Biochimie 92 (2010) 1689–1704

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Review

Phage display as a powerful tool to engineer protease inhibitors

Marie-Louise Zani, Thierry Moreau*

From Inserm U618 "Protéases et Vectorisation Pulmonaires", IFR 135 Imagerie Fonctionnelle, University of Tours, 37000 Tours, France

ARTICLE INFO

Article history:
Received 17 February 2010
Accepted 5 May 2010
Available online 12 May 2010

Keywords:
Protease inhibitors
Proteases
Affinity selection
Phage display
Protein engineering

ABSTRACT

Since its introduction by Georges Smith some 25 years ago, phage display has proved to be a powerful molecular technique for selecting proteins with desired biological properties from huge libraries. Early on, various protease inhibitor scaffolds were displayed at the surface of filamentous phages to select new inhibitors with shifted specificities and enhanced affinities towards one or more target protease(s). The past two decades have seen a number of natural protease inhibitors subjected to phage display, mostly to shift and increase their inhibitory specificity, but also to explore the molecular mechanisms by which they interact with their cognate enzymes with low or very high selectivity. This review focuses on the major uses of phage display in the field of protein protease inhibitors. The exquisite molecular mechanisms by which natural protease inhibitors prevent unwanted or excessive proteolysis in cells and tissues are also examined along with some of the general principles underlying the way phage display is applied to these molecules.

1. Introduction

Proteases do not simply break down proteins, they are also involved in a wide range of key physiological processes in cells and tissues. By cleaving peptide bonds [1], proteases may lead to a loss, a gain or even a switch of function depending on the nature of their substrates. Their activity, which is tightly controlled at several levels in the normal physiological state, can be dysregulated in such a way that inappropriate (excessive or insufficient) proteolysis may occur. Indeed many proteases are associated with disorders such as cancer, osteoporosis, inflammatory and neurodegenerative disorders, as well as parasitic, bacterial or viral diseases. The fact that proteases are linked to many diseases has stimulated efforts by both publicly and privately financed researchers to identify and design drugs that target proteases. Currently at least 50 proteases are considered to be potential therapeutic targets [2] and drugs specific for some of them have been approved for clinical use. The most important efforts have focused on finding small molecules that inhibit proteases, either reversibly or irreversibly. However protein inhibitors, in their natural or engineered form, may have great therapeutic value. A good example is hirudin, a potent protein inhibitor of thrombin that was isolated from the medicinal leech *Hirudo medicinalis* and is available in recombinant form (Leptirudin or Desirudin) for clinical use as anticoagulant. Apart from a few exceptions, the specificity of natural inhibitors is usually quite broad towards homologous protease targets. They therefore need to be modified using various protein engineering techniques to increase their selectivity and/or affinity. Phage display was first invented for the affinity selection of protein fragments synthesized expressed from cDNA fragments [3] but rapidly the technique evolved to generate antibodies recognizing specific antigens with a high affinity [4–6]. Successive advances in this technique, together with a better understanding of the principles underlying protein–protein interactions, have led to the emergence of many phage display projects based on a range of protein scaffolds, including protease inhibitors. This directed evolution technique is based on the fact that a protein of interest once fused to one type of phage coat protein can be displayed at the surface of phages. Interestingly, one or several regions of the protein are randomized so that each phage particle will display a unique variant, keeping in mind that the same phage contains the genetic information encoding the fusion protein. There is therefore a direct physical link between the phenotype of the displayed protein and its genotype. Therefore instead of having to genetically engineer several protein variants, expressing, purifying and then analyze each of them, phage display libraries that contain up to 10^{10} individual proteins can be designed using standard molecular biology techniques and screened. The final aim of most phage display applications is to identify polypeptides that bind to a given molecule, preferably with high affinity. Phages containing such binding proteins can then be extracted by affinity selection. The phage

Abbreviations: P1, P2, P3 etc and P1′, P2′, P3′ etc, designate inhibitor (or substrate) residues amino-terminal and carboxy-terminal of the scissile bond respectively; S1, S2, S3 etc and S1′, S2′, S3′ etc, the corresponding subsites of the cognate protease; WT, wild-type.

* Corresponding author. Tel.: +33 2 47 36 61 77; fax: +33 2 47 36 60 46.
E-mail address: moreaut@univ-tours.fr (T. Moreau).

0300-9084/$ – see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.biochi.2010.05.003

library is screened with one or several immobilized targets in a process called "panning" or "biopanning" (cf gold panning) in which only binding phages are captured. The bound phages are then eluted and the phage population is amplified by re-infection of bacterial cells before the next round of panning. Phage library screening generally entails three consecutive rounds of panning and amplification before analysis of the selected clones. The genotype–phenotype coupling allows the amino acid sequence of the specific binding proteins to be rapidly determined by sequencing the DNA of the phage vector. The last step is dedicated to the production of the protein variants of interest in an appropriate expression system to analyze their biological properties using conventional biochemical techniques. A possible limitation of phage display is that the use of a bacterial host to generate phage libraries may preclude the correct folding of the protein to be displayed. A protein that is detrimental to the host bacterium may also be poorly produced.

This review focuses on the various applications of phage display to select functional protease inhibitors that have been published in the litterature.

2. The world of natural protein protease inhibitors

A total of at least 1068 proteases and their homologs have been identified to date (January 2010) in the human genome [7]. Most of them belong to one of five families: the serine proteases, the cysteine proteases, the aspartic proteases, the metalloproteases and the threonine peptidases. Over the past 15–20 years, it has become clear that proteases are not simply non-specific digestive enzymes; they are involved in almost all biological processes in all living organisms. However, the nature of the chemical reaction they catalyse (hydrolysis of peptide bond in proteins) means that their activity must be tightly regulated to avoid excessive, hazardous proteolysis in cells or tissues. This control occurs at several levels, including regulation of gene transcription, activation of the protease pro-forms and inhibition of their proteolytic activity through the more or less direct blockage of their active sites by protein inhibitors (Fig. 1). This last

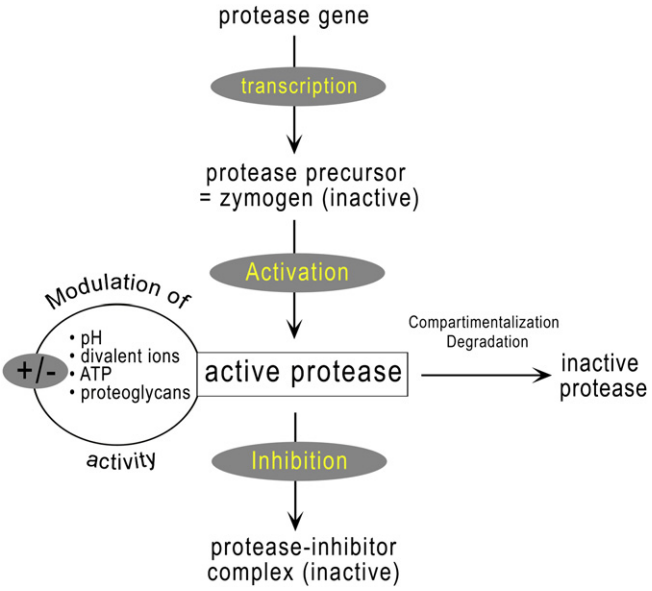

Fig. 1. Regulation of protease activity. The biological activities of proteases are controlled by multiple factors that regulate their synthesis, activate their zymogens and inhibit their activity. They may also be controlled by compartimentalization or degradation. Other factors, such as pH and specific ions are likely to modulate the activity of proteases within cells and tissues.

mode of regulation is believed to be most important in cells and tissues because protease inhibitors are present, sometimes in high concentration, in all biological fluids. There are at least 594 different inhibitor proteins (January 2010) distributed in 70 distinct families in humans [7]. The most selective and most potent naturally occurring protease inhibitors are small or large proteins or oligopeptides. Small molecules are sometimes secreted by micro-organisms to interfere with host proteases; one well-known example is the cysteine protease inhibitor E64 originally isolated from cultures of *Aspergillus japonicus* TPR-64.

An enormous number of protein-based protease inhibitors have been isolated and characterized so far. Their 3D structures, alone or complexed with their target proteases, have been elucidated over the past 30 years. These studies have revealed the many different structures/folds that finally all converge on a similar inhibitory archetype that blocks the active site with exquisite variations on this theme [8]. Most of known inhibitors are specific of a given mechanistic class of proteases although few examples of inhibitors exhibiting cross-class inhibition have been reported such as the serpins CrmA, PI9 and SCCA1 that inhibit cysteine proteases [9]. The general mechanism for protease inhibition is based on there being a good fit between the concave protease active site and the convex reactive site of the inhibitor, formed by one or more exposed loops. In some cases, either the N- or the C-terminus segment of the inhibitor also contributes to the formation of the inhibitory site with the exposed loop(s) [10]. Once formed, the tight, reversible, sometimes pseudo-irreversible, complex prevents the substrate gaining access to the protease active site.

The largest and most intensively investigated group of protease inhibitors is that comprising proteins targeting serine proteases. There are three distinct types of serine protease inhibitors. The first serine protease inhibitors that were discovered were directed against trypsin [11]. Latter on, many small proteins (30–200 residues) with inhibitory activity against trypsin- and chymotrypsin-like enzymes were characterized through successive investigation waves [12,13]. Elucidation of their 3D structures revealed that they could adopt different folds but the interacting region of the inhibitor always had virtually the same main-chain "canonical" conformation, hence the term canonical inhibitors (Fig. 2). For this reason, canonical inhibitors are thought to be outstanding examples of convergent protein evolution, where a similar biological function has been implemented in many distinct scaffolds by producing a single inhibitory loop with a canonical spatial conformation. Some researchers have suggested that this peculiar structural feature could be used to identify protease inhibitors in 3D structure databases [14,15]. Canonical inhibitors form a large group of inhibitors present in various cells and tissues of many organisms. Their concentrations can sometimes be very high, especially in plant seeds, bird's eggs and certain biological fluids. At least 18 distinct families of canonical inhibitors have been recognized, based on sequence similarities and disulfide bond topology [16]. This classification has now evolved into hierarchical, structure-based schemes such as the one used in the MEROPS database [7]. Their 3D structures have been extensively studied and have provided most of our insight into protein–protein recognition [17]. The mode of protease recognition by these inhibitors is almost always the same. In the complex, typically 10–18 residues of the inhibitory loop of the inhibitor (including those involved in both primary and secondary contacts) interact with 15–30 residues of the protease surface, mainly via Van der Waals interactions and 8–15 hydrogen bonds [8]. Canonical inhibitors all use the standard mechanism of inhibition [16]; the inhibitor residues bind to the protease active site in a substrate-like manner. The inhibitory loop may be cleaved at the scissile bond between residues P1 and P1′, but the equilibrium constant for this cleavage is often close to 1 at neutral pH. The average total area of

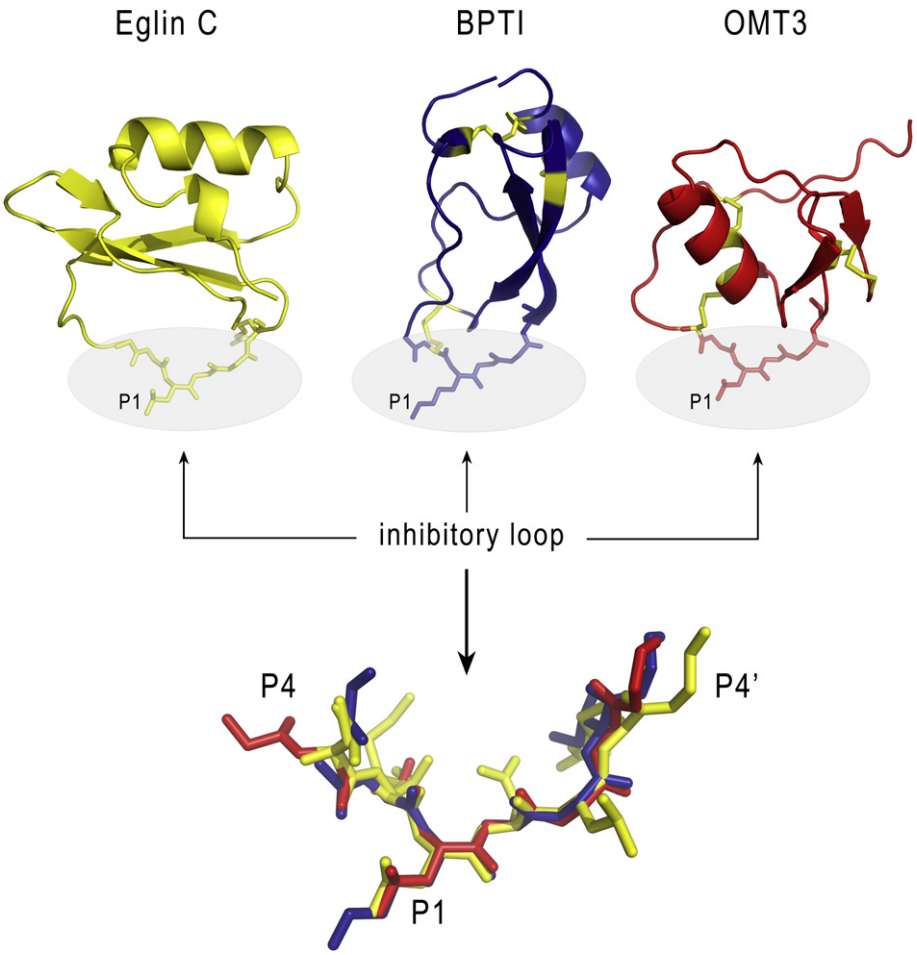

Fig. 2. Canonical inhibitors of serine proteases. Members of the most important serine protease inhibitor group, the canonical inhibitors, all have similar main-chain conformations of their exposed inhibitory loop in the P3 to P3′ region, but harbour different folds. This protruding loop fits directly into the protease active site and enables these inhibitors to act by a standard lock-and-key mechanism [16]. The 3D structures of eglin C, ovomucoid third domain (OMT3) and bovine pancreatic trypsin inhibitor (BPTI) extracted from PDB files 1PPF, 1ACB and 2PTC respectively are shown. The close-up view of their superimposed inhibitory loops illustrates the conformation found, with very few variations, in all canonical inhibitors. This figure was prepared with Pymol software (www.pymol.org).

contact at the protease–inhibitor interface is 1400 Å^2 [17]. This reflects the very high affinity of both partners, with typical K_i values $\leq 10^{-9}$M. Most canonical inhibitors are single domain inhibitory proteins but in some families (for example Kazal, BPTI, chelonianin), these single domains can be repeated from 2 to 15 times to form multi-domain, single-chain inhibitors, in which each single domain can interact independently with several identical or non identical (i.e with different specificities) protease molecules [16,18]. There is at least one exception to this behavior. Bikunin has two inhibitory sites that are so close together in the 3D structure that only one of them can bind to a target protease [19]. Ecotin is a canonical inhibitor that inhibits many pancreatic serine proteases, including trypsin, chymotrypsin and elastase. It was first identified in *Escherichia coli* and latter on in other gram-negative bacteria [20] and has a sophisticated architecture in which two single domains assemble each other to form an inhibitory active homodimer that can bind two protease molecules [21].

