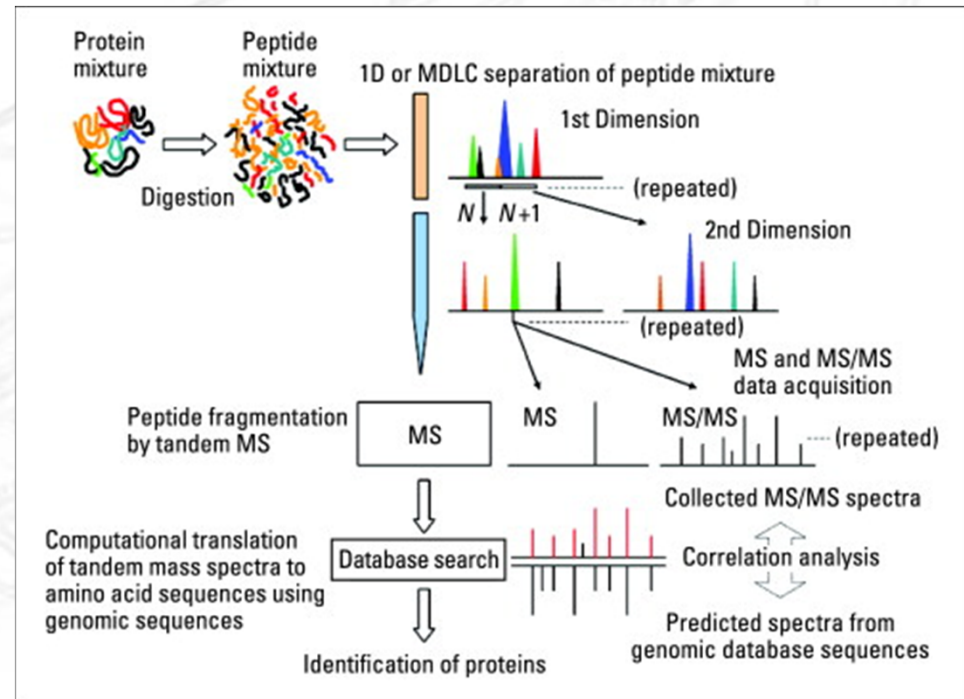
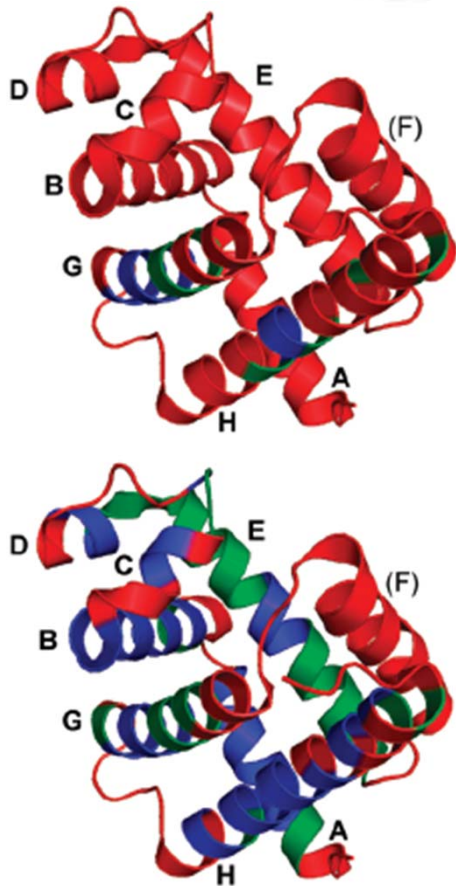
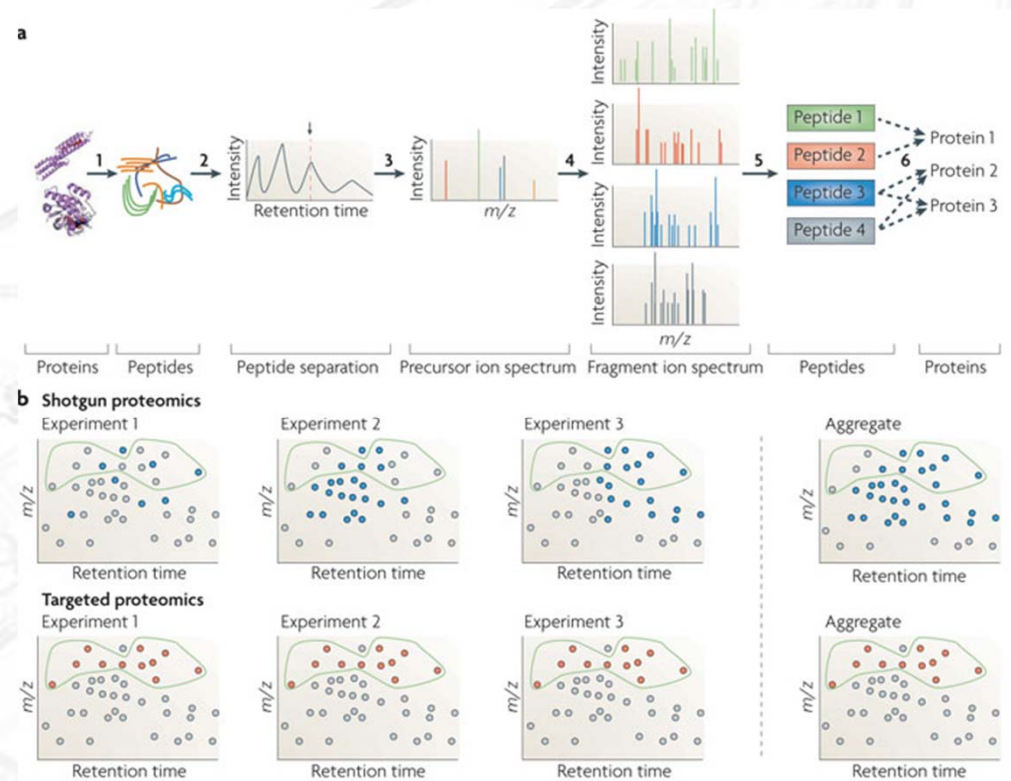


Week 10: Proteomics and Protein Function



Last Time...



Pawson Group: Cell Signaling and Cancer

Cell-Specific Information Processing in Segregating Populations of Eph Receptor Ephrin–Expressing Cells

Claus Jørgensen,¹ Andrew Shemman,^{1,2} Ginny I. Chen,^{1,2} Adrian Pasculescu,¹ Alexei Poliakov,³ Marilyn Hsiung,¹ Brett Larsen,¹ David G. Wilkinson,³ Rune Linding,^{4*} Tony Pawson^{1,2*}

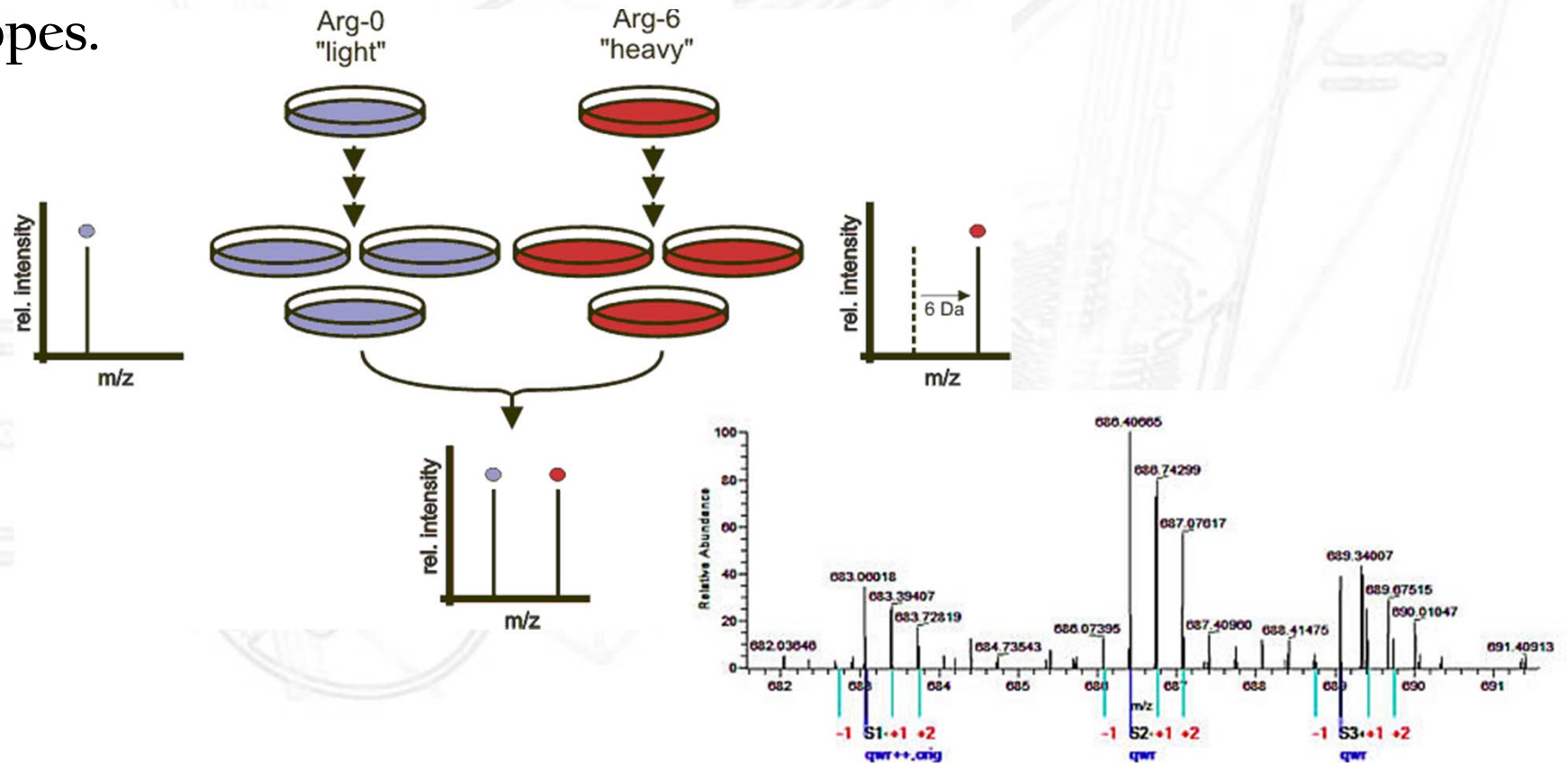
- This paper is concerned with cell signaling, particularly ‘clumping’ and ‘unclumping’ signals initiated by cell-to-cell contact.
- These interactions are governed by ‘Ephrin’ transmembrane proteins that become phosphorylated on the inside of the cell after interacting with each other on the outside of the cell.

