

Interpreting MQ Phosphopeptide Results

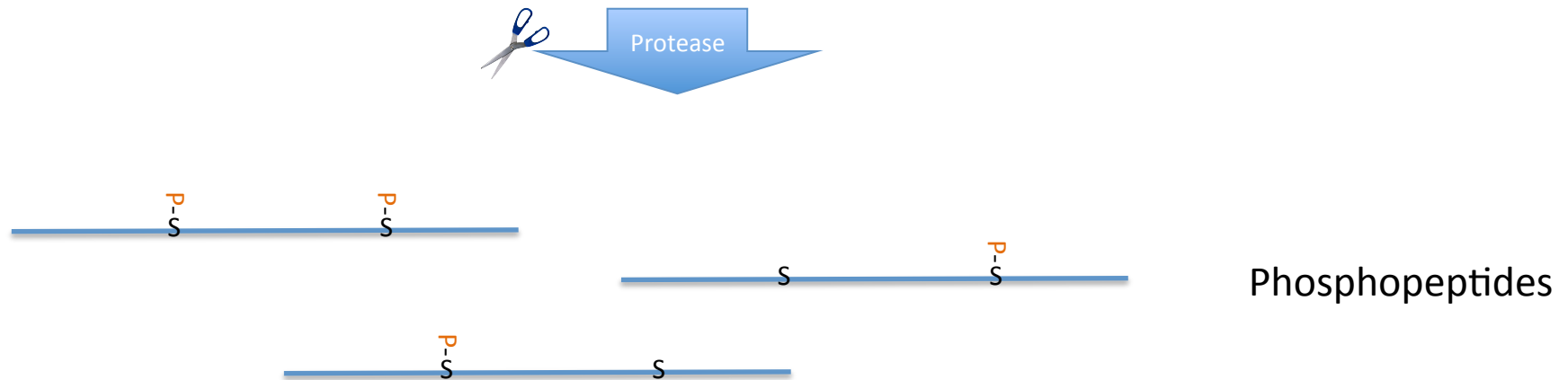
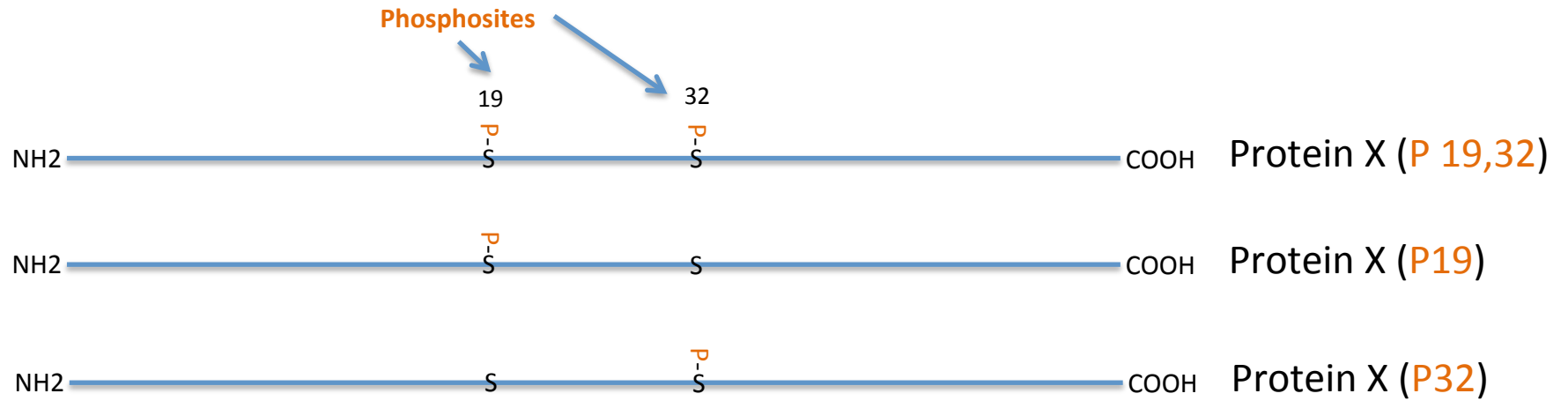
Format of results

