy chain are underlined.

The propeptides act as an important regulatory on/off switch as well as a folding catalyst in cathepsin activation. Not surprisingly, the nature of propeptides among cysteine cathepsins is highly divergent by both chain lengths and primary sequences. It is thought that this uniqueness is functionally relevant given its ubiquitous presence in most tissues and allows for the selective suppression of enzyme activity (hence unintended autoactivation) during the transport to the endolysosomal compartment.

In cathepsin L, two inhibitory propeptides, one containing 96 amino acid (Thr18–Glu113) and the other containing 3 amino acid (Glu289–Asp291) exist. A crystal structure of human procathepsin L revealed that the 96 amino acid inhibitory propeptide chain spans in the opposite directions of substrate binding and forms several high-affinity non-covalent interactions with the surrounding residues in active site [42,43]. Interestingly, this opposite direction binding of inhibitory propeptide segment is evolutionarily conserved in other members of cysteine cathepsins, including in cathepsin B.
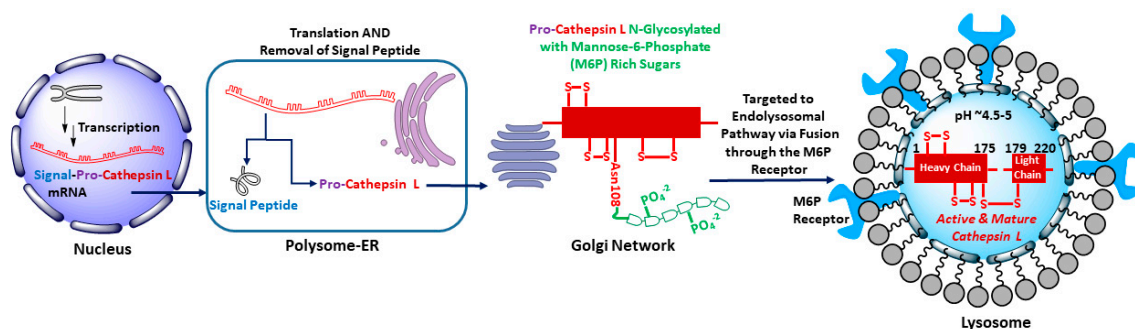
Figure 2. Biogenesis of human cathepsin L. After the full length cathepsin L mRNA is transcribed, it is translated in ribosomes. Following this, the full-length peptide enters the ribosomes-bound endoplasmic reticulum lumen where signal peptide is removed. Pro-cathepsin L then enters the Golgi network where it undergoes N-linked glycosylation at Asn108, followed by mannose phosphorylation and formation of appropriate disulfide linkages. In the last step, modified procathepsin L is shuttled to lysosome by endolysosomal pathways, generating the double chain form of active and mature human cathepsin L.

The dominant pathway of regulation of activated and mature cathepsin L is by endogenous protein inhibitors, cystatins, that like propeptide compete with the physiological substrates for binding to the enzyme active site (Table 1) [5,44]. Interestingly, protein inhibitory agents of cathepsin L have also been reported in other organisms. For example, Kotsyfakis M. et al. reported the existence of two cathepsin L inhibitory proteins in the carrier of the main vector of Lyme disease-carrying parasite, *Ixodus scapularis*. Named so because of its abilities to specifically inhibit cathepsin L (IC$_{50}$ = 4.68 nM; K$_i$ = 95 pM) activity, sialostatin L abrogates the protective proteolytic activity of host cells at the infestation sites, thereby promoting the tick's survival [45]. In addition, it also possess a potent anti-inflammatory and immunosuppressive activity by inhibiting cytotoxic killer T cells [46].

Table 1. Reported physiological inhibitory ligands of human cathepsin L with their inhibition constants.

Inhibitory Ligand	Inhibition Constant (K$_i$)
Recombinant Cathepsin L Propeptide	0.088 nM [47] (pH = 5.5)
Cystatin A	1.3 nM [48,49]
Cystatin B	0.23 nM [48,49]
Cystatin C	<0.005 nM [50]
Cystatin D	18 nM [51]
P41 of MHC Class II Molecule	2 pM [52]
L-Kinenogen	1.7 pM [53,54]
Cystatin F	0.31 nM [55]
Sialostatin L	95 pM [45]
Antimicrobial Peptide LL-37	150 nM [56]

2. Functional Biology of Cathepsin L and Its Role in Human Diseases

Different forms of cathepsin L enzyme have been isolated in distinct cell types and their intracellular and extracellular environment [4]. At specific cellular locations, they seem to play pivotal roles in a wide range of functional activities. In addition to the degradation of cellular proteins by the endolysosomal pathway, cathepsin L also participates in the autophagic degradation pathway; it specifically degrades two autophagosomal markers, LC3-II and GABARAP-II [57]. The specialized podocyte cells in the kidney play a critical role in the retention of various proteins required to be maintained in blood plasma [58]. It was shown that the cytoplasmic variant of active cathepsin L can degrade two key proteins, GTPase dynamin and synaptopodin, required for podocyte's proper glomerular filtration function [23,59,60]. Consistently, a higher podocyte cathepsin L expression was found in a variety of proteinuric kidney diseases, and its functional inhibition (e.g., by creating cathepsin L resistant dynamin) led to reversing of the symptoms [23,61]. Interestingly, a nuclear isoform of cathepsin L is associated with another unique function in kidney cells; it is involved in the proteolytic processing of Cux1, a cell cycle regulatory transcription factor that promotes cell proliferation and polycystic kidney disease (PKD) [62,63]. Indeed, in mouse model of PKD, a lower level of nuclear cathepsin L was corelated with an increased level of Cux1 protein in cysts [62].

Another type of cells where cathepsin L seems to play a major role in maintaining homeostasis is cardiomyocytes [17,64–67]. Early investigation by Stypmann et al. on a year old *Ctsl*–/– mice showed that they developed many key traits of dilated cardiomyopathy, such as interstitial myocardial fibrosis, cardiac chamber dilation, and impaired contraction. In addition, the newborn *Ctsl*–/– mice acquired increased number of acidic organelles, although with altered morphology and function. These studies indicate that the inhibition of cathepsin L activity in these cells is detrimental to the proper functioning of cardiac cells. This was further corroborated by an overexpression study of cathepsin L in mice cardiomyocytes that exhibited cardioprotective effect by inhibiting the Akt signaling pathway [68]. The function of secreted extracellular cathepsin L in cardiac remodeling and repair has also been studied extensively since it is known that many proteins of extracellular matrix (ECM) are also the physiological substrate of cathepsin L; these include laminin, collagen type I, IV and XVIII, and fibronectin [69–71]. Ischemic heart disease is major risk factor for diabetic patients, unable to maintain their sugar levels efficiently. During cardiac repair, endothelial progenitor cells containing elevated levels of cathepsin L home on ischemic cells, and initiate the process of neovascularization [72–74]. In uncontrolled diabetic condition however, cathepsin L function in ECM space during neovascularization is significantly impaired; this hampers the ischemic repair process [15,75].

Controlled release of trypsinogen from pancreatic acinar cells and its subsequent activation in duodenum is a critically important event and must be tightly regulated. Dysregulated release of activated trypsin results in an acute pancreatitis condition. While cathepsin B efficiently cleaves trypsinogen to produce enhanced levels of active trypsin promoting pancreatitis, cathepsin L acts a regulatory brake, and keeps the activation process under control. And it does so by cleaving trypsinogen at a distinct site, 3 amino acids C terminal to cathepsin B, thereby inactivating trypsinogen itself [76]. Consistently, *Ctsl*–/– mice showed increased trypsin activity in pancreas and surprisingly reduced pancreatitis due to diminished inflammation; this effect was attributed due to enhanced apoptosis of acinar cells (vs necrosis).

In neuro- and immuno-biology, cathepsin L has important documented roles as well [77–80]. For example, regulated secretory vesicles of the neuroendocrine system, such as chromaffin granules, house many mature function-ready enkephalin opioid peptides. Upon appropriate signaling stimuli, the activated enkephalin neuropeptides are secreted and released imparting their analgesic and immune-cell specific functions. Studies by Yasothornsrikul et al. revealed that cathepsin L is the proteolytic processing enzyme that renders active metenkephalin from proenkephalin [81]. In another studies, the severity of symptoms due to antigen-induced arthritis (AAI) was significantly reduced in *Ctsl*–/– mice due to a weakened T helper cell population in thymus [82].

The expression of cathepsin L is highly dysregulated in several human diseases, including diabetes, AAI, abdominal aortic aneurysm, liver fibrosis, and cancers [13,82–85]. Its involvement in numerous forms of highly invasive cancers is particularly noteworthy and is covered in great detail in recent reviews [25,26,86–88]. Early studies by Gottesman revealed that extracellular cathepsin L levels can increase up to 200-fold in transformed cells [89]. In the ECM environment, they can rapidly degrade several structural components and significantly increase the migratory/invasive potential of these cells, promoting metastasis [86,87]. Not surprisingly, cathepsin L inhibition by small molecule is considered a viable strategy for the development of novel anti-cancer agents.

3. Small Molecule Inhibitors

Over the years, several classes of physiological and synthetic inhibitors have been discovered targeting cathepsin L. Herein, the focus is on non-physiological inhibitors that could broadly be classified as reversible and irreversible inhibitors. Reversible and irreversible inhibitor could be distinguished by means of the mechanistic approaches they utilize for enzyme inactivation. Reversible inhibitors 'generally' engage with the target protein using non-covalent interaction; however, in certain cases, they form a quasi-covalent bond which eventually disengages from the active site, say, upon dilution. Irreversible inhibitors, on the other hand, permanently modify the protein of interest via the formation of a stable covalent bond. While both reversible and irreversible inhibitors have their pros and cons, it is generally believed that reversible inhibitors are preferred candidates in drug discovery and irreversible inhibitors not so due to their adverse immune responses [90]. There are, however, several examples of successful drugs being used in clinics that work by an irreversible inhibitory mechanism [91]. Irreversible inhibitors are also widely utilized in functional biology (e.g., in the development of activity-based probes). The majority of synthetic cathepsin L inhibitors have recognizable peptide sequences, often derived from its physiological substrates, that bind to the active site of the protein and often contain strategically placed electrophilic warheads that trap the nucleophilic Cys25 residue (Figure 3) for activity. Knowledge of enzyme structure and its associated substrate specificity thus has played an important role in designing selective inhibitors of cathepsins [92,93]. In a seminal study, Choe et al. studied substrate specificity using a highly diversified positional scanning synthetic combinatorial library comprised of 160,000 fluorogenic tetrapeptides; this allowed to differentiate individual enzymes' binding propensity based on their distinct amino acid preferences [93]. By capitalizing on this strategy, they successfully developed a selective substrate and substrate-based inhibitor of cathepsin K. The other important contribution to our understanding of subsite binding preferences of cathepsin L enzyme stems from several timely crystal structure studies that helped reveal key difference in the structural landscape of enzyme subsites [94–96]. In an important study, Shenoy et al. solved the crystal structure of ligand bound cathepsin L and documented the structural compositions of ligand binding sites (Figure 4) [96]. Their analysis of Z-Phe-Tyr (O-*tert*-Butyl)-C(O)C(H)O bound cathepsin L revealed that (a) S1 subsite is relatively wide and unrestricted and composed of Asp162, Ser24, and Cys25, (b) S1′ subsite is guided by Leu144, Trp189, Ala138 and Gly139 where Trp189 associates with Trp193 and Phe143 and forms an aromatic cluster that accommodates the *tert*-butyl group (c) side chains of Leu69 and Met70 help form the S2 subsite that engages in non-polar interactions with the phenyl side chain, and finally, (d) the carboxybenzyl group finds interaction with the Gly68 residue of S3 subsite (Figure 4). However, it turns out that different ligands may find slightly altered binding interactions with the enzyme subsites, depending on their structural features. For example, Shenoy et al. showed that although Z-Phe-Tyr (O-*tert*-Butyl)-DMK binds to the same subsites of cathepsin L as observed for the Z-Phe-Tyr (O-*tert*-Butyl)-C(O)C(H)O ligand and finds some alternative interactions with residues from the active site pocket. This finding has been corroborated by other studies as well that suggest that structural features of ligands can influence the subsite composition of the cathepsin L enzyme [94,95,97,98]. Such information has been duly capitalized to develop different classes of inhibitors that are discussed next. In the following section, we discuss the types of cathepsin L-targeting chemotypes that have been utilized for development of small molecule inhibitors.
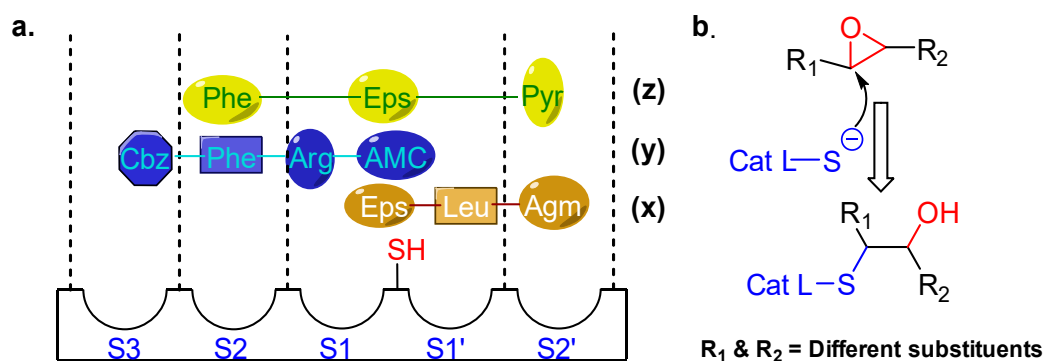
Figure 3. (**a**) Proposed binding mode of (**x**) E-64, (**y**) a putative cathepsin L substrate, and (**z**) CLIK148 to the active site of cathepsin L. (**b**) Inactivation mechanism of cathepsin L by epoxy-containing inhibitors.
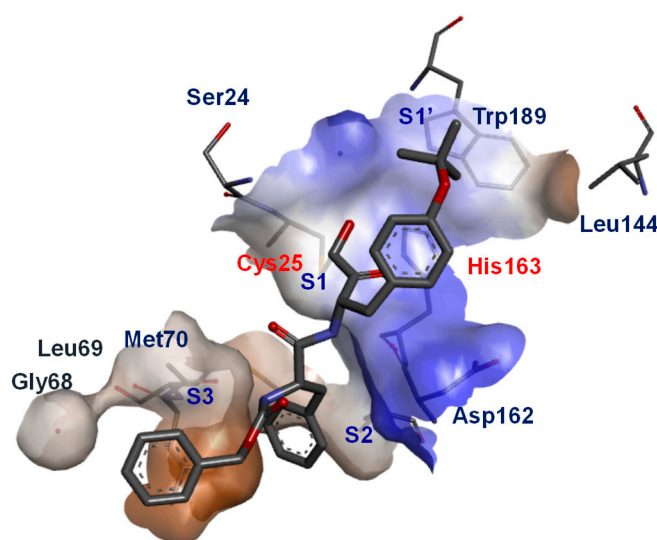
Figure 4. Schematic representation of Z-Phe-Tyr (O-tert-Butyl)-C(O)C(H)O bound within the cathepsin L subsites. The key subsite-forming amino acid residues with their corresponding numbers are shown in blue whereas the catalytic diad forming residues are depicted in red. (The authors used PDB:3OF8 file to construct the figure [96]).

3.1. Epoxysuccinates

Epoxysuccinate inhibitors have historically played a crucial role in deciphering the cysteine protease biology. This class of compounds contains an epoxide ring as an electrophilic warhead that traps the active site catalytic cysteine residue of the protein. E-64 (L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane) (Entry 1; Table 2), perhaps the most studied commercially available universal cathepsin inhibitor of this class, showed only a marginal selectivity towards cathepsin L compared to cathepsin B and a moderate 24-fold selectivity over cathepsin H, as reported by Barrett et al. [99]. Although they developed several synthetic analogs of E-64 with improved potency, they lacked desirable selectivity toward cathepsin L. They proposed the probable binding orientation of E-64 that follows a non-substrate-like orientation, i.e., it occupies only non-prime sites (Figure 3) at the enzyme pocket.

In another study, Gour-Salin et al. reported a number of epoxysuccinyl amino acid benzyl esters in which they systematically varied the amino acid attached to the epoxide ring. This was intended to investigate its effect in determining selectivity toward cathepsin L or S [100]. Their results surprisingly indicate that the specificity of these analogs did not follow the trend, generally observed for substrate, possibly due to E-64 like binding orientation at the enzyme pocket. Among all synthesized compounds,

HO-Eps-Arg-OBzl (Entry 2, Table 2) showed a significant 89-fold selectivity for cathepsin L over cathepsin S.

Notably in a seminal study, Katanuma et al. developed several cathepsin L-selective E-64 analogs (Cathepsin L Inhibitor by Katunuma: CLIK) with the help of computational modeling based on the stereo-structure [101]. Three of the developed CLIK inhibitors were hydrolytically stable and showed highly selective inhibition for hepatic cathepsin L in vivo. Further, they elucidated the inhibition mechanism of this class of compounds based on the crystal structure of papain-CLIK148 (Entry 3, Table 2) complex [102]. This crystal structure revealed that CLIK148, unlike E-64, binds to both prime and non-prime sites of the active site pocket. Notably, the specificity toward cathepsin L was attributed to the existence of phenylalanine residue at the S2 site, and a hydrophobic interaction mediated by N-terminal pyridine ring (Figure 3).

3.2. Peptidyldiazomethane and Peptidylchloromethane

The other classes of alkylating agents that initially played an important role in deciphering the mechanistic aspects of cathepsin inhibition were peptidyldiazomethane (Entry 4, Table 2) and peptidylchloromethane (Entry 5, Table 2). Crawford et al. reported highly potent inhibitors from these two chemotypes that spans the active site and shows significant selectivity improvement over other cysteine proteases [103]. However, these classes of inhibitors suffer from stability issues and have found limited utility in in vivo assays.

3.3. Peptidylhydroxylamine and Peptidylhydroxamates

Peptidylhydroxylamines were first introduced as mechanism-based inhibitors of serine and cysteine proteinases [104–107]. Bromme et al. adapted this scaffold and prepared a library of *N*-peptidyl-*O*-acyl hydroxylamines which exhibited rapid and selective inactivation of several lysosomal cysteine proteinases [108]. This class of inhibitors occupied the active site of the enzyme and irreversibly inactivated cysteine cathepsins as the free enzyme activity was not recovered when the enzyme-inhibitor complex was exposed to exhaustive ultrafiltration (up to an enzyme/free inhibitor ratio of <1:0.05) or chromatography on Sephadex G-10. Among developed inhibitors, Z-Phe-Phe-NHO-MA (Entry 6, Table 2) inhibited cathepsin L with most potency and showed significant selectivity, 58-fold and 436-fold, over cathepsin S and cathepsin B, respectively. Interestingly the stability and efficacy of this class of inhibitors were determined by the nature of substitutions on hydroxylamine oxygen as their electron-withdrawing tendencies showed a positive correlation with inactivation kinetics. Although this class of inhibitors exhibited desirable traits with respect to both potency and selectivity, they suffered majorly from aqueous stability issue; half-lives ($t_{1/2}$) only in the range of 45209 min in aqueous solution. In another study, Bromme et al. developed a series of *N*-peptidyl-*O*-acyl hydroxamates with lysine in P1 position with improved inhibitory profile for cysteine proteases over their serine counterparts [109]. The maximum inhibition was observed by Z-Phe-Lys-NHO-NBz (Entry 8, Table 2), with 7-fold selectivity over cathepsin S, and >100-fold selectivity over cathepsin B. The authors postulated that the active site Cys residue attacks the carbonyl of the hydroxamate and forms a tetrahedral intermediate, whereas the nitrogen of the hydroxamate, which primarily remains deprotonated (pKa < 5), presumably engages in an electrostatic interaction with the active site His159 residue (Figure 5).