Ephrins

- Ephrins are **receptor tyrosine kinases**, meaning they are receptors that **directly phosphorylate** proteins in the cell when they are phosphorylated.
- Ephrins are expressed at very low level in adult cells. Their primary role is in development:
 - Embryonic development (segmentation)
 - Neuronal development (axon guidance)
 - Cell migration (mainly during development)
 - Angiogenesis (formation of blood vessels)
 - Cancer (survival factor, help initiate metastasis?)

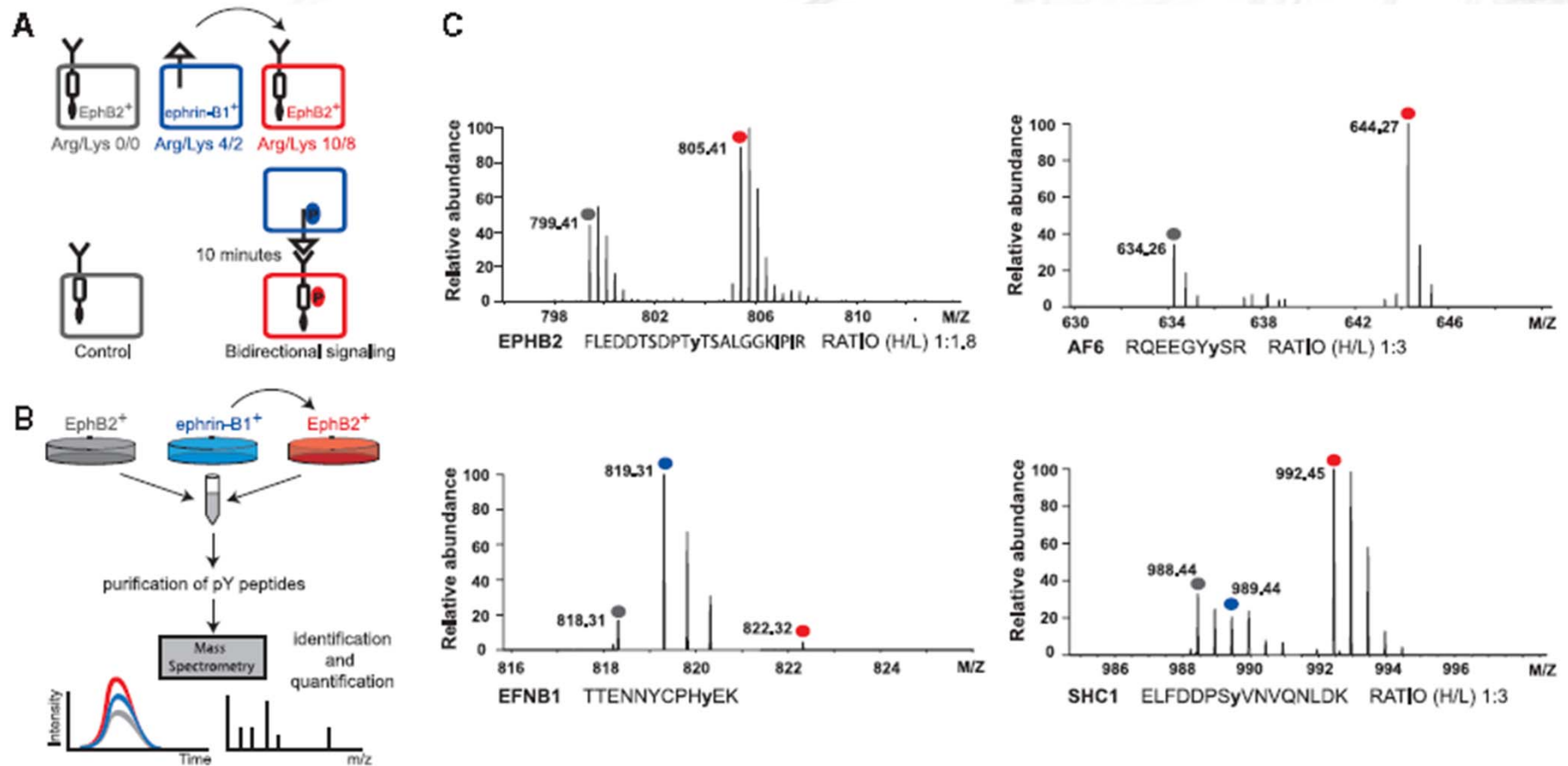
Quantitative Proteomics with Stable Isotope Labeling of Amino Acids in Cell Culture (SILAC)

- The study is designed to map phosphorylation on all proteins regulated by ephrin-B1 → ephrin-B2 interaction
- To do this, they use a quantitative proteomics technique called **SILAC** in which Arg and Lys residues are labeled with heavy isotopes.



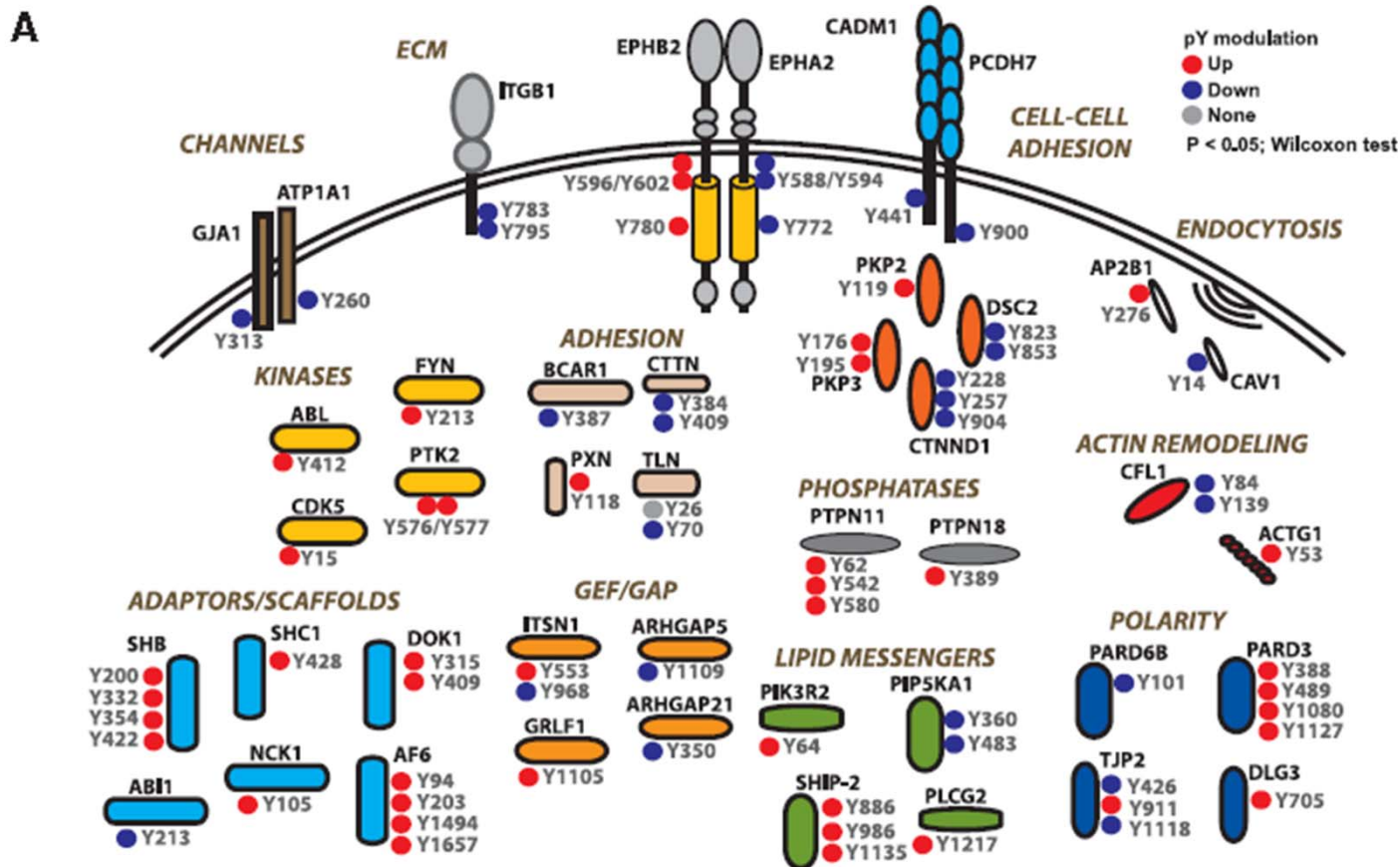
Ephrin Experiment Setup

- This is the experimental workflow for studying cell/cell interactions. MS = Orbitrap