- Usually we analyse phosphoproteomic data with MaxQuant
- You get the following result files either as single files or in a combined Excel file:
 1. XX-XXX¹_ProteinGroups
 2. XX-XXX¹_Phosphosites (STY)
- On request we can also provide other MQ output files

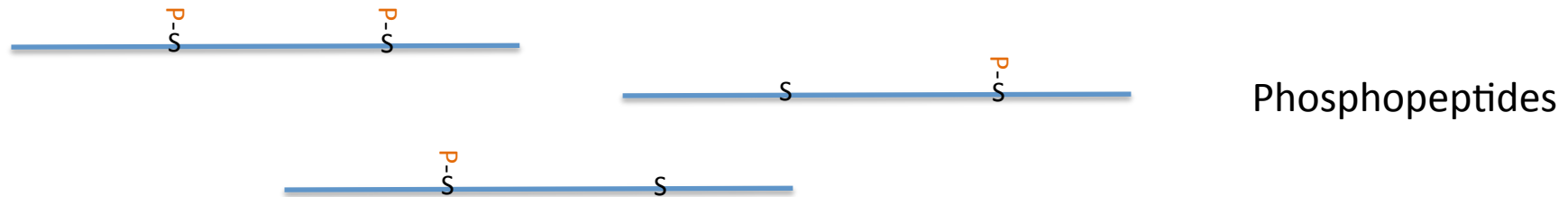
Phosphosites vs. Phosphopeptides

- Phosphosites are **NOT** Phosphopeptides!!!
- A phosphosite is a serine/threonine/tyrosine with a distinct location in the protein sequence, which is phosphorylated
- Quantitative information for each site is derived from ALL detected peptides, which are phosphorylated at this position (=combined/condensed information)

Example



Example (evidence table)

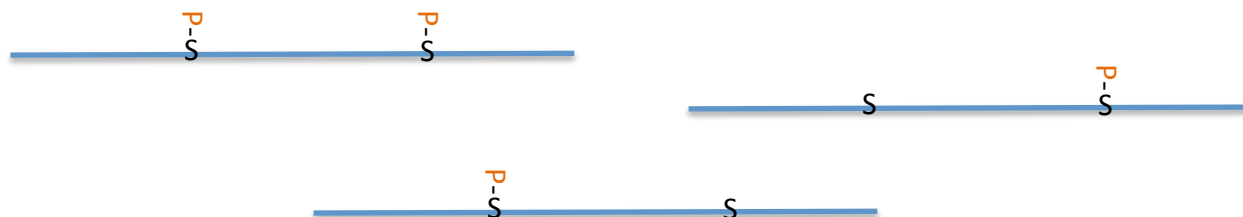


Quantitative
peptide
information

Peptide	Intensity	OR for SILAC based approaches →	Ratio H/L (normalized)
RILILS(P)GAAPS(P)EE	300000		2.3
RILILS(P)GAAPSEE	100000		1.2
RILILSGAAPS(P)EE	500000		1.7

This information can be found in the MQ output file:
evidence.txt

Example (phosphosite table)



Phosphopeptides

Quantitative
site
information

Position s within proteins	Intensit y	Intensit y____1	Intensit y____2	Intensit y____3	OR for SILAC based approac hes →	Ratio H/ L (norm)	Ratio H/ L (norm)____1	Ratio H/ L (norm)____2	Ratio H/ L (norm)____3
19	400000	100000	300000	0		1.2	1.2	2.3	NaN
32	800000	500000	300000	0		1.7	1.7	2.3	NaN

This information can be found in the MQ output file:

Phospho (STY)Sites.txt

Phospho (STY) table

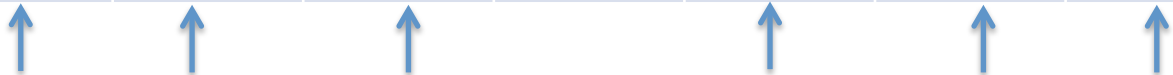
- This table (+/-the proteingroups table) is usually sufficient to make qualified assumptions about changes in the phosphorylation status of proteins between samples
- Key quantitative values are:
 - Occupancy (mainly applicable for SILAC data)
 - Intensity__1/2/3 OR Ratio H/L__1/2/3



