PC235: 2009 Lecture 2 Sample Preparation for MS Analysis

Robert Chalkley

chalkley@cgl.ucsf.edu

Purification… then Analysis

- Mass Spectrometry essentially detects the most abundant components in your sample. If the most abundant component is antibody, or a detergent, then this is what you will detect, even if your component of interest is also present (albeit at a lower level).

- This talk is going to largely focus on methods for sample purification that are suitable / compatible with MS.

- MS Analysis will be discussed in the next lecture…
- Preparation methods specific to post-translational modifications will be discussed in lecture 4…

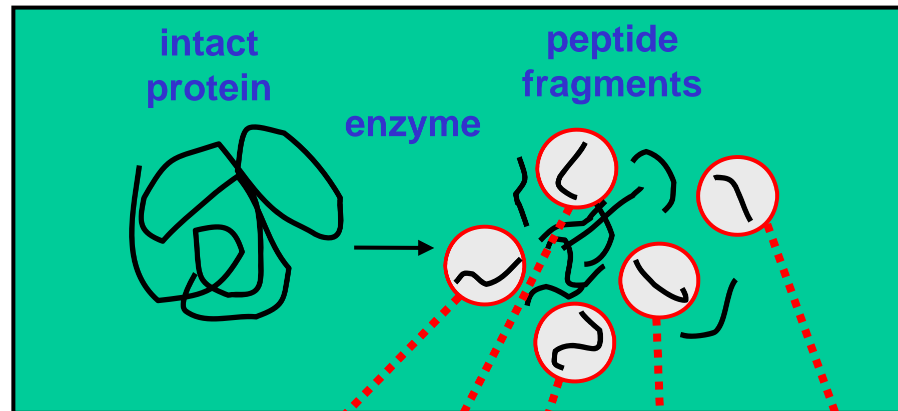
Why do we normally analyze peptides rather than proteins?

- MW of protein is not informative enough:
  - e.g. in NCBInr database 115 proteins mass 20000 +/- 2 Da
  - Human: 5 proteins.
- Sequence in database may be wrong: 1 amino acid difference will change the mass, such that you will not identify the protein
- Sequence in the database may be of the pre-protein with signal peptide still attached. May be alternatively spliced product.
- Protein may be post-translationally modified.
- Protein may be a degradation product.

- Analysis of intact proteins is less sensitive than peptides.
- Fragmentation of intact proteins by CID is inefficient in comparison to peptides (ECD/ETD fragmentation is better but still insensitive)

- Solution: Digest protein into peptides, then analyze peptides.

Peptides from Proteins



MEMEKEFEQIDKSGSWAAIYQDIRHEASDFPCRVAK**LPKNKNRNRYRDVS**
PFDHSRIKLHQEDNDYINASLIKMEEAQRSYILTQGPLPNTCGHFWEMVW
EQKSRGVVMLNRVMEKGSLKCAQYWPQKEEKEMIFEDTNLKLTLISEDIK
SYYTVRQLELENLTTQETREILHFHYTTWPDFGVPESPASFLNFLFKVRE
SGSLSPEHGPVVVHCSAGIGR**SGTFCLADTCLLLMDKRK**DPSSVDIKKVL
LEMRKFR**MGLIQTADQLRFSYLAVIEGAK**FIMGDSSVQDQWKELSHEDLE
PPPEHIPPPPRPPKRILEPHNGKCREFFPNHQWVKEETQEDKDCPIKEEK
GSPLNAAPYGIESMSQDTEVRSRVVGGSLRGAQAASPAKGEPSLPEKDED
HALSYWKPFLVNMCVATVLTAGAYLCYRFLFNSNT

Proteolytic Enzymes

- Trypsin
- Arg-C
- Lys-C
- Lys-N
- Glu-C
- Asp-N
- Chymotrypsin
- Pepsin

What Makes a Good Digestion Enzyme?

•Produces 'convenient' sized peptides: 8-20 residues long

Digest with Chymotrypsin

IVGGY TCAANSIPY QVSLNSGSHF CGGSLINSQW VVSAAHCY KSRIQVRL
GEHNIDVLEGNEQF INAAKIITHPNF NGNTLDNDIMLIKLSSPATLNSRV
ATVSLPRSCAAAGTECLISGW GNTKSSGSSY PSLLQCLKAPVLSDSSCKS
SY PGQITGNMICVGF LEGGKDSCQGDSGGPVVCNGQLQGIVSW GY
GCAQKNKPGVY TKVCNY VNW IQQTIAAN

•Cleavage Specificity – Must be able to predict where the enzyme will cut:

- •Glu-C: Cuts after E, some D.
- •Chymotrypsin: Cuts after F, W, Y, some L, occasionally V…

•Should not produce too many autolysis products: want to see peptides from your sample not from your enzyme!

Why We Generally Use Trypsin

- Majority of peptides 7 - 20 amino acids in length
- High Enzyme Specificity – cuts all Lys and Arg (to lesser extent followed by Pro).
- Produces peptides with basic C-terminus – give good fragmentation series

- Can be modified so it doesn't produce many peptides due to self-digestion
  - Methylate Lysine Residues

IVGGYTCAANSIPYQVSLNSGSHFCGGSLINSQWVVSAAHCYKSR IQVR
LGEHNIDVLEGNEQFINAAK IITHPNFNGNTLDNDIMLIKLSSPATLNSR
VATVSLPR SCAAAGTECLISGWGNTKSSGSSYPSLLQCLKAPVLSDSSCK
SSYPGQITGNMICVGFLEGGKDSCQGDSGGPVVCNGQLQGIVSWGYGCAQK
NKPGVYTKVCNYVNWIQQTIAAN

Tryptic digest of hemoglobin b chain

```
VHLTPEEKSA VTALWGKVNV DEVGGEALGR LLVVYPWTQR FFESFGDLST
PDAVMGNPKV KAHGKKVLGA FSDGLAHLDN LKGTFATLSE LHCDKLHVDP
ENFRLLGNVL VCVLAHHFGK EFTPPVQAAY QKVVAGVANA LAHKYH
```

MH+ (mono.)	Residues	Sequence
147.11	66 - 66	(K) K (V)
246.18	60 - 61	(K) VK (A)
319.14	145 - 146	(K) YH ( )
412.23	62 - 65	(K) AHGK (K)
932.52	9 - 17	(K) SAVTALWGK (V)
952.51	1 - 8	( ) VHLTPEEK (S)
1126.56	96 - 104	(K) LHVDPENFR (L)
1149.67	133 - 144	(K) VVAGVANALAHK (Y)
1274.73	31 - 40	(R) LLVVYPWTQR (F)
1314.67	18 - 30	(K) VNVDEVGGEALGR (L)
1378.70	121 - 132	(K) EFTPPVQAAYQK (V)
1421.67	83 - 95	(K) GTFATLSELHCDK (L)
1669.89	67 - 82	(K) VLGAFSDGLAHLDNLK (G)
1719.97	105 - 120	(R) LLGNVLVCVLAHHFGK (E)
2058.95	41 - 59	(R) FFESFGDLSTPDAVMGNPK (V)

Trypsin not always best choice

Rhodopsin (Membrane Protein)

MNGTEGPNFYVPFSNATGVVRSPFEYPQYYLAEPWQFSMLAAYMFLLIVLGFPINFLTLY
VTVQHKKLRTPLNYILLNLAVADLFMVLGGFTSTLYTSLHGYFVFGPTGCNLEGFFATLG
GEIALWSLVVLAIERYVVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPPLAGWSRYIP
EGLQCSCGIDYYTLKPEVNNESFVIYMFVVHFTIPMIIIFFCYGQLVFTVKEAAAQQQES
ATTQKAEKEVTRMVIIMVIAFLICWVPYASVAFYIFTHQGSNFGPIFMTIPAFFAKSAAI
YNPVIYIMMNKQFRNCMLTTICCGKNPLGDDEASATVSKTETSQVAPA

