Supplementary material

Time-resolved Global and Chromatin Proteomics During Herpes Simplex Virus Type 1 (HSV-1) Infection

Katarzyna Kulej ¹,², Daphne C. Avgousti ¹,², Simone Sidoli ³,⁴, Christin Herrmann ²,⁵, Ashley N. Della Fera ², Eui Tae Kim ¹,², Benjamin A. Garcia ³,⁴ *, and Matthew D. Weitzman ¹,² *

¹ Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA USA
² Division of Cancer Pathobiology, Children’s Hospital of Philadelphia, Philadelphia, PA USA
³ Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA USA
⁴ Epigenetics Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA USA
⁵ Cell and Molecular Biology Graduate Group, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA USA

* Correspondence: weitzmanm@email.chop.edu; bgarci@mail.med.upenn.edu

Key words: Herpes simplex virus (HSV-1) infection / phosphoproteomics / histone modifications / chromatin dynamics / post-translational modifications (PTMs) / viral life cycle / lytic infection
1. Time course changes in host proteomic response to infection

To further define the 16 clusters of the total proteome, we used GeneGo's MetaCore software (FDR < 0.05) to determine the pathways represented in each cluster. The 9 major categories found to be enriched overall were RNA metabolic processes, DNA metabolic processes, gene expression, chromatin organization, cell cycle, intracellular transport, response to stress, cell death, and other (Figure 2B). The mRNA and ncRNA processing pathways were highly enriched in both early time points (Figure S3; clusters 1, 5, 7, 9) and late time points (clusters 2, 8, 11) of HSV-1 infection. HSV-1 proteins have been shown to enhance degradation of some mRNA levels while stabilizing or delaying degradation of other host mRNA levels, consistent with the two distinct groups identified here (100). Regulation of mRNA stability was enriched in multiple clusters with different trends (cluster 1, 8 and 9). Ribosome biogenesis and regulation of translation were mainly present in clusters peaking at 9 hpi (clusters 2, 8, 11), which corresponds to late gene expression. Other translation related-proteins were enriched at 3-6 hpi (clusters 1, 4), which may represent events regulated by HSV-1 to inhibit host protein translation. For instance, down regulation of host RNA splicing was observed late during infection (clusters 1, 12 and 15), a function known to be regulated by the viral protein ICP27 (108, 109). Nuclear protein transport and localization were GO functions enriched at both early and very late time points of infection (clusters 3, 9, 13, 14 and 16), which may reflect import and export of virions. Proteins involved in regulation of immune response, including antigen processing and presentation, were highly enriched in clusters 3, 10 and 12 (p-value enrichment < E-22, < E-19, and < E-18, respectively). The two latest clusters showed monotonic decrease across six time points, while cluster 3 decreases after 6 hpi. The decrease of pro-inflammatory proteins may be regulated by the ICP34.5 protein, a key viral factor interfering with the interferon β (IFN-β) pathway (75, 110). Within cluster 3 we identified proteins that act as inhibitors of early stages of viral life cycle e.g. IFIT3, HLA-A, HLA-C, NOTCH1, RELA or LYN. The categories described here are examples of the depth of our data that are consistent with expected changes during infection. Since each cluster contains approximately 270 proteins, the known proteins can serve as a guide to investigate viral regulation of other host proteins in the same clusters.

The cellular DNA damage response and DNA repair pathways have been shown to be manipulated by multiple HSV-1 proteins (111, 112). DNA damage response and cellular response to stress were over-represented within cluster 1 which rises early in infection, and then subsequently appeared in clusters 3, 4 and 15, which all show decreases at late times. Previous reports have demonstrated that the MRN complex which has multiple essential functions within the DNA damage and repair pathway is recruited into HSV-1 viral replication compartments where it favors HSV-1 replication (22, 24, 113, 114). In our current study, the levels of Rad50 and MRE11A were found in clusters 1 and 4, increased at 3-6 hpi compared to mock and then decreased at the late time points of infection. Thus our study reveals further details of the temporal regulation of DNA repair pathway proteins by HSV-1.