Table 2. Reported small molecule inhibitory agents of human cathepsin L, their mechanism of inhibition, inhibitory parameters and selectivity profile, and their demonstrated utilities.
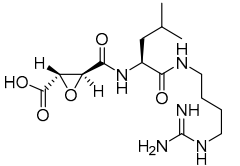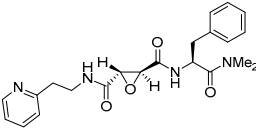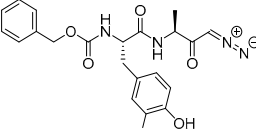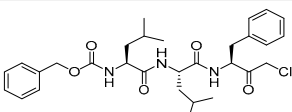
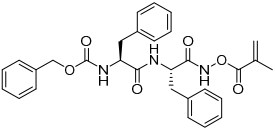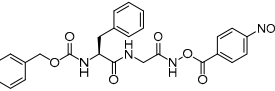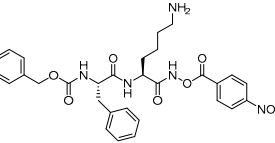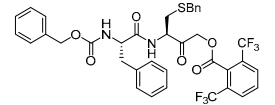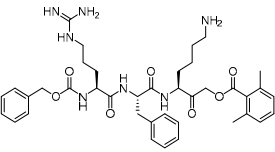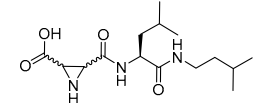| #	Chemical Structure	Inhibitor Class	Inhibitor Efficacy (Cathepsin L)	Selectivity Factor [#]	Mechanism of Inhibition (Demonstrated Utilities)	References
1		Epoxy Succinyl	2nd Order ($M^{-1} s^{-1}$) 96250	Cat B: 1.1 Cat H: 24	Irreversible (In vitro)	Barrett et al., 1982 [99]
2		Epoxy Succinyl	2nd Order ($10^3 M^{-1} s^{-1}$) 73	Cat S: 89	Irreversible (In vitro)	Gour-Salin et al., 1994 [111]
3		Epoxy Succinyl	10^{-6} M (100% inhibition)	Cat B: NI Cat S: 3 Cat K: NI	Irreversible (In vitro cellular)	Katunuma et al.,1999 [101]
4		Peptidyldiazomethane	2nd Order ($M^{-1} s^{-1}$) 1128000	Cat B: 41 Calpain: NI	Irreversible (In vitro)	Crawford et al., 1988 [103]
5		Peptidylchloromethane (chloromethylketone)	2nd Order ($M^{-1} s^{-1}$) 21500000	Cat B: 113 Calpain: <200	Irreversible (In vitro)	Crawford et al., 1988 [103]

Table 2. *Cont.*

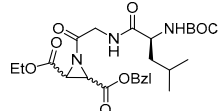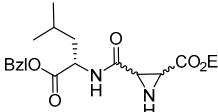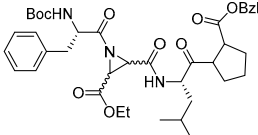| #	Chemical Structure	Inhibitor Class	Inhibitor Efficacy (Cathepsin L)	Selectivity Factor [#]	Mechanism of Inhibition (Demonstrated Utilities)	References
6		N-peptidyl-O-acyl hydroxylamines	2nd Order ($M^{-1} s^{-1}$) 1222000	Cat B: 436 Cat S: 58 Cat H: ND	Irreversible (In vitro)	Bromme et al., 1989 [108]
7		N-peptidyl-O-carbamoyl amino acid hydroxamates	2nd Order ($M^{-1} s^{-1}$) 931800	Cat B: 59 Cat S: 10 Papain: 184	Irreversible (In vitro)	Bromme et al., 1993 [110]
8		N-peptidyl-O-acyl hydroxamates	2nd Order ($M^{-1} s^{-1}$) 3538000	Cat S: 7 Cat B: 101 Cat H: 4655 SP [a]: II	Irreversible (In vitro)	Bromme et al., 1993 [109]
9		Peptidyl (acyloxy)methanes	2nd Order ($M^{-1} s^{-1}$) 10700000	Cat S: 7 Cat B: 4	Irreversible (In vitro)	Krantz, 1994 [112]
10		Peptidyl acyloxymethyl ketone	2nd Order ($M^{-1} s^{-1}$) 875000	Cat B: 11	Irreversible (In vitro)	Torkar et al., 2013 [113]
11		Peptidyl aziridine	2nd Order ($M^{-1} s^{-1}$) (L) ~ 20000 (D) ~ 110000	Cat B ~ 2(L) Cat B ~ 2(D)	Irreversible (In vitro)	Martichonok et al., 1995 [114]

Table 2. *Cont.*

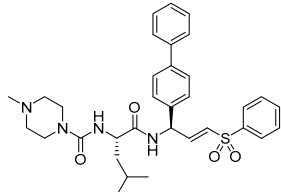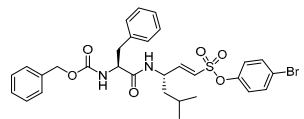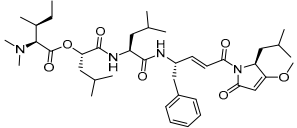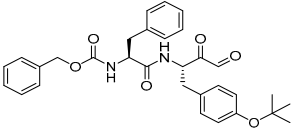| #	Chemical Structure	Inhibitor Class	Inhibitor Efficacy (Cathepsin L)	Selectivity Factor [#]	Mechanism of Inhibition (Demonstrated Utilities)	References
12		N-acylated aziridines (type I)	2nd Order (M^{-1} min^{-1}) (S,S+R,R): 3227	Cat B: 13 Papain: 22	Irreversible (In vitro)	Schirmeister, 1999 [115]
13		N-unsubstituted aziridines (type II)	2nd Order (M^{-1} min^{-1}) (R,R): 16261 (S,S): 3130	Cat B: 10(R,R) Cat B: 76(S,S)	Irreversible (In vitro)	Schirmeister, 1999 [115]
14		N-acylated bispeptidyl aziridines (type III)	2nd Order (M^{-1} min^{-1}) (R,R): 5896 (S,S): 1210	Cat B: 3(R,R) Cat B: 3(S,S)	Irreversible (In vitro)	Schirmeister, 1999 [115]
15		Peptidyl aryl vinyl sulfones	IC_{50} (nM) 2.6	Cat B: 404	Irreversible (In vitro)	Mendieta et al., 2010 [116]
16		Peptidyl arylvinylsulfonate	2nd Order (M^{-1} s^{-1}) 4300000	Cat K: 100 Cat B: 44000 Cat S: 13 Cat H: NI Cat D: NI Cat G: NI hPTP1B: NI Trypsin: NI	Irreversible (In vitro cellular)	Dana et al., 2014 [117]

Table 2. *Cont.*
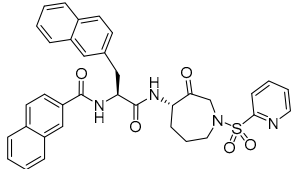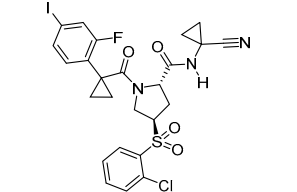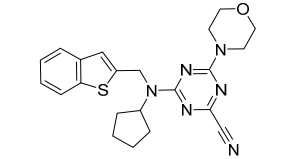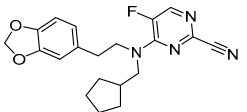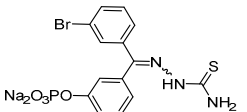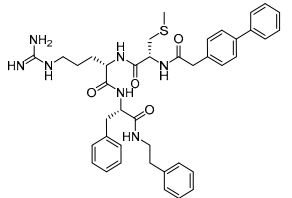
#	Chemical Structure	Inhibitor Class	Inhibitor Efficacy (Cathepsin L)	Selectivity Factor [#]	Mechanism of Inhibition (Demonstrated Utilities)	References
17		Gallinamide A-analog	2nd Order ($M^{-1} s^{-1}$) 8730000	ND	Irreversible (In vitro cellular)	Boudreau et al., 2019 [118]
18		Peptidyl aldehydes	IC_{50} (nM) 0.6	Cat B: 357	Covalent and reversible (In vitro)	Lynas et al., 2000 [119]
19		Azepanone-based Inhibitors	$K_{i,app}$ (nM) 0.43	Cat K>20000 Cat S: 36 Cat B: 349	Covalent and reversible (In vitro)	Marquis et al., 2005 [120]
20		Nitrile group containing inhibitor	IC_{50} (nM) 22	NA	Covalent and reversible (In vitro)	Hardegger et al., 2011 [121]
21		Nitrile group containing inhibitor	Ki (nM) 4	ND	Covalent and reversible (In vitro)	Giroud et al., 2017 [122]

Table 2. *Cont.*

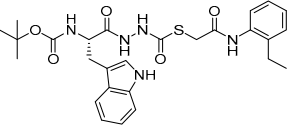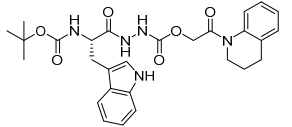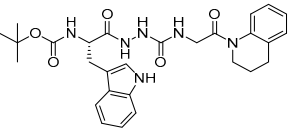| #	Chemical Structure	Inhibitor Class	Inhibitor Efficacy (Cathepsin L)	Selectivity Factor [#]	Mechanism of Inhibition (Demonstrated Utilities)	References
22		Nitrile group containing inhibitor	K_i (nM) 12	ND	Covalent and reversible (In vitro)	Kuhn et al., 2017 [123]
23		Thiosemicarbazone	IC_{50} (nM) 189 (active inhibitor) [b]	Cat B: >50	Covalent and reversible (In vitro cellular)	Parker et al., 2017 [124], Parker et al., 2015 [125] Kumar et al., 2010 [126] Kumar et al., 2010 [127]
24		Cat L propeptide mimic	IC_{50} (nM) 19	Cat K: 310 Cat B: 210	Reversible (In vitro)	Chowdhury et al., 2002 [95] Shenoy et al., 2009 [128]
25		Thiocarbazate	IC_{50} (nM) * 1	Cat V: 11 Cat S: 14 Cat B: 50 Cat K: 137	Reversible (In vitro)	Myers et al., 2008 [129–131]
26		Oxocarbazate	IC_{50} (nM) * 0.4	Cat B: 714	Reversible (In vitro cellular)	Myers et al., 2008 [130] Shah et al., 2010 [132]
27		Aza-peptide	IC_{50} (μM) 3.0	ND	Reversible (In vitro)	Myers et al., 2008 [130]

[#] Fractions were not considered while calculating selectivity factor; Inhibitory efficacy has been reported in whole numbers only; [a]: Serine Proteases; II: Insignificant Inhibition; ND: Not Determined; [b] Active form of inhibitor is devoid of the phosphate group; * Preincubation for 4 h.
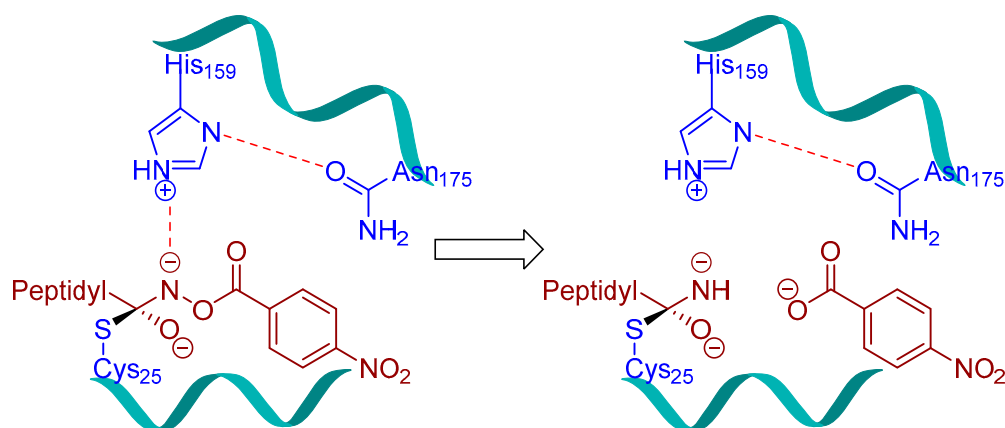
Figure 5. Postulated active-site interactions between papain, a prototypical cysteine protease, and peptidyl hydroxamate inhibitor. Active-site cysteine forms a covalent bond with the carbonyl group of the inhibitor and the negatively charged nitrogen engages in electrostatic interaction with the positively charged histidine residue.

In a follow-up study, Bromme et al. synthesized and tested a series of new inhibitors, with the general formula of Z-Phe-Gly-NHO-CO-Aa (Aa: amino acid), against papain class of enzymes [110]. This class of inhibitors covalently modified active-site Cys residue via sulfenamidation, like in their *N*-peptidyl-*O*-acyl hydroxamate counterparts, as shown by the mass spectrometric analysis. These inhibitors further provided the option of varying the leaving group that targets S2′ site; more hydrophobic substituents were preferred at this position as comparison of k_i/K_i values of the inactivation exhibited the following trend for Aa: Gly < Ala < Val < Leu < Phe < 4-NO$_2$-Ph. The nitrophenyl analog (Entry 7, Table 2) which incur both hydrophobicity and electron-withdrawing property exerted the maximum potency and selectivity for cathepsin L over other tested cysteine proteases.

3.4. Peptidyl Acyloxymethanes/Acyloxymethyl Ketones

Another class of peptidyl inhibitors that spans the active site and utilizes the catalytic machinery of cathepsin L for effective attenuation of enzyme activity is peptidyl acyloxymethanes. Krantz et al. developed a library of inhibitory compounds with a general sequence, Z-Phe-X-CH$_2$OCO-R; here they systematically varied the amino acid residue at P1 (denoted as X) and P1′ (denoted as R) positions [112]. Among the synthesized compounds, Z-Phe-Cys(SBn)-CH$_2$OCO-2,6-(CF$_3$)$_2$-Ph (Entry 9, Table 2) exhibited almost a diffusion-controlled inactivation kinetics toward cathepsin L, however it showed only marginal selectivity over cathepsin B and cathepsin S. By analyzing the structure-activity relationships of the library, the author elucidated the importance of S1′ site in cathepsin L that could potentially be used to harness selectivity among other cysteine cathepsins. In a separate study, Torkar et al. designed a library of peptides for cathepsin L that spanned the active site and attenuated the enzyme activity via a non-covalent interaction [113]. They initially evaluated the compounds' activities against cathepsin L and cathepsin B, and compared their hits against cathepsin K and S. The authors discovered five most selective non-covalent, peptidyl inhibitors of cathepsin L, and transformed them into irreversible inhibitors by strategically appending electrophilic warhead (Figure 6)—acyloxymethyl ketone (AOMK) groups (Entry 10, Table 2). However, the attachment of the AOMK group drastically impacted the selectivity profiles of these inhibitors, suggesting the importance of the adjuvant effect of prime site targeting in determining the efficacy and selectivity of this class of compounds. This class of inhibitors found wide-spread utilities in detecting protease activity and were utilized to develop activity-based probe; this will be discussed in the later sections.
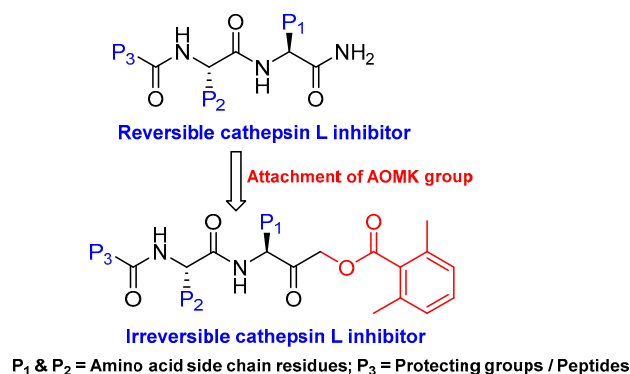
Figure 6. A reversible cathepsin L inhibitor was transformed to an irreversible one by strategically galvanizing an electrophilic warhead.

3.5. Peptidyl Aziridine

A very interesting class of inhibitors, the concept of which was derived from E-64, is peptidyl aziridines. Martichonok et al. developed a series of aziridine derivatives of E-64 and tested them against papain and cathepsin L and B [114]. Contrarily to E-64, in which (L)-diastereomer is more potent than (D)-isomer, aziridine analogs exhibited the opposite trend while still inactivating the enzymes. More importantly, the efficacy of this class of inhibitors was strongly pH-dependent and showed maximal inhibitory potency at pH4; this is attributed to the protonated form of aziridine ring that is more susceptible to nucleophilic attack by catalytic Cys residue. Among the developed library, HO-(*D*)-Az-Leu-NH-iAm (Entry 11, Table 2) analog exhibited maximal inhibitory potency towards cathepsin L with moderate selectivity over cathepsin B; the corresponding L isomer, HO-(*L*)-Az-Leu-NH-iAm (Entry 11) was almost 10-fold less activity while maintaining the similar trend in general. Since only protonated form of aziridine ring undergoes nucleophilic attack, presumably it does not involve water molecule-mediated ring-opening like its epoxide counterpart E-64. Although promising, a full potential of this class of compounds could presumably be achieved only below pH 4 when aziridine ring nitrogen gets completely protonated. This class of inhibitors thus lacks practical utility as the catalytic cysteine of cathepsin L starts to lose its activity below pH 4 and the majority of the cell-based assays are performed mostly at a pH higher than 4.0. In a notable study, Schirmeister et al. developed three different classes of inhibitors with aziridine-2,3-dicarboxylic acid (Azi), an electrophilic warhead, installed at different positions of the peptide chain (Figure 7) [115]. They performed a thorough SAR analysis of this type of inhibitors, all off which exhibited time-dependent irreversible inactivation of cathepsin L with no-recovery of enzyme activity, even after extensive dialysis of the enzyme-inhibitor complex. Among type-I inhibitors, N-acylated aziridines with aziridine as C-terminal amino acid, a mixture of diastereomeric peptides with procathepsin B sequence Leu-Gly-Gly (Entry 12, Table 2), exhibited enhanced inhibition toward cathepsin L. This was attributed to an overall unique folding of cathepsin L with a shallowness of the S2 pocket due to the presence of an additional Met161 residue. Type II class of inhibitors (Figure 7) resemble classic aziridine scaffold, N-unsubstituted aziridines with aziridine as N-terminal amino acid, analogously to E-64 where nitrogen of aziridine remained unsubstituted. Among the type II inhibitors tested, EtO-(*R,R*)-Azi-Leu-OBzl (Entry 13, Table 2) inactivated cathepsin L with higher second-order rate constant than EtO-(*S,S*)-Azi-Leu-OBzl, although the latter showed better selectivity over cathepsin B.

The type III inhibitor class (Figure 7) is comprised of N-acylated bispeptidyl derivatives of aziridine, where aziridine ring rests in the middle of the peptide. BOC-Phe-(*R,R*)-(EtO)-Azi-Leu-Pro-OBzl (Entry 14, Table 2) of this series was 5-fold more potent than the (*S,S*) analog; however both exhibited only a marginal selectivity over cathepsin B with diminished eudysmic ratio. The authors then superimposed and analyzed the structures of certain epoxide and aziridines and postulated that these inhibitors can assume different orientations in the active site while still binding within the enzyme pockets.