The second type of serine protease inhibitors are members of a superfamily of quite large homologous proteins (≈400 amino acids), the serpins (serine protease inhibitors) [9,13]. They are abundant in human plasma. Like canonical inhibitors, they interact with their target enzymes in a substrate-like manner via an exposed loop which is cleaved as soon as the serpin interacts with the protease active site. This cleavage leads to the formation of a covalent ester between the active serine (Ser195) of the proteinase and the backbone carbonyl of the cleaved loop. The loop region then becomes inserted into a β-sheet of the serpin molecule. Once the loop is completely inserted, the covalently attached protease is translocated by over 70 Å from its initial position [22]. The distortion of the protease active site resulting from this dramatic conformational change inactivates the protease [8,9,22].

The third type of serine protease inhibitors includes the more specialized and therefore less abundant non-canonical inhibitors produced by blood-sucking organisms such as leeches and insects. These inhibitors are selective for thrombin or factor Xa, two key enzymes in the coagulation process. They form very tight complexes with their target enzymes as a result of numerous interactions involving both the N-terminal part of the inhibitor and secondary interactions outside the active site [10]. A classic example of this type of inhibitor is hirudin, a small protein (65 residues) which inhibits thrombin very strongly (K_i = 20 fM) [23]. The contact surface area of the complex formed with its target is 1800 Å^2, which is much larger than that formed by canonical inhibitors and their cognate enzymes [8].

A hallmark of canonical inhibitors is that the main structural determinant for protease binding is the P1 residue that fits directly into the S1 primary specificity site of the target protease [8,24,25]. For example, an inhibitor with an arginine or lysine at P1 will

inhibit trypsin-like enzymes. As a consequence, replacing the P1 residue of an inhibitor may dramatically alter its original specificity. For example, bovine pancreatic trypsin inhibitor (BPTI), an extensively studied member of the Kunitz family of canonical inhibitors, may be transformed into a potent inhibitor of neutrophil elastase (HNE) by replacing its P1 lysine by a valine residue. This single substitution shifts the inhibitory constant K_i for HNE from 3.6 μM for the WT BPTI (poor inhibitor) to 3 nM (potent inhibitor) [26]; affinity is thus increased 1200-fold. Its affinity can be further increased 12-fold by modifying the nature of the P3 residue in the BPTI inhibitory loop. Many other such examples have been reported, in which mutations of one or several residues within the inhibitory loop of a canonical inhibitor have greatly altered its specificity and binding strength towards different protease targets. The modified inhibitors include eglin C [27–29], ovomucoid turkey domain 3 [16,30], winged bean chymotrypsin inhibitor (WCI) [31], Streptomyces subtilisin inhibitor [32–34], chymotrypsin inhibitor 2 [35], pancreatic secretory trypsin inhibitor (PSTI) [36], sunflower trypsin inhibitor-1 (SFTI-1) [37,38], APPI [39], *Schistocerca gregaria* chymotrypsin inhibitor (SGCI) [40], ecotin [41–43], secretory leukocyte proteinase inhibitor (SLPI) [44] and trappin-2 [45]. Members of the serpin family of inhibitors, which like canonical inhibitors interact with their cognate enzymes in a substrate-like manner before forming an irreversible complex [9], have also been modified by site-directed mutagenesis. Some of the best known are: α1-PI [46–48], α1-antichymotrypsin [49,50], antithrombin [51–53], plasminogen activator inhibitor-1 (PAI-1) [54], and protein C inhibitor (PCI) [55,56].

Cysteine protease inhibitors (CPIs) were identified and characterized much more recently than serine protease inhibitors. The first CPIs that were discovered belong to the large cystatin superfamily, which is divided into three families: the stefins, the cystatins and the kininogens. They all form tight complexes with cysteine proteases through the interaction of a hydrophobic wedge-shaped edge formed by three distinct regions of the inhibitor, the N-terminal end and two hairpin loops, that fits perfectly into the V-shaped protease active site. This mode of interaction was first suggested by the crystal structure of chicken cystatin [57], and confirmed latter with the X-ray structure of the stefin B-papain complex [58]. Unlike canonical inhibitors of serine proteases, cystatins do not bind in a true substrate-like manner, but rather bind in the immediate vicinity of the protease active site (i.e does not interact directly with the catalytic residues) to obstruct access by protein substrates (Fig. 3). The two hairpin loops, one containing the conserved consensus sequence QVVAG and the other a conserved PW sequence, interact with the S1′–S4′ protease subsites, while the N-terminal end, which is rather flexible in the uncomplexed cystatin, runs through the S3 to S1 subsites in such a way that cleavage at the P1–P1′ bond is impossible [8,58]. Other cysteine protease inhibitors may have folds that are unrelated to those of cystatins. Thyropins have a wedge-shaped edge like cystatins, but interact with cysteine proteases more selectively than do cystatins because additional contacts are formed between the inhibitor and the protease surface [59]. The chagasin family members, such as chagasin isolated from the protozoan *Trypanosoma cruzi*, share a comparable mode of inhibition where an inhibitory wedge composed of three loops (L2, L4, L6) forms a high number of contacts with the protease and are responsible for this inhibitor being bound tightly to its host protease, cathepsin L or cruzipain, the major cysteine protease secreted by *T. cruzi* [60,61]. The chagasin immunoglobulin-like fold (Fig. 3) is different from those of cystatins and thyropins, suggesting that all these cysteine protease inhibitors arose by a convergent evolution mechanism taking profit of the interaction of a tripartite loop with their target proteases. Many other new families of CPIs have been identified over the past decade. Only a single family of inhibitors (the cystatins) was known in the early 1990s, while up to ten different small or large families of CPIs are now recognized (see [62] for a review) based on the distinct protein fold and mode of action of their members.

The proteolytic activity of matrix metalloproteases (MMPs) is regulated by their endogenous inhibitors, the tissue inhibitors of metalloproteases (TIMPs). The TIMP family has four members (TIMP-1–4). They interact with MMPs in a substrate-like manner through the first four N-terminal residues, but they are not cleaved because the TIMP causes the catalytic water molecule of the

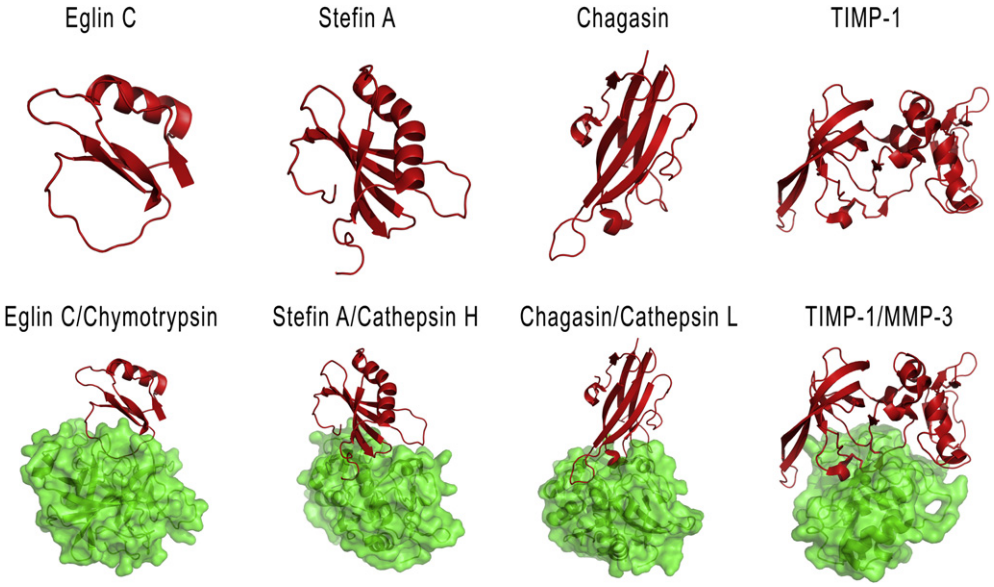

Fig. 3. Representative inhibitors of serine, cysteine and metalloproteases. Ribbon drawings of inhibitors and their inhibitory complexes. From left to right: eglin C with chymotrypsin (PDB code 1ACB); stefin A with cathepsin H (PDB code 1NB3); chagasin with cathepsin L (PDB code 2NQD); TIMP-1 with MMP-3 (PDB code 1UEA). For each inhibitor-protease pair, the structure of the inhibitor alone on the top of the figure is that extracted from the corresponding complex shown at the lower part of the figure. Inhibitors are coloured red and proteases with their molecular surfaces are shown in green. The figure was prepared with Pymol (www.pymol.org).

enzyme to be displaced. Together with the N-terminal residues 1–4, five loop segments of the inhibitor form an elongated, continuous edge (Fig. 3) whose binding in the protease active site prevents substrate access [8]. TIMPs bind tightly to MMPS but they are not very selective in that they interact with most MMPs (see [63] for a review). There are three other unrelated groups of metalloprotease inhibitors, one includes the potato carboxypeptidase A inhibitor (CPI) and the leech carboxypeptidase inhibitor (LCI), both of which interact with the protease active site via the C-terminal end of the inhibitor to form an enzyme–product type of complex in which the C-terminal residue is cleaved off in the complex [64]. The serralysin group of bacterial metalloproteases inhibitors such as the *Erwinia chrysanthemi* inhibitor use their N-termini, like TIMPs, to coordinate the catalytic Zn^{2+} ion in the protease active site [65]. The third group has at least one member, Streptomyces metalloprotease inhibitor (SMPI) that was isolated from *Streptomyces nigrescens.* Rather surprisingly, it is thought to inhibit the metalloproteases of the gluzincin family by a standard mechanism similar to that of canonical inhibitors of serine proteases [66].

Naturally occurring protein inhibitors of aspartic proteases seem to be very rare; only two have been identified to date [7]. One, a small protein from *Saccharomyces cerevisiae*, has a unique mode of inhibition since it is completely unfolded in the uncomplexed state and adopts a helical conformation once it is bound to the yeast protease A active site [67]. The structure of this complex is different from that found in zymogens of aspartic proteases, suggesting that this latter mode of proteolysis regulation is likely to be dominant in this class of enzymes.

Although the structural basis for tight binding by many protease inhibitors is now well understood, site-directed mutagenesis of one or a few of the amino acids involved in the interacting region(s) of the inhibitor to generate novel inhibitors may prove time-consuming and labor-intensive because only a limited number of variants can be studied simultaneously. Moreover, even when the 3D structures of the inhibitors or the protease–inhibitor complexes are known, deciding which amino acids are appropriate targets for mutagenesis and the type of substitution to be performed is not straightforward.

The power of phage display makes it an ideal tool with which to engineer natural protease inhibitors to generate new molecules with enhanced selectivity/affinity and to provide new insights into the molecular mechanisms by which these highly evolved molecules function. Other applications of phage display technology, which will not be reviewed here, include the discovery of inhibitory-constrained or unconstrained- peptide sequences that can serve as leads for further pharmaceutical development of peptide protease inhibitors.

3. Structure and biology of filamentous phages

The filamentous bacteriophages (Genus *Inovirus*) are viruses that infect many gram-negative bacteria by using their pili as receptors. They contain a circular single-stranded DNA genome packaged in a long, somewhat flexible, cylinder-shaped capsid [68]. Most of the information about filamentous phages (Ff phages) has been obtained for the f1, fd and M13 phages that infect *E. coli*. They have been extensively studied and widely used for phage display applications. Unlike most bacterial phages, Ff phages are non-lytic and do not assemble in the cytoplasm of host cells. Instead, they are continually secreted across the bacterial membranes as they are assembled. Each Ff phage particle forms a rod with a diameter of about 6.5 nm and a length of about 1 μm for a total molecular weight of about 16.3 Mda [69,70]. Their capsids are made up of five structural proteins that are synthesized as integral membrane protein of the inner membrane of *E. coli* cells [71]. The major capsid protein pVIII (50 residues) forms a helical array of 2700 identical units that overlap like fish scales around the single-stranded circular DNA molecule of about 6400 nucleotides. On one end, the particle is capped by about five copies of proteins pVII (33 residues) and pIX (32 residues). This end emerges first from the bacterial cell during phage assembly. The assembly process terminates with the incorporation of 3–5 copies each of the 406-residue pIII and 112-residue pVI proteins at the other end (Fig. 4) [68,72,73].

All five coat proteins have been used to display proteins or peptides (see [74] for a review) but the most commonly used phage proteins for phage display are pVIII and pIII, which have both their N-terminal region exposed to the exterior of the phage particle. The sequences to be displayed at the phage surface are typically inserted at the N-terminus of pVIII between the signal sequence and the beginning of the mature protein coding sequence. However, only short sequences of 6–8 residues can be displayed on each pVIII molecule because longer sequences prevent the correct assembly of the phage [74]. Larger polypeptides can be displayed on pVIII, but this requires production of the fusion protein from a phagemid vector (*vide infra*) to generate hybrid virions bearing mainly wild-type pVIII and a few copies of the engineered pVIII.

The pIII coat protein with its 3–5 copies located at the end of the phage is generally the protein of choice for most phage display applications because it allows long sequences to be inserted [74] and it is compatible with monovalent display (see below). The 406 amino acid pIII protein consists of three domains (N1, N2 and CT) that are connected by flexible glycine-rich regions. Deletion analysis studies have shown that the N-terminal domain N1 is involved in membrane penetration, while domain N2 enables the phage particle to be adsorbed onto the F-pilus of the *E. coli* cell [75]. The CT domain, which anchors pIII in the phage coat, is essential for the structural integrity of the phage and may be involved in forming the entry pore for DNA translocation. Like pVIII display, the infectivity of the resulting virion made of modified pIII proteins can be dramatically reduced, depending on the protein displayed. Again, this problem may be overcome by using phagemid vectors to produce hybrid virions bearing both WT and modified pIIIs, with the WT pIIIs ensuring phage infectivity.

