Proteases as drug targets

Andy J.P. Docherty, Tom Crabbe, James P. O’Connell and Colin R. Groom

Celltech R&D, Slough SL1 4EN, U.K.

Abstract

The effective management of AIDS with HIV protease inhibitors, or the use of angiotensin-converting enzyme inhibitors to treat hypertension, indicates that proteases do make good drug targets. On the other hand, matrix metalloproteinase (MMP) inhibitors from several companies have failed in both cancer and rheumatoid arthritis clinical trials. Mindful of the MMP inhibitor experience, this chapter explores how tractable proteases are as drug targets from a chemistry perspective. It examines the recent success of other classes of drug for the treatment of rheumatoid arthritis, and highlights the need to consider where putative targets lie on pathophysiological pathways – regardless of what kind of therapeutic entity would be required to target them. With genome research yielding many possible new drug targets, it explores the likelihood of discovering proteolytic enzymes that are causally responsible for disease processes and that might therefore make better targets, especially if they lead to the development of drugs that can be administered orally. It also considers the impact that biologics are having on drug discovery, and in particular whether biologically derived therapeutics such as antibodies are likely to significantly alter the way we view proteases as targets and the methods used to discover therapeutic inhibitors.

Introduction

Sales of AstraZeneca’s angiotensin-converting enzyme (ACE) inhibitor Zestril in the year 2000 totalled $1228 million. In the same year Roche and Abbott reported sales of $303 million and $210 million respectively for their HIV protease inhibitors Viracept and Kaletra. Clearly, proteolytic enzymes can be profitable drug targets. Furthermore, the speed with which the HIV inhibitors were developed indicates that proteases can be highly tractable for chemists seeking to make orally active inhibitors. Discovery of the HIV pro-

*To whom correspondence should be addressed (e-mail adocherty@celltech.co.uk).
tease in 1985 resulted from analysis of the viral nucleotide sequence [1]. Extensive analysis of its catalytic properties followed which, when combined with an early knowledge of its crystal structure, facilitated the rapid development of highly potent and selective inhibitors [2]. By 1997, following demonstration that the inhibitors could reduce viral load in patients, several such inhibitors received approval from the Food and Drug Administration in the U.S.A. for anti-viral use in humans. This example is significant not only because of the brief 12 years that it took to go from target discovery to licensed product, but also because it is an example of the way applied genomics can first reveal the existence of possible new drug targets. One might therefore expect the vast quantity of genome sequence information available to us today to be a very rich hunting ground for further protease drug targets. As Alan Barrett and his colleagues have pointed out, analysis of all the available genome sequences suggests that some 2% encode identifiable peptidase motifs [3]. First, though, let us examine what it is that makes proteolytic enzymes so tractable as drug targets.

**Tractability**

Proteolytic enzyme active sites are often composed of a deep pocket. This increases the likelihood of small drug-like molecules that mimic the substrate being able to have many possible points of interaction within the pocket. This facilitates tight binding. A variety of methods for discovering short peptide substrates for a particular protease are available, even if the physiological substrate of the protease is a very large protein, or if the true physiological substrate is unknown [4]. The discovery of good substrates is an essential first step, because it allows the development of assays suitable for screening for inhibitors. The susceptibility of such substrates to cleavage, and information derived from the subsequent kinetic analysis, helps define what structures are most readily accommodated within a binding pocket. Early acquisition of this information helps chemists to understand what is required to build enzyme specificity into inhibitor structures – an important consideration when working within large protease families where the therapeutic benefit is expected to derive from inhibiting one or some, but not all, of the family members. Knowledge of substrates derived in this way is an excellent starting point for the rational design of inhibitors that mimic substrates in their binding characteristics, but which have been modified in order to avoid cleavage. This rational approach may reduce the need for large-scale screening for inhibitor leads at the early stages of a drug discovery programme. This is an important point, because the early discovery of a series of leads for which a structure–activity relationship can be derived can be a major hurdle for some other classes of drug target. Examples include receptors for large protein ligands, or kinases where the structure of the various substrates within the active site are all very similar. Another feature that helps make proteases tractable drug targets is the fact that they are very often amenable to structure determination, and crystallographers have a very good track record for solving protease catalytic domain structures at high atomic resolution [5]. Combined with an ability to co-crystallize lead inhibitors, or to ‘soak’ inhibitors into pre-existing protease crystals, the result-
ing structural information further facilitates rational inhibitor development. In the case of HIV protease, crystal structures were available very soon after the discovery of the enzyme, and the resulting information impacted on inhibitor design from the very early stages of the project ([6] and references therein). The early acquisition of such information is a trend that is set to increase with the high-throughput crystallography and structure determination methods that are being brought to the fore today [7,8].

