EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Contact: simon.hackl@uni-tuebingen.de

PTMVision: An Interactive Visualization Platform for Post-Translational Modifications of Proteins

Hackl, S.[†], Jachmann, C.[†], Harbig, T.A.[†], Witte Paz, M.[†], Nieselt, K.[†]

Motivation

Post-translational modifications (PTMs) (chemical modifications after protein translation) **play a key role in biological processes**[1,2,3]. Analysis of PTMs in large protein datasets revealed that **proteins are modified** more **frequently** than previously thought[4,5,6]. Visual analysis of these data using **traditional visualization methods**[7,8] **suffers several issues**, especially when spatial information is important.

Lollipop Plot PTMs along primary sequence
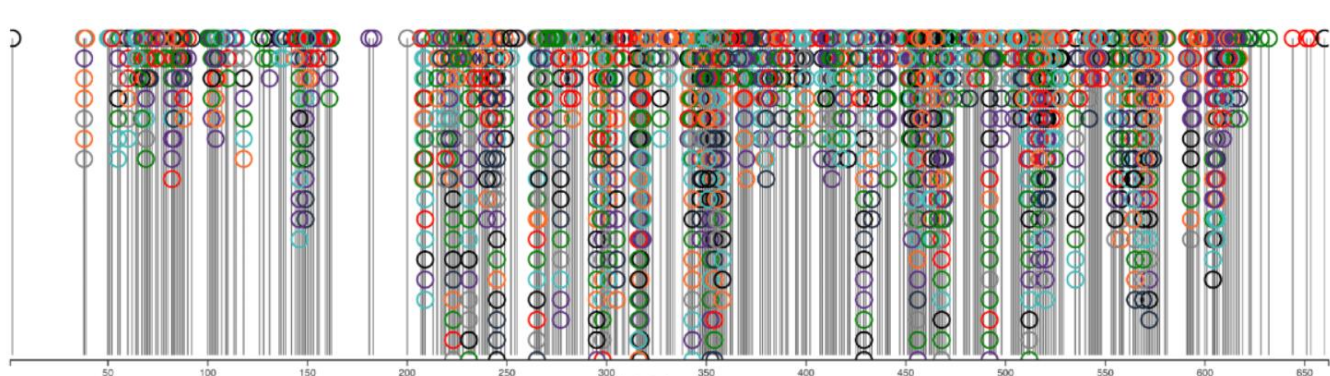
3D Structure Color coded number of PTMs.
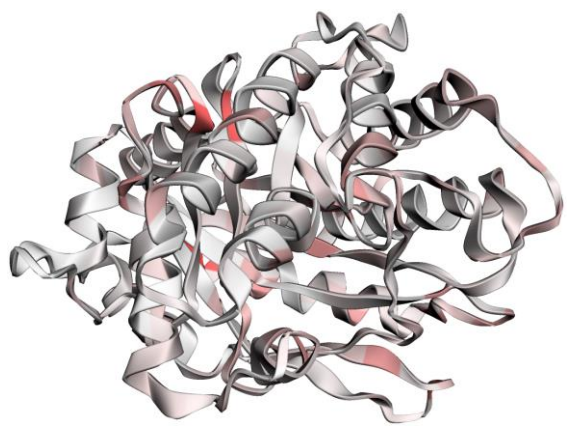
Issues

Visual Overload (ol)
Per Site Restriction (sr)
No Spatial Info (si)

Poor Overview (ov)
Hard to Interact (in)

The challenge in visualization is to find a clear representation of PTMs along the primary sequence of a single protein, taking into account its 3D structure.

Design

Together with experts we defined user tasks[9,10] to find a new way to visualize PTMs and tackle above issues:

- **Identify** sites with PTMs (IS) or of specific PTMs (IM). Highlight sites along the protein sequence (HS) and filter PTMs according to specified properties (FM).
- **Summarize** PTM compositions of a protein site (SS).
- **Compare** occurrence of modifications across the entire protein sequence (CM) or PTM composition of (contacting) protein sites (CS).