This scaffold was further explored by Vicik et al. who extended the previous work and developed a series of compounds in which Boc-(*S*)-Leu-(*S*)-Azy-(*S*,*S*)-Azi(OBn)$_2$ (Figure 8), Type I analog, spanned from S2 to S2′ pocket and inactivated cathepsin L with more than 700-fold selectivity over cathepsin B [133]. This motif was also used for affinity labeling of cathepsin L, which will be discussed in the later sections. These classes of compounds, especially N-unsubstituted aziridinyl peptides and in special cases N-acylated ones, exhibited a high selectivity and potency and provided the premise for the further development of chemical biology tools much needed for functional studies.
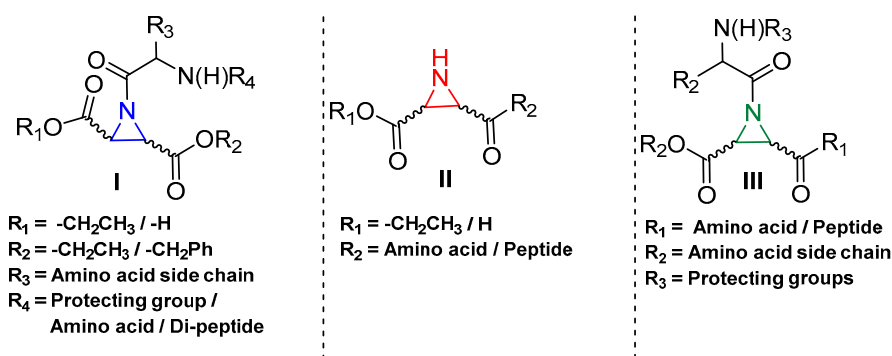
Figure 7. Types of N-substituted and unsubstituted aziridines (color-coded for visual clarity) that bind to the active site of cathepsin L and assume different orientations depending on the favorable interactions incurred by the R-groups.
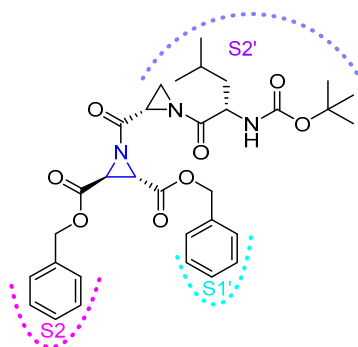
Figure 8. Binding mode of Boc-(*S*)-Leu-(*S*)-Azy-(*S*,*S*)-Azi(OBn)$_2$ in cathepsin L active site, as predicted by docking studies (Vicik et al., *ChemMedChem* **2006**, *1*, 1021–1028).

3.6. Peptidyl Aryl Vinylsulfones

Another promising scaffold that acts as a Michael acceptor and hijacks the catalytic residue of cysteine cathepsins is peptidyl aryl vinylsulfones. This scaffold was first introduced by Palmer et al. as an irreversible inhibitor of cysteine cathepsins [134]. Subsequently, they extended their study by performing an SAR analysis of this class of inhibitors which showed pan-cathepsin inhibition with occasional selectivity towards cathepsin S for certain scaffolds [135]. However, in a separate study by Mendieta et al., the authors developed a structurally novel library of twenty peptidyl 3-aryl vinylsulfones (Figure 9) in which they introduced extensive diversity at the R$_1$ position. Subsequently, they also varied the R$_2$ position while keeping either morpholine or N-methyl piperazine group intact [116]. Docking studies with most active and selective inhibitor (Entry 15, Table 2) of this class revealed that the inhibitor extends from S2 to S2′ sites of cathepsin L and the β-vinylsulfone moiety resides in a close proximity of Cys-25 residue thereby favoring the formation of Michael adduct. The authors postulated that considering the efficacy of peptidyl aryl vinyl sulfones, strong anti-cancer candidates could be harnessed by cultivating this scaffold.

X = O / N-methyl
R₁ = Phenyl / 1-Napthyl / 2-Napthyl / Mesityl / 4'-Biphenyl
R₂ = Benzyl / Iso-butyl
[Michael-acceptor vinylsulfone is shown in red]
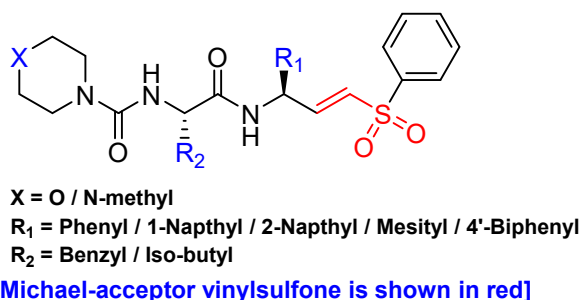
Figure 9. The general structure of the library of peptidyl aryl vinylsulfones where substituents (R_1 and R_2) have been varied to assess the favorable interactions within the cathepsin L active site.

3.7. Peptidyl Aryl Vinylsulfonate

One other very potent class of inhibitors that also includes a Michael acceptor is peptidyl aryl vinylsulfonate esters, a superior Michael acceptor than vinyl sulfone. They served as potent inhibitors of cruzain—a parasitic cysteine protease from *T. cruzi* that is homologous to cathepsin L [136,137]. This scaffold was explored by Dana et al., who determined the superiority of aryl vinylsulfonate ester over aryl vinylsulfone and aryl vinylsulfonamide counterparts towards cathepsin L inhibition [117]. Thus, they synthesized and screened the efficacy of a library of aryl vinylsulfonate ester compounds against cathepsin L; 4-bromo phenyl vinylsulfonate was found to be the champion ligand presumably due to favorable interactions between the 4-bromo phenyl moiety with the prime site residues of cathepsin L. They further designed a hybrid inhibitor, KD-1, (Entry 16, Table 2) by strategically appending the 4-bromophenyl vinylsulfonate moiety as electrophilic warhead to a modestly potent reversible cathepsin L inhibitor [113]; this design was based on the hypothesis that the developed compound will target both the prime site and the non-prime site residues for interaction. KD-1 indeed exhibited almost a diffusion-controlled inactivation kinetics while maintaining an excellent selectivity profile toward cathepsin L (Figure 10). Furthermore, KD-1 was cell-permeable and inhibited the intracellular activity of cathepsin L in human MDA-MB-231 breast cancer cell lines. KD-1 also enhanced the integrity of cell–cell junctions by effectively attenuating the migratory potential of the cells, as demonstrated by the scratch assay. The authors anticipated that this class of inhibitors may find extensive usage in deciphering context-specific cathepsin L biology.
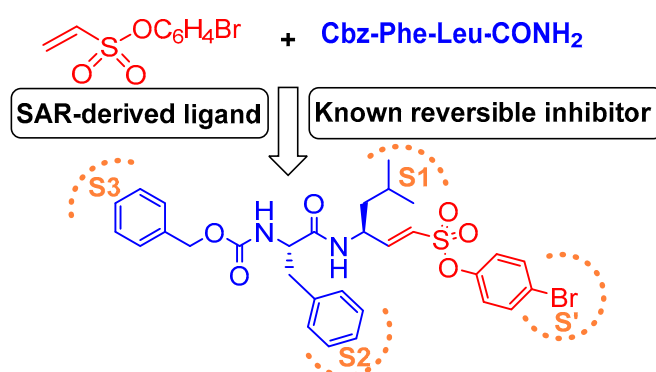
Figure 10. 4-bromophenylvinylsulfonate moiety was tethered to a known reversible inhibitor of only modest potency to achieve a highly potent and selective irreversible peptidyl cathepsin L-selective inhibitor, KD-1; this inhibitor likely binds to both prime- and non-prime sites of the enzyme.

3.8. Gallinamide A Analogs

Recently, another interesting discovery that provided a wealth of information on cathepsin L-inhibitor interactions came from marine cyanobacterial extracts. Miller et al. first reported gallinamide A as potent irreversible inhibitor of cathepsin L with an IC₅₀ value of 5 nM and a 28-

to 320-fold greater selectivity over cathepsin V and B, respectively [138]. They further performed molecular docking and molecular dynamics simulations and learned that the peptidyl backbone of the inhibitor spans the active site whereas the side chains engage in favorable interactions with different active site pockets, placing the Michael acceptor enamide in close proximity to the catalytic Cys residue. In a follow-up study, Boudreau et al. performed molecular docking studies to predict the potential modifications of a gallinamide A scaffold that would harness favorable enzyme–inhibitor interactions and enable the development of compounds with improved inhibitory efficacy [118]. They synthesized a panel of compounds by retaining gallinamide A and only varying the amino acids at P1, P1′, and P2′ positions. (Figure 11). This led to the discovery of the most potent analog of this series (Entry 17, Table 2) with sub-nanomolar IC_{50} value (94 pM) and fast time dependent inactivation kinetics, suggesting an improved binding and reactivity of the inhibitor with the enzyme active site. The authors found that this class of compounds effectively inactivated cruzain, a homologous cysteine protease from *T. cruzi*, using cell-based assay. Gallinamide A and its analogs thus provide a remarkable inhibitory scaffold that could potentially be harnessed to build selective enzyme inhibitors for a variety of therapeutic applications.
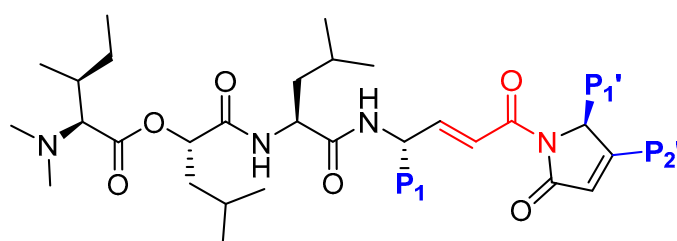
P_1 = L-Ala / L-Phe / D-Phe; P_1′ = L-Ala / L-Phe / D-Phe; P_2′ = -OCH$_3$ / -O(CH$_2$)$_4$CCH

Figure 11. Gallinamide was reconstructed with altered amino acid sequence at P1, P1′, and P2′ positions. The acrylic group, the Michael acceptor, is shown in red.

3.9. Peptidyl Aldehydes

One classical inhibitor that has been used over a long period of time to dissect cysteine proteinase activity is Leupeptin, a microbial product. Leupeptin is a peptidyl aldehyde that occupies the active site cleft of cysteine cathepsins and forms a thiohemiacetal intermediate by trapping catalytic cysteine residue; this complex hydrolyzes over time, thus showing the covalent and reversible nature of the inhibitor (Figure 12a). Leupeptin unfortunately suffers from non-specific inhibition of both serine and cysteine proteinases, thus making it unfavorable for clinical usage and chemical biology applications. This issue was, however, addressed by Woo et al. who designed and synthesized six peptidyl aldehyde analogs that were more potent than Leupeptin (IC_{50} = 70.3 nM) and exhibited improved selectivity towards cathepsin L over cathepsin B and calpain II [139]. The most potent cathepsin L inhibitor of this series was Z-Phe-Phe-H (IC_{50} = 0.74 nM) (Figure 12b) that showed more than 90-fold selectivity over cathepsin B. Interestingly; their data demonstrated the importance of aromatic amino acids, such as phenylalanine and tyrosine, at the P1 position in determining the potency and selectivity towards cathepsin L; O-alkylation of tyrosine group diminishes the inhibitory efficiency as in Z-Phe-Tyr(Bu)-H (IC_{50}: 6.96 nM). In a follow-up publication, they further tested the efficacy of Z-Phe-Tyr-H (IC_{50}: 0.85 nM, 100-fold selective over cathepsin B) (Figure 12b) in vitro and in vivo [140]. This compound effectively inhibited parathyroid hormone-stimulated osteoclastic bone resorption in pit formation assays, and suppressed bone weight loss of ovariectomized mouse in a dose-dependent manner when administered intraperitoneally.
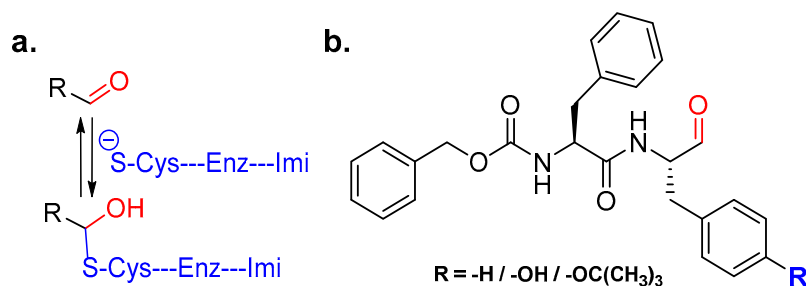
Figure 12. (**a**) Proposed mode of cathepsin L inactivation by aldehyde inhibitors. (**b**) Chemical structure of inhibitors, as discussed in text, are shown here.

In another interesting study, Yasuma et al. developed a library of compounds by varying the amino acid substituents at P1, P2, P3 position of the inhibitor and carried out a thorough SAR study [21]. Their study revealed that the configuration of the stereogenic center (S-configuration is favored over R-configuration) at the P1 position, and not the steric factor, was key to the inhibitory efficacy. Apparently, the substituent at P1 position does not interact with S1 position residues; rather, proper stereogenicity allows the placement of the inhibitor in vicinity of the catalytic cysteine residue for interaction. S2 subsite of cathepsin L, on the other hand, preferred a hydrophobic and moderate-size group; α-branched alkyl chains but not the bulkier groups like phenylalanine was favorable. Further, the S3 subsite showed a preference for hydrophobic and bulky moieties such as 1- and 2-naphthalenylsulfonyl substituents. Among synthesized compound, *N*-(1-naphthalenylsulfonyl-ʟ-isoleucyl-ʟ-tryptophanal (IC_{50} = 1.9 nM, 789-fold selective over cathepsin B; Figure 13) attenuated the release of Ca^{2+} and hydroxyproline from bone in an in vitro bone culture system and further restricted bone loss in ovariectomized mice dosed orally.
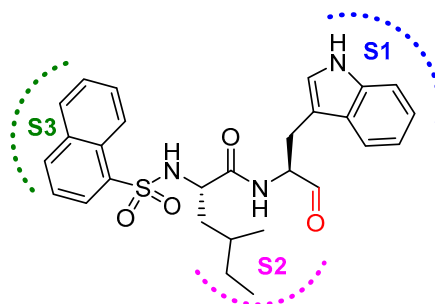
Figure 13. *N*-(1-naphthalenylsulfonyl-ʟ-isoleucyl-ʟ-tryptophanal orients itself into the active site of cathepsin L and finds favorable interactions within the S1, S2, and S3 pockets of the enzyme.

A further modification of this scaffold was reported by Lynas et al. [119] Here, authors designed and developed di- and tri-peptidyl α-keto-β-aldehydes, based on substrate and inhibitor specificity profiles of cathepsin L. The compound Z-Phe-Tyr(OBut)-COCHO (Entry 18, Table 2) turned out as highly potent and selective inhibitor of cathepsin L with K_i value of 0.6 nM. This molecule was further adapted by Shenoy et al. to assess the structural basis for cathepsin L inhibition [96]. In their study, the authors crystallized the glyoxal inhibitor with cathepsin L; the β-aldehyde forms a tetrahedral thiohemiacetal and α-keto oxygen atom is stabilized by the oxyanion hole. The Tyr(OBut) group was found to occupy S1 site while phenyl and carboxybenzyl groups occupied S2 and S3 sites, respectively. This class of inhibitors has successfully been deployed in the functional biology of cathepsin L.

3.10. Azepanone-based Inhibitors

Azepanone-based compounds were first reported as orally bioavailable and extremely potent inhibitors of cathepsin K, as shown by the pharmacokinetic studies in the rat [141]. Marquis et al. subsequently adopted the template and extended their work to acquire a selective inhibitor of cathepsin

L with similar potency [120]. This class of inhibitors are armored with keto functional group that act as an electrophilic warhead and traps cysteine proteases by forming a transient covalent bond with the active-site Cys residue, rendering inactivated enzyme. The authors initiated their work by scrupulously studying the cathepsin K-inhibitor complex that revealed the influence of P2 and P3 substituents of the inhibitor in determining the efficacy and selectivity profile of the compound. Based on these observations, they designed and synthesized a series of compound and secured a highly potent cathepsin L inhibitor ($K_{i,app}$: 0.43 nM; Entry 19, Table 2) that exerted remarkable selectivity over cathepsin K and fairly modest selectivity over both cathepsin B and S. Interestingly, SAR showed that replacement of P2 leucine and P3 benzofuran of cathepsin K inhibitor with bulkier hydrophobic aromatic groups yielded an improved potency and the selectivity towards cathepsin L (Figure 14). Molecular docking studies further supported this observation as cathepsin K was found to have a shallower S2 pocket than cathepsin L, thus incorporation of bulkier napthyl group at P2 position favored cathepsin L inhibition but not cathepsin K. On the other hand, inclusion of another napthyl group at P3 position promoted a steric clash rather than furthering the desired hydrophobic interactions within the S3 pocket of cathepsin K, thus incurring a better selectivity profile towards cathepsin L over cathepsin K. This template has proven to be an important tool to study cysteine cathepsins as it has been further extended to achieve potent cathepsin S-selective inhibitor with cellular activity [142].
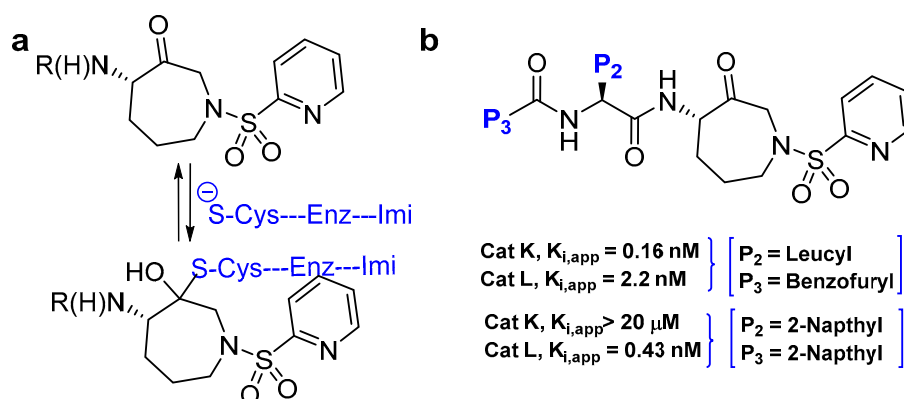
Figure 14. (**a**) Pictorial representation of mechanism of azepinone-mediated cathepsin L inhibition. (**b**) Generic structure of azepinone inhibitors. The inclusion of bulky aromatic groups at P2 and P3 positions leads to cathepsin L inhibitors with improved potency and selectivity.

3.11. Nitrile-Containing Inhibitors

Nitrile group containing inhibitors have been widely recognized as covalent and reversible inhibitors of a certain class of enzymes that depend on cysteine-mediated nucleophilic attack for catalysis; the nitrile residue traps the sulfur and forms a thioimidate bond (Figure 15) that hydrolyzes over the time rendering free enzyme. Odanacatib is one of the prime examples of this class of compounds that has been evaluated as a clinical agent, although with limited success [143,144]. Because of nitrile's tunable target engagement nature, this scaffold has been adapted to target other relevant enzymes, including cathepsin. Hardegger et al. utilized nitrile warhead and examined the effect of halogen bonding in protein–ligand interactions [121]. They developed a series of compounds and performed a thorough SAR analysis in which the nitrile electrophile faced towards S1 site and trapped the catalytic cysteine. In the developed analogs, the substituents that occupied the S3 site were systematically varied by strategically altering the substituents at the para-position of the phenyl group (Figure 16). The authors observed an improvement in inhibition profile with the placement of halogen at the para-position of phenyl ring which followed a trend Cl < Br < I (Entry 20, Table 2), with the F substituent being an outlier. Further analysis of the enzyme-inhibitor co-crystal structures revealed that halogen at the para position of the phenyl ring suitably interacted with Gly61 at the S3 site; fluorine analog pointed away to avoid the electronic repulsion from the oxygen lone pairs

of Gly61. The authors have also performed computational analysis which taken together with the crystal data suggests O· X-C angle and the distance between the interacting atoms primarily influenced the protein-ligand interaction. This work provides an important roadmap for developing improved chemical biology tools where a halogen-protein interaction has successfully been utilized [97].
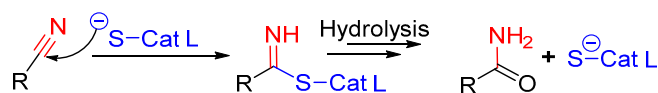
Figure 15. Nitrile group-containing inhibitors forms hydrolytically an unstable thioimidate bond with the catalytic cysteine residue of the enzyme.