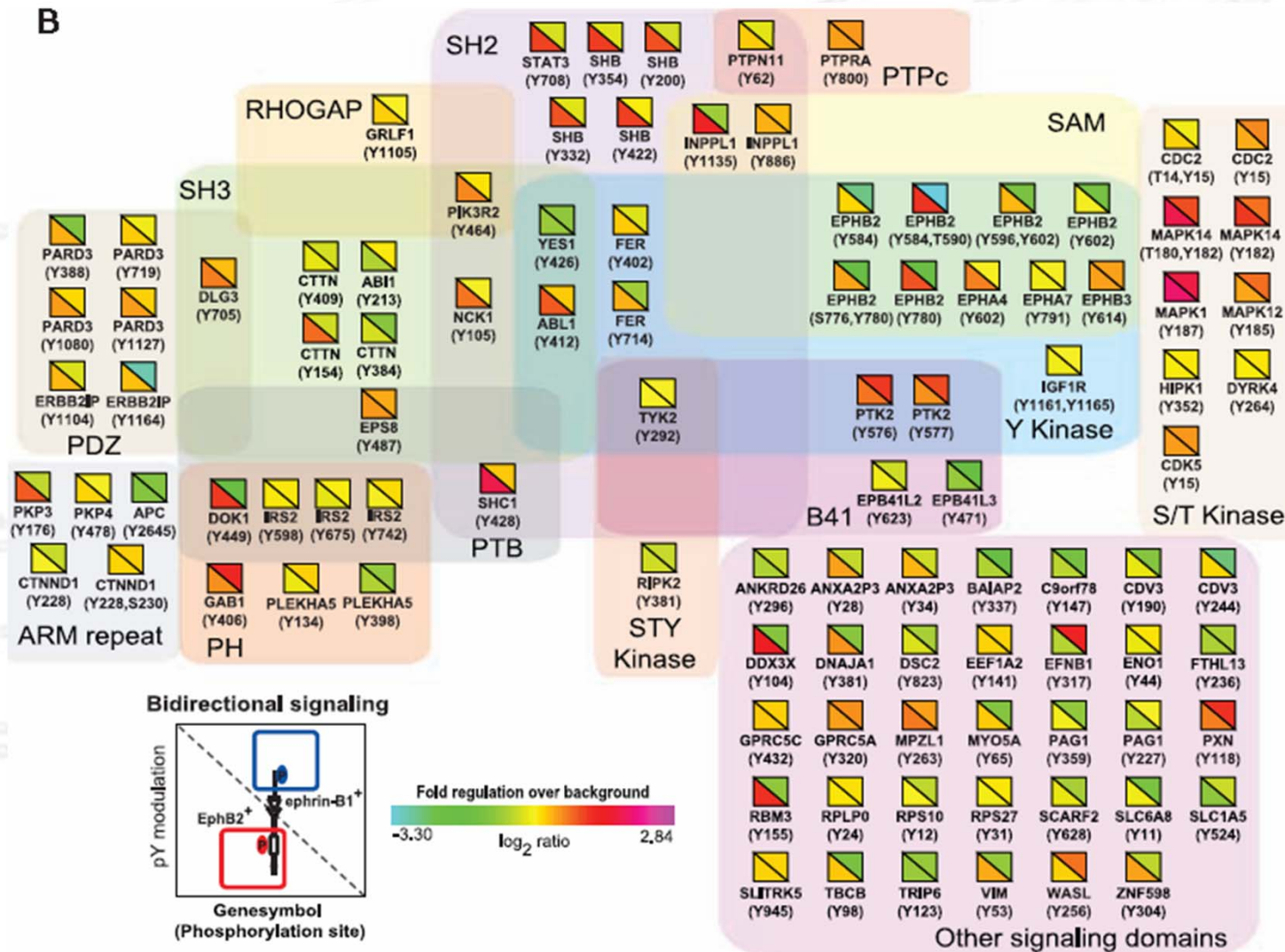
Ephrin Signaling Results

- In typical proteomics fashion, you end up with a huge map of proteins that are more-or less phosphorylated after cell/cell interactions (Eph-B2+ cells):



More Ephrin Signaling Results

- And an even broader map:



Example 2: The Siu Group

- Why stray far from home? A perfectly good example of large scale proteomics is right here in the Siu group, looking for biomarkers in various cancers...
- The Siu group is linked to ABSciex and so they use their iTRAQ or mTRAQ labeling technology for quantitation.

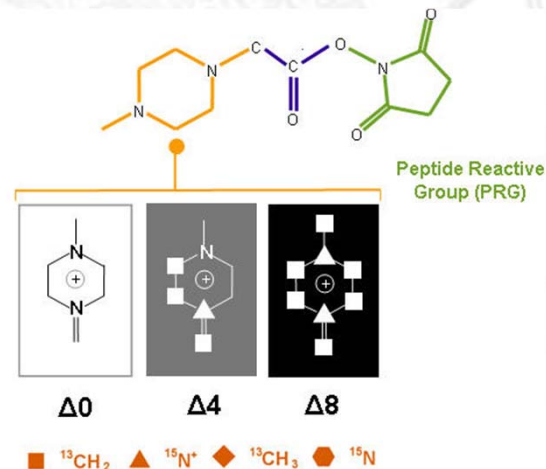
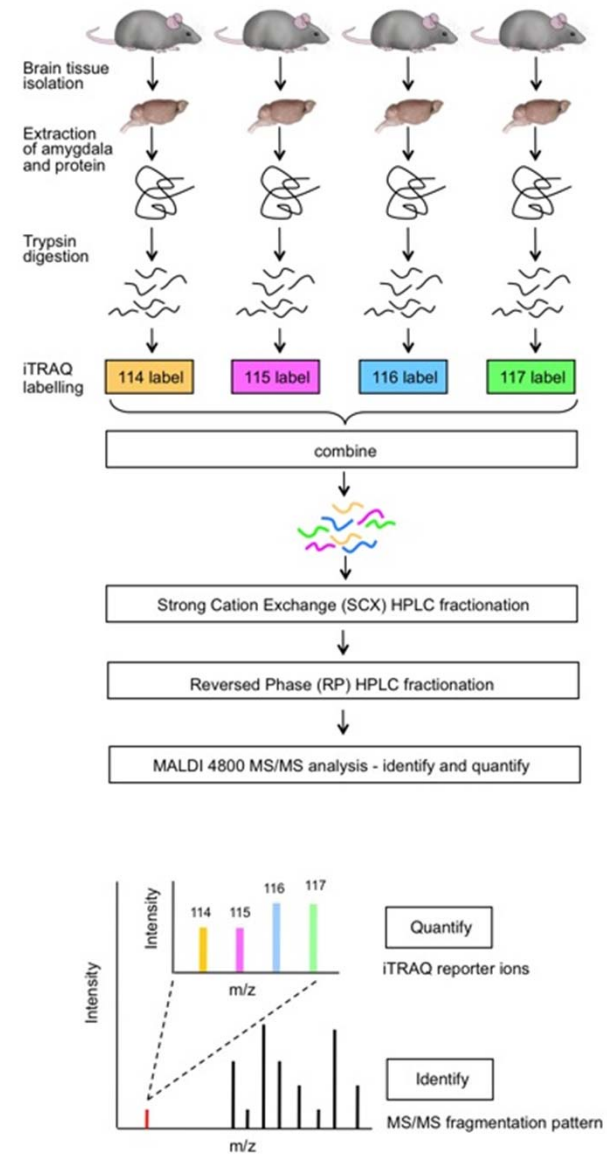


Figure 1: Chemical Structure of mTRAQ® Reagents

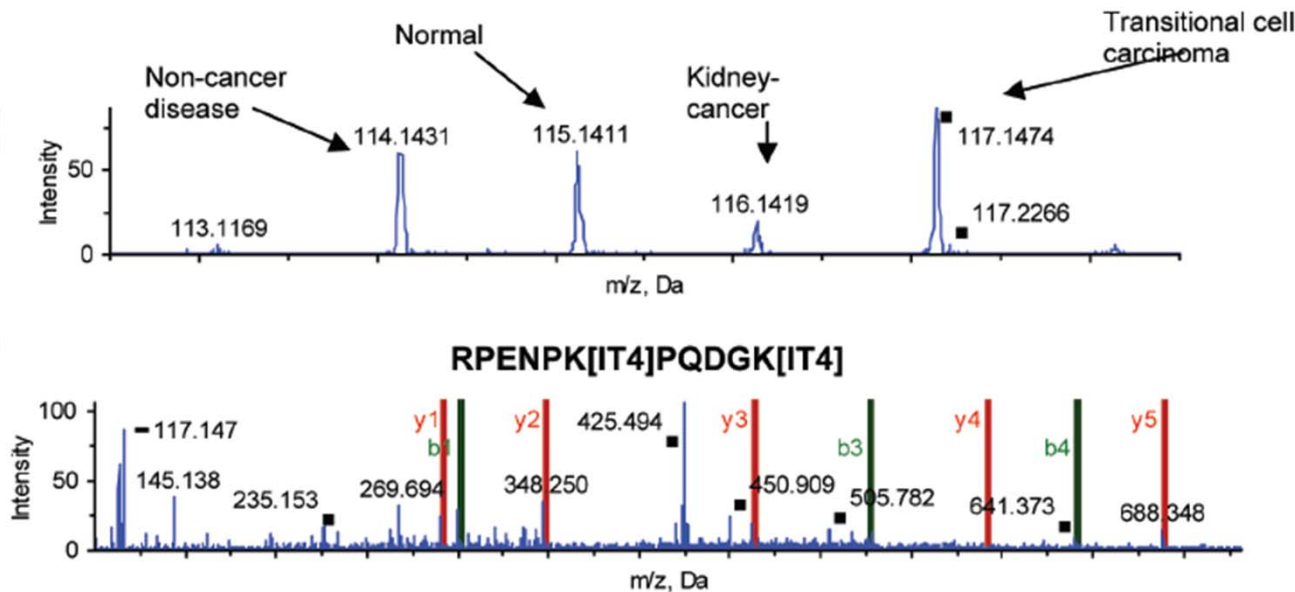


Cancer Biomarker Discovery with iTRAQ

Differential Protein Expressions in Renal Cell Carcinoma: New Biomarker Discovery by Mass Spectrometry

K. W. Michael Stu,[†] Lerol V. DeSouza,[†] Andreas Scorilas,[‡] Alexander D. Romaschin,^{§,||}
R. John Honey,[⊥] Robert Stewart,[⊥] Kenneth Pace,[⊥] Youssef Youssef,[§] Tsz-fung F. Chow,[§] and
George M. Yousef^{*,§,||}

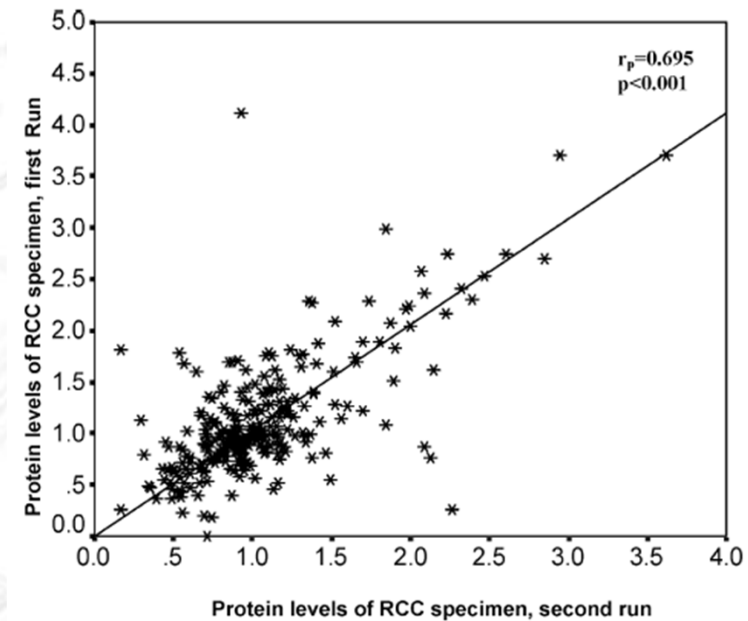
- Here's a quick example of iTRAQ use to discover biomarkers for renal cell carcinoma



Renal Cancer Cont.

