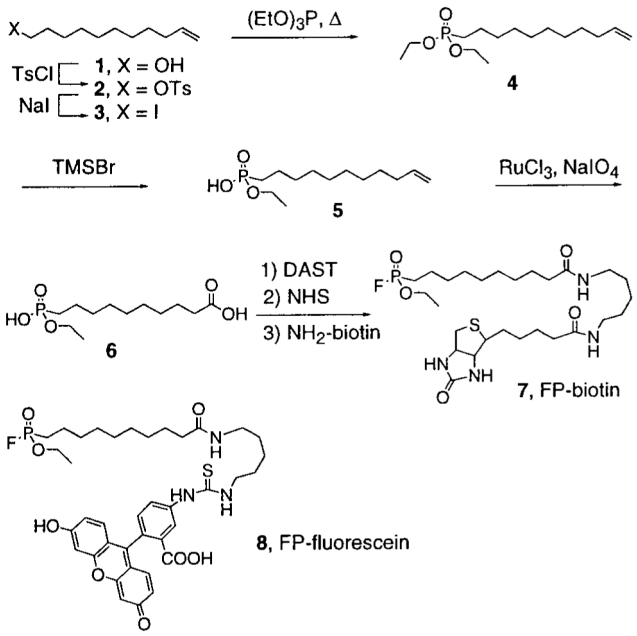
Proteome Wide Screening of Serine Protease Activity
Proc Natl Acad Sci **1999**, *97*, 14694; *Proteomics* **2001**, *1*, 1067; *Proc Natl Acad Sci* **2002**, *99*, 10335; *Biochemistry* **2001,** *40,* 4005; *J. Am. Chem. Soc*., **2005**, *127*, 10018.

Introduction & Hypothesis
• Serine hydrolases are one of the largest and most diverse classes of enzymes in the human proteome, ~1% of all predicted gene products.
• Could global proteome screening predict higher-order cellular properties?

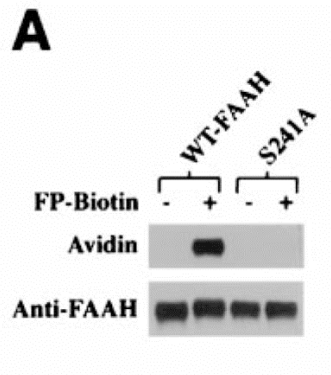
Used panel of human cancer cell lines because:
(*i*) diverse range (ii) well characterized (iii) differences in hormone responsiveness, (iv) differences in invasiveness, (*v*) derived from at least two distinct types of cancer to permit comparison of proteomic expression patterns both within and between cancer classes; (*vi*) previously been analyzed with gene expression microarrays, therefore allowing comparison between proteomic data and transcriptional profiles

Method Development
• Develop & synthesize reagents that will react with serine protease irreversibly. If these reagents bear either *labels* (e.g. fluorescent such as fluorescein but others could be used) or *tags* (e.g. a small molecule such as biotin that can be used for affinity separation NB K_D for biotin and corresponding protein avidin is ~10^{-15} M) then the enzymes that react can be assessed.

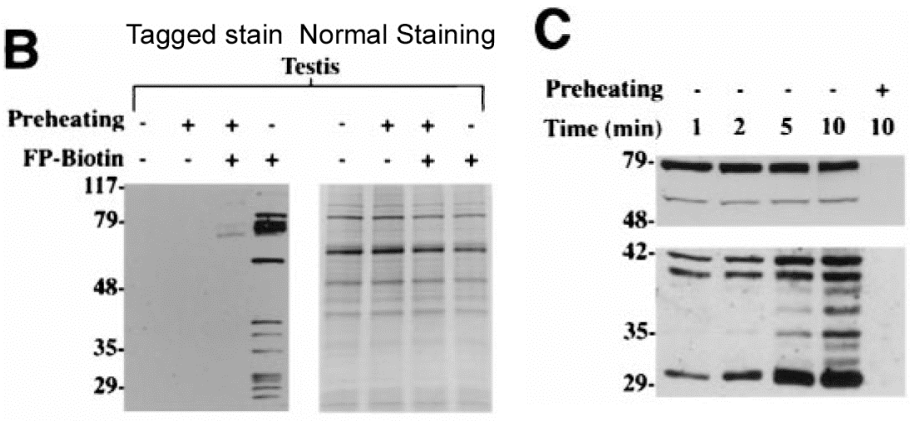


The method is based on long known fluorophosphonate suicide inhibitors. These are **broadly selective for serine hydrolases** because mimic TI (see sheet M2).

