

Targeting proteases: successes, failures and future prospects

Boris Turk

Abstract | Until fairly recently, proteases were considered primarily to be protein-degrading enzymes. However, this view has dramatically changed and proteases are now seen as extremely important signalling molecules that are involved in numerous vital processes. Protease signalling pathways are strictly regulated, and the dysregulation of protease activity can lead to pathologies such as cardiovascular and inflammatory diseases, cancer, osteoporosis and neurological disorders. Several small-molecule drugs targeting proteases are already on the market and many more are in development. The status of human protease research and prospects for future protease-targeted drugs are reviewed here, with reference to some key examples where protease drugs have succeeded or failed.

Zymogen

A zymogen or pro-enzyme is an inactive enzyme precursor. A zymogen requires a biochemical change (such as a hydrolysis reaction revealing the active site, or changing the conformation to reveal the active site) for it to become an active enzyme.

Proteases, also known as proteolytic enzymes, are enzymes that catalyse the breakdown of proteins by hydrolysis of peptide bonds. Using bioinformatic analysis of the mouse and human genomes, at least 500–600 proteases (~2% of the genomes) have been identified, many of which are orthologous^{1,2}.

Through evolution, proteases have adapted to the wide range of conditions found in complex organisms (variations in pH, reductive environment and so on) and use different catalytic mechanisms for substrate hydrolysis; their mechanism of action classifies them as either serine, cysteine or threonine proteases (amino-terminal nucleophile hydrolases), or as aspartic, metallo and glutamic proteases (with glutamic proteases being the only subtype not found in mammals so far). Proteases specifically cleave protein substrates either from the N or C termini (aminopeptidases and carboxypeptidases, respectively) and/or in the middle of the molecule (endopeptidases) (FIG. 1) and their primary role was long considered to be protein degradation relevant to food digestion and intracellular protein turnover³. However, the pioneering work of Davie and Neurath on trypsinogen activation⁴, followed by that of Davie and Ratnoff⁵ and MacFarland⁶ on the mechanism of blood clotting, led to the concept of protein activation by limited proteolysis. It was discovered that precise cleavage of proteins by proteases leads to a very subtle means of regulation. The blood coagulation cascade has since remained a prototype for a proteolytic cascade in which a signal is passed through a pathway by the sequential activation of protease zymogens. Protease signalling differs from most other signalling pathways by being irreversible, although

exceptions are known (for example, cathepsin C (DPPI) converts to a transferase at higher pH, thereby catalysing the opposite reaction³). By cleaving proteins (BOX 1), proteases are involved in the control of a large number of key physiological processes such as cell-cycle progression, cell proliferation and cell death, DNA replication, tissue remodelling, haemostasis (coagulation), wound healing and the immune response.

To understand why proteases are such important drug targets, and to be able to validate them as targets and identify or design efficient drugs against them, one has to understand the complexity of the biological processes they participate in, the mechanisms of protease activity and regulation and the biochemistry that relates their structure to function. Having obtained this information about a protease in its normal physiological setting, it is then important to understand how these protease properties are modified in a disease state.

Complexity of protease signalling

To identify the role of a given protease in a given biological process and to understand protease signalling in health and disease one needs to know the identity of a protease's physiological substrates — that is, the protease degradome⁷. However, this is not always sufficient for seeing the whole picture of the action of a specific protease, because in addition to the immediate effect on the direct substrate there is usually a downstream effect at the end of the signalling pathway. Understanding the downstream effects and complexity of protease cascades is therefore a challenging but crucial part of protease target selection and validation, because unexpected drug

Department of Biochemistry and Molecular Biology, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia. e-mail: boris.turk@ijs.si doi:10.1038/nrd2092

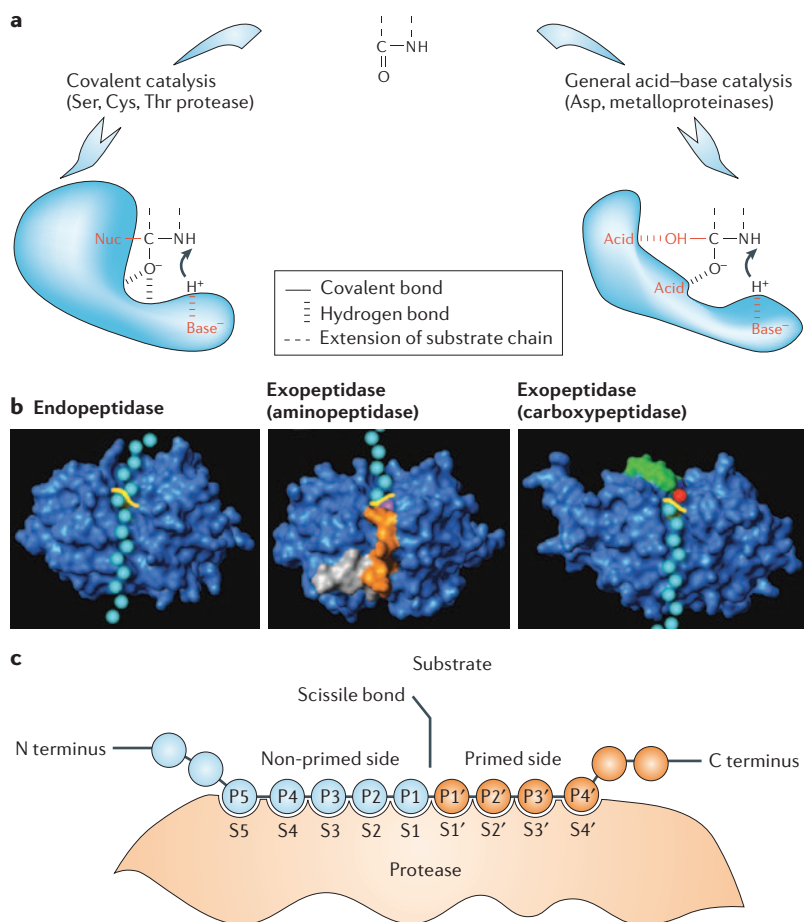


Figure 1 | Protease basics. **a** | Catalytic mechanisms of mammalian proteases. The five major catalytic classes of proteases use two fundamentally different catalytic mechanisms to stabilize the tetrahedral intermediate. In the serine, cysteine and threonine proteases the nucleophile of the catalytic site is part of an amino acid (covalent catalysis), whereas in the metalloproteinases and aspartic proteases the nucleophile is an activated water molecule (non-covalent catalysis). In covalent catalysis, histidines normally function as a base, whereas in non-covalent catalysis Asp or Glu residues and zinc (metalloproteinases) serve as acids and bases. A further difference between the two groups is apparent in the formation of the reaction products from the tetrahedral intermediate, which for cysteine and serine proteases requires an additional intermediate step (acyl-enzyme intermediate). **b** | Modes of substrate cleavage by peptidases (cysteine cathepsins were used as examples): endopeptidase (cathepsin L) and exopeptidases (left, cathepsin H, an aminopeptidase; and right, cathepsin X, a carboxypeptidase). Peptide substrate (schematically represented by cyan balls) runs through the entire length of the active site of an endopeptidase framework (blue) and is cleaved in the middle of the molecule (scissile bond marked yellow). In exopeptidases, substrate binding is structurally constrained (mini-chain in cathepsin H, orange; mini-loop in cathepsin X, green). In cathepsin exopeptidases these additional structural elements also provide negative charge (cathepsin H) to bind the positively charged amino terminus (blue) of the substrate, or positive charge (cathepsin X) to bind the negatively charged carboxyl terminus (red) of the substrate. **c** | Schematic representation of a protein substrate binding to a protease. The surface of the protease that is able to accommodate a single side chain of a substrate residue is called the subsite. Subsites are numbered S1–Sn upwards towards the N terminus of the substrate (non-primed sites), and S1'–Sn' towards the C terminus (primed sites), beginning from the sites on each side of the scissile bond. The substrate residues they accommodate are numbered P1–Pn, and P1'–Pn', respectively. The structure of the active site of the protease therefore determines which substrate residues can bind to specific substrate binding sites of the protease (known as the intrinsic subsite occupancy), thereby determining substrate specificity of a protease. Part a reproduced, with permission, from REF. 129 © (2004) Macmillan Magazines Ltd.

interactions with other proteases in the cascade can have devastating effects on the safety profile of a drug later in development (see section on matrix metalloproteinase (MMP) inhibitors as an example).

Most proteases signal by catalytic activity alone, with the simplest mechanism being direct protease cleavage of the substrate, such as cytokine processing by pro-inflammatory caspases⁸, ectodomain shedding (that is, release of extracellular domains of integral membrane proteins from the cell surface by members of the ADAM (a disintegrin and metalloproteinase) family⁹) or pro-protein maturation in the secretory pathway by furin¹⁰. Proteolysis can also proceed in several consecutive proteolytic steps, in which the substrate is not a protease (a linear pathway). Alternatively, a protease might be involved in a limited number of steps within a cascade-like process that are either not sequential or involve sequential modification/processing of non-protease proteins. Examples of these are regulated intramembrane proteolysis (RIP), which has a fundamental role in cell regulation and signalling, and is promoted by intramembrane cleaving proteases (ICLiPs)¹¹, and the ubiquitin–proteasome system, which is responsible for the proteolytic degradation of all intracellular proteins¹².

The most complex process in protease signalling is a cascade organization, with the most simple cascade involving at least two consecutive proteolytic steps with one protease zymogen being the substrate for an active protease. In principle, all zymogen activations that are not autocatalytic fall into this category. The most studied proteolytic cascade is the blood coagulation cascade, which involves the sequential activation of a series of serine proteases with thrombin activation as its final step, leading to fibrinogen activation and blood clot formation¹³. Sometimes, however, it is difficult to decide whether a pathway is a cascade or not. In several processes, it is simpler to talk about protease networks involving multiple classes of proteases that form a system, such as in extracellular matrix degradation in cancer¹⁴.

Proteases and drug discovery

During cell and tissue development and organism homeostasis, the protease signalling pathways work normally and are tightly controlled. But what happens when the regulation of protease signalling fails? At the substrate-cleavage level, there is either too little or too much proteolysis. Diminished proteolysis as a result of insufficient protease activity mostly originates from genetic irregularities (endogenous proteases¹), excessive inhibitory activity or insufficient activation (both often used by pathogens^{15–17}) and will not be extensively discussed here. By contrast, excessive or inappropriate proteolysis is seldom a result of genetic aberrations but most often results from numerous endogenous and/or exogenous factors, which result in unwanted activation of protease signalling pathways, such as the effect of atherosclerotic plaque formation or blood vessel injury on the blood coagulation cascade, which leads to the appearance of intravascular thrombi¹⁸. So far, inappropriate proteolysis has been found to have a major role in cancer as well as cardiovascular, inflammatory, neurodegenerative, bacterial, viral and parasitic

Box 1 | Protease substrates and their processing in health and disease

Although every single protein is destined to be processed by proteases during its recycling or degradation, a genuine protease substrate is not simply defined as a protein that contains a recognition sequence for a protease that is cleaved in an *in vitro* assay; the protein has to be found simultaneously with the active protease in the same cellular compartment or extracellular location. Proteases involved in protein recycling and degradation are therefore physically separated from the majority of other proteins by being kept in lysosomes or in a proteasome to prevent unwanted proteolysis. Apart from degradation, which is essential for some functions (such as protease signalling in antigen processing and presentation to major histocompatibility class I and II molecules as part of the immune response), processing (cleavage) of a protein by proteases can have the following different consequences for the function of a specific protein:

- Gain of function (for example, activation by limited proteolysis).
- Loss of function (for example, cleavage of inhibitor of caspase-activated DNase (ICAD) by caspase 3 during apoptosis, thereby releasing caspase-activated DNase (CAD) and enabling DNA fragmentation¹¹⁸).
- Switch of function (for example, chemokine processing by matrix metalloproteinases and other proteases⁹³).