Design Choice	Tasks	Solves
Interactive 2D representations	-	in
Contact map (structure encoding)	FM, CS	sr, si
Presence/absence heat map	IM, CM	ov
Line charts	IS, SS	ol
Single axis charts	IS, HS	ol
Sunburst charts	CS	ol

We showcased an initial **prototype at the IEEE Vis 2022 Bio+MedVis Challenge**. Our current **visualization** is available as an interactive dashboard in the form of a **web app**.

Prospects

Based on the feedback received during the Bio+MedVis Challenge and a second evaluation panel of domain experts, we identified the remaining challenges:
- In terms of Shneiderman's *details-on-demand* paradigm, the visualization of
  - the full structural context (i.e., one row/column of the contact map) of neighboring amino acids in relation to PTMs
  - non-successive sites in a comparative manner.
- Third evaluation by domain experts.

Visualization

❶ Presence/absence map
PTMs (x-axis) along the primary protein sequence (y-axis). Identify sites where specific PTMs occur and compare PTM patterns along the primary sequence. Zooming the x-axis allows a detailed view.

❷ Contact map
Projection of protein 3D structure to 2D. Filled cells indicate positions in close proximity.
Upper triangle: Filtered based on PTM related properties, e.g. shared modifications (red → higher share). Linked x- any y-axis zooming to view sites in detail.

❸ Single axis chart:s
Linked to the x- and y axis of contact map. Highlight and identify (user defined) sites of interest.
❹ Line charts
Linked to the x- and y axis of contact map. Summarize no. PTMs per site.
❺ Sunburst charts
Compare sites. Circles show classification of (inner) and shared PTMs (outer, red).

