

treatment are self-regulatory, through the activation of CD4⁺CD25⁺ lymphocytes, which inhibit the activity of tumor-specific CD8⁺ T lymphocytes and rapidly attenuated the response [5]. The generation of autoimmunity has not proven an overwhelming barrier to the deployment of molecular chaperone-based immunotherapy in general, although further studies are required to investigate the potential complications of therapy [10,21].

Conclusions

Overall therefore, this novel approach to the treatment of malignant melanoma may herald new ways in which antigen presenting cells can be stimulated and anti-tumor immunity generated (5). Although much is yet to be learned regarding the mechanisms of Hsp70 in tumor immunotherapy, these and other recent studies also illustrate the versatility of molecular chaperones in stimulating multiple aspects of the anti-tumor immune response.

References

- 1 Pardoll, D. (2003) Does the immune system see tumors as foreign or self? *Annu. Rev. Immunol.* 21, 807–839
- 2 Chouaib, S. *et al.* (2002) Tumor escape from killing: role of killer inhibitory receptors and acquisition of tumor resistance to cell death. *Tissue Antigens* 60, 273–281
- 3 Moller, P. and Hammerling, G. (1992) The role of surface HLA A, B, C molecules in tumor immunity. *Cancer Surv.* 13, 101–127
- 4 Larsson, M. *et al.* (2001) Dendritic cells resurrect antigens from dead cells. *Trends Immunol.* 22, 141–148
- 5 Daniels, G.A. *et al.* (2004) A simple method to cure established tumors by inflammatory killing of normal cells. *Nat. Biotechnol.* 22, 1125–1132
- 6 Engelhard, V.H. *et al.* (2002) Antigens derived from melanocyte differentiation proteins: self-tolerance, autoimmunity, and use for cancer immunotherapy. *Immunol. Rev.* 188, 136–146
- 7 Srivastava, P. (2002) Interaction of heat shock proteins with peptides and antigen presenting cells: chaperoning of the innate and adaptive immune responses. *Annu. Rev. Immunol.* 20, 395–425

- 8 Castelli, C. *et al.* (2004) Heat shock proteins: biological functions and clinical application as personalized vaccines for human cancer. *Cancer Immunol. Immunother.* 53, 227–233
- 9 Manjili, M.H. *et al.* (2002) Development of a recombinant HSP110-HER-2/neu vaccine using the chaperoning properties of HSP110. *Cancer Res.* 62, 1737–1742
- 10 Todryk, S.M. *et al.* (2003) Facets of heat shock protein 70 show immunotherapeutic potential. *Immunology* 110, 1–9
- 11 Millar, D.G. *et al.* (2003) Hsp70 promotes antigen-presenting cell function and converts T-cell tolerance to autoimmunity *in vivo*. *Nat. Med.* 9, 1469–1476
- 12 Todryk, S. *et al.* (1999) Heat shock protein 70 induced during tumor cell killing induces Th1 cytokines and targets immature dendritic cell precursors to enhance antigen uptake. *J. Immunol.* 163, 1398–1408
- 13 Noessner, E. *et al.* (2002) Tumor-derived heat shock protein 70 peptide complexes are cross-presented by human dendritic cells. *J. Immunol.* 169, 5424–5432
- 14 Becker, T. *et al.* (2002) CD40, an extracellular receptor for binding and uptake of Hsp70-peptide complexes. *J. Cell Biol.* 158, 1277–1285
- 15 Asea, A. *et al.* (2000) HSP70 stimulates cytokine production through a CD14-dependant pathway, demonstrating its dual role as a chaperone and cytokine. *Nat. Med.* 6, 435–442
- 16 Asea, A. *et al.* (2002) Novel signal transduction pathway utilized by extracellular HSP70: role of toll-like receptor (TLR) 2 and TLR4. *J. Biol. Chem.* 277, 15028–15034
- 17 Srivastava, P. (2003) Hypothesis: controlled necrosis as a tool for immunotherapy of human cancer. *Cancer Immun.* 3, 4
- 18 Melcher, A. *et al.* (1998) Tumor immunogenicity is determined by the mechanism of cell death via induction of heat shock protein expression. *Nat. Med.* 4, 581–587
- 19 Gallucci, S. *et al.* (1999) Natural adjuvants: endogenous activators of dendritic cells. *Nat. Med.* 5, 1249–1255
- 20 Huang, X.F. *et al.* (2003) A broadly applicable, personalized heat shock protein-mediated oncolytic tumor vaccine. *Cancer Res.* 63, 7321–7329
- 21 Belli, F. *et al.* (2002) Vaccination of metastatic melanoma patients with autologous tumor-derived heat shock protein gp96-peptide complexes: clinical and immunologic findings. *J. Clin. Oncol.* 20, 4169–4180
- 22 Lindquist, S. and Craig, E.A. (1988) The heat shock proteins. *Annu. Rev. Genet.* 22, 631–637