4. Phage or phagemid vector

Proteins of interest can be displayed at the phage surface using either a phage vector where the heterologous sequence is inserted directly into the coding sequence of pIII or another coat protein, or using a plasmid-based so-called phagemid vector that contains only the fusion protein gene (see [74] for a review). Phagemid vectors are now widely used because they are small and cloning proteins of interest is easier than cloning in phage vectors. In phagemid vectors, the cDNA of the protein to be displayed is cloned into a small plasmid under the control of a weak promoter. Typically a phagemid vector carries both a plasmid origin of replication and a Ff origin of replication, the latter allowing the synthesis of a single-stranded DNA and the subsequent formation of virions. Virions are produced from *E. coli* cells harbouring such a phagemid vector by infecting the cells with a helper phage, a Ff phage with a defective Ff origin that prevents its efficient replication. Infected cells then express both wild-type coat proteins encoded by the helper phage genome and a small amount of the fusion protein encoded by the phagemid vector. Finally, most virions contain the phagemid genome, because the helper phage genome has altered packaging efficiency. Thus the link between the phenotype and genotype of the displayed protein is preserved.

Depending on the vector (phage or phagemid) and on the coat protein retained for display, the valency of display at the phage surface may be either polyvalent or monovalent. For example, phage vectors for pVIII display will display polyvalently, but if a phagemid vector is used to display the protein of interest on pIII, display will be

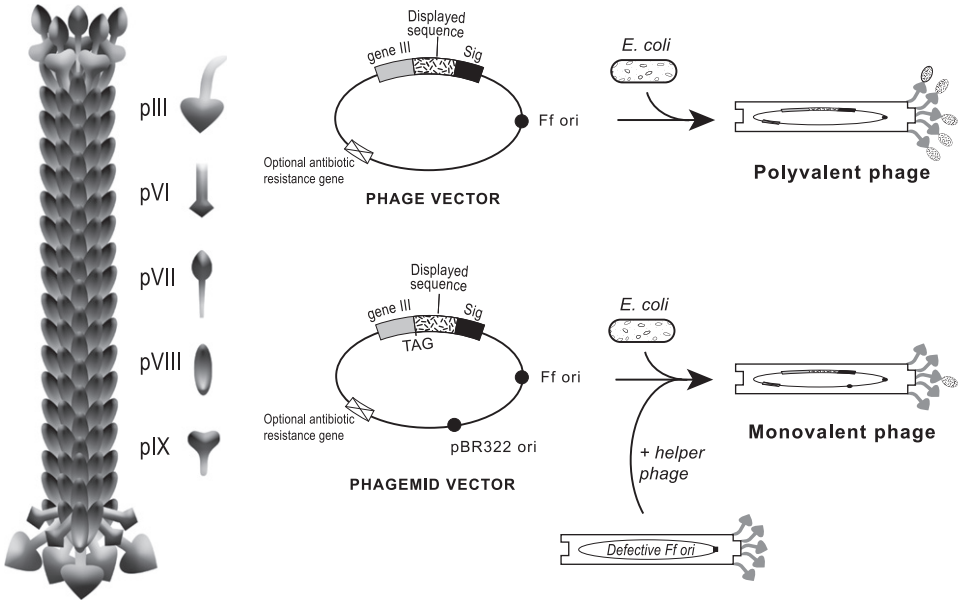

Fig. 4. The structures of filamentous phages and vectors used to display proteins at the phage surface. Schematic structure of a M13 filamentous phage widely used for phage display studies. The phage particle is about 1 μm long and 6.5 nm across. The capsid, containing the single-stranded DNA genome, is made up of about 2700 copies of the major coat protein pVIII, which is often used to display proteins polyvalently at the phage surface. One end of the particle is capped with about five copies each of the coat proteins pVII and pIX, while two other minor proteins pIII and pVI are incorporated (3–5 copies each) at the other end of the phage. Two strategies may be used to display a protein at the surface of phages. The gene encoding the protein may be cloned in frame with the phage gene encoding the phage proteins pIII or pVIII directly in the DNA phage vector. This results in more than one copy of the modified coat protein being incorporated into the phage particle. Hence, the protein display will be polyvalent. The second, more popular, strategy uses a phagemid vector, which is a plasmid-like vector containing both a phage and an *E. coli* origin of replication and the sequence encoding pIII coat protein plus cloning sites and an antibiotic-resistance gene for selecting bacteria. However, phagemid cannot be used to produce complete phages because the phagemid vector does not have all the structural and non-structural genes. A phage helper, such as M13K07, which has a slightly defective origin of replication, provides all the other structural proteins needed to generate complete phage particles. This system is mainly used for monovalent phage display.

monovalent, i.e one fusion protein per phage. The amount of fusion protein expression will follow a Poisson distribution if a pIII phagemid display vector and a helper phage are used: <10% of the virions will display 1 copy of the fusion protein, very few will display two copies and the remaining majority of phages will display only wild-type pIII. This means that the major displaying particles are monovalent while most virions do not display any protein at all [74]. This is important to understand when designing any phage display project because it will affect the selection steps. Polyvalency leads to mutivalency binding events and confers a high apparent affinity (avidity) on weak-binding virions, while monovalent display allows selection based on intrinsic affinity. It is therefore generally preferred in applications where the major goal is to select the tightest binding variants from a displayed library. This has allowed for example the selection of hGH variants with as little as three-fold higher receptor binding affinity than WT hGH [76].

It may be advantageous to start with a polyvalent display strategy when the initial binding affinity of the WT protein is low ($K_i > 1$ μM) because it allows a far greater representation of rare library variants. Every virion displays several copies of the protein, which is much better than the <10% of phages displaying the protein of interest in a monovalent display. Monovalent libraries may be constructed later for further affinity optimization.

5. Library design

The gene of interest is either cloned or synthesized from oligonucleotides whether the chosen vector is a phage or a phagemid. It is then inserted between a signal peptide and the gene encoding protein pIII or pVIII, depending on whether the display is monovalent or polyvalent. The degree of diversity in the region of interest of the protein to be displayed may vary. The simplest and most complete randomization scheme, “hard randomization” [77], involves the introduction of degenerate codons that encode all twenty possible natural amino acids at amino acid positions where variation is desired. The NNS (N = A, C, G, T, S=C, G) or NNK (K = T, G) codon is used for this, but there is always some redundancy because some amino acids are encoded by two or three of the 32 unique codons while others have only one codon. However, the number of required transformants increases so rapidly (Table 1) that the theoretical diversity will exceed the actual diversity of the library, depending on the number of positions randomized. There are two ways to circumvent this problem. One is to use a “tailored randomization” [77] in which only a subset of natural amino acids is introduced at any given position. For example, the TDK degenerate codon (T, D = A, G, T, K = G, T) represents six unique codons encoding five hydrophobic residues (C, F, L, W, Y). The other, even more restrictive way is to use “soft randomization” or “doping”. In this method, a small, defined variation is allowed within each

Table 1
Theoretical number of required clones as a function of the number of amino acid positions to be randomized.

Number of randomized positions	Amino acid diversity (number of proteins)	DNA diversity (number of clones)	Number of transformants required for a confidence level of:	
			90%	99%
1	20	32	74	149
2	400	1000	2.4×10^3	4.8×10^3
3	8000	33000	7.6×10^4	1.5×10^5
4	160000	1.1×10^6	2.4×10^6	4.9×10^6
5	3.2×10^6	3.4×10^7	7.7×10^7	1.6×10^8
6	6.4×10^7	1.1×10^9	2.5×10^9	5.0×10^9
7	1.28×10^9	3.4×10^{10}	7.9×10^{10}	1.6×10^{11}
8	2.56×10^{10}	1.1×10^{12}	2.5×10^{12}	5.1×10^{12}

selected codon, so that each is strongly biased towards one particular amino acid. The library can be constructed with either single-stranded oligonucleotide-directed mutagenesis or double-stranded cassette mutagenesis. Oligonucleotide-directed mutagenesis may be preferable because it does not require unique restriction sites at the boundaries of the gene segment to be replaced. Molecular diversity can also be generated using error-prone PCR [78–82] and DNA shuffling [83–86]. In error-prone PCR, amplification is performed under conditions that decrease the fidelity of the DNA polymerase, leading to random amino acid substitutions. However, some nucleotide substitutions are preferred, and only single-nucleotide substitutions occur, so that not all the possible twenty amino acids are represented in such libraries. DNA shuffling starts with a set of highly homologous genes to ensure DNA recombination between all the genes after they have been broken up by an endonuclease. This method has not been used as such in the field of protease inhibitors. Instead DNA shuffling-like techniques have been used to generate a second generation library of ecotin variants (see below) starting from clones selected from 3 initial libraries.

It is important to evaluate the size and diversity of the library so that all possible clones will be represented and sampled during the selection step. The size of the library is usually designed to be at least 5–10 larger than the theoretical size so that rare clones can be recovered. The upper limit for a library is generally considered to be approximately 10^9 clones using standard cloning techniques. This theoretically ensures that all possible combinations of six amino acids positions are displayed (Table 1). Thus, a library of 10^9 clones will be incomplete if more than six positions are varied with all 20 amino acids. A highly efficient, electrocompetent bacterium like *E. coli* SS320 should be used because of the large number of required bacterial transformants. This commercially available strain has been optimized for both efficient phage infection and its electroporation properties. Completeness of the library is generally assessed by direct DNA sequencing of clones. It may be calculated using the equation: $N = \ln(1-p)/\ln[1-(1/n)]$, where N is the total number of individual clones in the library, n is the number of possible combinations and p the probability that any clone can be found in the library given a library size of N [87]. An alternative strategy, iterative optimization, has been proposed to avoid dealing with huge libraries [88]. A first region of the displayed protein is variegated so as to design a library of manageable size ($\leq 10^7$) that will allow the selection of a range of good binders while eliminating poor binders. Then a second library is designed in which clones selected from the first library are randomized in a second region that is likely to interact with the desired target. This approach was used to select tight-binding inhibitors of target proteases based on the scaffold of the first Kunitz domain of lipoprotein-associated coagulation inhibitor [88,89]. The obvious advantage of an iterative protocol is that it allows the selection of a large number of displayed proteins, with a medium-sized library (about 10^7 or less) for each round of randomization and screening.

It is also desirable to get rid of the wild-type form of the inhibitor from the pool of displayed proteins when constructing a library designed to select high-affinity protease inhibitors because the WT inhibitor can interfere with selection, especially if its starting affinity is already high (low K_i). The final library will contain significant amounts of wild-type template because mutagenic oligonucleotides are not incorporated into the template gene with 100% efficiency. Since the WT inhibitor is the most prevalent sequence, it will dominate the selection steps. Display of the wild-type inhibitor can be easily prevented by using an inhibitor template that includes one or up to three stop codon(s) within the chosen randomized region so that the translation of such a sequence will provide only truncated proteins (i.e undisplayed inhibitors). TAA is generally used since it is the most efficient termination codon in *E. coli,* according to the codon use table for this bacterium. Successful incorporation of the mutagenic oligonucleotides will lead to the replacement of this stop codon while simultaneously introducing appropriate degenerate sequences. This simple procedure ensures that only mutated variants of the chosen inhibitor will be displayed.

A trick to consider when setting up a phage display project is to include in frame an amber codon (TAG) between the protein to be displayed and the phage coat protein so that it can mediate a switch between phage surface expression provided a suppressor *E. coli* strain is used (e.g. XL1-Blue) to propagate the phage or soluble expression in a bacterial expression system if a nonsuppressor strain, which recognizes the amber codon as a stop codon, is used.

6. The region of the protease inhibitors to be randomized

Most if not all phage display applications using protease inhibitors are based on affinity selection, as are most successful applications of this technique. Consequently the priority for randomization will be the region(s) of the inhibitor that is/are involved in the close interaction with the protease active site. Deciding which amino acids to mutate can only be done in the light of the knowledge of the 3D structure of the inhibitor alone or better, complexed with its target protease. The number of key residues contributing to the inhibitory specificity of the two main classes of serine protease inhibitors (canonical inhibitors and serpins) is very limited. They all lie along a linear sequence that forms the inhibitory loop. This loop is therefore the primary target for randomization of a limited number of positions, with particular emphasis on the P1 position, as it makes the major contribution to the total binding energy of the inhibitor-protease complex (Table 2). The inhibitory loop of canonical inhibitors is usually stabilized by a P2 or P3 cysteine residue which forms a disulfide bond with another Cys of the protein scaffold [90]. This Cys must be strictly conserved in the randomization scheme of residues belonging to the inhibitory loop so that the canonical conformation cannot be modified.

Surprisingly, no affinity selection studies using phage display techniques have been done on serpin inhibitors. Although their selectivity may be enhanced through appropriate substitutions within their reactive site loop, its fusion with the phage coat protein may hamper the rearrangements of the serpin structure that occur when it is complexed with its target protease. Pannekoek's group at the University of Amsterdam used human plasminogen activator inhibitor-1 (PAI-1) as a model to localize the PAI-1 interaction sites with a limited set of monoclonal antibodies [82], or with a variable region of its cognate protease tPA (see below) [91]. They designed a library based on DNA shuffling whereby the PAI-1 cDNA was fragmented with an endonuclease and then reassembled the full-length gene under PCR conditions favoring error-prone mutations. They found that 114 different positions distributed over the entire sequence length were mutated with either a moderate of relatively high frequency, but that there were no mutations in the reactive inhibitory loop, although the selection procedure included panning against tPA. It has been suggested that such a global randomization scheme may compensate in some cases the lack of an available 3D structure.

The amino acid positions within one or more of the regions that interact with the protease can be randomized for inhibitors like ecotin and cystatins (Table 2), provided the library has the required diversity.

7. Selection

Once a protease inhibitor of interest has been fused to either pIII or another coat protein, the fusion protein must be checked to

Table 2
Protease inhibitors used in phage display studies.