__1, __2, __3 ? What does that mean?

___1, ___2, ___3 ???

Positi ons withi n prote ins	Amin o acid	Intensity untreated	Intensity untreated_ __1	Intensity untreated_ __2	Intensity untreated_ __3	Intensity treated	Intensity treated__ 1	Intensity treated__ 2	Intensity treated__ 3
19	S	400000	100000	300000	0	125000	50000	75000	0



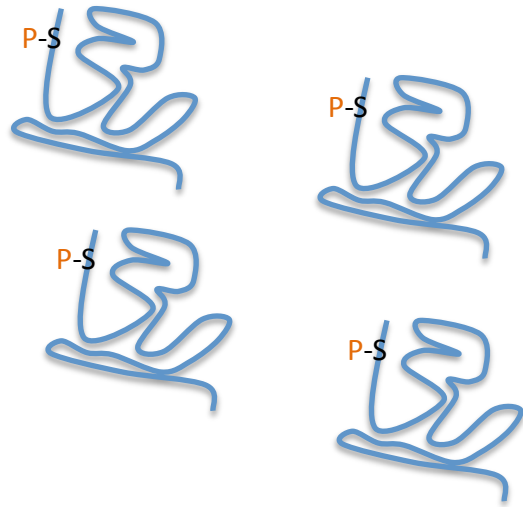
- There is quantitative information derived from single, double or triple(&higher) phosphorylated peptides
- Using this intensities is MUCH better then looking just at the summed site intensities, because in a biological system the same protein may have distinct functions when it is differentially phosphorylated.

Occupancy

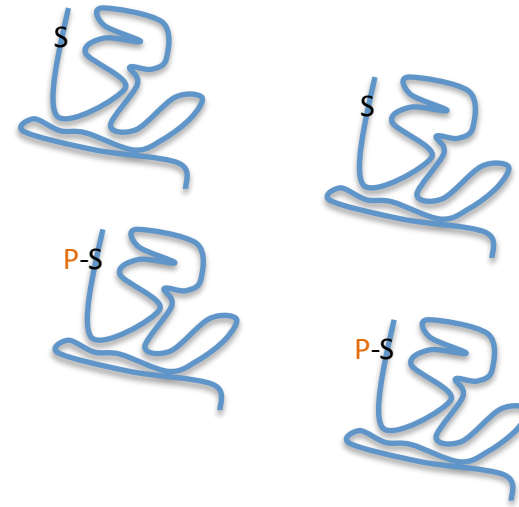
- Fraction of protein for which the site is modified (Value from 0 - 1 → 0 -100% modified)
 - This value is calculated from three separate ratios:
 - Ratio H/L proteingroup
 - Ratio H/L modified peptide (here: phosphorylation)
 - Ratio H/L unmodified counterpart peptide
 - Occupancy is calculated for every labeling state if all three ratios are present

Occupancy example

Untreated (light)



Treated (heavy)



Quantitative
site
information

Positions within proteins	Occupancy L	Occupancy H
19	1	0.5

No occupancy values...

- If I don't have occupancy values I must look at two types of data:

1. Phosphosite quant data (Phosphosites (STY).txt)

A. Intensities___1/2/3 → label free data

B. Ratio H/L (norm)___1/2/3 → SILAC data



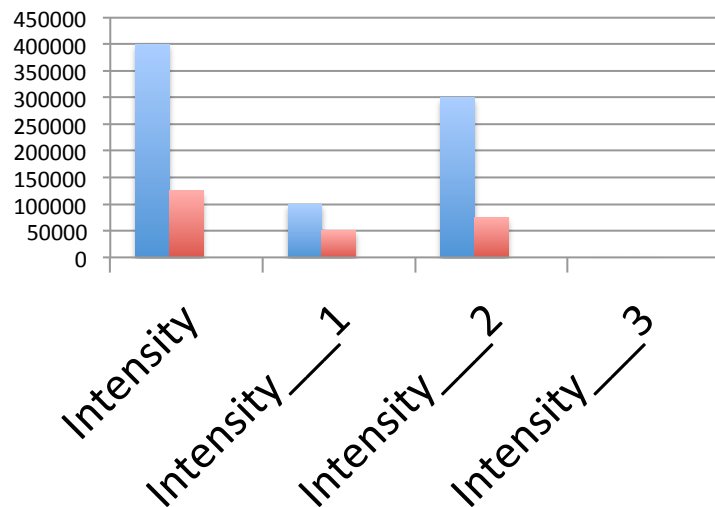
2. ProteinGroup quant data (ProteinGroups.txt)

