Spatial and Temporal Effects in Protein Post-translational Modification Distributions in the Developing Mouse Brain

ARTICLE *in* JOURNAL OF PROTEOME RESEARCH · AUGUST 2013

Impact Factor: 4.25 · DOI: 10.1021/pr4002977 · Source: PubMed

5 AUTHORS, INCLUDING:

Alistair Vincent Gordon Edwards
University of Southern Denmark
17 PUBLICATIONS **286** CITATIONS
SEE PROFILE

Veit Schwämmle
University of Southern Denmark
59 PUBLICATIONS **946** CITATIONS
SEE PROFILE

Martin R Larsen
University of Southern Denmark
160 PUBLICATIONS **7,400** CITATIONS
SEE PROFILE

Journal of proteome research

Article

pubs.acs.org/jpr

Spatial and Temporal Effects in Protein Post-translational Modification Distributions in the Developing Mouse Brain

Alistair V. G. Edwards,[†] Gregory J. Edwards,[‡] Veit Schwämmle,[†] Henrik Saxtorph,[§] and Martin R. Larsen*,[†]

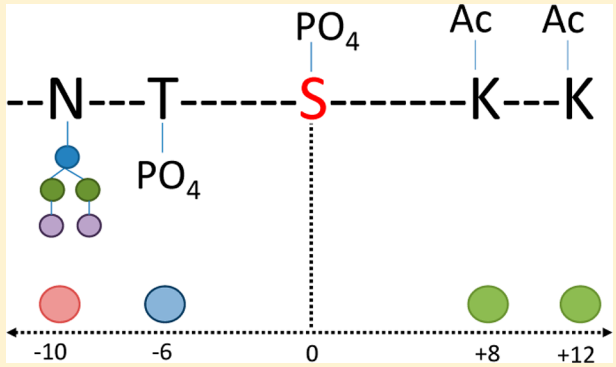
[†]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, Odense, DK 5230, Denmark

[‡]Port Jackson Bioinformatics, Annandale, Sydney, NSW 2038, Australia

[§]Biomedical Laboratory, Odense University Hospital, Winsløwparken 23, Odense, DK 5000, Denmark

Supporting Information

ABSTRACT: Protein post-translational modification (PTM) is a powerful way to modify the behavior of cellular proteins and thereby cellular behavior. Multiple recent studies of evolutionary trends have shown that certain pairs of protein post-translational modifications tend to occur closer to each other than expected at random. This type of observation may form the basis of a proposed "PTM code", whereby protein function is controlled by complex patterns of multiple PTMs. This code could provide an additional, powerful level of regulatory control for protein function and is a plausible explanation for observations of increasingly frequent and diverse protein modification in cell biology. In this study, we use mass spectrometry and proteomic strategies to present biological data showing spatiotemporal PTM co-localization across multiple PTM categories, which display changes over development of the brain. This may be an indication of the existence of a PTM-based functional coding mechanism, which would significantly expand our view of the ways in which cells use protein PTMs in complex signaling networks.



KEYWORDS: *protein post-translational modification, PTM code, mouse brain development*

INTRODUCTION

Protein post-translational modifications, encompassing dozens of different chemical additions to specific residues in a protein sequence that can be added to target proteins through the action of other proteins (such as kinases and transferases), have a well-documented role to play in controlling cellular behavior. The ways in which the effects of PTMs are felt range from the simple, such as the addition of a modification group reducing access to active sites in a protein,[1,2] to more complex consequences as seen in, for example, the modulation of p53 by acetylation, phosphorylation, and ubiquitination.[3] It appears, in fact, that many proteins are regulated at a site-specific level by a range of PTMs,[3−5] which offers the possibility of increasing the range of differing functional outcomes in a manner dependent on precisely which PTMs are present.[6] This idea is attracting increasing attention (e.g., in databases aimed specifically at cataloguing PTM associations[7]) at least partly because it is strongly reminiscent of occurrences in the nucleus, where histone protein tails are multiply and differentially modified to control DNA transcription.[4,8−10] The existence of these reproducible modification patterns in histone proteins is known as the 'histone code' and justifiably attracts much attention as a means of gaining insight into PTM-driven regulation of gene expression.

Recent research has shown, however, that this kind of multi-PTM-driven control of protein function may not be limited to histone proteins. Minguez et al.[11] have demonstrated using evolutionary meta-analysis that certain PTMs (e.g., phosphorylation and acetylation) are generally located closer to each other in space and sequence across entire proteomes than a random distribution can explain and that these relationships hold true across multiple species. This strongly suggests that these multiple, closely spaced PTM distributions have some biological role to play that is retained across evolutionary time. In a similar vein, Beltrao et al.[12] have shown that acetylation and phosphorylation events in particular are found significantly closer together than expected by chance across 11 species, which suggests the existence of 'regulatory hotspots' through which much information could be routed, strengthening the argument for a possible biological role for these patterns of PTMs. Given this recently documented existence of physical distribution relationships between PTMs in proteins beyond histones, the next stage is to ask whether there exists a more generalized 'PTM code' that extends to all cellular proteins (as

Special Issue: Chromosome-centric Human Proteome Project

Received: April 3, 2013
Published: August 15, 2013

has been outlined for specific non-histone proteins, such as tubulin[13,14]). Such a code could be used to vastly increase the information-carrying capacity of a signaling network, by finely regulating the function of its constituent parts;[6,15] as the network is made more complex through increasing the number of component parts (that is, by post-translationally modifying existing components), so the range of possible end points or functional outcomes also increases. This would be particularly impactful in situations where close control of cellular processes is absolutely necessary, and where a great deal of information is being distributed in a short period of time. Developmental processes provide an example of complex events and pathways that require this exceedingly precise control and regulation in order to avoid amplification effects of relatively minor deviations (an assertion intuitively understood and also supported by, for example, developmental trends in RNA splicing complexity[16]). These are situations wherein a PTM code could reasonably be expected to be clearly evident and may be the ideal places to search for evidence of a code of this nature.