NB
DAST = Diethylaminosulphur trifluoride
NHS = *N*-hydroxysuccinimide
TMS = trimethylsilyl

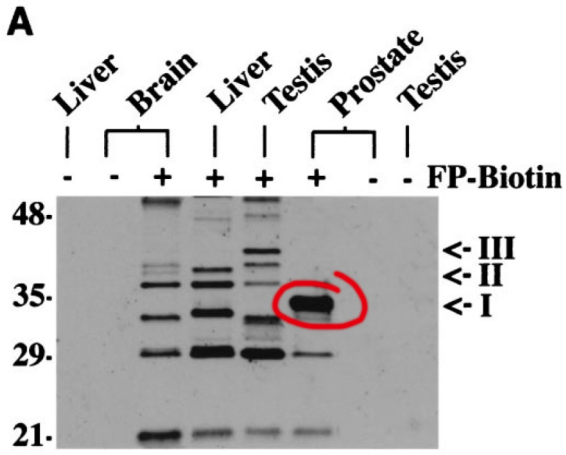


• Test the system against known mammalian enzyme FAAH (fatty acid amide hydrolase)
These are highlighted bands from PAGE 'gels' that have been stained using a protein avidin-HRP crosslink. HRP (horse radish peroxidase is an enzyme that when it reacts with hydrogen peroxide along with substrates such as dioxetanes / luminol give rise to chemiluminescence that appears as a band on the gel). So if a protein has been tagged with **7**, avidin will bind to that part of the gel, that band will have HRP activity and will be marked when 'stained'. When the Ser in the active site is mutated to Ala (S241A): no band.



Here in a tissue sample the staining using the avidin-biotin dependent method is lost when preheated at 80˚ C for 5 min. This denatures that serine protease activity and so stops the 'tagging'. The normal staining shows that the amount of protein there is the same for all samples. Note that there are several types of serine proteases in this mixed sample and that staining is tagging reaction time dependent.

