



Carbene Protein Footprinting

BMSS Biomacromolecular SIG Training

27 July 2020

Neil Oldham

University of Nottingham

Overview

- Introduction and Background
- Protein and Probe
- Photochemical activation
- Proteolysis and MS analysis
- Data processing

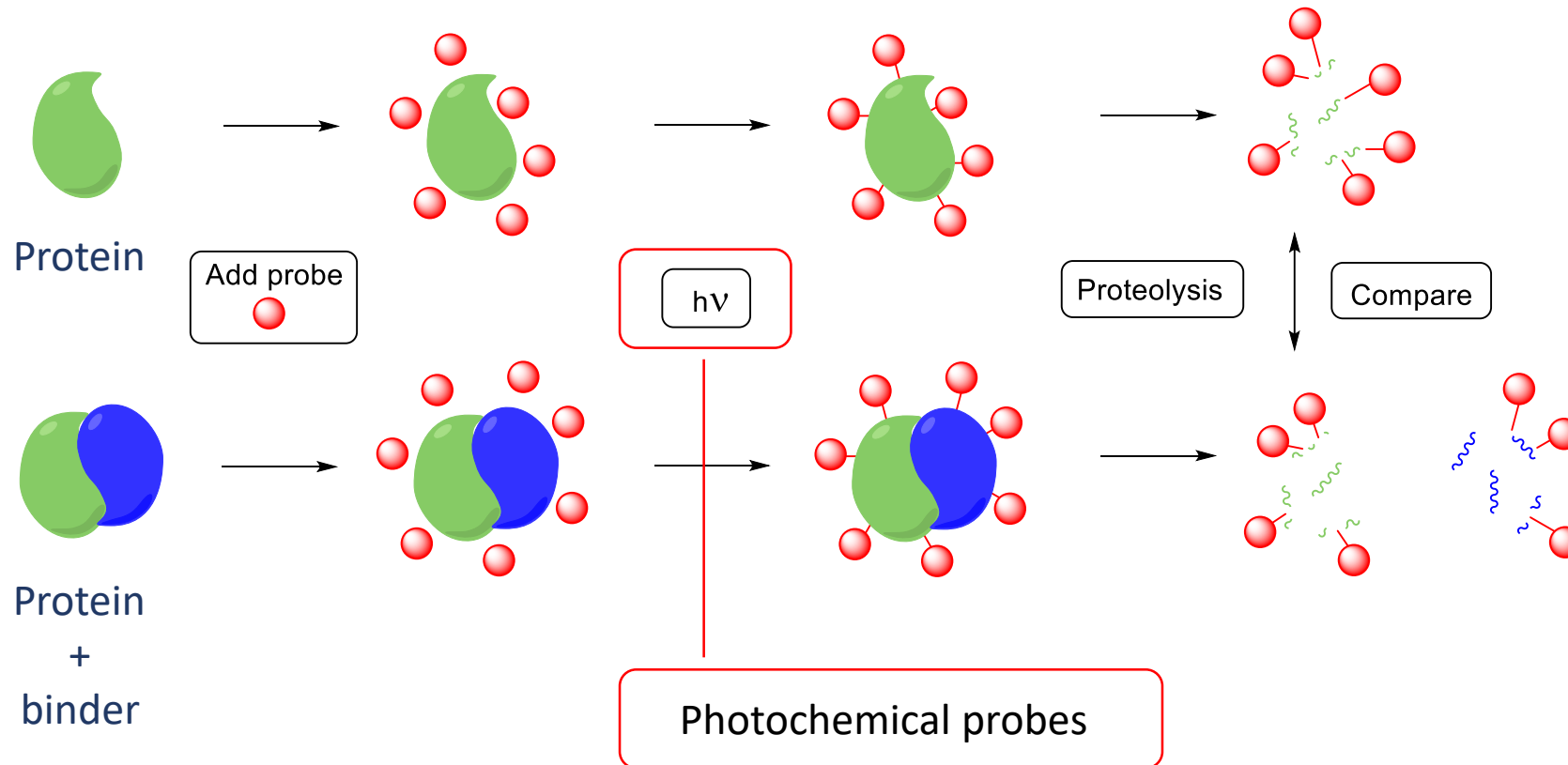
Introduction and Background

Concept

- Covalent labelling of a protein can be used to reveal solvent accessibility of the protein's surface
- Quantitative changes in labelling associated with masking by a binding partner allow binding sites on the protein to be identified

Introduction and Background

Concept



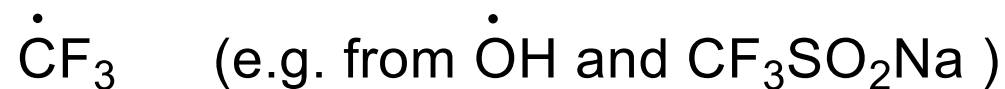
Introduction and Background

Probes

Radicals

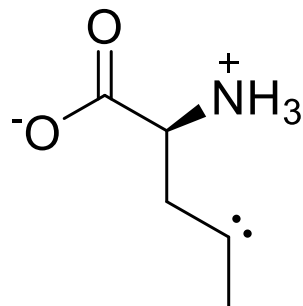


Gross et al. JACS, 18724 (2012)

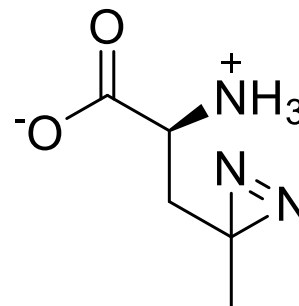


Gross et al. Angew. Chem., 14007 (2017)