Alternatively, processing (cleavage) of a protein by proteases can have no direct consequence for the function or activity of protein:

- Original function of the protein is already lost (for example, during degradation and recycling of inactivated proteins).
- Potential gain or loss of function (for example, stepwise processing of a protein precursor by different proteases, in which the first cleavage is necessary for subsequent cleavage(s) to occur¹¹⁹).
- Ligand removal (for example, deubiquitination¹²⁰).
- Trimming of the amino or carboxyl terminus of the protein (for example, cathepsin C (DPP1); trimming of the amino termini of cysteine cathepsins¹²¹).

Functions 1–3 (function 3 being a combination of functions 1 and 2) are of crucial importance in disease.

diseases. Because excessive proteolysis can be prevented by blocking the appropriate proteases, this area is widely explored by pharmaceutical companies¹⁹.

The history of drugs designed to suppress protease activity dates back to the 1950s. It has been more than 50 years since the first two drugs that affect protease signalling — heparin and the vitamin K analogue warfarin, which target the blood coagulation cascade — were brought into clinical practice to treat thrombosis. However, neither is a direct protease inhibitor: warfarin prevents vitamin K recycling and thereby inhibits post-translational γ -carboxylation of the N-terminal regions of vitamin K-dependent proteases (**Factor II, VII, IX and Xa**) that are required for their biological activity, whereas heparin is an allosteric regulator of antithrombin activity, which indirectly influences the activities of Factor Xa and thrombin. However, the development of antithrombotic drugs has been the subject of several excellent reviews^{18,20,21} and will therefore not be covered in detail here. Nevertheless, coagulopathies, thrombolysis and cardiovascular diseases (see below) remain one of the key disease areas in which proteases are considered to be major targets.

A story of success: ACE inhibitors

With current annual sales exceeding US\$6 billion, the angiotensin-converting enzyme (**ACE**) inhibitors are definitely the major protease inhibitor success story.

ACE inhibitors have been on the market for more than 20 years, with 13 of them currently approved for clinical use (TABLE 1) and several others in clinical trials. They are widely used to treat cardiovascular conditions including hypertension, heart failure and heart attack. ACE (peptidyl dipeptidase A) is a zinc metalloproteinase that acts as a carboxy dipeptidase, and one of the central proteases in the renin–angiotensin system. It catalyses the conversion of **angiotensin I** into angiotensin II (FIG. 2a), a step required for angiotensin receptor activation²².

Development of the current generation of ACE inhibitors (FIG. 2b) is a special story, because 30 years ago neither the sequence nor three-dimensional structure of the enzyme was available. Instead, the structure of carboxypeptidase A, now known to be considerably different from that of ACE, was used for drug design. The ACE inhibitors share a number of common features: low molecular mass; a metal-chelating group (phosphinate, thiol ligand or carboxylate) that binds to the zinc moiety in the enzyme's active site; and they target only the S1, S1' and S2' subsites of ACE with high affinity ($IC_{50} = 0.4\text{--}23\text{ nM}$ for the active metabolite). To achieve good oral bioavailability, most of the inhibitors (except captopril, lisinopril and enalapril) are synthesized as ester prodrugs. Most, except captopril, require only a single dose per day, although they are only efficacious in a limited number of people and have some side effects, including coughing and angioedema^{19,23}.

When the crystal structure of ACE was eventually solved, it revealed that the enzyme is composed of two highly homologous but not identical domains with somewhat different functions and specificities²⁴ (FIG. 2c). This subsequent finding could explain some of the side effects seen with early ACE inhibitors, which were relatively non-selective and inhibited both domains with similar affinities (one- to tenfold difference in K_i ; reviewed in REF. 23). However, the C-terminal domain seems to be primarily responsible for the conversion of angiotensin I to angiotensin II with a major effect on blood pressure regulation, whereas the N-terminal domain seems to be involved in haemoregulation, a hypothesis recently confirmed in a precontracted porcine coronary microarteries model using a new generation of specific inhibitors of the C-terminal (RXPA380) and N-terminal (RXP407) domains²⁵. This new set of inhibitors, designed on the basis of structural differences between the two domains, is much more selective (K_i (C-term)/ K_i (N-term) = 25,000 for RXPA380 (REF. 26)) and should offer improved control of different physiological processes while improving the safety profile.

Another approach to ACE inhibition is the generation of dual inhibitors of ACE and neutral endopeptidase (**NEP**). NEP is a zinc metalloproteinase similar to ACE that is responsible for the degradation of various natriuretic peptides (**atrial natriuretic peptide**, brain-derived natriuretic peptide and C-type natriuretic peptide), and also cleaves angiotensin I and bradykinin. Therefore simultaneous potentiation of atrial natriuretic peptide via NEP inhibition and attenuation of angiotensin II via ACE inhibition leads to complementary effects in

Autocatalytic

The proteolytic conversion of a zymogen protease molecule into its mature form by molecules of the same kind.

Extracellular matrix degradation

The extracellular matrix are non-cellular components of tissues, and are primarily composed of various glycoproteins, proteoglycans and hyaluronic acid. In various disease states such as cancer and osteoarthritis, ECM is degraded by various proteases, including matrix metalloproteinases and the cathepsins.

Table 1 | Protease inhibitors approved for clinical use

Indication	Compound	Company	Target	Protease class
Hypertension, myocardial infarction	Captopril	Bristol-Myers Squibb	ACE	Metallo
	Enalapril	Merck		
	Lisinopril	AstraZeneca		
	Trandolapril	Abbott		
	Zofenopril	Menarini group		
	Ramipril	Aventis		
	Moexipril	Boehringer Mannheim		
	Imidapril	Trinity Pharmaceuticals		
	Perindopril	Daiichi Pharmaceutical, Servier/Solvay		
	Qinapril	Pfizer		
	Fosinopril	Bristol-Myers Squibb		
	Benazepril	Novartis		
	Cilazapril	Roche		
Periodontitis	Periostat	CollaGenex	MMP1, MMP2	
AIDS	Ritonavir	Abbott	HIV protease	Aspartic
	Amprenavir	Vertex Pharmaceuticals, GlaxoSmithKline		
	Fosamprenavir	GlaxoSmithKline		
	Atazanavir	Bristol-Myers Squibb		
	Lopinavir	Abbott		
	Indinavir	Merck		
	Saquinavir	Hoffmann-La Roche		
	Nelfinavir mesylate	Pfizer		
Thrombosis	Ximelagatran*	AstraZeneca	Thrombin	Serine
	Argatroban	Mitsubishi Pharma		
	Lepirudin	Aventis (Hoechst Marion Roussel)		
	Desirudin	Novartis		
	Fondaparinux sodium (indirect)	Sanofi Synthelabo		
Thrombosis, unstable angina	Bivalirudin	The Medicines Company		
Respiratory disease	Sivelestat	Ono	Human neutrophil elastase	
Pancreatitis	Camostat mesilate	Ono	Trypsin-like	
Pancreatitis, inflammation	Nafamostat mesilate	Japan Tobacco	Broad-spectrum	
Cancer	Bortezomib	Millennium	Proteasome	Threonine

*Withdrawn from production because of liver toxicity. ACE, angiotensin-converting enzyme; MMP, matrix metalloproteinase.

the management of hypertension and congestive heart failure. Omapatrilat (Bristol-Myers Squibb), a bicyclic thiazepinone peptidomimetic²⁷, is the most advanced of the dual ACE/NEP inhibitors in development¹⁹. However, although it was shown to be superior over existing agents in reducing hypertension, its registration was halted by the FDA because of increased side effects, such as severe angioedema, in Phase III trials^{23,28} and additional testing has been requested.

A story of failure: MMP inhibitors

The MMPs were the first protease targets seriously considered for combating cancer because of their role in extracellular matrix degradation. Several MMP inhibitors, which can be generally divided into two major groups, hydroxamates and non-hydroxamates, were therefore developed¹⁹. After encouraging pre-clinical results in various cancer models, several of the MMP inhibitors, such as the hydroxamates batimastat

(BB94; British Biotech), marimastat (BB2516; British Biotech) and prinomastat (AG-3340; Aguron), and the non-hydroxamates neovastat (AE-941; Aeterna), which is extracted from shark cartilage, rebimastat (BMS-275291; Bristol-Myers Squibb) and tanomastat (Bay-12-9566; Bayer), were tested in advanced clinical

trials but all failed because of severe side effects and/or no major clinical benefit.

There were several reasons for this failure. First, these compounds are mostly broad-spectrum MMP inhibitors that were developed at a time when only a few metalloproteinases were known and almost nothing was

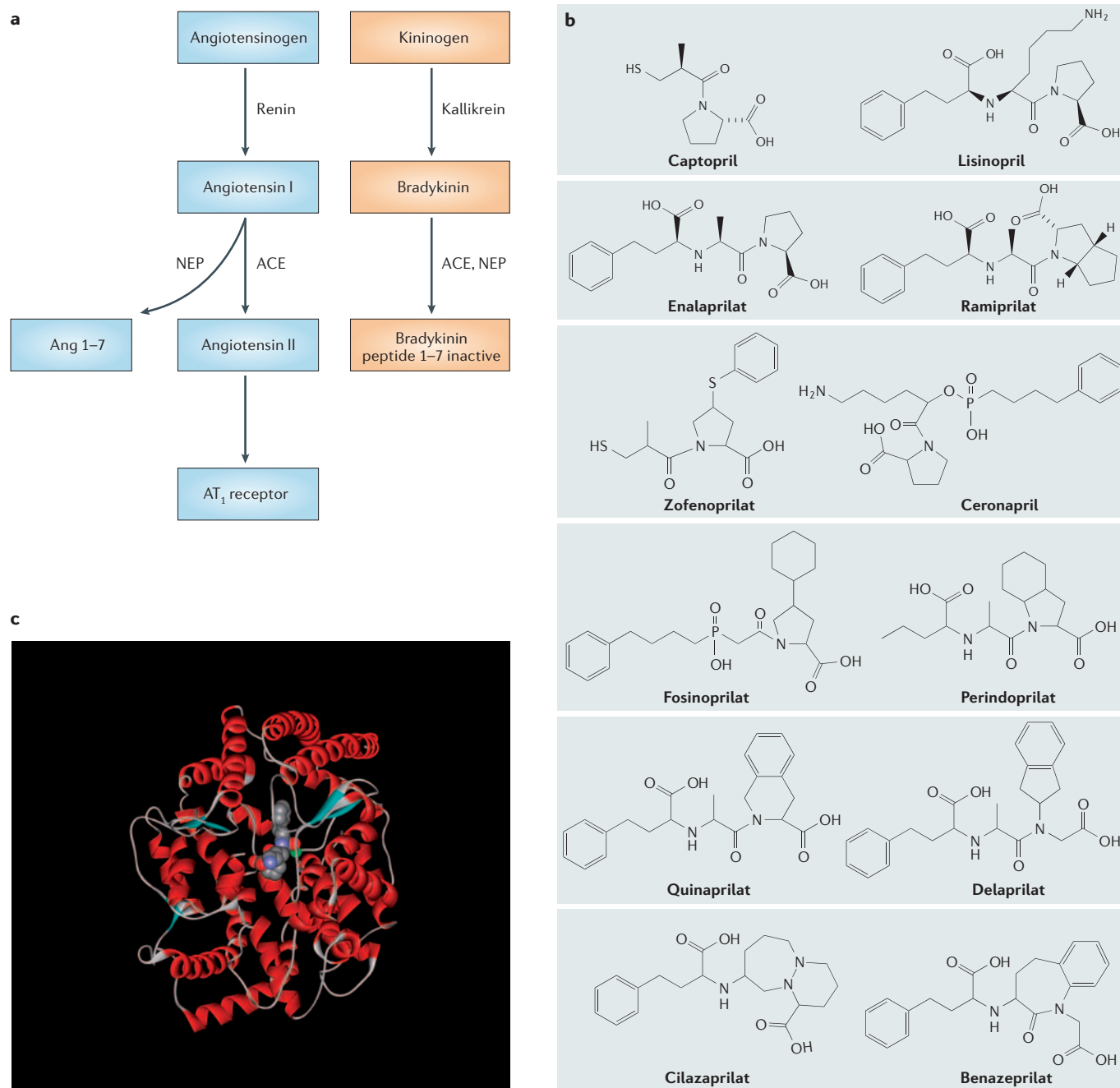


Figure 2 | The renin–angiotensin system and drug discovery. **a** | Schematic representation of proteolysis in the renin–angiotensin system (RAS). Renin is the first and the rate-limiting protease in the RAS, responsible for angiotensinogen cleavage to generate angiotensin I (AngI). AngI is further cleaved by angiotensin-converting enzyme (ACE), resulting in angiotensin II (AngII), which binds to the AngII receptor 1 (AT_1), resulting in AT_1 -mediated increase in blood pressure. Neutral endopeptidase (NEP) is another protease in addition to various natriuretic peptides that can cleave AngI to form Ang 1–7 peptide (blue). Another substrate of ACE is bradykinin, which is a vasoactive peptide, generated from kininogen by kallikrein and related kininogenases. ACE cleavage, however, leads to inactivation of bradykinin (orange). **b** | Chemical structures of ACE inhibitors currently used in the clinic. **c** | Structure of ACE–lisinopril complex. Coordinates were taken from the Protein Data Bank (entry code 1O86).