As many of the authors of the above-mentioned papers point out, what is now needed is biological evidence in order to begin to confirm or refute the existence of a general PTM code of the type proposed; that is, the observed relationships need to be demonstrated in a biological system over biological time frames and shown to be connected to cellular events. With this in mind, we have performed a study focused on the physical and temporal relationships between PTMs during mouse brain development, aiming to elucidate the possible connections between protein PTMs and cell- and organism-level biology. This has first revealed that phosphorylation, acetylation, and glycosylation all share the clustering characteristics previously described in cross-sectional multispecies studies, but in a spatiotemporal biological context. Second, we show potentially biologically relevant changes in this clustering behavior over the development of the mouse brain. Both of these observations may suggest the existence of a global, PTM-based regulatory system.

MATERIAL AND METHODS

All chemicals were purchased from Sigma (St. Louis, MO, USA) unless stated otherwise. Experiments were performed in biological triplicate, and soluble samples also had a fourth replicate. See Figure 1 for a summary of methods.

Animal Model

Mice of strain NMR1 were sourced from the Biomedical Laboratory at Odense University Hospital. Animals were euthanized within 12 h after birth, at 8 days old, 21 days old, or 80 days old. Whole brains were excised rapidly, flushed in 0.1% saline, and snap frozen in liquid nitrogen. All animal work was performed in accordance with the guidelines of the University of Southern Denmark

Sample Preparation

Brains were dounce homogenized in 0.1 M Na_2CO_3 (ice cold) containing protease and phosphatase inhibitors (Roche, Meylan, France) and then tip-probe sonicated 2 × 20 s on ice. Homogenates were ultracentrifuged at 150,000*g* for 1.5 h. Pellets (membrane fractions) were resuspended in 6 M urea/2 M thiourea. Supernatants (soluble fractions) were precipitated using 20% trichloroacetic acid, and pellets from this precipitation were resuspended in 6 M urea/2 M thiourea. Samples were reduced in 10 mM dithiothreitol and then

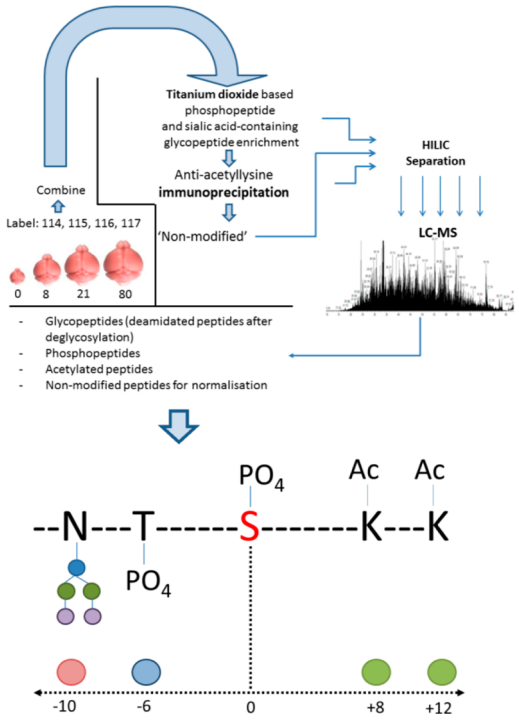


Figure 1. Methodology used to conduct study. See Materials and Methods for detailed description. Briefly, brains were homogenized, and protein was extracted and labeled. Peptides were subjected to a comprehensive enrichment strategy. Once peptide sequences and confident PTMs had been assigned, locations of PTM sites on protein sequences were mapped.

alkylated using 20 mM iodoacetamide. Lysyl-endopeptidase (Wako, Osaka, Japan) was added for 3 h, and then samples were diluted 6-fold in 50 mM triethyl ammonium bicarbonate. Trypsin (Promega, Madison, WI, USA) was added at a ratio of 1:40 (w/w) and left for 18 h to digest. Digests were acidified with formic acid (2% final concentration), and aliquots were quantified using amino acid analysis. Equal amounts of protein from each were iTRAQ fourplex labeled according to manufacturer's instructions (AB Sciex, MA, USA) and combined.

Peptide Enrichments

Combined peptide samples were diluted in 80% acetonitrile, 5% trifluoroacetic acid, and 1 M glycolic acid and enriched using titanium dioxide (GL Science, Japan) essentially as per published techniques.[17−19] Briefly, beads were added to samples at 0.6 mg/100 μg (bead/peptide), incubated, and washed with 80% ACN/1% TFA and then 10% ACN/0.1% TFA. Phosphopeptides were eluted with pH 11 ammonia solution (Merck & Co., Inc., NJ, USA). Flow through from this enrichment was cleaned up on Oasis HLB columns (Waters, MA, USA), dried down, and then immunoprecipitated for lysine acetylation following published procedures.[20] Elutions from TiO_2 were separated into multi- and monophosphorylated fractions using SIMAC.[21] Multiphosphorylated elutions were cleaned up on homemade R3 columns and dried down for MS analysis. Monophosphorylated elutions were adjusted to 70% ACN/2% TFA and enriched again using TiO_2, cleaned up on R3, and then dried down for chromatography. All peptide samples were then resuspended in pH 7.4 ammonium bicarbonate and deglycosylated using PNGase F (Roche) and

dx.doi.org/10.1021/pr4002977 | J. Proteome Res. 2014, 13, 260−267

sialidase A (Prozyme, CA, USA) before being desalted once more and lyophilized.