Carbenes



from

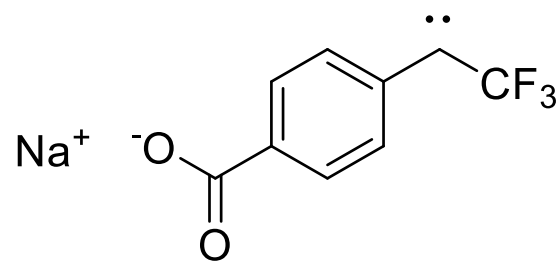


and $h\nu$ @352 nm

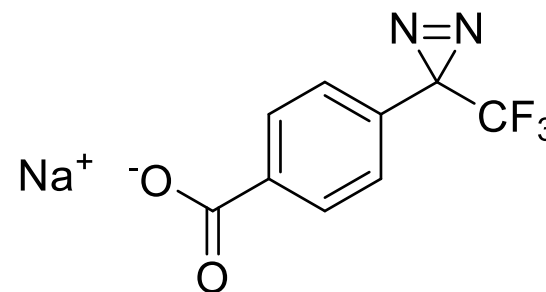
Schriemer et al. Anal. Chem., 2913 (2011)

Introduction and Background

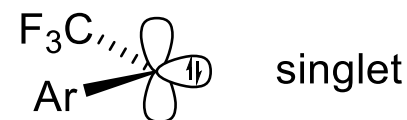
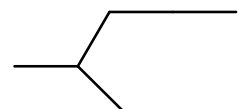
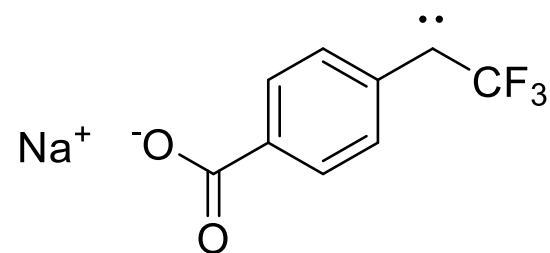
Carbene Probe



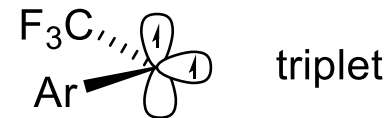
from



and $h\nu$ @352 nm



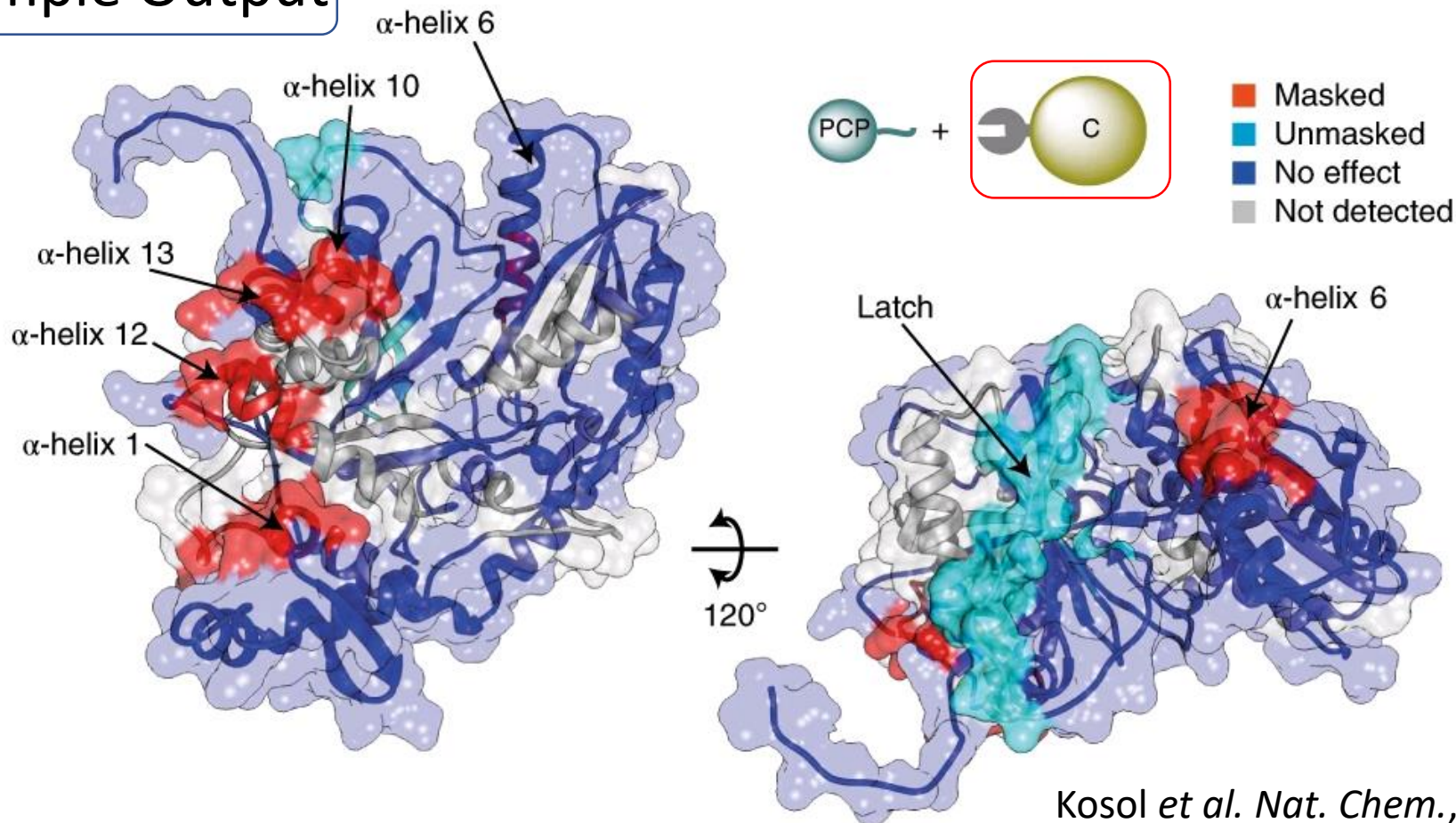
singlet



triplet

Introduction and Background

Example Output



Kosol *et al.* *Nat. Chem.*, 913 (2019)