Proteolytic enzymes are clearly tractable in terms of lead generation, but the very fact that the leads are often peptide-derived can make their further development a challenge. For ease of use, it is clearly desirable for a drug to be administered orally. This requires that it should be absorbed through the gut and be made bioavailable to the tissues in which it is to act for sufficient periods of time to provide therapeutic benefit. Guidelines for good oral bioavailability have been described by Lipinski et al. [9] and are summarized in Table 1. Generally speaking, compounds that fall within these guidelines compete with ligands or substrates for binding to structurally defined pockets. Proteolytic enzymes usually have such pockets, and they are therefore often referred to as being ‘druggable’ targets. Unfortunately, inhibitor leads based on peptides tend to be highly susceptible to proteolysis when administered in vivo, especially if administered orally. In order to turn a peptide-based inhibitor into an orally bioavailable drug that is not rapidly metabolized and cleared, at the very least it is generally necessary to decrease the $M_r$ and remove most, if not all, of the peptide bonds. This can lead to loss of the potency and selectivity characteristics on the basis of which the compound was first selected. With this in mind, the following section examines just how much effort the pharmaceutical industry has already devoted to protease targets, and the frequency of success.

Success to date

Within chemistry-based drug discovery research, the total number of individual drug targets across the industry (excluding non-human targets, for example in the anti-infectives field) has been estimated to be approx. 480 [10,11]. These targets fall into 130 protein families, of which we estimate that 10% are peptidases [12]. This indicates that a significant amount of the pharmaceutical industry’s research effort is devoted to protease targets. Furthermore, this effort appears to be highly productive in terms of yielding compounds that meet the Lipinski guidelines for orally active compounds [9]. For example,

Table 1 Guidelines for oral bioavailability. These guidelines are based on [9]. Note that compounds derived from natural products, or classes that are substrates for biological transporters, will not follow the listed guidelines.

<table>
<thead>
<tr>
<th>Poor absorption or permeation of a compound is likely when it has:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;5 hydrogen-bond donors (defined as NH or OH)</td>
<td></td>
</tr>
<tr>
<td>&gt;10 hydrogen-bond acceptors (defined as O or N atoms)</td>
<td></td>
</tr>
<tr>
<td>$M_r$&gt;500</td>
<td></td>
</tr>
<tr>
<td>High lipophilicity: (clogP&gt;5 or MlogP&gt;4.15)</td>
<td></td>
</tr>
</tbody>
</table>
analysis of the industry portfolio of targets reveals that approx. 400 of them 
have a corresponding drug candidate, as defined by Lipinski et al. [9], that has 
an IC₅₀ in a relevant assay of 10 μM or less [12]. Again, proteases are well re-
presented. For example, if one just takes endopeptidases, we estimate that there 
are about 14 metallo-, 12 serine, six cysteine and five aspartate endopeptidase 
drug targets for which a potentially bioavailable ‘Lipinski-like’ compound 
exists. This suggests that approx. 10% of all the ‘Lipinski-compliant’ targets are 
in fact proteases. However, the sobering fact is that, despite this large amount 
of research effort, and apparent successful lead generation, only a few of the 
120 or so marketed classes of drug target proteases – HIV protease, ACE, 
thrombin and a few others. Clearly, ‘druggable’ does not imply ‘good drug tar-
et’. Other parameters obviously come into play, not least the issue of whether 
a putative protease target is actually causal in a disease process, and whether its 
point of activity in a pathway is amenable to therapeutic intervention. These 
points are considered in the following case study.