0167-7799/\$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tibtech.2004.12.011

Papa's got a brand new tag: advances in identification of proteases and their substrates

Alan B. Marnett and Charles S. Craik

Program in Chemistry and Chemical Biology, Department of Pharmaceutical Chemistry, The University of California, San Francisco, California 94143, USA

Characterization of proteolytic enzymes and their substrates presents a formidable challenge in the context of biological systems. Despite the fact that an estimated 2% of the human genome codes for proteases, only a small fraction of these enzymes have well-characterized functions. Much of the difficulty in understanding protease biology is a direct result of the complexity of

regulation, localization and activation exhibited by this class of enzymes. Here, we focus on several recently developed techniques representing crucial advances toward identification of proteases and their natural substrates.

The protease problem

Currently, a total of 686 proteases and their homologs are defined in the human genome, most of which belong to one

Corresponding author: Craik, C.S. (craik@cgl.ucsf.edu).

Available online 18 December 2004

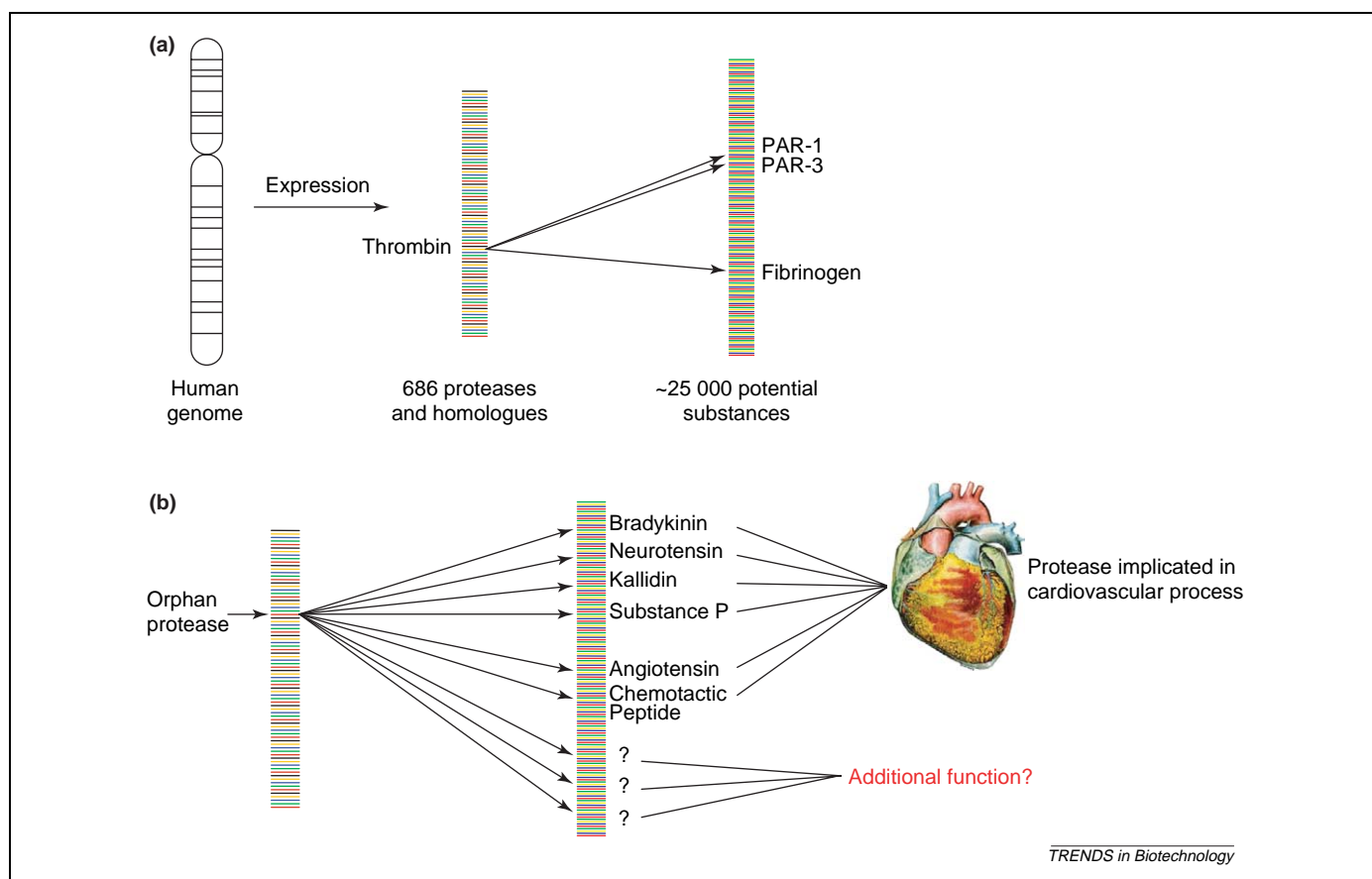


Figure 1. The protease problem. (a) The human genome encodes an estimated 25 000 proteins, 686 of which are proteases and their homologs. Thrombin, a serine protease, cleaves fibrinogen and protease-activated receptors-1 and -3 (PAR-1 and PAR-3). (b) Identification of natural substrates might implicate an orphan protease in a physiological process. Discovery of all natural substrates of an enzyme is crucial, because many proteases are involved in multiple processes.

of five groups: the metallo, serine, cysteine, threonine or aspartyl proteases [1]. These enzymes are categorized by their mechanism of peptide bond hydrolysis in which either an amino acid residue in the serine, cysteine or threonine proteases or an activated water molecule in the metallo or aspartyl proteases, acts as the catalytic nucleophile. Although originally regarded as tenacious digestive enzymes, proteases have crucial roles in several exquisitely regulated physiological processes, including development, innate and adaptive immunity, cell cycle regulation and apoptosis. Clearly proteolysis is a