



**HPC**  
WALES • CYMRU

Powering Success  
Pweru Llwyddiant

# HPC User Presentation

Bela Paizs, Bangor University

**HPC Wales User Group Meeting**

**Wednesday 22<sup>nd</sup> October 2014**

Ewrop & Chymru:  
Buddsoddi yn eich dyfodol  
Cronfa Datblygu Rhanbarthol Ewrop

Europe & Wales:  
Investing in your future  
European Regional Development Fund

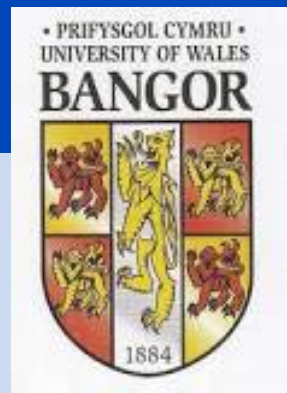


Llywodraeth Cymru  
Welsh Government

# Computational Proteomics

Bela Paizs

School of Chemistry, Bangor University, Bangor, LL57 2UW, UK



# Overview

- Proteomics
- Mass spectrometry in proteomics
- HPC & proteomics

# Why is Research on Proteins Important?

Determination of the genome (hereditary information encoded in DNA) is only the first step towards understanding biology.

# Why is Research on Proteins Important?

Determination of the genome (hereditary information encoded in DNA) is only the first step towards understanding biology.

The same genome



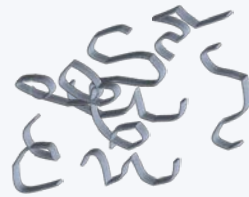
different expressed proteins

# Typical Protein Identification Workflow



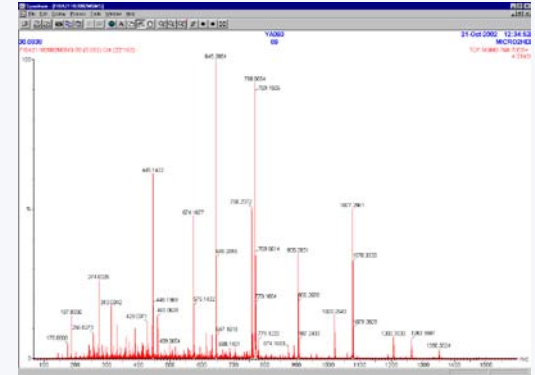
Protein  
mixture

Digestion  
→

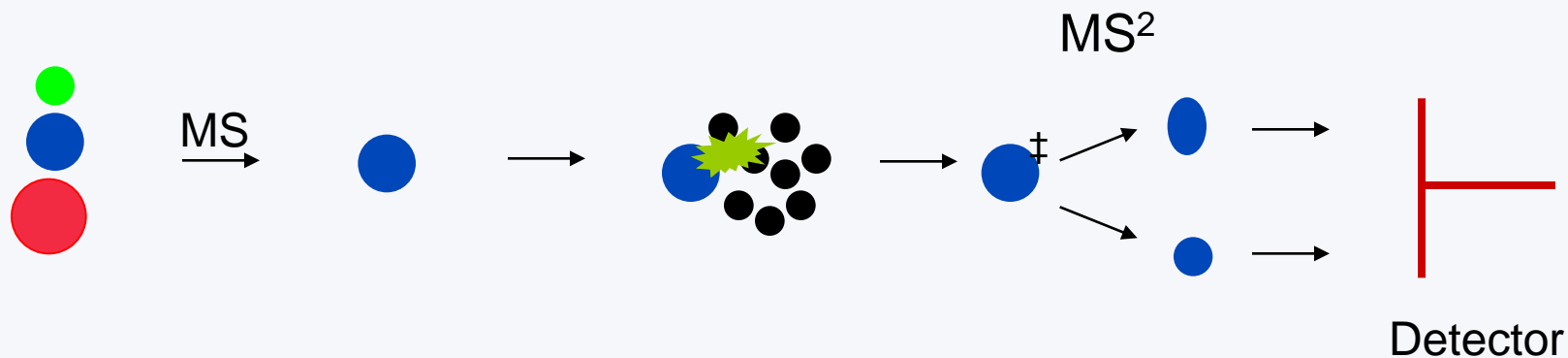
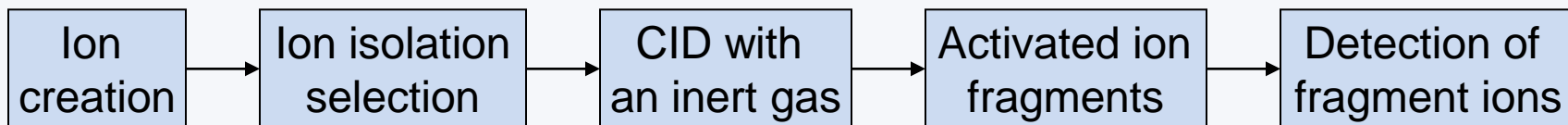