- One of the challenges of quantitative proteomics is significant run-to-run fluctuations in protein ‘intensity’:

Protein	Swiss Prot ID	Craven ²²	Perego ²³	Sarto ¹⁸	Shi ⁵
60 kDa heat shock protein	P10809				
60S acidic ribosomal protein P2	P05387				
78 kDa glucose-regulated protein	P11021				
Albumin	P02768				
Alpha Crystallin B chain	P02511				
Alpha enolase	P06733				
Aminoacylase	Q03154				
ATP synthase D chain	P30049				
Calreticulin	P27797				
Cathepsin B	P07858				
Cathepsin D	P07339				
Elongation factor 1-beta	P24534				
Elongation factor 2	P13639				
Glutamate dehydrogenase	P00367				
Glutathione S transferase	P09211				
Heat shock 27 kDa protein	P04792				
Heat shock cognate 71 kDa protein	P11142				
Heterogeneous nuclear ribonucleoprotein A2/B1	P22626				
Lactate dehydrogenase A	P00338				
Lactate dehydrogenase B	P07195				
Peptidyl-prolyl cis-trans isomerase B	P23284				
Protein disulfide isomerase	P07237				
Septin 2	Q15019				
Tropomyosin alpha 3 chain	P06753				
Total Proteins identified:	243	34	28	27	23



- Agreement with other studies, however... would **anything** agree?

Lots of Proteins

- Over 937 proteins were identified... a partial list of 'dysregulated' ones is shown here:

no.	protein name	gene symbol	Swiss-Prot ID	fold change ^a	regulation ^a
1	KIAA1865 protein	C14orf4	Q9H1B7	4.8624	UP
2	tyrosine 3-hydroxylase 5-monoxygenase activation protein	YWHAH	Q04917	4.2912	UP
3	CAPNS1 protein	CAPNS1	P04632	3.9777	UP
4	Complement component 1 inhibitor	SERPINC1	Q96FE0	3.66505	UP
5	UDP-glucose 6-dehydrogenase	UGDH	Q60701	3.3497	UP
6	L-lactate dehydrogenase A chain	LDHA	P00338	3.32095	UP
7	Nicotinamide N-methyltransferase	NNMT	P40261	3.2572	UP
8	Hypothetical protein DKFZp666H13168	GSTO1	Q7Z3T2	3.1252	UP
9	Poly(RC)-binding protein 2, isoform b	PCBP2	Q6PKG5	3.0866	UP
10	ADP-ribosylation factor 3	ARF3	P61204	3.0753	UP
11	Calnexin precursor	CANX	P27824	2.8414	UP
12	3'(2',5'-bisphosphate nucleotidase 1	EPN1	Q95861	2.7755	UP
13	Hypothetical protein DKFZp547T2313	FABP7	Q9H047	2.7722	UP
14	Catechol O-methyltransferase, membrane-bound form	COMT	P21964	2.7512	UP
15	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	GAPDH5	O14556	2.6716	UP
16	Hypothetical protein	ANKA4	Q6P452	2.6156	UP
17	Hypothetical protein DKFZp666I04222	SERPINC6	Q7Z2Y7	2.5934	UP
18	Echinoderm microtubule-associated protein-like 4	EMIL4	Q9HC35	2.575	UP
19	vimentin-human	VIM	P08670	2.5615	UP
20	Cytoplasmic dynein intermediate chain 2C	DYNC1I2	Q7Z4X1	2.5234	UP
21	Calpain small subunit 1	CAPNS1	P04632	2.5198	UP
22	Glutathione S-transferase	GSTA2	P09210	2.49785	UP
23	Alpha Crystallin β	CRYAB	P02511	2.48945	UP
24	ALDOC protein	ALDOC	Q6P0L5	2.4321	UP
25	Rab GDP dissociation inhibitor alpha	GDI1	P31190	2.4198	UP
26	PRKAR2A protein	PRKAR2A	Q9BUB1	2.4146	UP
27	Chloride intracellular channel protein 1	CLIC1	O0C299	2.4112	UP
28	Pre-B-cell colony enhancing factor 1, isoform b	PBEF1	Q8WVW5	2.4046	UP
29	Annexin A5	ANXA5	P08758	2.3679	UP
30	glyceraldehyde-3-phosphate dehydrogenase	GAPDH	P04406	2.3475	UP
31	Endothelial cell growth factor 1 (platelet-derived)	ECGF1	P19971	1.834	UP
32	Major vault protein	MVP	Q14764	1.6983	UP
33	Adipose differentiation-related protein	ADFP	Q95541	1.6629	UP
34	60 kDa heat shock protein	HSPD1	P10809	0.4896	DOWN
35	ATP synthase delta chain	ATP5D	P30040	0.4879	DOWN
36	Coronin 1A	CORO1A	P31146	0.4878	DOWN
37	GCSH protein	GCSH	Q6IAT2	0.4873	DOWN
38	TAGLN protein	TAGLN	Q6FIS2	0.4862	DOWN
39	Ksp-cadherin	CDH16	Q6UW93	0.4818	DOWN
40	splicing factor, arginine/serine-rich 2	SFRS2IP	Q95590	0.4802	DOWN
41	Zinc finger protein 207	ZNF207	Q43670	0.4726	DOWN
42	40S ribosomal protein S17	RPS17	P08708	0.4702	DOWN
43	Secreted cement gland protein XAG-2 homologue	AGR2	Q95994	0.46395	DOWN
44	Thymosin beta-4	TM6B4X	P62328	0.4631	DOWN
45	CKB protein	CKB	Q6FG40	0.4589	DOWN
46	Calmodulin	CALM1	Q13942	0.4562	DOWN
47	Hypothetical protein FLJ46684	C9orf58	Q6ZK40	0.4479	DOWN
48	Elongation factor Tu	TUFM	P49411	0.4476	DOWN
49	Tumor protein p53 inducible protein 3	TP53IIP3	Q9BWB8	0.4383	DOWN
50	Calmodulin	CALM1	P62158	0.4245	DOWN
51	Hypothetical protein DKFZp586K2222	TPM1	Q9Y427	0.4187	DOWN
52	creatine kinase-B	CKB	P12277	0.4162	DOWN
53	ATP synthase beta chain	ATP5B	P06576	0.41455	DOWN
54	Hemoglobin beta	HBB	Q6R7N2	0.4127	DOWN
55	Histone H2A	HISTH2A	Q7L7L0	0.4112	DOWN
56	AP endonuclease 1	APEX1	P27695	0.4085	DOWN
57	Acyl-CoA dehydrogenase, medium-chain specific	ACADM	P11310	0.3938	DOWN
58	Nonhistone chromosomal protein HMGB-17	HMGB2	P05204	0.39295	DOWN
59	HES1 protein	C21orf33	P30042	0.3925	DOWN
60	Eukaryotic translation initiation factor 3 subunit 3	EIF3H	O15372	0.3894	DOWN
61	Calreticulin	CALR2	Q96BK4	0.3772	DOWN
62	Programmed cell death protein 5	PDCD5	O14737	0.3623	DOWN
63	Chromosome 10 open reading frame 65	C10orf65	Q66XE5	0.3616	DOWN
64	adenylate kinase 3 alpha	AK3	Q9UIJ7	0.3613	DOWN
65	LOC112817 protein	C10orf65	Q66EV5	0.3554	DOWN
66	Ubiquinol-cytochrome-c reductase complex core protein I	UQCRC1	P31930	0.3531	DOWN
67	Plastin 3	PLS3	Q66V16	0.3394	DOWN
68	Membrane associated protein SLP-2	STOML2	Q9UIZ1	0.3341	DOWN
69	MHC class II antigen	HLA-DRB1	Q9MYD9	0.3294	DOWN
70	Reticulocalbin 1 precursor	RCN1	Q15293	0.312	DOWN
71	DNA-binding protein B	YBX1	P67809	0.31	DOWN
72	Cytochrome c	CYCS	Q6NUR2	0.2901	DOWN
73	NADH-ubiquinone oxidoreductase 13 kDa-A subunit	NDUF6	O75380	0.2759	DOWN
74	Lupus La protein	SSE	P05455	0.2709	DOWN
75	Pyruvate dehydrogenase E1 component beta subunit	PDHB	P11177	0.2662	DOWN
76	FERM, RhoGAP, and pleckstrin domain protein 1, isoform 1	FARP1	Q9Y4F1	0.2628	DOWN
77	Mitochondrial aldehyde dehydrogenase 2	ALDH2	Q6IV71	0.2603	DOWN