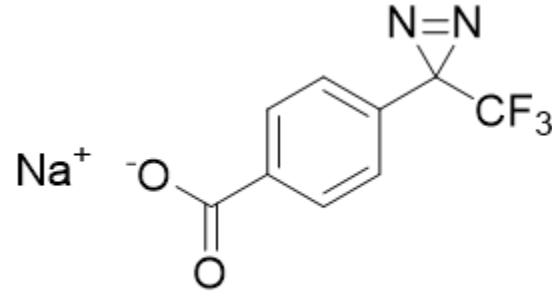
Protein and Probe

Protein

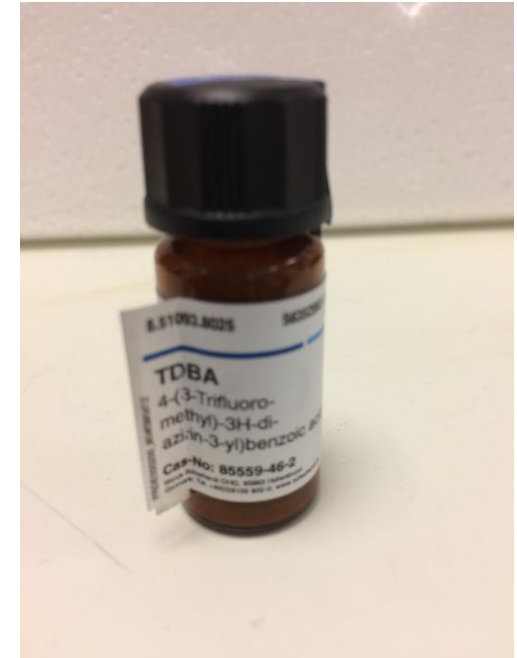
- Typical concentration 10-50 μM , typical volumes 20-100 μL
- Suitable buffers include Tris-HCl, PBS etc.
- No special considerations – just keep conditions constant across experiments