Apoptosis, or programmed cell death, is an important physiological process for host defense against viral infection, which occurs at early stages of viral infection in order to limit viral propagation (115). Viruses may
also trigger apoptosis at the late stage of viral replication to facilitate viral release and spread. Apoptotic events were regulated in a dynamic manner across all time-points of HSV-1 infection but can divided into two groups: (I) pro-apoptotic and (II) anti-apoptotic processes. Positive regulation of apoptotic processes was highly enriched in clusters 1, 3, 4 and 14. This group of proteins increases primarily at early time points (3-6 hpi) with a dramatic decrease after 6 hpi. Enrichment of pro-apoptotic proteins was also detected in cluster 13, 14 and 16 which peaked at 15 hpi. Proteins involved in positive regulation of cell death processes were also enriched in cluster 8 and cluster 11. Moreover, cluster 8 was unique in high enrichment of proteins involved in apoptotic nuclear changes. The increase of apoptosis category specifically after 6 hpi, when viral replication reaches maximum efficiency, may support cellular chromatin margination and re-localization of viral replication compartments (115, 116). To overcome host resistance, HSV-1 carries anti-apoptotic factors e.g. ICP27, (117), which we found primarily enriched in cluster 1 and 3, decreasing after 6 hpi.

2. Comparison of the described dataset with other virus proteomic studies

A number of prior studies have provided insight into the dynamic proteome during HSV-1 infection. Berard et al. (100) performed analysis of cytosolic and nuclear proteins (Human Embryonic Kidney cell line, HEK293) across a time-course (4, 10 and 24 hpi) of HSV-1 infection using quantitative stable isotope labeling by amino acids in cell culture (SILAC) approach in combination with liquid chromatography – tandem MS (LC-MS/MS). In total, our study covered 77% (242 proteins) of the differentially regulated proteins identified in their study (Table S1A). The remaining 23% (72 proteins) may be due to differences in cell types and time points used in each study. Other HSV-1 proteomic research focused on specific cellular compartments such as ribosomes, microsomes, and the nucleus (118-120). Prior to our work, the whole cell extract analysis of HSV-1 infection was performed by Antrobus et al. (81) on HEp-2 cells (human epithelial larynx carcinoma cell line). Using 2-DE and LC-MS/MS they identified 68 differentially regulated host proteins at 6 hpi, of which 64 were also detected in our study. They confirmed up-regulation of some proteins such as SFQ, NDRG1, KHSRP, HSPA5, down-regulation of CBX3 and MCM4, and no-change for VIM and MATR3 protein by immunoblotting. These results are consistent with our data, with the exception of MATR3 which we found up-regulated at 6 hpi. Our clustering analysis showed no-change for VIM protein up to 6 hpi, while after 6 hpi its abundance gradually increased (Figure S3; cluster 6). Berard et al. (100) also confirmed up-regulation of VIM at 24 hpi, especially in the cytosolic fraction. It has been reported that the phosphorylation of MATR3 at T150 or S188 affects its intranuclear and nuclear localization, particularly in late HSV-1 infection (121), which may account for the difference with our dataset. It is possible that MATR3 compartmentalization may limit its solubility and thus detection. Loret et al. described potential incorporation of up to 59 distinct cellular proteins (including protein subtypes e.g. α, β, γ) in mature extracellular HSV-1 virions (68, 122). In our study we identified 49 of those proteins (e.g. PPIA, NME1, CSNK2A1, UBE2L3, HSPA8 and CFL1) and the majority were present in clusters with the highest intensity at 15 hpi (Table S1A; Figure S3, clusters 13 and 14). Other proteins enriched at the late time points include peroxiredoxin-1 (PRDX1) and protein S100-A11 (S100A11) (cluster 16), macrophage migration
inhibitory factor (MIF), 14-3-3 protein zeta/delta (YWHAZ) and Ras-related proteins Rab-1A/5A (RAB1A/RAB5A) (cluster 8). S100A11 and MIF have inflammatory functions in extracellular space and potentially facilitate HSV-1 access into the host cell (123), while Rab-1A and Rab-5A are involved in virion assembly and viral RNA genome replication, respectively (124). This suggests that a sudden increase in protein abundance at the latest time points of HSV-1 infection is not an artifact of MS-quantification or due to cell death, but may be highly regulated by HSV-1 in order to incorporate selected proteins into mature virions to accelerate early viral gene expression during subsequent infections. Taken together, our analysis of the host and viral proteome faithfully reproduces known events during infection, overlaps with previous proteomic studies, and provides novel temporal information and depth to HSV-1 infection proteomics.

Supplementary figure legends

Figure S1: Qualitative analysis of biological replicates for proteome and phosphoproteome. Three dimensional PCA of all replicates of the proteome (A) and phosphoproteome (B) analysis of host proteins. Pearson correlation of all replicates of the host (C) and HSV-1 viral (D) datasets. The darker colors represent higher correlation between two analyzed conditions.