X = -F / -Cl / -Br / -I / -CF₃ / -H / -CH₃
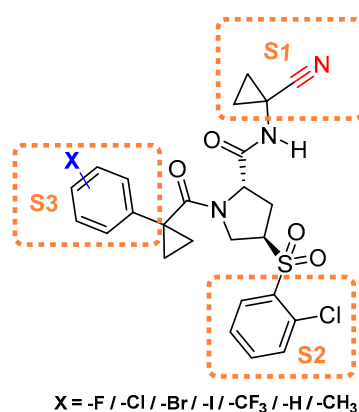
Figure 16. Substituted aryl groups were screened and evaluated to harness favorable halogen-protein interaction by targeting S3 pocket of cathepsin L enzyme.

To examine what effect amide···heteroarene π-stacking interactions may have on chalcogen bonding in the S3 pocket of cathepsin L, Giroud et al. utilized triazine-nitrile scaffold [145]. The authors synthesized a diverse set of triazine-nitrile compounds with a diversified heteroarenes targeting S3 pocket; the S1 and S2 substituents were kept constant. Among the developed compound library, 2-benzothienyl analog (Entry 21, Table 2) exhibited maximum inhibitory potential; 2-benzofuranyl, 2-benzothiazolyl, and 2-imidazopyridinyl, which are of similar geometry, also followed a similar inhibitory pattern (Figure 17). Molecular modelling based on co-crystal structures showed favorable chalcogen interaction to the backbone carbonyl of Asn66 (d(S···O = CAsn66) = 3.5 Å and the angle α(OAsn66···S-C) = 158°) at the S3 pocket; this was further supported by a conformational strain analysis, as chalcogen–enzyme interactions compensated for higher torsional strain in the S-containing ligands when compared to the benzofuranyl and imidazopyridinyl ligands. Their study demonstrated the importance of both intermolecular interactions and conformational strain in assessing the effect of heterobicyclic ligands at the S3 pocket that could be potentially be utilized to develop cathepsin L selective inhibitors.

In a subsequent study, Kuhn et al. systematically compared the effectiveness of four different approaches: (a) selection by a medicinal chemist (b) manual modeling (c) docking followed by manual filtering, and (d) free energy calculations (FEP). This systematic protocol enabled them to prioritize building blocks for effective targeting of cathepsin L enzyme [123]. The authors developed a series of 36 analogs by varying only S2 substituents and keeping S1 and S3 fixed (Figure 18). After analyzing the affinity by enzyme kinetics, they found that the FEP method was superior over other well-established methodologies; this method not only predicted the most relevant ligands but also identified the topological requirements of the substituents for a more effective engagement in the S2 pocket. Among the developed compounds, cyclopentylmethyl substituent in the S2 pocket (Entry 22, Table 2) incurred the most favorable interaction as it optimally filled the front part of the pocket. This strategy certainly provided an edge over other conventional methodologies in predicting the optimal ligands for the S2

pocket targeting. These findings could benefit the ongoing effort of achieving a suitable therapeutic candidate for cathepsin L enzyme.
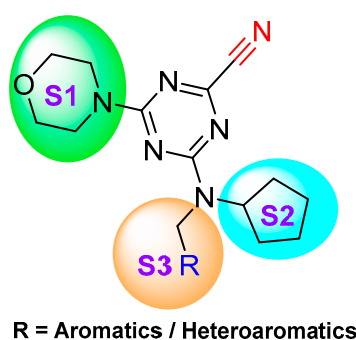
Figure 17. Triazine nitrile scaffold was utilized with varying substituents that explored the S3 pocket on the enzyme for cathepsin L inhibition.
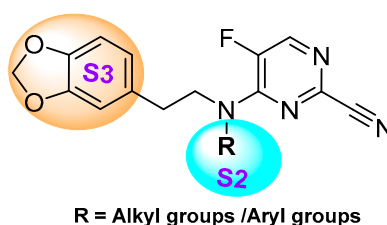
Figure 18. Diverse set of pyrimidine-based nitrile Inhibitors, with ligands targeting the S2 pocket of cathepsin L for inhibition.

The readers are encouraged to refer to studies by Falgueyret et al. and Asaad et al. where, although not selective, nitrile-containing inhibitors targeting cathepsin L have also successfully been reported [146,147].

3.12. Thiosemicarbazone

The thiosemicarbazone moiety was first recognized as a relevant covalent and reversible warhead of cathepsin L homologous enzyme cruzain, a protease from *T. Cruzi*. The mechanism of inactivaction involves the formation of a transient covalent bond with the catalytic Cys residue (Figure 19) [148]. In an interesting study, Kishore Kumar et al. first utilized this idea and synthesized a small library of compounds in which the most active class of inhibitors were comprised of one meta-bromo substituted aryl ring along with another one with optimally substituted functionalities [127]. The inhibitor places itself in the active site cleft of cathepsin L where meta-bromo substituted aryl ring occupies the S2 site and thiosemicarbazone motif lies near the active site cysteine.
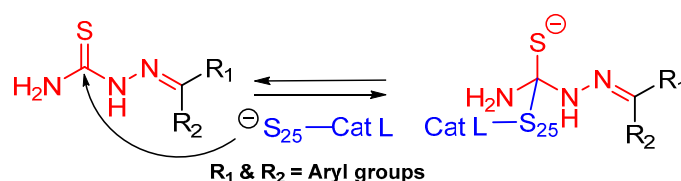
Figure 19. Catalytic cysteine residue of cathepsin L forms a covalent and reversible bond with the thiosemicarbazone group rendering the inhibited enzyme.

However, when the motif was extended to capture S1' site interaction by placing the aryl/alkyl group at the terminal nitrogen of thiosemicarbazone, the inhibitory potency was completely diminished. Overall, this class of inhibitors showed a good selectivity over cathepsin B and exhibited low cytotoxicity

when tested on human cancer cell lines. In follow up studies, Kishore et al. and Parker et al. further expanded the scope of thiosemicarbazone scaffold and developed diversely functionalized analogs that exhibited an enhanced inhibitory potency and promising cellular activities while still retaining the selectivity over cathepsin B [125,126]. In their latest study, Parker et al. strategically transformed an active inhibitor with limited aqueous solubility into a water-soluble prodrug (Entry 23, Table 2), by phosphorylation of phenolic hydroxy group; this group was readily hydrolyzable by alkaline phosphatases, rendering the active pharmacophore [124]. The phosphate prodrug exhibited a remarkable 600-fold increase in solubility over the parent drug and did not disintegrate in aqueous solution, even after prolonged exposure at the physiological temperature. Furthermore, this compound did not show any significant cytotoxicity on normal primary HUVECs cells in comparison to other FDA-approved cytotoxic drugs, Doxorubicin and Paclitaxel. This prodrug thus far has shown promise to be a desirable clinical candidate and the authors have proposed to evaluate its in vivo efficacy in a preclinical setup.

3.13. Propeptide Mimics

As noted earlier, cathepsin L, like other cysteine cathepsins, contains an inhibitory propeptide domain that spans the active site of the enzyme in the inverse direction to the regular substrate binding mode. Chowdhury et al., in their seminal study, exploited this concept by examining the effect of a series of synthesized tripeptidyl compounds that mimicked cathepsin L inhibitory propeptide [95]. Importantly, the developed tripeptidyl motifs also exhibited nanomolar potency; however, a moderate truncation of the full-length propeptide drastically lost all activities [47]. Notably, while the full-length propeptide showed only 2-fold selectivity over cathepsin K, the most potent analog of this series (Entry 24, Table 2) demonstrated a far-improved selectivity (310-fold). The authors further investigated the binding mode of this class of inhibitors by means of co-crystal structure and molecular modeling. This revealed that (a) arginine residue of the inhibitor occupied the S1 pocket, (b) phenyl alanine residue found favorable hydrophobic interactions within the S2 pocket, (c) 2-phenylethyl group pointed toward S3 pocket, (d) the methionine residue showed optimal interaction within S1′ pocket, and (e) the biphenyl acetyl group extended to the S3′ pocket for favorable interactions [95,128]. This class of inhibitors has shown resistance to enzyme-dependent hydrolysis and demonstrates the reversible mode of enzyme inactivation. Overall, this inhibitor class provides a wealth of information on inhibitor binding to cathepsin L and provides a general template for the development of therapeutic candidates for other relevant enzymes as well.

3.14. Thiocarbazate, Oxocarbazate and Azapeptides

In an effort to discover small molecule inhibitors of cathepsin L, Myers et al. performed high throughput screening (HTS) of the NIH Molecular Libraries Small Molecule Repository (MLSMR); they identified 2,5-disubstituted oxadiazoles (Figure 20a) as potent hit compounds [129]. Surprisingly, upon synthesis and purification of the putative inhibitory lead compounds, a complete loss of activity was observed. The authors then investigated the compounds' integrity from NIH MLSMR library by LC-MS; this showed the presence of additional impurities. To trace back the active impurity, the authors hypothesized the presence of impurities resulting from an acid-catalyzed ring-opening reaction of thiocarbazate. The resulting azapeptides was likely the active pharmacophore that inhibited enzyme via acylation of active site Cys; this was validated by synthesis of azapeptides and performing the enzyme assay [149]. The (*S*)-stereoisomer of newly synthesized compound (Entry 25, Table 2; Figure 20b) indeed attenuated the activity of cathepsin L with an IC$_{50}$ value of 56 nM. In follow up studies, the authors further developed a series of compounds with structural diversity and performed a computational analysis to recognize the basis of potent enzyme inhibition [130,150] One of the thiocarbazate analogs developed this way (Figure 20c) showed improved potency over the parent compound. Molecular modeling studies performed with parent compound (Entry 25, Table 2) in complex with papain indicated that indole motif preferably bound to S2 subsite, -NHBoc group engaged

in favorable hydrophobic interactions within the S3 subsite and 2-ethylphenyl anilide extended to S1′ pocket. To further probe the importance of thiocarbazate moiety, the authors synthesized compound containing oxocarbazate (Entry 26, Table 2) and azapeptide (Entry 27, Table 2) motifs. The oxocarbazate showed a fairly improved IC_{50} value (7 nM) towards cathepsin L, whereas the azapeptide was at best only a modest inhibitor (IC_{50} = 3 μM). Consistently, the binding mode of oxocarbazate exerted similarity to that of thiocarbazate when investigated by molecular modeling studies [150].
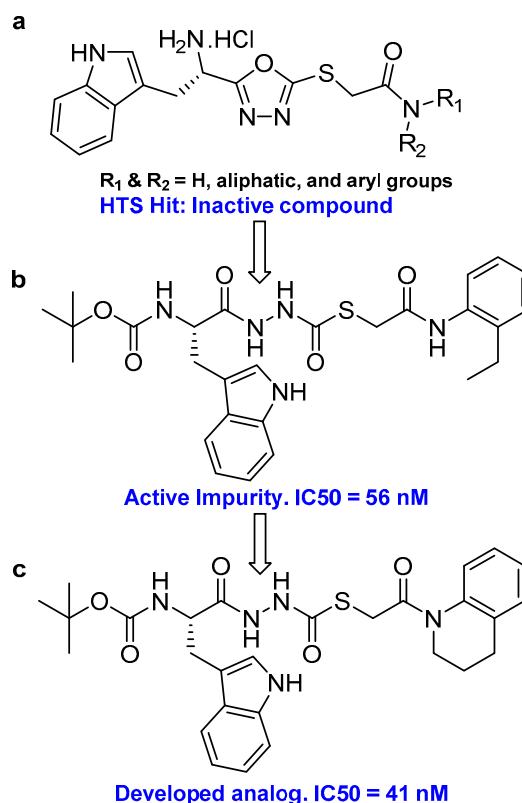
Figure 20. (**a**) Structure of HTS hit that was a false positive. (**b**) Structure of the impurity that turned out to be an active pharmacophore. (**c**) The most potent analog of this series, achieved by SAR analysis.

In a separate study, Shah et al. carried out a thorough enzymatic analysis of the champion thiocarbazate compound (Entry 25, Table 2) that showed a time-dependent improvement in the inhibition profile; the IC_{50} value went down to 1 nM when preincubated with cathepsin L for 4 h [131]. LC-MS and kinetic analysis of enzyme-inhibitor complex (inhibition rate constants: k_{on} = 24,000 $M^{-1}s^{-1}$ and k_{off} = 2.2×10^{-5} s^{-1}, and binding constant: K_i = 0.89 nM) demonstrated a slow-binding kinetics and reversibility of inhibition. The selectivity over other members of the enzyme family was modest. Interestingly, the compound inhibited propagation of malaria parasite *Plasmodium falciparum* [IC_{50} = 15.4 μM], and *Leishmania major* [IC_{50} = 12.5 μM], and did not exhibit any significant toxicity against human aortic endothelial cells and zebrafish. Although thiocarbazate motif showed promise as an inhibitory scaffold, the lack of reasonable stability (it decomposes even in DMSO) and only modest inhibitory activity in cell-based assays probably ceased any further development of the scaffold [129]. The authors also extended their studies to evaluate the potential of oxocarbazate inhibitor that showed an improved IC_{50} value of 0.4 nM upon 4 h preincubation with the enzyme. Like as in the case of thiocarbazate, they performed an enzyme kinetic analysis of the enzyme-inhibitor complex and obtained the following parameters: inhibition rate constants: k_{on} = 153,000 $M^{-1}s^{-1}$ and k_{off} = 4.4×10^{-5} s^{-1}, and binding constant: K_i = 0.29 nM [132]. The inhibitor blocked SARS-CoV (IC_{50} = 273 ± 49 nM) and Ebola virus (IC50 = 193 ± 39 nM) entry into the human embryonic kidney (HEK) 293T cells, a process that utilizes cathepsin L-mediated proteolysis for host cell infection. The oxocarbazate,

when treated with HEK 293T lysate in the presence of DCG-04, an activity-based cysteine cathepsin probe, showed reduced cathepsin L labeling when assessed by a Western-blot analysis; this further corroborated the results obtained from the virus pseudotype infection assay. Overall, oxocarbazate inhibitor not only provided a promising template for further exploitation but also rendered a new direction for intervening SARS and Ebola virus infections.

4. Molecular Probes

Ubiquitous expression of human cathepsin L in most human tissues possesses a significant challenge in targeting this enzyme for therapeutic development. This problem is further exacerbated with recent findings that alternative spliced isoforms could exist at distinct cellular locations (e.g., nucleus, cytosol, and ECM space) [37,38,151]. While several unique functional roles of cathepsin L are known, it has also been reported that some of its function can also be accomplished by other members of the cathepsin family (i.e., functional redundancy); for example, both cathepsin L and B can mediate a mutually compensatory role in the inflammatory response signaling pathways [152]. In this regard, accurate function of cathepsin L must be first determined in different cell types individually before considering significant investment in drug development. Since the function of an enzyme, such as cathepsin L, depends primarily on its activity profile and given that the activity profiles of differentially processed cathepsin L isoforms may be very different, probes capable of reporting accurate activity status in different cell types (and cellular location) are anticipated to advance our understanding if cathepsin L biology. Over the years, the concept of Activity-Based Probes (ABPs) has emerged as a valuable chemical biology tool for monitoring the enzyme activity (not just the expression profile alone) in cells at the proteome levels [153–156].

In most cases, existing covalent inhibitors containing a recognition motif are adopted and transformed to ABPs by conjugating optimal detection modalities; these include fluorescent probes, affinity labels, radiotracer, and many others (Figure 21). Indeed, the use of ABPs has rather established cathepsins as key diagnostic marker for various disease conditions, and have even enabled optical surgical navigation, leading to an improved surgical precision [157,158]. In the following sections, we discussed cathepsin L-selective molecular probes that have been developed and utilized for monitoring its activity. A thorough overview of cathepsin probe development for imaging purpose could be found elsewhere [10,159–161].
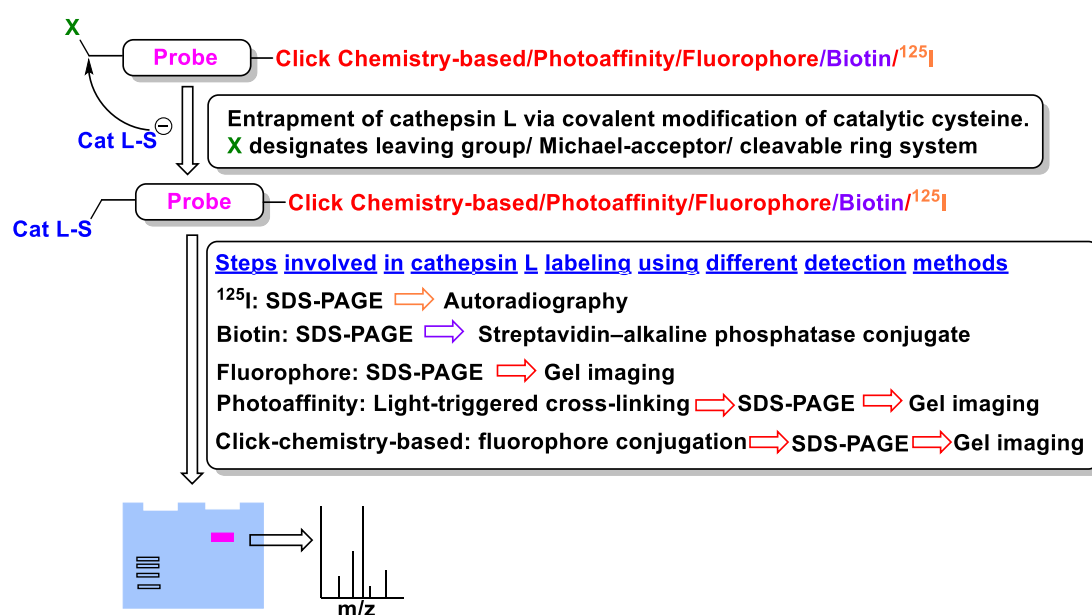
Figure 21. Schematic representation of enzyme-labeling and proteome-wide activity profiling studies using probes for selective detection of active and functional cathepsin L enzyme.

4.1. Radio-Labelled

Radio-labeled inhibitors have long been used as a primary mode for detecting active cysteine proteases both in vitro and in vivo. Docherty et al. first used a chloromethyl inhibitor containing radioactive iodine, ^{125}I-Tyr-Ala-Lys-ArgCH$_2$Cl, and detected cathepsin B in crude granule fraction of islet cells [162]. In their follow up study, they also presumably identified cathepsin L in insulin secretory granule using the same radio-isotopically labeled inhibitor [163]. The mechanism of detection involves covalent modification of the target protein that shows up as a distinct band upon performing autoradiography. This chloromethyl containing inhibitor turned out to be non-selective due to its reactivity towards trypsin, a serine protease. This was followed by the discovery of a radioactive-peptidyldiazomethane compound (Entry P1, Table 3), a selective cysteine proteinase inhibitor [164]. Mason et al. adopted the peptidyldiazomethane scaffold, which potently inhibited both cathepsin L and B. This scaffold showed improved inhibition profile upon iodination, as demonstrated by Crawford et al. [103] This inhibitory agent was then transformed to a radio-labeled probe via incorporation of [125]I [164]. The developed probe was utilized to detect both cathepsin L and B in Kirsten-virus-transformed KNIH 3T3 cells. The incubation of cellular extracts with P1 followed by gel electrophoresis showed the presence of two protein bands at 30 and 23 kDa, showing two active forms of cathepsin L. Active cathepsin B was also detected at around 33–35 kDa. Interestingly, pulse-chase experiments with [35S]methionine-labeled proteins only detected two separate bands at 36 kDa and 39 kDa, which correspond to the intracellular inactive precursors of cathepsin L and B respectively. This indicated that the inactive precursor proteins did not react with P1, demonstrating its unique ability to quantify only active protein. Further, P1 was utilized to probe active cathepsin L and B in different human tissues as well as in lysosomes and whole cells [165,166]. In a follow-up study, Xing et al. developed Fmoc-[I$_2$]Tyr-Ala-CHN$_2$ (Entry P2, Table 3) that selectively detected cathepsin L and B over cathepsin S, exhibiting a faster rate of inactivation towards cathepsin L [167]. The developed compound successfully probed the amount of active cathepsin L and B in different cell-lines; two unknown proteins also got labeled in certain cases. Overall, these probes enabled the detection of active cathepsin enzymes with their cellular location, thereby advancing the knowledge of cathepsin L biology.