Chromatographic Separation

All peptide samples excepting those multiphosphorylated were separately resuspended in 90% ACN/0.1% TFA and loaded onto an Agilent 1200 (Agilent, CA, USA) chromatography system running in HILIC mode. HILIC resin was TSK-gel Amide 80 (Tosoh Bioscience, Japan). Samples were separated over a gradient of 100% to 60% organic solvent over 30 min, then 60% to 0% over 15 min. Fractions were collected every minute at absorbances over 500 AU on a UV detector and every 5 min otherwise. All fractions were lyophilized.

Mass Spectrometry

Dry fractions were resuspended in 0.1% FA and loaded onto a Thermo EasyLC. Peptides were eluted using a 0% to 60% organic gradient over 70 min, then 60% to 100% over 20 min. Multiphosphorylated fractions were separated over a gradient proportionally twice this length. All LC−MS runs were performed on 75 μm inner diameter fused silica columns, packed with C18 material (Dr. Maisch, Ammerbuch-Entringen, Germany). Mass spectrometry was performed using higher energy collision fragmentation (HCD) fragmentation on a Thermo LTQ Velos (Thermo Fisher Scientific, Bremen, Germany). Quantification was performed on reporter tags at m/z 114, 115, 116, and 117. MS settings: A full MS scan in the mass area of 400−1800 Da was performed in the Orbitrap with a resolution of 30,000 fwhm and a target value of 1×10^6 ions. For each full scan the seven most intense ions (>+1 charge state) were selected for higher energy collision dissociation (HCD) and detected at a resolution of 7500 fwhm. The settings for the HCD were as follows: threshold for ion selection was 20,000, the target value of ions used for HCD was 1×10^5, activation time was 10 ms, isolation window was 2.5 Da, and normalized collision energy was 48.

Data Analysis

Data was searched using Thermo Proteome Discoverer (version 1.3.0.339) allowing for variable phosphorylation, asparagine deamidation and acetylation, as well as methionine oxidation. iTRAQ labeling was also set as a variable modification in order to detect lysine modification, while cysteine carbamidomethylation was set as a fixed modification. Data were searched against a Swiss-prot rodent database (UniProtKB/Swiss-Prot 2012_10), precursor mass tolerance of 10 ppm, fragment mass tolerance of 0.05 Da. Data were filtered to a 1% false discovery rate using Proteome Discoverer's percolator function, and peptide identifications were filtered to remove those with a Mascot scores less than 18. Data was exported from Proteome Discoverer and manually normalized such that the log2 of iTRAQ ratios displayed a median value of zero for all peptides in a given protein. This was performed across an entire labeling experiment to correct for protein abundance variation. Peptides were then entered into ReportSites software.[22] ReportSites is a Perl program of ∼16,000 lines. Its central function is to overlay and merge the peptide sequences from MS/MS data into a nominated protein database and thereby eliminate redundant modified sites that are discovered more than once in peptide sequences from the same protein. In doing this, the program also retains information about the relative position of each unique PTM site and any quantitative information regarding the peptide and/or site. From this it is possible to centralize each PTM and describe the distributions of other PTMs around it. When this is repeated over many tens of thousands of peptides, maps of relative positions of PTMs are produced. All PTM sites were filtered for site localization (pRS score >75%,[23] only putative glycosylation sites within the eukaryotic glycosylation sequon allowed, internal acetylation sites only[20]), in addition to normal false discovery rate and additional score filtering (Mascot score >18). Glycosylation sites are noted as 'putative' as they are observed after deglycosylation as deamidation sites. When incorporating quantitation each peptide contributes its iTRAQ data to each PTM site that it carries, and these values are averaged for each identified site as part of the redundancy elimination process that is the central function of ReportSites, to produce a net quantitation value for each unique site. In cases where tags were not detected for a particular peptide, that peptide contributes nothing to quantitative analysis for PTM sites it carries. Where peptides mapped to multiple different proteins, priority was given to the highest scoring match. The code also measures the distribution of all nonmodified residues around each site to produce graphs of background residue frequency from which to calculate site occupancy, etc. ReportSites is freely available and useable at ptmtools.portjackson.org. R was used to perform calculations and produce graphs of PTM distribution for heatmaps and PTM flux mapping, and a 95% confidence interval was used to assess all significant changes (in, for example, Figure 3).

DATA DEPOSITION

Raw data is available at: http://mouse-brain-ptmome.s3.amazonaws.com/index.html.

RESULTS AND DISCUSSION

Basic PTM Mapping

One observation made in previous studies of PTM distribution is the need for a specific investigation of PTM relationships in space and time.[6] This work, specifically investigating relationships between multiple PTMs and the dynamics of these relationships in a biological situation, has revealed several observations of interest to PTM biology. Our approach centered around using a custom-made tool to first map PTM sites onto a protein database and remove redundancy from the identifications, and subsequently centralize specific residues (in this case, modified ones) and describe the distribution of all other residues in the data set around these; some will be distant, some close. This allows positional relationships between subsets of residues or PTMs in data sets to be clearly drawn out. When the distributions of confidently detected peptides and PTMs (1% false discovery rate) were mapped using published techniques[22] onto the protein database searched against, some 203,000 redundant modification sites were found, which reduced to 114,000 passing stringent filters for site confidence (see Materials and Methods), which in turn reduced to 14,324 confidently localized, filtered, and unique PTM sites, covering phosphorylation, N-linked glycosylation, and acetylation, distributed across 4,547 proteins (see Figure 2). Of these proteins, 706 bore multiple categories of PTM, and 1,558 individual peptides carried multiple sites of modification. Approximately 90% of identified peptides passing the above-mentioned filters mapped to a single protein (86.4% to *Mus musculus*, 12.6% to *Rattus norvegicus*). Of those peptides with multiple accessions, 2.5% did not include a *Mus musculus*