Inhibitor displayed	Target for selection	Type of fusion	Randomization	Desired property	Reference
Serine protease inhibitors: canonical inhibitors					
Bovine pancreatic trypsin inhibitor (BPTI)	Bovine chymotrypsin and porcine pancreatic elastase	Monovalent with pIII	Any of 20 AA at P1, P2′ and P3′ A, G, V, D at P1′	Selective tight-binding inhibitors of targets	[102]
Bovine pancreatic trypsin inhibitor (BPTI)	Bovine trypsin and chymotrypsin, human leukocyte elastase, porcine pancreatic elastase, human azurocidin	Monovalent with pIII	Any of 20 AA at P1	Selective tight-binding inhibitors of targets	[136]
Bovine pancreatic trypsin inhibitor (BPTI)	Human neutrophil elastase	Polyvalent with pIII	Limited set of AA (2–6) at P1, P1′, P2′, P3′ and P4′	Selective tight-binding inhibitors of targets	[100]
Ecotin	Bovine trypsin, urokinase-like plasminogen activator (uPA)	Monovalent with pIII	Any of 20 AA at P1–P1′ (84–85)	Selective tight-binding inhibitors of uPA	[110]
Ecotin	Bovine trypsin, rat trypsin and urokinase-like plasminogen activator (uPA)	Monovalent with pIII	Focused library: limited set of AA at positions 108,110,112 and 113 within the 100s secondary binding loop Complete library: any of 20 AA at positions 108,110,112 and 113	Selective tight-binding inhibitors of uPA	[111]
Ecotin	Bovine trypsin, rat trypsin and urokinase-like plasminogen activator (uPA)	Monovalent with pIII	Any of 20 AA at four positions (67–70) in the 60s secondary binding loop and/or P1 Met replaced by Arg	Bidentate tight-binding inhibitors of trypsin and uPA	[87]
Ecotin	Plasma kallikrein, Fxa, FXIa, FXIIa, uPA and MT-SP1	Monovalent with pIII	Limited set of 4 AA within four surface contact loops at positions 51–54 (50s loop), 67–71 (60s loop), 81–86 (80s lop) and 108–112 (100s loop)	Selective tight-binding inhibitors of plasma kallikrein, MT-SP1 and Factor XIIa	[85]
Schistocerca gregaria proteinase inhibitor-1 and -2 (SGPI-1 and SGPI-2)	Bovine trypsin and chymotrypsin, crayfish typsin	Monovalent with pVIII	Complete chimeric SGPI-1/ SGPI-2 library involving shuffling of 18 critical residues	Elucidation of the structural determinants for taxon-specificity of group I pacifastins	[116]
Inhibitor domain of Alzheimer's amyloid β-protein precursor (APPI)	Bovine trypsin and chymotrypsin	Monovalent with pIII	Any of 20 AA at P5, P3, P1, P2′ and P4′	Exploring binding specificity of APPI	[137]
Inhibitor domain of Alzheimer's amyloid β-protein precursor (APPI)	Human plasma kallikrein	Monovalent with pIII	Library I: Any of 20 AA at P5, P3, P1, P2′ and P4′ Library II: Any of 20 AA at P5, P4, P3, 39, R or K at P1 Library III: Any of 20 AA at P1′, P2′, P3′, 34, R or K at P1	Selective tight-binding inhibitors of target	[104]
Inhibitor domain of Alzheimer's amyloid β-protein precursor (APPI)	Human Factor VIIa	Monovalent with pIII	Library I: Any of 20 AA at P5, P3, P1, P2′ and P4′ Library II: Any of 20 AA at P5, P4, P3, 39, R or K at P1 Library III: Any of 20 AA at P1′, P2′, P3′, 34, R or K at P1	Selective tight-binding inhibitors of target	[99, 103]
Pancreatic Trypsin Secretory Inhibitor (PSTI)	NS3 protease of hepatitis C virus	Monovalent with pIII	Any of 20 AA at 6, 7 or 8 positions within the inhibitory loop (17–23)	Selective tight-binding inhibitors of target	[94, 138]
First Kunitz domain of Lipoprotein-Associated Coagulation Inhibitor (LACI-D1)	Human plasmin, plasma kallikrein, thrombin	Polyvalent with pIII	Limited set of AA within inhibitory loop (10–21) and second binding loop (31–39)	Selective tight-binding inhibitors of targets	[88, 89]
Leech-Derived Tryptase Inhibitor (LDTI)	Human thrombin, neutrophil elastase, plasmin	Monovalent with pIII	Any of 20 AA at P1, P3′ and P4′ Limited set of AA at P1′ and P2′	Selective tight-binding inhibitors of targets	[105, 106]
Mustard Trypsin Inhibitor (MTI-2)	Bovine trypsin and chymotrypsin	Monovalent with pIII	Any of 20 AA at P3, P2, P1, P1′ and P2′	Protease inhibitor toxic against aphids	[125]
Proteinase Inhibitor II (PI2)	Bovine trypsin and chymotrypsin	Monovalent with pIII	First set: Replacement of P2, P1 and P1′ of each inhibitory loop by Ala Second set: substitution of P1 Leu by Arg	Analysis of the relative contribution of the two inhibitory domains to proteinase binding	[95]
Serine protease inhibitors: serpins					
Plasminogen Activator Inhibitor-1 (PAI-1)	Tissue plasminogen activator (tPA)	Monovalent with pIII	Hypermutated library of PAI-1 obtained by DNA shuffling	Identification of stabilizing mutations for the transition from the unstable, active form to stable, inactive (latent) form of PAI-1	[82]
Plasminogen Activator Inhibitor-1 (PAI-1)	Thrombin, Tissue plasminogen activator (tPA)	Monovalent with pIII	Hypermutated library of PAI-1 obtained by DNA shuffling	Identification of residues involved in interaction with thrombin and tPA	[91]

Table 2 (*continued*)

Inhibitor displayed	Target for selection	Type of fusion	Randomization	Desired property	Reference
Serine protease inhibitors: non-canonical inhibitors					
Hirudin	Thrombin	Monovalent with pIII	Library I: Any of 20 AA at the first five N-terminal positions (1–5) Library II: Any of 20 AA in the protease sensitive region 56–64 Library III: Any of 20 AA at position 56 Library IV: Any of 20 AA in the protease sensitive region 49–50	Selective tight-binding inhibitors of thrombin resistant to proteolytic degradation	[118]
Cysteine protease inhibitors: cystatins					
Soyacystatin	Papain	Monovalent with pIII	Any of 20 AA in hairpin loop I (48–52) and in hairpin loop II (78–79)	Identification of optimal residues in phytocystatin for interaction with papain	[92]
Chicken cystatin	Papain	Monovalent with pIII	Functional display of WT cystatin	–	[139]
Onchocystatin from the nematode *Onchocerca volvulus*	Cysteine proteinases from larvae of the common bean weevil, *Acanthoscelides obtectus* and papain	Monovalent with pIII	Any of 20 AA at hairpin loop I (65–69)	Selective tight-binding inhibitors for insect (*A. obtectus*) cysteine proteases	[124]
Cystatin A	Papain	Monovalent with pIII	Any of 20 AA at the first 4 N-terminal positions	Analysis of the relative contributions of the first four AA of cystatin A for protease selectivity	[117]

ensure that it is indeed directed to the phage surface and is both stable and functional. The fusion protein may be detected by Western blotting using antibodies against the inhibitor [92,93], against the coat protein to which the inhibitor is fused [94], or against a tag sequence inserted between the inhibitor and the coat protein [95]. The inhibitor must be fused to the chosen coat protein in such a way that there is no interference with inhibitor-target protease complex formation. This is especially important for inhibitors in which all or part of their inhibitory site lies near their N- or C-terminal end. The type of fusion may also influence the expression level of the whole fusion protein. This was shown to occur when eglin C was fused by its C-terminal to pIII: the whole protein was poorly displayed [96]. An alternative is to fuse the opposite end of the inhibitor to another coat protein, as was done by fusing pVI to the N-terminus of eglin C [96]. Another frequent solution is to introduce a short flexible spacer between the inhibitor and the phage coat protein to preserve full protease access to the inhibitory site. This strategy was used to fuse the C-terminus of ecotin to the C-terminus of pIII via a Gly–Gly–Gly linker so that the formation of an active inhibitory ecotin dimer was not blocked [93].

Selection is generally performed by incubating the phage library *in vitro* with the desired target protease which has been immobilized on ELISA plates, for example, – a process referred to as "biopanning" or "panning". The plates are then washed to remove unbound phages and the bound phages are eluted under conditions that exploit the remarkably high stability of filamentous phages to extreme conditions of pH, ionic strength and/or denaturing agents. This *in vitro* selection step typically enriches the desired phages at least 10-fold but it may be up to 1000-fold in most favorable cases. Further rounds of selection will amplify the eluted phage(s) so that enough material is available for sequencing. The selection of highly selective and potent inhibitors from a pool of phages require a pure target protease, which is generally immobilized on microtiter plates or beads and incubated with the phage library. However the active site of the immobilized protease may be shielded so that the phage-exposed inhibitor cannot bind to the target protease. A soluble target protease can be used to overcome this problem. Once the complex is formed, the phage–protease complexes can be immobilized on beads coated with streptavidin or neutravidin, provided the target protease has been biotinylated [85,97,98]. This approach has additional advantages. First, it requires less purified protease; second, it ensures that the protease has the proper conformation to bind inhibitors; and third, it allows the concentration of the target protease to be controlled, which has a direct effect on the affinity-based selection. It is better to use a low protease concentration to preferentially select inhibitors with K_i values in the nanomolar or subnanomolar range, as this ensures that the E/K_i ratio is $\geq$10, so favoring the binding of high-affinity phages rather than low-affinity phages (Table 3). Although these kinetic rules do not apply to immobilized target proteases, a greater fraction of bound phages generally indicates that the displayed inhibitor has a high affinity.

Highly selective inhibitors of the protease target can be identified by including competing protease(s) of undesired specificity in the panning step. This will remove poorly selective inhibitors [85,99]. Two additional features must be evaluated before the immobilized target protease is incubated with the phage library [98]. First, the non-specific binding to the immobilized target must be as low as possible. It can be evaluated in binding assays using phages that display no inhibitor (WT phages), or phages that display an inhibitor that does not bind to the target protease. This is important to keep in mind since phages have a high non-specific affinity for various surfaces. Second, the potentially deleterious effects of proteases on phage proteins must be checked because they are active enzymes.

Table 3
Effect of the $[E]_0/K_i$ ratio and target protease concentration on the theoretical proportion of an inhibitor-exposed phage bound to the target protease. For a given affinity (K_i) of an inhibitor exposed at the surface of a phage particle, the proportion of bound phages is calculated from the following equation for tight-binding inhibition [140]: % bound inhibitor $= (([E]_0 + [I]_0 + K_i) - \sqrt{([E]_0 + [I]_0 + K_i)^2 - 4[E]_0[I]_0})/(2[E]_0)$ where $[E]_0$ is the initial enzyme concentration and $[I]_0$ the initial concentration of free inhibitor-displayed phage. Calculations were done assuming the inhibitor and enzyme concentrations are equimolar i.e complex formation obeys true second order conditions.

	K_i (nM) of inhibitor-exposed phage for target protease		
	0.1	1	1000
Low target protease concentration (1 nM)	73% ($[E]_0/K_i = 10$)	38% ($[E]_0/K_i = 1$)	0.1% ($[E]_0/K_i = 10^{-3}$)
High target protease concentration (1 μM)	99% ($[E]_0/K_i = 10^4$)	97% ($[E]_0/K_i = 10^3$)	38% ($[E]_0/K_i = 1$)

One way to overcome such a problem is to reduce the time that the phage is incubated with the target protease, especially if phage infectivity is reduced or lost in all the selection steps.

8. Phage display applied to protease inhibitors

8.1. Affinity selection of canonical serine protease inhibitors

One of the most popular and most successful applications of phage display has been the generation of tight-binding inhibitors of any desired protease target by altering the selectivity of protease inhibitors. The earliest inhibitors tested, in the 1990s, included libraries of canonical serine protease inhibitors of the Kunitz family including bovine pancreatic trypsin inhibitor (BPTI), lipoprotein-associated coagulation inhibitor and Alzheimer's amyloid β-protein precursor (APPI), each with various levels of randomization within the inhibitor (Table 2). BPTI is a bovine inhibitor that was used clinically (Trasylol®) for which no human ortholog has been found. It has been widely used in a wealth of protein engineering studies because it is very stable and its inhibitory profile is well known. The first phage display library designed with BPTI was rather small since it included ~1000 variants in which a restricted set of amino acids were incorporated at critical positions (P1–P4′) within the BPTI inhibitory loop [100]. Panning against immobilized human neutrophil elastase, which is poorly inhibited by BPTI ($K_i = 3.6 \times 10^{-6}$ M [101]), yielded a series of BPTI variants which strongly inhibited neutrophil elastase. The variant in which the IAFFP sequence replaced the wild-type sequence KARII at P1–P4′ inhibited neutrophil elastase with a K_i of 1 pM, i.e its affinity for elastase was 3.6×10^6 times greater than that of WT BPTI. In similar, more recent experiments, Otlewski's group improved the ability of BPTI to inhibit chymotrypsin 420-fold and its ability to inhibit porcine pancreatic elastase 7×10^6 fold by randomizing the P1, P1′, P2′ and P3′ positions in its inhibitory loop [102].

The scaffold of the 58-residue APPI, which is structurally similar to BPTI, was used to select APPI variants that specifically inhibited a variety of serine proteases. The randomization scheme applied to APPI was almost the same as that used for BPTI, modifying residues within the inhibitory loop, except that residues 34 and 39 involved in secondary interactions were also modified. This approach produced an APPI variant bearing 4 mutations (T11P, M17L, S19L and G39Y) that inhibited factor VIIa 150-times better ($K_i = 2$ nM) than WT APPI. However it was not very selective; it also inhibited plasmin and factor Xa in the nanomolar range [103]. A second set of experiments included protease partners in the selection step to increase selectivity for factor VIIa by removing phages displaying APPI variants with affinity for factor XIa and/or plasma kallikrein and plasmin. Selective APPI-derived inhibitors of factor VIIa were then obtained in which Lys was highly preferred at P4′ instead of Leu found in direct selection of the APPI phage library [99]. This group used the same APPI library to select a variant APPI with 6 mutations (T11D, P13H, M17A, I18H, S19P, F34Y). This was a potent inhibitor of plasma kallikrein with a K_i of 15 pM, 10,000-times better than WT APPI [104].

Markland et al. [88,89,98] used another Kunitz domain, LACI-D1 the first domain of human lipoprotein-associated coagulation factor, as a scaffold to select high-affinity binders of plasmin, plasma kallikrein and thrombin. The LACI-D1 homologous sequences and the 3D structure of the trypsin–BPTI complex were analyzed to select thirteen positions involved in interactions at the protease–inhibitor interface for randomization. They avoided building a huge library to ensure representation of all amino acids combinations by using an iterative strategy to screen small libraries (*vide ante*, §5). Panning with plasmin yielded a LACI-D1 variant having nine mutations that was a 12,500-times better inhibitor ($K_i = 87$ pM) than WT LACI-D1. They also selected a LACI-D1 variant from the same library that was a potent inhibitor of plasma kallikrein, with a K_i value of 40 pM similar to that of APPI variants. Panning the libraries with thrombin identified a set of binders with a narrower distribution of amino acid mutations but that did not inhibit thrombin. The most likely explanation is that library screening selected variants that bound to a region of the protease other than the active site [89]. This suggests that the scaffolds of Kunitz inhibitor could also be used to select ligands that bind to targets other than proteases.