Protein and Probe

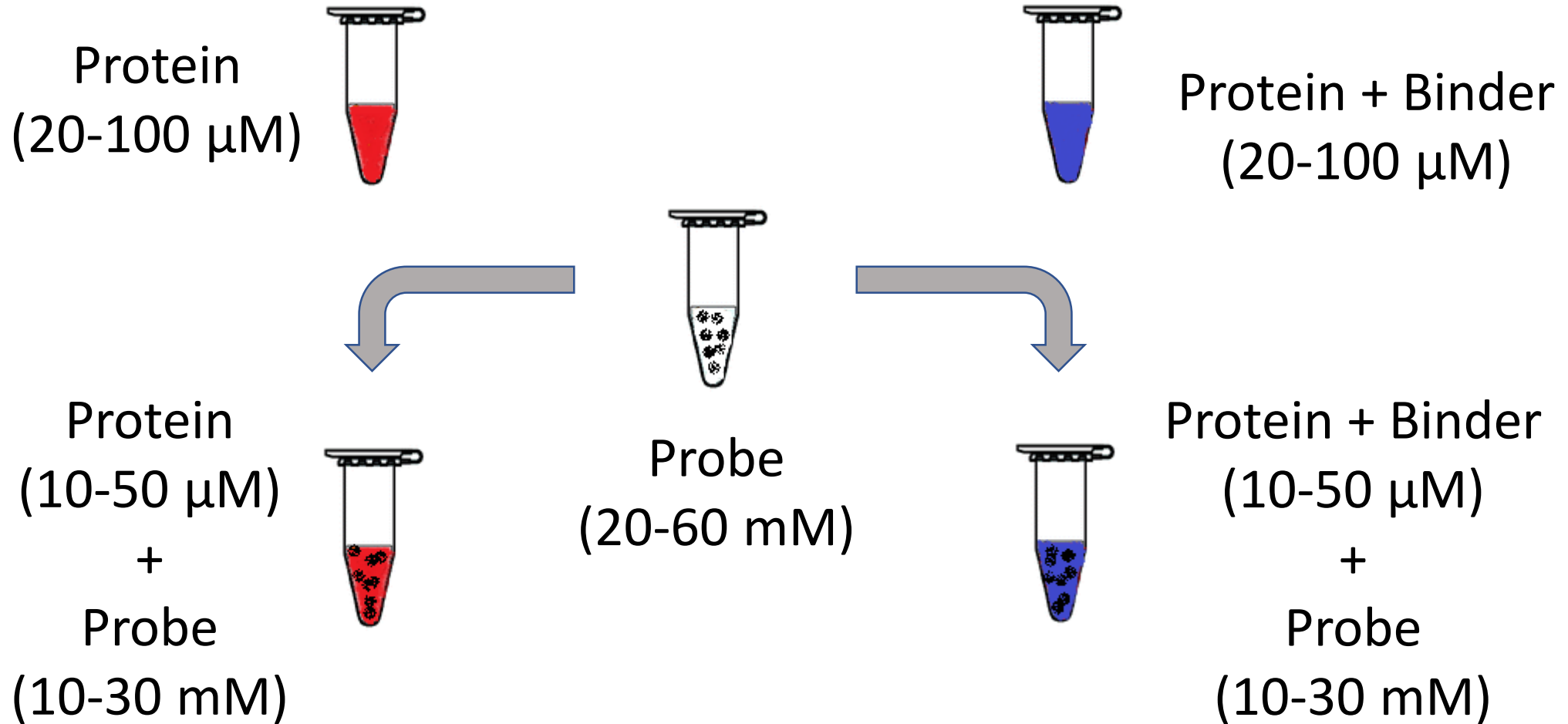
Probe



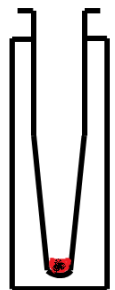
- Free acid (TDBA) is commercially available (Merck (Novabiochem) 25 mg - £ 180)
- To make sodium salt – add 1.1 mol. excess. of TDBA to NaOH solution of desired conc. (e.g. 40 mM) and centrifuge to remove undissolved free acid. Solution can be freeze dried, but is stable at 4 °C



Protein and Probe



Photochemical Activation

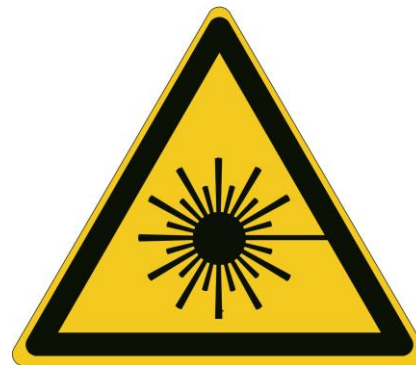


5 μ L

Flash freeze

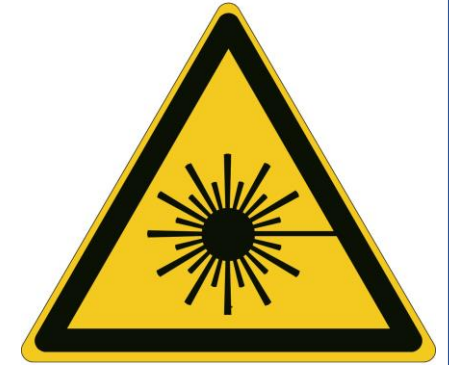
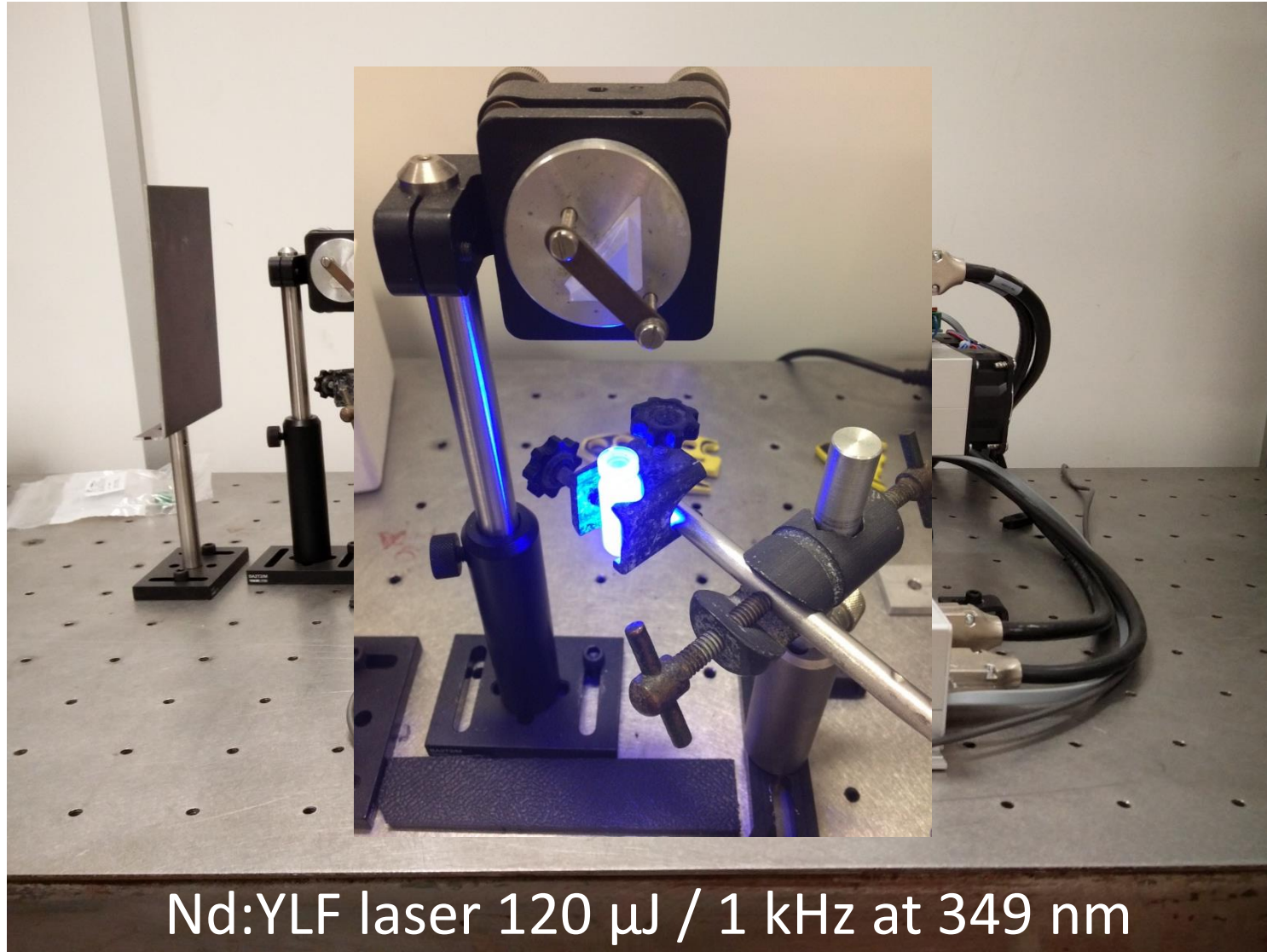