dx.doi.org/10.1021/pr4002977 | J. Proteome Res. 2014, 13, 260−267

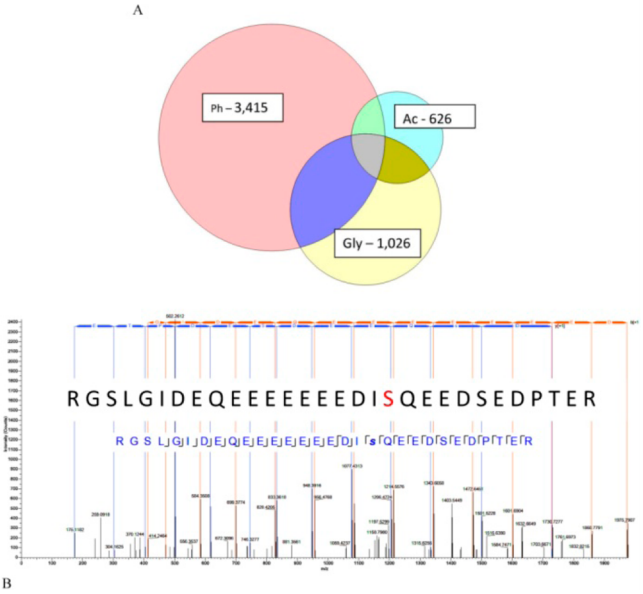


Figure 2. (A) Incidence of and overlap between proteins bearing different PTM types confidently identified in this study. Numbers represent all proteins bearing the given category of PTM. Ph = phosphorylation, Gly = putative glycosylation, Ac = acetylation. All sites passed our confidence filters. (B) Example spectrum of a phosphorylated peptide from IGSF3_MOUSE (Q6ZQA6).

annotation. *Mus musculus* proteins were preferentially used when present.

When the distributions of confidently detected PTMs were mapped in the described way onto the protein database searched against, a list of 14,324 confidently localized, filtered, and unique PTM sites, covering phosphorylation, N-linked glycosylation, and acetylation, was produced (see Supplementary File 1 for peptide sequences). Specifically, our final data included 10,987 phosphorylation sites (pS, pT or pY, pSTY as a group), 1,830 putative N-linked glycosylation sites (glyN), and 1,507 acetylation sites (acK). ReportSites allowed us to map unique PTM sites onto proteins and then to map the distributions of all other PTM sites around each of them, so as to produce a residue frequency-adjusted PTM distribution curve around a centralized residue type. General notation for this is 'surrounding PTM category/central PTM category', e.g. pSTY/glyN for phosphorylation around putative glycosylation.

Within the distribution of this data, evidence of PTM clustering was observed (see Figure 3 for collated data for PTMs around phosphorylation sites; figures with other centralized PTMs in Supplementary File 2), that is, the number of PTMs observed in regions immediately surrounding other PTMs was significantly increased (>2 standard deviations versus background when comparing the proximal 5 vs proximal 30 residues) and followed a well-defined decay pattern with distance from any centralized PTM. The pattern can be clearly seen in Figure 3 as a concentration of green loci around the central point of the distribution. All targeted PTMs displayed clustering tendencies; however, it appears that the pattern is very dependent on the central residue: for instance, glycosylation behaves very differently depending on whether it is analyzed around acK, pSTY, or itself. The strongest clustering was displayed by putative glyN sites around pSTY residues, with a 5.4-fold increase in background-adjusted modification rates when comparing sites 25 residues versus 5 residues away from a PTM, whereas glyN only clustered relatively weakly around itself (see Figure 4 for summary of clustering strength for different central residues). This may be due to steric hindrance from bulky glycan structures, as well as the required consensus sequence for N-linked glycosylation. We generally noted that the effect for all PTM pairs was at its most pronounced within ±5 residues of the centrally aligned PTM. No one PTM as a central site has the same pattern of other PTMs around it over time, suggesting a complex and context-dependent phenomenon. These effects are not unexpected; processes involving many cell types and many thousands of protein isoforms would be expected to display patterns of modification that do not follow simple trends and may reflect the global nature of interdependent PTM deployment irrespective of effects specific to particular cell types or proteins. In line with this, there were no over-