Tanaka et al. [105] used the scaffold of a member of the Kazal family of canonical inhibitors, leech-derived tryptase inhibitor or LDTI, a small inhibitor (46 residues) of tryptase, trypsin and chymotrypsin. Mutations were restricted to the P1–P4′ positions of the inhibitory loop of LDTI. LDTI-based thrombin binders inhibited only thrombin and trypsin, and not factor Xa, plasma kallikrein or neutrophil elastase. This study was later extended by selecting LDTI variants that bound tightly to immobilized plasmin or neutrophil elastase [106]. They found new powerful LDTI-derived inhibitors that inhibited neutrophil elastase, plasmin, or both enzymes with affinities in the nanomolar range. Most of the mutations did not fit the nature of amino acids found at the corresponding positions in other natural elastase inhibitors. These various studies emphasize the power of the combinatorial approach provided by the phage display technique. In most cases, there is no other way that most, if not all, the mutations necessary to produce a high affinity inhibitor from a low-affinity one can be identified by rational design, even when the 3D structure of the engineered inhibitor is known.

8.2. Engineering ecotin, a unique serine protease inhibitor

Ecotin is a canonical inhibitor of serine proteases that was first isolated from *E. coli* [107] and later found in other gram-negative bacteria [20]. It is present in the periplasmic space where it presumably inhibits host proteases. It has a remarkably broad specificity, inhibiting almost all serine proteases with a chymotrypsin fold, but it does not inhibit *E. coli* serine proteases. Compared to other canonical inhibitors, it has several additional structural features that make it a unique and refined protein inhibitor. First it has an unusual P1 Met residue instead of the Lys or Arg residue found in good trypsin inhibitors at that position. As suggested by the 3D structures of ecotin complexed to various proteases, the Met side chain is flexible enough to fit the active site pocket of any serine protease [108]. It thus plays a key role in determining the panspecificity of ecotin. Second, ecotin differs from other canonical inhibitors in that it forms an active dimer through a swapping-like mechanism whereby the extended C-terminal end of each monomer (142 residues) allow the head-to-tail arrangement of the two monomers (Fig. 5). As a consequence, the partner ecotin monomer forms a second generic binding site for proteases. This adds a further surface for interaction with the protease of about 900 Å^2, in addition to the primary interaction site involving the ecotin inhibitory loop which has an area of approximately 1900 Å^2 [108]. Thus, the protease–inhibitor complex has one ecotin molecule that interacts with both the protease active site and with a second ecotin molecule to form an heterotetramer in which two protease molecules are bound (Fig. 5). This unique combination of substrate-like interactions at the primary site and antibody-like interactions at the secondary site enables ecotin to bind to various proteases, since non-optimal interactions at the primary site may be compensated for by the secondary binding site. The relative contribution of each interface to the total binding energy may thus depend on the shape and primary specificity of the target protease [109]. With its panspecificity, it seems that ecotin has evolved to be a good inhibitor of many proteases ($K_i \sim 10^{-9}$ M) and optimal

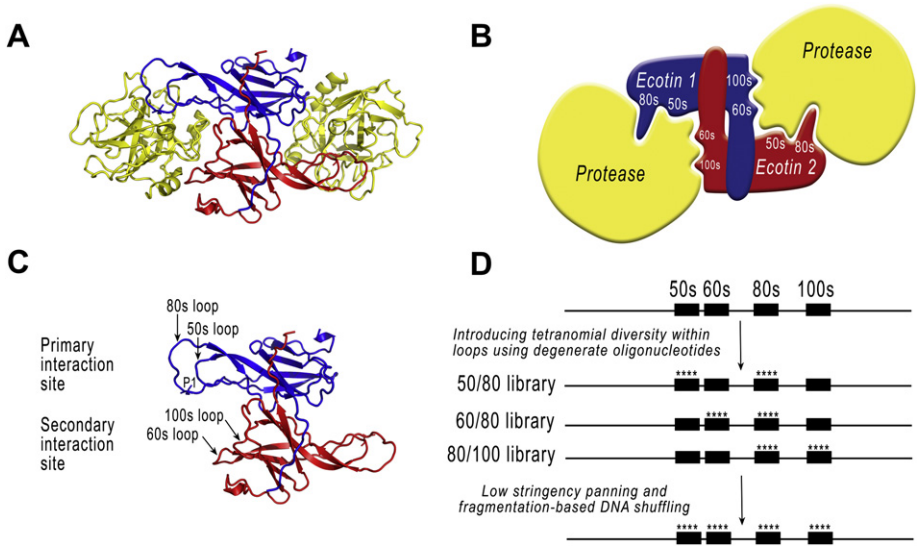

Fig. 5. Ecotin structure and strategy for its phage display. A: ribbon representation of the ecotin dimer complexed to fiddler crab collagenase (PDB code 1AZZ). Ecotin monomers are shown in red and blue and the two complexed collagenase molecules in yellow. B: Diagram illustrating the mode of interaction of ecotin dimer with its protease target. The primary interaction site is composed of the primary binding loop, the 80s loop, which has a canonical conformation that allows docking into the protease active site, and the supporting 50s loop. The secondary binding site includes the 60s and 100s loops contributed by the second ecotin monomer, which both interact with a protease region distal from the active site. C: ribbon representation of the ecotin dimer extracted from the above complex, showing the four loops interacting with the protease surface. D: strategy for randomizing the four ecotin loops that interact with the protease surface. As described in the text, sub-libraries were designed so that tetranomial amino acid diversity was restricted to two of the four ecotin loops. The three libraries were then combined using a DNA shuffling-like technique to give a single library after a low-stringency panning step [85].

($K_i << 10^{-9}$ M) for a restricted set of proteases. These extra contacts in ecotin, which are missing in other canonical inhibitors, have provided opportunities to increase levels of specificity/selectivity of ecotin towards various predetermined target proteases using phage display approaches to modify both interaction sites.

Craik's group at the University of California has developed a phage display system in which functional pIII-fused ecotin is displayed monovalently at the surface of M13 phages [93]. They have used this system, in a series of studies, to generate ecotin variants with enhanced potency towards various target proteases. For example, they used uPA as a target to engineer new ecotin variants because it has the same substrate specificity as bovine trypsin but is inhibited by WT ecotin ≈10,000 fold less than trypsin, despite the fact that the structures of the two catalytic domains are similar. In addition, because it is involved in cancer metastasis and tumor invasion, uPA is a good target for use in the design of inhibitors using the power of phage display techniques to identify ecotin variants that are potent inhibitors of uPA but difficult to predict in a rational manner. Previous studies with small canonical inhibitors showed that a dramatic shift in specificity could be obtained by modifying the P1 residue or residues in its immediate vicinity. An initial phage library was then designed that allowed full randomization (any of 20 amino acids) at the P1 (Met 84) and P1′ (Met 85) residues of the ecotin inhibitory 80s loop. A strong consensus sequence of Arg or Lys at P1 and Arg at P1′ quickly emerged for uPA-binding phages. Subsequent expression of various P1–P1′ ecotin mutants in *E. coli* demonstrated that the affinity for uPA was increased (decreased K_i) about 700–2800 fold, depending on the mutants, compared to WT ecotin ($K_i = 2.8$ μM). Although the P1 residue was clearly the main specificity determinant, the double ecotin mutant M84R/M85R had a much greater inhibitory activity ($K_i = 1$ nM) than single mutants at position 84 or 85, suggesting that the effects of the two mutations were synergistic. At that time, ecotin M84R/M85R was the most potent uPA inhibitor ever identified [110].

The group later designed even more potent inhibitors of uPA by selecting displayed ecotin variants in which only a few positions were randomized within the 60s loop [87] or the 100s loop [111] that form the secondary binding site (Fig. 5). The phage library in which the 60s loop was randomized included the M84R mutation at P1 plus any substitution with one of 20 possible amino acids at positions 67–70. Panning this library against uPA generated the consensus sequence WGY(R/P) at positions 67–70, while panning against rat trypsin revealed no consensus sequence. However, panning a similar library in which the M84R mutation was absent against rat trypsin resulted in the consensus sequence WG(FL)P, confirming the hypothesis that binding regions distal to the primary binding site in ecotin contribute to inhibitory potency and/or protease selectivity. The ecotin mutant M84R + D70R was found to have a very low K_i (50 pM) towards uPA, a 56,000 fold greater binding potency than that of WT ecotin.

Two libraries were designed to investigate how ecotin residues within the 100s loop influence uPA inhibition. One was a complete library in which positions 108, 110, 112 and 113, previously identified as critical in the ecotin–protease interaction [112], were randomized with any of the 20 natural amino acids. The second library focused on the M84R mutation at P1 and the same four positions (108, 110, 112 and 113) of the 100s loop. Each was randomized with the five amino acids that a computational study predicted would interact best with uPA [111]. Selected mutants from the two libraries generated independent consensus sequences for tight uPA binders; one was R108-R110-W112-S113 from the focused library and the other was R(R/N)QL at the same positions from the complete library panned with a low percentage (15%) of RRWS sequence among the sequenced clones. The molecular modeling studies used to guide the design of the focused library successfully predicted a preference for Arg at positions 108 and 110, but did not accurately predict the amino acid preference at positions 112 and 113. This was attributed to limitations in the scoring function used to calculate the energy for the binding of uPA to various modeled ecotin variants [111]. However, using results from computational studies to design an appropriate library is quite attractive in that it may reduce drastically the numbers of

permutations, and hence the number of transformants, that must be obtained in the library. The results were combined with previously identified substitutions in the 60s loop to generate the ecotin variant D70R/M84R/N110R/K112Q that is a potent uPA inhibitor ($K_i = 50$ pM) [111].

In parallel to these studies by Craik's group, several proteases related to the trypsin fold (S1A subfamily in MEROPS) were discovered and shown to be involved in various diseases like cancer. They are thus potential targets for antiprotease-based inhibition strategies. However, the highly conserved 3D structure of S1A serine proteases makes the design of an inhibitor that targets a single protease a quite difficult, complex task. Craik's group extended their previous phage display studies with ecotin to generate an ecotin library in which there was diversity in all four surface loop regions of ecotin involved in its interaction with the protease [85]. The structural studies of ecotin complexed to various proteases suggested that sculpting the ecotin contact area using a combinatorial approach would allow to identify highly selective inhibitors of a given S1A protease and more specifically of plasma kallikrein, matriptase (MT-SP1) and factor XIIa. Twenty positions within the four ecotin loops were to be varied, but the number of theoretical combinations ($4^{20} = 1.1 \times 10^{12}$), even with just four mutations at each position, was too high for a single library to be complete. This problem was circumvented by randomizing the four surface loops in a two-step procedure (Fig. 5). This generated three libraries combining mutations in two loops at a time, first the 50s and 80s loops, then the 60s and 80s loops and lastly the 80s and 100s loops. Based on amino acids found at the corresponding positions in ecotin orthologs (*Pseudomonas aeruginosa*, *Yersinia pestis*, *Trypanosoma brucei*) on one hand and on the nature of amino acids found at hot spots in protein–protein interfaces on the other hand, only four amino acids including that of WT ecotin were incorporated at each concerned position [85]. Thus the tetranomial diversity of each library kept their sizes reasonable ($\sim 4 \times 10^7$ clones). Each of the three libraries then underwent rounds of low-stringency selection for clones that bound plasma kallikrein. The second step in library design consisted of recombining all previous selected libraries into a single library. This was done by a DNA shuffling technique in which all three libraries were fragmented with DNAse I. The complete ecotin genes were then reassembled into a single library. To ensure selective inhibitors of plasma kallikreins to be retained from the shuffled library, competitive panning was performed by pre-incubating ecotin-displaying phages with five related S1A proteases (factor Xa, factor XIa, factor XIIa, uPA and matriptase). Among the six unique ecotin variants with selective inhibitory activity towards plasma kallikrein, ecotin-Pkal (ecotin H53P, S82W, T83N, M84R, M85R, A86S, R108S, N110S) was found to be the most potent and contained mutations in the 50s, 80S and 100s loops. This highly selective molecule inhibits plasma kallikrein six times better than WT ecotin, with a K_i as low as 11 pM, while at the same time, the affinity of ecotin-Pkal lower for the other five related ecotin-inhibited serine proteases was decreased by at least two orders of magnitude [85]. In addition to its potential therapeutic value, ecotin-Pkal is a valuable tool with which to decipher the biological roles of plasma kallikrein *in vivo*. Other ecotin variants selective for matriptase and factor XIIa were obtained from the same shuffled library, but the increase in selectivity was less marked than for inhibitors of plasma kallikrein. This is probably due to the fact that the tetranomial diversity incorporated in the library does not allow a sufficient space sequence to be explored. Nevertheless these impressive phage display studies with ecotin clearly demonstrate that the ecotin scaffold is uniquely adaptable to many proteases because of the complex interface it forms with target proteases.

These studies demonstrate the power of a well-designed phage display project for dramatically improving the inhibitory potency of a molecule for a given protease target through successive optimization steps.

8.3. Understanding the structure–function relationships of protease inhibitors

The pacifastins are small, cysteine-rich, canonical inhibitors of about 35 amino acids that were discovered in arthropods. They have a very small core region, a relatively large number of disulfide bridges, and they are extremely thermostable [113]. There are two groups. Group I pacifastins have a protein core that is governed by a Lys10–Trp26 interaction, whereas group II pacifastins have a core organized around Phe10, which is surrounded by small hydrophobic residues. They inhibit various types of serine proteases. The group I pacifastins are almost inactive against vertebrate proteases but the group II pacifastins are not. For example, SGPI-1 isolated from the locust *Schistocerca gregaria* inhibits arthropod trypsins, such as crayfish trypsin, five orders of magnitude more efficiently than it inhibits mammalian trypsins [114]. The 3D structure of the SGTI-1/crayfish trypsin complex revealed that, in addition to the inhibitory loop, two regions of the inhibitor (P4–P12 and P4′–P5′) form extended interactions with crayfish trypsin [115]. Although these additional interactions clearly provide significant more binding energy for crayfish trypsin than for bovine trypsin, the structural basis for taxon selectivity of group I inhibitors is not fully understood. Replacing the P1 Leu of SGPI-2 by Arg switched SGPI-2 from a chymotrypsin inhibitor to a potent inhibitor of both bovine and arthropod trypsins, suggesting that the taxon selectivity of pacifastin is related to the group I fold.