Targets in rheumatoid arthritis (RA)

Recent years have seen the introduction of new biological drugs that have 
dramatically changed the treatment of RA and other inflammatory diseases such 
as Crohn’s disease. The two leading classes of drug both target the inflammatory 
mediator tumour necrosis factor α (TNFα), either with a neutralizing antibody 
or with a soluble form of the natural receptor [13]. They are the first true dis-
eesase-modifying drugs for this indication, which, despite intense research 
initiatives, had represented a very large unmet medical need. It is worth examin-
ing why this particular target has proven to be such an excellent point for 
therapeutic intervention compared with other targets, including several pro-
teases that, over many years, have received a similar level of attention. The 
answer is largely due to where TNFα lies on the pathophysiological pathways 
that lead to RA. As shown in Figure 1, several features of the RA-affected joint 
represent points of intervention where one would anticipate therapeutic benefit.

Several classes of endopeptidase thought to be responsible for the thin-
ning of cartilage and the erosion of bone have received attention from the 
pharmaceutical industry. Much attention has been focused on the matrix metal-
loproteinases (MMPs), and in particular MMP-1 and MMP-3. These enzymes 
are thought to be responsible for the turnover of collagen and proteoglycan 
respectively (both of which are essential structural components of the cartilage 
extracellular matrix), and received much of the early interest [14–16]. In the 
case of MMP-3, despite strong evidence for its pathophysiological involvement 
in cartilage breakdown, MMP-3⁻/⁻ mice proved to be just as susceptible to car-
tilage loss as wild-type mice [17]. In the case of MMP-1, the potent orally 
bioavailable inhibitor Trocade (Kᵢ, ~3 nM; in vitro IC₅₀ ~60 nM), developed by 
Roche, was found to demonstrate good efficacy in rodent disease models and 
was well tolerated in human Phase 1 and 2 studies, but was ultimately discon-
tinued because of an ‘unfavourable risk/benefit ratio’ ([15,18] and Roche 
Holding AG news release). This is generally interpreted to mean either that no 
benefit was observed or that administration of the drug made matters worse.
Despite these setbacks, protease inhibitor development continues, focused largely on more recently discovered members of the MMP family such as MMP-13 and various ADAMTS (ADAM with thrombospondin motifs, where ADAM denotes a disintegrin and metalloproteinase) family members that recent studies have shown to be capable of cleaving aggrecan – the polypeptide backbone of proteoglycan [19]. Given the large repertoire of MMPs and ADAMTSs that we now know to be capable of participating in cartilage remodelling, the challenge is to identify the principal contributors in disease and to make appropriately selective inhibitors. Inhibiting the ‘wrong’ proteases will probably be therapeutically counterproductive, because it has already been shown that knockouts of several MMPs lead to disabilitating phenotypes that one would not wish to recapitulate in humans [20–22]. In addition, there are difficulties with modelling such scenarios in rodent species, because the repertoire of MMPs used by mice, for example, appears to differ from that in humans. On the other hand, redundancy in which proteases are used in particular physiological settings may require that broad-spectrum inhibitors are developed. For example, aggrecan is susceptible to cleavage at various positions by a great many MMPs and ADAMTSs, and it may be that several enzymes contribute to pathological remodelling [19]. However, the use of broad-spectrum inhibitors runs against the apparent need for at least some degree of selectivity to avoid inhibition of normal physiological processes [20,23,24]. Indications that this may present a problem derive from the many broad-spectrum MMP inhibitors that have been evaluated in cancer trials, many of which have suffered from unacceptable side-effect profiles, as well as a lack of efficacy [25].

Taking a step back from the proteolytic enzymes responsible for cartilage and bone resorption, it is apparent that a great many soluble mediators orchestrate the cellular events that occur within the joint in RA. This includes
proliferation and vascularization of the synovium, inflammatory cell infiltra-

tion, and the consequential activation of bone, cartilage and synovial cells and 
release of further inflammatory mediators, proteolytic enzymes and their activ-
ators. Known as the cytokine network, these mediators include not only the 
so-called ‘destructive’ cytokines such as TNFα, interleukin-1β (IL-1β) and IL-
17 (among others), but also ‘modulatory’ cytokines such as IL-10 and growth 

factors such as transforming growth factor β [26]. Genome mining to find fur-
ther cytokines, for example IL-10 family members, has been very successful, 
and has yielded several that are currently being explored for their possible roles 
in inflammatory disease [27–29]. Several cytokines, such as IL-1β, IL-6, IL-15 