Mass	Residues	Sequence
903.4424	340 – 348	(K)TETSQVAPA (-)
1357.574	315 - 325	(R)NCMLTTICCGK (N)
1403.665	326 - 339	(K)NPLGDDEASATVSK (T)
1490.709	232 - 245	(K)EAAAQQQESATTQK (A)
1499.75	136 - 147	(R)YVVVCKPMSNFR (F)
1727.887	297 - 311	(K)SAAIYNPVIYIMMNK (Q)
2257.071	1 - 21	(-)MNGTEGPNFYVPFSNATGVVR (S)
3231.569	148 - 177	(R)FGENHAIMGVAFTWVMALACAAPPLAGWSR (Y)
5058.603	253 - 296	(R)MVIIMVIAFLICWVPYASVAFYIFTHQGSNFGPIFMTIPAFFAK (S)
5388.76	22 - 66	(R)SPFEYPQYYLAEPWQFSMLAAYMFLLIVLGFPINFLTLYVTVQHK (K)
6503.206	178 - 231	(R)YIPEGLQCSCGIDYYTLKPEVNNESFVIYMFVVHFTIPMIIIFFCYGQLVFTVK (E)
7182.745	70 - 135	(R)TPLNYILLNLAVADLFMVLGGFTSTLYTSLHGYFVFGPTGCNLEGFFATLGGEIALWSLVVLAIER(Y)

- Also, protein solubility may be an issue for membrane proteins

In-Gel Digestion

1. Cut band out of gel.
2. Wash gel pieces to remove SDS and destain coomassie / fluorescent stained bands.
3. Reduce disulfide bonds using DTT.
4. Alkylate free cysteine side-chains with iodoacetamide.
5. Wash to remove reducing / alkylating reagents.
6. Dehydrate gel pieces.
7. Add enzyme and digest for 4h or overnight.
8. Extract peptides from gel pieces.

Shevchenko, A. et al. *Anal Chem.* (1996) **68** p.850-8.

In-Solution Digestion

- Volatile salt, low concentration:
  - Usually 25 mM ammonium bicarbonate
- Protein may need to be denatured in order to get efficient digestion
  - ACN: most enzymes can tolerate up to 20-30% ACN
  - Urea or guanidine HCl: as much as tolerated by the enzyme:
    - Trypsin works in 2M urea or 1M guanidine HCl
    - Lys-C tolerates 2M guanidine HCl
- Add a detergent only if necessary
- 1-10% enzyme (w/w), ~ 4 h, @37°C

Carbamylation and Urea

•Upon storage, cyanates form in urea. These cause carbamylation of primary amines.

•Cyanates can be removed from urea using a mixed bed ion exchange resin.

Other Useful Digestion Information

- Modified enzymes e.g. TPCK (methylated) trypsin, are generally more stable, so more tolerant of heat, chaotropes…

- Enzymes isolated from different sources may display very different activity (Roche vs WAKO Lys-C)

- Asp-N:
  - has not been sequenced, so you will not identify Asp-N peptides in database searching.
  - is a metallo-protease, thus chelating agents will inactivate it.

Other Useful Digestion Information

- Covalent modifications may slow down or prevent cleavages: Lys(Me), Lys(Me_2), Lys(Ac)
- Glycosylation may sterically hinder proteolysis
  - N-linked sugars can be removed by peptide-N-glycosidase F (PNGase F)
- Endoproteases are poor exoproteases – implications for RRKR
- Enzymes may act as ligases – moving a few residues from one terminus to the other → transpeptidation
- Enzymes can be cheated:
  - Lys-C and trypsin will cleave Lys-imitations[1]
    - Beta elimination of phospho/glyco/sulfoSer, then SH-CH2-CH2-NH2 addition

[1]Rusnak, F. et al. (2002) *J. Biomol. Tech.* 13, 228-237.

Chemical Methods

- CNBr in 0.1M HCl (or neat formic acid) - Met-C
- Asn↓Gly - 2M hydroxylamine, 2M Guanidine HCl (pH 9), 45°C, 4h
- Asp↓Xxx: 10 mM HCl, 108°C, 2 h.

- Effective for membrane proteins.
  - Will work in-gel; on the blot; in a slurry
  - May access sites when protease cannot.

Protein Level Separation

- 1D Gels
- 2D Gels
- Ion Exchange Chromatography
- Size Exclusion Chromatography
- Reverse Phase Chromatography

- Mw cut-off filters

1D SDS-PAGE Gels

•SDS-PAGE: separation by protein size(mw).
•Probably the highest resolution single dimension separation available for proteins. Very popular for protein purification or for separation of a complex mixture prior to MS analysis.
•For analysis of complex mixture, the whole gel lane may be cut into 10-20 bands, then each band is analyzed to characterize the whole mixture: 'geLC-MS'.

•Advantages:
- •High resolution.
- •Removes detergents, salts…
- •Relatively easy.

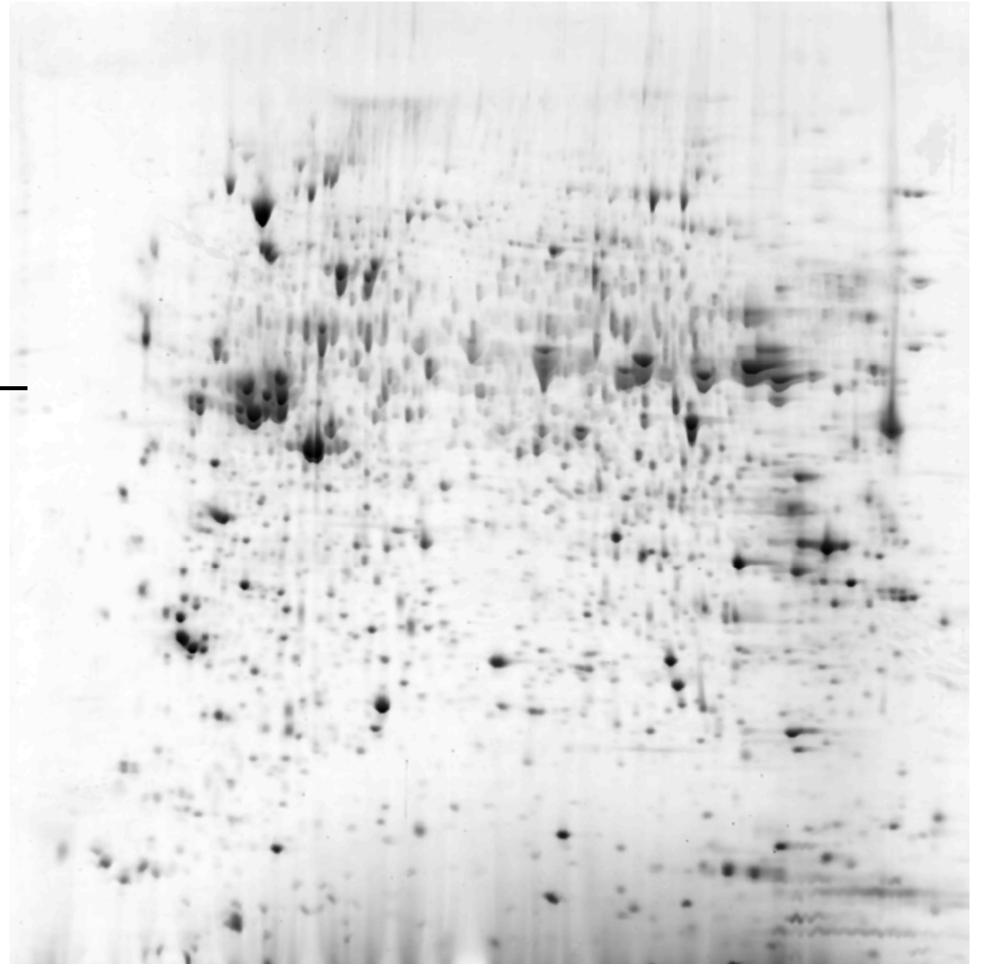
•Disadvantages:
- •Limited sample loading capacity.
- •Can lose small proteins.