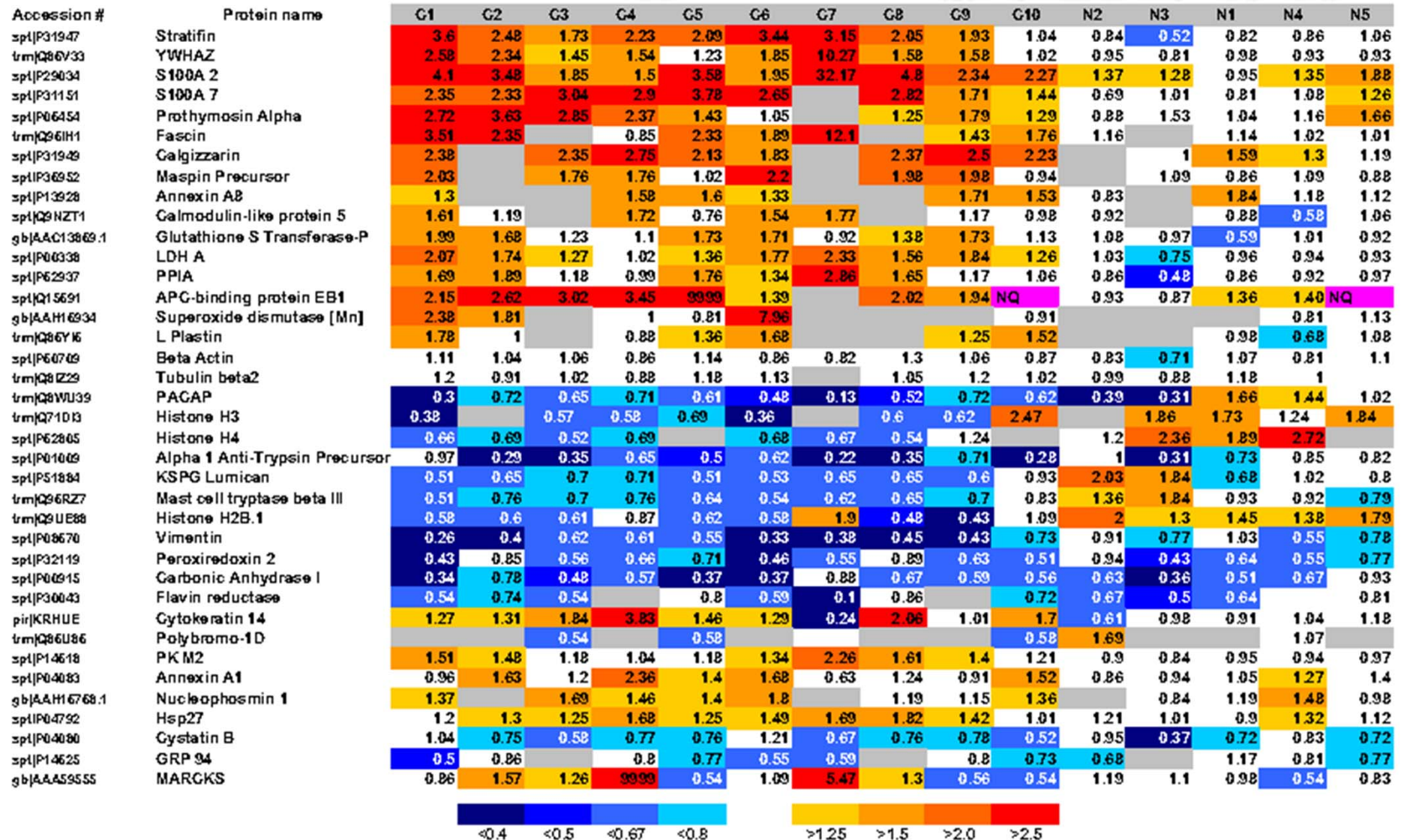
Siu Group Example...

Discovery and Verification of Head-and-neck Cancer Biomarkers by Differential Protein Expression Analysis Using iTRAQ Labeling, Multidimensional Liquid Chromatography, and Tandem Mass Spectrometry*

Ranju Ralhan†§¶||, Leroi V. DeSouza‡§, Ajay Matta¶**, Satyendra Chandra Tripathi¶, Shaun Ghanny§‡‡, Siddhartha Datta Gupta§§, Sudhir Bahadur¶¶, and K. W. Michael Siu‡§|||

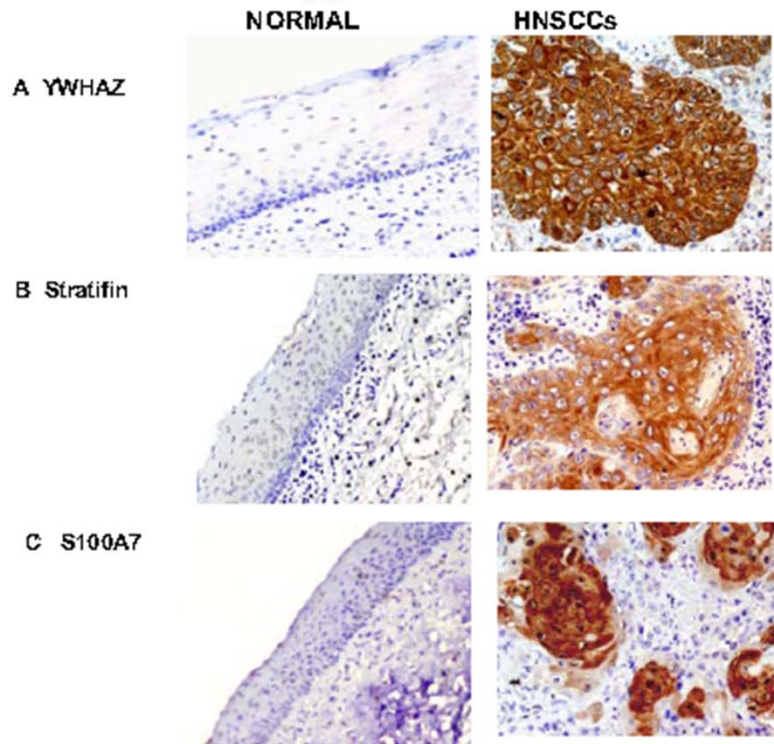
- This is a similar study with a better representation of the data as a ‘heat map’.

The Head and Neck Cancer Heat Map



Histological Proof

- Immunohistochemistry was used to stain cells for the first three overexpressed proteins on the list...



- But are these proteins specific to head and neck cancer? Or are they upregulated in other cancers?

Head and Neck Specific Biomarkers

- So lets do the same immunohistochemistry on other cancers:

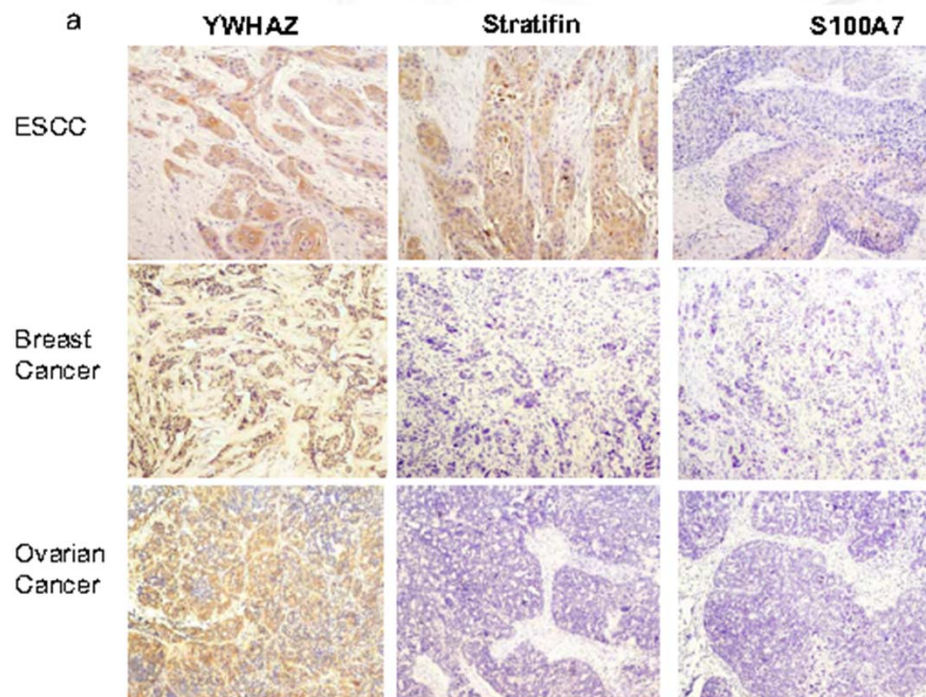


TABLE III

Receiver operating characteristics from the IHC scores of a panel of the three best performing biomarkers, YWHAZ, stratifin, and S100-A7, individually and as a panel

Biomarkers	Sensitivity	Specificity	PPV	NPV	AUC
YWHAZ	1.00	0.71	0.71	1.00	0.90
Stratifin	0.92	0.60	0.62	0.91	0.85
S100-A7	0.96	0.71	0.71	0.96	0.90
YWHAZ, stratifin, S100-A7	0.92	0.87	0.83	0.94	0.91

TABLE IV

Comparison of receiver operating characteristics from the iTRAQ ratios of the panel of the three best performing biomarkers: non-paired non-cancerous tissues give better sensitivity and specificity as a comparator than paired non-cancerous tissues

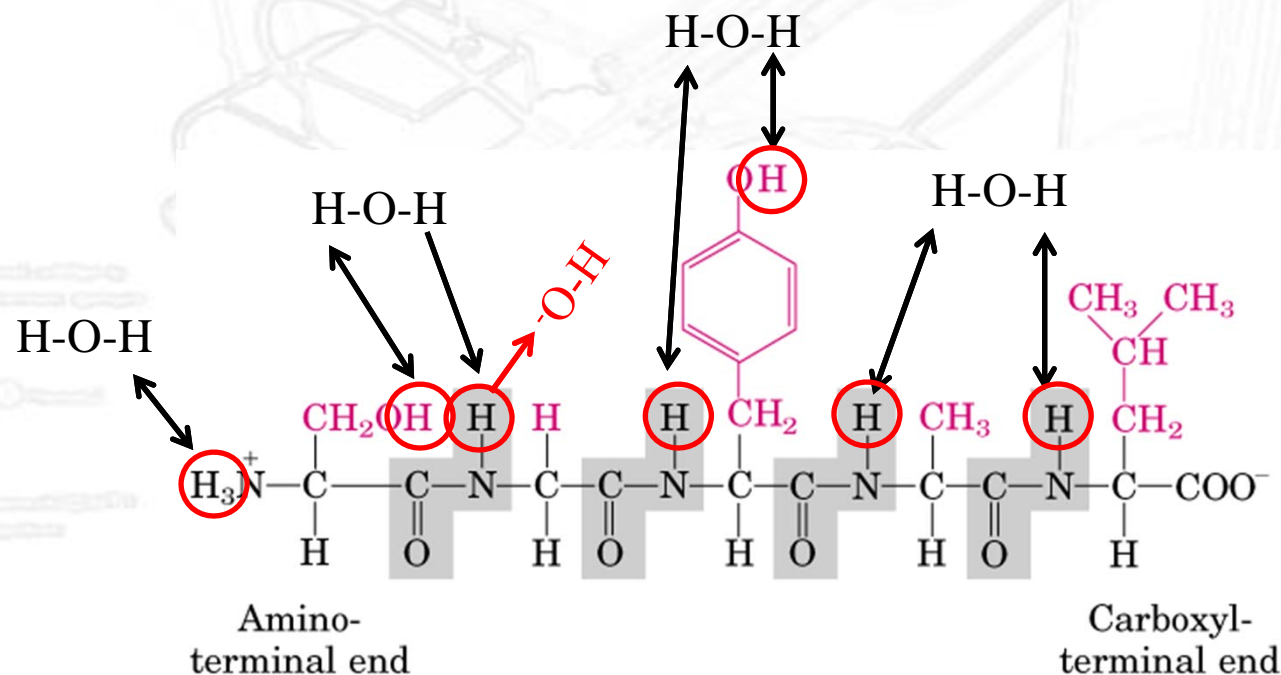
Comparison	Sensitivity	Specificity	PPV	NPV	AUC
Cancer versus paired normal	0.92	0.83	0.85	0.92	0.89
Cancer versus non-paired normal	0.96	0.96	0.98	0.90	0.97

And Now for Something Completely Different:

- So we've talked about finding proteins – *i.e.* **which ones are there** and **to what extent...**
- We might subsequently ask – How do these proteins do what they do. How do they function or mis-function?
- It turns out that integral to the question of how proteins function is the question of how proteins **move**, which is called **conformational dynamics**.
- Don't believe me? Try getting an enzyme from a hyperthermophile to work at room temperature!
- We might also be concerned with **protein aggregation diseases**, which are associated with conformational dynamics.