Peptide  
mixture

Mass  
spectrometry  
→



# Collision Induced Dissociation (CID)



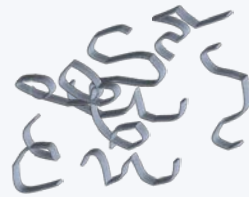
Two or more stages of MS:  
tandem mass spectrometry

# Typical Protein Identification Workflow



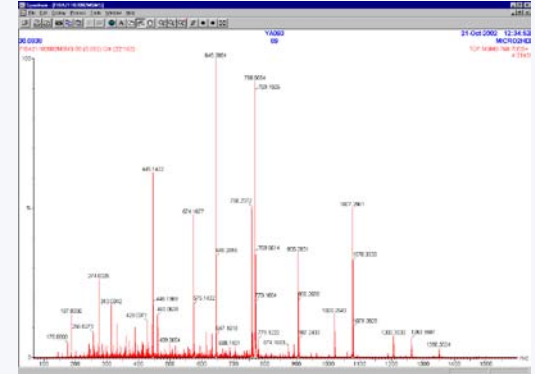
Protein mixture

Digestion



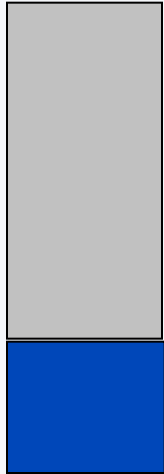
Peptide mixture

Mass spectrometry





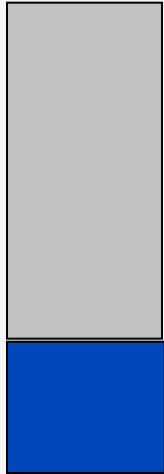
# Current Sequencing Software Performs Poorly



~50 % unassigned  
Related biological information is lost

~50 % assigned, but with some error rate

# Current Sequencing Software Performs Poorly



~50 % unassigned  
Related biological information is lost

~50 % assigned, but with some error rate

## NewsFocus

### ‘Proteomics Ponders Prime Time’

„... most errors in MS-based analyses arise not because the technology cannot spot the proteins researchers are looking for but because software programs often misidentify them.“

Science, 26 September 2008, page 1758.

# Typical Protein Identification Workflow



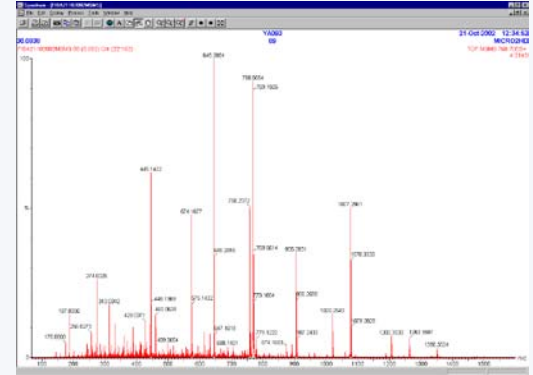
Protein mixture

Digestion



Peptide mixture

Mass spectrometry



# Need for Deeper Understanding of Peptide Fragmentation

- Current software uses *oversimplified* fragmentational models
- We need to develop a new generation of score functions that utilizes advanced peptide fragmentation chemistry
- Fragmentation chemistry is computed at HPC Wales