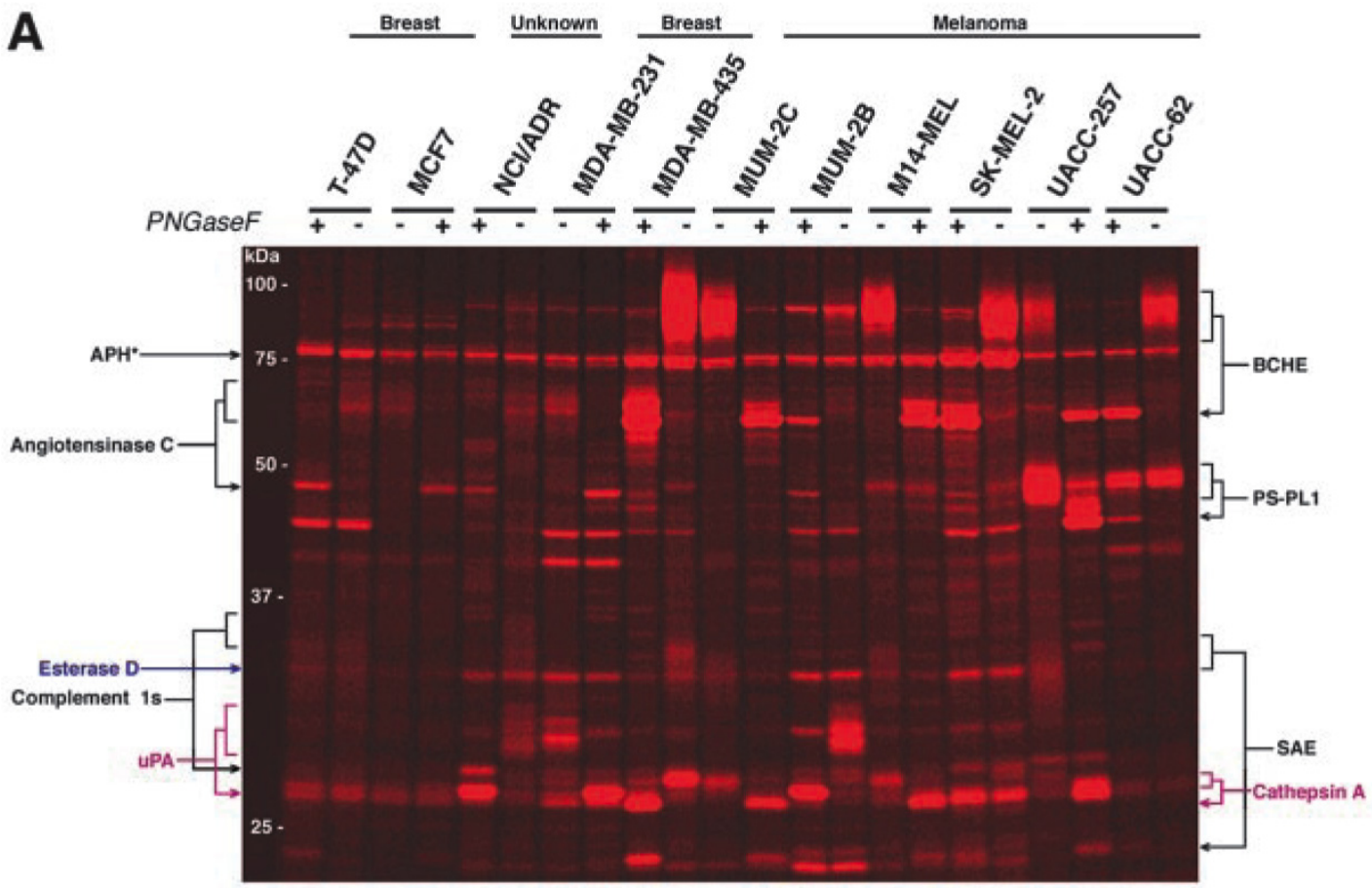
• This demonstrated that serine proteases could be detected at subnanomolar levels.
• The biotin tag not only allows the 'in gel' staining method but also allows the potential for 'fishing' for biotin bearing proteins using e.g. avidin selective affinity columns.
• Tagged proteins can be identified using protein sequencing methods (see sheet T2) – in particular proteolytic digest then MS.



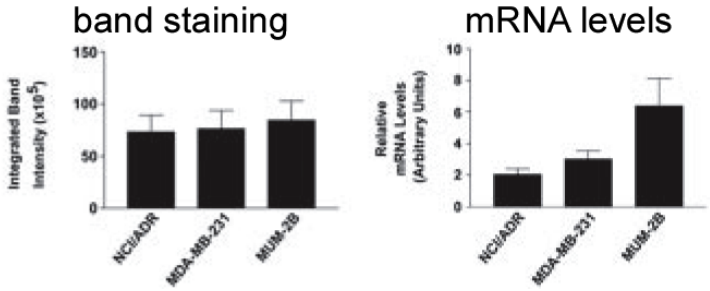
In these samples from several mammalian tissues several serine proteases are seen, including one that has similar mass to human prostate specific antigen.