finely choreographed event *in vivo* ensuring cleavage of the appropriate substrate at a precise time, location and cellular context. Aberrant regulation of endogenous protease activity results in a variety of life-threatening illnesses, such as hemophilia, cancer and heart disease. As a result, there is significant interest in proteases as therapeutic targets. However, lack of knowledge of the explicit roles of proteases has hindered these efforts, often resulting in unforeseen physiological effects following enzyme inhibition. Even well-studied enzymes with known functions can have major roles in unexpected processes.

The possibility of multiple functions in proteases can complicate the situation further, as seen in the case of thrombin. The proteolytic activity of thrombin had been studied for decades in the selective processing of blood coagulation factors. It was then shown that thrombin

could activate platelets and regulate the behavior of other cells by means of G protein-coupled protease activated receptors (PARs) [2]. PARs provide a mechanism by which a protease, such as thrombin, can act as a hormone and communicate directly with cells in addition to its role in the blood coagulation cascade.

Hence we are presented with the daunting challenge of understanding the physiological roles of all proteases, and the substrates they process, in their entirety (Figure 1). Recent advances in activity-based protease labeling reagents, substrate profiling and differential protein labeling methods provide promising methodologies for identification of active proteases and their natural substrates, and increase the number of tools available to the protease biologist, biochemist and cell biologist.

Identifying proteases in a cellular context

A major challenge in elucidating the mechanism of protease involvement in a physiological process is the determination of all enzymes present. However, because of the tightly regulated controls over protease activity, mRNA levels frequently show little correlation to active protein. Even quantification of proteases at the protein level can be misleading because enzymatic activity can vary as a result of differential levels of endogenous protease inhibitors. Therefore, the problem evolves from identification of proteases to identification of active proteases.