Hydrogen Deuterium Exchange (HDX)

- How do we learn about protein structure, folding and conformational dynamics using MS? HDX!!
- The idea behind HDX is simple: Amide protons on the peptide backbone are constantly exchanging with solvent...

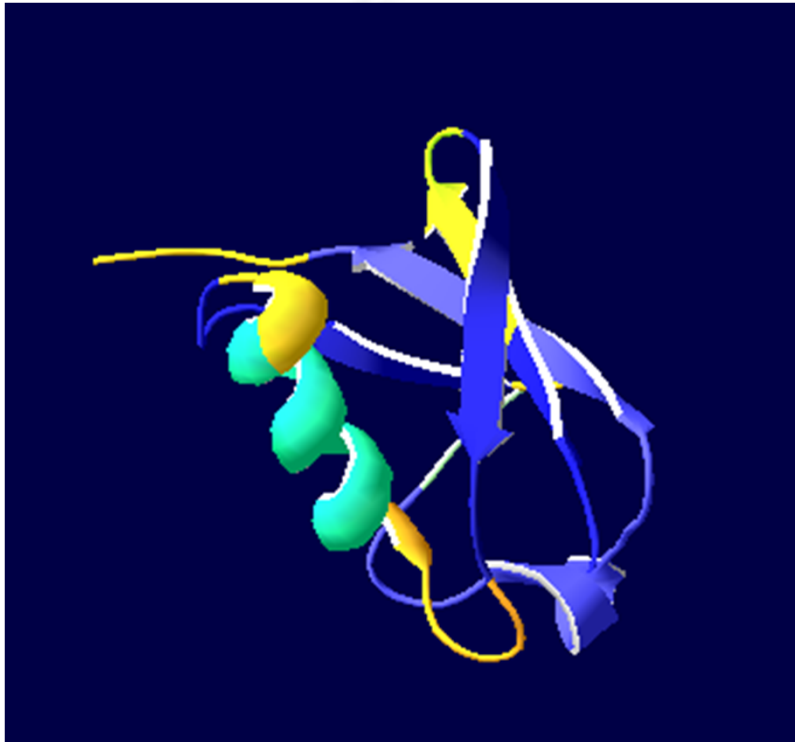


HDX and Protein Structure

- So if we put the protein in D₂O instead of water, exchange of backbone and side-chain protons for deuterium will make the protein heavier...which I can measure with MS...
- So what? No structural info there... but here's the kicker: In order for H/D exchange to occur **we must first break any hydrogen bonds that might be present.**
- What holds protein secondary structures (helices and beta-sheets) together? H-bonds!! That means, where there is a lot of secondary structure, HDX is going to be **slow**...

HDX Cont.

- Here's an example on a small protein called ubiquitin:

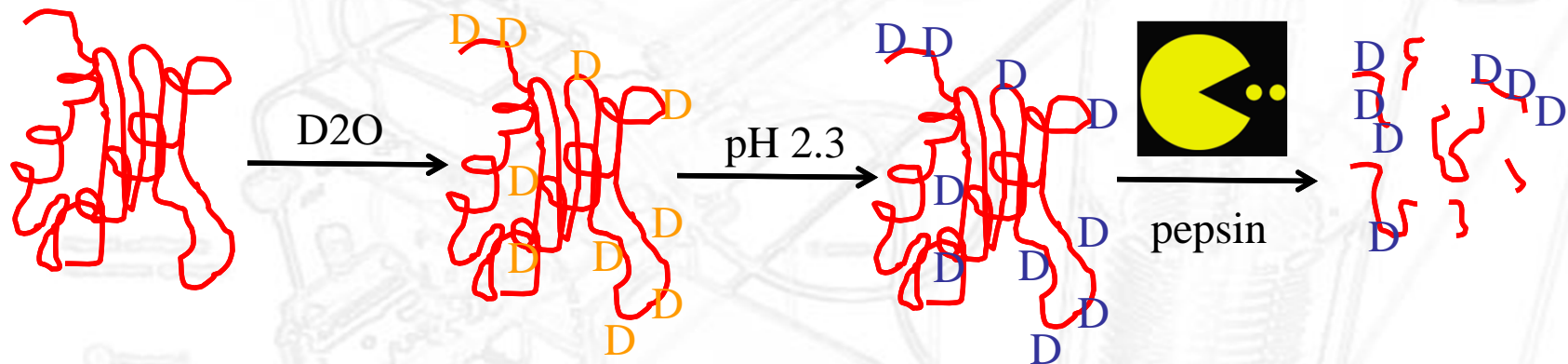


LOW-----MED-----HIGH

- Note that where there is secondary structure, there tends to be a **low** level of exchange. Where there are loops, there tends to be a **higher** level of exchange...

Spatially Resolved HDX

- But HDX would increase the mass on the **whole protein**. How do we know **where** the D is going?
- Time to break out the ‘spatially resolved HDX by mass spec’ workflow:

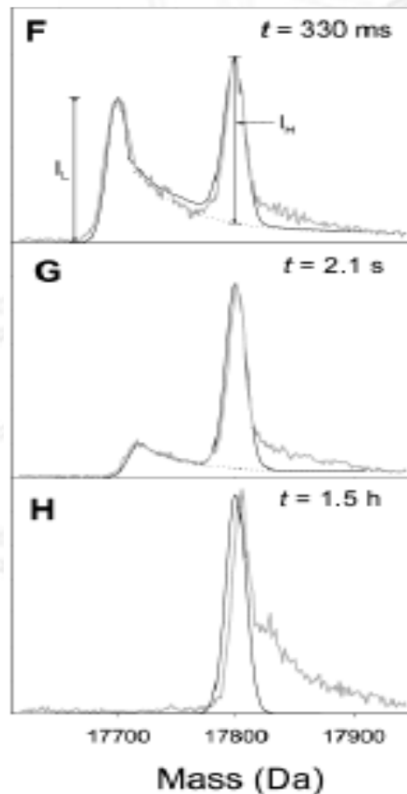


- We cannot simply label the protein and do CAD due to **proton scrambling**. Structural info is lost!

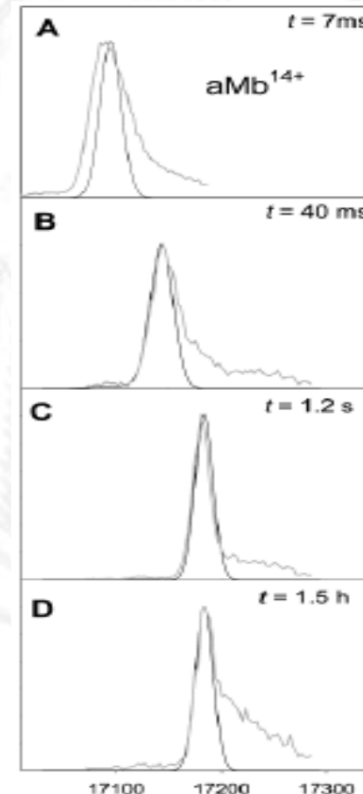
Global HDX: EX1 and EX2 exchange

- Just for the sake of completeness, HDX generally occurs in one of 2 regimes, called **EX1** and **EX2**.

EX1: Protein motion is **slow** relative to exchange. Motion occurs, **ALL** exchangeable sites opened up undergo exchange at once.



EX2: Protein motion is **fast** relative to exchange. Motion occurs, there is a **probability** of available sites exchanging.



Protein Folding by HDX: Example 1

- When we're doing HDX, we have to study protein folding (because we can't study folding in acid). The ultimate objective is to have a time-resolved 'movie' of how the protein folds up.
- We're not there yet, but this paper comes pretty close:

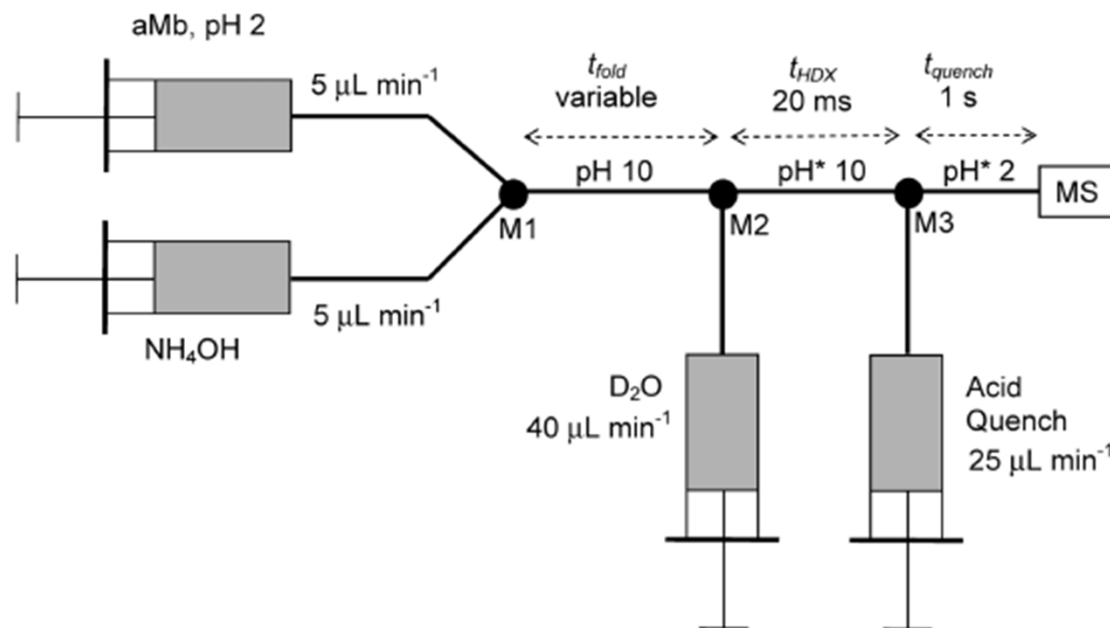
Characterizing Short-Lived Protein Folding Intermediates by Top-Down Hydrogen Exchange Mass Spectrometry