Summarize and Overview along Primary Sequence, Account for 3D Neighborhood
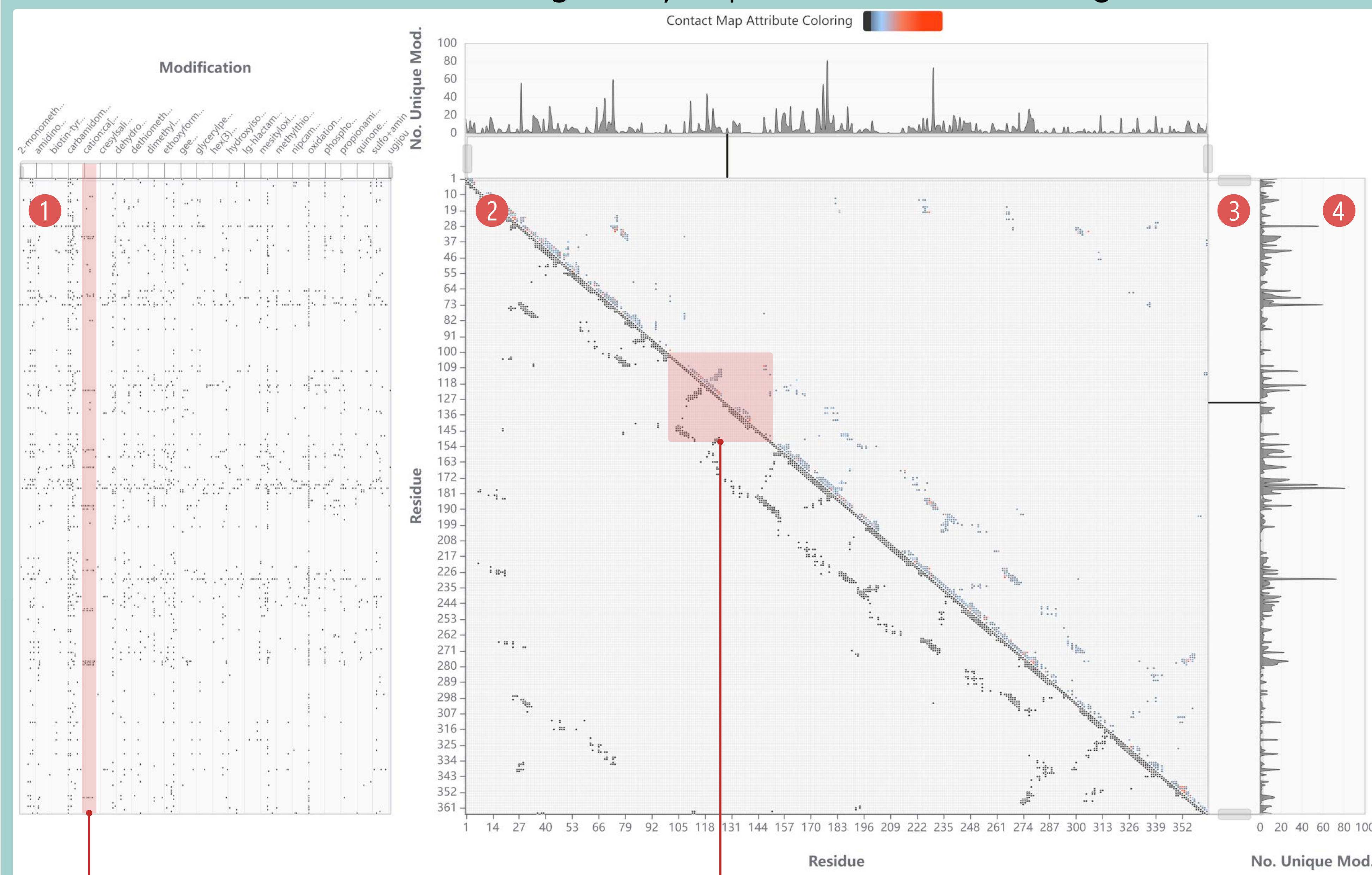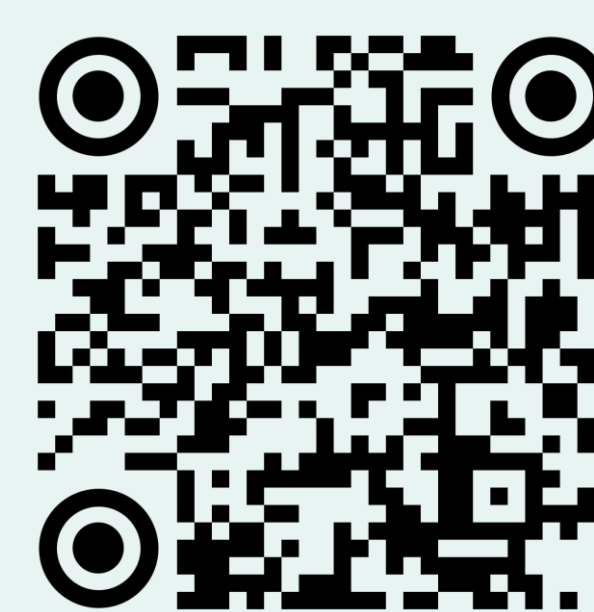

