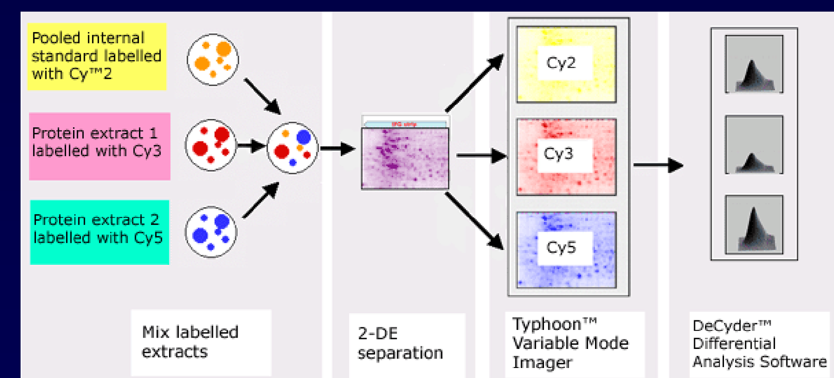
The Art of DIGE

Kathy Stone, Terence Wu, Erol Gulcicek, Mark Shifman, Jean Kanyo, Christian Collin-Hansen, Mary LoPresti, Chris Colangelo, Angus Nairn and Ken Williams

Abstract: Differential (fluorescence) gel electrophoresis (DIGE) is a powerful tool which utilizes 2D SDS PAGE for locating proteins that change in expression due to treatment / disease state. By incorporating mass and charge matched fluorescent dyes, three samples can be compared in one 2D gel thereby eliminating spot misalignment. Using the DeCyder software (GE Healthcare) protein spots are detected and quantitated on a pair (or group) of fluorescent images leading to the determination of which protein spots change abundance. Selected proteins are then identified by robotic excision of the gel spot, tryptic digestion and mass spectrometric analysis. One challenge with DIGE is efficient presentation of the gel image, spot location and protein identification. Open source software called YPED (Yale Protein Expression Database) was developed within our resource to help solve this challenge. YPED names protein spots from MALDI-Tof/Tof analysis (AB 4700 /4800), and links gel images with the proteins identified from the mass spectrometric analysis. Users log on a web interface using their unique Logon ID and password. By choosing the experiment of interest, they can visually see the gel image with identified spots indicated. Clicking on the spot lists the protein spot number from DeCyder, the protein identified, the corresponding peptide sequences, Mascot scores, Cy dye ratios etc. Data is exportable in an Excel format. A link to the Panther Classification System aids the researcher in studying pathways and functions. The power and benefit of DIGE is most clearly evident in looking at sites of post translational modifications (PTM). By utilizing a metal oxide enrichment step prior to DIGE analysis, we are able to locate phosphoproteins, as well as changes in the phosphoproteome. DIGE can also detect changes in protein glycosylation – particularly when comparing a sample before and after de-glycosylation. Examples of both types of PTMs will be shown.

Differential (fluorescence) Gel Electrophoresis (DIGE) - Based Protein Profiling



- Fluorescence-based detection provides high sensitivity with 3 dyes (Cy2, Cy3, and Cy5) available
- Dyes attach via a N-Hydroxysuccinimide (NHS) linkage to primary amines
- Minimal dyes (~ 3% of protein will have 1 dye molecule attached to lysines)
- Limit of detection= 1.3 fold above background