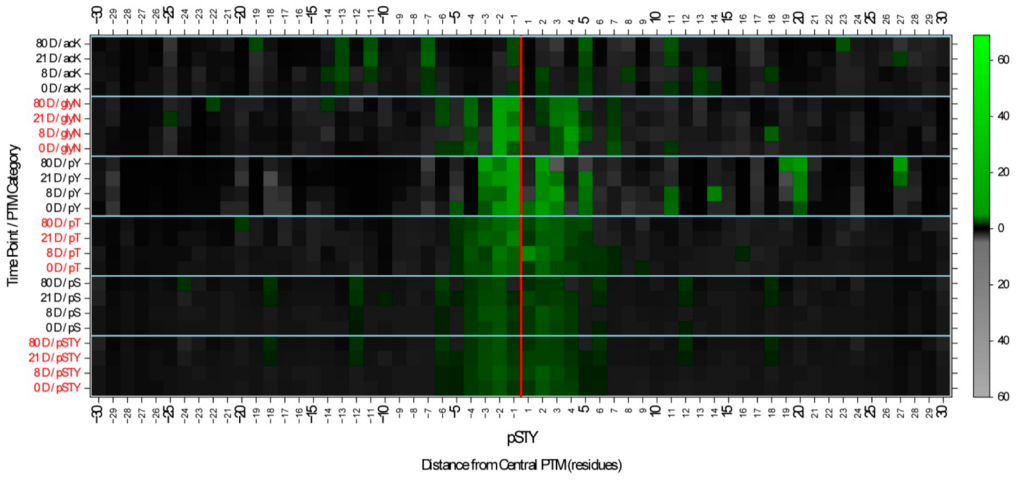


Figure 3. Heat map of PTM distributions around phosphorylation sites. Modified sites in ±30 residue window around centralized modified residues (in this case, phosphorylation, pSTY). *Y*-axis shows the surrounding residues split by PTM category and time point (0, 8, 21, and 80 days). *X*-axis shows the positional loci around the central site of interest (itself marked in red). Green represents a significant increase versus background modification rates (>2 SD), and degree of this increase is shown as *n*-fold change according to the scale shown. Equivalent graphs for remaining PTM relationships are in Supplementary File 2.

dx.doi.org/10.1021/pr4002977 | *J. Proteome Res.* 2014, 13, 260−267

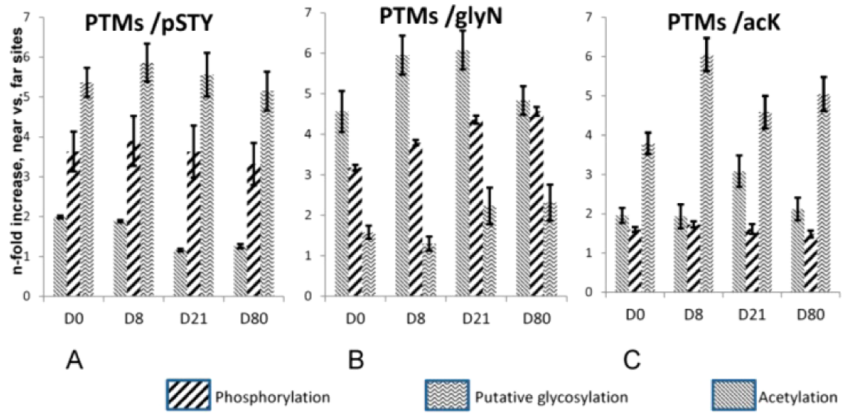


Figure 4. Clustering patterns over time and divided by central residue type showing degree of clustering at different time points and around different PTM categories, showing context-specific patterns. Sites are spread around (A) phosphorylation (/pSTY), (B) putative glycosylation (/glyN), or (C) acetylation (/acK). Values are adjusted for background residue frequency, as well as normalized to protein median abundance. Error bars represent ± 1 standard deviation. Dots indicate changes beyond a 70% confidence interval versus time point 0; asterisks, versus preceding time point. *X*-axis shows time point.

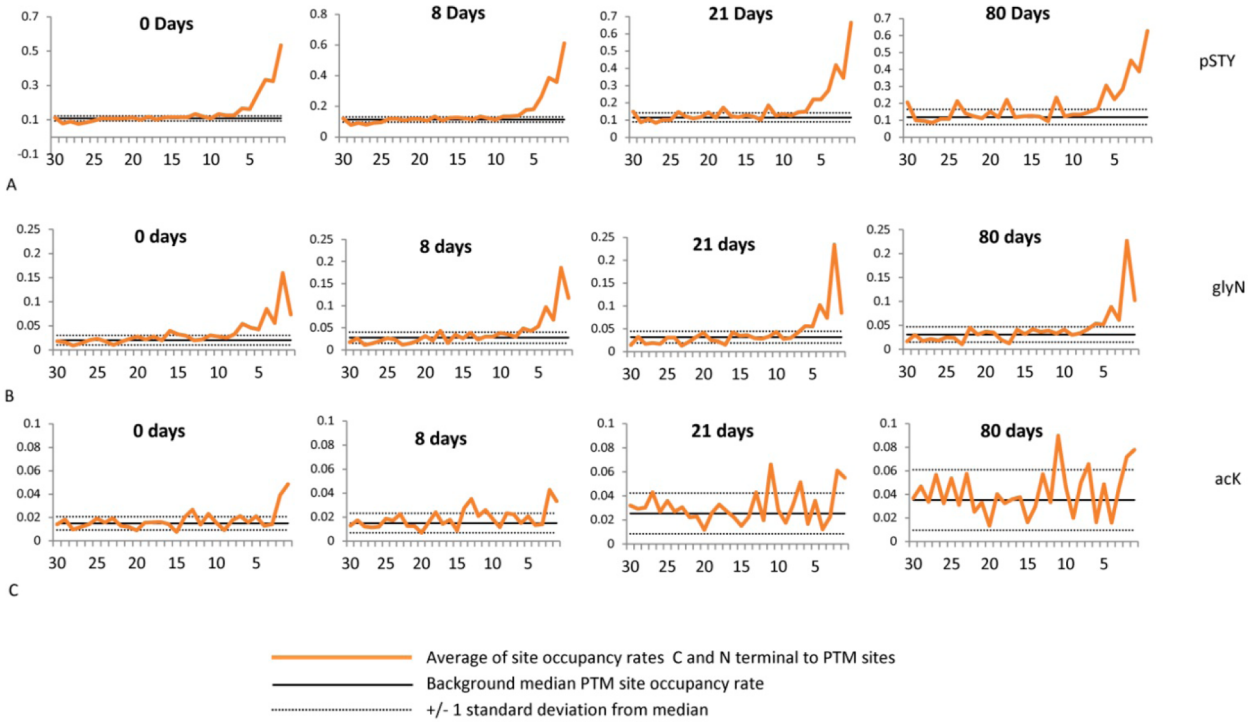


Figure 5. Temporal clustering profiles for PTM sites around phosphorylation sites, showing alterations in degree of PTM clustering over time. (A) Distribution of phosphorylation sites around phosphorylation sites, (B) distribution of glycosylation sites around phosphorylation sites, and (C) distribution of acetylation sites around phosphorylation sites, all assessed by site occupancy rates (number of observed sites divided by number of possible sites) over windows of ±30 residues from the central PTM site. *X*-axis is absolute distance from site, and *Y*-axis is site occupancy rate. See key for color codes. See Supplementary Figure 1 for remaining profiles.

whelmingly clear functional separations between proteins containing multiply modified peptides and singly modified proteins. Gene ontology[24] functional annotations associated with protein or small-molecule binding were well represented in proteins containing multiply modified peptides but no more so than in the data set as a whole, which taken together suggests that the causative process does not target any particular physical or functional proteome regions.