Liquid nitrogen



Laser irradiation

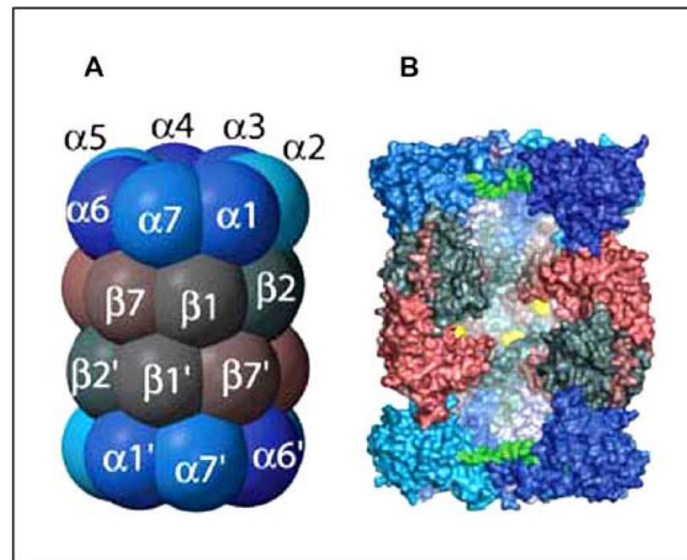
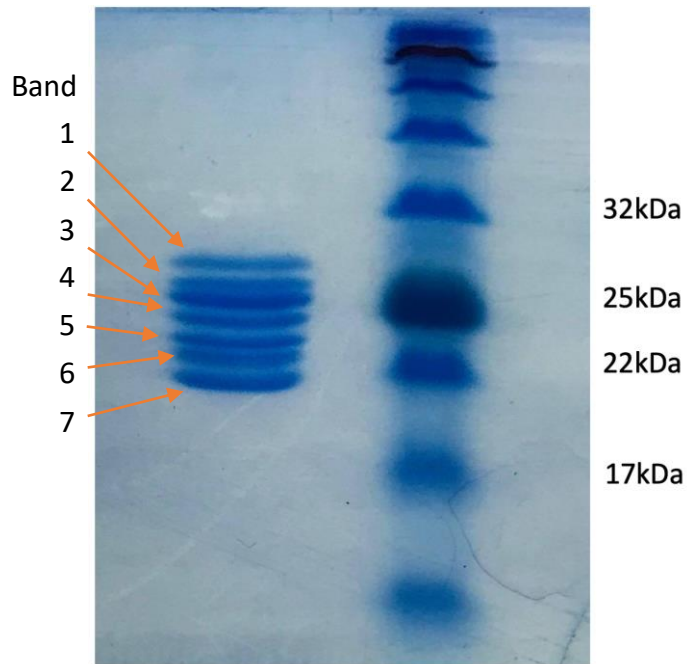
Photochemical Activation



For normal operation the beam should be fully enclosed

Proteolysis and MS analysis

- Covalent labelling is irreversible
- Protein sample can be separated by SDS-PAGE



20S proteasome

Proteolysis and MS analysis

- Protein digestion can be in gel or in solution
- A range of enzymes can be used:

Trypsin

AspN

GluC

Pepsin

Chymotrypsin

Elastase



Run digests (in silico and actual) on the protein(s) before labelling to guide enzyme choice