Application to Human Cancer Cell Lines

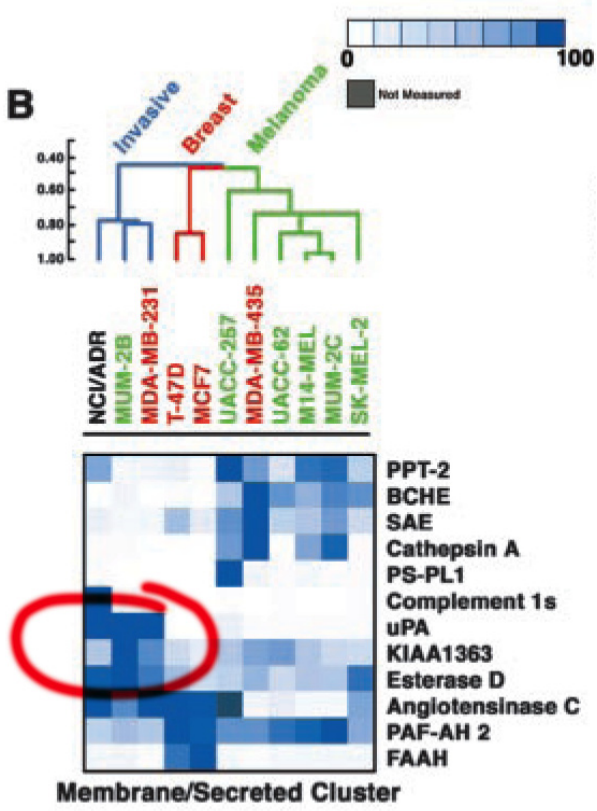
• The same method but using rhodamine fluorescent label instead of biotin as a marker was employed.
• Proteomes from each cell line were separated into three fractions: 1. secreted, 2. membrane-associated, and 3. soluble and then reacted
•These were analysed by PAGE and visualized in-gel by using a flatbed laser-induced fluorescence scanner.
• In parallel experiments, biotinylated FP probes were used to affinity isolate the active enzymes, which allowed for their molecular identification by mass spectrometry methods (see sheet T2).



This gel shows the results for the *secreted* proteases. Note the variation in the enzyme uPA. uPA is already linked to invasiveness of cancers but note how the enzymes activity does not exactly correlate with mRNA levels (below). This highlights the importance of activity-based assays and the difference in the products of transcription and translation



To analyse the correlation between enzyme activity and 'invasiveness', a 'cluster analysis' was performed (right). This simply groups activity levels according to origin or property. While most simply show activities that depend on cell line source. UPA and KIAA1363 show a fair link with invasiveness of