We acknowledge that the methodology used does not require that clustered PTMs be observed in the same scan of the same peptide and that this allows for the possibility of sequence co-localization but separation in physical space (due to compartmentalization, etc.). If this is assumed, however, the question of why it would be necessary for PTMs to co-localize in sequence at all still remains, and the existence of the 3,372 sites on 1,558 peptides where simultaneous PTM was observed is also unexplained. It can generally be seen that even though PTMs are in many cases deployed as a consequence of protein−protein interactions, the driving mechanisms of which are broadly understood, their distributions follow patterns that require further explanation. We also note that while the observations made here are derived from many cell types, they pertain to global effects and as such should not be obscured by the presence of a range of cell types; indeed, it could be argued

dx.doi.org/10.1021/pr4002977 | *J. Proteome Res.* 2014, 13, 260−267

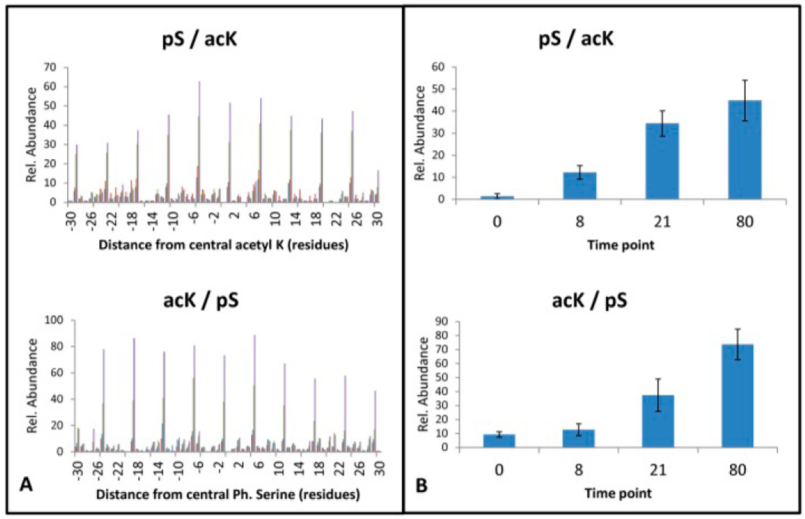


Figure 6. Patterns of acetylation and phosphorylation changing in concert in neurofilament heavy polypeptide (P19246), showing increases in degree of protein modification by two PTMs in concert. Error bars in panel B represent 1 standard deviation above and below the mean.

that any PTM clustering effect would be diluted by diverse cell types, making those that we have observed all the more significant.

Temporal Data

When protein-normalized iTRAQ-labeled peptide data were incorporated into the above analysis, temporal trends in these clustering relationships that could be correlated to developmental stages were seen. For example, Figure 5A shows the progression of PTM clustering around phosphorylation sites from day 0 (newborn) up to day 80 (mature). PTM occurrence fluctuates closely around the background median value at points more distant from the central pSTY, until within 5−10 residues away from the centralized site it begins to rise significantly (>2 standard deviations above the median), indicating an increased frequency of phosphorylation near to other phosphorylation sites. Over time this pattern displays an increase in background 'noise', or background phosphorylation rates, relative to the central peak. This increase or decrease in background PTM frequency, which necessarily corresponds to an increase or decrease in the strength of the clustering effect, is difficult to quantify; however, Figure 4 shows the incidence of what we term 'near' versus 'far' PTM sites with reference to a central PTM (with these definitions being as described above) and shows that the over-representation of phosphorylation sites near to the central phosphorylation site indeed decreases over time. Figure 4 also shows that there are many instances where the over-representation of other PTMs in 'near' areas changes with time by amounts that exceed a 95% confidence interval with respect variously to time point zero, intermediate time points, or both, which altogether demonstrates significant alterations in PTM co-localization over the developmental time frame.

Figure 5B and C shows the same distributions around pSTY for glyN and acK, respectively. It can be seen that acetylation (5C) displays a clear tendency to cluster around phosphorylation and that this effect is decreased over time, as evidenced by a widening of the standard deviation spread to exceed any increase in PTM rates near to the central PTM. With respect to glycosylation distribution around phosphorylation sites, the same clustering pattern is observed, but there does not appear to be any clear temporal trend in glycosylation site frequency around phosphorylation sites over time. This could reflect the functional separation of these two modifications; N-linked glycosylation is thought to be primarily a membrane-associated PTM, whereas phosphorylation is primarily intracellular. It is also known that mechanisms to add or remove phosphate groups are less common in the extracellular environment (though not absent[25]), with the possible consequence that PTM patterns around phosphorylation with respect to predominantly extracellular PTMs such as glycosylation are relatively static. Developmental temporal progressions for other pairs (e.g., glyN around a central acK) are shown in Supplementary Figure 1, and though some relationships do not show strong clustering and some do not show clear effects over development, when examined together with Figure 4 a general trend for increased PTM occurrence close to other PTMs that changes over time can be seen. Overall, these clustering effects suggest a role for deployment of multiple PTMs in 'hotspots' and that these areas are targets of PTM flux over development. This may have significant implications for control of protein function on a large scale.