DIGE Analysis General Procedure

1) Amino Acid Analysis (AAA – ion exchange, post column ninhydrin detection on a Hitachi Model L-8900PH) is used to quantitate the sample in the lysis buffer (7 M Urea, 2 M Thiourea, 4% CHAPS and 30 mM Tris, pH 8.8)
2) Based on the amino acid analysis, 50 to 100 µg of protein of each sample is labeled by adding 1 µl Cy2, Cy3 or Cy5-N-hydroxysuccinimidyl ester dye (400 pmol/µl in 99.8% anhydrous DMF) followed by a 30 min. incubation on ice in the dark. The reaction is stopped by adding 1 µl 10 mM lysine.
3) Two (or three if an internal standard has also been labeled) samples are combined and brought to a final volume of 450 µl by adding the appropriate volume of rehydration buffer containing 8M Urea, 4% CHAPS, 13 mM DTT and 1% of the pH 3-10 ampholine solution.
4) 24 cm Immobiline (IPG) Drystrips are re-hydrated for 10-24 hrs, and the sample is subjected to isoelectric focusing.
5) The SDS polyacrylamide gel electrophoresis (second) dimension is carried out on a 22 cm wide by 24 cm tall by 1.0 mm thick gel purchased from Jule Inc. Currently, we routinely use a 12.5% polyacrylamide gel which will optimally separate 12-200 kD proteins.
6) Immediately following SDS PAGE, the gel is scanned at 2 (or 3) wavelengths sequentially on a GE Healthcare Typhoon 9400 Imager. After scanning, 16 bit GEL files of each color channel are exported for image analysis using the GE Healthcare DeCyder software.
7) The DeCyder software is used to quantify the gel image and to identify a "pick list" of differentially expressed protein spots to be excised and subjected to in–gel tryptic digestion on the ETTAN TA Digester and MALDI-TOF/TOF (AB 4800) or LTQ Orbitrap mass spectrometry for protein identification. For MALDI-TOF/TOF, the digest is dried, re-dried from 50µl water, dissolved in 0.7µl 3mg/ml α-cyano-4-hydroxycinnamic acid matrix (Waters, part number in 0.05% TFA, 50% acetonitrile. 0.6µl is spotted on target. For LTQ Orbitrap analysis, the digest supernatant is loaded directly into a vial and injected.

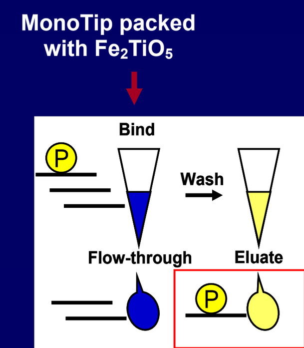
AB MALDI-TOF/TOF	4700	4800
# gel protein spots analyzed	11983	9478
# wells missing gel piece	799	621
# protein spots identified	7839	7199
Average % Identified	70%	81.3%

MALDI-TOF/TOF Protein Identifications

- MS: 1500 laser shots
- MS/MS: 10,000 laser shots (required for very low level samples)
- Up to 10 MS/MS collected
- Approx. 7 minutes/sample
- Identification done using Mascot/AB GPS Explorer with a combined peptide mass fingerprint and MS/MS search
- 50 ppm mass accuracy for peptide mass fingerprint
- 0.2Da mass accuracy for peptide fragments
- ~ 6% of gel pieces are lost in the digester

Phosphoprotein Enrichment

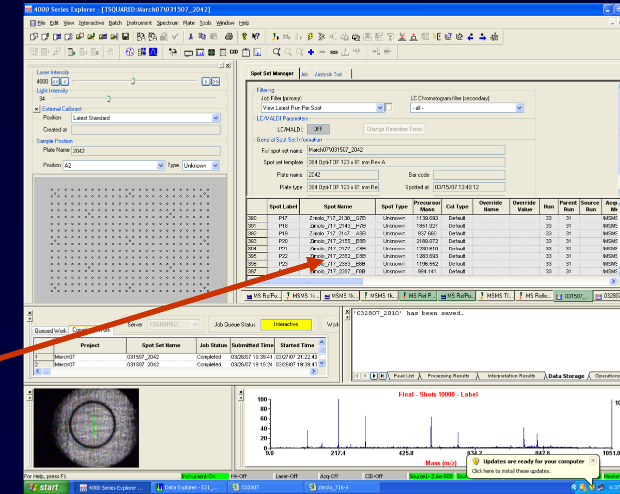
- MonoTip is packed with Fe_2TiO_5 and equilibrated with 0.5% TFA, 50% acetonitrile (equilibration buffer)
- Acidified protein sample is loaded onto tip and washed with the equilibration buffer
- Phosphoproteins are eluted off the tip using 28% ammonium hydroxide
- Eluted samples are dried and re-dried from water prior to further analysis