known about their substrates and biology. Since then, metalloproteinases, including ADAMs and ADAM-TS, both discovered after clinical trials had begun, have been shown to have other important cellular functions and are generally pro-survival. Musculoskeletal side effects observed could be at least partially attributed to the crossreactivity of the inhibitors, as well as to poor understanding of their complex biology, which is still an issue²⁹. Second, MMP inhibitors were ineffective against advanced or late-stage tumours, which were present in most of the patients. This finding was later confirmed in mouse models (for example, a mouse model of pancreatic islet cell carcinogenesis³⁰), but in the initial animal studies inhibitors were given during the early stages of tumour progression. Third, the design of clinical trials was problematic and Phase I was often followed immediately by combined Phase II/III trials without sufficient time for good data analysis (reviewed in REF. 31). Essential, although expensive, lessons about protease inhibitors were learnt from these trials, and the further development of MMP inhibitors as therapeutic agents for cancer was largely abandoned.

Strategies for protease drug discovery

Identifying and validating protease targets. Generally, a protein target should have altered activity in the disease of interest and can often be connected to differential expression of the protein. But this is not always the case with proteases: deregulation of activity resulting from increased activation (which is seen, for example, with caspases) or loss of a key inhibitor would not change the expression profile of the candidate protease. Therefore, analysis of the activity of the proteases *in vivo* is currently one of the hottest areas of protease research^{32,33}. In addition to offering real-time monitoring of protease activity, which was recently achieved for several cysteine cathepsins in a cancer model³⁴, this technology is also highly amenable to testing the efficacy of potential therapeutic inhibitors in living cells, which has advantages over conventional methods of monitoring activity in cell or tissue extracts.

A key part of understanding protease signalling in both health and disease is to identify the endogenous substrates of proteases and many strategies have been applied to achieve this. The so-called 'bottom-up' approach — that is, searching for the processing protease for a known physiological substrate — has been used successfully to identify the substrates of several proteases including, for example, those involved in the renin–angiotensin system³⁵, ACE³⁶ and, more recently, the ICLiPs¹¹.

The advent of *in silico* biology, genetics and proteomics has also helped in the search for physiological protease substrates⁷. For example, once the extended substrate specificities of proteases have been determined by using phage display peptidic libraries, combinatorial fluorescent substrate libraries or positional scanning libraries based on covalent inhibitors^{37–40}, bioinformatics can be carried out on a genome-wide basis to find potential substrates according to the protease substrate specificities determined. However, because the specificities of proteases

on small synthetic substrates are often different from the specificities on protein or long peptidic substrates, this method requires additional confirmation by other techniques. Probably the best example of this approach is the successful identification of caspase substrates⁴¹, but it has also worked for granzyme B³⁸ and for metalloproteinases⁴².

Genetic approaches to substrate searching generally consist of classical knockin/knockout animal models or, more recently, RNA interference (RNAi) gene knockdown. The use of knockout animals has been somewhat limited because of challenges such as embryonic lethality and upregulation of compensatory protease functions that provide misleading phenotypes. Nevertheless, the approach has been useful in several cases, such as the identification of various metalloproteinase substrates⁹ and, of particular interest to the pharmaceutical industry, the confirmation of an essential role for β -site APP-cleaving enzyme (BACE) in the production of β -amyloid in mouse, which validated BACE as a target for **Alzheimer's disease**⁴³. Another approach, which is often used in target validation, is to combine an animal disease model with a classical protease mouse knockout model, which enables one to address specifically the role of a protease of interest. This approach has been used to study the role of cysteine cathepsins in different cancer cell models, such as the Rip1–Tag2 mouse pancreatic islet tumour model⁴⁴. RNAi, although a powerful tool in target validation, has been less frequently used in substrate identification, although one example is the use of RNAi to explore the roles of different MMPs in cancer⁴⁵.

It should be noted that genetic studies in model organisms such as yeast, *Caenorhabditis elegans*, *Drosophila*, *Xenopus* and zebrafish (*Brachydanio rerio*), although mostly devoted towards general identification of signalling pathways and not specifically towards protease identification, provided some major discoveries in the protease field. An example was the discovery that the proapoptotic *CED-3* gene in *C. elegans* is highly homologous to the cysteine protease interleukin-1 β converting enzyme (**caspase 1**), which revolutionized the entire apoptosis field⁴⁶.

The yeast two-hybrid system, which is widely used in the identification of physiological ligands of proteins, was shown to work also in the protease field and led to the identification of monocyte chemoattractant protein 3 (**MCP3**) as an MMP2 substrate using the haemopexin domain of MMP2 as a bait⁴⁷. However, this method has not been widely adopted by the protease research community.

Advances in proteomics tools (for example, the development of new fluorescence probes for two-dimensional electrophoresis with enhanced sensitivity, and of new methods for identifying neo-amino termini of peptides by liquid chromatography and mass spectrometry), which can be used in combination with genetic methods or in specific cellular or *ex vivo* models, are being applied in the hunt for protease substrates. Several protease substrates have been recently identified in this way, including MMP substrates such as chemokines and

Positional scanning library
A combinatorial chemistry approach in which individual positions in the molecule are kept defined with the other positions being degenerate to yield 'positional libraries'. These compound mixtures are then tested. The most active moieties of the defined (and therefore known) positions are subsequently combined to yield the screening result, the active molecule.

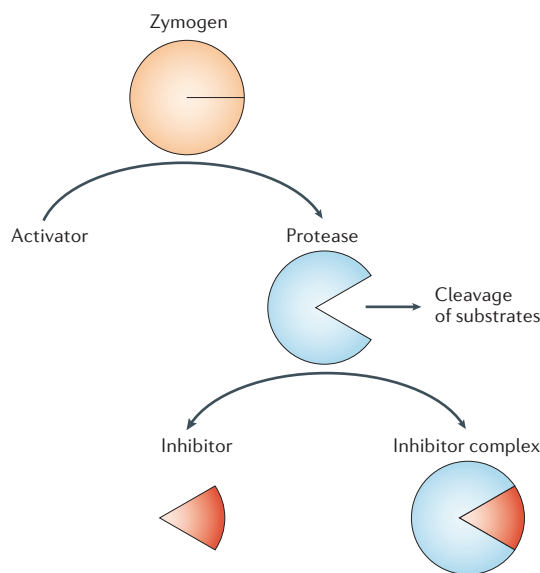


Figure 3 | Regulation of protease activity. The fundamental mechanisms governing activity are conserved in most proteases. Latent protease zymogens await an activation signal, which can come from an allosteric activator or another protease. Once active, substrate and inhibitor compete for protease binding, and the outcome is defined by the local concentration of inhibitor. The double-headed arrow depicts reversible inhibition. Adapted from REF. 129 © (2004) Macmillan Magazines Ltd.

cytokines (substrates of MT1-MMP and MMP2 (REF. 48)) and granzyme B substrates (procaspase 3 and HOP (Hsp70/Hsp90 organizing protein))⁴⁹. The power of proteomics therefore lies in the possibility of resolving complex problems, such as the identification of 92 different caspase substrates among >1,800 different proteins identified in Fas-mediated apoptosis in Jurkat cells⁵⁰. Proteomics is therefore highly suitable for the identification and validation of potential protease drug targets.

Some validated protease targets, such as renin, ACE, blood coagulation proteases, γ - and β -secretases and dipeptidyl peptidase IV (DPPIV), have been identified through searching for the processing protease for a known physiological substrate. However, often we can only see the final phenotype of the disease and we have to start searching for protease targets from that point. Because of the complexity of (protease) signalling pathways, it is important to identify, if possible, the earliest event in disease progression, because the earlier in a cascade of events that an inhibitor acts, the more effective at blocking the ultimate effects the inhibitor will be. Blocking a step at the end of a cascade would require larger doses of an inhibitor to achieve comparable effectiveness, and this might lead to safety and toxicity issues. However, there might be some trade off here with selectivity: blocking an early step could potentially require much tighter regulation and as such the development of more specific compounds. The anticoagulant pathway provides a good example of this, because it is necessary to decide whether to target Factor Xa, which is a mediator, or thrombin,

which is the final effector, considering that one molecule of Factor Xa within the prothrombinase complex can generate more than 130 thrombin molecules per minute²⁰. Nevertheless, different companies develop direct inhibitors against both thrombin and Factor Xa. However, insufficient understanding of the biological role of a protease target can lead to the development of a drug with unexpected and severe side effects, as happened with the first-generation MMP inhibitors developed for cancer and rheumatoid arthritis⁵¹.

Regulation of protease activity

The action of proteases is tightly controlled to prevent improper cleavage of signalling molecules (FIG. 3). Protease activities are regulated at the transcriptional level by differential expression, and at the protein level by activation of inactive zymogens and by the binding of inhibitors and cofactors. Activation can be either auto-catalytic or catalysed by other proteases, as described above. Alternatively, proteases are sometimes activated with the assistance of an activation complex, such as the death-inducing signalling complex (DISC), which is involved in the activation of various caspases⁵¹. As such, preventing protease activation could be an attractive area for potential new therapeutics, and is particularly amenable to RNAi approaches⁵².

Protease activity is also regulated by cofactors, proteins that reversibly bind to proteases and/or inhibitors and affect their final activity, often in an allosteric manner. The blood coagulation cascade provides the most well-known examples, with tissue factor protein regulation of Factor VIIa activity⁵³ being the best-known cofactor to be explored by the pharmaceutical industry²⁰. Although they are not proteins and therefore not cofactors in the usual sense, glycosaminoglycans (GAGs) are extremely important allosteric regulators of proteases and their physiological inhibitors. Their best-characterized interaction is with antithrombin, the major inhibitor of several blood coagulation proteases and a subject of therapeutic intervention⁵⁴ (TABLE 1). Heparin (one of the best-characterized GAGs) and its derivatives are therefore often referred to as indirect and allosteric inhibitors of blood coagulation proteases, primarily Factor X.

The other major regulators of protease activity are their endogenous or exogenous inhibitors. In this regard, nature has been economical because the number of endogenous human inhibitors identified is substantially smaller than the number of proteases identified (only 105 in humans), with only a slightly higher number identified in mice¹²⁴ (see MEROPS database, Further information). This might be explained by the low specificity of the inhibitors for their target proteases, meaning that one inhibitor can inhibit several proteases. For example, the main inhibitors for all the metallo-endopeptidases, of which there are more than 180, consist of only four proteins: tissue inhibitor of metalloproteinase 1 (TIMP1), TIMP2, TIMP3 and TIMP4. It is common to classify an endogenous protein inhibitor according to the structural mechanism by which it inhibits its target protease^{55,56}. The second method is to examine the physiological relevance of inhibition⁵⁷ (BOX 2).