2D Gels

Separate in one dimension based on pI/charge and in the second dimension by size.

Advantages:

- Highest resolution separation available.
- Very visual and semi-quantitative – good for identifying changes.
- Can separate differently modified versions of the same protein.



Disadvantages:

- Slow.
- Difficult to get highly reproducible separation.
- Biased against proteins of high pI, low pI, small proteins.
- Limited loading capacity.

Other 1D Gel Approaches

Native Gel Electrophoresis:

•Separation by size under non-denaturing conditions. Used for separating protein complexes. Can also be performed in 2D.

Isolectric focusing/Immobilized pH gradient gels:

•Separation by pI. Can be high resolution if over a narrow pH range.

Protein Staining

•Coomassie: Sensitivity: 10-50ng. Binds to SDS. Quick, easy and highly compatible with MS.

•Sypro Ruby: Fluorescent stain. Sensitivity: 1-2ng. Binds to SDS. Sensitive and highly compatible with MS, but bands not visible to the naked eye, so how do you cut them out?

•Silver Staining: Sensitivity 1-2ng. Binds directly to protein. Binding can be reversed by strong oxidizing agents. Less compatible with MS (staining protocols that use glutaraldehyde are incompatible).

•Zinc Imidazole: Negative stain. Sensitivity 5-20ng. Metal precipitate forms in gel causing gel to turn cloudy/opaque except where protein present. Fast (10 min), reversible staining. Stain 'fades' relatively quickly, so bands/spots need to be cut straight away. Highly compatible with MS.

Protein Liquid Chromatography

Ion Exchange Chromatography

•Separate on basis of charge (and sometimes also hydrophobicity/hydrophilicity). Reasonably high capacity, but low resolution.

Size-Exclusion Chromatography

•Separate on basis of size. Can be performed under denaturing or non-denaturing conditions. Can only be performed on relatively large amounts of sample and is not concentrating, so resolution is dependent on sample loading volume (but is never high resolution).

Reverse Phase Chromatography

•Elute components according to increased hydrophobicity. Relatively high resolution, but only good for smaller proteins (although newer non-porous silica resins have improved performance with larger proteins).

•Hydrophilic Interaction Chromatography (HILIC)

•Effectively the inverse of reverse phase, where components elute according to increased hydrophilicity.

Molecular Weight Cut-off Filters

Relatively quick and effective way to desalt large amounts of sample

- Cut-offs are very inaccurate:

    e.g. you can probably find a given 50kDa protein in both the retained and flow-through of a 30kDa MWCO filter.

- You will get sample loss to the filter membrane.
    - May not be significant, depending on the amount of sample.

Peptide-level Separation

In general it is possible to achieve higher resolution separation at the peptide level than the protein level.

Reverse phase Chromatography

- Normally performed at low pH. High resolution and often interfaced online with mass spectrometry.
- Can also be performed at high pH, giving an orthogonal separation to at low pH.

Ion Exchange Chromatography

- Separation according to charge. Most common ion exchange chromatography for peptides is strong cation exchange (SCX), where components elute with increasing charge.
    - Slight drawback: if you have digested your protein with trypsin, most peptides are going to have two charges (free N-terminal amine group and the side-chain of the lysine/arginine). Hence, SCX provides relatively poor resolution of tryptic peptides.

Affinity Purification

- Immunoprecipitation:
  - High specificity.
  - Problems caused by the introduction of large amounts of antibody in sample.
    - Cross-linking antibody to column reduces leaching, but some always leaks off.
    - Often necessary to perform 1D gel separation afterwards to separate antibody chains from other components in the sample.

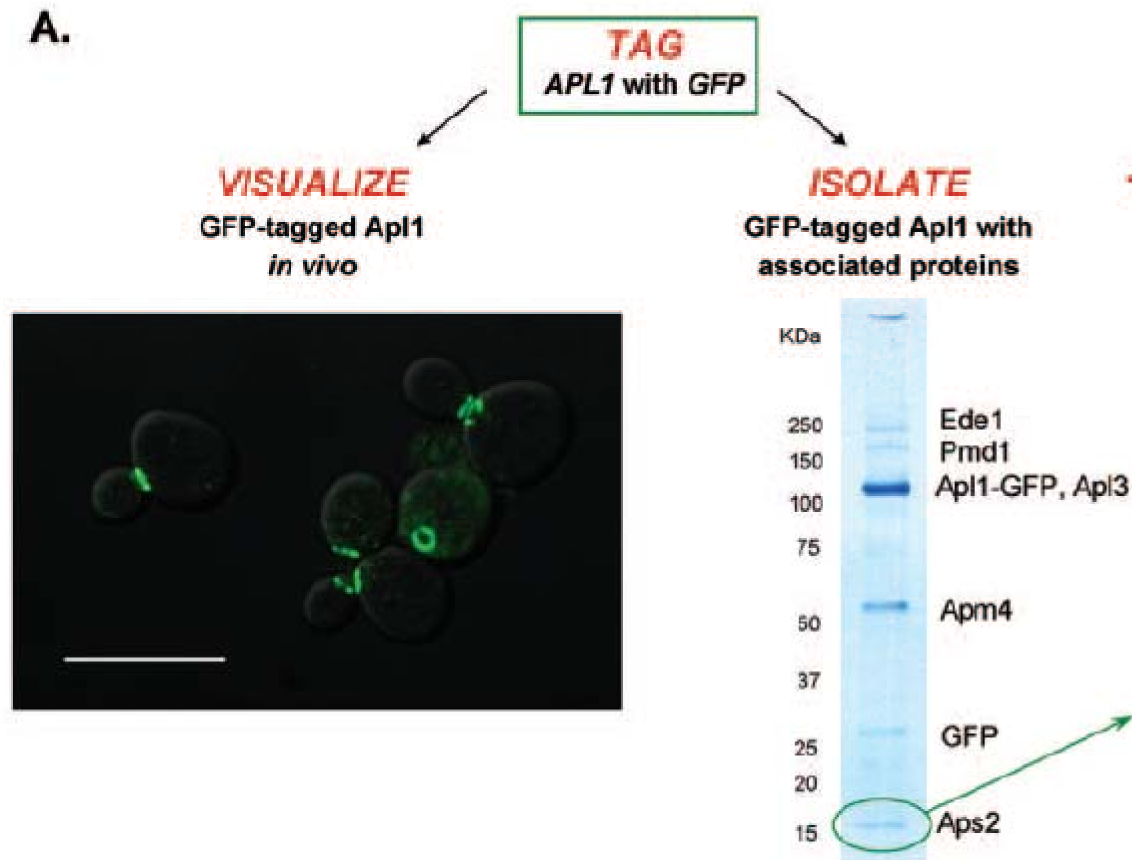
Protein Tags

- FLAG-tag, HA-tag, MYC-tag: small peptide tags that can then be used for immunoprecipitation
- His-tag: poly-His tag allows purification by nickel metal ion affinity chromatography.
- Biotin: Biotin-avidin/streptavidin interaction very strong, allowing stringent washing during purification protocol. Can have problems with successfully eluting biotinylated components from avidin/streptavidin columns, so monomeric avidin columns may be preferable.
- GST, MBP: Whole protein sequences that can be added to the termini of a protein. GST is purified by a glutathione column, MBP is purified with an amylose column.