Proteolysis and MS analysis

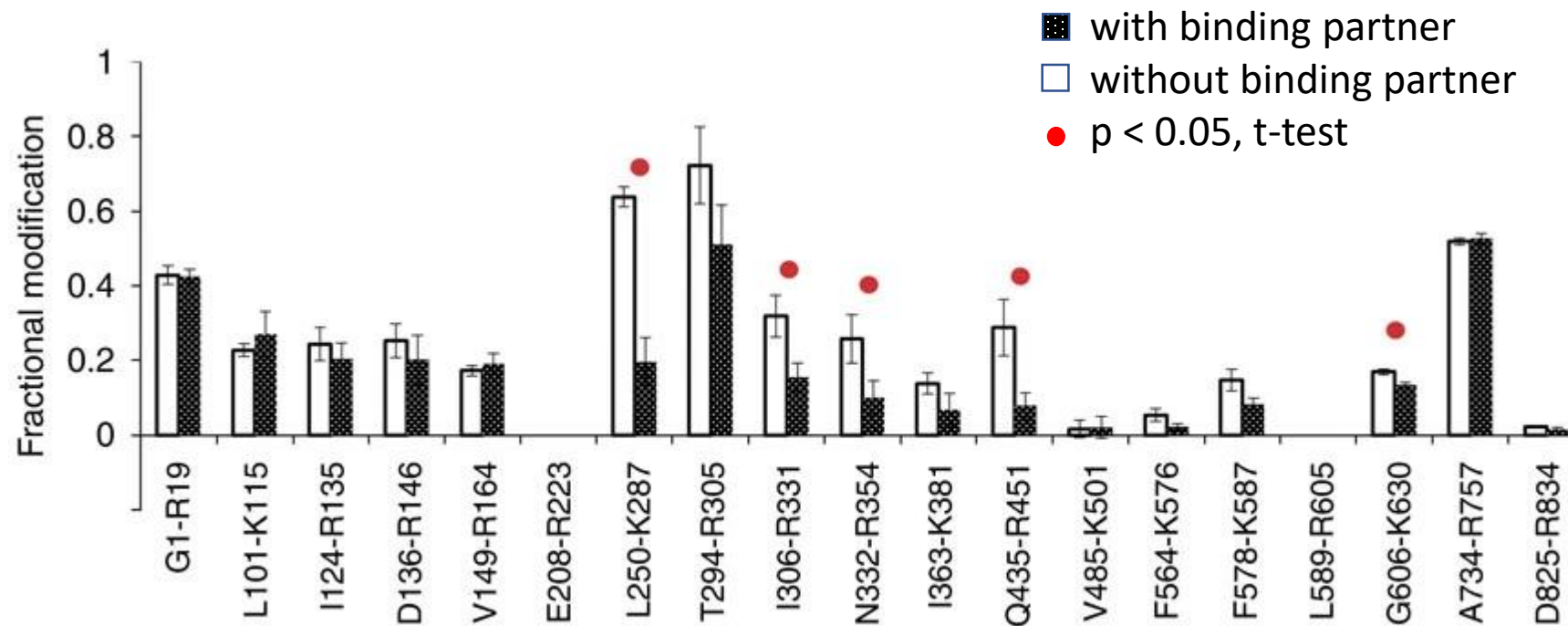
- Nano or cap LC-ESI-MS preferred for digest analysis
- FT-MS or TOF-MS (HR and accurate mass) beneficial
- MS/MS (CID or ECD/ETD) confirms ID of unlabelled peptides; and can be used to map labelling at the sub-peptide level
- No special requirements for MS analysis beyond those typically adopted for peptide analysis/bottom-up proteomics

Data Processing

- Output of experiment – ~10 LC-MS files containing data for modified and unmodified peptides (usually in 2 *groups*)
- Data can be analysed *manually* -
 - generate EICs for each labelled and unlabelled peptide
 - integrate peaks - PA
 - determine fractional modification $f_m = \frac{PA_{mod}}{PA_{mod} + PA_{unmod}}$
 - Plot f_m for each peptide and compare between *groups*

Data Processing

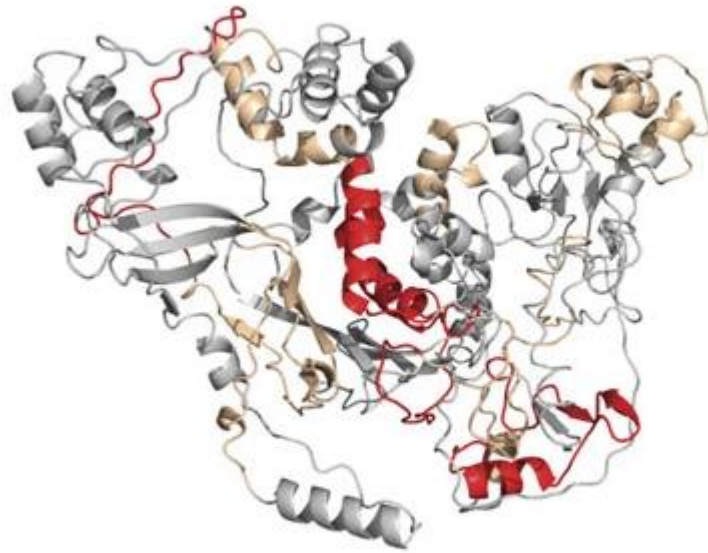
- Differences between *groups* report areas of masking and unmasking



Manzi *et al. Nat. Comm.*, 13288 (2016)

Data Processing

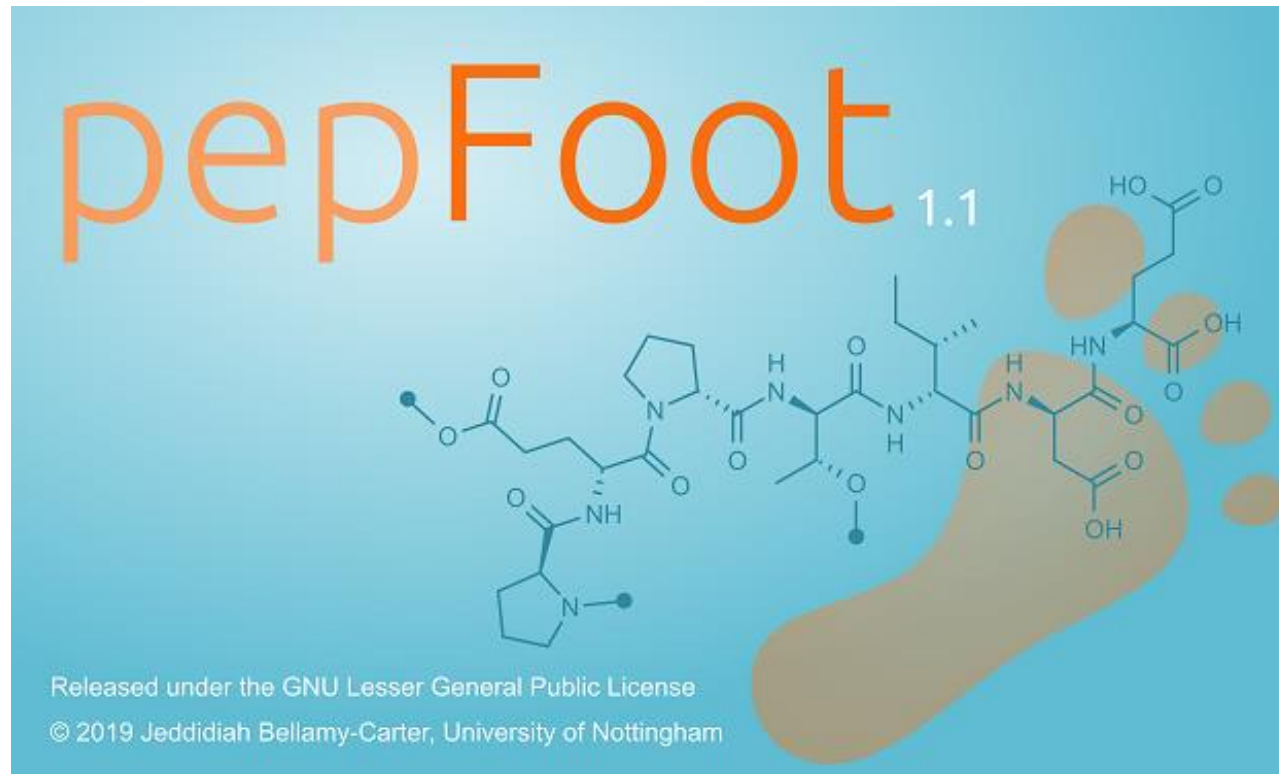
- Differences can then be plotted on protein structures (X-ray, NMR, cryoEM or homology models)