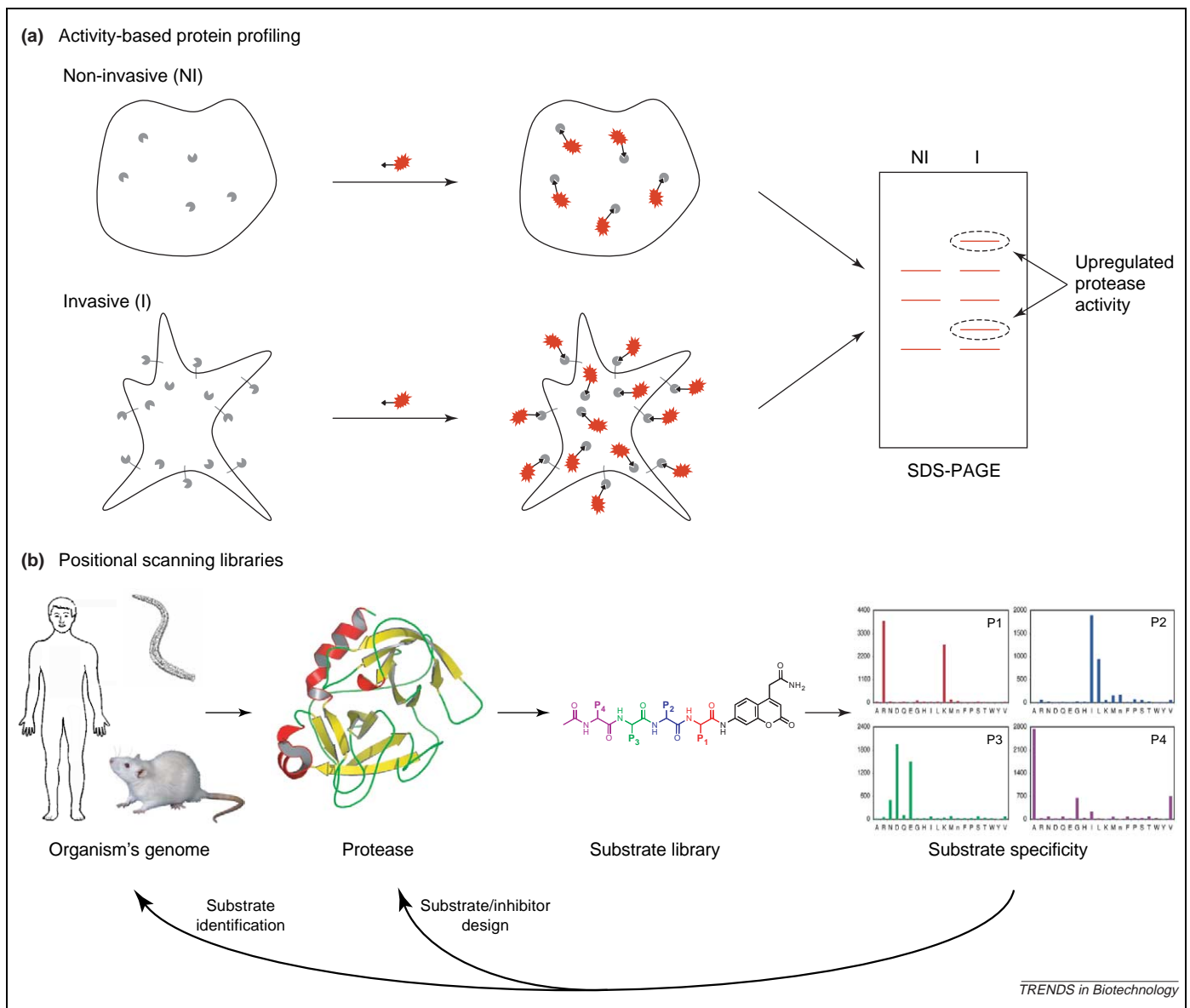


Figure 2. Recently developed techniques for protease and substrate identification. (a) Activity-based profiling identifies active proteases (gray) using fluorescently labeled (red) inhibitors. Lysates from non-invasive (NI) and invasive (I) cells are analyzed by gel electrophoresis. (b) Positional scanning libraries are used to determine substrate specificity of proteases from many organisms. Substrate information can be used to discover candidate protease substrates or to develop specific enzyme inhibitors.