Table 3. Reported probes of cathepsin L, their detection module, efficacy, and utilities.
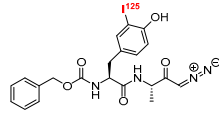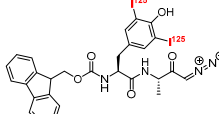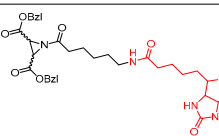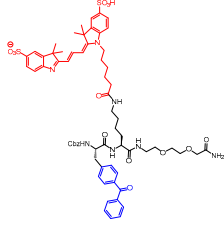
#	Chemical Structure	Probe Class	Efficacy (Cathepsin L)	Selectivity Factor	Mechanism of Probe Action (Demonstrated Utilities)	References
P1		Radio-labelled	2nd order rate constant (M^{-1}s^{-1}) 240000	Cat B: 23	Covalent and irreversible (In vitro Cellular)	Mason et al., 1989 [165] Wilcox et al., 1992 [166]
P2		Radio-labelled	2nd order rate constant (M^{-1}s^{-1}) 60900	Cat B: 2.4 Cat S: NI [#]	Covalent and irreversible (In vitro Cellular)	Xing et al., 1998 [167]
P3		Affinity-labeled	Inhibitor Constant(μM) (S,S) 1.4	Cat B: 36 Papain: 4	Covalent and irreversible (In vitro)	Gelhaus et al., 2004 [168] Vicik et al., 2006 [169]

Table 3. *Cont.*

#	Chemical Structure	Probe Class	Efficacy (Cathepsin L)	Selectivity Factor	Mechanism of Probe Action (Demonstrated Utilities)	References
P4		Photoaffinity-based	Inhibitor Constant (μM) 3.6 (able to detect picomolar amount of protein)	Cat B: 9 Cat K: 3 Cat S: 0.3 Cat V: 0.15 [P4 did show remarkable selectivity in vivo and not in vitro]	Covalent and irreversible (In vitro)	Torkar et al., 2012 [170]
5		Two-photon FRET-based	N.D	N.D	Reversible (In vitro Cellular)	Na et al., 2012 [171]
6		Fluorescent	Cathepsin L specific (some degree of labelling was seen for cat V and B)	Cat V: 847 Cat B: 413 Cat S: 1431 Cat K: >1600	Covalent and irreversible (In vitro Cellular)	Poreba et al., 2018 [172]
7		Clickable and tagless	2nd Order (M^{-1} s^{-1}) 430000	Cat B:9556 Cat K: 29 Cat H: NI Cat D: NI Cat G: NI hPTP1B: NI Trypsin: NI	Covalent and irreversible (In vitro Cellular)	Dana et al., 2019 [173]
8		TOF-based	2nd Order (M^{-1} s^{-1}) 229300 (mix-Gd) 232500 (159-Tb) 227000 (175-Lu)	Selectivity > 150-fold: cat B, V, and S N.I: cat K and Legumain	Covalent and irreversible (In vitro Cellular)	Poreba et al., 2019 [174]

The detecting agents—radionuclide (P1, P2), biotin (P3), fluorophore (P4, P5, P6), clickable acetylene (P7), and lanthanide containing DOTA (P8)—is red-color coded and the photoactivatable benzoyl group (P4) and quencher (P5) is coded in blue. #: No Inhibition.

4.2. Affinity-Based

Gelhaus et al. first developed biotinylated aziridine-2,3-dicarboxylate and demonstrated its anti-plasmodial activity using cell-based studies [168]. Since, the biotinylated compound inhibited plasmodial protease falcipain and cathepsin L, authors suggested that this scaffold could be utilized for the development of cell-permeable, non-radioactive reagents that selectively labels enzymes involved in parasite pathogenicity. Later on, Vicik et al. adopted this motif (Entry P3, Table 3) and developed an affinity label to probe cathepsin L activity [133]. The aziridine analog irreversibly inactivates the enzyme via covalent modification, as discussed previously. The conjugated b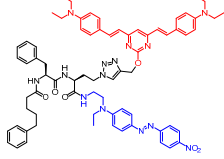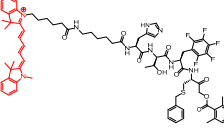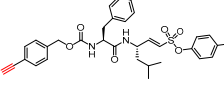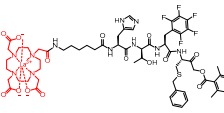iotin moiety is utilized for affinity pull down and target identification. When (*S*,*S*) isomer of the biotinylated probe was incubated with cathepsin L and subjected to gel electrophoresis, electro-transferred to a membrane, and exposed to streptavidin-alkaline phosphatase conjugate, a strong labeling of the enzyme-inhibitor complex was observed. However, when the enzyme was treated with E-64, an active site-directed competitive and irreversible cathepsin inhibitor, prior to incubation with P3, the labeling was diminished, clearly demonstrating that P3 competes for the active site of cathepsin L. In line with this observation, a desthiobiotinylated analog also exhibited the same trend but with reduced labeling due to its weaker binding affinity to streptavidin. Although P3 has a modest binding affinity (K$_i$ = 1.4 μM) to cathepsin L, it exerted a 36-fold selectivity over cathepsin B. Certainly, the affinity labeling technique not only

served as an ABP for cathepsin L but also provides a premise for developing aziridine-based chemical tools for functional proteomics.

4.3. Photoaffinity-Based

Although covalent modifiers of proteins have been vastly exploited as chemical tools for target identification and functional proteomics, photoaffinity probes offer unique mode of action. They bind proteins non-covalently (affinity based solely on non-covalent interactions) first and form a non-selective covalent bond with the closest amino acid residue only upon irradiation. Torkar et al. took advantage of this technique and developed the first photoaffinity-based probe (Entry P3, Table 3) to detect active cathepsin L selectively over other members of the family [170]. The photoaffinity-based probe was designed based on existing peptidyl acetyloxymethyl ketone (AOMK), a known covalent modifier of cysteine proteases. During the design, the AOMK group at the C-terminus was replaced by a short di(ethylene glycol) moiety that increased the aqueous solubility and altered the character of the inhibitor from irreversible to a reversible one. The probe was comprised of a lysine residue that was strategically placed to append fluorescent cyanine-3 (Cy3) group for detection. A photoactivatable benzoylphenylalanine amino acid was placed to accommodate the S2 pocket of cathepsin L. The developed probe detected recombinant cathepsin L upon incubation and subsequent irradiation for 40 min at 365 nm as demonstrated by SDS-PAGE. The protein band became invisible when enzyme was incubated with known cathepsin L inhibitors, E-64 and GB111-NH$_2$, respectively, prior to the probe treatment [175,176]. The probe also showed preferential selectivity towards cathepsin L when compared to other two fluorescent-based probes. The probe P3 exhibited a remarkable selectivity for cathepsin L over all other cathepsins in light-mediated labeling experiments with recombinant proteins. The relative labeling percentages (cathepsin B and K: 4%; cathepsin S: 1%) were insignificant, except for cathepsin V (27%), the closest homolog of cathepsin L, relative to cathepsin L. Interestingly, U87-MG glioma cell extracts did not present any cathepsin L for detection with the P3 probe. However, when the same cell extracts were treated with recombinant cathepsin L added externally and subjected to a labeling experiment, the probe selectively detected the desired protein in complex proteome. The mechanism of protein labeling by P3 was attributed to putative bond formation between benzophenone and non-conserved Met161 at the S2 site of cathepsin L. Although the developed probe lacks cell-penetrability, the authors envisioned that the technique might have diagnostic and prognostic value where cathepsin L overexpression is high, such as in malignant tissues.

4.4. Fluorescence-Based

4.4.1. Two-Photon Based

Although several fluorescent-based molecular probes have been reported for cysteine cathepsins, the superiority of two-photon-based over one-photon based imaging technique inspired Na et al. to develop probes with a better cellular imaging profile [171]. Notably, the two-photon fluorescence imaging technique provides increased tissue penetration depth with reduced photobleaching, and a lower tissue autofluorescence. The authors first fabricated a microarray with 105 different peptidyl aldehydes and screened against GFP-labeled cathepsin L enzyme; this led to the identification of two inhibitory hits that potently inactivated enzyme with IC$_{50}$ values of 14.5 and 20.8 nM. The lead inhibitors were further structurally modified to include (a) a two-photon dye, DL-1 (a 4,6-bis(4-hydroxystyryl)pyrimidine derivative), at P1 position, and (b) Disperse Red 1 dye, a fluorescence quencher, in the place of aldehyde moiety. The resulting imaging probes (Entry P5 in Table 3 is one of such examples) showed a time-dependent increase in fluorescence signals when treated with HepG2 cell lysates, a mammalian liver cancer cell-line known to overexpresses cysteine cathepsins. Enhanced fluorescence signal is due to the release of the quencher upon successful proteolytic cleavage. Furthermore, to assess the suitability of P5 as imaging agents for live cells, HepG2 cells were incubated with the probe and subjected to a live-cell imaging analysis. There were strong fluorescence signals

from endolysomal compartments that disappeared completely when cells were pretreated with E-64, validating the target specificity of the probe. As the probe was developed using cathepsin L as a model enzyme, it likely will lack specificity and perhaps interact with other members of the family. Still, however, this motif could potentially be utilized to develop selective ABPs targeting respective cathepsins and assess their activities in a tissue environment, as proposed by the authors.

4.4.2. One-Photon Based

Activity-based probes with a single photon fluorescent tag have been successfully utilized for the functional analysis of target proteins both in vitro and in vivo. Poreba et al., in their pursuit of developing cathepsin L selective imaging agent, took advantage of this technology [172]. They realized the importance of developing selective cathepsin L substrate that will bind to the active site of the enzyme over other homologous proteins. To do so, they employed Hybrid Combinatorial Substrate Library (HyCoSuL) technology that provides information on optimal chemical space inside the enzyme active sites by strategically scanning diverse peptide library containing both natural and unnatural amino acids. This led to the acquisition of a panel of compounds with desired properties. The potent hits discovered by HyCoSuL technology were further optimized to gain selectivity over other cathepsins while identifying the most efficient substrate based on Michaelis Menten parameters. Unfortunately, when the chosen peptide substrate transformed to an activity-based probe by appending an acyloxymethylketone warhead and a biotin tag, it showed cross-reactivity with cathepsin B. Despite biotinylated probe's somewhat compromised selectivity, the authors still assessed its activity-based labeling profile in HEK293T cells; as expected, cathepsin L labeling was primarily observed with minor amounts of cathepsin B. Further optimization by swapping the biotin and Arg with cyanine-5 and Cys(Bzl) (Entry P6, Table 3) groups was carried out next. The newly developed fluorescent ABP, P6, showed an enhanced selectivity (in comparison to pan-cathepsin probes) against a panel of other recombinant cathepsins (cathepsin V, B, S, and K), as well as cellular extracts derived from HEK293T and MDA-MB-231 cells. The developed probe served as an effective imaging agent for cellular cathepsin L activity in human MDA-MB-231 breast cancer cells when incubated for 8 h. The selectivity of the probe started to recede with longer incubation time. The importance of optimizing the probe concentration and time-course of the reaction was evident from these experiments. Interestingly, P6 only detected active cathepsin L and not procathepsin L; this certainly signifies the effectiveness of the developed probes as an activity-based probe. The authors further examined the colocalization of the probe with both cathepsin L and B in MDA-MB-231 cells by treating the cells with respective cathepsin antibodies and performing a quantitative pixel analysis from a set of fluorescence microscopy images. This supported the previous observation as the highest weighted colocalization coefficient was obtained for cathepsin L and not for cathepsin B. The developed compound certainly harbors the key traits of an effective activity-based probe, as the authors duly envisioned its significance in deciphering cathepsin L biology in the coming years.

4.5. Clickable and Tagless

As noted above, fluorescence-based imaging probes have been successfully developed to gain access to unknown functionalities of cathepsin enzymes. However, the bulkiness and often multiple charges associated with the fluorophore and/or quencher structures on these probes likely render them poorly cell-permeable and reduce their target affinity. To address this issue, Dana et al. adopted a previously reported peptidyl vinylsulfonate inhibitor KD-1—a highly potent, selective, covalent and irreversible inhibitor of cathepsin L discussed in Section 3.7—and tactically appended a small alkynyl group at the *para*-position of the Cbz group [117]. This led to the development of a clickable and tagless activity-based probe (catABP) of cathepsin L that retained the key desirable traits of KD-1; i.e., cell permeability, charge profile, molecular weight, potency and selectivity profile [170]. This strategy eliminated the requirements of including bulky and charged fluorophore/quencher moiety to the probe. One of the inherent advantages of this approach is that the labeling can be performed in live

cellular environment with high efficiency. After cell lysis, the labeled cathepsin L can be quantified by performing click chemistry protocol with a fluorescent azide containing dye, resolving the protein using gel-electrophoresis, and directly scanning the gel for fluorescence signal. As anticipated, the developed KDP-1 probe (Entry P7, Table 3), exhibited rapid inactivation kinetics, retained selectivity for cathepsin L, and labeled recombinant active cathepsin L in an activity-dependent manner; the heat-denatured and E-64 treated cathepsin L showed no labeling when subjected to the identical labeling protocol. A mass-spectrometric analysis of the enzyme-probe complex concluded that the probe was active-site directed, and covalently modified the catalytic Cys residue for inactivation. Since the KDP-1 probe was developed with the intention of capturing active cathepsin L in vivo and in human live cell culture, the probe was tested for its cytotoxicity in MDA-MB-231 cells; no cytotoxicity was observed in these cells, even at a concentration as high as 10 μM. Incubation of MDA-MB-231 cells overexpressing cathepsin L with KDP-1 attenuated the intracellular cathepsin L activity in a dose-dependent manner; this was demonstrated by live-cell imaging and wound healing assay. Finally, KDP-1 also interfered with the hatching process of post-fertilized zebrafish embryos, further validating probe's in vivo activity; notably, cathepsin L serves as the key regulator of the hatching process [177–179]. In conclusion, KDP-1 demonstrated many desirable attributes of a good probe and is anticipated to find extensive applications in probing cathepsin L function in cells from diverse origins.

4.6. Mass Cytometry-Compatible Activity-Based Probes

Although fluorophores-containing molecules have been appreciated as useful imaging probes, commonly used fluorophores often suffer from spectral overlapping that limits the number of targets that can be analyzed concomitantly [180]. To address this issue, Poreba et al. recently developed protease-selective lanthanide-labeled probes compatible with mass cytometry which allows subsequent analysis by both mass and imaging mass cytometry (IMC) [171]. These metal-tagged, time of flight activity-based probes (TOF) allowed them to determine cellular activities and location of three lysosomal proteases. Thus, using cathepsin L, B, and legumain as the model systems, they elegantly crafted an activity-based probe by incorporating (a) a protease-selective peptide sequence for specific enzyme recognition, (b) the acyloxymethylketone as electrophilic warhead to trap the target enzyme, and (c) the dodecanetetraacetic acid (DOTA) for metal chelation that is tethered to the peptide sequence via a linker (Entry P8, Table 3). They further incorporated three different lanthanides, ^{159}Tb, ^{175}Lu, and Gd (a mixture of naturally occurring six isotopes), to validate their approach and to further evaluate the influence of isotopes on enzyme binding specificity. The newly developed probes exerted promising selectivity toward both recombinant proteases and proteases from cancer cell lines, HCT-116 and MDA-MB-231.The authors were able to simultaneous detect the activities of proteases in HCT-116 cells. Moreover, each of the protease-specific probes exerted a similar labelling efficiency in HCT-116 cells, regardless of their metal counterparts, which further reinforces the compatibility of this class of probes as cytometry-labelling agents. They extended their investigation to THP-1 cells, a non-adherent monocyte-like cell line, that expresses both cathepsin B and L but contains very low levels of legumain enzyme. Probe treated THP-1 showed a clear labeling of both cathepsin B and L with no detectable activity of legumain. This finding was consistent with their transcriptional data. Finally, the developed probes not only allowed to detect the activome of cathepsin L, B, and legumain in peripheral blood mononuclear cells (PBMC) but also enabled to categorize NK cells in two distinct populations based on protease activome levels. This strategy, unlike many existing technologies, thus allowed the simultaneous detection of target proteases, thereby providing a more holistic understanding of the activome. This certainly merits the future development of TOF-based probes for multiplexed enzyme activity detection.

5. Final Perspectives

The inhibition of cathepsin L has continue to emerge at the forefront of drug development for several human diseases. Yet, no inhibitory agents targeting cathepsin L have advanced to clinical

trials. While inhibitors of cathepsin B and S are currently being evaluated in clinical trials, recent failure of Odanacatib, a cathepsin K inhibitor for osteoporosis, in late stage clinical trial has made pharmaceutical industries wary of targeting cathepsins. The key challenge remains gaining inhibitor selectivity with respect to the other members of cathepsins and directing them to the targeted cell types for selective functional perturbation. With its ubiquitous expression profile, the function of cathepsin L in individual cell types must be precisely defined first; this is especially important since several isoforms of cathepsin L have been reported in the distinct cellular locations, and their activity and functional profile in individual cell types still remain poorly documented. While recently developed cathepsin inhibitors and probes have significantly advanced our understanding of cathepsin L function in both normal and disease cells, more efforts are needed for the development of isoform-selective reagents for further advancement; perhaps new allosteric modules on cathepsin L enzyme can be explored and exploited for precision targeting. Fortunately, several structural coordinates for cathepsin L enzyme forms, some with a diverse set of complexing ligands and some without, are now available (Table 4). These could aid in the development of isoform-selective inhibitory probes that will enable researchers to assess the context-specific needs of targeting cathepsin L in different cellular states.

In recent years, disease-associated protease activatable prodrugs have gained much recognition and garnered breakthrough therapies in the area of antibody-drug conjugation (ADC) [181]. The majority of the ADCs are constructed by tethering a cytotoxic drug with an antibody via a protease-sensitive module for targeted delivery. Cathepsin B-specific module, for example, has been successfully implemented for this role, which led to the development of FDA approved therapies; for example, Brentuximab vedotin (Adcetris®) for CD30-positive relapsed or refractory Hodgkin's lymphoma. In addition, protease cleavable prodrug strategy has also inspired the development of cathepsin B selective probe and even enabled real-time monitoring of drug release [182,183]. Interestingly, Ueki et al. slightly maneuvered this approach to acquire a prodrug which gets serially activated by histone deacetylase (HDAC) and cathepsin L, and subsequently delivers the cytotoxic payload, puromycin, to cancer cells [184]. This strategy has thus enabled selective targeting of cancer cells, specifically with high HDAC and cathepsin L activities. Taken together, these developments surely lay the foundation for the development of future cathepsin L-based therapies. Much excitement remains as new cathepsin L functions continue to emerge from specific cell types in the coming years.

Table 4. Reported structural studies of cathepsin L enzyme.