For subsequent MS analysis smaller tags are best, otherwise you can spend more time analyzing your tag than the protein of interest!

Fluorescent Tag

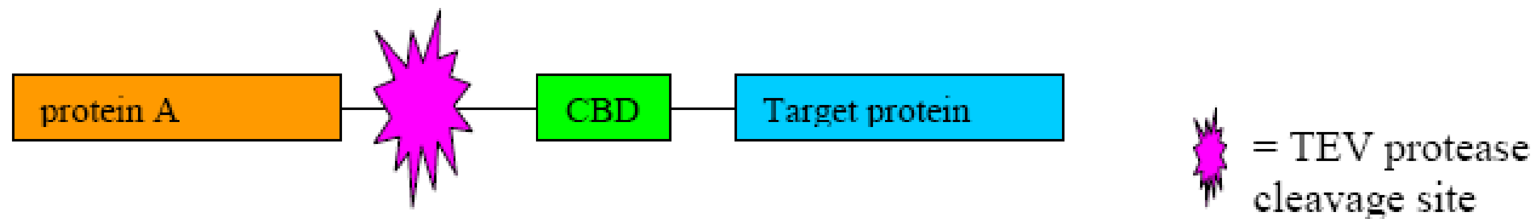
- Use GFP (or YFP) as tag, allowing protein visualization *in vivo* and IP using anti-GFP antibody



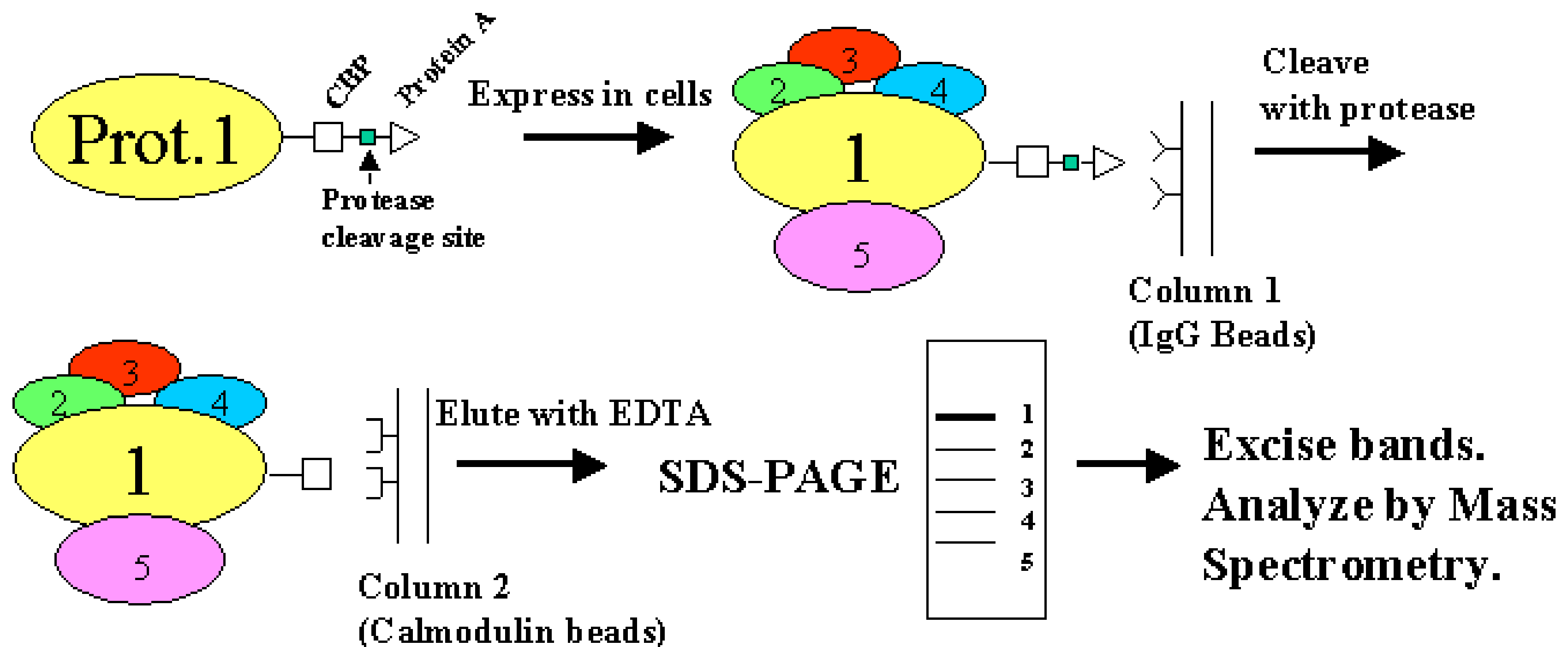
Cristea, I.M. et al. *Mol Cell Proteom* (2005) 4 1933-1941

Tandem Affinity Tag Purification (TAP-Tags)

- Why use one tag when you can use two?
  - Gives much lower non-specific binding.



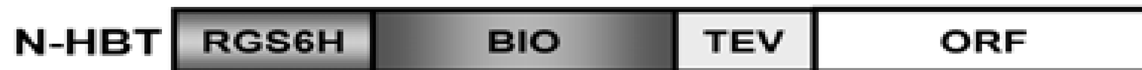
Tandem Affinity Purification



TAP-Tag variants

•Many combinations of tags have been used (and cleavage linkers).

Glu-Asn-Leu-Tyr-Phe-Gln↓Gly	TEV protease
Leu-Val-Pro-Arg↓Gly-Ser	Thrombin
Leu-Glu-Val-Leu-Phe-Gln↓Gly-Pro	PreScission™ protease



•Combination of a His-tag and a sequence that becomes *in vivo* biotinylated in prokaryotes and eukaryotes.

- •Allows purification under completely denaturing conditions.

Tagwerker, C. et al. *Mol Cell Proteomics* (2006) **5** 737 - 748.

The Usual Suspects

- Even with tandem affinity purification strategies there are still non-specific interactors, so you need to know which these are.
- The non-specific interactors are often the same for many different IP strategies:

- Structural proteins: Actin, Myosin, Tubulin, Vimentin, Spectrin (if from brain)…
- Glycolytic enzymes: Pyruvate kinase, Glyceraldehyde-3-phosphate dehydrogenase, Triose phosphate isomerase…
- Ribosomal Proteins: 40S Ribosomal protein…
- HSP/Chaperones: HSP70, HSP60, HSP…

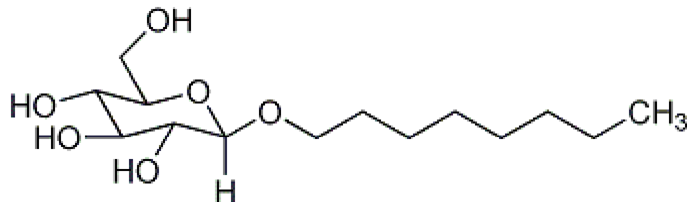
Other 'Friends'

- Keratins (1,2,6,9,10): The most studied proteins in proteomics experiments! Skin and hair.
    - Wear gloves.
    - Stain gels in a box with a lid.
    - Don't lean over gel, including when cutting bands out.
- Bovine Serum Albumin: From medium, from enzyme preps, western blotting…
- Casein: Major milk protein – western blots.
- Avidin, IgG: affinity columns leak!
- Rubber tree proteins: from latex gloves
- Dermicidin: added to gloves