Box 2 | **Classification of protease inhibitors**

Protease inhibitors can be classified according to their mechanism of inhibition or by the physiological outcome or relevance of the inhibition. Four different mechanisms for protease inhibition have been described so far⁵⁶:

- Canonical inhibitors (lock-and-key) bind to their targets in a substrate-like manner to form an almost Michaelis-type complex, a principle used by serine protease inhibitors (serpins)¹²².
- Exosite-binding inhibitors directly block the active site by binding adjacent to it and covering it partially in a substrate-like manner. This principle was observed initially with cystatins and cysteine cathepsins¹²³.
- Inhibitors might also bind in a quasi-substrate-like manner that is a combination of the canonical and exosite-binding mechanisms (for example, as with tissue inhibitor of metalloproteinases¹²⁴).
- Allosteric inhibitors — such as X-linked inhibitor of apoptosis protein, which inhibits caspase 9 — prevent dimerization by binding in the interdomain region away from the active site¹²⁵.

With the exception of allosteric inhibitors all other endogenous protein inhibitors are mostly competitive.

The physiological inhibition of proteases depends on temporal and spatial co-localization of the protease and its inhibitor, their relative concentrations and binding kinetics; this leads to two major categories of physiological inhibitors⁵⁷:

- Emergency type inhibitors (characterized by a large excess concentration of inhibitor, rapid binding and no co-localization with proteases) with the major function to block any protease activity in an inappropriate compartment (for example, cystatins⁵⁷).
- Regulatory type inhibitors (often co-localized with proteases) which are responsible for the fine regulation of protease activity.

Regulatory type inhibitors can be subdivided into the following categories:

- Threshold inhibitors: low concentration and rapid binding, their major function is to neutralize accidental protease activation (for example, inhibitors of apoptosis¹²⁶).
- Buffer-type inhibitors: weak binders, these temporarily block proteases to prevent inappropriate activity and can be easily displaced from the complex (for example, propeptides of cathepsins¹²⁷).
- Delay-type inhibitors: often irreversible, slow binders that enable protease activity for a limited amount of time (for example, antithrombin⁵⁴).
- Pro-inhibitors: synthesized as inactive, and require proteolytic processing to become active (for example, invariant chain p41 fragment resulting from the proteolytic processing of major histocompatibility class II molecules¹²⁸).

Of the several approaches to targeting protease signalling, protease inhibition is the most widely explored so far. In addition to directly blocking the activity of the target protease, protease inhibitors also block activation of the downstream proteases, or affect protease activity in complexes with cofactors, thereby covering all three mechanisms of protease regulation reasonably well. The rest of this review is therefore devoted to alternative approaches to inhibit mature proteases.

Direct inhibition: large versus small

Large molecules. Once a protease target is validated, using an exogenous modulator to mimic its physiological inhibition is an attractive therapeutic approach, but unfortunately it is only seldom possible to do this for protease targets. Nevertheless, there has been some success in the blood coagulation area, in which novel large-molecule inhibitors based on proteases used by various blood-sucking animals have been developed.

The first examples were derivatives of hirudin, a potent 65-amino-acid reversible thrombin inhibitor purified from the medicinal leech. Both recombinant hirudin

(desirudin) and bivalirudin (a synthetic 20-amino-acid polypeptide based on the inhibitory sequence of hirudin with an improved safety profile compared with hirudin) were approved as anticoagulants for clinical use (TABLE 1). A second example is recombinant nematode anticoagulant protein (rNAP) from the hookworm *Ancylostoma caninum*, initially developed by Corvas Pharmaceuticals, and later acquired by Nuvelo. Nuvelo recently reported that rNAP has an acceptable safety profile and is well tolerated in doses up to 10 µg per kg body weight in patients being treated for acute coronary syndromes (ACS) in a Phase IIa trial⁵⁸. Other examples, such as the natural peptidic inhibitors of Factor X, tick anticoagulant peptide (TAP)⁵⁹ and antistasin⁶⁰, have served as prototypes for inhibitor design.

The Bowman–Birk inhibitor concentrate isolated from soybean⁶¹ is a mixture of several serine protease inhibitors that primarily has chymotrypsin inhibitory activity but also some trypsin inhibitory activity, and was shown to have a positive effect in treating oral cancer in Phase II clinical trials several years ago. However, not much research has been done on this agent since. It is thought to mediate its effects by inhibiting the chymotryptic activity of the proteasome⁶², and so works in a similar way to bortezomib (Velcade; Millennium), a proteasome inhibitor that was recently approved for cancer treatment (TABLE 1).

Another large-molecule approach is to use humanized antibodies, a method initially developed for other protein target classes⁶³. Although research into neutralizing antibodies against protease targets is at an early stage, there are already some examples of this drug class in preclinical and early clinical development. These include neutralizing antibodies against uPA (urokinase type plasminogen activator)⁶⁴ and cathepsin B⁶⁵ (Krka Pharmaceuticals) for potential cancer treatment, against tissue factor (hATR-5; Chugai Pharmaceuticals) for a wide range of thrombotic disorders and against Factor VIIa (mAb-LE2E9; ThromboGenics) for venous thrombosis treatment²⁰.

Small molecules. The majority of drug discovery efforts in the field of protease inhibition are still devoted to small molecules. Ideally, a small-molecule inhibitor would be completely selective for its target, have excellent absorption, distribution, metabolism, excretion and toxicology (ADMET) parameters for acute and chronic use, be orally available and be amenable to daily (or less frequent) administration. Obviously, no such compound exists for any target, but for protease targets there are several specific challenges. First, research with protease-knockout mice has shown that even the most simple protein-processing events often have complex consequences, and that protease signalling in its broadest sense is often only a small part of disease biology in the whole organism and not even always the best aspect to target. Second, there is no inhibitor that is completely selective for a single protease target. Finally, many potent protease inhibitors have poor safety profiles. Indeed, many of the best protease inhibitors have been a compromise in which a degree of efficacy has had

Mature protease
The catalytically active form of a protease that results from proteolytic processing of its zymogen (inactive pro-enzyme).

to be sacrificed to develop a compound with improved bioavailability and tolerable side effects, depending on the disease to be treated.

Several general strategies for small-molecule screening are also applicable to the design of protease inhibitors. High-throughput screening procedures using a company's entire compound collection are seldomly used nowadays. More often, screening for protease inhibitors involves virtual screening, fragment-based screening and screening of small focused libraries of compounds (typically several thousand compounds). The methods can be also combined, as exemplified by Novartis in its screen for inhibitors of the Dengue virus NS3 protease⁶⁶. However, caution is required with this approach because it can result in the generation of hits that are non-selective for the target protease⁶⁷.

One limitation with screening for protease inhibitors is that some of these methods, such as virtual screening and fragment-based screening, require the three-dimensional structure of the target protease. However, knowledge of the three-dimensional structure of the protease is, although not essential, beneficial in all screening methods and for inhibitor design in general. For example, the development of drugs against the aspartic protease of the HIV-1 virus was only really enabled once the three-dimensional protease structure was available; as a result the first drugs advanced to the market within 10 years of initial target identification^{19,68} (TABLE 1).

One of the major factors for pharmaceutical companies to consider in protease drug discovery is whether to develop reversible or irreversible inhibitors. Although designing an irreversible inhibitor is fairly simple for some protease classes, and can be achieved by attaching the substrate recognition sequence to a warhead, the question is whether such an inhibitor would be sufficiently selective. Although both reversible and irreversible inhibitors would largely ablate the activity of the target, they would have differential activity against other off-target proteases. A reversible inhibitor, especially one with good selectivity, would only partially block the activities of other proteases; however, an irreversible inhibitor would, sooner or later (depending on the binding kinetics), block all proteases that it binds. Indeed, several irreversible inhibitors that have been developed have had this selectivity problem, such as the chloro- and fluoromethylketones, which were originally developed against caspases but also non-selectively inhibited most of the cysteine cathepsins and legumain⁶⁹. So, for this reason, irreversible inhibitors are most useful as tools for elucidating the catalytic mechanisms of a protease, whereas they are not the most suitable strategy for a therapeutic that is intended to be used chronically to treat a disease. However, in acute situations, when only a single dose (or a small number of doses) would be needed, an irreversible inhibitor with reasonable selectivity could be acceptable if it shows benefit in treating a life-threatening disease and if it could be cleared rapidly enough to minimize unfavourable side effects.

The ideal inhibitor, therefore, would be a non-covalent reversible inhibitor, because such an inhibitor would generally provide better selectivity and should cause

less side effects than covalent inhibitors. However, many of the protease inhibitors in development are of the covalent kind, and have been developed with the same rationale as used for the design of irreversible inhibitors¹⁹. Non-covalent inhibitors are much more difficult to design. Possible methods for generating non-covalent inhibitors include designing transition-state analogues (mimicking the transition state during substrate hydrolysis) or substrate variants, but not substrate analogues, which generally produce covalent inhibitors⁷⁰, or using fragment-based screening. So far, a few non-covalent reversible inhibitors have been developed and some have reached the market, such as the serine protease inhibitor argatroban (GlaxoSmithKline)¹⁹ (TABLE 1), whereas non-covalent inhibitors of, for example, cysteine proteases are still in the early experimental phase^{71,72}.

Bioavailability remains one of the key issues in protease inhibitor optimization. Diminishing the peptidic character of the inhibitor often improves its bioavailability and possibly prevents nonspecific degradation of the compound by endogenous proteases (which leads to rapid clearance of the compound and thereby lower efficacy). Nevertheless, there is no general solution to this problem. Good bioavailability can be observed in animal models, but then is not reproducible in humans (this problem is not unique to protease inhibitors). One strategy to solve this problem is to make a prodrug that is converted to the active metabolite in the body. Using this approach improves both bioavailability and the clearance time, making the drug more efficient. The ACE inhibitors, which are largely manufactured as ester prodrugs (see above), embody this approach.

Another factor in protease drug design is whether the inhibitor is competitive or non-competitive. With the exception of the first antiprotease therapeutic — heparin, an allosteric indirect inhibitor of blood coagulation proteases — all other protease inhibitors currently in use or in development are competitive inhibitors. Discovering competitive inhibitors is generally easier because of the simplicity of activity tests used, but there are also disadvantages with this approach. For example, competitive inhibitors can cause saturation of substrate competition, which necessitates the use of substantially higher doses of the compounds and narrows the safety window.

Although much more difficult to develop, allosteric small-molecule inhibitors could be useful against many proteases by, for example, binding to protease exosites and preventing protein substrate binding or recognition. A recent breakthrough in allosteric protease inhibitor design has been achieved with the development of the first allosteric caspase inhibitors by Sunesis. Using a method called tethering, which is based on the formation of a reversible disulphide bond between a free cysteine on the protease and a thiol-containing fragment from the screening library⁷³, two inhibitors were identified: 5-fluoro-1*H*-indole-2-carboxylic acid (2-mercapto-ethyl)-amide (FICA) and 2-(2,4-dichlorophenoxy)-*N*-(2-mercapto-ethyl)-acetamide (DICA). These were shown to bind to a cysteine residue in the vicinity of the active site cleft of caspases 3 and 7, respectively, locking the specificity loops of the protease into

Warhead

A reactive functional group, which covalently binds to amino-acid residues of the target enzyme.

a zymogen-like conformation, thereby abolishing enzymatic activity⁷⁴. A similar approach could potentially also be used for other protease classes.

Upregulating proteases: inhibiting the inhibitors

In some cases, it might be necessary to upregulate the activity of a protease rather than inhibit it, and an attractive strategy for achieving this is to block the activity of endogenous protease inhibitors. A group from the Burnham Institute used compounds based on a polyphenylurea scaffold to inhibit the inhibitor of apoptosis (IAP) family of proteins, which repress caspase activity. These compounds therefore de-repressed the caspases and stimulated apoptosis of cancer cells. Initial studies in cell culture showed a significant apoptotic effect and good selectivity for cancer cells, including their sensitization towards other chemotherapeutic drugs. Moreover, the compounds were also effective in tumour xenograft models in mice and showed no toxicity, thereby validating protease inhibitors, as well as proteases themselves, as relevant targets⁷⁵. These compounds are currently in preclinical development.

Another approach for blocking endogenous protease inhibitors is to use antisense molecules. Aegera Therapeutics recently announced the initiation of Phase I clinical trials for AEG35156, a second-generation antisense oligonucleotide inhibitor of X-linked IAP protein (XIAP), which potentiates cancer cell apoptosis. Monoclonal antibodies might also have potential; the now commercially available antiplasminogen activator inhibitor 1 (PAI1) monoclonal antibody 33H1F7 (REF. 76) was shown to transform active PAI1 to a tPA substrate, thereby accelerating thrombolysis in a rat mesenteric artery model⁷⁷.