A sophisticated phage display approach was used to decipher the structural determinants of taxon selectivity. Sequences of taxon-selective SGPI-1 and taxon-non-selective SGPI-2 were both randomized by introduction of 2 types of amino acids (as found in SGPI-1 or SGPI-2) at the 18 positions by which the 2 inhibitors differ to design a complete chimeric SGPI-1/SGPI-2 phage library [116]. This library was then selected *in vitro* for binding to immobilized crayfish trypsin, and binding to bovine trypsin. Such a phage display strategy has two advantages, first it avoids the Herculean task of analyzing each of 2^{18} (131,072) variants, and second it excludes *a priori* the bias related to the preconceived ideas about which amino acids play a key role in governing the protease selectivity [116]. Sequence analysis of bovine trypsin-selected clones revealed that the nature of amino acid side chain at some positions was directly correlated with the core sequence, suggesting that there is intra-molecular functional coupling between the two regions. A pair-wise covariance analysis confirmed that wild-type residues P21 and T22 were excluded from those positions when the clones harbour a K10-W26 core type as found in WT SGPI-1. This suggests that the K10-W26 core interacts with the T20-P21-T22 region (P8–P10 region), which forms a beta-turn in the SGPI-1 structure, which then hampers binding to bovine trypsin [115]. In addition, introducing five surface residues of the taxon-non-selective SGPI-2 into the SGPI-1 structure allowed the mutated SGPI-1 to inhibit both bovine and crayfish trypsins while retaining its native taxon-specific K10-W26 core. Thus there is a complex network of interactions between residues at the surface of the inhibitor and between them and residues of the protein core. This tremendous phage display-based work sheds new light on the structure–function features of a family of canonical inhibitors by suggesting that residues of surface loop(s) are not the sole determinants of inhibitory selectivity.

The cystatins are another example of the use of phage display to improve our understanding of structure–function relationships in protease inhibitors. A library was designed with cystatin A (also know as stefin A) in which the four N-terminal residues were fully

randomized [117]. Panning the library with papain as target protease yielded cystatin variants that all contained a glycine at position 4 and various hydrophobic residues at positions 1–3. This not only confirmed previous evidence that glycine 4 is essential for the tight binding of cystatins to cysteine proteases, it also suggests that the selectivity of cystatins can be increased by appropriate substitutions at the first three residues, provided they are mainly hydrophobic. Another study designed to identify the optimal residues for plant cystatins (or phytocystatins) to interact with cysteine proteases used soyacystatin N that was fully randomized at four positions (48–51) in the QxVxG-type first hairpin loop and two positions (78–79) in the second hairpin loop [92]. These two loops, together with the N-terminal end, form the wedge-shaped edge that is the inhibitory site of cystatins (Fig. 3). Cystatin variants selected from the library using their affinity for papain were found to have conserved the QVVAG motif in the second loop, while the glutamate residue at position 78 in the third loop, could be replaced by a hydrophobic or basic residue, but not the tryptophan at position 79. Kinetic studies with soyacystatin E78L demonstrated that this variant had a moderately increased affinity for papain; it was only 2 fold due to a greater association rate constant [92].

8.4. Other applications

In addition to the affinity selection of potent inhibitors, phage display has also been used to study protein–protein interactions. A PAI-1 library containing predominantly random, single mutants of PAI-1 was generated by error-prone PCR to identify residues of this serpin that interact with thrombin or with a positively-charged loop (VR1) at the surface of its target protease tPA [91]. A two-step selection was performed by first incubating the phage library with a thrombin mutant whose VR1 region had been replaced by that of tPA (negative selection), and then incubating the remaining phages with tPA before panning with an immobilized anti-tPA antibody (positive selection). Sequence analysis of the selected clones revealed that three positions P2, P1′ and P4′ were crucial for the interaction of PAI-1 with either thrombin or VR1 in tPA. The same group used this hypermutated PAI-1 library (§ 6) to identify dominant amino acids in PAI-1 that are involved in its binding to four distinct monoclonal antibodies by analyzing the sequence of defective clones that did not bind to three of the four monoclonal antibodies [86]. Yet another study used the library to identify residues implicated in the latency transition by locating mutations that held PAI-1 in its stable, inactive (latent) conformation [82].

The highly specific thrombin-inhibitor hirudin, like many proteinaceaous inhibitors, may be sensitive to unwanted proteolytic degradation. A combination of phage display selection and high-throughput screening methods was used to show that several hirudin variants with substitutions in the first few N-terminal residues (residues 1–5) and at the C-terminal end (residues 49, 50, 56, 62–64) were resistant to degradation by pepsin and chymotrypsin. However, the capacity of hirudin to inhibit thrombin was somewhat altered as the IC-50 was 100-time higher than that of WT hirudin [118]. However that was not a major issue in that special case because the affinity of WT hirudin for thrombin is in the subpicomolar range.

Phage display has been used in an original way to identify serine protease inhibitors from the dog hookworm *Ancylostoma caninum* [96]. The cDNA sequences from a cDNA library derived from this parasite were fused to the 3′ end of the gene encoding the pVI coat protein of M13 phages in each of the three possible reading frames. The displayed proteins were then panned with two proteases to identify cDNAs encoding inhibitors. The purification of a Kunitz-type inhibitor from a parasite extract confirmed the validity of this attractive approach to finding novel inhibitors.

Although not directly related to protease inhibitors, phage display has been widely used to map protease specificity so as to identify consensus substrate sequences for the design of synthetic substrates. This special application, called substrate phage display, uses a very elegant library design in which a random peptide substrate sequence is fused between a phage coat protein and a protein (e.g. hGH, human growth hormone) for which a tight-binding partner may be immobilized (e.g. hGH binding protein). Protease-mediated release of bound phages resulting from the hydrolysis within the randomized peptide sequence allows the identification of amino acid sequences that behave as good or bad substrates of the protease under study (see [119] for review). Some recent examples have allowed the specificity of various proteases to be elucidated including MASP-2 [120], C1s protease [121], human kallikreins KLK1 and KLK6 [122] or Factor Xa [123].

Phage display has also been used to modify natural inhibitors of cysteine and serine proteases in order to improve plant defenses against the digestive proteases of insects. For example, a cystatin derived from the nematode *Onchocerca volvulus* has been variegated with all 20 possible amino acids in the QVVAG motif of its second loop [124] that forms part of the inhibitory site (Fig. 3). Onchocystatin variants were selected by panning the library with purified cysteine proteases from the bean weevil *Acanthoscelides obtectus*. Their QVVAG sequence was found to be replaced by either DVVSA or NTSSA. The two variants bound to *A. obtectus* cysteine proteases more tightly than to papain. Although they were not further characterized, their genes could well be good candidates for producing transgenic plants resistant to this insect pest. In another attempt to use proteinase inhibitors for plant defense, MTI-2, a canonical inhibitor of trypsin isolated from mustard seeds, was randomized at five positions (P3–P2′) within its inhibitory loop and subjected to phage display selection against trypsin and chymotrypsin [125]. Nanomolar variant inhibitors of chymotrypsin and picomolar typsin inhibitors were obtained although WT MTI-2 does not inhibit chymotrypsin. One MTI-2 variant, Chy8, was found to be eight times more toxic for aphids than WT MTI-2. The authors suggest that MTI-2 variants could be incorporated into transgenic plants to increase their resistance to sucking insect pests like aphids.

9. Concluding remarks

Twenty five years have elapsed since the first demonstration that functionally active proteins could be displayed at the surface of filamentous phages. This needle-in-a-haystack technology has proved that *in vitro* directed evolution mimicking in some ways how natural proteins evolve is clearly achievable. Many different protein scaffolds have subsequently been used to tackle a wide diversity of biological issues [126–130]. There will, undoubtedly be further improvements in all aspects of phage display technology [131], including the development of other new display systems using yeast, bacteria, or even ribosomes, as well as other families of phages (T4, T7, lambda), although they have not yet been used in the field of proteinase inhibitors.

The most successful and attractive phage display applications for engineering protease inhibitors have triggered the generation of many new, highly selective inhibitors of a rather restricted set of proteases, mainly serine proteases. Recent advances in the field of proteases have led to the identification of new validated target proteases of other classes, including MMPs and cysteine proteases [2]. However, the numerous attempts to target MMPs with small inhibitors have failed due to the severe side-effects of the broad spectrum inhibitors developed to date in addition to providing minor clinical benefits [2]. These detrimental off-target effects could be minimized by the development of highly selective inhibitors of a single MMP. Engineering natural protein inhibitors of MMPs like

TIMPs by phage display to produce this selectivity is theoretically attractive, but is likely to be a challenging task because MMPs all have extremely similar catalytic domains. To circumvent this, antibody-based approaches to target protease active site may represent a valuable alternative in generating neutralizing antibodies against proteases from libraries of phage-displayed antibodies as shown recently. For example, monoclonal antibodies specific to the murine catalytic domains of MMP-1, MMP-2 and MMP-3 have been obtained by a phage display approach. These novel antibodies have been demonstrated to selectively inhibit their corresponding target protease with an affinity in the low nanomolar range [132]. A second example of this promising approach is the selection of a highly selective and potent antibody inhibitor of a transmembrane serine protease, matriptase or MT-SP1, from a phage-displayed human scFv library [133]. Analysis of the X-ray structure of the inhibitory antibody/matriptase complex has shown that, in addition to interacting in the substrate-like manner as do many natural inhibitors, the variable regions of the antibody-based inhibitor gain selectivity by binding to surface loops surrounding the protease active site, thus preventing substrates from reaching the protease active site [134]. A last recent example is the selection from a phage library of a heavy chain fragment that neutralized the zinc–metalloproteinase activity of botulinum neurotoxin type A [135]. This combinatorial, antibody-based approach is still at an early stage but it should provide inhibitors that are specific for a single member of a family of closely related proteases, such as the MMPs when protease inhibitor scaffolds are not able to provide the required selectivity.

Acknowledgments

The work on proteases and protease inhibitors in our laboratory is supported by institutional funding from the Institut National de la Santé et de la Recherche Médicale (INSERM) and grants from Vaincre la Mucoviscidose, ANR and Région Centre. The authors thank Prof. Philippe Minard, Université de Paris Sud for helpful discussions. The English text was edited by Dr. Owen Parkes.

References

[1] I. Schechter, A. Berger, On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. 27 (1967) 157–162.
[2] B. Turk, Targeting proteases: successes, failures and future prospects. Nat. Rev. Drug Discov. 5 (2006) 785–799.
[3] G.P. Smith, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228 (1985) 1315–1317.
[4] T. Clackson, H.R. Hoogenboom, A.D. Griffiths, G. Winter, Making antibody fragments using phage display libraries. Nature 352 (1991) 624–628.
[5] J. McCafferty, A.D. Griffiths, G. Winter, D.J. Chiswell, Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348 (1990) 552–554.
[6] G.P. Smith, J.K. Scott, Libraries of peptides and proteins displayed on filamentous phage. Methods Enzymol. 217 (1993) 228–257.
[7] N.D. Rawlings, A.J. Barrett, A. Bateman, MEROPS: the peptidase database. Nucleic Acids Res. 38 (2010) D227–D233.
[8] W. Bode, R. Huber, Natural protein proteinase inhibitors and their interaction with proteinases. Eur. J. Biochem. 204 (1992) 433–451.
[9] G.A. Silverman, P.I. Bird, R.W. Carrell, F.C. Church, P.B. Coughlin, P.G. Gettins, J.A. Irving, D.A. Lomas, C.J. Luke, R.W. Moyer, P.A. Pemberton, E. Remold-O'Donnell, G.S. Salvesen, J. Travis, J.C. Whisstock, 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 (2001) 33293–33296.
[10] J. Otlewski, F. Jelen, M. Zakrzewska, A. Oleksy, The many faces of protease–protein inhibitor interaction. EMBO J. 24 (2005) 1303–1310.
[11] M. Laskowski, M. Laskowski Jr., Naturally occurring trypsin inhibitors. Adv. Protein Chem. 9 (1954) 203–242.
[12] M. Laskowski Jr., I. Kato, Protein inhibitors of proteinases. Annu. Rev. Biochem. 49 (1980) 593–626.
[13] J. Travis, G.S. Salvesen, Human plasma proteinase inhibitors. Annu. Rev. Biochem. 52 (1983) 655–709.
[14] R.M. Jackson, R.B. Russell, The serine protease inhibitor canonical loop conformation: examples found in extracellular hydrolases, toxins, cytokines and viral proteins. J. Mol. Biol. 296 (2000) 325–334.
[15] R.M. Jackson, R.B. Russell, Predicting function from structure: examples of the serine protease inhibitor canonical loop conformation found in extracellular proteins. Comput. Chem. 26 (2001) 31–39.
[16] M.J. Laskowski, M.A. Qasim, S.M. Lu, Interaction of standard mechanism, canonical protein inhibitors with serine proteinases. in: C. Kleanthous (Ed.), Protein Protein Recognition: The Frontiers in Molecular Biology Series. Oxford University Press, Oxford, 2000, pp. 228–279.
[17] J. Janin, C. Chothia, The structure of protein–protein recognition sites. J. Biol. Chem. 265 (1990) 16027–16030.
[18] D. Krowarsch, T. Cierpicki, F. Jelen, J. Otlewski, Canonical protein inhibitors of serine proteases. Cell. Mol. Life Sci. 60 (2003) 2427–2444.
[19] Y. Xu, P.D. Carr, J.M. Guss, D.L. Ollis, The crystal structure of bikunin from the inter-alpha-inhibitor complex: a serine protease inhibitor with two Kunitz domains. J. Mol. Biol. 276 (1998) 955–966.
[20] C.T. Eggers, I.A. Murray, V.A. Delmar, A.G. Day, C.S. Craik, The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase. Biochem. J. 379 (2004) 107–118.
[21] M.E. McGrath, T. Erpel, C. Bystroff, R.J. Fletterick, Macromolecular chelation as an improved mechanism of protease inhibition: structure of the ecotin–trypsin complex. EMBO J. 13 (1994) 1502–1507.
[22] J.A. Huntington, R.J. Read, R.W. Carrell, Structure of a serpin–protease complex shows inhibition by deformation. Nature 407 (2000) 923–926.
[23] S.R. Stone, J. Hofsteenge, Kinetics of the inhibition of thrombin by hirudin. Biochemistry 25 (1986) 4622–4628.
[24] H. Czapinska, J. Otlewski, Structural and energetic determinants of the S1-site specificity in serine proteases. Eur. J. Biochem. 260 (1999) 571–595.
[25] J. Otlewski, M. Jaskolski, O. Buczek, T. Cierpicki, H. Czapinska, D. Krowarsch, A.O. Smalas, D. Stachowiak, A. Szpineta, M. Dadlez, Structure–function relationship of serine protease-protein inhibitor interaction. Acta Biochim. Pol. 48 (2001) 419–428.
[26] J.A. Kraunsoe, T.D. Claridge, G. Lowe, Inhibition of human leukocyte and porcine pancreatic elastase by homologues of bovine pancreatic trypsin inhibitor. Biochemistry 35 (1996) 9090–9096.
[27] A. Desilets, J.M. Longpre, M.E. Beaulieu, R. Leduc, Inhibition of human matriptase by eglin c variants. FEBS Lett. 580 (2006) 2227–2232.
[28] D.W. Heinz, S.G. Hyberts, J.W. Peng, J.P. Priestle, G. Wagner, M.G. Grutter, Changing the inhibitory specificity and function of the proteinase inhibitor eglin c by site-directed mutagenesis: functional and structural investigation. Biochemistry 31 (1992) 8755–8766.
[29] T. Komiyama, R.S. Fuller, Engineered eglin c variants inhibit yeast and human proprotein processing proteases, Kex2 and furin. Biochemistry 39 (2000) 15156–15165.
[30] M.A. Qasim, S.M. Lu, J. Ding, K.S. Bateman, M.N. James, S. Anderson, J. Song, J.L. Markley, P.J. Ganz, C.W. Saunders, M. Laskowski Jr., Thermodynamic criterion for the conformation of P1 residues of substrates and of inhibitors in complexes with serine proteinases. Biochemistry 38 (1999) 7142–7150.
[31] S. Khamrui, J. Dasgupta, J.K. Dattagupta, U. Sen, Single mutation at P1 of a chymotrypsin inhibitor changes it to a trypsin inhibitor: X-ray structural (2.15 Å) and biochemical basis. Biochim. Biophys. Acta 1752 (2005) 65–72.
[32] S. Kojima, I. Kumagai, K. Miura, Effect of inhibitory activity of mutation at reaction site P4 of the *Streptomyces subtilisin* inhibitor, SSI. Protein Eng. 3 (1990) 527–530.
[33] S. Kojima, Y. Nishiyama, I. Kumagai, K. Miura, Inhibition of subtilisin BPN' by reaction site P1 mutants of *Streptomyces subtilisin* inhibitor. J. Biochem. 109 (1991) 377–382.
[34] S. Kojima, S. Obata, I. Kumagai, K. Miura, Alteration of the specificity of the *Streptomyces subtilisin* inhibitor by gene engineering. Biotechnol. (N Y) 8 (1990) 449–452.
[35] C. Longstaff, A.F. Campbell, A.R. Fersht, Recombinant chymotrypsin inhibitor 2: expression, kinetic analysis of inhibition with alpha-chymotrypsin and wild-type and mutant subtilisin BPN', and protein engineering to investigate inhibitory specificity and mechanism. Biochemistry 29 (1990) 7339–7347.
[36] J. Collins, M. Szardenings, F. Maywald, H. Blocker, R. Frank, H.J. Hecht, B. Vasel, D. Schomburg, E. Fink, H. Fritz, Human leukocyte elastase inhibitors: designed variants of human pancreatic secretory trypsin inhibitor (hPSTI). Biol. Chem. Hoppe Seyler 371 (Suppl.) (1990) 29–36.
[37] K. Hilpert, G. Hansen, H. Wessner, R. Volkmer-Engert, W. Hohne, Complete substitutional analysis of a sunflower trypsin inhibitor with different serine proteases. J. Biochem. 138 (2005) 383–390.
[38] E. Zablotna, A. Jaskiewicz, A. Legowska, H. Miecznikowska, A. Lesner, K. Rolka, Design of serine proteinase inhibitors by combinatorial chemistry using trypsin inhibitor SFTI-1 as a starting structure. J. Pept. Sci. 13 (2007) 749–755.
[39] S. Sinha, J. Knops, F. Esch, E.D. Moyer, T. Oltersdorf, Conversion of the Alzheimer's beta-amyloid precursor protein (APP) Kunitz domain into a potent human neutrophil elastase inhibitor. J. Biol. Chem. 266 (1991) 21011–21013.
[40] Z. Malik, S. Amir, G. Pal, Z. Buzas, E. Varallyay, J. Antal, Z. Szilagyi, K. Vekey, B. Asboth, A. Patthy, L. Graf, Proteinase inhibitors from desert locust, *Schistocerca gregaria*: engineering of both P(1) and P(1)' residues converts a potent chymotrypsin inhibitor to a potent trypsin inhibitor. Biochim. Biophys. Acta 1434 (1999) 143–150.
[41] G. Pal, G. Sprengel, A. Patthy, L. Graf, Alteration of the specificity of ecotin, an *E. coli* serine proteinase inhibitor, by site directed mutagenesis. FEBS Lett. 342 (1994) 57–60.
[42] I.S. Seong, H.R. Lee, J.H. Seol, S.K. Park, C.S. Lee, S.W. Suh, Y.M. Hong, M.S. Kang, D.B. Ha, C.H. Chung, The P1 reactive site methionine residue of