Preparation of DIGE samples for MALDI-TOF/TOF Analysis and entering into the Yale Protein Expression Database (YPED): Naming of Sample Spots

4800 Sample Naming Excel Table

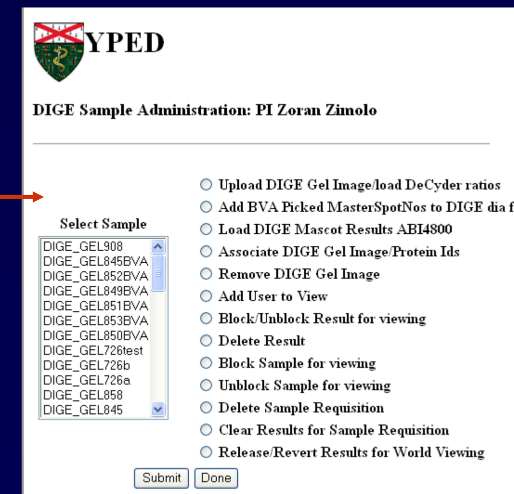
PI Last	Zimolo	
Gel #	625	
Spot #	96 well Plate position	4800 Sample Name
108	A1	Zimolo_625_108__A1
140	B1	Zimolo_625_140__B1
316	C1	Zimolo_625_316__C1
364	D1	Zimolo_625_364__D1
421	E1	Zimolo_625_421__E1
524	F1	Zimolo_625_524__F1
527	G1	Zimolo_625_527__G1
649	H1	Zimolo_625_649__H1
679	A2	Zimolo_625_679__A2



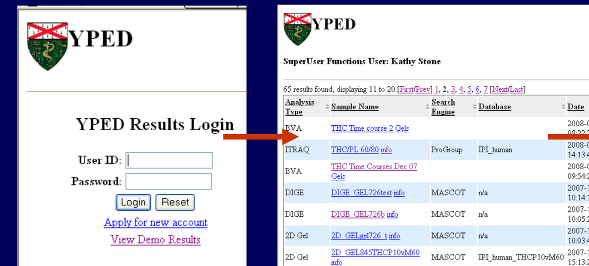
An Excel table is used to name each sample spot grabbing the spot number from the DeCyder generated pick list *.txt file and linking it with the position in the 96 well plate from the ETTAN Spot Picker. The user inputs the PI name, gel number and spot number. The sample name is generated and then copied into the 4800 sample list.

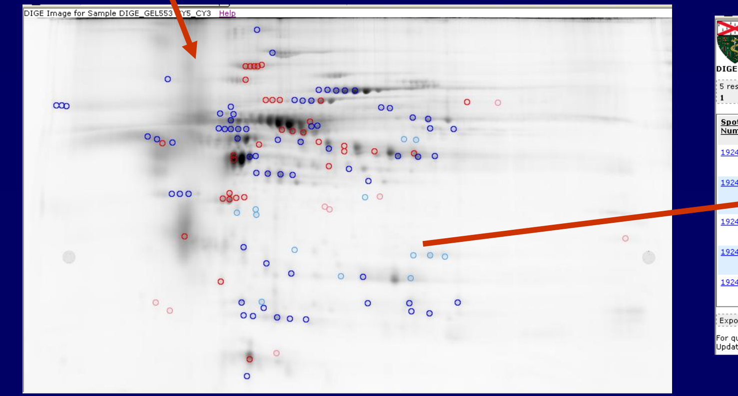
Yale Protein Expression Database (YPED)

The Yale Protein Expression Database (YPED) is an integrated web-accessible software system that addresses the need for storage, retrieval and integrated analysis of large amounts of data from high throughput proteomic technologies. YPED has been particularly valuable in DIGE analysis where it links the gel image to the protein identification, also handling BVA (biological variation analysis) results. The proteomics staff member uses YPED to log in sample details (in a submission form format), and to link the submission form with the Typhoon generated *TIFF image and then with the spot details from the DeCyder *XML file. This in turn is then linked to the protein identifications from the AB GPS Explorer by scanning for files with the unique name generated. The process is very automated and takes less than 5 minutes to generate the YPED entry. For the investigator, there are compare tools (e.g. DIGE results to an iTRAQ analysis) and a link to the Panther Classification System to aid the researcher in studying pathways and molecular and biological functions. The below shows the YPED flow for a researcher's DIGE experiment.