New drugs in development

Many new protease inhibitors are currently in development, with at least 50 different proteases being considered as potential targets. Many of these targets were identified using animal models, particularly rodent knockouts, and were then used to screen for inhibitors, whereas others were rationally developed using structure-based drug design. Because different proteases use different reaction mechanisms (FIG. 1), completely different strategies for inhibitor design are often required. Below, the drug development efforts against four different types of protease target, one from each mechanistic class, is discussed.

Fighting hypertension with renin inhibitors. Renin is an aspartic protease responsible for activating angiotensinogen to generate angiotensin I (FIG. 2a), and as such offers an alternative to ACE inhibitors in the treatment of hypertension and in end-organ damage²². Renin inhibition might be advantageous over ACE inhibition because renin has specificity for a single substrate (angiotensinogen) and therefore could potentially cause fewer adverse effects⁷⁸. However, early renin inhibitors were substrate analogues and tended to be degraded quickly and to have low bioavailability. Zankiren (A-72517; Abbott), although a peptidic transition-state analogue

inhibitor, was the first to show improved proteolytic stability and bioavailability in several animal models (maximum of 53% in the dog) and also showed good efficacy in humans⁷⁹.

However, the latest generation of renin inhibitors are non-peptidic. Currently, Speedel is most active in this field. Together with Novartis, Speedel has developed the alkanecarboxamide aliskiren (SPP100) (in-licensed from Novartis in 1999)⁸⁰ (FIG. 4a). Aliskiren was found to be superior to other renin inhibitors and at least equivalent to ACE inhibitors and AT₁-receptor blockers in animal models⁸¹. Phase III data for aliskiren either as monotherapy or as a combination therapy were disclosed in September 2005 and showed excellent results in outcome measures of blood pressure reduction, safety and potential for improved end-organ protection resulting from efficient blockade of the renin-angiotensinogen system. On the basis of these results, Novartis filed a New Drug Application (NDA) with the FDA for aliskiren in April 2006.

Speedel also acquired the renin programme from Roche, which is thought to focus on 3-alkoxy-4-aryl-piperidines and their further optimization. One such compound, the 3,4,5-trisubstituted piperidine Ro 66-1132 (FIG. 4a), was effective in lowering blood pressure and reducing end-organ damage in double transgenic rats engineered to express human renin and angiotensinogen genes⁸². In October 2005, Speedel announced that two of these inhibitors (SPP630 and SPP635) had been successfully tested in a microdosing trial in human volunteers and in preclinical studies in the aforementioned double transgenic rat model had showed excellent bioavailability (>30% in humans; 70–90% in rats) and half-life (30 h in humans). SPP635 entered Phase I trials in October 2005.

The third group of renin inhibitors from Speedel was developed in collaboration with Locus Pharmaceuticals, and in June 2005 Speedel announced the discovery of a new series of promising lead compounds from this series (SPP800), although no structural data were disclosed.

DPPIV inhibitors for type 2 diabetes. Type 2 diabetes is a major debilitating disease that is caused by abnormal glucose homeostasis and is commonly associated with obesity. In the late 1980s, the incretins, including glucagon-like peptide 1 (GLP1) and glucose-dependent insulinotropic peptide (GIP), were identified as hormones capable of lowering blood glucose levels, which makes them potential targets for antidiabetic therapy. GLP1 and related peptides were also found to be rapidly degraded by the serine aminopeptidase DPPIV (CD26) *in vitro*⁸³ and *in vivo*⁸⁴, which suggested that there was therapeutic potential for DPPIV inhibitors. The GLP1-stabilizing effect of DPPIV and validation of it as an antidiabetic target was confirmed by the observation that DPPIV-deficient animals experienced improved glucose tolerance^{85,86}.

DPPIV is predominantly a membrane-bound protease and primarily cleaves dipeptides after proline residues, a feature exploited in the development of DPPIV inhibitors^{87–90}. Currently, there are several inhibitors in advanced clinical trials (FIG. 4b). The most advanced is vildagliptin

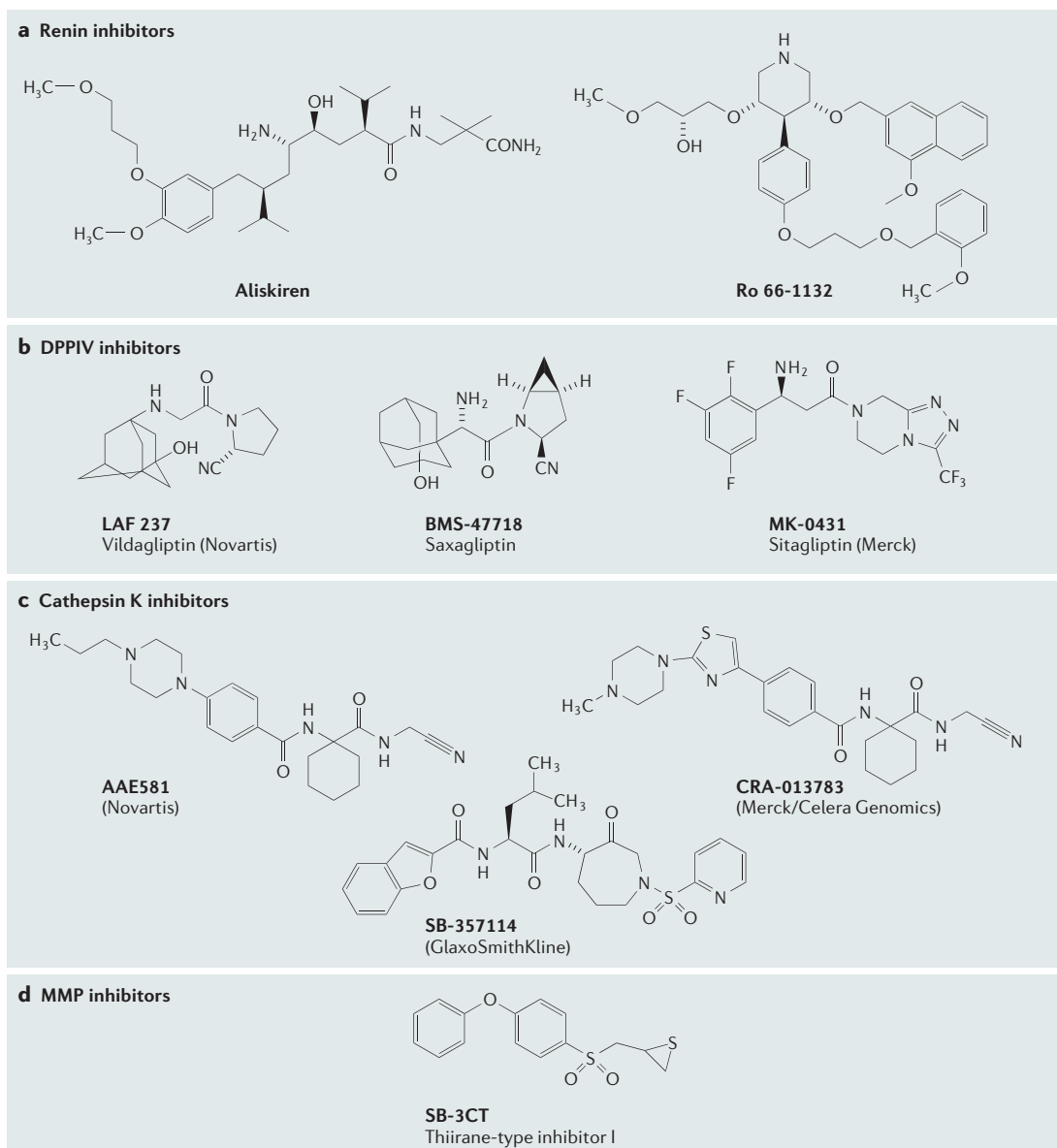


Figure 4 | **New protease inhibitors in development.** **a** | Renin inhibitors (aspartic proteases). **b** | Dipeptidyl peptidase IV (DPPIV) inhibitors (serine proteases). **c** | Cathepsin K inhibitors (cysteine proteases). **d** | Matrix metalloproteinase (MMP) inhibitors.

(LAF237; Novartis), which forms a reversible covalent bond between the nitrile of the inhibitor and the hydroxylate of the enzyme. In Phase III trials, vildagliptin was well tolerated without weight gain and significantly decreased blood glucose levels. In addition, vildagliptin was reported to improve and sustain pancreatic islet cell function and insulin sensitivity over a 1-year period. An NDA for vildagliptin was filed in March 2006 by Novartis.

Other DPPIV inhibitors currently in development include sitagliptin (MK-0431), a non-peptidic heterocyclic compound developed by Merck that entered Phase III trials in 2004 and for which relatively few data have been disclosed so far. Nevertheless, in February 2006 Merck filed an NDA for sitagliptin (Januvia). This non-covalent reversible inhibitor was found to be

selective against other related proteases and showed good efficacy in preclinical studies⁹¹. Sitagliptin was shown to be efficient in improving glycaemic control and was well tolerated in three Phase II studies involving more than 1,000 people. Another company working with reversible non-covalent inhibitors is Probiobdrug, which has developed several compounds, one of which (P93/01, also known as PSN9301) is currently in Phase II trials conducted by Prosidion⁸⁹. Finally, saxagliptin (BMS-477118; Bristol-Myers Squibb) is a covalent reversible DPPIV inhibitor currently in Phase III trials⁹² that showed good selectivity and oral bioavailability in a Phase II study.

All the DPPIV inhibitors developed so far have been well tolerated in trials. However, DPPIV is also involved

in chemokine processing⁹³ and therefore some precautions regarding long-term treatment should be considered, especially as DPPiV-knockout animals experienced minor immune defects (including decreased cytokine levels and a decreased number of T cells and natural killer cells⁹⁴).

Cathepsin K and osteoporosis. Cathepsin K is a lysosomal cysteine cathepsin predominantly located in osteoclasts and is the major enzyme involved in bone resorption. The first evidence for this role came from a genetic study on **pycnodysostosis**, a rare genetic disorder associated with severe defects in bone growth, which revealed that an inactivating mutation in the gene encoding cathepsin K is a causative factor⁹⁵. Further evidence then came from the studies on cathepsin K-deficient mice, which developed osteopetrosis from lack of bone resorption^{96,97}. This was later supported by a study on transgenic mice that overexpressed cathepsin K and which experienced reduced trabecular bone volume as a result of increased bone resorption⁹⁸.

Recent clinical data suggest that cathepsin K inhibitors are advantageous over existing osteoporosis therapies, including bisphosphonates, because they not only prevent bone loss but also allow bone reformation. However, the development of cathepsin K inhibitors actually preceded the aforementioned clinical studies and started as soon as cathepsin K was discovered as the major cysteine protease in osteoclasts⁹⁹. Development of cathepsin K inhibitors was particularly difficult because of species-specific variation in key residues involved in substrate and inhibitor recognition, which resulted in substantially different affinities of the compounds for the target in rodents and humans, thereby making rodent disease models redundant¹⁰⁰. Various *in vitro* cathepsin K models were therefore developed and the high degree of similarity between human cathepsin K and that of non-human primates enabled researchers at GlaxoSmithKline to generate a cynomolgus monkey model for testing cathepsin K inhibitors¹⁰¹.

There are now four cathepsin K inhibitors in clinical trials: AAE581 (balicatib (FIG. 4); passed Phase II) and AFG-495 (no structural data disclosed; Phase I) developed by Novartis; SB-462795 developed by GlaxoSmithKline (Phase II); and an undisclosed inhibitor resulting from a joint Merck/Celera programme, possibly similar to the CRA-013783 compound¹⁰² (Phase I). All these inhibitors are presumably reversible and form a covalent bond with the thiol group of the active-site cysteine. As the substrate binding sites S1, S2 and S1' are almost identical among different cathepsins¹⁰³, the inhibitor selectivity must be obtained from the cumulative contribution of multiple interactions with the enzyme, with at least some of them extending to the S3 subsite.