Figure S2: Radar plot of Gene Ontology biological process enrichment for proteins detected in four or fewer conditions. Proteins identified in fewer than five time points of HSV-1 infection (286 proteins) were excluded from fuzzy c-means clustering. The majority of those proteins were involved in cellular immune response processes. The circular axis represents p-value (-Log10) enrichment obtained from MetaCore analysis (FDR < 0.05).

Figure S3: Cluster analysis of host proteins during viral infection. Fuzzy c-means clustering calculated 16 as the ideal number of cluster groups to demonstrate protein trends. The y-axis represents z-score of protein abundance. Color coding from purple to pale blue represents density of lines, i.e. number of proteins. Only proteins that belong to a specific cluster in a significant manner are represented. Proteins with trends that could fit multiple groups did not pass the significance threshold. The empty group (cluster 9) did not achieve any protein that was significantly enriched in the cluster trend.

Figure S4: Absolute intensities for host and viral proteins. The absolute raw intensity (iBAQ) was calculated for host (blue) and viral (red) proteins, and is displayed ranked on the x-axis from the most abundant to the least intense. As expected, viral proteins become higher in the rank at later time points.

Figure S5: Abundance and distribution of phosphosites on host and viral proteins. Distribution of absolute intensity (iBAQ) of viral and host proteins across the six time points. The blue central line represents the median; the pale blue border marks the first and third quartile (25% and 75%, respectively). Plots are shown for viral proteins (A), viral phosphosites (B), host proteins (C) and host phosphosites (D).

Figure S6: Clustering analysis of phosphorylation sites identified on host proteins during HSV-1 infection. Fuzzy c-means clustering identified 14 cluster groups for the phosphorylation trends. The y-axis
Figure S7: Clustering analysis of phosphorylation sites identified on HSV-1 proteins during viral infection. Fuzzy c-means clustering calculated 8 ideal cluster groups to divide the phosphorylation trends. The y-axis represents z-score of phosphorylation abundance. Color coding from purple to pale blue represents density of lines, i.e. number of phosphosites. Only phosphosites that belong to a specific cluster in a significant manner are represented. Those with trends that could fit multiple groups did not pass the significance threshold. The empty group (cluster 4) did not achieve any phosphosite that was significantly enriched in the cluster trend.

Figure S8: Proteins identified as enriched in high salt chromatin fraction. (A) Log₂ transformed absolute intensity (iBAQ) of HSV-1 proteins enriched in the high salt purification. Color coding represents protein intensity. (B and C) Volcano plot representing the fold change and significance of total identified host chromatin protein abundance at 4 hpi vs mock (B) and 8 hpi vs mock (C). The standard significant threshold (t-test p-value < 0.05) transformed in –Log₁₀ is 1.30. ‘Light grey dots’ highlight t-test p-value > 0.05.

Figure S9: Histone peptide distribution and representative example of modified forms across infection time points. (A) Multi-scatterplot comparing all conditions displaying the correlation of the different conditions. Replicates were averaged. Blue number represents Pearson correlation. (B) 3D PCA representing the distance between conditions and replicates. The PCA was generated by using the relative abundance (in %) of the histone peptides.

Figure S10: Heatmaps of all modified histone peptides during HSV-1 infection. Histone PTMs are divided into histone H3 and H4 (left) and H2A, H2B and H1 (right). The relative abundance (in %) of the modified peptides was z-score normalized prior to representation in order to simplify the visualization of how peptides change in abundance across the analyzed time points.

Figure S11. Integration of epi-proteomics data. Bar plots highlighting the time-resolved regulation patterns of selected host and viral proteins and their PTMs during HSV-1 infection. (A) LMNA. (B) UL31. (C) UL31(pS24), UL31(pS26), UL31(pS27). Error bars represent standard deviation (n=3).