and IL-17, are already being pursued as drug targets [30–32]. As extracellular 
gene products, they are all amenable to evaluation as targets in vivo with anti-

bodies that specifically modulate activity by binding to the ligand, or blocking 
the corresponding receptor. In addition, innovations in the technology sur-
rounding the ‘humanization’ and manufacturing of antibodies has led to their 
therapeutic use in a variety of disease settings [33–35]. Innovations whereby 
cytokine receptors can be manufactured as soluble proteins fused to 
immunoglobulin Fc domains to give them long circulating half-lives in vivo has 
similarly led to their application as both proof-of-concept reagents and thera-
peutics for use in humans [36]. As a consequence, although TNFα is not in 
itself tractable as a traditional chemistry target, it has proven to be an excellent 
target for biologics.

The two leading products are Remicade, an anti-TNFα antibody from 
Centocor that is licensed for use in RA and Crohn’s disease, and Enbrel, a p75-
TNF receptor–Fc from Immunex, which is licensed for the treatment of RA and 
psoriatic RA. Sales of these drugs reported for the first three-quarters of 2002 
were $918 million and $567 million respectively. The established pharmaceuti-
cal industry therefore cannot ignore the impact of biologics in large, chronic 
indications where orally absorbed chemically derived products were previously 
thought to be the only kind of product that would meet with patient compli-
ance. In fact, one might have expected that the cost and intravenous 
administration required for Remicade and Enbrel might mitigate against them 
being highly successful drugs. It turns out, however, that patients given anti-
TNF treatment benefit from an immediate anti-inflammatory effect, as well as 
longer-term disease-modifying benefits [37]. The success of Remicade and 
Enbrel is therefore due not only to the lack of any previously existing disease-
modifying treatments for RA, but also to the fact that they deliver a clinically 
demonstrable benefit to patients very soon after their first administration. 
Although there are good reasons to believe that inhibitors of some of the prote-
olytic enzymes described above may have helped to maintain bone and cartilage 
integrity, it is unlikely that they would have the same profound anti-inflamma-
tory properties of the anti-TNFα treatments. In consequence, the clinical trial 
design for Trocade required much more complex read-outs for efficacy, involv-
ing the use of MRI (magnetic resonance imaging) to measure changes to bone 
and cartilage integrity over long periods of time. TNFα therefore turns out to be 
a far superior target for therapeutic intervention because of its upstream posi-
tion in the disease process and the fact that (probably) small changes in levels of
its activity are sufficient to change the balance of cytokines, leading to a readily measurable clinical benefit (Figure 2). Nonetheless, this approach has required the development of brand new classes of therapeutic agent.

One obvious outcome of the success with anti-TNFα agents is that cytokine-mediated pathways, in particular the TNFα pathway, are now considered to be validated in terms of therapeutic intervention for inflammatory disease in humans. The whole pathway has therefore been examined for possible additional points of intervention, and proteolytic enzymes feature highly [13]. Instead of focusing on the matrix-degrading proteases induced downstream by TNFα and other inflammatory mediators such as IL-1β, a search for chemistry targets has focused on the proteolytic events that give rise to these mediators. Interleukin-1β-converting enzyme (ICE) is a cysteine endopeptidase belonging to the caspase family, and is responsible for the intracellular processing of an IL-1β precursor prior to secretion of the cytokine in active form [38]. TNFα-converting enzyme (TACE) is a member of the adalysin family of metallo-endopeptidases, and is responsible for the cleavage of transmembrane-domain-anchored TNFα, leading to its rapid shedding from inflammatory cells following acute inflammatory

![Diagram of the TNFα pathway](image_url)