A. LFQ Intensity → label-free data

B. Ratio H/L (norm) → SILAC data

Example

Position with in proteins	Amino acid	Intensity untreated	Intensity untreated__1	Intensity untreated__2	Intensity untreated__3	Intensity treated	Intensity treated__1	Intensity treated__2	Intensity treated__3
19	S	400000	100000	300000	0	125000	50000	75000	0

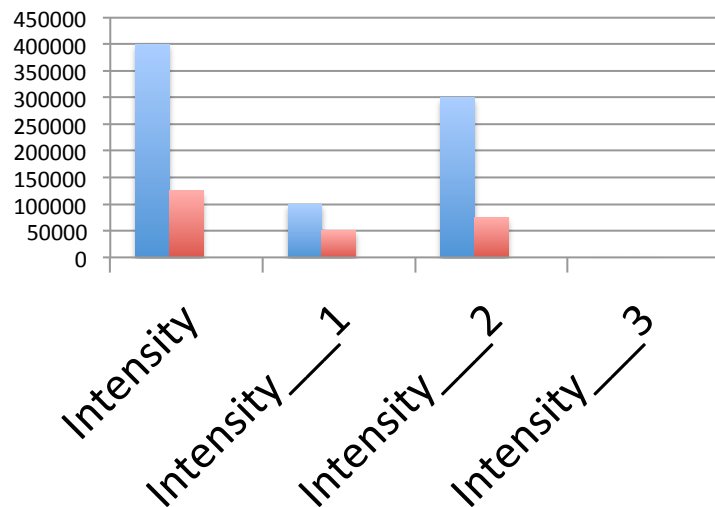


■ untreated
■ treated

➔ From this we could assume that serine at position 19 is more phosphorylated under the **untreated** condition

Example

Position with in proteins	Amino acid	Intensity untreated	Intensity untreated__1	Intensity untreated__2	Intensity untreated__3	Intensity treated	Intensity treated__1	Intensity treated__2	Intensity treated__3
19	S	400000	100000	300000	0	125000	50000	75000	0

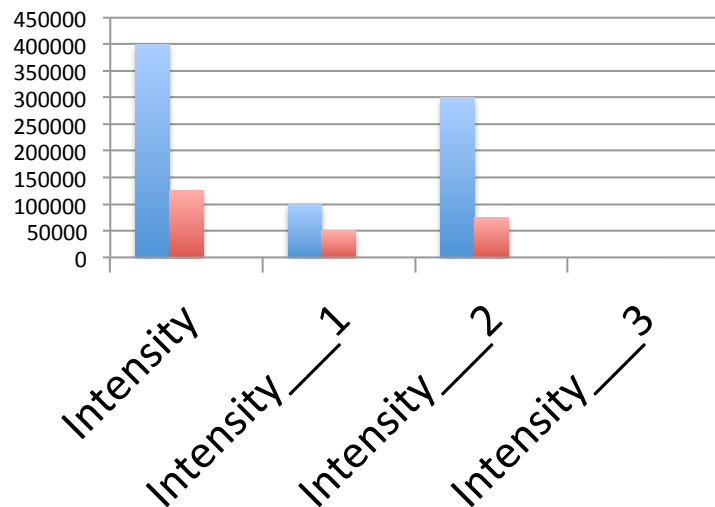


■ untreated
■ treated

➔ But what about the protein level? Maybe the whole protein is less abundant in the **treated** sample?

Example

Position with in proteins	Amino acid	Intensity untreated	Intensity untreated__1	Intensity untreated__2	Intensity untreated__3	Intensity treated	Intensity treated__1	Intensity treated__2	Intensity treated__3
19	S	400000	100000	300000	0	125000	50000	75000	0



■ untreated
■ treated

So you need to also look at the quant protein data (proteingroups.txt) before making any assumptions!!!