Supplementary table legends

Table S1. Time-course monitoring of changes in host protein abundance during HSV-1 infection. (A) Total list of identified and quantified host proteins across six time points of virus infection. Protein IDs refer to UniProt database; Peptide counts highlights the number of razor and unique peptides used for protein quantification; Sequence coverage [%] represents the percentage of the protein’s sequence covered by the peptides identified in the MS run; Cluster number refers to Figure S3 which highlights the
trend of protein regulation across six time points of HSV-1 infection (N/A – non-clustered proteins); 

\textit{isClusterMember} determines whether the protein belongs significantly to the cluster assigned; \textit{ANOVA p-value} describes the ANOVA p-value of protein levels; \textit{Protein iBAQ intensity} highlights the MS-measured protein abundance (average of biological replicates, n=3); \textit{Comparison to previous studies} highlights host proteins identified in previous large-scale proteomics studies (81, 100) (see also \textbf{Supplementary information section 2}) as significantly regulated during HSV-1 infection ('(+)' if identified). (B) List of host proteins subjected to fuzzy c-means clustering analysis. Only proteins identified in minimum five conditions were subjected to fuzzy c-means clustering analysis. (C) List of non-clustered host proteins. Proteins highlighted in ‘red’ were identified uniquely in control (mock) and at the early time points of HSV-1 infection (3-6 hpi); in ‘green’ are shown proteins identified specifically at late time points (9-15 hpi) (D) List of proteins involved in chromatin organization based on the annotation retrieved from MetaCore and UniProt database knowledge. Proteins with ANOVA p-value < 0.05 (highlighted in ‘red’) were considered for further analysis in this study.

\textbf{Table S2. Time-course monitoring of changes in HSV-1 protein abundance across six time points of HSV-1 infection in HFF cells.} (A) Total list of identified and quantified HSV-1 proteins across five time points of HFF cell infection. \textit{Protein IDs} refer to UniProt database; \textit{Peptide counts} highlights the number of razor and unique peptides used for protein quantification; \textit{Sequence coverage [%]} represents the percentage of the protein’s sequence covered by the peptides identified in the MS run; \textit{ANOVA p-value} describes the ANOVA p-value of protein levels; \textit{Protein iBAQ intensity} highlights the measured protein abundance (average of biological replicates, n=3); \textit{Gene category} describes the classification of HSV-1 proteins within the IE, E or L type of genes; \textit{Structural components} highlights proteins identified as a part of viral capsid, tegument or envelope; \textit{Mature virions} highlights proteins identified in mature extracellular virions based on the study of Loret et al. (68) (N/A – uncertain localization); \textit{Monotonic increase across five time points} – describes proteins with continuously increasing abundance across all analyzed time points. (B) The HSV-1 proteins not identified in this study (referring to the UniProt reviewed \textit{Human herpesvirus 1 (strain 17; 17syn+) (HHV-1)} database).

\textbf{Table S3. Time-course monitoring of changes in host phosphorylation site abundance during HSV-1 infection.} (A) Total list of identified and quantified host phosphosites across six time points of virus infection. \textit{Protein IDs} refer to UniProt database; \textit{Mod site} highlights the phosphorylated S/T or Y residue within protein amino acid (aa) sequence; \textit{Sequence window} highlights the 15 aa residue within protein aa sequence prior and after modified site; \textit{Peptide sequence} visualizes the identified phosphopeptide sequence; \textit{Position peptide} highlights the phosphorylated S/T or Y residue within identified phosphopeptide sequence; \textit{Localization prob} is the confidence score for site localization of the phosphorylation (100% means fully unambiguous); \textit{Cluster number} refers to \textbf{Figure S6} which highlights the trend of phosphosite regulation across six time points of HSV-1 infection (N/A – non-clustered phosphosites); \textit{isClusterMember} determines whether the protein belongs significantly to the cluster assigned; \textit{ANOVA p-value} describes the ANOVA p-value of phosphorylation levels; \textit{Phosphosite intensity} highlights the MS-measured phosphosite
abundance (average of biological replicates, n=3). (B) List of host phosposites subjected to fuzzy c-means clustering analysis. Only phosposites identified in minimum four conditions were subjected to fuzzy c-means clustering analysis. (C) List of non-clustered host phosposites.