Detergents and Mass Spectrometry

Wash your ions carefully and use fabric softener only with multiply charged ions for best effect…

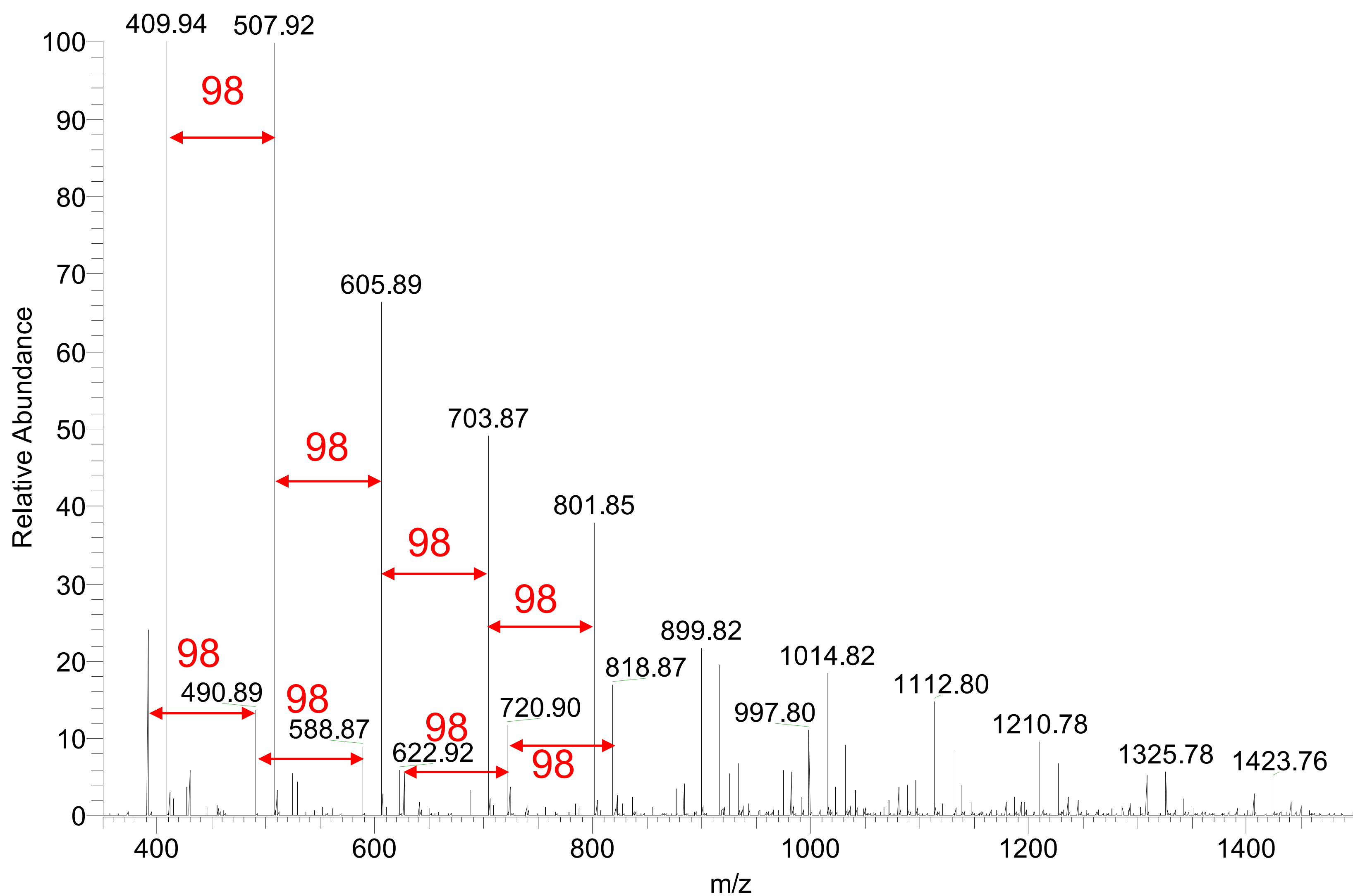
- Detergents can be ionic, or non-ionic and they can be single components or polymers.
- Ionic detergents are readily detected in mass spectrometry, so are not compatible with mass spectrometry (if you want to see anything else in your sample).
- Polymeric detergents are bad because they may be separated by chromatography, so versions can elute throughout an LC-MS run.
- The best detergents to use are therefore, non-ionic single components:
    - e.g. N-Octyl-beta-D-Glucopyranoside (N-Octyl glucoside)



Salts and Mass Spectrometry

- Molecules in mass spectrometry are typically analyzed in a protonated state; i.e. $[M+H]^+$
- If metal salts are present, then metal adducts can be formed; e.g. $[M+Na]^+$ or $[M+K]^+$.
    - Having protonated and metal adducts makes the spectrum more complicated to interpret.
    - Metal adducted peptides do not fragment as readily as protonated, making identification by fragmentation analysis difficult.
- If a salt crystallizes in the electrospray capillary it can block flow, meaning it has to be replaced and sample is lost. It can also block the orifice into the mass spectrometer.
- Some salts form clusters (most notably phosphate), and these can drown out the signal from all other components in the sample.

Phosphate Clusters



Salts and Mass Spectrometry

The Bottom Line

- Salts are bad for mass spectrometry, so should be avoided whenever possible.
    - If they are required, then volatile salts are best; e.g. ammonium bicarbonate, ammonium formate…
- Most salts can be removed prior to MS analysis using a reverse phase clean-up step.

Other ‘Problem’ Components

Trifluoroacetic acid (TFA)

- Ion pairing agent used in chromatography to ‘sharpen’ peaks.
    - In ESI, ionizes well, forms clusters and causes ion suppression.

Glycerol

- Stabilizing agent for protein storage. Viscous, involatile.
    - In ESI, viscosity affects ESI spray/plume formation.
    - In MALDI, sample never dries, preventing matrix crystallization.

Inhibitors

Your sample can continue to change after you isolate it.

- Proteases may still be active.
  - In the MS analysis of peptides you are assuming peptides are formed by the protease you add.
    - Identifying ‘non-specific’ cleaved peptides is more difficult.
- If you are interested in phosphorylation, you should add phosphatase inhibitors.

Variability Introduced by Sample Handling

- Experimental design is very important in determining biological differences between samples.
- Many studies (especially clinical) are carried out to compare two sets of samples (e.g. normal vs cancer; + vs – drug treatment…)
  - There have been several studies where results have been shown to group more by 'person who prepared sample'; 'number of freeze/thaw cycles'; 'order in which samples were run' rather than by the expected groupings.

Prakash, A. *Mol Cell Proteom* (2007) **6** 1741-1748

What else to do prior to MS?

Desalting

- Almost invariably required for solution digests
- Not necessary for LC-MS analysis of in-gel digests
- May be necessary for MALDI analysis of in-gel digests.

- Usually performed using pipette tips with C18 resin packed into end of tip – ‘zip tips’

Concentrating

- We typically inject 1-5ul of sample, so if your sample is in a larger volume, it will need concentrating

- Concentrating and desalting can be performed at the same time using a zip tip.

Common ‘Complaints’

But you didn’t find my protein of interest! and I know it is there because I have a band on a western…

(Variant 2): You didn’t find my phosphorylation site! and I know it is phosphorylated because I have a band on a western…

- There is a great deal of variety in the quality of antibodies.
    - Some are more specific than others
    - Some are much more sensitive than others

- A western blot can detect very low levels of protein; below the level that mass spectrometry can detect.

Mass spectrometry can reliably tell you if something is there, but it cannot tell you if something is not there.