To determine the presence of active proteases within a lysate, cell or tissue homogenate, mechanism-based inhibitors targeting the active site of the protease have been modified to incorporate a detection moiety such as biotin, a fluorophore or a radioactive probe [3–4]. Referred to as activity-based probes, these inhibitors predominantly label the active form of the enzyme through covalent bond formation with the catalytic center of the protease. They label a particular class of enzyme based on the reactivity of the electrophilic moiety but will not distinguish among different members of the class because specificity determinants are not present. Treatment of a given biological sample with an inhibitor, followed by SDS-PAGE analysis, detection and mass spectrometry, rapidly identifies the active enzymes (Figure 2a). The activity-based protein profiling approach has been applied to serine and cysteine proteases as a result of the formation of a covalent enzyme-inhibitor adduct during

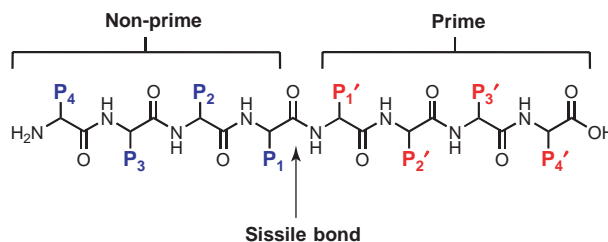
inactivation. However, analogous mechanism-based inhibitors are ineffective against the large class of metalloproteases and the important aspartyl proteases because water is the active nucleophile and no covalent intermediate is formed with the enzyme.

Recent work by Saghatelian *et al.* circumvented the requirement for covalent bond formation during mechanism-based inhibition of metalloproteases (MMP) [5]. A broad-spectrum zinc chelating hydroxamate was modified to include a benzophenone photocrosslinker in the S₂' position of the inhibitor, transforming the compound from a reversible to irreversible inhibitor upon irradiation at 365 nm. Incorporation of the crosslinker only minimally affected the IC₅₀ values against MMP-2, -7, and -9 and resulted in an inhibitor capable of discriminating between the active or zymogen form of the enzyme. Profiling both invasive and non-invasive melanoma cell lines for active metalloproteases, the

Table 1. Currently available chemical-based substrate profiling methods^a

Specificity	Mechanism	Method	Detection	Refs
Non-prime side	Coumarin-based peptide substrates	Solution, 96-well plate	Fluorescence	[8,13]
	Peptide-based epoxide inhibitors	Solution inhibition, SDS-PAGE	Radioactivity	[7]
	Coumarin-based peptide substrates	Solid phase, glass slide	Fluorescence	[10]
	PNA-tagged peptidyl inhibitors	Solution inhibition, array hybridization	Fluorescence	[11,15]
Prime side	Lanthanide ion chelating substrates	Solution, 96-well plate	Fluorescence	[9]
	Biotinylated peptide substrates	Solution cleavage	Sequencing	[12]
Both	Peptide mixtures	Solution cleavage	Mass spectrometry	[14]

^aLibraries are categorized according to the substrate information obtained from each method. P4-P4' represent amino acids of the substrate where cleavage occurs at the scissile bond, between P1 and P1'.



authors discovered the dramatic upregulation of nepri-lysin activity, a peptide-hormone-processing enzyme, in invasive melanomas. Interestingly, the enzyme shares no sequence homology with known MMPs, and reflects the versatility of the probe in labeling many subgroups within the metalloprotease superfamily.

With the expansion of activity-based protein profiling technology to include metalloproteases, in theory >90% of proteolytic enzymes might now be monitored in many complex cellular contexts. However, because currently available chemistries do not label all members of a given class, further development of the electrophilic moieties will be necessary to profile all proteases.

From protease to substrates

Another challenge is the identification of the natural substrates of a protease among thousands of cellular proteins, which is critical for defining the role of 'orphan' proteases with unknown function. Perhaps the most widely applied methodology in the search for protease substrates is the candidate approach, in which the substrate-recognition sequence of the protease is determined and used to search the proteome for potential matches. Originally, determining substrate specificity was an arduous process that involved monitoring cleavage of commercially available individual peptides by HPLC. The emergence of substrate phage display in the early 1990s provided an unbiased method for selecting optimized protease cleavage sequences, spanning both the prime and non-prime sides, from millions of potential substrates [6]. Although significantly more exhaustive in searching sequence space than screening individual substrates, the entire process can be technically challenging and time consuming even for a single enzyme.