PDB Entry	Method	Resolution (Å)	Reference
1CJL	X-ray	2.20	[42]
1CS8	X-ray	1.80	[42]
1ICF	X-ray	2.00	[42]
1MHW	X-ray	1.90	[95]
2NQD	X-ray	1.75	[185]
2VHS	X-ray	1.50	[186]
2XU1	X-ray	1.45	[121]
2XU3	X-ray	0.90	[121]
2XU4	X-ray	1.12	[121]
2XU5	X-ray	1.60	[121]
2YJ2	X-ray	1.15	[97]
2YJ8	X-ray	1.30	[97]
2YJ9	X-ray	1.35	[97]
2YJB	X-ray	1.40	[97]
2YJC	X-ray	1.14	[97]

Table 4. *Cont.*

PDB Entry	Method	Resolution (Å)	Reference
3BC3	X-ray	2.20	[187]
3H89	X-ray	2.50	[128]
3H8B	X-ray	1.80	[128]
3H8C	X-ray	2.50	[128]
3HHA	X-ray	1.27	[147]
3HWN	X-ray	2.33	[147]
3IV2	X-ray	2.20	[188]
3K24	X-ray	2.50	[188]
3KSE	X-ray	1.71	[189]
3OF8	X-ray	2.20	[96]
3OF9	X-ray	1.76	[96]
4AXL	X-ray	1.92	[190]
4AXM	X-ray	2.80	[190]
5F02	X-ray	1.43	[191]
5I4H	X-ray	1.42	[192]
5MAE	X-ray	1.00	[145]
5MAJ	X-ray	1.00	[145]
5MQY	X-ray	1.13	[123]
6EZP	X-ray	1.37	[98]
6EZX	X-ray	2.34	[98]
6F06	X-ray	2.02	[98]

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

MA	Methacrylate
Phe	Phenylalanine
Gly	Glycine
Z/Cbz	Carboxybenzyl
His	Histidine
NBz	Nitrobenzyl
iAm	isoamyl
SAR	Structure Activity Relationship
AOMK	Acyloxymethanes/Acyloxymethyl Ketones
Leu	Leucine
Met	Methionine
Azi	Aziridine
FRET	Fluorescence Resonance Energy Transfer
catABP	Clickable and Tagless Activity-based Probe

References

1. de Duve, C. The lysosome turns fifty. *Nat. Cell Biol.* **2005**, *7*, 847–849. [CrossRef] [PubMed]
2. Schroder, B.A.; Wrocklage, C.; Hasilik, A.; Saftig, P. The proteome of lysosomes. *Proteomics* **2010**, *10*, 4053–4076. [CrossRef] [PubMed]
3. Rossi, A.; Deveraux, Q.; Turk, B.; Sali, A. Comprehensive search for cysteine cathepsins in the human genome. *Biol. Chem.* **2004**, *385*, 363–372. [CrossRef] [PubMed]
4. Reiser, J.; Adair, B.; Reinheckel, T. Specialized roles for cysteine cathepsins in health and disease. *J. Clin. Investig.* **2010**, *120*, 3421–3431. [CrossRef]
5. Turk, V.; Stoka, V.; Vasiljeva, O.; Renko, M.; Sun, T.; Turk, B.; Turk, D. Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochim. Biophys. Acta* **2012**, *1824*, 68–88. [CrossRef]
6. Olson, O.C.; Joyce, J.A. Cysteine cathepsin proteases: Regulators of cancer progression and therapeutic response. *Nat. Rev. Cancer* **2015**, *15*, 712–729. [CrossRef]
7. Vizovisek, M.; Fonovic, M.; Turk, B. Cysteine cathepsins in extracellular matrix remodeling: Extracellular matrix degradation and beyond. *Matrix Boil. J. Int. Soc. Matrix Biol.* **2019**, *75*, 141–159. [CrossRef]
8. Kramer, L.; Turk, D.; Turk, B. The Future of Cysteine Cathepsins in Disease Management. *Trends Pharmacol. Sci.* **2017**, *38*, 873–898. [CrossRef]
9. Siklos, M.; BenAissa, M.; Thatcher, G.R. Cysteine proteases as therapeutic targets: Does selectivity matter? A systematic review of calpain and cathepsin inhibitors. *Acta Pharm. Sin. B* **2015**, *5*, 506–519. [CrossRef]
10. Loser, R.; Pietzsch, J. Cysteine cathepsins: Their role in tumor progression and recent trends in the development of imaging probes. *Front. Chem.* **2015**, *3*, 37. [CrossRef]
11. Jakos, T.; Pislar, A.; Jewett, A.; Kos, J. Cysteine Cathepsins in Tumor-Associated Immune Cells. *Front. Immunol.* **2019**, *10*, 2037. [CrossRef]
12. Nakagawa, T.; Roth, W.; Wong, P.; Nelson, A.; Farr, A.; Deussing, J.; Villadangos, J.A.; Ploegh, H.; Peters, C.; Rudensky, A.Y. Cathepsin L: Critical role in Ii degradation and CD4 T cell selection in the thymus. *Science* **1998**, *280*, 450–453. [CrossRef]
13. Manchanda, M.; Das, P.; Gahlot, G.P.S.; Singh, R.; Roeb, E.; Roderfeld, M.; Datta Gupta, S.; Saraya, A.; Pandey, R.M.; Chauhan, S.S. Cathepsin L and B as Potential Markers for Liver Fibrosis: Insights From Patients and Experimental Models. *Clin. Transl. Gastroenterol.* **2017**, *8*, e99. [CrossRef]
14. Maehr, R.; Mintern, J.D.; Herman, A.E.; Lennon-Dumenil, A.M.; Mathis, D.; Benoist, C.; Ploegh, H.L. Cathepsin L is essential for onset of autoimmune diabetes in NOD mice. *J. Clin. Investig.* **2005**, *115*, 2934–2943. [CrossRef]
15. Huang, X.; Vaag, A.; Carlsson, E.; Hansson, M.; Ahren, B.; Groop, L. Impaired cathepsin L gene expression in skeletal muscle is associated with type 2 diabetes. *Diabetes* **2003**, *52*, 2411–2418. [CrossRef]
16. Kitamoto, S.; Sukhova, G.K.; Sun, J.; Yang, M.; Libby, P.; Love, V.; Duramad, P.; Sun, C.; Zhang, Y.; Yang, X.; et al. Cathepsin L deficiency reduces diet-induced atherosclerosis in low-density lipoprotein receptor-knockout mice. *Circulation* **2007**, *115*, 2065–2075. [CrossRef]
17. Sun, M.; Chen, M.; Liu, Y.; Fukuoka, M.; Zhou, K.; Li, G.; Dawood, F.; Gramolini, A.; Liu, P.P. Cathepsin-L contributes to cardiac repair and remodelling post-infarction. *Cardiovasc. Res.* **2011**, *89*, 374–383. [CrossRef]
18. He, W.; McCarroll, C.S.; Nather, K.; Elliott, E.B.A.; Nicklin, S.A.; Loughrey, C.M. The Cathepsin-L Inhibitor Caa0225 Protects against Myocardial Ischaemia-Reperfusion Injury. *Heart* **2015**, *101*, A1. [CrossRef]
19. Potts, W.; Bowyer, J.; Jones, H.; Tucker, D.; Freemont, A.J.; Millest, A.; Martin, C.; Vernon, W.; Neerunjun, D.; Slynn, G.; et al. Cathepsin L-deficient mice exhibit abnormal skin and bone development and show increased resistance to osteoporosis following ovariectomy. *Int. J. Exp. Pathol.* **2004**, *85*, 85–96. [CrossRef]
20. Furuyama, N.; Fujisawa, Y. Distinct roles of cathepsin K and cathepsin L in osteoclastic bone resorption. *Endocr. Res.* **2000**, *26*, 189–204. [CrossRef]
21. Yasuma, T.; Oi, S.; Choh, N.; Nomura, T.; Furuyama, N.; Nishimura, A.; Fujisawa, Y.; Sohda, T. Synthesis of Peptide Aldehyde Derivatives as Selective Inhibitors of Human Cathepsin L and Their Inhibitory Effect on Bone Resorption. *J. Med. Chem.* **1998**, *41*, 4301–4308. [CrossRef] [PubMed]
22. Onishi, K.; Li, Y.; Ishii, K.; Hisaeda, H.; Tang, L.; Duan, X.; Dainichi, T.; Maekawa, Y.; Katunuma, N.; Himeno, K. Cathepsin L is crucial for a Th1-type immune response during Leishmania major infection. *Microbes Infect.* **2004**, *6*, 468–474. [CrossRef] [PubMed]

23. Sever, S.; Altintas, M.M.; Nankoe, S.R.; Moller, C.C.; Ko, D.; Wei, C.; Henderson, J.; del Re, E.C.; Hsing, L.; Erickson, A.; et al. Proteolytic processing of dynamin by cytoplasmic cathepsin L is a mechanism for proteinuric kidney disease. *J. Clin. Investig.* **2007**, *117*, 2095–2104. [CrossRef]

24. Reiser, J.; Oh, J.; Shirato, I.; Asanuma, K.; Hug, A.; Mundel, T.M.; Honey, K.; Ishidoh, K.; Kominami, E.; Kreidberg, J.A.; et al. Podocyte migration during nephrotic syndrome requires a coordinated interplay between cathepsin L and alpha3 integrin. *J. Biol. Chem.* **2004**, *279*, 34827–34832. [CrossRef]

25. Sudhan, D.R.; Siemann, D.W. Cathepsin L targeting in cancer treatment. *Pharmacol. Ther.* **2015**, *155*, 105–116. [CrossRef]

26. Lankelma, J.M.; Voorend, D.M.; Barwari, T.; Koetsveld, J.; Van der Spek, A.H.; De Porto, A.P.; Van Rooijen, G.; Van Noorden, C.J. Cathepsin L, target in cancer treatment? *Life Sci.* **2010**, *86*, 225–233. [CrossRef]

27. Mohamed, M.M.; Sloane, B.F. Cysteine cathepsins: Multifunctional enzymes in cancer. *Nat. Rev. Cancer* **2006**, *6*, 764–775. [CrossRef]

28. Allan, E.R.; Yates, R.M. Redundancy between Cysteine Cathepsins in Murine Experimental Autoimmune Encephalomyelitis. *PLoS ONE* **2015**, *10*, e0128945. [CrossRef]

29. Cermak, S.; Kosicek, M.; Mladenovic-Djordjevic, A.; Smiljanic, K.; Kanazir, S.; Hecimovic, S. Loss of Cathepsin B and L Leads to Lysosomal Dysfunction, NPC-Like Cholesterol Sequestration and Accumulation of the Key Alzheimer's Proteins. *PLoS ONE* **2016**, *11*, e0167428. [CrossRef]

30. Jung, M.; Lee, J.; Seo, H.Y.; Lim, J.S.; Kim, E.K. Cathepsin inhibition-induced lysosomal dysfunction enhances pancreatic beta-cell apoptosis in high glucose. *PLoS ONE* **2015**, *10*, e0116972. [CrossRef]

31. UniProt, C. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515.

32. Lang, L.; Reitman, M.; Tang, J.; Roberts, R.M.; Kornfeld, S. Lysosomal enzyme phosphorylation. Recognition of a protein-dependent determinant allows specific phosphorylation of oligosaccharides present on lysosomal enzymes. *J. Biol. Chem.* **1984**, *259*, 14663–14671.

33. Braun, M.; Waheed, A.; von Figura, K. Lysosomal acid phosphatase is transported to lysosomes via the cell surface. *EMBO J.* **1989**, *8*, 3633–3640. [CrossRef] [PubMed]

34. Hashimoto, Y.; Kondo, C.; Katunuma, N. An Active 32-kDa Cathepsin L Is Secreted Directly from HT 1080 Fibrosarcoma Cells and Not via Lysosomal Exocytosis. *PLoS ONE* **2015**, *10*, e0145067. [CrossRef] [PubMed]

35. Jerala, R.; Zerovnik, E.; Kidric, J.; Turk, V. pH-induced conformational transitions of the propeptide of human cathepsin L. A role for a molten globule state in zymogen activation. *J. Biol. Chem.* **1998**, *273*, 11498–11504. [CrossRef] [PubMed]

36. Nishimura, Y.; Kawabata, T.; Furuno, K.; Kato, K. Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. *Arch. Biochem. Biophys.* **1989**, *271*, 400–406. [CrossRef]

37. Abudula, A.; Rommerskirch, W.; Weber, E.; Gunther, D.; Wiederanders, B. Splice variants of human cathepsin L mRNA show different expression rates. *Biol. Chem.* **2001**, *382*, 1583–1591. [CrossRef]

38. Jean, D.; Guillaume, N.; Frade, R. Characterization of human cathepsin L promoter and identification of binding sites for NF-Y, Sp1 and Sp3 that are essential for its activity. *Biochem. J.* **2002**, *361*, 173–184. [CrossRef]

39. Goulet, B.; Nepveu, A. Complete and limited proteolysis in cell cycle progression. *Cell Cycle* **2004**, *3*, 986–989. [CrossRef]

40. Goulet, B.; Sansregret, L.; Leduy, L.; Bogyo, M.; Weber, E.; Chauhan, S.S.; Nepveu, A. Increased expression and activity of nuclear cathepsin L in cancer cells suggests a novel mechanism of cell transformation. *Mol. Cancer Res. MCR* **2007**, *5*, 899–907. [CrossRef]

41. Goulet, B.; Baruch, A.; Moon, N.S.; Poirier, M.; Sansregret, L.L.; Erickson, A.; Bogyo, M.; Nepveu, A. A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. *Mol. Cell* **2004**, *14*, 207–219. [CrossRef]

42. Coulombe, R.; Grochulski, P.; Sivaraman, J.; Menard, R.; Mort, J.S.; Cygler, M. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J.* **1996**, *15*, 5492–5503. [CrossRef] [PubMed]

43. Coulombe, R.; Li, Y.; Takebe, S.; Menard, R.; Mason, P.; Mort, J.S.; Cygler, M. Crystallization and preliminary X-ray diffraction studies of human procathepsin L. *Proteins* **1996**, *25*, 398–400. [CrossRef]

44. Abrahamson, M.; Alvarez-Fernandez, M.; Nathanson, C.M. Cystatins. *Biochem. Soc. Symp.* **2003**, *70*, 179–199.

45. Kotsyfakis, M.; Sa-Nunes, A.; Francischetti, I.M.; Mather, T.N.; Andersen, J.F.; Ribeiro, J.M. Antiinflammatory and immunosuppressive activity of sialostatin L, a salivary cystatin from the tick Ixodes scapularis. *J. Biol. Chem.* **2006**, *281*, 26298–26307. [CrossRef]

46. Horka, H.; Staudt, V.; Klein, M.; Taube, C.; Reuter, S.; Dehzad, N.; Andersen, J.F.; Kopecky, J.; Schild, H.; Kotsyfakis, M.; et al. The tick salivary protein sialostatin L inhibits the Th9-derived production of the asthma-promoting cytokine IL-9 and is effective in the prevention of experimental asthma. *J. Immunol.* **2012**, *188*, 2669–2676. [CrossRef]

47. Carmona, E.; Dufour, E.; Plouffe, C.; Takebe, S.; Mason, P.; Mort, J.S.; Menard, R. Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry* **1996**, *35*, 8149–8157. [CrossRef]

48. Green, G.D.; Kembhavi, A.A.; Davies, M.E.; Barrett, A.J. Cystatin-like cysteine proteinase inhibitors from human liver. *Biochem. J.* **1984**, *218*, 939–946. [CrossRef]

49. Bromme, D.; Rinne, R.; Kirschke, H. Tight-binding inhibition of cathepsin S by cystatins. *Biomed. Biochim. Acta* **1991**, *50*, 631–635.

50. Barrett, A.J.; Davies, M.E.; Grubb, A. The place of human gamma-trace (cystatin C) amongst the cysteine proteinase inhibitors. *Biochem. Biophys. Res. Commun.* **1984**, *120*, 631–636. [CrossRef]

51. Balbin, M.; Hall, A.; Grubb, A.; Mason, R.W.; Lopez-Otin, C.; Abrahamson, M. Structural and functional characterization of two allelic variants of human cystatin D sharing a characteristic inhibition spectrum against mammalian cysteine proteinases. *J. Biol. Chem.* **1994**, *269*, 23156–23162. [PubMed]

52. Bevec, T.; Stoka, V.; Pungercic, G.; Dolenc, I.; Turk, V. Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L. *J. Exp. Med.* **1996**, *183*, 1331–1338. [CrossRef] [PubMed]

53. Salvesen, G.; Parkes, C.; Abrahamson, M.; Grubb, A.; Barrett, A.J. Human low-Mr kininogen contains three copies of a cystatin sequence that are divergent in structure and in inhibitory activity for cysteine proteinases. *Biochem. J.* **1986**, *234*, 429–434. [CrossRef] [PubMed]

54. Mullere-Esterl, W.; Iwanaga, S.; Nakanishi, S. Kininogens Revisited. *Trends Biochem. Sci.* **1986**, *11*, 336–339. [CrossRef]

55. Ni, J.; Fernandez, M.A.; Danielsson, L.; Chillakuru, R.A.; Zhang, J.; Grubb, A.; Su, J.; Gentz, R.; Abrahamson, M. Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor. *J. Biol. Chem.* **1998**, *273*, 24797–24804. [CrossRef] [PubMed]

56. Andrault, P.M.; Samsonov, S.A.; Weber, G.; Coquet, L.; Nazmi, K.; Bolscher, J.G.; Lalmanach, A.C.; Jouenne, T.; Bromme, D.; Pisabarro, M.T.; et al. Antimicrobial Peptide LL-37 Is Both a Substrate of Cathepsins S and K and a Selective Inhibitor of Cathepsin L. *Biochemistry* **2015**, *54*, 2785–2798. [CrossRef]

57. Ueno, T.; Takahashi, K. A cathepsin L-specific inhibitor preferentially inhibits degradation of autophagosomal LC3 and GABARAP in HeLa and Huh-7 cells. *Autophagy* **2009**, *5*, 878–879. [CrossRef]

58. Reiser, J.; Sever, S. Podocyte biology and pathogenesis of kidney disease. *Annu. Rev. Med.* **2013**, *64*, 357–366. [CrossRef]

59. Sever, S.; Chang, J.; Gu, C. Dynamin rings: Not just for fission. *Traffic* **2013**, *14*, 1194–1199. [CrossRef]

60. Faul, C.; Donnelly, M.; Merscher-Gomez, S.; Chang, Y.H.; Franz, S.; Delfgaauw, J.; Chang, J.M.; Choi, H.Y.; Campbell, K.N.; Kim, K.; et al. The actin cytoskeleton of kidney podocytes is a direct target of the antiproteinuric effect of cyclosporine A. *Nat. Med.* **2008**, *14*, 931–938. [CrossRef]

61. Cocchiaro, P.; De Pasquale, V.; Della Morte, R.; Tafuri, S.; Avallone, L.; Pizard, A.; Moles, A.; Pavone, L.M. The Multifaceted Role of the Lysosomal Protease Cathepsins in Kidney Disease. *Front. Cell Dev. Biol.* **2017**, *5*, 114. [CrossRef] [PubMed]

62. Alcalay, N.I.; Sharma, M.; Vassmer, D.; Chapman, B.; Paul, B.; Zhou, J.; Brantley, J.G.; Wallace, D.P.; Maser, R.L.; Vanden Heuvel, G.B. Acceleration of polycystic kidney disease progression in cpk mice carrying a deletion in the homeodomain protein Cux1. *Am. J. Physiol. Ren. Physiol.* **2008**, *295*, F1725–F1734. [CrossRef] [PubMed]

63. Alcalay, N.I.; Vanden Heuvel, G.B. Regulation of cell proliferation and differentiation in the kidney. *Front. Biosci.* **2009**, *14*, 4978–4991. [CrossRef] [PubMed]

64. Stypmann, J.; Glaser, K.; Roth, W.; Tobin, D.J.; Petermann, I.; Matthias, R.; Monnig, G.; Haverkamp, W.; Breithardt, G.; Schmahl, W.; et al. Dilated cardiomyopathy in mice deficient for the lysosomal cysteine peptidase cathepsin L. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6234–6239. [CrossRef] [PubMed]

65. Petermann, I.; Mayer, C.; Stypmann, J.; Biniossek, M.L.; Tobin, D.J.; Engelen, M.A.; Dandekar, T.; Grune, T.; Schild, L.; Peters, C.; et al. Lysosomal, cytoskeletal, and metabolic alterations in cardiomyopathy of cathepsin L knockout mice. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2006**, *20*, 1266–1268. [CrossRef] [PubMed]