■ Significant masking by binding partner

Data Processing

Semi-automated data analysis



- Batch processing
- Data visualisation

github.com/jbellamycarter/pepfoot

Bellamy-Carter and Oldham. *J. Proteome Res.*, 2925 (2019)

Data Processing

File Edit Help

Modifications Enzymes

Peptide Level Analysis NGL Viewer

Peptides

- NERDTAR
- RSNGDGVGGSISYEYEGFIVGAYGAADR
- SNGDGVGGSISYEYEGFIVGAYGAADR
- TNLQEAQPLGNGK
- TNLQEAQPLGNGK
- KAEQWATGLK
- AEQWATGLK
- AEQWATGLKYDANNIYLAANYGETR
- YDANNIYLAANYGETR**
- YDANNIYLAANYGETR
- YDANNIYLAANYGETR

Predicted 100 %
Experimental 95 %

Charge State 221-236 2

Unmodified m/z 924.43
Modified m/z 1025.44

Unmodified m/z error 3.06 mmu
Modified m/z error -6.68 mmu

Unmodified Area 2258413651
Modified Area 226751858
Modification 0.09

Remove Assignment

Zoom Extract Integrate

Data Files

- LM_AD_OmpF_1_040816.mz5
- LM_AD_OmpF_2_040816.mz5
- LM_AD_OmpF_3_040816.mz5

Sequence

```
GAEIYNKDGKVDLYGKAVGLHYFSKGNSENSYGGNDMTYA
RLGFKGETQINSDLTGYGOWEYFQGNNEGADAQTGNKTRL
AFAGLKYADVGSFDYGRNYGVYDALGYTDLPEFGDGTAYS
DDFFVGRVGGVATYRNSNFFGLVDGLNFVQYLKNERDTAR
RSNGDGVGGSISYEYEGFIVGAYGAADRNLQEAQPLGNGK
KAEQWATGLKYDANNIYLAANYGETRNPITNKFNTSFGA
NKTQDVLVAQYQDFGLRPSIAYTKSAKDVEGIDVDLVN
YFEVGATYFKNMSTYVDYIINQIDSDNKLGVGSDDTVAVG
IVYQF
```

Modifications

- Acetyl
- Acrylamide adduct
- Amidation
- Aryldiazirine-TDBA
- Biotinylation
- Carbamidomethyl
- Carbamylation
- Aryldiazirine-TDBA