A major advance toward increasing the throughput of substrate-specificity determination was the development of positional scanning synthetic combinatorial libraries (Figure 2b) [7–15]. These libraries systematically profile each subsite of an enzyme, resulting in either non-prime or prime side preferred substrate sequences (see Table 1 for substrate nomenclature [16]). The tandem use of

these libraries provides amino acid sequence data spanning P4-P4'. To date, substrate libraries are the fastest method for determining the specificity of an enzyme and have proven effective in identifying substrates of highly selective proteases. Although pooled libraries do not address cooperativity, arrays of single substrates and inhibitors printed on glass chips provide insight into cooperativity and might prove useful in understanding substrate specificity. With the development of positional scanning libraries, the time required to obtain protease specificity information has been reduced from days to minutes.

Substrate-specificity information is proving crucial not only for identifying potential substrates, but also for incorporation of specificity determinants into inhibitors, providing highly selective activity-based probes. These are useful in examining the role of an enzyme in a complex biological solution as well as dissecting structure–function relationships in a given protease. For example, Marnett *et al.* used specific activity-based probes to determine the intramolecular communication between two independent active sites of a herpesvirus protease thereby identifying a novel site for therapeutic intervention [15].

Natural substrate discovery by protein labeling

Ultimately, defining the preferred substrate specificity to scan the proteome for potential substrates is an indirect method. Many factors, such as exosites and extended substrate interactions are not accounted for and might therefore render the libraries less applicable in certain cases. Accordingly, a methodology for monitoring substrates directly is desirable. Proteomic approaches to substrate identification have been reviewed and critically evaluated by Lopez–Otin and Overall [16].

Recently, two reports described differential protein-labeling techniques for rapid identification of variations in protein levels as a result of protease cleavage that were used to identify both intra- and extracellular substrates [17,18]. Tam *et al.* employed an isotope-coded affinity tag (ICAT) labeling strategy to monitor the relative abundance of proteins present in conditioned medium