AAE581, a peptidic nitrile and the most advanced of the aforementioned drugs, is highly specific for human cathepsin K ($K_i = 0.7\text{--}1.4\text{ nM}^{104,105}$), and has >1,000-fold selectivity for cathepsin K over other cathepsins *in vitro*. However, a recent study by researchers from Merck has shown that AAE581 and related basic inhibitors, such

as the tri-ring aminonitrile CRA-013783, are 10–100-fold more potent against other cathepsins in cellular assays because of their lysosomotropic character and subsequent lysosomal entrapment. This resulted in an increased off-target effect in a cellular assay of cathepsin S-mediated antigen presentation¹⁰⁵. Nevertheless, AAE581 successfully reduced collagen breakdown and bone resorption in Phase II trials, which were conducted on 140 post-menopausal women receiving once-a-day treatment. No adverse side effects were reported and Novartis is currently recruiting for Phase III trials to begin in 2006.

During the past two years, researchers at GlaxoSmithKline have published a large amount of work on the development of cathepsin K inhibitors, in which they have explored all the substrate-binding subsites and different warheads. This work leads to the conclusion that the most promising compounds are the nitriles (developed by Novartis and Merck) and ketones, such as the GlaxoSmithKline compound SB-462795, which is an α -heteroatom cyclic ketone^{106,107}. Although several non-covalent cathepsin K inhibitors, such as arylaminoethyl amides⁷², were developed, none of them has progressed beyond the experimental phase.

In addition to treating osteoporosis, cathepsin K inhibitors might also be useful in other bone-related pathologies, such as **Paget's disease**, osteoarthritis and even bone metastasis¹⁰⁰.

New strategies for MMP inhibitors. The failure of early MMP inhibitors in cancer and arthritis (discussed earlier) means that interest in the development of new MMP inhibitors has been severely dampened. Most companies began to investigate the possibilities of targeting MMPs for other indications, and have generally abandoned the hydroxamates because of their lack of specificity.

CollaGenex Pharmaceuticals has focused on modified non-antibiotic doxycycline derivatives. Periostat (TABLE 1), a broad-spectrum inhibitor of MMPs, was approved by the FDA for the treatment of periodontal disease in 1998 and is currently in Phase III trials for chronic adult periodontitis. Following a positive result from two Phase III double-blinded, placebo-controlled clinical trials, CollaGenex has recently filed an NDA for another doxycycline derivative, Oracea, which is the first orally administered, systemically delivered drug to treat rosacea, a skin disease characterized by the appearance on the face of inflammatory lesions, redness and telangiectasia (spider veins). CollaGenex also announced positive results from a Phase II study of incyclinide (also known as Col-3, CMT-3 or metastat) for the treatment of acne. Although incyclinide previously failed in cancer studies, it showed positive effect in preclinical studies as an anti-inflammatory agent in rosacea. However, it should be noted that it is not completely clear whether MMPs are the only targets of doxycycline.

Other diseases for which MMPs are considered to be relevant targets are primarily associated with inflammation and the cardiac system; for example, Procter & Gamble has a compound PG-116800 in Phase II trials

for treating patients after a heart attack, but results presented at the 2005 American College of Cardiology Annual Meeting showed that the treatment has so far shown no benefit over placebo. MMP inhibitors are also being investigated for vascular disease, myocardial infarction, stroke, acute lung injury and chronic obstructive pulmonary disease (COPD); however, research is mostly in the preclinical phases⁴⁵.

Meanwhile, cancer has not been completely forgotten. A group at the University of Notre Dame has developed novel thirane-type mechanism-based MMP inhibitors specific for gelatinases (MMP2 and MMP9). The prototype compound SB-3CT¹⁰⁸ (FIG. 4), although having relatively low aqueous solubility, increased survival in an aggressive mouse model of T-cell lymphoma¹⁰⁹ and also rescued neurons in a transient focal cerebral ischaemia model¹¹⁰, thereby showing good therapeutic potential. The thirane inhibitors are being further developed and new selective MMP inhibitors with improved solubility were recently reported¹¹¹.

Proteases as biomarkers

One of the biggest areas of research at the moment is the identification and validation of biomarkers for drug development and diagnostics. Here, protease biology could also come into its own. Because of their differential expression in disease, proteases and their inhibitors could be used as biomarkers for early diagnostics and in certain cases as prognostic markers to enable selection of the appropriate therapy, regardless of the status of an individual protease as a drug target. This is especially true in cancer diagnostics: for example, the serine protease plasminogen activator and its inhibitor PAI1 are among the most relevant diagnostic and prognostic markers for breast cancer, and are also now used in some other forms of cancer¹¹².

Another example is the serine protease kallikrein 3, better known as PSA (prostate-specific antigen), which has been the major diagnostic marker for prostate cancer for years¹¹³. More recently, cysteine cathepsins have also been suggested as relevant biomarkers for cancer¹¹⁴ and one of them, cathepsin B, as a marker for arthritis as well¹¹⁵. Proteases have also been found to be useful as diagnostic markers for parasitic infections, such as the cysteine protease cruzipain (member of the CA family of proteins), which was found to be the major immunogenic protein of *Trypanosoma cruzi* in Chagas disease¹¹⁶. Improved knowledge of protease biology will therefore not only aid drug discovery efforts against protease targets, but can also be used to provide phenotypic read outs of the efficacy of other drug classes.

Conclusions and future perspectives

Looking back at the progress made with protease-targeted therapies we cannot say that we have been tremendously successful. In the cardiovascular and anticoagulant areas, in which proteases have been validated targets for more than half a century, only recently has an orally available direct protease inhibitor made it on to the market (Ximelagatran; TABLE 1), which has since been withdrawn because of hepatotoxicity.

However, the withdrawal of ximelagatran is not a reflection of a problem with protease inhibitors *per se* and therefore it is still encouraging that the drug was developed in the first place.

Past drug development failures provide not only invaluable lessons but are also a useful resource. Many of these failed compounds, especially specific inhibitors, could still be used as powerful tools in the target validation process and in the evaluation and discovery of protease signalling pathways by both academic and industrial researchers, if industry is willing to disclose them.

As the costs and time involved with drug discovery and development show no immediate sign of decreasing, despite the advent of many new technologies, most savings will be made by parallelizing processes. For example, instead of waiting for a protease to be validated as a disease-relevant target, a vast amount of information, including structure, substrate specificity, assay development and initial screening for the interacting compounds, can be gathered beforehand and potentially extrapolated to other, related disease-relevant proteases. The Protease Platform at Novartis is already an example of such thinking.

Judging by the new compounds currently being tested in advanced clinical trials and the number of NDAs for agents that are directed at protease targets submitted so far in 2006, a boom of new therapies based on protease inhibition can be expected in the coming years. Encouragingly, these include completely new therapeutic areas such as osteoporosis and type 2 diabetes. In addition, second- and third-generation protease drugs are expected to significantly improve on existing protease-targeted therapies, certainly in the cardiovascular area (hypertension and thrombosis) but probably in other disease areas as well. This is also predicted to almost double the size of the current market — ~US\$11 billion — for protease-targeted drugs by 2009.

The major areas of interest for protease-targeted therapies are likely to remain the cardiovascular, inflammatory and infectious diseases areas, but discovery efforts will probably increase for cancer and neurodegenerative disorders. A more detailed look at the proteases currently considered as potential targets shows that endogenous proteases are often linked to chronic diseases, and are therefore attractive to pharmaceutical companies. However, there has recently been a revived interest in protease-targeted drugs for infectious diseases, although only a few (AIDS, hepatitis C) were considered seriously until recently.

Finally, we should remember the success stories. Although it was almost 30 years from when ACE was validated as a target until the first ACE inhibitors appeared on the market, it was only 10 years after the discovery of HIV protease that the first inhibitors against that target were on the market. With the emergence of new technologies, and as the structure and physiological function of more proteases are revealed and updated, it should hopefully become easier both to identify and validate proteases as relevant drug targets and to develop effective and safe drugs against them.