**Figure 2 Targets within the TNFα pathway.** The clinical success of antibodies or soluble receptors that modulate the activity of TNFα highlights several possible protease targets on the same pathway. The inhibition of downstream targets such as proteases involved in cartilage destruction may have less impact than the inhibition of targets that are additionally involved in inflammatory processes. ICE and TACE are the subject of inhibitor development programmes, but currently inhibitors aimed at the non-protease targets type IV phosphodiesterase (PDE4) and p38 MAP kinase (p38 MAPK) appear to be more advanced. Points of intervention more readily targeted by biologics are also apparent, and include blockers of the receptors for TNF (TNFR) or IL-1β (IL-1R), or IL-1β itself, for which a drug (Anakinra, a recombinant IL-1β receptor antagonist) has recently been launched [74].
Since being cloned by a number of pharmaceutical companies, both ICE and TACE have been pursued as drug targets. Several groups have reported ICE inhibitors, but only Vertex Pharmaceuticals appear to have a compound in clinical development for RA ([40] and company news release). TACE inhibitors have emerged from the various series of compounds developed originally for the MMPs (structural similarities place TACE and the MMPs within the metzincin family), but to date none appear to have entered clinical development [41]. As targets for synthetic orally bioavailable chemistry programmes, both ICE and TACE present significant challenges in terms of selectivity if inappropriate inhibition of related family members is to be avoided. Furthermore, the relative contributions in disease of cell-bound compared with shed TNFα are not entirely clear, and what could be achieved therapeutically with a TACE inhibitor that only prevents shedding is hard to predict. In the case of ICE, it is not clear how critical its activity is to the occurrence of IL-1β in the RA joint. However, it does seem likely that the therapeutic use of neutralizing antibodies, natural antagonists or soluble receptors significantly lowers the levels of cytokine in the inflamed joint. Whether a similar decrease can be achieved with ICE or TACE inhibitors is an issue that has yet to be addressed.

In fact, the most promising targets on the TNFα pathway for chemically derived drugs appear not to be proteases (Figure 2). Upstream on the pathway, orally active inhibitors of type IV phosphodiesterase are potent inhibitors of TNFα release from lipopolysaccharide-stimulated human macrophages, and a number of companies have entered compounds into clinical development for the treatment of asthma [42]. Elsewhere on the pathway, orally active inhibitors of p38 mitogen-activated protein kinase (MAP kinase) are looking promising in a variety of inflammation models [43,44]. Such inhibitors appear to act upstream to prevent synthesis of TNFα (and IL-1β) following inflammatory cell stimulation. They also have some downstream activity on p38 MAP kinase-dependent signalling pathways that operate when TNFα or IL-1β bind to their receptors, although other pathways also exist at this point [45]. Many pharmaceutical companies are very active in this area, including Vertex Pharmaceuticals, Boehringer Ingelheim Pharmaceuticals and Scios Inc., who all have compounds in clinical development for RA (company news releases).

The following section considers how one might apply what has been learnt about ‘good’ targets in the RA field to a better understanding of how we should mine genome sequence information for new proteolytic enzyme targets.

**Genome mining**

The human genome is estimated to encode around 30000 gene products [46,47]. If we assume that where the protein corresponding to one member of a gene family has been proven to bind a drug-like molecule, other members will also be able to do so, we can estimate the number of ‘druggable’ targets. Through homology searches and analysis of gene family sizes, it has been estimated that, using this definition, some 3000 gene products are likely to be druggable (Figure 3) [12]. However, only those that play causal roles in disease processes, or which when modulated provide symptomatic benefit, can become
viable drug targets. Estimates of the number of disease-related genes vary widely depending on whether the number is based solely on expression data in disease tissue or on more substantial data such as evidence from knockout experiments. Estimates range between 3000 and 10,000 [48–50]. How many causally involved gene products fall into the druggable gene product camp is even less clear, but one estimate based on the number of anti-fungal targets arising from the yeast genome puts this as low as 2–5%, which is in broad agreement with data derived from knockout experiments [12,51] (Figure 3). This indicates that the number of useful new drug targets, as defined by the traditional requirements of chemistry programmes aimed at orally bioavailable drugs, may be a rather modest several hundred.

However, this overlooks the fact that the 30,000 or so predicted gene products do not represent all the possible targets. Splice variation, post-translational modification, variation in glycosylation, functional synapses composed of multimeric protein complexes, and non-catalytic domains in the case of proteases, very often exist. If these are considered to be potential molecular targets, the total number of such targets becomes greater than the number of gene products (Figure 3). Most of these targets are likely to be intractable to traditional chemistry strategies. However, a number of them may make excellent starting points for the development of biologics in the form of therapeutic proteins, soluble receptor-based drugs, antibody therapeutics, or components of therapeutic strategies based on gene therapy, vaccination or antisense technology, etc. The success of Remicade and Enbrel tells us that at least some of these therapeutic approaches are eminently feasible.