ecotin is not crucial for its specificity on target proteases. A potent inhibitor of pancreatic serine proteases from *Escherichia coli*. J. Biol. Chem. 269 (1994) 21915–21918.
[43] J.L. Seymour, R.N. Lindquist, M.S. Dennis, B. Moffat, D. Yansura, D. Reilly, M.E. Wessinger, R.A. Lazarus, Ecotin is a potent anticoagulant and reversible tight-binding inhibitor of factor Xa. Biochemistry 33 (1994) 3949–3958.
[44] S.P. Eisenberg, K.K. Hale, P. Heimdal, R.C. Thompson, Location of the protease-inhibitory region of secretory leukocyte protease inhibitor. J. Biol. Chem. 265 (1990) 7976–7981.
[45] M.L. Zani, K. Baranger, N. Guyot, S. Dallet-Choisy, T. Moreau, Protease inhibitors derived from elafin and SLPI and engineered to have enhanced specificity towards neutrophil serine proteases. Protein Sci. 18 (2009) 579–594.
[46] T. Sulikowski, B.A. Bauer, P.A. Patston, alpha(1)-Proteinase inhibitor mutants with specificity for plasma kallikrein and C1s but not C1. Protein Sci. 11 (2002) 2230–2236.
[47] A. Tsuji, H. Kanie, H. Makise, K. Yuasa, M. Nagahama, Y. Matsuda, Engineering of alpha1-antitrypsin variants selective for subtilisin-like proprotein convertases PACE4 and PC6: importance of the P2' residue in stable complex formation of the serpin with proprotein convertase. Protein Eng. Des. Sel. 20 (2007) 163–170.
[48] M. Watanabe, A. Hirano, S. Stenglein, J. Nelson, G. Thomas, T.C. Wong, Engineered serine protease inhibitor prevents furin-catalyzed activation of the fusion glycoprotein and production of infectious measles virus. J. Virol. 69 (1995) 3206–3210.
[49] B.S. Cooperman, E. Stavridi, E. Nickbarg, E. Rescorla, N.M. Schechter, H. Rubin, Antichymotrypsin interaction with chymotrypsin. Partitioning of the complex. J. Biol. Chem. 268 (1993) 23616–23625.
[50] H. Rubin, M. Plotnick, Z.M. Wang, X. Liu, Q. Zhong, N.M. Schechter, B.S. Cooperman, Conversion of alpha 1-antichymotrypsin into a human neutrophil elastase inhibitor: demonstration of variants with different association rate constants, stoichiometries of inhibition, and complex stabilities. Biochemistry 33 (1994) 7627–7633.
[51] Y.J. Chuang, R. Swanson, S.M. Raja, S.C. Bock, S.T. Olson, The antithrombin P1 residue is important for target proteinase specificity but not for heparin activation of the serpin. Characterization of P1 antithrombin variants with altered proteinase specificity but normal heparin activation. Biochemistry 40 (2001) 6670–6679.
[52] W.P. Sheffield, M.A. Blajchman, Site-directed mutagenesis of the P2 residue of human antithrombin. FEBS Lett. 339 (1994) 147–150.
[53] A.W. Stephens, A. Siddiqui, C.H. Hirs, Site-directed mutagenesis of the reactive center (serine 394) of antithrombin III. J. Biol. Chem. 263 (1988) 15849–15852.
[54] P.M. Sherman, D.A. Lawrence, A.Y. Yang, E.T. Vandenberg, D. Paielli, S.T. Olson, J.D. Shore, D. Ginsburg, Saturation mutagenesis of the plasminogen activator inhibitor-1 reactive center. J. Biol. Chem. 267 (1992) 7588–7595.
[55] S.T. Cooper, F.C. Church, Reactive site mutants of recombinant protein C inhibitor. Biochim. Biophys. Acta 1246 (1995) 29–33.
[56] J.E. Phillips, S.T. Cooper, E.E. Potter, F.C. Church, Mutagenesis of recombinant protein C inhibitor reactive site residues alters target proteinase specificity. J. Biol. Chem. 269 (1994) 16696–16700.
[57] W. Bode, R. Engh, D. Musil, U. Thiele, R. Huber, A. Karshikov, J. Brzin, J. Kos, V. Turk, The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. EMBO J. 7 (1988) 2593–2599.
[58] M.T. Stubbs, B. Laber, W. Bode, R. Huber, R. Jerala, B. Lenarcic, V. Turk, The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. EMBO J. 9 (1990) 1939–1947.
[59] G. Guncar, G. Pungercic, I. Klemencic, V. Turk, D. Turk, Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. EMBO J. 18 (1999) 793–803.
[60] A. Ljunggren, I. Redzynia, M. Alvarez-Fernandez, M. Abrahamson, J.S. Mort, J.C. Krupa, M. Jaskolski, G. Bujacz, Crystal structure of the parasite protease inhibitor chagasin in complex with a host target cysteine protease. J. Mol. Biol. 371 (2007) 137–153.
[61] S.X. Wang, K.C. Pandey, J. Scharfstein, J. Whisstock, R.K. Huang, J. Jacobelli, R.J. Fletterick, P.J. Rosenthal, M. Abrahamson, L.S. Brinen, A. Rossi, A. Sali, J.H. McKerrow, The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. Structure 15 (2007) 535–543.
[62] G. Dubin, Proteinaceous cysteine protease inhibitors. Cell. Mol. Life Sci. 62 (2005) 653–669.
[63] K. Maskos, W. Bode, Structural basis of matrix metalloproteinases and tissue inhibitors of metalloproteinases. Mol. Biotechnol. 25 (2003) 241–266.
[64] J. Vendrell, E. Querol, F.X. Aviles, Metallocarboxypeptidases and their protein inhibitors. Structure, function and biomedical properties. Biochim. Biophys. Acta 1477 (2000) 284–298.
[65] U. Baumann, M. Bauer, S. Letoffe, P. Delepelaire, C. Wandersman, Crystal structure of a complex between *Serratia marcescens* metallo-protease and an inhibitor from *Erwinia chrysanthemi*. J. Mol. Biol. 248 (1995) 653–661.
[66] S. Tate, A. Ohno, S.S. Seeram, K. Hiraga, K. Oda, M. Kainosho, Elucidation of the mode of interaction of thermolysin with a proteinaceous metalloproteinase inhibitor, SMPI, based on a model complex structure and a structural dynamics analysis. J. Mol. Biol. 282 (1998) 435–446.
[67] M. Li, L.H. Phylip, W.E. Lees, J.R. Winther, B.M. Dunn, A. Wlodawer, J. Kay, A. Gustchina, The aspartic proteinase from *Saccharomyces cerevisiae* folds its own inhibitor into a helix. Nat. Struct. Biol. 7 (2000) 113–117.
[68] D.A. Marvin, Filamentous phage structure, infection and assembly. Curr. Opin. Struct. Biol. 8 (1998) 150–158.
[69] D.W. Banner, C. Nave, D.A. Marvin, Structure of the protein and DNA in fd filamentous bacterial virus. Nature 289 (1981) 814–816.
[70] S. Cabilly, The basic structure of filamentous phage and its use in the display of combinatorial peptide libraries. Mol. Biotechnol. 12 (1999) 143–148.
[71] T.J. Henry, D. Pratt, The proteins of bacteriophage M13. Proc. Natl. Acad. Sci. USA 62 (1969) 800–807.
[72] M. Russel, Filamentous phage assembly. Mol. Microbiol. 5 (1991) 1607–1613.
[73] M. Russel, P. Model, Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. J. Virol. 63 (1989) 3284–3295.
[74] M. Russel, H.B. Lowman, T. Clackson, Introduction to phage biology and phage display. in: T. Clackson, H.B. Lowman (Eds.), Phage Display: a Practical Approach. Oxford University Press, 2004, pp. 1–26.
[75] L.W. Deng, P. Malik, R.N. Perham, Interaction of the globular domains of pIII protein of filamentous bacteriophage fd with the F-pilus of *Escherichia coli*. Virology 253 (1999) 271–277.
[76] H.B. Lowman, S.H. Bass, N. Simpson, J.A. Wells, Selecting high-affinity binding proteins by monovalent phage display. Biochemistry 30 (1991) 10832–10838.
[77] S.S. Sidhu, G.A. Weiss, Constructing phage display libraries by oligonucleotide-directed mutagenesis. in: T. Clackson, H.B. Lowman (Eds.), Phage Display: a Practical Approach. Oxford University Press, 2004, pp. 27–41.
[78] F.M. Aslan, Y. Yu, S.C. Mohr, C.R. Cantor, Engineered single-chain dimeric streptavidins with an unexpected strong preference for biotin-4-fluorescein. Proc. Natl. Acad. Sci. USA 102 (2005) 8507–8512.
[79] J. Light, R.A. Lerner, Random mutagenesis of staphylococcal nuclease and phage display selection. Bioorg. Med. Chem. 3 (1995) 955–967.
[80] H. Miyakubo, A. Sugio, T. Kubo, R. Nakai, K. Wakabayashi, S. Nakamura, Phage display of xylan-binding module of xylanase J from alkaliphilic *Bacillus sp*. strain 41M-1. Nucleic Acids Symp. Ser. (2000) 165–166.
[81] H. Pannekoek, M. van Meijer, R.R. Schleef, D.J. Loskutoff, C.F. Barbas 3rd, Functional display of human plasminogen-activator inhibitor 1 (PAI-1) on phages: novel perspectives for structure-function analysis by error-prone DNA synthesis. Gene 128 (1993) 135–140.
[82] A.A. Stoop, E. Eldering, T.R. Dafforn, R.J. Read, H. Pannekoek, Different structural requirements for plasminogen activator inhibitor 1 (PAI-1) during latency transition and proteinase inhibition as evidenced by phage-displayed hypermutated PAI-1 libraries. J. Mol. Biol. 305 (2001) 773–783.
[83] R. Azriel-Rosenfeld, M. Valensi, I. Benhar, A human synthetic combinatorial library of arrayable single-chain antibodies based on shuffling in vivo formed CDRs into general framework regions. J. Mol. Biol. 335 (2004) 177–192.
[84] C. Fermer, I. Andersson, K. Nilsson, O. Nilsson, Specificity rescue and affinity maturation of a low-affinity IgM antibody against pro-gastrin-releasing peptide using phage display and DNA shuffling. Tumour Biol. 25 (2004) 7–13.
[85] A.A. Stoop, C.S. Craik, Engineering of a macromolecular scaffold to develop specific protease inhibitors. Nat. Biotechnol. 21 (2003) 1063–1068.
[86] A.A. Stoop, L. Jespers, I. Lasters, E. Eldering, H. Pannekoek, High-density mutagenesis by combined DNA shuffling and phage display to assign essential amino acid residues in protein–protein interactions: application to study structure–function of plasminogen activation inhibitor 1 (PAI-I). J. Mol. Biol. 301 (2000) 1135–1147.
[87] S.Q. Yang, C.S. Craik, Engineering bidentate macromolecular inhibitors for trypsin and urokinase-type plasminogen activator. J. Mol. Biol. 279 (1998) 1001–1011.
[88] W. Markland, A.C. Ley, S.W. Lee, R.C. Ladner, Iterative optimization of high-affinity proteases inhibitors using phage display. 1. Plasmin. Biochemistry 35 (1996) 8045–8057.
[89] W. Markland, A.C. Ley, R.C. Ladner, Iterative optimization of high-affinity protease inhibitors using phage display. 2. Plasma kallikrein and thrombin. Biochemistry 35 (1996) 8058–8067.
[90] M. Laskowski, M.A. Qasim, What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? Biochim. Biophys. Acta 1477 (2000) 324–337.
[91] M. van Meijer, Y. Roelofs, J. Neels, A.J. Horrevoets, A.J. van Zonneveld, H. Pannekoek, Selective screening of a large phage display library of plasminogen activator inhibitor 1 mutants to localize interaction sites with either thrombin or the variable region 1 of tissue-type plasminogen activator. J. Biol. Chem. 271 (1996) 7423–7428.
[92] H. Koiwa, M.P. D'Urzo, I. Assfalg-Machleidt, K. Zhu-Salzman, R.E. Shade, H. An, L.L. Murdock, W. Machleidt, R.A. Bressan, P.M. Hasegawa, Phage display selection of hairpin loop soyacystatin variants that mediate high affinity inhibition of a cysteine proteinase. Plant J. 27 (2001) 383–391.
[93] C.I. Wang, Q. Yang, C.S. Craik, Phage display of proteases and macromolecular inhibitors. Methods Enzymol. 267 (1996) 52–68.
[94] P. Rottgen, J. Collins, A human pancreatic secretory trypsin inhibitor presenting a hypervariable highly constrained epitope via monovalent phagemid display. Gene 164 (1995) 243–250.