Screen shot of the YPED page linking the various DIGE and protein identification results together.



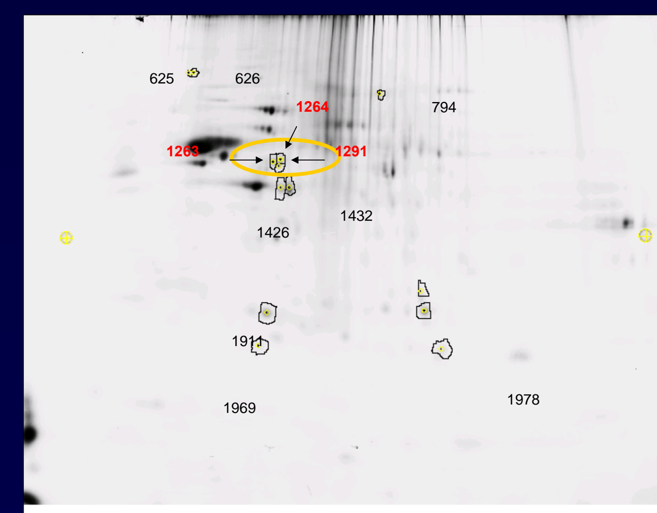


Shifman, M.A., Li, Y., Colangelo, C.M., Stone, K.L., Wu, T.L., Cheung, K., Miller, P.L., and Williams, K.R. (2007) YPED: A Web-Accessible Database System for Protein Expression Analysis, J. Proteome Research, 6, 4019-4024.

Utilizing DIGE for Locating Sites of Post Translational Modifications

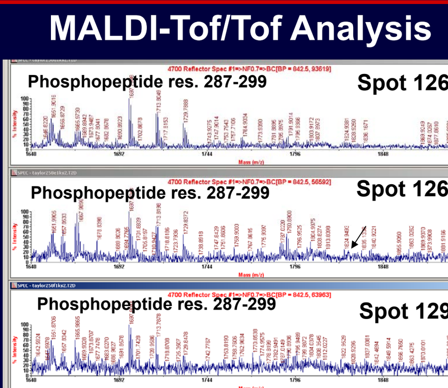
- Phosphorylated proteins migrate with different isoelectric points in DIGE
- Phosphopeptides can often be located based on MALDI-MS fingerprint analysis of the tryptic digest (Mascot)
- MS/MS analysis can often locate the site of phosphorylation
- Metal oxide enrichment can be used to locate phosphoproteins
- DIGE can also detect changes in protein glycosylation,
- acetylation, and methylation
- Comparison of samples before and after de-glycosylation is a powerful approach to locating glycopeptides

Relative Protein Expression in Nucleus Accumbens saline vs cocaine treated (300µg total load)- Location of Phosphopeptide Differences



- Labeled spots were differentially regulated and selected for spot picking, tryptic digestion and protein identification.
- Spot numbers in red all contain the Glial fibrillary acidic protein (GFAP) based on MALDI-Tof/Tof analysis.

MALDI-Tof/Tof Analysis



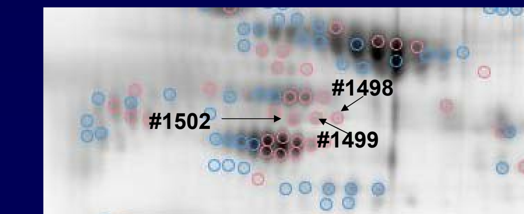
Phosphopeptides

Spot #	m/z	# phosphosites
1263	854.05	1
	1665.95	2
	1790.88	Not seen
1264	854.05	Not seen
	1665.96	2
	1790.88	2
1291	854.05	Not seen
	1665.97	2
	1790.88	Not seen

Location of differing sites of acetylation utilizing DIGE

Sample:
Cy5: 100µg of pre-dexamethasone treated plasma Female pool of 3 patients
Cy3: 100µg of post 24 hour dexamethasone treated plasma, Female pool of 3 patients
** both samples after IgY14 depletion (GenWay Biotech Inc.)