Biological Samples can be ‘messy’ – a real case example

Proteomic studies on *Schistosoma mansoni*

* “waterborn” parasitic disease, 200 M people infected worldwide through their skin
* Intermediate host – *Biomphalaria glabrata*

* Enzymes in cercaria (worm) secretion are the key to infection

It’s a well-characterized, simple mixture…

1. cercaria released from snails, rinsed
2. secretion triggered with skin lipid
3. solution collected, concentrated
4. 1D SDS-PAGE protein-fractionation
5. in-gel digestion with trypsin
6. LC/MS analysis

What was Found

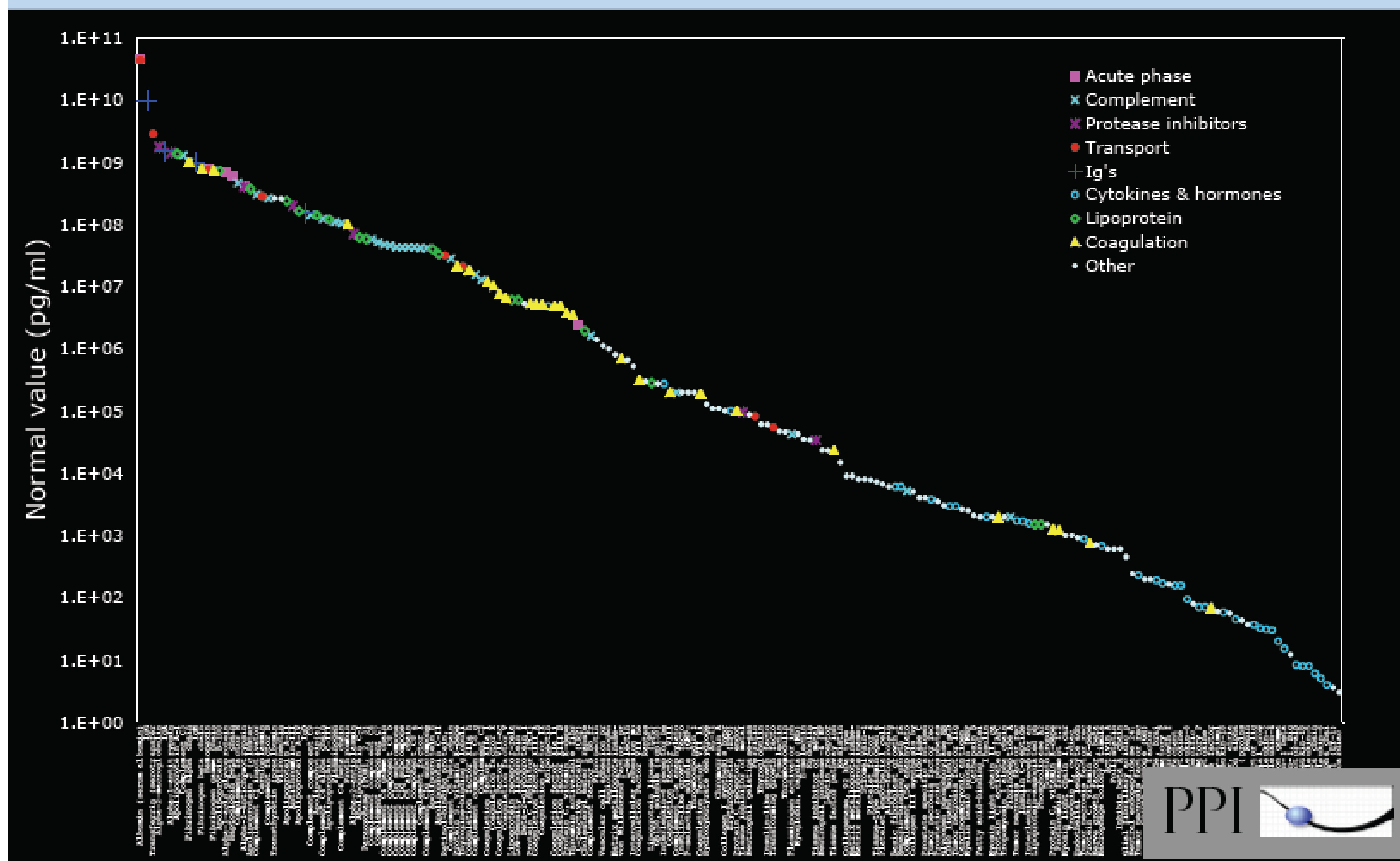
- cercaria surface proteins
- cercaria elastases, and glycolytic enzymes
- snail immune response proteins
- photosynthetic proteins
- human keratin
- bovine serum albumin!

- Contamination from the worm
- SECRETION – the real stuff!
- Contamination from the host
- Lettuce - should we starve the snails?
- Technician
- Lab – too many westerns….

Improved isolation protocol, 40-fold increase in elastase activity!

Knudsen et al., *Mol. Cell. Proteomics,* (2005) **4**: 1862 - 1875.

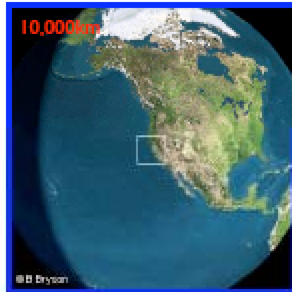
Proteomics from Serum



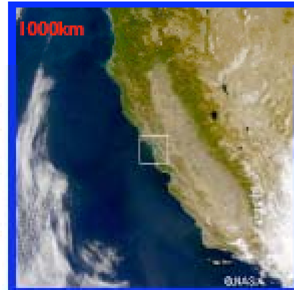
Slide from L. Andersen PPI

10^{10} Really Is Wide Dynamic Range

(Here on a linear scale)



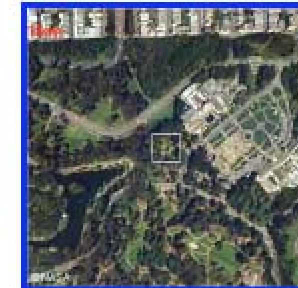
10



9

8

7



6

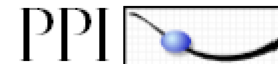
5

4

3

2

1



Slide courtesy Bruno Domon, ETH Zurich

Albumin (serum albumin)
IgG
Transferrin (serotransfer
Alpha-2-macroglobulin
IgA
Alpha-1-antitrypsin
Apolipoprotein A-I
Complement C3
Fibrinogen alpha chain
IgM
Fibrinogen beta chain
Hemopexin
Fibrinogen gamma chain
Apolipoprotein B-100
Alpha-1-acid glycoprotein
Haptoglobin beta chain
Complement factor H
Alpha-1-antichymotrypsin
Apolipoprotein J (cluster
Complement C1 inactivator
Ceruloplasmin
Complement factor B
Transthyretin (prealbumin
Vitronectin
Apolipoprotein A-II
Antithrombin-III
Apolipoprotein A-IV
IgD
Complement component C9
Apolipoprotein(a)
Complement C4 alpha chain
Apolipoprotein C-III
Plasminogen
Complement C4 beta chain
Prothrombin

95% of serum protein is Albumin.

99% of serum protein consists of 20 proteins.

How do you find interesting proteins in this background?

- Multiple affinity antibody columns have been developed to remove 5, 8, 10, 11… most abundant proteins.
  - When you remove these, what else do you remove at the same time? (e.g. albumin is designed to bind and transport proteins).

- Use an affinity step that selectively enriches a subset of the proteins (not including the most abundant); e.g. glycoprotein enrichment.

- If you know the protein/peptide you are looking for, there are mass spectrometry methods (MRM) that can allow detection of these at low levels in a complex background.