Identify Sites of Interest

Display Details on Demand

PTMs of residue 147, LYS

PTMs of residue 108, LYS
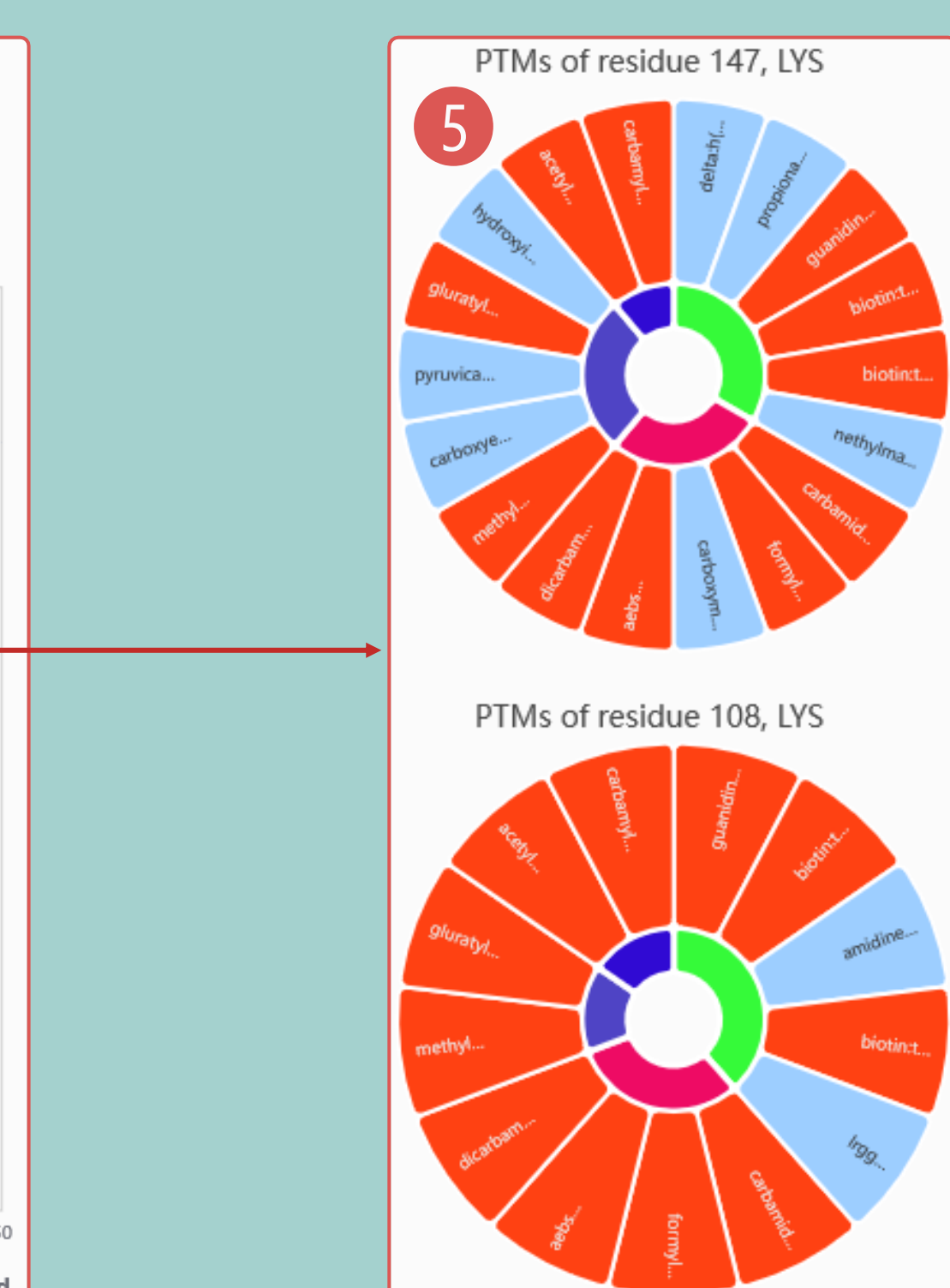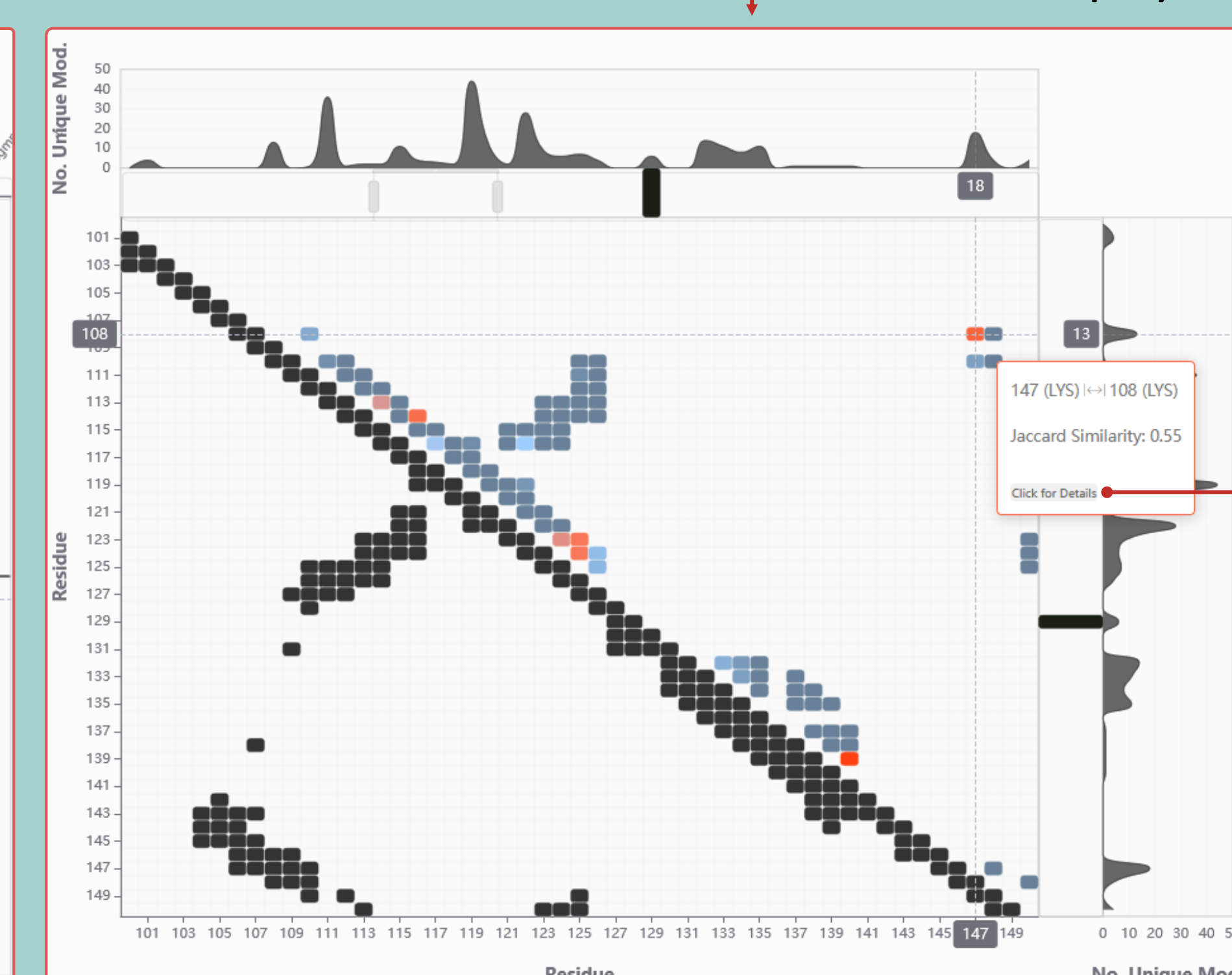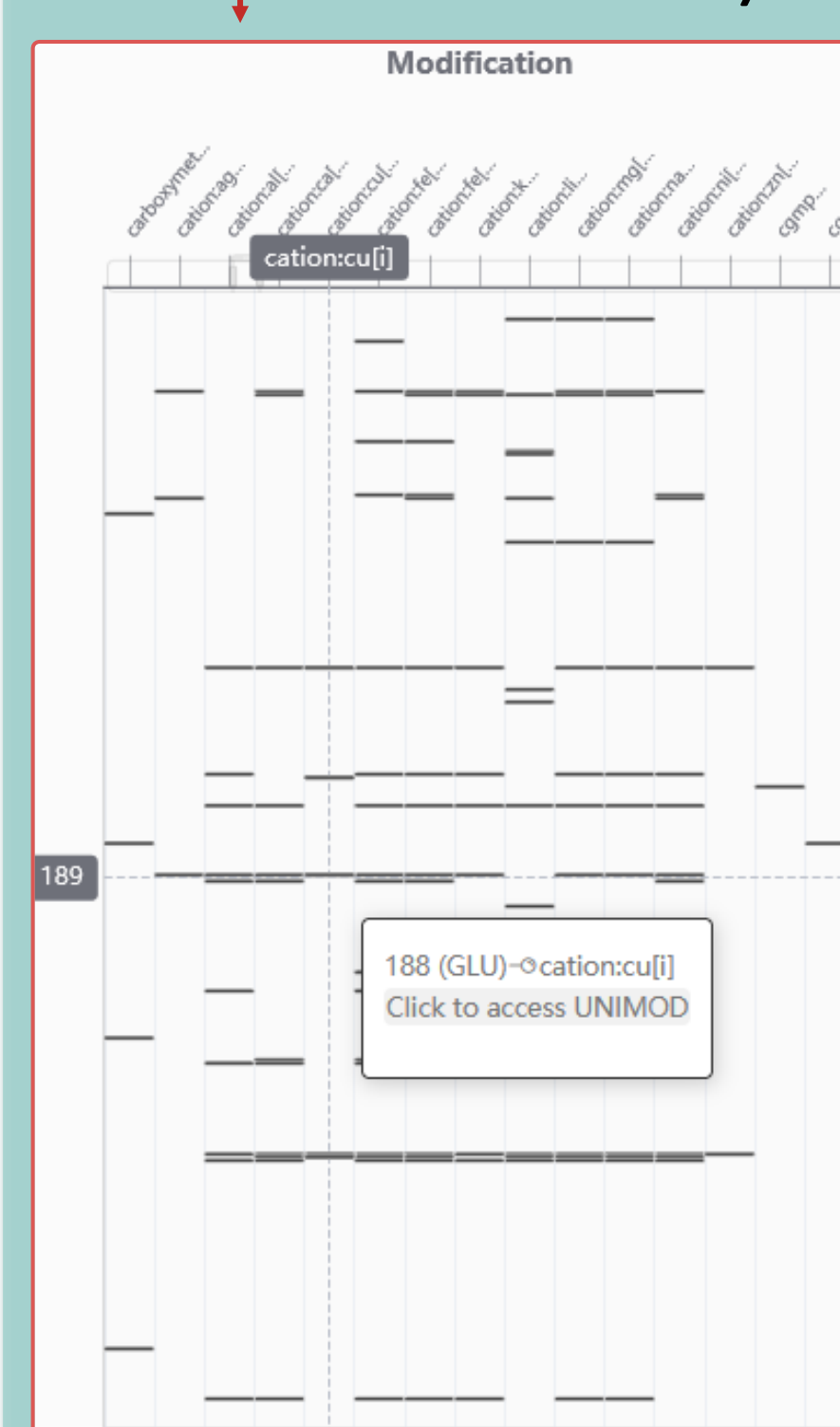
Web-App