66. Spira, D.; Stypmann, J.; Tobin, D.J.; Petermann, I.; Mayer, C.; Hagemann, S.; Vasiljeva, O.; Gunther, T.; Schule, R.; Peters, C.; et al. Cell type-specific functions of the lysosomal protease cathepsin L in the heart. *J. Biol. Chem.* **2007**, *282*, 37045–37052. [CrossRef]

67. Mehra, S.; Kumar, M.; Manchanda, M.; Singh, R.; Thakur, B.; Rani, N.; Arava, S.; Narang, R.; Arya, D.S.; Chauhan, S.S. Clinical significance of cathepsin L and cathepsin B in dilated cardiomyopathy. *Mol. Cell. Biochem.* **2017**, *428*, 139–147. [CrossRef]

68. Tang, Q.; Cai, J.; Shen, D.; Bian, Z.; Yan, L.; Wang, Y.X.; Lan, J.; Zhuang, G.Q.; Ma, W.Z.; Wang, W. Lysosomal cysteine peptidase cathepsin L protects against cardiac hypertrophy through blocking AKT/GSK3beta signaling. *J. Mol. Med.* **2009**, *87*, 249–260. [CrossRef]

69. Dalet-Fumeron, V.; Guinec, N.; Pagano, M. In vitro activation of pro-cathepsin B by three serine proteinases: Leucocyte elastase, cathepsin G, and the urokinase-type plasminogen activator. *FEBS Lett.* **1993**, *332*, 251–254. [CrossRef]

70. Felbor, U.; Dreier, L.; Bryant, R.A.; Ploegh, H.L.; Olsen, B.R.; Mothes, W. Secreted cathepsin L generates endostatin from collagen XVIII. *EMBO J.* **2000**, *19*, 1187–1194. [CrossRef]

71. Everts, V.; Korper, W.; Hoeben, K.A.; Jansen, I.D.; Bromme, D.; Cleutjens, K.B.; Heeneman, S.; Peters, C.; Reinheckel, T.; Saftig, P.; et al. Osteoclastic bone degradation and the role of different cysteine proteinases and matrix metalloproteinases: Differences between calvaria and long bone. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* **2006**, *21*, 1399–1408. [CrossRef] [PubMed]

72. Chavakis, E.; Dimmeler, S. Homing of progenitor cells to ischemic tissues. *Antioxid. Redox Signal.* **2011**, *15*, 967–980. [CrossRef] [PubMed]

73. Chavakis, E.; Koyanagi, M.; Dimmeler, S. Enhancing the outcome of cell therapy for cardiac repair: Progress from bench to bedside and back. *Circulation* **2010**, *121*, 325–335. [CrossRef] [PubMed]

74. Chavakis, E.; Urbich, C.; Dimmeler, S. Homing and engraftment of progenitor cells: A prerequisite for cell therapy. *J. Mol. Cell. Cardiol.* **2008**, *45*, 514–522. [CrossRef] [PubMed]

75. Urbich, C.; Dernbach, E.; Rossig, L.; Zeiher, A.M.; Dimmeler, S. High glucose reduces cathepsin L activity and impairs invasion of circulating progenitor cells. *J. Mol. Cell. Cardiol.* **2008**, *45*, 429–436. [CrossRef]

76. Wartmann, T.; Mayerle, J.; Kahne, T.; Sahin-Toth, M.; Ruthenburger, M.; Matthias, R.; Kruse, A.; Reinheckel, T.; Peters, C.; Weiss, F.U.; et al. Cathepsin L inactivates human trypsinogen, whereas cathepsin L-deletion reduces the severity of pancreatitis in mice. *Gastroenterology* **2010**, *138*, 726–737. [CrossRef]

77. Li, L.; Wang, X.; Fei, X.; Xia, L.; Qin, Z.; Liang, Z. Parkinson's disease involves autophagy and abnormal distribution of cathepsin L. *Neurosci. Lett.* **2011**, *489*, 62–67. [CrossRef] [PubMed]

78. Uchiyama, Y.; Koike, M.; Shibata, M.; Sasaki, M. Autophagic neuron death. *Methods Enzymol.* **2009**, *453*, 33–51.

79. Ohkouchi, S.; Shibata, M.; Sasaki, M.; Koike, M.; Safig, P.; Peters, C.; Nagata, S.; Uchiyama, Y. Biogenesis and proteolytic processing of lysosomal DNase II. *PLoS ONE* **2013**, *8*, e59148. [CrossRef]

80. Sugita, S.; Horie, S.; Nakamura, O.; Maruyama, K.; Takase, H.; Usui, Y.; Takeuchi, M.; Ishidoh, K.; Koike, M.; Uchiyama, Y.; et al. Acquisition of T regulatory function in cathepsin L-inhibited T cells by eye-derived CTLA-2alpha during inflammatory conditions. *J. Immunol* **2009**, *183*, 5013–5022. [CrossRef]

81. Yasothornsrikul, S.; Greenbaum, D.; Medzihradszky, K.F.; Toneff, T.; Bundey, R.; Miller, R.; Schilling, B.; Petermann, I.; Dehnert, J.; Logvinova, A.; et al. Cathepsin L in secretory vesicles functions as a prohormone-processing enzyme for production of the enkephalin peptide neurotransmitter. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9590–9595. [CrossRef]

82. Schurigt, U.; Eilenstein, R.; Gajda, M.; Leipner, C.; Sevenich, L.; Reinheckel, T.; Peters, C.; Wiederanders, B.; Brauer, R. Decreased arthritis severity in cathepsin L-deficient mice is attributed to an impaired T helper cell compartment. *Inflamm. Res.* **2012**, *61*, 1021–1029. [CrossRef] [PubMed]

83. Hsing, L.C.; Kirk, E.A.; McMillen, T.S.; Hsiao, S.H.; Caldwell, M.; Houston, B.; Rudensky, A.Y.; LeBoeuf, R.C. Roles for cathepsins S, L, and B in insulitis and diabetes in the NOD mouse. *J. Autoimmun.* **2010**, *34*, 96–104. [CrossRef] [PubMed]

84. Lv, B.J.; Lindholt, J.S.; Wang, J.; Cheng, X.; Shi, G.P. Plasma levels of cathepsins L, K, and V and risks of abdominal aortic aneurysms: A randomized population-based study. *Atherosclerosis* **2013**, *230*, 100–105. [CrossRef]

85. Sudhan, D.R.; Rabaglino, M.B.; Wood, C.E.; Siemann, D.W. Cathepsin L in tumor angiogenesis and its therapeutic intervention by the small molecule inhibitor KGP94. *Clin. Exp. Metastasis* **2016**, *33*, 461–473. [CrossRef] [PubMed]

86. Gocheva, V.; Joyce, J.A. Cysteine cathepsins and the cutting edge of cancer invasion. *Cell Cycle* **2007**, *6*, 60–64. [CrossRef]

87. Gocheva, V.; Zeng, W.; Ke, D.; Klimstra, D.; Reinheckel, T.; Peters, C.; Hanahan, D.; Joyce, J.A. Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. *Genes Dev.* **2006**, *20*, 543–556. [CrossRef]

88. Joyce, J.A.; Hanahan, D. Multiple roles for cysteine cathepsins in cancer. *Cell Cycle* **2004**, *3*, 1516–1619. [CrossRef]

89. Gottesman, M.M. Transformation-dependent secretion of a low molecular weight protein by murine fibroblasts. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 2767–2771. [CrossRef]

90. Johnson, D.S.; Weerapana, E.; Cravatt, B.F. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* **2010**, *2*, 949–964. [CrossRef]

91. Singh, J.; Petter, R.C.; Baillie, T.A.; Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **2011**, *10*, 307–317. [CrossRef] [PubMed]

92. Turk, D.; Guncar, G. Lysosomal cysteine proteases (cathepsins): Promising drug targets. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2003**, *59*, 203–213. [CrossRef] [PubMed]

93. Choe, Y.; Leonetti, F.; Greenbaum, D.C.; Lecaille, F.; Bogyo, M.; Bromme, D.; Ellman, J.A.; Craik, C.S. Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. *J. Biol. Chem.* **2006**, *281*, 12824–12832. [CrossRef] [PubMed]

94. Fujishima, A.; Imai, Y.; Nomura, T.; Fujisawa, Y.; Yamamoto, Y.; Sugawara, T. The crystal structure of human cathepsin L complexed with E-64. *FEBS Lett.* **1997**, *407*, 47–50. [CrossRef]

95. Chowdhury, S.F.; Sivaraman, J.; Wang, J.; Devanathan, G.; Lachance, P.; Qi, H.; Menard, R.; Lefebvre, J.; Konishi, Y.; Cygler, M.; et al. Design of noncovalent inhibitors of human cathepsin L. From the 96-residue proregion to optimized tripeptides. *J. Med. Chem.* **2002**, *45*, 5321–5329. [CrossRef] [PubMed]

96. Shenoy, R.T.; Sivaraman, J. Structural basis for reversible and irreversible inhibition of human cathepsin L by their respective dipeptidyl glyoxal and diazomethylketone inhibitors. *J. Struct. Biol.* **2011**, *173*, 14–19. [CrossRef] [PubMed]

97. Hardegger, L.A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; et al. Halogen bonding at the active sites of human cathepsin L and MEK1 kinase: Efficient interactions in different environments. *ChemMedChem* **2011**, *6*, 2048–2054. [CrossRef]

98. Giroud, M.; Dietzel, U.; Anselm, L.; Banner, D.; Kuglstatter, A.; Benz, J.; Blanc, J.B.; Gaufreteau, D.; Liu, H.; Lin, X.; et al. Repurposing a Library of Human Cathepsin L Ligands: Identification of Macrocyclic Lactams as Potent Rhodesain and Trypanosoma brucei Inhibitors. *J. Med. Chem.* **2018**, *61*, 3350–3369. [CrossRef]

99. Barrett, A.J.; Kembhavi, A.A.; Brown, M.A.; Kirschke, H.; Knight, C.G.; Tamai, M.; Hanada, K. L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochem. J.* **1982**, *201*, 189–198. [CrossRef]

100. Goursalin, B.J.; Lachance, P.; Bonneau, P.R.; Storer, A.C.; Kirschke, H.; Broemme, D. E-64 Analogs as Inhibitors of Cathepsin L and Cathepsin S: Importance of the S2-P2 Interactions for Potency and Selectivity. *Bioorg. Chem.* **1994**, *22*, 227–241. [CrossRef]

101. Katunuma, N.; Murata, E.; Kakegawa, H.; Matsui, A.; Tsuzuki, H.; Tsuge, H.; Turk, D.; Turk, V.; Fukushima, M.; Tada, Y.; et al. Structure based development of novel specific inhibitors for cathepsin L and cathepsin S in vitro and in vivo. *FEBS Lett.* **1999**, *458*, 6–10. [CrossRef]

102. Tsuge, H.; Nishimura, T.; Tada, Y.; Asao, T.; Turk, D.; Turk, V.; Katunuma, N. Inhibition mechanism of cathepsin L-specific inhibitors based on the crystal structure of papain-CLIK148 complex. *Biochem. Biophys. Res. Commun.* **1999**, *266*, 411–416. [CrossRef] [PubMed]

103. Crawford, C.; Mason, R.W.; Wikstrom, P.; Shaw, E. The design of peptidyldiazomethane inhibitors to distinguish between the cysteine proteinases calpain II, cathepsin L and cathepsin B. *Biochem. J.* **1988**, *253*, 751–758. [CrossRef] [PubMed]

104. Fischer, G.; Demuth, H.U.; Barth, A. N,O-Diacylhydroxylamines as enzyme-activated inhibitors for serine proteases. *Die Pharm.* **1983**, *38*, 249.

105. Demuth, H.U.; Baumgrass, R.; Schaper, C.; Fischer, G.; Barth, A. Dipeptidylpeptidase IV–inactivation with N-peptidyl-O-aroyl hydroxylamines. *J. Enzym. Inhib.* **1988**, *2*, 129–142. [CrossRef]

106. Demuth, H.U.; Schönlein, C.; Barth, A. Potent and selective inactivation of proteinases with N-peptidyl-O-acylhydroxylamines. *Biochim. Biophys. Acta* **1989**, *996*, 19–22. [CrossRef]

107. Demuth, H.U.; Neumann, U.; Barth, A. Reactions between dipeptidyl peptidase IV and diacyl hydroxylamines: Mechanistic investigations. *J. Enzym. Inhib.* **1989**, *2*, 239–248. [CrossRef]

108. Brömme, D.; Schierhorn, A.; Kirschke, H.; Wiederanders, B.; Barth, A.; Fittkau, S.; Demuth, H.U. Potent and selective inactivation of cysteine proteinases with N-peptidyl-O-acyl hydroxylamines. *Biochem. J.* **1989**, *263*, 861–866. [CrossRef]

109. Broemme, D.; Neumann, U.; Kirschke, H.; Demuth, H.-U. Novel N-peptidyl-O-acyl hydroxamates: Selective inhibitors of cysteine proteinases. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* **1993**, *1202*, 271. [CrossRef]

110. Brömme, D.; Kirschke, H. N-peptidyl-O-carbamoyl amino acid hydroxamates: Irreversible inhibitors for the study of the S2′ specificity of cysteine proteinases. *FEBS Lett.* **1993**, *322*, 211–214. [CrossRef]

111. Gour-Salin, B.J.; Lachance, P.; Magny, M.C.; Plouffe, C.; Menard, R.; Storer, A.C. E64 [trans-epoxysuccinyl-L-leucylamido-(4-guanidino)butane] analogues as inhibitors of cysteine proteinases: Investigation of S2 subsite interactions. *Biochem. J.* **1994**, *299*, 389–392. [CrossRef] [PubMed]

112. Krantz, A. Peptidyl (acyloxy)methanes as quiescent affinity labels for cysteine proteinases. *Methods Enzymol.* **1994**, *244*, 656. [PubMed]

113. Torkar, A.; Lenarcic, B.; Lah, T.; Dive, V.; Devel, L. Identification of new peptide amides as selective cathepsin L inhibitors: The first step towards selective irreversible inhibitors? *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2968–2973. [CrossRef] [PubMed]

114. Martichonok, V.; Plouffe, C.; Storer, A.C.; Menard, R.; Jones, J.B. Aziridine Analogs of [[trans-(Epoxysuccinyl)-L-leucyl]amino]-4-guanidinobutane (E-64) as Inhibitors of Cysteine Proteases. *J. Med. Chem.* **1995**, *38*, 3078. [CrossRef] [PubMed]

115. Schirmeister, T. New Peptidic Cysteine Protease Inhibitors Derived from the Electrophilic α-Amino Acid Aziridine-2,3-dicarboxylic Acid. *J. Med. Chem.* **1999**, *42*, 560–572. [CrossRef] [PubMed]

116. Mendieta, L.; Pico, A.; Tarrago, T.; Teixido, M.; Castillo, M.; Rafecas, L.; Moyano, A.; Giralt, E. Novel peptidyl aryl vinyl sulfones as highly potent and selective inhibitors of cathepsins L and B. *ChemMedChem* **2010**, *5*, 1556–1567. [CrossRef] [PubMed]

117. Dana, D.; De, S.; Rathod, P.; Davalos, A.R.; Novoa, D.A.; Paroly, S.; Torres, V.M.; Afzal, N.; Lankalapalli, R.S.; Rotenberg, S.A.; et al. Development of a highly potent, selective, and cell-active inhibitor of cysteine cathepsin L-A hybrid design approach. *Chem. Commun.* **2014**, *50*, 10875–10878. [CrossRef]

118. Boudreau, P.D.; Miller, B.W.; McCall, L.-I.; Almaliti, J.; Reher, R.; Hirata, K.; Le, T.; Siqueira-Neto, J.L.; Hook, V.; Gerwick, W.H. Design of Gallinamide A Analogs as Potent Inhibitors of the Cysteine Proteases Human Cathepsin L and Trypanosoma cruzi Cruzain. *J. Med. Chem.* **2019**, *62*, 9026–9044. [CrossRef]

119. Lynas, J.F.; Hawthorne, S.J.; Walker, B. Development of peptidyl α-keto-β-aldehydes as new inhibitors of cathepsin L—Comparisons of potency and selectivity profiles with cathepsin B. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1771–1773. [CrossRef]

120. Marquis, R.W.; James, I.; Zeng, J.; Trout, R.E.; Thompson, S.; Rahman, A.; Yamashita, D.S.; Xie, R.; Ru, Y.; Gress, C.J.; et al. Azepanone-based inhibitors of human cathepsin L. *J. Med. Chem.* **2005**, *48*, 6870–6878. [CrossRef]

121. Hardegger, L.A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; et al. Systematic investigation of halogen bonding in protein-ligand interactions. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 314–318. [CrossRef] [PubMed]

122. Giroud, M.; Harder, M.; Kuhn, B.; Haap, W.; Trapp, N.; Schweizer, W.B.; Schirmeister, T.; Diederich, F. Fluorine Scan of Inhibitors of the Cysteine Protease Human Cathepsin L: Dipolar and Quadrupolar Effects in the pi-Stacking of Fluorinated Phenyl Rings on Peptide Amide Bonds. *ChemMedChem* **2017**, *12*, 257–270. [CrossRef] [PubMed]

123. Kuhn, B.; Tichý, M.; Wang, L.; Robinson, S.; Martin, R.E.; Kuglstatter, A.; Benz, J.; Giroud, M.; Schirmeister, T.; Abel, R.; et al. Prospective Evaluation of Free Energy Calculations for the Prioritization of Cathepsin L Inhibitors. *J. Med. Chem.* **2017**, *60*, 2485–2497. [CrossRef] [PubMed]

124. Parker, E.N.; Odutola, S.O.; Wang, Y.; Strecker, T.E.; Mukherjee, R.; Shi, Z.; Chaplin, D.J.; Trawick, M.L.; Pinney, K.G. Synthesis and biological evaluation of a water-soluble phosphate prodrug salt and structural analogues of KGP94, a lead inhibitor of cathepsin L. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 1304–1310. [CrossRef]

125. Parker, E.N.; Song, J.; Kishore Kumar, G.D.; Odutola, S.O.; Chavarria, G.E.; Charlton-Sevcik, A.K.; Strecker, T.E.; Barnes, A.L.; Sudhan, D.R.; Wittenborn, T.R.; et al. Synthesis and biochemical evaluation of benzoylbenzophenone thiosemicarbazone analogues as potent and selective inhibitors of cathepsin L. *Bioorg. Med. Chem.* **2015**, *23*, 6974–6992. [CrossRef]

126. Kumar, G.D.K.; Chavarria, G.E.; Charlton-Sevcik, A.K.; Yoo, G.K.; Song, J.; Strecker, T.E.; Siim, B.G.; Chaplin, D.J.; Trawick, M.L.; Pinney, K.G. Functionalized benzophenone, thiophene, pyridine, and fluorene thiosemicarbazone derivatives as inhibitors of cathepsin L. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6610–6615. [CrossRef]

127. Kishore Kumar, G.D.; Chavarria, G.E.; Charlton-Sevcik, A.K.; Arispe, W.M.; Macdonough, M.T.; Strecker, T.E.; Chen, S.E.; Siim, B.G.; Chaplin, D.J.; Trawick, M.L.; et al. Design, synthesis, and biological evaluation of potent thiosemicarbazone based cathepsin L inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1415–1419. [CrossRef]

128. Shenoy, R.T.; Chowdhury, S.F.; Kumar, S.; Joseph, L.; Purisima, E.O.; Sivaraman, J. A Combined Crystallographic and Molecular Dynamics Study of Cathepsin L Retrobinding Inhibitors. *J. Med. Chem.* **2009**, *52*, 6335–6346. [CrossRef]

129. Myers, M.C.; Shah, P.P.; Diamond, S.L.; Huryn, D.M.; Smith, A.B., 3rd. Identification and synthesis of a unique thiocarbazate cathepsin L inhibitor. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 210–214. [CrossRef]