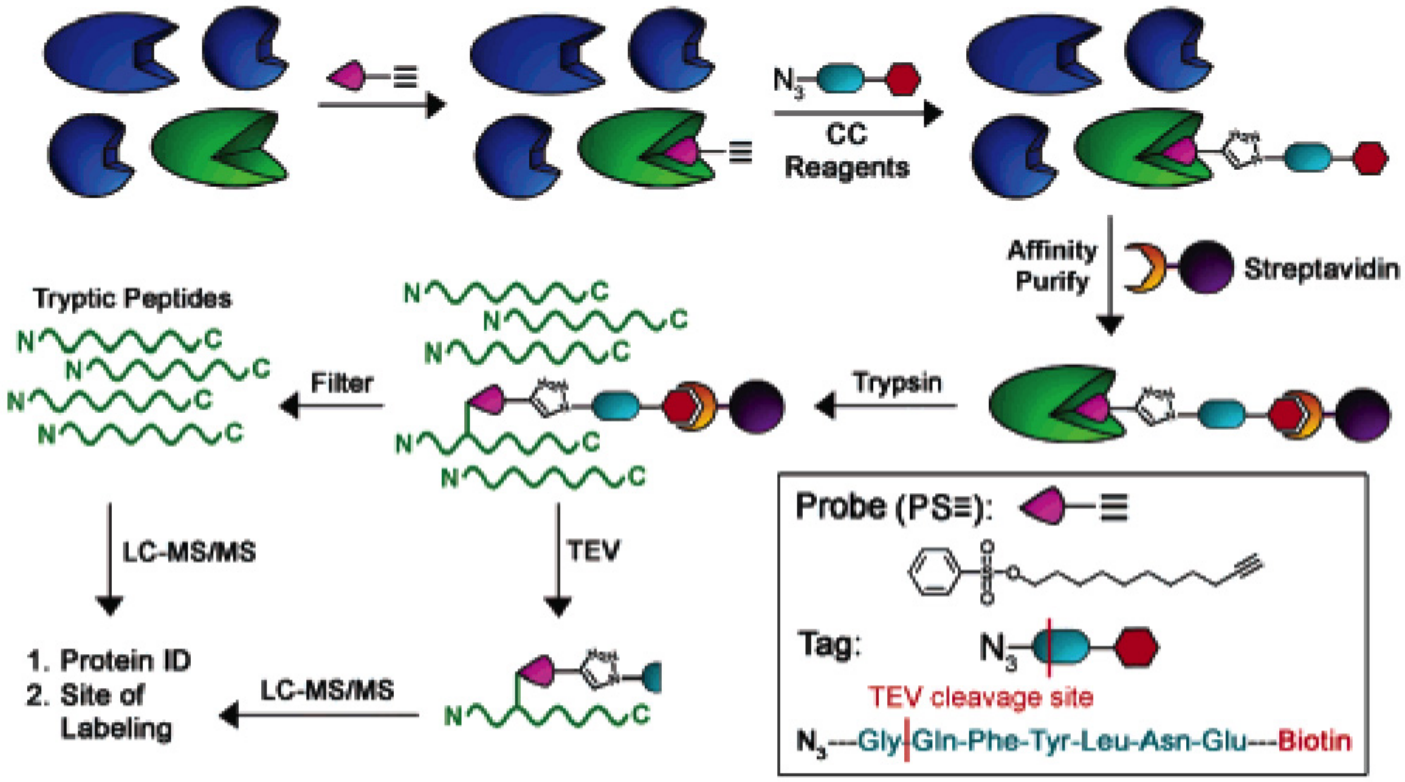


the corresponding cell lines. This provides a direct link from chemistry (activity) to biology (phenotype or pathology).

Problems?

• Enzyme active sites have evolved in various different ways – look at the probe – can you see features that might be incompatible with a) hydrophilic substrate or b) peptide specific enzymes?
• Consider the other classes of acyl transferases on sheet M2 and that we have encountered. Assess how these might react.
• Using you knowledge of enzyme kinetics, what sort of 'activity' is this really assessing?

Further Applications and Developments



The same ideas have been applied to a more generalized method for 'tagging' enzymes with active site nucleophiles.

Pointers to interpreting this figure:

• Note the sulfonate ester as a simple method for allowing S_N2 reactions with protein nucleophiles
• Note the aqueous compatible 3+2 cycloaddition between azide and alkyne
• Note the use again of biotin, here for affinity purification
• Note the 'on column' digestion of the tagged enzyme using trypsin to give peptide fragments for MS (see sheet T2)
• Note how only the tagged peptide remains and can then be separately identified by clevage with another protease TEV.
• The table below lists the nucleophiles in the protein that react with the sulfonate ester – note the broad coverage but that the majority 'hit' active site nucleophiles in this sample set.

Predicted Sites of Labeling for Select Enzymes[a]

enzyme	labeled peptide	site(s) of labeling	catalytic residue	active site residue	tryp digest % coverage
ECH1	K.EVDMGLAA**D*****VGTLQR.L	D204	yes	yes	69
VLCAD	R.IF**E*****GANDILR.L	E463	yes	yes	36
	K.ELGAFGLQVPSELGGLGLSNTQ **Y*****AR.L	Y161	no	yes	
LCAD	K.GFYYLMQELPQ**E*****R.L	E291	yes	yes	62
MCAD	K.IYQ**IY*****E*****GTAQIQR.L	Y400/E401	yes (E)	yes	42
ALDH6	R.**C*****MALSTAILVGEAK.K	C317	yes	yes	48
Thiolase	K.DGGQYALVAA**C*****AAGGQGHAM IVEAYPK.-	C459	yes	yes	42
QR2	K.VLAPQISFGL**D*****VSSEEER.K	D194	no	yes	29
Isocitrate DH (m)	K.SSGGFVWA**C*****K.N	C308	no	yes	46
Isocitrate DH (c)	K.SEGGFIWA**C*****K.N	C269	no	yes	15
His triad protein	R.ISQAE**E*****DD*****QQLLGHLLLVAK.K	E105/D107	unknown	unknown	31
LOC67914	R.AVLAGI**Y*****NTTE*****LVMMQDSSPD FEDTWR.F	Y246/E250	unknown	unknown	13

[a] Predicted sites of probe labeling are indicated with an asterisk (∗). In cases where Sequest did not unambiguously distinguish between two potential sites, both are given.

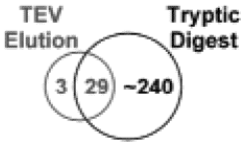


Figure. Proteins identified by TOP−ABPP. Proteins found only in TEV and tryptic digests were regarded as false assignments and background, respectively.

• The ***Figure*** shows a Venn diagram for the proteins identified using this method – only 29 were found in both TEV and Tryptic digests. Only these were considered valid identifications. Why? Trace through the Scheme to rationalize this.