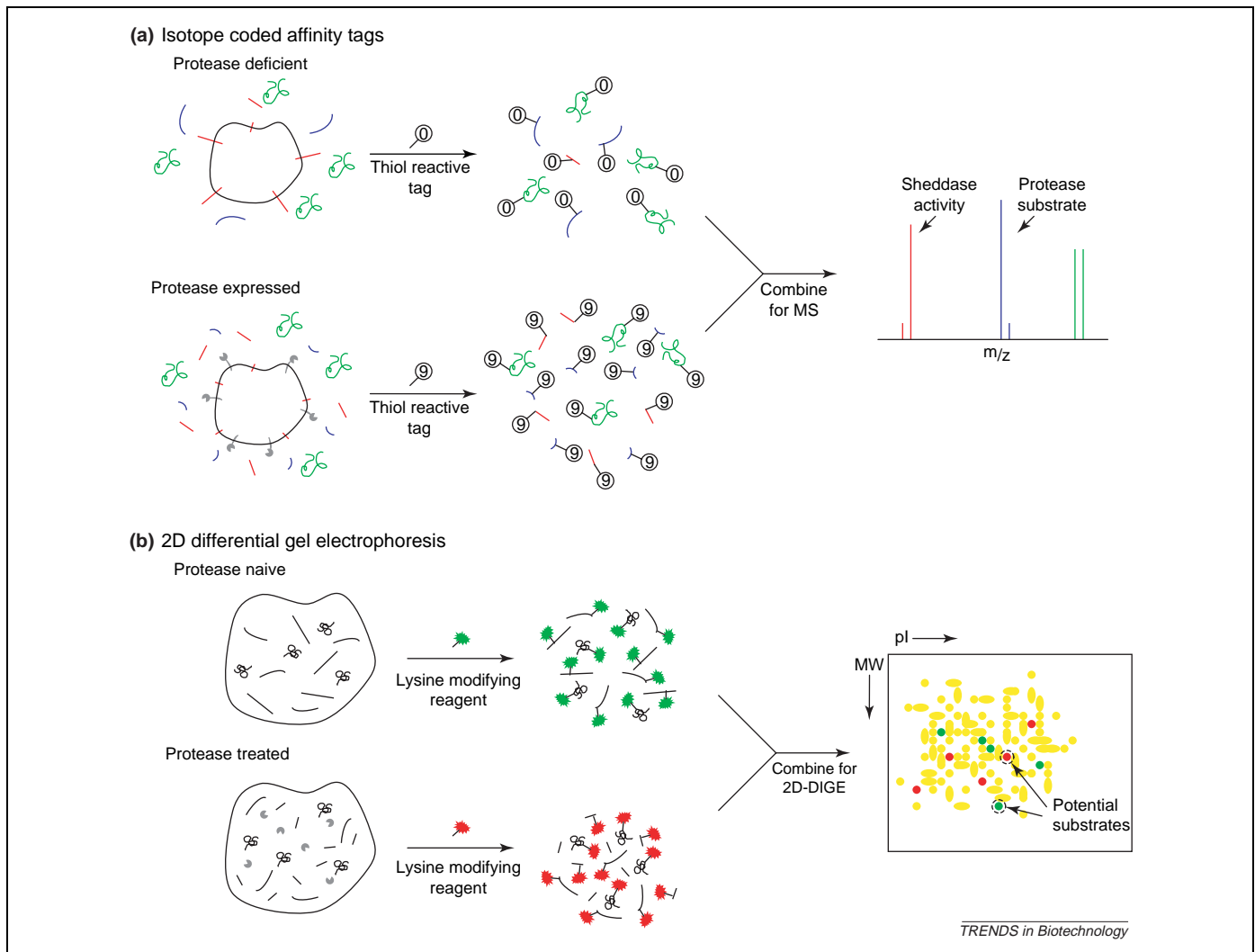


Figure 3. Mass spectrometry-based approaches to substrate discovery. (a) Isotope coded affinity tags containing $^{13}\text{C}_0$ (0) or $^{13}\text{C}_9$ (9) label thiols in differentially treated protein mixtures and are analyzed by mass spectrometry. (b) Differential gel electrophoresis is used to analyze differences in protein content following treatment with Cy3 (green) or Cy5 (red). Abbreviations: m/z, mass to charge ratio; MW, molecular weight; pl, isoelectric point.

resulting from a human breast carcinoma cell line transfected with a matrixmetalloprotease (MT1-MMP) or vector control [17]. Decreased protein levels in the presence of MT1-MMP were attributed to direct substrate cleavage by the enzyme, whereas increased levels were attributed to sheddase activity of the enzyme, releasing membrane-bound substrate molecules from the cell into the medium (Figure 3a).

To examine the consequences of proteolysis on intracellular substrates, Bredemeyer *et al.* subjected mouse lymphoma cell lysates to serine proteases Granzyme A or B [18]. Lysine residues of control or enzyme-treated lysates were subsequently modified by 1-(5-carboxypentyl)-1'-propylindocarbocyanide halide *N*-hydrosuccinimidyl ester (Cy3), or 1-(5-carboxypentyl)-1'-methylindocarbocyanide halide *N*-hydrosuccinimidyl ester (Cy5), respectively, and combined for global protein analysis (Figure 3b). Fluorescence two-dimensional differential gel electrophoresis (FL-2D-DIGE) clearly identified several proteins with altered abundance as a result of protease cleavage. The procedure was used to identify novel substrates of Granzyme B and

the physiological consequence of cleaving these putative substrates is currently being pursued.

Prospects for the future of protease biology

Over the past 20 years it has become clear that proteases range from non-specific digestive enzymes to specific processing enzymes, and are involved in many diverse and elegant physiological processes. It is anticipated that ultimately proteases will be implicated in almost every biological process and understanding their role will become crucial in taking advantage of their tremendous therapeutic potential. Unfortunately, it is the breadth and variety of roles that make proteases extremely difficult to study because each protease varies in expression, activation and localization. Therefore, it is increasingly unlikely that a single methodology will be developed to identify every active protease, or every natural substrate. Substrate profiling of a protease is a rapid and powerful initial step towards de-orphaning a protease. Protein labeling has recently been shown to be a promising technique with diverse applications. Specifically, activity-based probes, ICAT and 2D-FL-DIGE have been

successfully employed for identification of active proteases and for the implication of presumed substrates in a biological context. Coupled with substrate profiling, we expect that continued advances in mass spectrometry and proteomic analysis will result in new methods for use in combination on a protease-specific basis.

However, the exciting recent advances in substrate identification highlight candidate substrates that must ultimately be validated in an actual biological setting. Frequently, the process of substrate validation is limited to traditional biochemical techniques and represents the rate-limiting step, requiring a specific inhibitor, antibody or knockout to address the issue. Although outside the scope of this article, the development of a methodology such as RNAi has decreased the time necessary for substrate validation relative to traditional techniques. It has been successfully used to reveal Taspase-1 cleavage of the mixed-lineage leukemia gene product [19]. Similar developments in the biological validation of substrates will complement the recent advances in identification of proteases and their natural substrates and help define the roles of proteases in all physiological processes.