**Figure 3 How many new drug targets will genomics generate?**

Intersection A indicates those targets that may be exploited by chemistry, while intersection B indicates those targets that may be exploited by biologics. The latter category makes available targets that have not traditionally been recognized to be suitable for chemistry.

© 2003 Biochemical Society
One feature of genome research is the speed and scale at which it identifies potential new drug targets. For example, following the discovery that TACE, the protease responsible for shedding TNFα, was an adamalysin and consisted of several domains, including a disintegrin and metallo-endopeptidase (metzincin) domain, intense genome mining rapidly extended the known members of this family to 30 or more [52–54]. Further genome mining subsequently uncovered the ADAMTS family [55]. Rapidly increasing in size, this family includes some enzymes that cleave aggrecan [56]. The striking point is that genome analysis revealed the existence of many new family members, and expression was demonstrated in a variety of tissues, long before most of them (with a few exceptions [57,58]) had been evaluated in ‘wet’ experiments as proteases. Indeed genetics, for example through the use of gene knockouts, may indicate which of the candidate proteases is responsible for a phenotype before physiologically meaningful enzyme assay results are available [51]. In fact, in the case of ADAM-33, genetics revealed an association with asthma in humans only months after the gene was first reported [59,60]. Such technological innovations are clearly very powerful. Nonetheless, it is often the case that the significance of a particular phenotype in a knockout mouse is hard to relate to disease in the human adult. Furthermore, expression profiling to reveal the existence of a new protease in disease tissues only demonstrates a disease association. It does not demonstrate that the protease is causal in the disease process, nor that the protease is tractable for drug discovery, or that its activity represents a feasible point for therapeutic intervention. There is, therefore, a need for powerful in vivo target validation strategies. If such strategies are to be dependent on selective bioavailable compounds, it would be desirable to have the means to generate compounds that modulate the activity of a relatively novel protease before detailed knowledge of its biochemical properties exists. Furthermore, given the size of most proteolytic enzyme families and the structural similarities within active sites, the need for inhibitor selectivity cannot be ignored, and the means to obtain such selectivity needs to be sought. These are difficult challenges for early chemistry programmes. In contrast, antibodies can make excellent target validation reagents, although not for cellular targets, where RNA-mediated interference technology may have a big role to play [61]. Furthermore, thanks to great progress in the ways in which antibodies can be engineered and manufactured cost-effectively (for example to remove inessential Fc domains or to create small monovalent Fabs expressed in *Escherichia coli* [62–64]), there are encouraging signs that they are likely to be useful for the long-term treatment of chronic diseases [65]. Could proteolytic enzymes be targets for inhibitory monovalent Fabs?

**Proteolytic enzymes as antibody drug targets?**

A clear limitation in the use of antibodies is the requirement for extracellular targets. With that exception, they bring potential opportunities for inhibition that do not arise readily from the use of chemically derived inhibitors. For example, the high-affinity, tight-binding and exquisite specificity properties that an antibody can possess arise from multiple interactions
over a relatively large surface area [66,67]. They do not share a dependency on the kind of deep binding pocket that would typically be recognized as necessary for low-molecular-mass inhibitor compounds. The fact that it has been possible to isolate an inhibitory single-domain camelid antibody with a $K_i$ towards $\alpha$-amylase of 10 nM which, although binding to the catalytic site, did not contact the catalytic residues, suggests that useful inhibitory antibodies directed against proteases should be possible [68]. Many proteolytic enzymes consist of multi-domain structures in which the catalytic domain is just one of many. Several examples exist where it has been shown that specificity for the natural (generally, in this case, high-molecular-mass) substrate is dependent on the involvement of an additional domain to the one that contributes the catalytic mechanism. Examples include the C-terminal domain of MMP-1 required for collagen cleavage, and the thrombospondin-like motif in ADAMTS-4 required for aggrecanase cleavage [69,70]. It is probable that antibodies directed against these domains would act as functional inhibitors of enzyme activity. In fact, recent new high-throughput ways in which antibodies are isolated mean that it is now feasible to consider screening for inhibitory antibodies from the entire immune repertoire of an immunized animal. The selected lymphocyte antibody method (SLAM) means that all that is required is an assay for detecting such functional antibodies [71]. The immune repertoire of an immunized animal, in the form of B cells plated out in microtitre wells, is screened and, following the identification of wells containing the required inhibitory activity, DNA encoding the corresponding antibody variable regions are recovered by PCR and expressed recombinantly. An important feature of this approach is that a precise knowledge of which part of a novel protease contributes to the proteolytic property one wishes to inhibit is not essential. In fact, the resulting antibodies can be used subsequently as powerful reagents to define the molecular components of the protease that contribute to its function both in vitro and in vivo.