Digestion

Trypsin

Peptide Length 5 - 40

Peptide Charge 1 - 4

Missed Cleavages 1

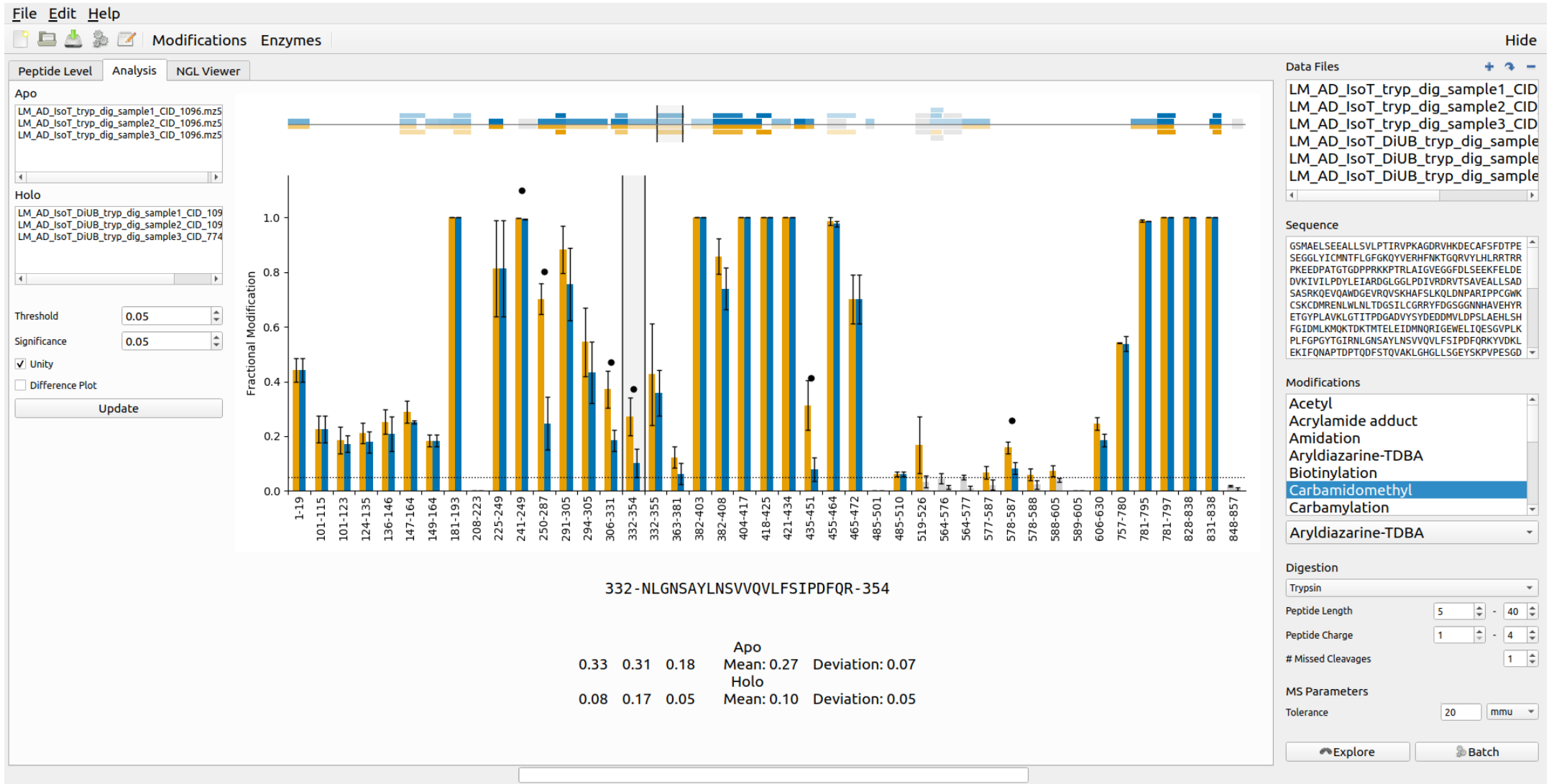
MS Parameters

Tolerance 20 mmu

Explore Batch

Chromatogram extracted!

Data Processing



Data Processing

The screenshot displays the NGL Viewer software interface. The central window shows a 3D ribbon model of a protein structure, colored by residue type (e.g., alpha-helices in orange, beta-sheets in green, loops in white). A specific residue is labeled "GLN 574".

The interface includes a menu bar (File, Edit, Help) and a toolbar with icons for file operations and analysis. The main window is divided into several panels:

- Peptide Level / Analysis / NGL Viewer:** The active view.
- PDB:** Shows the loaded file "USP5 (demo).pdb" and a "Load PDB" button.
- Colour Scaling:** Includes a checked "Continuous" option.
- Controls:** Lists various interaction controls such as "Cartoon Toggle", "Hetero Object Toggle", "Recenter view", "Spin Toggle", "Rock Toggle", and movement instructions (Left Move: Rotate around center, Right Move: Move in plane, Shift + Left or Middle: Zoom, Shift + Middle: Clip Plane).
- Data Files:** Lists several data files, including "LM_AD_IsoT_tryp_dig_sample1_CID", "LM_AD_IsoT_tryp_dig_sample2_CID", "LM_AD_IsoT_tryp_dig_sample3_CID", "LM_AD_IsoT_DiUB_tryp_dig_sample", "LM_AD_IsoT_DiUB_tryp_dig_sample", and "LM_AD_IsoT_DiUB_tryp_dig_sample".
- Sequence:** Displays the protein sequence:

```
GSMALSEEALLSVLPTIRVPKAGDRVHKDECAFSFDTPE  
SEGGLYICMNTFLGFGKQYVERHFNKTGRVYLHLRRTTR  
PKEEDPATGTGDPKPKPTRLAIGVEGFDLSEKFDLDE  
DVKIVILPDYLEIARDLGLGGLPDIVRDRVTSAVEALLSAD  
SASRKOEVQAWDGEVROVSKHAFSLKQLDNPARIPPCGWK  
CSKDMRENLLNLTDGSLCGRRYFDGSGGNHVAHEHYR  
ETGYPLAVKLGITIPDGADVYSYDEDDMVLDPSLAEHL  
FGIDMLKMKTDKTMTELEIDMNQRIGEWELIQESGVPLK  
PLFGPGYTGIRNLGNSAYLNSVQVLFIPDFQRKYVDKL  
EKIFQNAPTDPTQDFSTQVAKLGHGLSGEYSKVPVPSGD
```
- Modifications:** A list of modification types including Acetyl, Acrylamide adduct, Amidation, Aryldiazirine-TDBA, Biotinylation, Carbamidomethyl (highlighted), and Carbamylation. Below this is a dropdown menu for "Aryldiazirine-TDBA".
- Digestion:** Shows "Trypsin" as the selected enzyme. Parameters include Peptide Length (5 to 40), Peptide Charge (1 to 4), and # Missed Cleavages (1).
- MS Parameters:** Shows a Tolerance of 20 mmu.
- Buttons:** "Explore" and "Batch" buttons are located at the bottom right.

Summary

- Introduction and Background
- Protein and Probe
- Photochemical activation
- Proteolysis and MS analysis
- Data processing