Bottom line at the moment is:
No good strategy has been developed. A lot of energy has been put into biomarker discovery in plasma/serum with little tangible success.

Cross-Linking

•Cross-linking can be used for different goals, and the type of cross-linker to use is dependent on what you want to discover.

•Capture weak/transient interactors.

  - •Specificity of cross-linking not important.
  - •Ability to reverse cross-linking important.

•Identify binding interfaces between proteins in a complex.

  - •Few highly specific crosslinks.
  - •Non-reversible.
  - •Ability to enrich cross-linked peptides useful.

Weak/Temporary Interactions

Formaldehyde

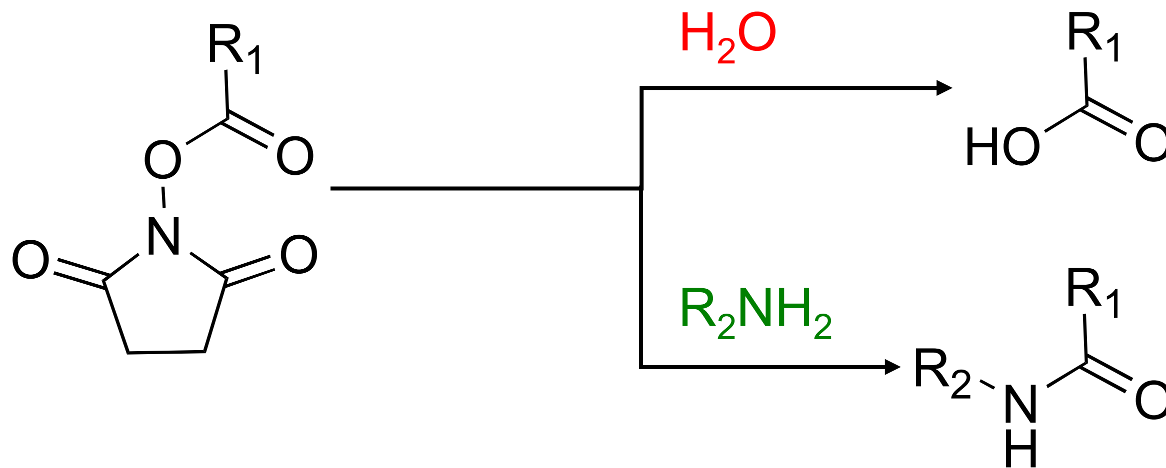
- Provides short cross-links
- Cheap
- Reversible (by heating)
- Soluble in water – can be used *in vivo*[1]

- Need to calibrate amount of formaldehyde to add to prevent formation of large, insoluble cross-linked pellets.

[1]Schmitt-Ulms G, et al, Time-controlled transcardiac perfusion cross-linking for the study of protein interactions in complex tissues.
Nat Biotechnol. (2004) **22**(6):724-31

Crosslinking for Protein Complex Interface Analysis

- Most reagents target primary amines (lysines and free N-termini).
- NHS-ester crosslinkers can undergo both aminolysis and hydrolysis.



- For cross-link formation you want aminolysis.
  - Problem: there is a lot of water around!

Cross-linking Products

•True Intermolecular Crosslinks are formed in very low yield.

•Reaction mixture consists largely of unmodified and "dead-end"-modified proteins.

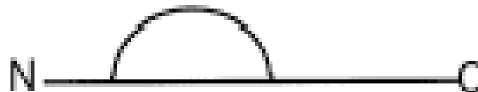
J Am Soc Mass Spectrom 2003, 14, 834–850

SCHILLING ET AL.

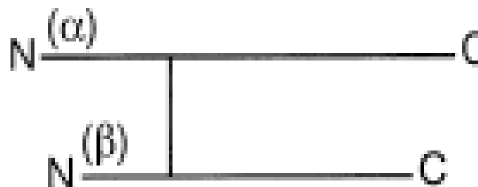
(a) Single modifications

Type 0
'deadend'

Type 1
'intrapeptide'



Type 2
'interpeptide'



Finding and Identifying Crosslinked Products

Not easy…

Can analyze sample before and after cross-linking to find peaks only present after crosslinking.

How do you find which ones are 'Type 2's'?

Software: MS-Bridge (part of Protein Prospector)

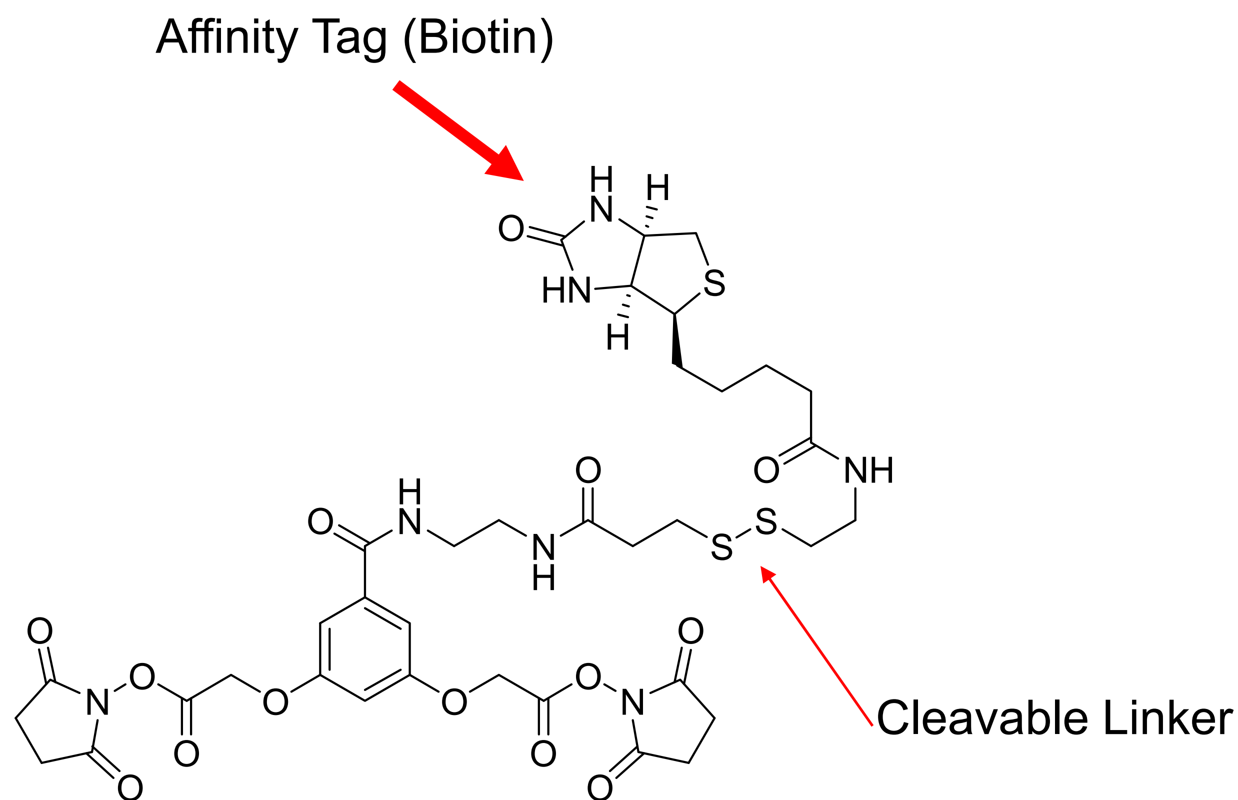
1. Give it the sequences of the proteins in the complex
2. Tell it the crosslinking reagent (so it knows the mass and specificity)
3. Give it a list of observed masses

It will return a list of potential cross-linked peptides for each mass.