Finally, as an example of modification of a specific, biologically relevant protein in line with the proposed PTM code, we highlight the PTM status and dynamics of neurofilament heavy polypeptide, NFH. This was the most heavily modified protein detected with 42 instances of PTM, is central to axonal growth, and is known to perform different functions at different stages of development.[26−28] NFH is furthermore known to be phosphorylated at multiple regions at repeated K−S units, entirely in agreement with our observations in which we noted a repeated pattern of serine phosphorylation around other phosphoserines in this protein (Figure 6A). PTMs in this pattern increased in abundance over development, again consistent with published data suggesting developmental phosphoregulation of NFH (Figure 6B). Of more interest to our clustering analysis, however, is that we also observed a nearly identical pattern of PTM distribution and dynamics when acetylated lysines were introduced to PTM analysis of this protein; modified lysines clustered around modified serines and *vice versa* with a separation of 6 residues, demonstrating a pattern of double modification at the repeated KS motif (Figure 6A). Very interestingly, it is clear that PTMs

dx.doi.org/10.1021/pr4002977 | *J. Proteome Res.* 2014, 13, 260−267

occurring within this repeating pattern of phosphoserines and acetylated lysines are increasing in abundance in concert over the time frame investigated (Figure 6B), indicating that the co-occurrence of phosphorylated serines and acetylated lysines in this protein is developmentally favored. The existence of consistent, developmentally targeted multiple modification has interesting implications for protein function when considered against the background of the already known role of phosphorylation for NFH over brain development.[29,30]

CONCLUSIONS

It seems from our data that PTMs tend to closely cluster, not just around other PTMs of the same type but also around PTMs of different types, and that this effect is altered over time. This is observed to occur over a biologically meaningful time frame, during a defined biological process (namely, brain development), and to act on both proteins relevant to that process as well as on a proteome scale. The implications of this are potentially profound; multiple, closely spaced sites of PTM deployment allow for the existence of regulatory hotspots where the function of the modified protein and its interaction partners can be altered in ways that depend on the combination of modifications that are present. This combinatorial coding would provide an extra level of 'computational strength' for a signaling network and may allow for finer control of developmental protein-driven processes in order to produce the appropriate output.[6]

We feel that the observed PTM clustering may be a reflection of the sought-after 'PTM code', a level of cellular regulatory machinery over and above those currently characterized. This effect could greatly expand the potential for close control of protein structure and function, and while further work is needed to more fully describe the extent of the observed patterns, it is extremely interesting to consider the potential impact of such a phenomenon on cell biology.

ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Ph: +45 6550 234. E-mail: mrl@bmb.sdu.dk.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

A.V.G.E. is the recipient of a Lundbeck Foundation project grant in biomedicine and an Amazon Web Services AWS in Education Research Grant. V.S. acknowledges financial support from the Danish Council for Independent Research, Natural Sciences. M.R.L. is the recipient of a Lundbeck Foundation Junior Group Leader Fellowship. The Lundbeck Foundation provided funds to purchase mass spectrometry infrastructure.

ABBREVIATIONS

PTM, post-translational modification; iTRAQ, isobaric tags for relative and absolute quantitation; ACN, acetonitrile; TFA, trifluoroacetic acid; FA, formic acid; HILIC, hydrophilic interaction chromatography; SIMAC, sequential elution from IMAC; pS, T, or Y, serine/threonine or tyrosine phosphorylation; AcK, lysine acetylation; GlyN, putative N-linked glycosylation