Video

References

[1]GRAVES J. D., KREBS E. G.: Protein phosphorylation and signal transduction. Pharmacology & therapeutics 82, 2-3 (1999), 111–121.

[2]HOCHSTRASSER M.: Ubiquitin-dependent protein degradation. Annual review of genetics 30, 1 (1996), 405–439.

[3]DRAZIC A., MYKLEBUST L. M., et al.: The world of protein acetylation. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1864, 10 (2016), 1372–1401.

[4]BOUWMEESTER R., GABRIELS R., et al.: Deeplc can predict retention times for peptides that carry as-yet unseen modifications. Nature Methods 18, 11 (2021), 1363–1369.

[5]DEGROEVE S., GABRIELS R., et al.: ionbot: a novel, innovative and sensitive machine learning approach to lc-ms/ms peptide identification. bioRxiv (2021).

[6]GABRIELS R., MARTENS L., et al.: Updated ms2pip web server delivers fast and accurate ms2 peak intensity prediction for multiple fragmentation methods, instruments and labeling techniques. Nucleic acids research 47, W1 (2019), W295–W299.

[7]BLUDAU I., WILLEMS S., et al. : The structural context of posttranslational modifications at a proteome-wide scale. PLOS Biology 20, 5 (2022), 1–23.

[8]CRUZ E. R., NGUYEN H., et al.: Functional analysis tools for post-translational modification: a post-translational modification database for analysis of proteins and metabolic pathways. The Plant Journal 99, 5 (2019), 1003–1013.

[9]NUSRAT S., HARBIG T., et al.: Tasks, techniques, and tools for genomic data visualization. In Computer Graphics Forum (2019), vol. 38, Wiley Online Library, 781–805.

[10]BREHMER M., MUNZNER T.: A multi-level typology of abstract visualization tasks. IEEE transactions on visualization and computer graphics 19, 12 (2013), 2376–2385.

Funding and Acknowledgement

[†]University of Tübingen, Institute for Bioinformatics and Medical Informatics, Integrative Transcriptomics, Tübingen, Germany