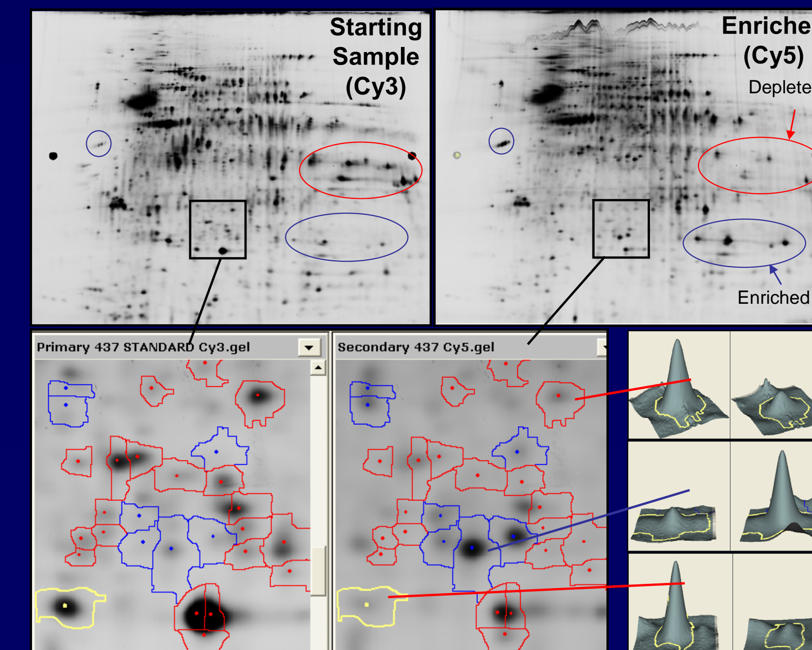
LC-MS/MS Analysis was done on a LTQ-Orbitrap (Thermo Scientific) equipped with a Waters nanoAcquity UPLC system. The column used was a Waters 75µm x 250 mm 1.7µm eluted at 300 nl/min. A 60 minute linear gradient of buffer A (0.1% formic acid/water) and buffer B (0.075% formic acid/



Protein spots # 1498, #1499, and #1502 were identified as Inter-alpha-trypsin inhibitor heavy chain H4 precursor using Mascot Daemon and Mascot v 2.2. Spot positions varied due to differing sites of lysine acetylation.

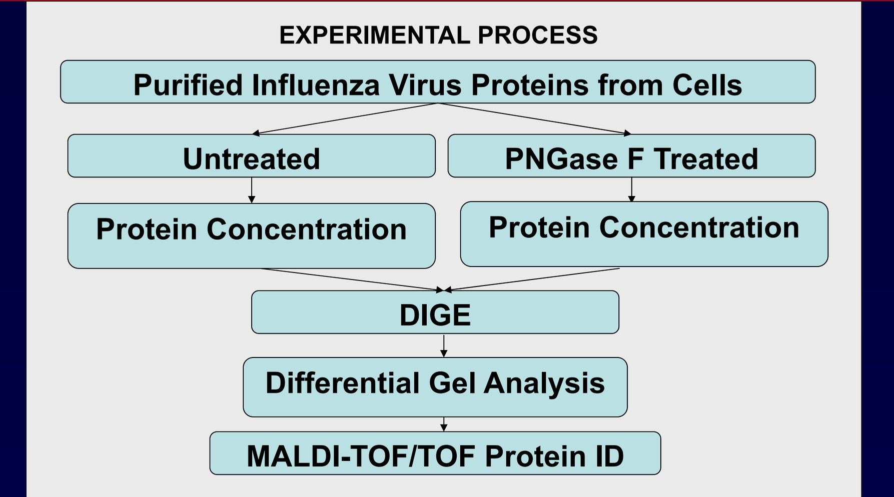
Spot #	Mass of acetylated peptides
1498	1000.59
1499	1000.59
	1523.89
1502	none