Acknowledgements

We are grateful to Carly R.K. Loeb for critical review of the manuscript. This work was supported by a University of California President's Dissertation Year Fellowship (A.B.M) and NIH Grants GM56531 and CA72006 (C.S.C.).

References

- Rawlings, N.D. *et al.* (2004). MEROPS: the peptidase database. *Nucleic Acids Res.* 32 Database issue: D160-D1634
- Vu, T.K. *et al.* (1991) Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell* 64, 1057–1068
- Greenbaum, D. *et al.* (2000) Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol.* 7, 569–581
- Liu, Y. *et al.* (1999) Activity-based protein profiling: the serine hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14694–14699
- Saghatelian, A. *et al.* (2004) Activity-based probes for the proteomic profiling of metalloproteases. *Proc. Natl. Acad. Sci. U. S. A.* 101, 10000–10005
- Matthews, D.J. and Wells, J.A. (1993) Substrate phage: selection of protease substrates by monovalent phage display. *Science* 260, 1113–1117
- Greenbaum, D.C. *et al.* (2002) Small molecule affinity fingerprinting. A tool for enzyme family subclassification, target identification, and inhibitor design. *Chem. Biol.* 9, 1085–1094
- Harris, J.L. *et al.* (2000) Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc. Natl. Acad. Sci. U. S. A.* 97, 7754–7759
- Barrios, A.M. and Craik, C.S. (2002) Scanning the prime-site substrate specificity of proteolytic enzymes: a novel assay based on ligand-enhanced lanthanide ion fluorescence. *Bioorg. Med. Chem. Lett.* 12, 3619–3623
- Salisbury, C.M. *et al.* (2002) Peptide microarrays for the determination of protease substrate specificity. *J. Am. Chem. Soc.* 124, 14868–14870
- Winssinger, N. *et al.* (2002) Profiling protein function with small molecule microarrays. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11139–11144
- Turk, B.E. *et al.* (2001) Determination of protease cleavage site motifs using mixture-based oriented peptide libraries. *Nat. Biotechnol.* 19, 661–667
- Thornberry, N.A. *et al.* (1997) A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. *J. Biol. Chem.* 272, 17907–17911
- Zhu, L. *et al.* (2003) The role of dipeptidyl peptidase IV in the cleavage of glucagon family peptides: *in vivo* metabolism of pituitary adenylate cyclase activating polypeptide-(1-38). *J. Biol. Chem.* 278, 22418–22423
- Marnett, A.B. *et al.* (2004) Communication between the active sites and dimer interface of a herpesvirus protease revealed by a transition-state inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6870–6875
- Lopez-Otin, C. and Overall, C.M. (2002) Protease degradomics: a new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* 3, 509–519
- Tam, E.M. *et al.* (2004) Membrane protease proteomics: Isotope-coded affinity tag MS identification of undescribed MT1-matrix metalloproteinase substrates. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6917–6922
- Bredemeyer, A.J. *et al.* (2004) A proteomic approach for the discovery of protease substrates. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11785–11790
- Hsieh, J.J. *et al.* (2003) Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* 115, 293–303

0167-7799/\$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tibtech.2004.12.010

The growing importance of fat in regenerative medicine

Brian M. Strem^{1,2} and Marc H. Hedrick^{1,2}

¹Macropore Biosurgery Inc., 6740 Top Gun Street, San Diego, California 92121, USA

²Departments of Bioengineering and Surgery, University of California, Los Angeles, USA

A recent publication by Michael Longaker and colleagues represents a landmark for the use of adipose tissue as a source of cells for tissue regeneration. The authors investigated the ability of adipose tissue-derived cells (ADCs) to regenerate critical size calvarial (superior portion of the skull) defects in mice by using a novel osteoconductive apatite-coated Poly-lactic-co-

glycolic acid (PLGA) scaffold for cell delivery. Direct comparison of this osteogenic ability was performed with bone marrow stromal cells and juvenile calvarial-derived osteoblasts.

Introduction

Although we and others have previously reported the ability of adipose tissue-derived cells (ADCs) to undergo

Corresponding author: Strem, B.M. (bstrem@macropore.com).
Available online 13 December 2004