New chemistry strategies for inhibitor discovery

In addition to the breakthroughs in antibody technology, developments in chemistry also promise to enhance the drug discovery process. As with SLAM, developments in what has been called ‘chemical genomics’ also offer the prospect of discovering protease inhibitor leads without a great deal of biochemical characterization of the target, or a bias towards a conventional substrate-binding pocket. In a strategy called the automated ligand identification system (ALIS), pools of combinatorially synthesized compounds are screened for their ability to bind to a target protein of interest [72]. Compounds that bind tightly (depending on the stringency of the conditions used) are separated from the rest of the pool by size-exclusion chromatography and subsequently identified by mass spectrometry. Tight-binding compounds are only then tested for their inhibitory and selectivity properties. Another strategy, based on identifying individual compounds among libraries arrayed on to a solid surface on the basis of their binding to a protein of interest, has also been described [73]. If applied to protease targets, this approach opens up
the prospect of discovering novel leads without a detailed knowledge of the enzyme’s substrate specificity pocket. Furthermore, it offers the possibility of identifying leads that act through binding to alternative sites, such as allosteric sites or non-catalytic exosites. New classes of inhibitor that act though such sites may provide routes to more selective inhibitors than has sometimes been possible by targeting the substrate-binding pocket with competitive antagonists.

Conclusions

Experience shows that proteolytic enzymes are tractable targets for lead discovery and are ‘druggable’, in that they can give rise to orally active drugs. Thanks to readily identifiable peptidase motifs, a multitude of potential protease targets have arisen from genome mining. Some of the questions that we would ask of such a new target are summarized in Table 2. Clearly, generating data to show whether a particular protease is truly causal in a disease process is not easy. In some cases it is apparent that their presence in disease tissue is a consequence of the disease rather than a significant cause. This may be the case with MMPs found in the presence of the stroma around proliferating tumour cells in cancer tissues. In inflammatory diseases such as RA, targets other than matrix-degrading proteases have clearly been shown to represent superior points for intervention. This in turn has required the development of non-chemically derived biologics as the therapeutic agent of choice – in this case for neutralizing the activity of TNFα. In the near future, a significant number of new ‘post-genomic’ targets may not in fact be tractable for chemistry-based programmes, and will instead be reliant on the application of biologics technology for target validation and, very likely, for drug development. Close scrutiny of how well such drugs work, and the pathways in which they act, should help highlight better chemistry targets, including proteases. Armed with some of the new methods for identifying novel leads, such proteolytic enzymes should be suitable targets for orally administered drugs.

Table 2 Proteolytic enzyme target checklist.

| Is it a druggable target? | ✓ |
| Is it at a critical point on a pathophysiological pathway? | ✓ |
| Is it of a proven therapeutic class? | ✓ |
| Peptide hormone release | ✓ |
| Viral processing enzymes | ✓ |
| Matrix remodelling | ✓ |
| Amyloid processing | ✓ |
| Cell surface shedding or activation events | ✓ |
| Intracellular processing events | ✓ |
| Are there other potential drug/antibody binding sites? | ✓ |
| Would non-chemical intervention be feasible? | ✓ |
| Does its cleavage product represent a surrogate for therapeutic efficacy? | ✓ |

© 2003 Biochemical Society
We thank Alan Barrett at the Wellcome Trust Sanger Institute for the MEROPS database – a resource that has facilitated many of our protease mining expeditions. We also gratefully acknowledge our colleagues within the Chemistry, Biology, Pharmacology and Clinical Departments at Celltech, Gill Murphy at the Cambridge Institute for Medical Research, Vera Knäuper at the University of York and Vim Van den Berg at the University Medical Centre in Nijmegen for the many fruitful discussions that culminated in this chapter.

References

© 2003 Biochemical Society