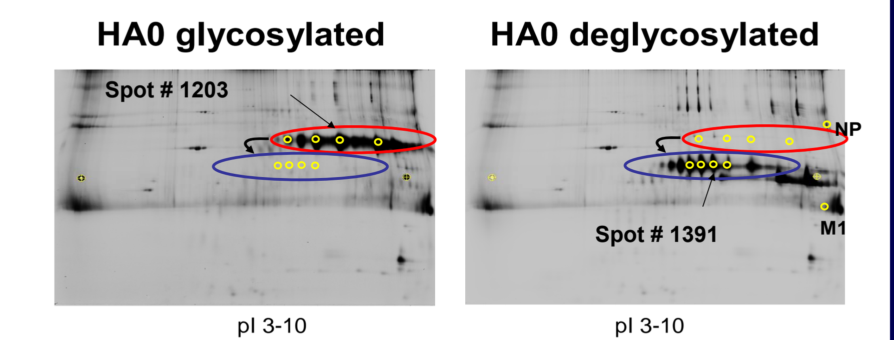
Phosphoprotein Enrichment by Metal Oxides on Synaptoneurosomes Fe (III) titanate: Fe_2TiO_5



The gel image on the left top is the starting material and was labeled with Cy3. The same sample was subjected to phosphoprotein enrichment (see methods) and labeled with Cy5 (right top). The bottom images are the expanded regions with the DeCyder view shown in the right bottom. A total of 5µg of rat brain synaptoneurosomes were used for each dye.

Locating Glycoproteins and sites of Glycosylation



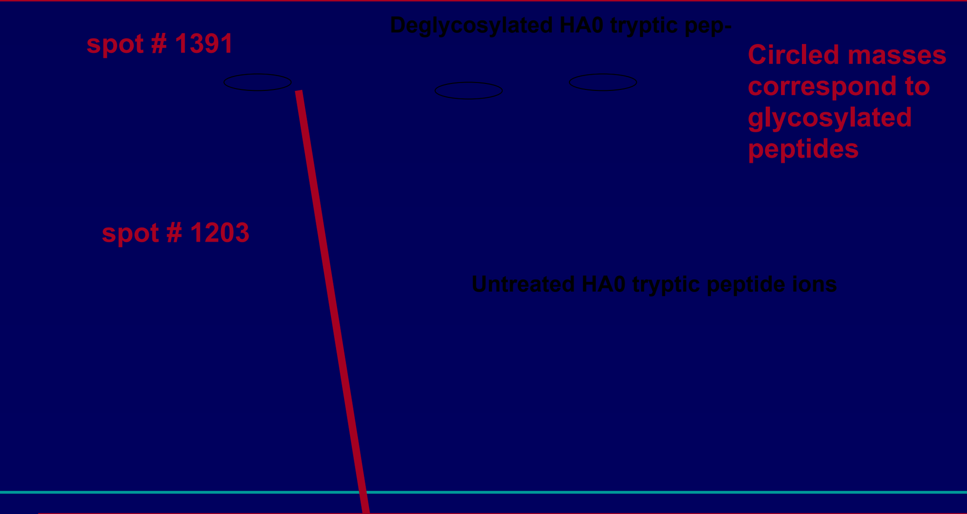


Cy3 Image: untreated virus proteins

Cy5 Image: PNGase F treated virus proteins

Protein spots in yellow were identified by MALDI-TOF/TOF

MALDI-TOF/TOF Spots #1391 and 1203: Expanded mass range



The N residue that the PNGaseF removes the sugar from is deamidated to D.

Conclusions:
DIGE is a very powerful tool for looking at changes in protein expression levels and changes in the post translational proteome.

Acknowledgments
Some of the methods described above were used in studies funded with Federal funds from NHLBI/NIH contract N01-HV-28186, NIDA/NIH grant 1 P30 DA018343-01, and NIAD/NIH grant 5 U54 AI057158-02 (Northeast Biodefense Center - Regional Centers of Excellence).

Visit our web site at http://keck.med.yale.edu/ for more information