REFERENCES

(1) Xin, F.; Radivojac, P. Post-translational modifications induce significant yet not extreme changes to protein structure. *Bioinformatics* **2012**, *28*, 2905−13.
(2) Riemen, A. J.; Waters, M. L. Dueling post-translational modifications trigger folding and unfolding of a beta-hairpin peptide. *J. Am. Chem. Soc.* **2010**, *132*, 9007−13.
(3) Brooks, C. L.; Gu, W. Ubiquitination, phosphorylation and acetylation: the molecular basis for p53 regulation. *Curr. Opin. Cell Biol.* **2003**, *15*, 164−71.
(4) Tweedie-Cullen, R. Y.; Reck, J. M.; Mansuy, I. M. Comprehensive mapping of post-translational modifications on synaptic, nuclear, and histone proteins in the adult mouse brain. *J. Proteome Res.* **2009**, *8*, 4966−82.
(5) Munton, R. P.; Tweedie-Cullen, R.; Livingstone-Zatchej, M.; Weinandy, F.; Waidelich, M.; Longo, D.; Gehrig, P.; Potthast, F.; Rutishauser, D.; Gerrits, B.; Panse, C.; Schlapbach, R.; Mansuy, I. M. Qualitative and quantitative analyses of protein phosphorylation in naive and stimulated mouse synaptosomal preparations. *Mol. Cell. Proteomics* **2007**, *6*, 283−93.
(6) Creixell, P.; Linding, R. Cells, shared memory and breaking the PTM code. *Mol. Syst. Biol.* **2012**, *8*, 598.
(7) Minguez, P.; Letunic, I.; Parca, L.; Bork, P. PTMcode: a database of known and predicted functional associations between post-translational modifications in proteins. *Nucleic Acids Res.* **2013**, *41*, D306−11.
(8) Mischerikow, N.; Spedale, G.; Altelaar, A. F.; Timmers, H. T.; Pijnappel, W. W.; Heck, A. J. In-depth profiling of post-translational modifications on the related transcription factor complexes TFIID and SAGA. *J. Proteome Res.* **2009**, *8*, 5020−30.
(9) Drogaris, P.; Wurtele, H.; Masumoto, H.; Verreault, A.; Thibault, P. Comprehensive profiling of histone modifications using a label-free approach and its applications in determining structure-function relationships. *Anal. Chem.* **2008**, *80*, 6698−707.
(10) Sims, R. J., 3rd; Reinberg, D. Is there a code embedded in proteins that is based on post-translational modifications? *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 815−20.
(11) Minguez, P.; Parca, L.; Diella, F.; Mende, D. R.; Kumar, R.; Helmer-Citterich, M.; Gavin, A. C.; van Noort, V.; Bork, P. Deciphering a global network of functionally associated post-translational modifications. *Mol. Syst. Biol.* **2012**, *8*, 599.
(12) Beltrao, P.; Albanese, V.; Kenner, L. R.; Swaney, D. L.; Burlingame, A.; Villen, J.; Lim, W. A.; Fraser, J. S.; Frydman, J.; Krogan, N. J. Systematic functional prioritization of protein posttranslational modifications. *Cell* **2012**, *150*, 413−25.
(13) Reed, N. A.; Cai, D.; Blasius, T. L.; Jih, G. T.; Meyhofer, E.; Gaertig, J.; Verhey, K. J. Microtubule acetylation promotes kinesin-1 binding and transport. *Curr. Biol.* **2006**, *16*, 2166−72.
(14) Verhey, K. J.; Gaertig, J. The tubulin code. *Cell Cycle* **2007**, *6*, 2152−60.
(15) Nussinov, R.; Tsai, C. J.; Xin, F.; Radivojac, P. Allosteric post-translational modification codes. *Trends Biochem. Sci.* **2012**, *37*, 447−55.
(16) Revil, T.; Gaffney, D.; Dias, C.; Majewski, J.; Jerome-Majewska, L. A. Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics* **2010**, *11*, 399.
(17) Engholm-Keller, K.; Birck, P.; Storling, J.; Pociot, F.; Mandrup-Poulsen, T.; Larsen, M. R. TiSH - a robust and sensitive global phosphoproteomics strategy employing a combination of TiO(2), SIMAC, and HILIC. *J. Proteomics* **2012**, *75*, 5749−61.
(18) Thingholm, T. E.; Jensen, O. N.; Larsen, M. R. Enrichment and separation of mono- and multiply phosphorylated peptides using sequential elution from IMAC prior to mass spectrometric analysis. *Methods Mol. Biol.* **2009**, *527*, 67−78 xi..

(19) Thingholm, T. E.; Larsen, M. R. The use of titanium dioxide micro-columns to selectively isolate phosphopeptides from proteolytic digests. *Methods Mol. Biol.* **2009**, *527*, 57−66 xi.
(20) Choudhary, C.; Kumar, C.; Gnad, F.; Nielsen, M. L.; Rehman, M.; Walther, T. C.; Olsen, J. V.; Mann, M. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **2009**, *325*, 834−40.
(21) Thingholm, T. E.; Jensen, O. N.; Robinson, P. J.; Larsen, M. R. SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Mol. Cell. Proteomics* **2008**, *7*, 661−71.
(22) Edwards, A. V. G.; Edwards, G.; Larsen, M. R.; Cordwell, S. J. ReportSites - a computational method to extract positional and physico-chemical information from large-scale proteomic post-translational modification datasets. *J. Proteomics Bioinform.* **2012**, *5*, 104−07.
(23) Taus, T.; Kocher, T.; Pichler, P.; Paschke, C.; Schmidt, A.; Henrich, C.; Mechtler, K. Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* **2011**, *10*, 5354−62.
(24) Ashburner, M.; Ball, C. A.; Blake, J. A.; Botstein, D.; Butler, H.; Cherry, J. M.; Davis, A. P.; Dolinski, K.; Dwight, S. S.; Eppig, J. T.; Harris, M. A.; Hill, D. P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J. C.; Richardson, J. E.; Ringwald, M.; Rubin, G. M.; Sherlock, G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25−9.
(25) Yalak, G.; Vogel, V. Extracellular phosphorylation and phosphorylated proteins: not just curiosities but physiologically important. *Sci. Signaling* **2012**, *5*, re7.
(26) Lariviere, R. C.; Julien, J. P. Functions of intermediate filaments in neuronal development and disease. *J. Neurobiol.* **2004**, *58*, 131−48.
(27) Marszalek, J. R.; Williamson, T. L.; Lee, M. K.; Xu, Z.; Hoffman, P. N.; Becher, M. W.; Crawford, T. O.; Cleveland, D. W. Neurofilament subunit NF-H modulates axonal diameter by selectively slowing neurofilament transport. *J. Cell Biol.* **1996**, *135*, 711−24.
(28) Trinidad, J. C.; Thalhammer, A.; Specht, C. G.; Lynn, A. J.; Baker, P. R.; Schoepfer, R.; Burlingame, A. L. Quantitative analysis of synaptic phosphorylation and protein expression. *Mol. Cell. Proteomics* **2008**, *7*, 684−96.
(29) Ackerley, S.; Thornhill, P.; Grierson, A. J.; Brownlees, J.; Anderton, B. H.; Leigh, P. N.; Shaw, C. E.; Miller, C. C. Neurofilament heavy chain side arm phosphorylation regulates axonal transport of neurofilaments. *J. Cell Biol.* **2003**, *161*, 489−95.
(30) Tonnaer, E. L.; Peters, T. A.; Curfs, J. H. Neurofilament localization and phosphorylation in the developing inner ear of the rat. *Hear. Res.* **2010**, *267*, 27−35.