1. Puente, X. S., Sanchez, L. M., Overall, C. M. & Lopez-Otin, C. Human and mouse proteases: a comparative genomic approach. *Nature Rev. Genet.* **4**, 544–558 (2003).
2. Rawlings, N. D., Tolle, D. P. & Barrett, A. J. MEROPS: the peptidase database. *Nucleic Acids Res.* **32** (Database issue), D160–D164 (2004).
3. Barrett, A. J., Rawlings, N. D. & Woessner, J. F. Jr (eds) *Handbook of Proteolytic Enzymes* 2nd edn (Academic, Amsterdam, 2004).
An essential book for every protease lab. Contains a lot of information about the majority of known proteases from various organisms.
4. Davie, E. W. & Neurath, H. Identification of a peptide released during autocatalytic activation of trypsinogen. *J. Biol. Chem.* **212**, 515–529 (1955).
5. Davie, E. W. & Ratnoff, O. D. Waterfall sequence for intrinsic blood clotting. *Science* **145**, 1310–1312 (1964).
6. MacFarland, R. G. An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier. *Nature* **202**, 498–499 (1964).
References 4–6 are classical works describing protease activation by limited proteolysis.
7. Lopez-Otin, C. & Overall, C. M. Protease degradomics: a new challenge for proteomics. *Nature Rev. Mol. Cell Biol.* **3**, 509–519 (2002).
A valuable collection of different approaches for system-wide analysis of proteases.
8. Thornberry, N. A. *et al.* A novel heterodimeric cysteine protease is required for interleukin-1 β processing in monocytes. *Nature* **356**, 768–774 (1992).
9. Blobel, C. P. ADAMs: key components in EGFR signalling and development. *Nature Rev. Mol. Cell Biol.* **6**, 32–43 (2005).
10. Thomas, G. Furin at the cutting edge: from protein traffic to embryogenesis and disease. *Nature Rev. Mol. Cell Biol.* **3**, 756–766 (2002).
11. Weihofen, A. & Martoglio, B. Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides. *Trends Cell Biol.* **13**, 71–78 (2003).
12. Glickman, M. J. & Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* **82**, 373–428 (2002).
13. Davie, E. W., Fujikawa, K. & Kisiel, W. The coagulation cascade: initiation, maintenance, and regulation. *Biochemistry* **30**, 10363–10370 (1991).
14. Koblinski, J. E., Ahram, M. E. & Sloane, B. F. Unraveling the role of proteases in cancer. *Clin. Chim. Acta* **291**, 113–135 (2000).
15. Stennicke, H. S., Ryan, C. A. & Salvesen, G. S. Retrieval from execution: the molecular basis of caspase inhibition. *Trends Biochem. Sci.* **27**, 94–101 (2002).
16. Thome, M. *et al.* Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors. *Nature* **386**, 517–521 (1997).
17. Cuconati, A. & White, E. Viral homologs of BCL-2: role of apoptosis in the regulation of virus infection. *Genes Dev.* **16**, 2465–2478 (2002).
18. Gustafsson, D. *et al.* A new oral anticoagulant: the 50-year challenge. *Nature Rev. Drug Discov.* **3**, 649–659 (2004).
19. Abbenante, G. & Fairlie, D. P. Protease inhibitors in the clinic. *Med. Chem.* **1**, 71–104 (2005).
An excellent collection of the current data on protease inhibitors in clinical and preclinical research.
20. Rueff, J. & Katus, H. A. New antithrombotic drugs on the horizon. *Expert Opin. Investig. Drugs* **12**, 781–797 (2003).
21. Hirsch, J. Current anticoagulant therapy — unmet clinical needs. *Thromb. Res.* **109**, S1–S8 (2003).
22. Zaman, M. A., Oparil, S. & Calhoun, D. A. Drugs targeting the renin–angiotensin–aldosterone system. *Nature Rev. Drug Discov.* **1**, 621–636 (2002).
23. Acharya, K. R., Sturrock, E. D., Riordan, J. F. & Ehlers, M. R. ACE revisited: a new target for structure-based drug design. *Nature Rev. Drug Discov.* **2**, 891–902 (2003).
24. Natesh, R., Schwager, S. L., Sturrock, E. D. & Acharya, K. R. Crystal structure of the human angiotensin-converting enzyme–lisinopril complex. *Nature* **421**, 551–554 (2003).
25. van Esch, J. H. *et al.* Selective angiotensin-converting enzyme C-domain inhibition is sufficient to prevent angiotensin I-induced vasoconstriction. *Hypertension* **45**, 120–125 (2005).
26. Georgiadis, D. *et al.* Roles of the two active sites of somatic angiotensin-converting enzyme in the cleavage of angiotensin I and bradykinin. *Circ. Res.* **93**, 148–154 (2003).
27. Robl, J. A. *et al.* Dual metalloprotease inhibitors: mercaptoacetyl-based fused heterocyclic dipeptide mimetics as inhibitors of angiotensin-converting enzyme and neutral endopeptidase. *J. Med. Chem.* **40**, 1570–1577 (1997).
28. Kostis, J. B. *et al.* Omapatrilat and enalapril in patients with hypertension: the Omapatrilat Cardiovascular Treatment vs. Enalapril (OCTAVE) trial. *Am. J. Hypertens.* **17**, 103–111 (2004).
29. Overall, C. M. & Lopez-Otin, C. Strategies for MMP inhibition in cancer: innovations for the post-trial era. *Nature Rev. Cancer* **2**, 657–672 (2002).
30. Bergers, G., Javaherian, K., Lo, K. M., Folkman, J. & Hanahan, D. Effects of angiogenesis inhibitors on multistage carcinogenesis in mice. *Science* **284**, 808–812 (1999).
31. Coussens, L. M., Fingleton, B. & Matrisian, L. M. Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science* **295**, 2387–2392 (2002).
An excellent critical analysis of failure of MMP inhibitors in clinical trials.
32. Neeffes, J. & Dantuma, N. P. Fluorescent probes for proteolysis: tools for drug discovery. *Nature Rev. Drug Discov.* **3**, 58–69 (2004).
33. Baruch, A., Jeffery, D. A. & Bogoy, M. Enzyme activity — it's all about image. *Trends Cell Biol.* **14**, 29–35 (2004).
34. Blum, G. *et al.* Dynamic imaging of protease activity with fluorescently quenched activity based probes. *Nature Chem. Biol.* **1**, 203–209 (2005).
A demonstration of in vivo monitoring protease activity in a living mouse, an extremely valuable technique in evaluating the physiological role of a protease in health and disease.
35. Page, I. H. & Helmer, O. M. A crystalline pressor substance (angiotensin) resulting from the interaction between renin and renin activator. *J. Exp. Med.* **71**, 29 (1940).
36. Skeggs, L. T., Kahn, J. R. & Shumway, N. P. The preparation and function of the hypertensin-converting enzyme. *J. Exp. Med.* **103**, 295–299 (1956).
37. Matthews, D. J. & Wells, J. A. Substrate phage: selection of protease substrates by monovalent phage display. *Science* **260**, 1113–1117 (1993).
The first report on the use of phage display for determining substrate specificity of a protease.
38. Harris, J. L., Peterson, E. P., Hudig, D., Thornberry, N. A. & Craik, C. S. Definition and redesign of the extended substrate specificity of granzyme B. *J. Biol. Chem.* **273**, 27364–27373 (1998).
39. Harris, J. L. *et al.* Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc. Natl Acad. Sci. USA* **97**, 7754–7759 (2000).
40. Berger, A. B., Vitorino, P. M. & Bogoy, M. Activity-based protein profiling: applications to biomarker discovery, in vivo imaging and drug discovery. *Am. J. Pharmacogenomics* **4**, 371–381 (2004).
41. Thornberry, N. A. *et al.* A combinatorial approach defines specificities of members of the caspase family and granzyme B. *J. Biol. Chem.* **272**, 17907–17911 (1997).
The first paper to demonstrate the power of combinatorial chemistry in defining the substrate specificity of a protease.
42. Turk, B. E., Huang, L. L., Piro, E. T. & Cantley, L. C. Determination of protease cleavage site motifs using mixture-based oriented peptide libraries. *Nature Biotechnol.* **19**, 661–667 (2001).
43. Roberds, S. L. *et al.* BACE knockout mice are healthy despite lacking the primary β -secretase activity in brain: implications for Alzheimer's disease therapeutics. *Hum. Mol. Genet.* **10**, 1317–1324 (2001).
An excellent example of target validation using a mouse knockout model.
44. Gocheva, V. *et al.* Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. *Genes Dev.* **20**, 543–556 (2006).
A good example how knockout mice models can be combined with a disease model in order to identify the appropriate targets.
45. Fingleton, B. Matrix metalloproteinases as valid clinical targets. *Curr. Pharm. Des.* (in the press).
46. Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M. & Horvitz, H. M. The *C. elegans* cell death gene *ced-3* encodes a protein similar to mammalian interleukin-1 β -converting enzyme. *Cell* **75**, 641–652 (1993).
47. McQuibban, G. A. *et al.* Inflammation dampened by gelatinase A cleavage of monocyte chemoattractant protein-3. *Science* **289**, 1202–1206 (1999).
48. Tam, E. M., Morrison, C. J., Wu, Y. I., Stack, S. M. & Overall, C. M. Membrane protease proteomics: isotope-coded affinity tag MS identification of undescribed MT1-matrix metalloproteinase substrates. *Proc. Natl Acad. Sci. USA* **101**, 6917–6922 (2004).
49. Bredemeyer, A. J. *et al.* A proteomic approach for the discovery of protease substrates. *Proc. Natl Acad. Sci. USA* **101**, 11785–11790 (2004).
50. Van Damme, P. *et al.* Caspase-specific and nonspecific in vivo protein processing during Fas-induced apoptosis. *Nature Meth.* **2**, 771–777 (2005).
This paper demonstrates the power of proteomics in determining in vivo protease substrates.
51. Medema, J. P. *et al.* FLICE is activated by association with the CD95 death-inducing signaling complex (DISC). *EMBO J.* **16**, 2794–2804 (1997).
52. Reed, J. C. Apoptosis based therapies. *Nature Rev. Drug Discov.* **1**, 111–121 (2002).
53. Eigenbrot, C. & Kirchhofer, D. New insight into how tissue factor allosterically regulates Factor VIIa. *Trends Cardiovasc. Med.* **12**, 19–26 (2002).
54. Olson, S. T., Björk, I. & Bock, S. C. Identification of critical molecular interactions mediating heparin activation of antithrombin. *Trends Cardiovasc. Med.* **12**, 198–205 (2002).
55. Bode, W. & Huber, R. Natural protein protease inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **204**, 433–451 (1992).
56. Bode, W. & Huber, R. Structural basis of the endoprotease–protein inhibitor interaction. *Biochim. Biophys. Acta* **1477**, 241–252 (2000).
57. Turk, B., Turk, D. & Salvesen, G. S. Regulating cysteine protease activity: essential role of protease inhibitors as guardians and regulators. *Curr. Pharm. Des.* **8**, 1623–1637 (2002).
58. Giugliano, R. P. *et al.* Addition of a tissue-factor/factor VIIa inhibitor to standard treatments in NSTE-ACS managed with an early invasive strategy: results of the phase 2 ANTHEM-TIMI 32 double-blind randomized clinical trial. *Am. Heart Assoc. Sci. Sessions Dallas, Texas, USA* 13–16 Nov (2005).
59. Vlasuk, G. P. Structural and functional characterization of tick anticoagulant peptide (TAP): a potent and selective inhibitor of blood coagulation Factor Xa. *Thromb. Haemost.* **70**, 212–216 (1993).
60. Tuszynski, G. P., Gasic, T. B. & Gasic, G. T. Isolation and characterization of antistasin. *J. Biol. Chem.* **262**, 9718–9723 (1987).
61. Armstrong, W. B. *et al.* Clinical modulation of oral leukoplakia and protease activity by Bowman–Birk inhibitor concentrate in a phase IIa chemoprevention trial. *Clin. Cancer Res.* **6**, 4684–4691 (2000).
62. Chen, Y. W., Huang, S. C., Lin-Shiau, S. Y. & Lin, J. K. Bowman–Birk inhibitor abates proteasome function and suppresses the proliferation of MCF7 breast cancer cells through accumulation of MAP kinase phosphatase-1. *Carcinogenesis* **26**, 1296–1306 (2005).
63. Weimann, L. M. Fully human therapeutic monoclonal antibodies. *J. Immunother.* **29**, 1–9 (2005).
64. Reuning, U. *et al.* Urokinase-type plasminogen activator (uPA) and its receptor (uPAR): development of antagonists of uPA/uPAR interaction and their effects in vitro and in vivo. *Curr. Pharm. Des.* **9**, 1529–1543 (2003).
65. Premzl, A., Zavašnik-Bergant, V., Turk, V. & Kos, J. Intracellular and extracellular cathepsin B facilitate invasion of MCF-10A neoT cells through reconstituted extracellular matrix in vitro. *Exp. Cell Res.* **283**, 206–214 (2003).
66. Novartis Institute for Tropical Diseases. New technologies for high-throughput screening and lead discovery of anti-viral compounds. *Dengue Digest* [online], <http://www.nitd.novartis.com/downloads/dengue_digest_vol_2_no_1_april.pdf> (2005).
67. McGovern, S. L., Helfand, B. T., Feng, B. & Shoichet, B. K. A specific mechanism of nonspecific inhibition. *J. Med. Chem.* **46**, 4265–4272 (2003).
68. De Clercq, E. Antiviral drugs in current clinical use. *J. Clin. Virol.* **30**, 115–133 (2004).
69. Rozman-Pungerčar, J. *et al.* Inhibition of papain-like cysteine proteases and legumain by 'caspase-specific' inhibitors: when reaction mechanism is more important than specificity. *Cell Death Diff.* **10**, 881–888 (2003).
A good example of the selectivity problems of irreversible protease inhibitors.
70. Smyth, T. P. Substrate variants versus transition state analogues as noncovalent, reversible enzyme inhibitors. *Bioorg. Med. Chem.* **12**, 4081–4088 (2004).
71. Thurmond, R. L. *et al.* Identification of a potent and selective noncovalent cathepsin S inhibitor. *J. Pharmacol. Exp. Ther.* **308**, 268–276 (2004).