Jingxi Pan,[†] Jun Han,[‡] Christoph H. Borchers,[‡] and Lars Konermann^{*†}

Department of Chemistry, The University of Western Ontario, London, Ontario, N6A 5B7, Canada, and University of Victoria-Genome BC Proteomics Centre, Victoria, British Columbia, V8Z 7X8, Canada

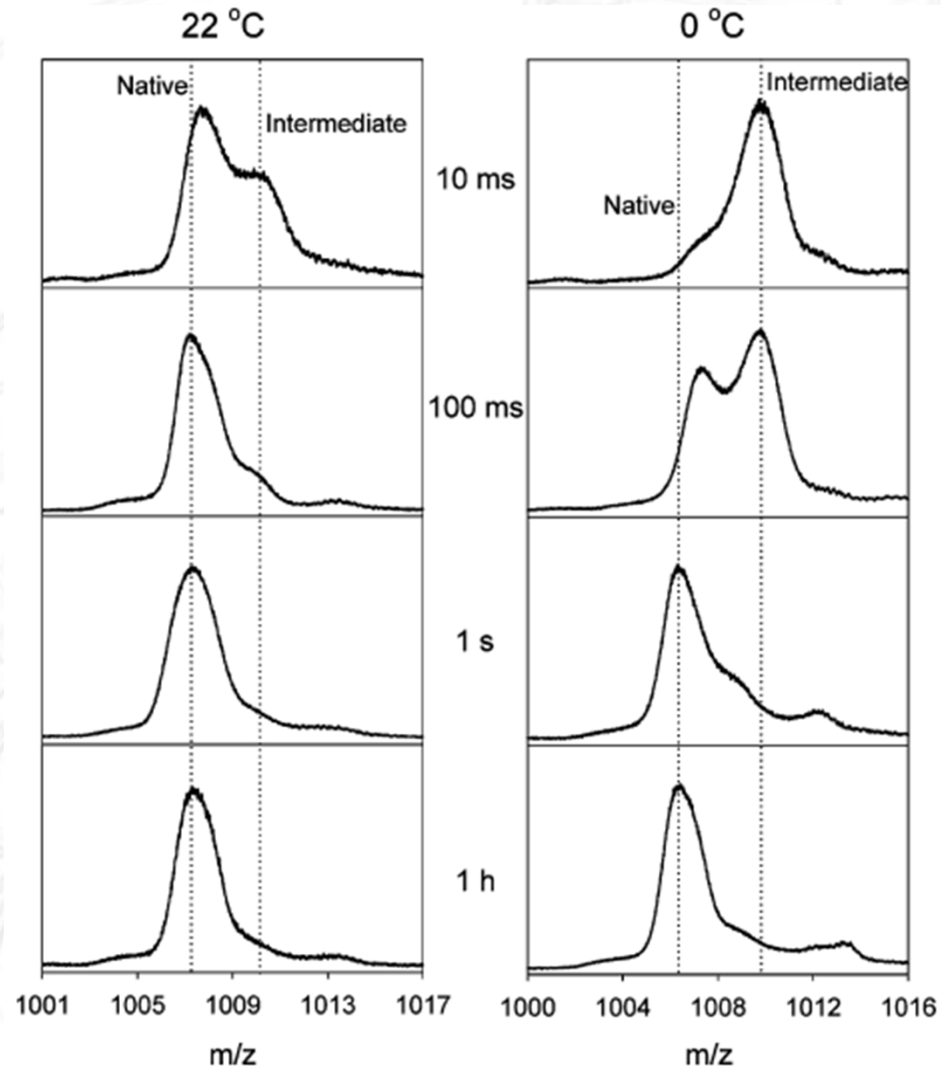
Spatially Resolved Protein Folding

- The crux of this paper is that they don't use the proteolysis: Instead they use 'top-down' ECD of the whole protein...
- They are also using a 'pulse labeling' approach in which you let the protein fold to a certain extent, then label, then use ECD to get your peptides:



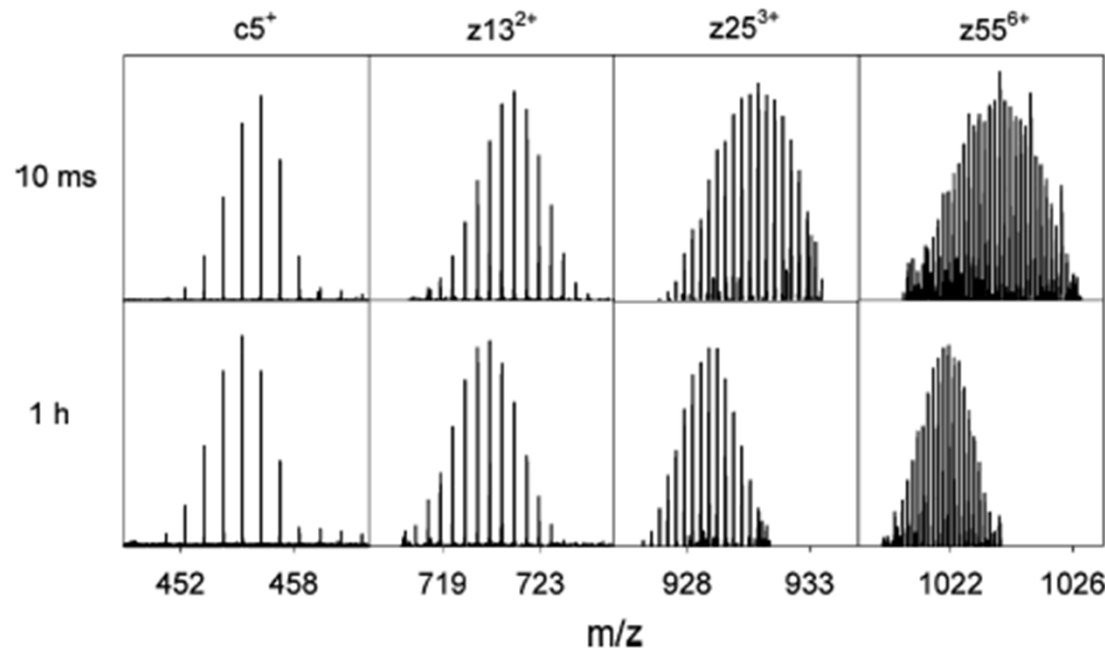
Time-resolved Protein Folding Cont.

- They also had to cool down their apparatus to get the folding within their available time-window.



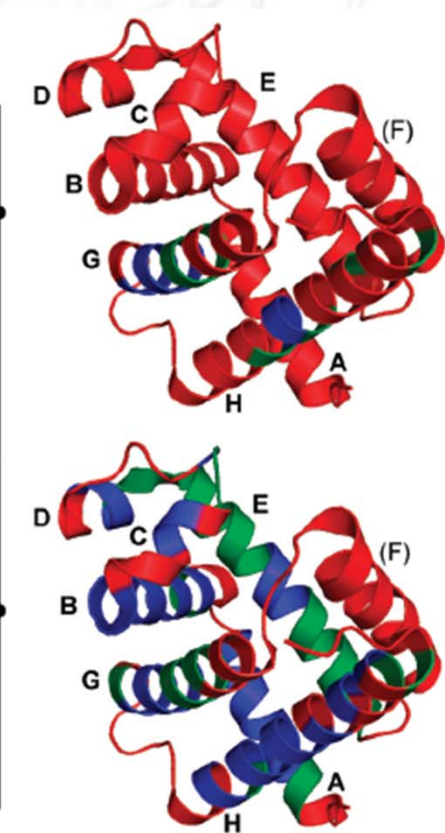
Protein Folding HDX Time-resolved...

- Here are some of the D-labeled peptides that they are getting from their ECD



- The resolution is so good here because they're doing this on an FT-ICR

from this paper are as follows:

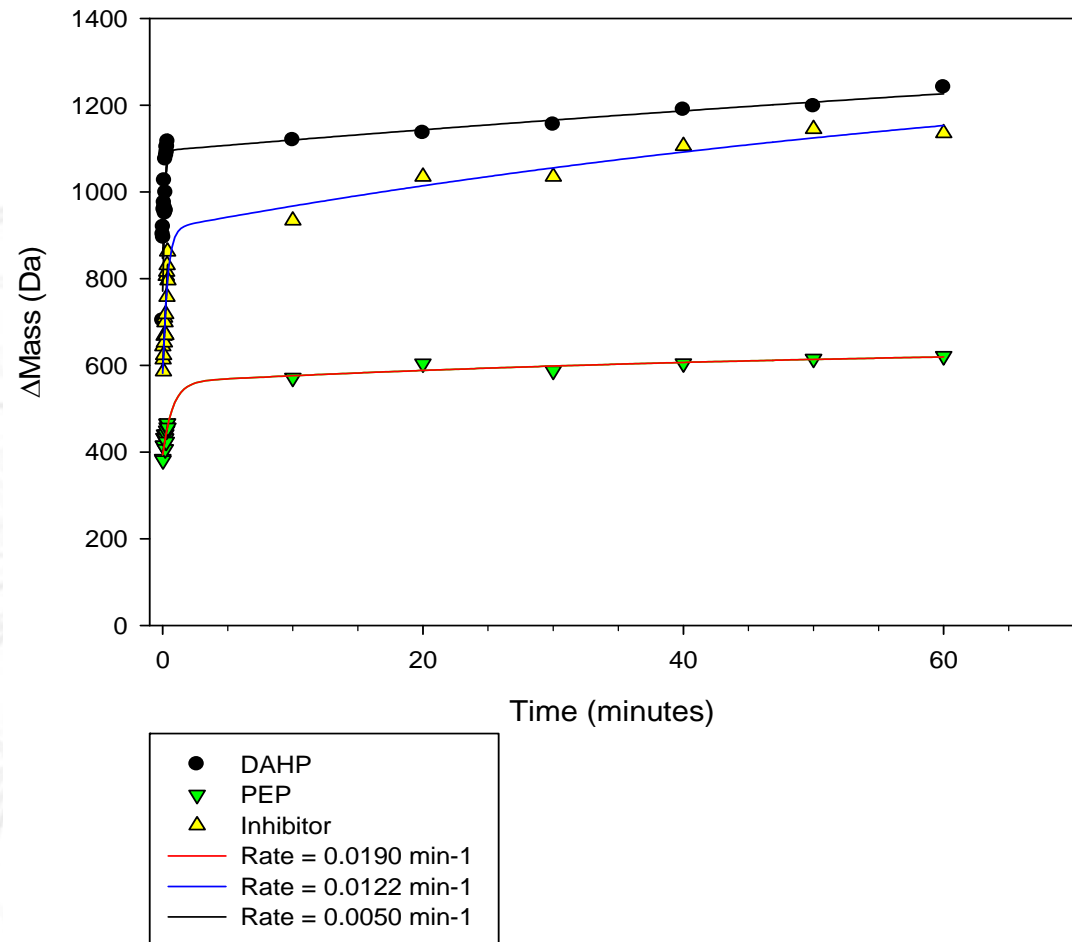


Example 2: Ligand Binding

- Many proteins function through binding of small molecules. This binding, and in some cases the binding surface can be studied by HDX-MS.