[95] M. Jongsma, P. Bakker, W. Stiekema, D. Bosch, Phage display of a double-headed proteinase inhibitor: analysis of the binding domains of potato proteinase inhibitor II. Mol. Breed 1 (1995) 181–191.
[96] L.S. Jespers, J.H. Messens, A. De Keyser, D. Eeckhout, I. Van den Brande, Y.G. Gansemans, M.J. Lauwereys, G.P. Vlasuk, P.E. Stanssens, Surface expression and ligand-based selection of cDNAs fused to filamentous phage gene VI. Biotechnol. (NY) 13 (1995) 378–382.
[97] R.E. Hawkins, S.J. Russell, G. Winter, Selection of phage antibodies by binding affinity. Mimicking affinity maturation. J. Mol. Biol. 226 (1992) 889–896.
[98] W. Markland, B.L. Roberts, R.C. Ladner, Selection for protease inhibitors using bacteriophage display. Methods Enzymol. 267 (1996) 28–51.
[99] M.S. Dennis, R.A. Lazarus, Kunitz domain inhibitors of tissue factor-factor VIIa. II. Potent and specific inhibitors by competitive phage selection. J. Biol. Chem. 269 (1994) 22137–22144.
[100] B.L. Roberts, W. Markland, A.C. Ley, R.B. Kent, D.W. White, S.K. Guterman, R.C. Ladner, Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. Proc. Natl. Acad. Sci. USA 89 (1992) 2429–2433.
[101] J.P. Vincent, M. Lazdunski, Trypsin-pancreatic trypsin inhibitor association. Dynamics of the interaction and role of disulfide bridges. Biochemistry 11 (1972) 2967–2977.
[102] L. Kiczak, M. Kasztura, K. Koscielska-Kasprzak, M. Dadlez, J. Otlewski, Selection of potent chymotrypsin and elastase inhibitors from M13 phage library of basic pancreatic trypsin inhibitor (BPTI). Biochim. Biophys. Acta 1550 (2001) 153–163.
[103] M.S. Dennis, R.A. Lazarus, Kunitz domain inhibitors of tissue factor-factor VIIa. I. Potent inhibitors selected from libraries by phage display. J. Biol. Chem. 269 (1994) 22129–22136.
[104] M.S. Dennis, A. Herzka, R.A. Lazarus, Potent and selective Kunitz domain inhibitors of plasma kallikrein designed by phage display. J. Biol. Chem. 270 (1995) 25411–25417.
[105] A.S. Tanaka, M.M. Silva, R.J. Torquato, M.A. Noguti, C.A. Sampaio, H. Fritz, E.A. Auerswald, Functional phage display of leech-derived tryptase inhibitor (LDTI): construction of a library and selection of thrombin inhibitors. FEBS Lett. 458 (1999) 11–16.
[106] I.T. Campos, M.M. Silva, S.S. Azzolini, A.F. Souza, C.A. Sampaio, H. Fritz, A.S. Tanaka, Evaluation of phage display system and leech-derived tryptase inhibitor as a tool for understanding the serine proteinase specificities. Arch. Biochem. Biophys. 425 (2004) 87–94.
[107] C.H. Chung, H.E. Ives, S. Almeda, A.L. Goldberg, Purification from *Escherichia coli* of a periplasmic protein that is a potent inhibitor of pancreatic proteases. J. Biol. Chem. 258 (1983) 11032–11038.
[108] M.E. McGrath, S.A. Gillmor, R.J. Fletterick, Ecotin: lessons on survival in a protease-filled world. Protein Sci. 4 (1995) 141–148.
[109] S.A. Gillmor, T. Takeuchi, S.Q. Yang, C.S. Craik, R.J. Fletterick, Compromise and accommodation in ecotin, a dimeric macromolecular inhibitor of serine proteases. J. Mol. Biol. 299 (2000) 993–1003.
[110] C.I. Wang, Q. Yang, C.S. Craik, Isolation of a high affinity inhibitor of urokinase-type plasminogen activator by phage display of ecotin. J. Biol. Chem. 270 (1995) 12250–12256.
[111] M.C. Laboissiere, M.M. Young, R.G. Pinho, S. Todd, R.J. Fletterick, I. Kuntz, C.S. Craik, Computer-assisted mutagenesis of ecotin to engineer its secondary binding site for urokinase inhibition. J. Biol. Chem. 277 (2002) 26623–26631.
[112] S.Q. Yang, C.I. Wang, S.A. Gillmor, R.J. Fletterick, C.S. Craik, Ecotin: a serine protease inhibitor with two distinct and interacting binding sites. J. Mol. Biol. 279 (1998) 945–957.
[113] B. Breugelmans, G. Simonet, V. van Hoef, S. Van Soest, J. Vanden Broeck, Pacifastin-related peptides: structural and functional characteristics of a family of serine peptidase inhibitors. Peptides 30 (2009) 622–632.
[114] A. Patthy, S. Amir, Z. Malik, A. Bodi, J. Kardos, B. Asboth, L. Graf, Remarkable phylum selectivity of a *Schistocerca gregaria* trypsin inhibitor: the possible role of enzyme–inhibitor flexibility. Arch. Biochem. Biophys. 398 (2002) 179–187.
[115] K. Fodor, V. Harmat, C. Hetenyi, J. Kardos, J. Antal, A. Perczel, A. Patthy, G. Katona, L. Graf, Extended intermolecular interactions in a serine protease–canonical inhibitor complex account for strong and highly specific inhibition. J. Mol. Biol. 350 (2005) 156–169.
[116] B. Szenthe, A. Patthy, Z. Gaspari, A.K. Kekesi, L. Graf, G. Pal, When the surface tells what lies beneath: combinatorial phage-display mutagenesis reveals complex networks of surface–core interactions in the pacifastin protease inhibitor family. J. Mol. Biol. 370 (2007) 63–79.
[117] K. Ylinenjarvi, M. Widersten, I. Bjork, Hydrophobic sequences can substitute for the wild-type N-terminal sequence of cystatin A (stefin A) in tight binding to cysteine proteinases selection of high-affinity N-terminal region variants by phage display. Eur. J. Biochem. 261 (1999) 682–688.
[118] F. Wirsching, M. Keller, C. Hildmann, D. Riester, A. Schwienhorst, Directed evolution towards protease-resistant hirudin variants. Mol. Genet. Metab. 80 (2003) 451–462.
[119] D. Matthews, M.D. Ballinger, Substrate phage display. in: T. Clackson, H.B. Lowman (Eds.), Phage Display: a Practical Approach. Oxford University Press, 2004, pp. 117–133.
[120] F.K. Kerr, A.R. Thomas, L.C. Wijeyewickrema, J.C. Whisstock, S.E. Boyd, D. Kaiserman, A.Y. Matthews, P.I. Bird, N.M. Thielens, V. Rossi, R.N. Pike, Elucidation of the substrate specificity of the MASP-2 protease of the lectin complement pathway and identification of the enzyme as a major physiological target of the serpin, C1-inhibitor. Mol. Immunol. 45 (2008) 670–677.
[121] F.K. Kerr, G. O'Brien, N.S. Quinsey, J.C. Whisstock, S. Boyd, M.G. de la Banda, D. Kaiserman, A.Y. Matthews, P.I. Bird, R.N. Pike, Elucidation of the substrate specificity of the C1s protease of the classical complement pathway. J. Biol. Chem. 280 (2005) 39510–39514.
[122] H.X. Li, B.Y. Hwang, G. Laxmikanthan, S.I. Blaber, M. Blaber, P.A. Golubkov, P. Ren, B.L. Iverson, G. Georgiou, Substrate specificity of human kallikreins 1 and 6 determined by phage display. Protein Sci. 17 (2008) 664–672.
[123] H.J. Hsu, K.C. Tsai, Y.K. Sun, H.J. Chang, Y.J. Huang, H.M. Yu, C.H. Lin, S.S. Mao, A.S. Yang, Factor Xa active site substrate specificity with substrate phage display and computational molecular modeling. J. Biol. Chem. 283 (2008) 12343–12353.
[124] F.R. Melo, M.O. Mello, O.L. Franco, D.J. Rigden, L.V. Mello, A.M. Genu, M.C. Silva-Filho, S. Gleddie, M.F. Grossi-de-Sa, Use of phage display to select novel cystatins specific for *Acanthoscelides obtectus* cysteine proteinases. Biochim. Biophys. Acta 1651 (2003) 146–152.
[125] L.R. Ceci, M. Volpicella, Y. Rahbe, R. Gallerani, J. Beekwilder, M.A. Jongsma, Selection by phage display of a variant mustard trypsin inhibitor toxic against aphids. Plant J. 33 (2003) 557–566.
[126] A. Fernandez-Gacio, M. Uguen, J. Fastrez, Phage display as a tool for the directed evolution of enzymes. Trends Biotechnol. 21 (2003) 408–414.
[127] M. Paschke, Phage display systems and their applications. Appl. Microbiol. Biotechnol. 70 (2006) 2–11.
[128] S.S. Sidhu, S. Koide, Phage display for engineering and analyzing protein interaction interfaces. Curr. Opin. Struct. Biol. 17 (2007) 481–487.
[129] W.G. Willats, Phage display: practicalities and prospects. Plant Mol. Biol. 50 (2002) 837–854.
[130] G.P. Smith, V.A. Petrenko, Phage display. Chem. Rev. 97 (1997) 391–410.
[131] T. Bratkovic, Progress in phage display: evolution of the technique and its applications. Cell. Mol. Life Sci. (2009).
[132] S. Pfaffen, T. Hemmerle, M. Weber, D. Neri, Isolation and characterization of human monoclonal antibodies specific to MMP-1A, MMP-2 and MMP-3. Exp. Cell. Res. 316 (2010) 836–847.
[133] C.J. Farady, J. Sun, M.R. Darragh, S.M. Miller, C.S. Craik, The mechanism of inhibition of antibody-based inhibitors of membrane-type serine protease 1 (MT-SP1). J. Mol. Biol. 369 (2007) 1041–1051.
[134] C.J. Farady, P.F. Egea, E.L. Schneider, M.R. Darragh, C.S. Craik, Structure of an Fab–protease complex reveals a highly specific non-canonical mechanism of inhibition. J. Mol. Biol. 380 (2008) 351–360.
[135] J. Thanongsaksrikul, P. Srimanote, S. Maneewatch, K. Choowongkomon, P. Tapchaisri, S.I. Makino, H. Kurazono, W. Chaicumpa, A VHH that neutralizes the zinc-metalloproteinase activity of botulinum neurotoxin type A. J. Biol. Chem. (2010).
[136] L. Kiczak, K. Koscielska, J. Otlewski, M. Czerwinski, M. Dadlez, Phage display selection of P1 mutants of BPTI directed against five different serine proteinases. Biol. Chem. 380 (1999) 101–105.
[137] A.J. Scheidig, T.R. Hynes, L.A. Pelletier, J.A. Wells, A.A. Kossiakoff, Crystal structures of bovine chymotrypsin and trypsin complexed to the inhibitor domain of Alzheimer's amyloid beta-protein precursor (APPI) and basic pancreatic trypsin inhibitor (BPTI): engineering of inhibitors with altered specificities. Protein Sci. 6 (1997) 1806–1824.
[138] N. Dimasi, F. Martin, C. Volpari, M. Brunetti, G. Biasiol, S. Altamura, R. Cortese, R. De Francesco, C. Steinkuhler, M. Sollazzo, Characterization of engineered hepatitis C virus NS3 protease inhibitors affinity selected from human pancreatic secretory trypsin inhibitor and minibody repertoires. J. Virol. 71 (1997) 7461–7469.
[139] A.S. Tanaka, C.A. Sampaio, H. Fritz, E.A. Auerswald, Functional display and expression of chicken cystatin using a phagemid system. Biochem. Biophys. Res. Commun. 214 (1995) 389–395.
[140] J.G. Bieth, Theoretical and practical aspects of proteinase inhibition kinetics. Methods Enzymol. 248 (1995) 59–84.