Batch-Tag (Protein Prospector again)

Can search for peptides with unknown mass modification (e.g. the mass of a cross-linked peptide).

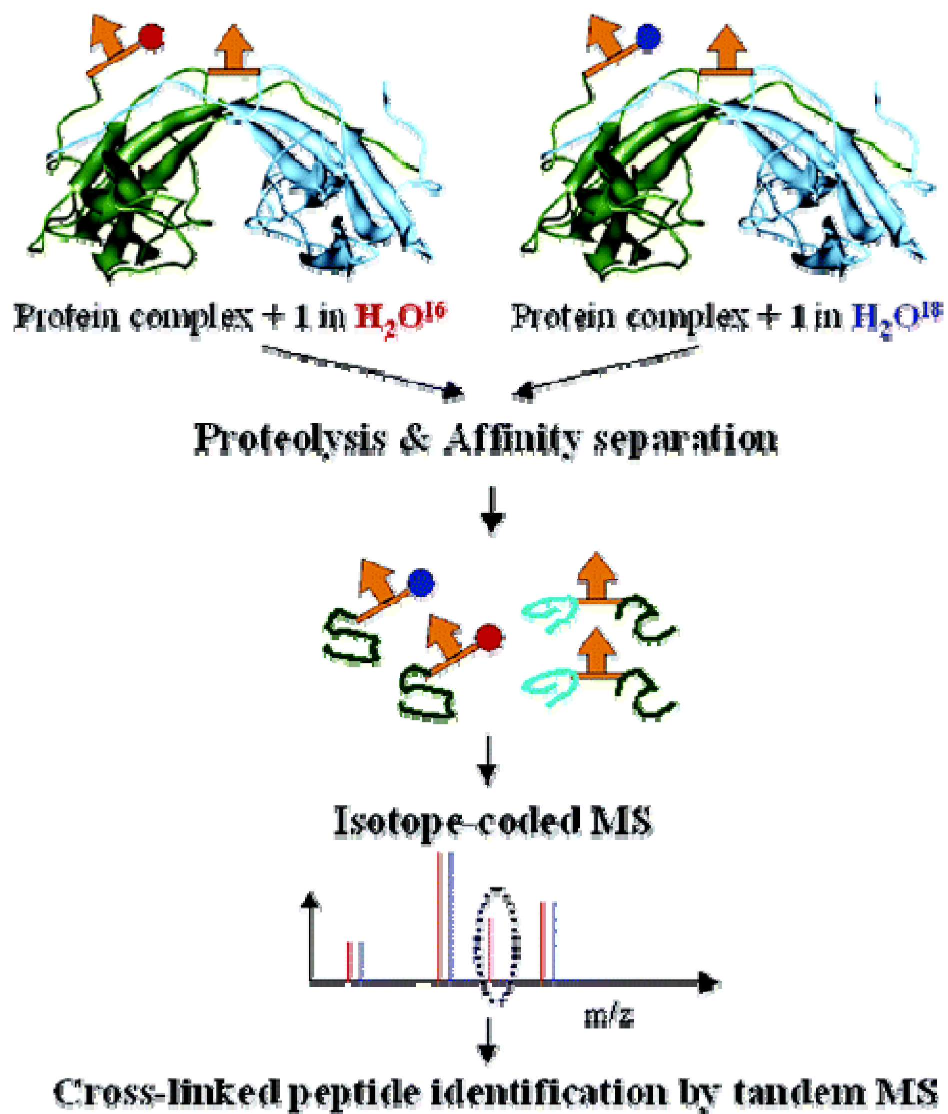
ICAT-XL, Affinity crosslinker



Chu, F. *et al. JACS* (2006) **128** 10362-10363

ICAT-XL: Identifying Crosslinked Peptides

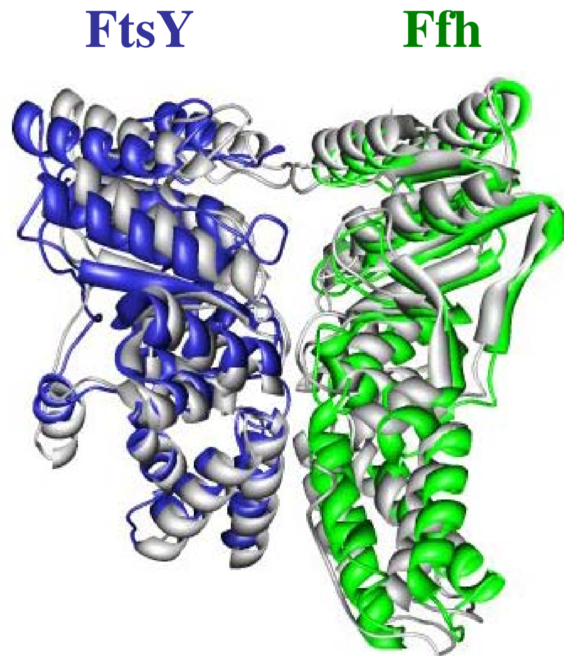
- Mixture of ^{16}O and ^{18}O incorporated into hydrolysis products (type 0 crosslinks)
    - Will give peaks with pairs 2 Da apart
- Crosslinked peptides do not incorporate oxygen, so will be a 'singlet'.



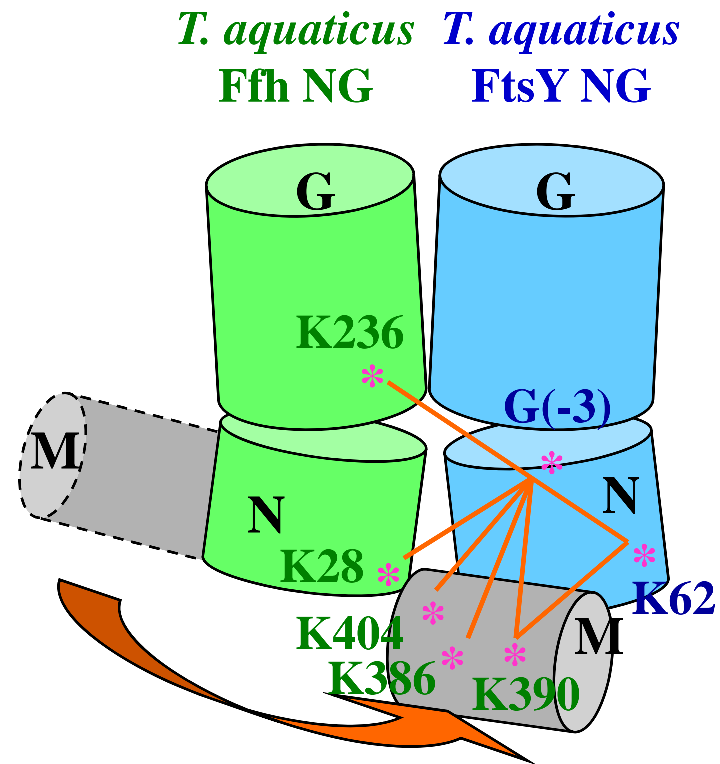
Chu, F. *et al. JACS* (2006) **128** 10362-10363

SRP complex w/ receptor

A **Model Overlaid w/ Xtal Struct.**



B **Domain Rearrangement**



Chu, F. *et al. PNAS* **101**, 16454-9 (2004).

Summary

- Mass spectrometry can be used to solve many proteomics problems.
- Compatible with many sample purification protocols.

- Can employ separation at both protein and peptide level.

- A cleaner sample to start with saves a lot of work later
    - When pulling down complexes, reduce non-specific interactors.
    - Try not to contaminate your samples.
    - Be careful of detergents.
    - Need to get rid of non-volatile salts.