130. Myers, M.C.; Shah, P.P.; Beavers, M.P.; Napper, A.D.; Diamond, S.L.; Smith, A.B., 3rd; Huryn, D.M. Design, synthesis, and evaluation of inhibitors of cathepsin L: Exploiting a unique thiocarbazate chemotype. *Bioorganic Med. Chem. Chem. Lett.* **2008**, *18*, 3646–3651. [CrossRef]

131. Shah, P.P.; Myers, M.C.; Beavers, M.P.; Purvis, J.E.; Jing, H.; Grieser, H.J.; Sharlow, E.R.; Napper, A.D.; Huryn, D.M.; Cooperman, B.S.; et al. Kinetic characterization and molecular docking of a novel, potent, and selective slow-binding inhibitor of human cathepsin L. *Mol. Pharmacol.* **2008**, *74*, 34–41. [CrossRef]

132. Shah, P.P.; Wang, T.; Kaletsky, R.L.; Myers, M.C.; Purvis, J.E.; Jing, H.; Huryn, D.M.; Greenbaum, D.C.; Smith, A.B., 3rd; Bates, P.; et al. A small-molecule oxocarbazate inhibitor of human cathepsin L blocks severe acute respiratory syndrome and ebola pseudotype virus infection into human embryonic kidney 293T cells. *Mol. Pharmacol.* **2010**, *78*, 319–324. [CrossRef] [PubMed]

133. Vicik, R.; Helten, H.; Schirmeister, T.; Engels, B. Rational design of aziridine-containing cysteine protease inhibitors with improved potency: Studies on inhibition mechanism. *ChemMedChem* **2006**, *1*, 1021–1028. [CrossRef] [PubMed]

134. Palmer, J.T.; Rasnick, D.; Klaus, J.L.; Bromme, D. Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J. Med. Chem.* **1995**, *38*, 3193–3196. [CrossRef] [PubMed]

135. Bromme, D.; Klaus, J.L.; Okamoto, K.; Rasnick, D.; Palmer, J.T. Peptidyl vinyl sulphones: A new class of potent and selective cysteine protease inhibitors: S2P2 specificity of human cathepsin O2 in comparison with cathepsins S and L. *Biochem. J.* **1996**, *315*, 85–89. [CrossRef] [PubMed]

136. Reddick, J.J.; Cheng, J.M.; Roush, W.R. Relative rates of Michael reactions of 2′-(phenethyl)thiol with vinyl sulfones, vinyl sulfonate esters, and vinyl sulfonamides relevant to vinyl sulfonyl cysteine protease inhibitors. *Org. Lett.* **2003**, *5*, 1967–1970. [CrossRef] [PubMed]

137. Roush, W.R.; Gwaltney, S.L.; Cheng, J.M.; Scheidt, K.A.; McKerrow, J.H.; Hansell, E. Vinyl sulfonate esters and vinyl sulfonamides: Potent, irreversible inhibitors of cysteine proteases. *J. Am. Chem. Soc.* **1998**, *120*, 10994–10995. [CrossRef]

138. Miller, B.; Friedman, A.J.; Choi, H.; Hogan, J.; McCammon, J.A.; Hook, V.; Gerwick, W.H. The marine cyanobacterial metabolite gallinamide A is a potent and selective inhibitor of human cathepsin L. *J. Nat. Prod.* **2014**, *77*, 92–99. [CrossRef]

139. Woo, J.-T.; Sigeizumi, S.; Yamaguchi, K.; Sugimoto, K.; Kobori, T.; Tsuji, T.; Kondo, K. Peptidyl aldehyde derivatives as potent and selective inhibitors of cathepsin L. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1501. [CrossRef]

140. Woo, J.T.; Yamaguchi, K.; Hayama, T.; Kobori, T.; Sigeizumi, S.; Sugimoto, K.; Kondo, K.; Tsuji, T.; Ohba, Y.; Tagami, K.; et al. Suppressive effect of N-(benzyloxycarbonyl)-L-phenylalanyl-L-tyrosinal on bone resorption in vitro and in vivo. *Eur. J. Pharmacol.* **1996**, *300*, 131–135. [CrossRef]

141. Marquis, R.W.; Ru, Y.; LoCastro, S.M.; Zeng, J.; Yamashita, D.S.; Oh, H.J.; Erhard, K.F.; Davis, L.D.; Tomaszek, T.A.; Tew, D.; et al. Azepanone-based inhibitors of human and rat cathepsin K. *J. Med. Chem.* **2001**, *44*, 1380–1395. [CrossRef] [PubMed]

142. Kerns, J.K.; Nie, H.; Bondinell, W.; Widdowson, K.L.; Yamashita, D.S.; Rahman, A.; Podolin, P.L.; Carpenter, D.C.; Jin, Q.; Riflade, B.; et al. Azepanone-based inhibitors of human cathepsin S: Optimization of selectivity via the P2 substituent. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4409–4415. [CrossRef] [PubMed]

143. Gauthier, J.Y.; Chauret, N.; Cromlish, W.; Desmarais, S.; Duong le, T.; Falgueyret, J.P.; Kimmel, D.B.; Lamontagne, S.; Leger, S.; LeRiche, T.; et al. The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 923–928. [CrossRef] [PubMed]

144. Stone, J.A.; McCrea, J.B.; Witter, R.; Zajic, S.; Stoch, S.A. Clinical and translational pharmacology of the cathepsin K inhibitor odanacatib studied for osteoporosis. *Br. J. Clin. Pharmacol.* **2019**, *85*, 1072–1083. [CrossRef] [PubMed]

145. Giroud, M.; Ivkovic, J.; Martignoni, M.; Fleuti, M.; Trapp, N.; Haap, W.; Kuglstatter, A.; Benz, J.; Kuhn, B.; Schirmeister, T.; et al. Inhibition of the Cysteine Protease Human Cathepsin L by Triazine Nitriles: Amide···Heteroarene π-Stacking Interactions and Chalcogen Bonding in the S3 Pocket. *ChemMedChem* **2017**, *12*, 257–270. [CrossRef] [PubMed]

146. Falgueyret, J.P.; Oballa, R.M.; Okamoto, O.; Wesolowski, G.; Aubin, Y.; Rydzewski, R.M.; Prasit, P.; Riendeau, D.; Rodan, S.B.; Percival, M.D. Novel, nonpeptidic cyanamides as potent and reversible inhibitors of human cathepsins K and L. *J. Med. Chem.* **2001**, *44*, 94–104. [CrossRef] [PubMed]

147. Asaad, N.; Bethel, P.A.; Coulson, M.D.; Dawson, J.E.; Ford, S.J.; Gerhardt, S.; Grist, M.; Hamlin, G.A.; James, M.J.; Jones, E.V.; et al. Dipeptidyl nitrile inhibitors of Cathepsin L. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4280–4283. [CrossRef]

148. Mallari, J.P.; Shelat, A.; Kosinski, A.; Caffrey, C.R.; Connelly, M.; Zhu, F.; McKerrow, J.H.; Guy, R.K. Discovery of trypanocidal thiosemicarbazone inhibitors of rhodesain and TbcatB. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2883–2885. [CrossRef]

149. Magrath, J.; Abeles, R.H. Cysteine protease inhibition by azapeptide esters. *J. Med. Chem.* **1992**, *35*, 4279. [CrossRef]

150. Beavers, M.P.; Myers, M.C.; Shah, P.P.; Purvis, J.E.; Diamond, S.L.; Cooperman, B.S.; Huryn, D.M.; Smith, A.B., 3rd. Molecular docking of cathepsin L inhibitors in the binding site of papain. *J. Chem. Inf. Mode.* **2008**, *48*, 1464–1472. [CrossRef]

151. Jean, D.; Rousselet, N.; Frade, R. Cathepsin L expression is up-regulated by hypoxia in human melanoma cells: Role of its 5′-untranslated region. *Biochem. J.* **2008**, *413*, 125–134. [CrossRef] [PubMed]

152. Orlowski, G.M.; Colbert, J.D.; Sharma, S.; Bogyo, M.; Robertson, S.A.; Rock, K.L. Multiple Cathepsins Promote Pro-IL-1beta Synthesis and NLRP3-Mediated IL-1beta Activation. *J. Immunol.* **2015**, *195*, 1685–1697. [CrossRef] [PubMed]

153. Cravatt, B.F.; Wright, A.T.; Kozarich, J.W. Activity-based protein profiling: From enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383–414. [CrossRef] [PubMed]

154. Bogyo, M. Finding enzymes that are actively involved in cancer. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2379–2380. [CrossRef]

155. Kumar, S.; Zhou, B.; Liang, F.; Wang, W.Q.; Huang, Z.; Zhang, Z.Y. Activity-based probes for protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7943–7948. [CrossRef]

156. Kumar, S.; Zhou, B.; Liang, F.; Yang, H.; Wang, W.Q.; Zhang, Z.Y. Global analysis of protein tyrosine phosphatase activity with ultra-sensitive fluorescent probes. *J. Proteome Res.* **2006**, *5*, 1898–1905. [CrossRef]

157. Yim, J.J.; Tholen, M.; Klaassen, A.; Sorger, J.; Bogyo, M. Optimization of a Protease Activated Probe for Optical Surgical Navigation. *Mol. Pharm.* **2018**, *15*, 750–758. [CrossRef]

158. Chan, C.H.; Liesenfeld, L.F.; Ferreiro-Neira, I.; Cusack, J.C., Jr. Preclinical Evaluation of Cathepsin-Based Fluorescent Imaging System for Cytoreductive Surgery. *Ann. Surg. Oncol.* **2017**, *24*, 931–938. [CrossRef]

159. Kato, D.; Boatright, K.M.; Berger, A.B.; Nazif, T.; Blum, G.; Ryan, C.; Chehade, K.A.; Salvesen, G.S.; Bogyo, M. Activity-based probes that target diverse cysteine protease families. *Nat. Chem. Biol.* **2005**, *1*, 33–38. [CrossRef]

160. Edgington, L.E.; Verdoes, M.; Bogyo, M. Functional imaging of proteases: Recent advances in the design and application of substrate-based and activity-based probes. *Curr. Opin. Chem. Biol.* **2011**, *15*, 798–805. [CrossRef]

161. Sanman, L.E.; Bogyo, M. Activity-based profiling of proteases. *Annu. Rev. Biochem.* **2014**, *83*, 249–273. [CrossRef] [PubMed]

162. Docherty, K.; Carroll, R.; Steiner, D.F. Identification of a 31,500 molecular weight islet cell protease as cathepsin B. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 3245–3249. [CrossRef] [PubMed]

163. Docherty, K.; Hutton, J.C.; Steiner, D.F. Cathepsin B-related proteases in the insulin secretory granule. *J. Biol. Chem.* **1984**, *259*, 6041. [PubMed]

164. Mason, R.W.; Wilcox, D.; Wikstrom, P.; Shaw, E.N. The identification of active forms of cysteine proteinases in Kirsten-virus-transformed mouse fibroblasts by use of a specific radiolabeled inhibitor. *Biochem. J.* **1989**, *257*, 125. [CrossRef]

165. Mason, R.W.; Bartholomew, L.T.; Hardwick, B.S. The use of benzyloxycarbonyl [^{125}I]iodotyrosylalanyldiazomethane as a probe for active cysteine proteinases in human tissues. *Biochem. J.* **1989**, *263*, 945–949. [CrossRef]

166. Wilcox, D.; Mason, R.W. Inhibition of cysteine proteinases in lysosomes and whole cells. *Biochem. J.* **1992**, *285*, 495–502. [CrossRef]

167. Xing, R.; Addington, A.K.; Mason, R.W. Quantification of cathepsins B and L in cells. *Biochem. J.* **1998**, *332*, 499–505. [CrossRef]

168. Gelhaus, C.; Vicik, R.; Hilgenfeld, R.; Schmidt, C.L.; Leippe, M.; Schirmeister, T. Synthesis and antiplasmodial activity of a cysteine protease-inhibiting biotinylated aziridine-2,3-dicarboxylate. *Biol. Chem.* **2004**, *385*, 435–438. [CrossRef]

169. Vicik, R.; Busemann, M.; Gelhaus, C.; Stiefl, N.; Scheiber, J.; Schmitz, W.; Schulz, F.; Mladenovic, M.; Engels, B.; Leippe, M.; et al. Aziridide-based inhibitors of cathepsin L: Synthesis, inhibition activity, and docking studies. *ChemMedChem* **2006**, *1*, 1126–1141. [CrossRef]

170. Torkar, A.; Bregant, S.; Devel, L.; Novinec, M.; Lenarcic, B.; Lah, T.; Dive, V. A novel photoaffinity-based probe for selective detection of cathepsin L active form. *Chembiochem Eur. J. Chem. Biol.* **2012**, *13*, 2616–2621. [CrossRef]

171. Na, Z.; Li, L.; Uttamchandani, M.; Yao, S.Q. Microarray-guided discovery of two-photon (2P) small molecule probes for live-cell imaging of cysteinyl cathepsin activities. *Chem. Commun.* **2012**, *48*, 7304–7306. [CrossRef] [PubMed]

172. Poreba, M.; Rut, W.; Vizovisek, M.; Groborz, K.; Kasperkiewicz, P.; Finlay, D.; Vuori, K.; Turk, D.; Turk, B.; Salvesen, G.S.; et al. Selective imaging of cathepsin L in breast cancer by fluorescent activity-based probes. *Chem. Sci.* **2018**, *9*, 2113–2129. [CrossRef] [PubMed]

173. Dana, D.; Garcia, J.; Bhuiyan, A.I.; Rathod, P.; Joo, L.; Novoa, D.A.; Paroly, S.; Fath, K.R.; Chang, E.J.; Pathak, S.K. Cell penetrable, clickable and tagless activity-based probe of human cathepsin L. *Bioorg. Chem.* **2019**, *85*, 505–514. [CrossRef] [PubMed]

174. Poreba, M.; Groborz, K.; Rut, W.; Pore, M.; Snipas, S.J.; Vizovisek, M.; Turk, B.; Kuhn, P.; Drag, M.; Salvesen, G.S. The Activome: Multiplexed probing of activity of proteolytic enzymes using mass cytometry-compatible activity-based probes (TOF-probes). *bioRxiv* **2019**. [CrossRef]

175. Blum, G.; von Degenfeld, G.; Merchant, M.J.; Blau, H.M.; Bogyo, M. Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes. *Nat. Chem. Biol.* **2007**, *3*, 668–677. [CrossRef]

176. Blum, G.; Mullins, S.R.; Keren, K.; Fonovic, M.; Jedeszko, C.; Rice, M.J.; Sloane, B.F.; Bogyo, M. Dynamic imaging of protease activity with fluorescently quenched activity-based probes. *Nat. Chem. Biol.* **2005**, *1*, 203–209. [CrossRef]

177. Vogel, A.M.; Gerster, T. Expression of a zebrafish Cathepsin L gene in anterior mesendoderm and hatching gland. *Dev. Genes Evol.* **1997**, *206*, 477–479. [CrossRef]

178. Crosier, P.S.; Bardsley, A.; Horsfield, J.A.; Krassowska, A.K.; Lavallie, E.R.; Collins-Racie, L.A.; Postlethwait, J.H.; Yan, Y.L.; McCoy, J.M.; Crosier, K.E. In situ hybridization screen in zebrafish for the selection of genes encoding secreted proteins. *Dev. Dyn.* **2001**, *222*, 637–644. [CrossRef]

179. Gardiner, M.R.; Daggett, D.F.; Zon, L.I.; Perkins, A.C. Zebrafish KLF4 is essential for anterior mesendoderm/pre-polster differentiation and hatching. *Dev. Dyn.* **2005**, *234*, 992–996. [CrossRef]

180. Kasperkiewicz, P.; Altman, Y.; D'Angelo, M.; Salvesen, G.S.; Drag, M. Toolbox of Fluorescent Probes for Parallel Imaging Reveals Uneven Location of Serine Proteases in Neutrophils. *J. Am. Chem. Soc.* **2017**, *139*, 10115–10125. [CrossRef]

181. Choi, K.Y.; Swierczewska, M.; Lee, S.; Chen, X. Protease-activated drug development. *Theranostics* **2012**, *2*, 156–178. [CrossRef] [PubMed]

182. Chowdhury, M.A.; Moya, I.A.; Bhilocha, S.; McMillan, C.C.; Vigliarolo, B.G.; Zehbe, I.; Phenix, C.P. Prodrug-inspired probes selective to cathepsin B over other cysteine cathepsins. *J. Med. Chem.* **2014**, *57*, 6092–6104. [CrossRef] [PubMed]

183. Weinstain, R.; Segal, E.; Satchi-Fainaro, R.; Shabat, D. Real-time monitoring of drug release. *Chem. Commun.* **2010**, *46*, 553–555. [CrossRef] [PubMed]

184. Ueki, N.; Lee, S.; Sampson, N.S.; Hayman, M.J. Selective cancer targeting with prodrugs activated by histone deacetylases and a tumour-associated protease. *Nat. Commun.* **2013**, *4*, 2735. [CrossRef]

185. Ljunggren, A.; Redzynia, I.; Alvarez-Fernandez, M.; Abrahamson, M.; Mort, J.S.; Krupa, J.C.; Jaskólski, M.; Bujacz, G. Crystal Structure of the Parasite Protease Inhibitor Chagasin in Complex with a Host Target Cysteine Protease. *J. Mol. Boil.* **2007**, *371*, 137–153. [CrossRef]

186. Fairhead, M.; Johnson, K.A.; Kowatz, T.; McMahon, S.A.; Carter, L.G.; Oke, M.; Liu, H.; Naismith, J.H.; van der Walle, C.F. Crystal structure and silica condensing activities of silicatein alpha-cathepsin L chimeras. *Chem. Commun.* **2008**, *15*, 1765–1767. [CrossRef]

187. Chowdhury, S.F.; Joseph, L.; Kumar, S.; Tulsidas, S.R.; Bhat, S.; Ziomek, E.; Ménard, R.; Sivaraman, J.; Purisima, E.O. Exploring Inhibitor Binding at the S' Subsites of Cathepsin L. *J. Med. Chem.* **2008**, *51*, 1361–1368. [CrossRef]

188. Adams-Cioaba, M.A.; Krupa, J.C.; Xu, C.; Mort, J.S.; Min, J. Structural basis for the recognition and cleavage of histone H3 by cathepsin L. *Nat. Commun.* **2011**, *2*, 197. [CrossRef]

189. Renko, M.; Türk, D. Unreduced Cathepsin L in Complex with Stefin A. 2010. Available online: https://www.rcsb.org/structure/3KSE (accessed on 4 February 2020).

190. Ehmke, V.; Winkler, E.; Banner, D.W.; Haap, W.; Schweizer, W.B.; Rottmann, M.; Kaiser, M.; Freymond, C.; Schirmeister, T.; Diederich, F. Optimization of Triazine Nitriles as Rhodesain Inhibitors: Structure-Activity Relationships, Bioisosteric Imidazopyridine Nitriles, and X-ray Crystal Structure Analysis with Human Cathepsin L. *ChemMedChem* **2013**, *8*, 967–975. [CrossRef] [PubMed]

191. Banner, D.; Benz, J.; Stihle, M.; Kuglstatter, A. CATHEPSIN L IN COMPLEX WITH (2S,4R)-4-(2-Chloro-4-methoxy-benzenesulfonyl)-1-[3-(5-chloro-pyridin-2-yl)-azetidine-3-carbonyl]-pyrrolidine-2-carboxylic acid (1-cyano-cyclopropyl)-amide. 2016. Available online: https://www.rcsb.org/structure/5F02 (accessed on 4 February 2020). [CrossRef]

192. Sosnowski, P.; Turk, D. Caught in the act: The crystal structure of cleaved cathepsin L bound to the active site of Cathepsin L. *FEBS Lett.* **2016**, *590*, 1253–1261. [CrossRef] [PubMed]