72. Altmann, E., Green, J. & Tintelnot-Blomley, M. Arylaminoethyl amides as inhibitors of the cysteine protease cathepsin K—investigating P1' substituents. *Bioorg. Med. Chem. Lett.* **13**, 1997–2001 (2003).
73. Hardy, J. A. & Wells, J. A. Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* **14**, 706–715 (2004).
74. Hardy, J. A., Lam, J., Nguyen, J. T., O'Brien, T. & Wells, J. A. Discovery of an allosteric site in the caspases. *Proc. Natl Acad. Sci. USA* **101**, 12461–12466 (2004).
- Excellent work identifying the first small-molecule allosteric protease inhibitor.**
75. Schimmer, A. D. *et al.* Small-molecule antagonists of apoptosis suppressor XIAP exhibit broad antitumor activity. *Cancer Cell* **5**, 25–35 (2004).
- The first report demonstrating that potentiating protease activity by inhibiting endogenous protease inhibitors by small molecules can be a useful therapeutic approach.**
76. Bijlens, A. P., Gils, A., Knockaert, I., Stassen, J. M. & Declercq, P. J. Importance of the hinge region between α -helix F and the main part of serpins, based upon identification of the epitope of plasminogen activator inhibitor type 1 neutralizing antibodies. *J. Biol. Chem.* **275**, 6375–6380 (2000).
77. Rupin, A., Martin, F., Vallez, M. O., Bonhomme, E. & Verbeuren, T. J. Inactivation of plasminogen activator inhibitor-1 accelerates thrombolysis of platelet-rich thrombus in rat mesenteric arterioles. *Thromb. Haemost.* **86**, 1528–1531 (2001).
78. Stanton, A. Therapeutic potential of renin inhibitors in the management of cardiovascular disorders. *Am. J. Cardiovasc. Drugs* **3**, 389–394 (2003).
79. Menard, J. *et al.* Dose-dependent effects of the renin inhibitor zankiren HCl after a single oral dose in mildly sodium-depleted normotensive subjects. *Circulation* **91**, 330–338 (1995).
80. Nussberger, J., Wuerzner, G., Jensen, C. & Brunner, H. R. Angiotensin II suppression in humans by the orally active renin inhibitor Aliskiren (SPP100): comparison with enalapril. *Hypertension* **39**, E1–E8 (2002).
81. Wood, J. M., Schnell, C. R., Cumin, F., Menard, J. & Webb, R. L. Aliskiren, a novel, orally effective renin inhibitor, lowers blood pressure in marmosets and spontaneously hypertensive rats. *J. Hypertens.* **23**, 417–426 (2005).
82. Mervaala, E. *et al.* Blood pressure-independent effects in rats with human renin and angiotensinogen genes. *Hypertension* **35**, 587–594 (2000).
83. Mentlein, R., Gallwitz, B. & Schmidt, W. E. Dipeptidyl-peptidase IV hydrolyses gastric inhibitory polypeptide glucagon-like peptide-1 (7–36) amide, peptide histidine methionine and is responsible for their degradation in human serum. *Eur. J. Biochem.* **214**, 829–835 (1993).
84. Kieffer, T. J., McIntosh, C. H. S. & Pederson, R. A. Degradation of glucose-dependent insulinotropic polypeptide and truncated glucagon-like peptide 1 *in vitro* and *in vivo* by dipeptidyl peptidase IV. *Endocrinology* **136**, 3585–3596 (1995).
85. Marguet, D. *et al.* Enhanced insulin secretion and improved glucose tolerance in mice lacking CD26. *Proc. Natl Acad. Sci. USA* **97**, 6874–6879 (2000).
86. Nagakura, T. *et al.* Improved glucose tolerance via enhanced glucose-dependent insulin secretion in dipeptidyl peptidase IV-deficient Fischer rats. *Biochem. Biophys. Res. Commun.* **284**, 501–506 (2001).
87. Deacon, C. F., Ahren, B. & Holst, J. J. Inhibitors of dipeptidyl peptidase IV: a novel approach for the prevention and treatment of type 2 diabetes? *Exp. Opin. Investig. Drugs* **13**, 1091–1102 (2004).
88. Nielsen, L. L. Incretin mimetics and DPP-IV inhibitors for the treatment of type 2 diabetes. *Drug Discov. Today* **10**, 703–710 (2005).
89. Demuth, H. U., McIntosh, C. H. & Pederson, R. A. Type 2 diabetes—therapy with dipeptidyl peptidase IV inhibitors. *Biochim. Biophys. Acta* **1751**, 33–44 (2005).
90. Augustyns, K., Van der Veken, P., Senten, K. & Haemers, A. The therapeutic potential of inhibitors of dipeptidyl peptidase IV (DPP IV) and related proline-specific dipeptidyl aminopeptidases. *Curr. Med. Chem.* **12**, 971–998 (2005).
91. Kim, D. *et al.* (2R)-4-oxo-4-[3-(trifluoromethyl)-5,6-dihydro[1,2,4]triazolo[4,3-a]pyrazin-7(8H)-yl]-1-(2,4,5-trifluorophenyl)butan-2-amine: a potent, orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J. Med. Chem.* **48**, 141–151 (2005).
92. Augeri, D. J. *et al.* Discovery and preclinical profile of Saxagliptin (BMS-477118): a highly potent, long-acting, orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J. Med. Chem.* **48**, 5025–5037 (2005).
93. Moser, B., Wolf, M., Walz, A. & Loetscher, P. Chemokines: multiple levels of leukocyte migration control. *Trends Immunol.* **25**, 75–84 (2004).
94. Yan, S., Marguet, D., Dobers, J., Reutter, W. & Fan, H. Deficiency of CD26 results in a change of cytokine and immunoglobulin secretion after stimulation by pokeweed mitogen. *Eur. J. Immunol.* **33**, 1519–1527 (2003).
95. Gelb, B. D., Shi, G.-P., Chapman, H. A. & Desnick, R. J. Pycnodysostosis is caused by a deficiency of cathepsin K. *Science* **273**, 1236–1238 (1996).
96. Saffig, P. *et al.* Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin K-deficient mice. *Proc. Natl Acad. Sci. USA* **95**, 13453–13458 (1998).
97. Gowen, M. *et al.* Cathepsin K knockout mice develop osteopetrosis due to a deficit in matrix degradation but not demineralization. *J. Bone Miner. Res.* **14**, 1654–1663 (1999).
98. Kiviranta, R. *et al.* Accelerated turnover of metaphyseal trabecular bone in mice overexpressing cathepsin K. *J. Bone Miner. Res.* **16**, 1444–1452 (2001).
99. Inaoka, T. *et al.* Molecular cloning of human cDNA for cathepsin K: novel cysteine protease predominantly expressed in bone. *Biochem. Biophys. Res. Commun.* **206**, 89–96 (1995).
100. Grabowska, U., Chambers, T. J. & Shiroo, M. Recent developments in cathepsin K inhibitor design. *Curr. Opin. Drug Discov. Devel.* **8**, 619–630 (2005).
101. Stroup, G. B. *et al.* Potent and selective inhibition of human cathepsin K leads to inhibition of bone resorption *in vivo* in a nonhuman primate. *J. Bone Miner. Res.* **16**, 1739–1746 (2001).
102. Palmer, J. T. *et al.* Design and synthesis of tri-ring P3 benzamide-containing aminonitriles as potent, selective, orally effective inhibitors of cathepsin K. *J. Med. Chem.* **48**, 7520–7534 (2005).
103. Turk, D., Gunčar, G., Podobnik, M. & Turk, B. Revised definition of substrate binding sites of papain-like cysteine proteases. *Biol. Chem.* **379**, 137–147 (1998).
104. Misbach, M. *et al.* AAE581, a potent and highly specific cathepsin K inhibitor, prevents bone resorption after oral treatment in rat and monkey. *Ann. Mtg Am. Soc. Bone Miner. Res.* Nashville, Tennessee, USA 23–27 Sept (2005).
105. Falgoutret, J. P. *et al.* Lysosomotropism of basic cathepsin K inhibitors contributes to increased cellular potencies against off-target cathepsins and reduced functional selectivity. *J. Med. Chem.* **48**, 7535–7543 (2005).
106. Deaton, D. N. & Kumar, S. Cathepsin K inhibitors: their potential as anti-osteoporosis agents. *Progr. Med. Chem.* **42**, 245–375 (2004).
- Excellent review paper focused on the development of cathepsin K inhibitors, which contains useful general information about medicinal chemistry.**
107. Tavares, F. X., Deaton, D. N., Miller, A. B., Miller, L. R. & Wright, L. L. Ketoheterocycle-based inhibitors of cathepsin K: a novel entry into the synthesis of peptidic ketoheterocycles. *Bioorg. Med. Chem. Lett.* **15**, 3891–3895 (2005).
108. Brown, S. *et al.* Potent and selective mechanism-based inhibition of gelatinases. *J. Am. Chem. Soc.* **122**, 6799–6800 (2000).
109. Krüger, A. *et al.* Antimetastatic activity of a novel mechanism-based gelatinase inhibitor. *Cancer Res.* **65**, 3523–3526 (2005).
110. Gu, Z. *et al.* A highly specific inhibitor of matrix metalloproteinase-9 rescues laminin from proteolysis and neurons from apoptosis in transient focal cerebral ischemia. *J. Neurosci.* **25**, 6401–6408 (2005).
111. Ikejiri, M. *et al.* Potent mechanism-based inhibitors for matrix metalloproteinases. *J. Biol. Chem.* **280**, 33992–34002 (2005).
112. Look, M. P. & Foekens, J. A. Clinical relevance of the urokinase plasminogen activator system in breast cancer. *Acta Pathol. Microbiol. Immunol. Scand.* **107**, 150–159 (1999).
113. Dixon, S. C., Knopf, K. B. & Figg, W. D. The control of prostate-specific antigen expression and gene regulation by pharmacological agents. *Pharmacol. Rev.* **53**, 73–91 (2001).
114. Kos, J. & Lah, T. T. Cysteine proteinases and their endogenous inhibitors: target proteins for prognosis, diagnosis and therapy in cancer. *Oncol. Rep.* **5**, 1349–1361 (1998).
115. Lai, W. T., Chang, C. H., Tang, Y., Bronson, R. & Tung, C. H. Early diagnosis of osteoarthritis using cathepsin B sensitive near-infrared fluorescent probes. *Osteoarthritis Cartilage* **12**, 239–244 (2004).
116. Cazzulo, J. J., Stoka, V. & Turk, V. The major cysteine proteinase of *Trypanosoma cruzi*: a valid target for chemotherapy of Chagas disease. *Curr. Pharm. Des.* **7**, 1143–1156 (2001).
117. Schechter, I. & Berger, A. On the size of the active site of proteinases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162 (1967).
118. Enari, M. *et al.* A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. *Nature* **391**, 43–50 (1998).
119. Rawson, R. B. *et al.* Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs. *Mol. Cell* **1**, 47–57 (1997).
120. Wilkinson, K. D. Ubiquitination and deubiquitination: targeting of proteins for degradation by the proteasome. *Semin. Cell. Dev. Biol.* **11**, 141–148 (2000).
121. Turk, D. *et al.* Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO J.* **20**, 6570–6582 (2001).
122. Silverman, G. A. *et al.* The serpins are an expanding superfamily of structurally similar but functionally diverse proteins: evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *J. Biol. Chem.* **276**, 33293–33296 (2001).
123. Stubbs, M. T. *et al.* The refined 2.4Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J.* **9**, 1939–1947 (1990).
124. Gomis-Ruth, F. X. *et al.* Mechanism of inhibition of the human matrix metalloproteinase stromelysin-1 by TIMP-1. *Nature* **389**, 77–81 (1997).
125. Shiozaki, E. N. *et al.* Mechanism of XIAP-mediated inhibition of caspase-9. *Mol. Cell* **11**, 519–527 (2003).
126. Deveraux, O., Takahashi, R., Salvesen, G. S. & Reed, J. C. X-linked IAP is a direct inhibitor of cell death proteases. *Nature* **388**, 300–304 (1997).
127. Turk, B., Turk, D. & Turk, V. Lysosomal cysteine proteases: more than scavengers. *Biochim. Biophys. Acta* **1477**, 98–111 (2000).
128. 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.* **183**, 1331–1338 (1996).
129. Nature Publishing Group. *Signalling scissors: new perspectives on proteases* [online], < <http://www.nature.com/horizon/proteases/index.html> > (2004).

Acknowledgements

Research in the Turk laboratory is supported by grants from Slovene Research Agency. I wish to thank primarily G. Salvesen for numerous discussions, similar way of thinking and permission to use information and figures from his summary of the Horizon Symposium 'Signalling scissors: new perspectives on proteases' (see Further information). Also thanks to H. P. Nestler for providing most of the data for Table 1, G. Gunčar for help with figure preparation, D. Turk and V. Turk for valuable discussions and critical reading of the manuscript, and D. Deaton, D. Percival, E. Altmann, R. Thurmond, N. Borkakoti, U. Grabowska, B. Gerhartz, V. Dive, and many others for sharing their data with me. I would also like to acknowledge the authors of the other ~300,000 papers found in Medline using protease/proteolysis/proteinase as keywords, whose work contributed to our current understanding of proteolysis but could not have been cited here because of space limitations.

Competing interests statement

The author declares no competing financial interests.

DATABASES

The following terms in this article are linked online to:

Entrez Gene:

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>
ACE | Angiotensin I | Atrial natriuretic peptide | Caspase 1 | DPPIV | Factor II | Factor VIII | Factor IX | Factor X | GIP | GLP1 | MCP3 | NEP | TIMP1 | TIMP2 | TIMP3 | TIMP4 | XIAP
OMIM: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>
Alzheimer's disease | Paget's disease | Pycnodysostosis | Rheumatoid arthritis | Type 2 diabetes

FURTHER INFORMATION

MEROPS database: <http://merops.sanger.ac.uk/>
Signalling scissors: new perspectives on proteases: <http://www.nature.com/horizon/proteases/index.html>
Access to this links box is available online.