Table S4. Time-course monitoring of changes in phosphorylation site abundance on HSV-1 proteins during HFF infection. (A) Total list of identified and quantified phosposites on HSV-1 proteins across five time points of virus infection. Protein IDs refer to UniProt database; Mod site highlights the phosphorylated S/T or Y residue within protein amino acid (aa) sequence; Sequence window highlights the 15 aa residue within protein aa sequence prior and after modified site; Peptide sequence visualizes the identified phosphopeptide sequence; Position peptide highlights the phosphorylated S/T or Y residue within identified phosphopeptide sequence; Localization prob is the confidence score for site localization of the phosphorylation (100% means fully unambiguous); Cluster number refers to Figure S7 which highlights the trend of phososite regulation across six time points of HSV-1 infection (N/A – non-clustered phosposites); isClusterMember determines whether the protein belongs significantly to the cluster assigned; ANOVA p-value describes the ANOVA p-value of phosphorylation levels; Phosphosite intensity highlights the MS-measured phosphosite abundance (average of biological replicates, n=3); Previously reported HSV-1 phosphosites - phosphorylation sites of HSV-1 proteins identified in previous large- and small-scale studies (‘(+)’ if identified) (14, 15, 18, 81, 107, 125-147). (B) List of HSV phosphosites normalized by protein intensity. Only phosposites identified in minimum two out of three biological replicates, and normalized by protein abundance were subjected to fuzzy c-means clustering analysis. (C) List of not-normalized and non-clustered HSV-1 phosphosites.

Table S5. Time-course monitoring of changes in chromatin associated host proteome during HSV-1 infection. (A) Total list of host proteins quantified during HSV-1 infection at mock, 4 hpi and 8 hpi by MS-analysis of high salt fractions. Table includes the log₂ fold change iBAQ values obtained for protein at 4 hpi vs mock and 8 hpi vs mock for highest salt fractions (600 mM NaCl). Proteins with paired sample t-test p-value smaller than 0.05 were considered as significantly altered between the two tested conditions. The number of razor and unique peptides used for quantification was also highlighted. Sequence coverage [%] represents the percentage of the protein’s sequence covered by the peptides identified in the MS run. (B) List of host proteins involved in chromatin organization processes retrieved from MetaCore database (FDR < 0.05).

Table S6. Time-course monitoring of changes in chromatin associated viral proteins during HSV-1 infection. Total list of HSV-1 proteins quantified during HSV-1 infection at 4 hpi and 8 hpi by MS-analysis of high salt fractions. Table includes the log₂ fold change iBAQ values obtained for protein at 8 hpi vs 4hpi for highest salt fraction (600 mM NaCl). Proteins with paired sample t-test p-value smaller than 0.05 were considered as significantly altered between the two tested conditions. The number of razor and unique peptides used for quantification was also highlighted. Sequence coverage [%] represents the percentage of the protein’s sequence covered by the peptides identified in the MS run. Gene category column describes
the classification of HSV-1 proteins within the IE, E or L type of genes. **Structural components** highlights proteins identified as a part of viral capsid, tegument or envelope. **Mature virions** highlights proteins identified in mature extracellular virions based on the study of Loret et al. (68) (N/A – uncertain localization).

**Table S7. Quantification of modified histone peptides during infection.** (A) Relative abundance of histone peptides detected across all time points and the calculated ANOVA p-value from three biological replicates. (B) Deconvoluted single marks from table A. For simplicity, the relative abundance of single PTMs was extracted by summing all peptides carrying the given mark.

**Supplementary figures**

**Figure S1.**
Figure S2.
Figure S3.
Figure S4.
Figure S5.
Figure S6.

Cluster 1
z-normalized phospho sites intensities
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
Cluster 13
Cluster 14
Figure S7.
Figure S8.

A

B

C
Figure S9.
Figure S10.
Figure S11.

A. UL31 protein intensity over time.

B. Phospho-UL31 protein intensity over time.

C. LMNA protein intensity over time.
References


75. Wang, Y., Yang, Y., Wu, S., Pan, S., Zhou, C., Ma, Y., Ru, Y., Dong, S., He, B., Zhang, C., and Cao, Y. (2014) p32 is a novel target for viral protein ICP34.5 of herpes simplex virus type 1 and facilitates viral nuclear egress. *Journal of Biological Chemistry* 289, 35795–35805


107. Mou, F., Wills, E., and Baines, J. D. (2009) Phosphorylation of the U(L)31 protein of herpes simplex virus 1 by the U(S)3-encoded kinase regulates localization of the nuclear envelopment complex and egress of nucleocapsids. *Journal of Virology* 83, 5181–5191


130. Potel, C., and Elliott, G. (2005) Phosphorylation of the herpes simplex virus tegument protein VP22 has no effect on incorporation of VP22 into the virus but is involved in optimal expression and virion packaging of ICP0. Journal of Virology 79, 14057–14068


137. O'Reilly, D., Hanscombe, O., and O'Hare, P. (1997) A single serine residue at position 375 of VP16 is critical for complex assembly with Oct-1 and HCF and is a target of phosphorylation by casein kinase II. EMBO J. 16, 2420–2430