- This shows a substantial difference in global dynamics when the protein binds: It's natural substrate, an inhibitor.

DAHP Synthase H/D Exchange Profile

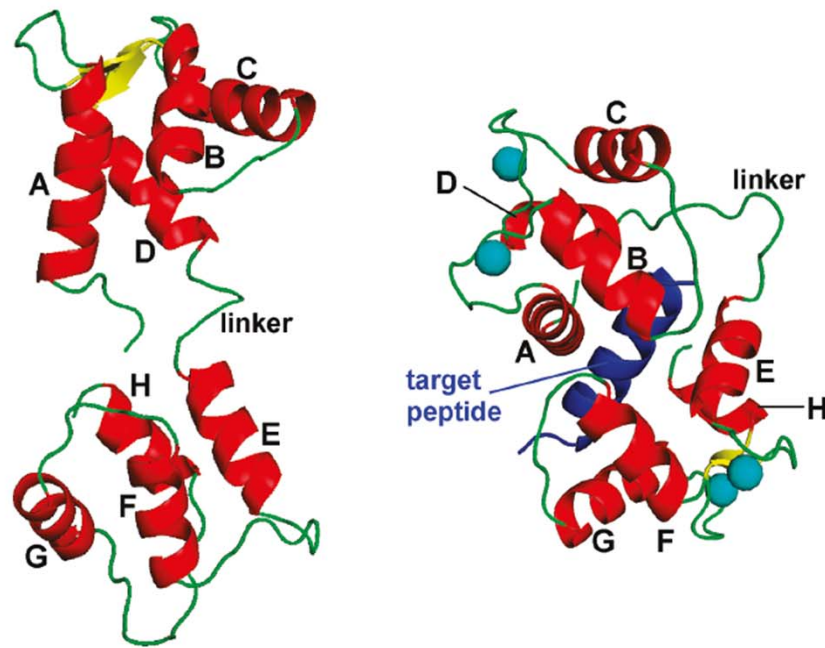


Calcium Binding to Calmodulin

- Here's a neat example of Calcium binding to Calmodulin in the presence and absence of a target peptide:

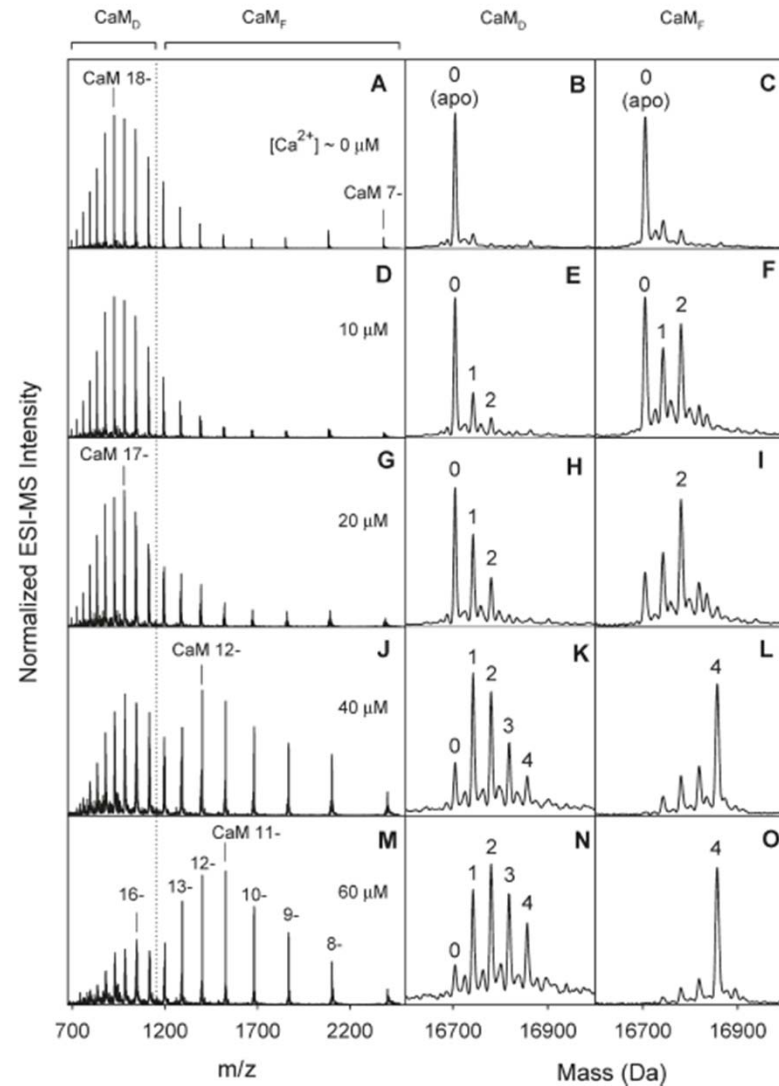
Calcium-Induced Structural Transitions of the Calmodulin–Melittin System Studied by Electrospray Mass Spectrometry: Conformational Subpopulations and Metal-Unsaturated Intermediates[†]

Jingxi Pan and Lars Konermann*



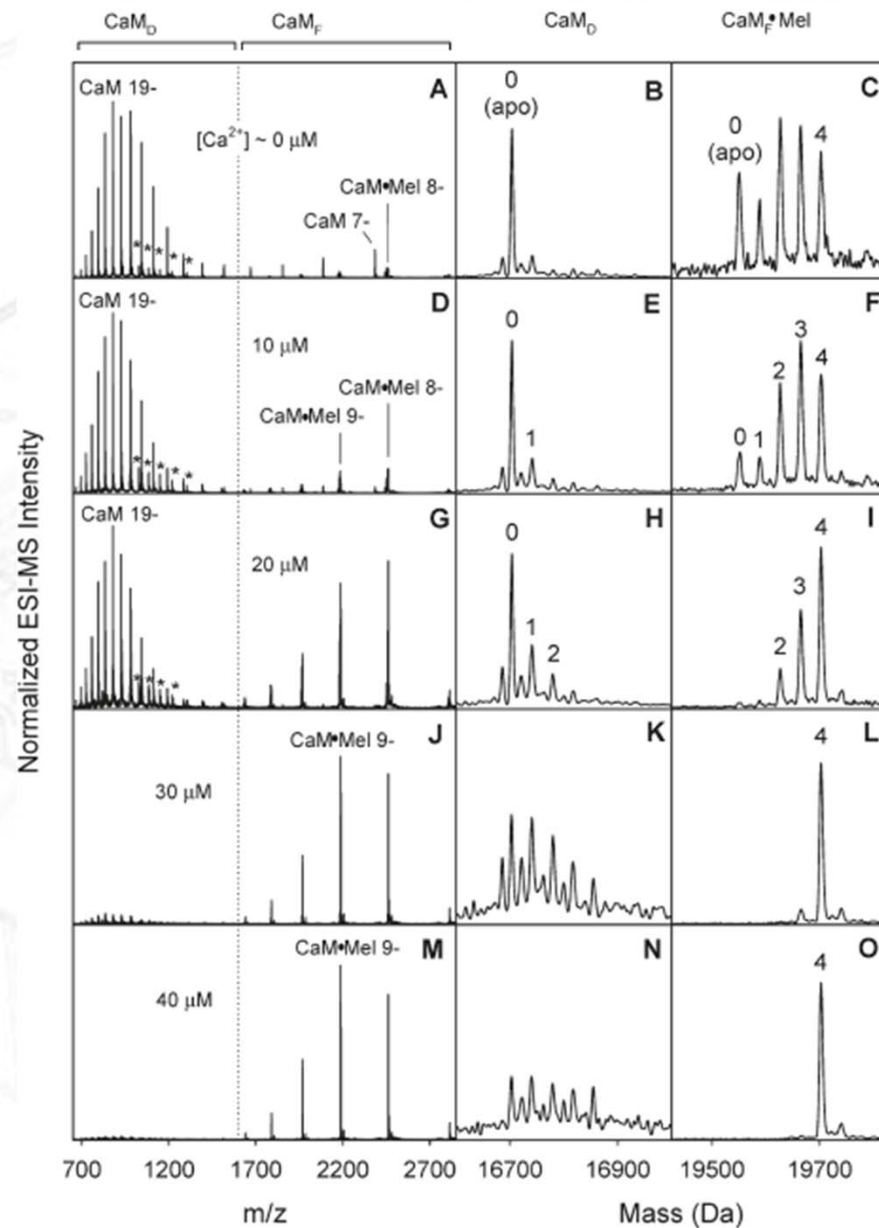
Head and Neck Specific Biomarkers

- Calcium binding in the absence of target peptide:
- CaM_D is denatured
- CaM_F is folded
- Calcium induces folding



More Calcium Binding

- Calcium binding in the presence of target protein
- Calmodulin showing higher affinity for calcium...
- Denatured Cam doesn't bind calcium or the target peptide...



Calcium Binding and Folding Kinetics

- Kinetic folding of Calmodulin in the presence of Ca^{2+} and target peptide...
- Allows us to watch binding and folding at the same time, which is pretty cool...

