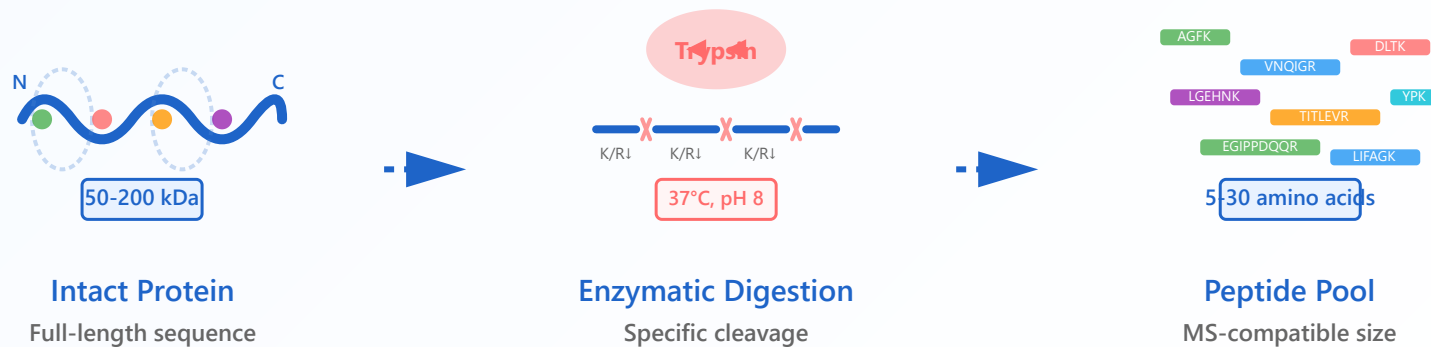


## Bottom-up Proteomics



### Protein Digestion

- Enzymatic cleavage into peptides
- 5-30 amino acid peptides
- Most common workflow

## 1. Protein Digestion: Breaking Down the Target

### 1.1 Overview and Purpose

Protein digestion is the critical first step in bottom-up proteomics where intact proteins

are enzymatically cleaved into smaller peptide fragments. This process transforms complex, large proteins into manageable peptides that are compatible with mass spectrometry analysis.

The goal is to generate peptides with optimal characteristics for mass spectrometry: typically 5-30 amino acids in length, containing appropriate charge states, and with predictable fragmentation patterns.

**Why Digestion is Necessary:** Intact proteins are too large and complex for efficient ionization and fragmentation in most mass spectrometers. Peptides provide better sensitivity, more consistent ionization, and interpretable MS/MS spectra.

## 1.2 Digestion Process

Enzymatic Digestion Workflow



K↓ R↓

### Trypsin Specificity

- Cleaves after K and R residues
- Predictable peptide generation
- Optimal MS-friendly peptides

### Peptide Separation

- Reverse-phase liquid chromatography
- Gradient elution
- Online LC-MS coupling



### Data Complexity

- Thousands of peptides
- Multiple charge states
- Requires computational analysis



## Example: Complete Digestion Protocol

Step-by-Step Trypsin Digestion Protocol:

## 2. Trypsin Specificity: Detailed Analysis

### 2.1 Real Protein Example

1

Denaturation

Add 8M urea in 100mM Tris-HCl (pH 8.5)  
Incubate: Room temp, 30 min  
Result: Unfolded protein structure

3

Alkylation

Add 55mM iodoacetamide (IAA)  
Incubate: Dark, room temp, 20 min  
Result: Cys residues carbamidomethylated

5

Digestion

Add trypsin (1:50 enzyme:protein ratio)  
Incubate: 37°C, overnight (12-16 hours)  
Optional: Add more trypsin at 6 hours

2

Reduction

Add 10mM DTT or 5mM TCEP  
Incubate: 56°C, 30 min (DTT) or RT, 20 min (TCEP)  
Result: Disulfide bonds broken (5-5 → 2 SH)

4

Dilution

Dilute with 50mM NH<sub>4</sub>HCO<sub>3</sub> to <2M urea  
Reason: High urea inhibits trypsin activity  
Final buffer: pH 8.0-8.5 is optimal

6

Quenching & Cleanup

Add formic acid to pH < 3 (stops enzyme)  
Desalt using C18 zip-tip or SPE cartridge  
Ready for LC-MS analysis

Quality Control Checkpoints:

Before Digestion:

✓ Protein concentration: 0.1-1 mg/ml (Bradford or BCA assay)  
✓ pH check: Should be 8.0-8.5 for optimal trypsin activity

Troubleshooting:

• Incomplete digestion → Extend time, add more enzyme, check pH  
• Over-digestion (peptides too short) → Reduce enzyme amount or time  
• Autolysis peaks (trypsin self-digestion) → Use sequencing grade trypsin

After Digestion:

✓ SDS-PAGE: Protein band should disappear  
✓ Peptide concentration: Measure by A280 or BCA

Let's examine how trypsin cleaves a real protein:  
Human Serum Albumin (HSA), one of the most  
abundant blood proteins.

## Trypsin Digestion of Human Serum Albumin Fragment

HSA Sequence Fragment (residues 125-145):

VHPEYAVSVLLRLAKEYEATLEECCAK

Trypsin Cleavage Sites (K and R):

VHPEYAVSVLLR↓LAK↓EYEATLEECCAK↓

Resulting Tryptic Peptides:

Peptide 1:

VHPEYAVSVLLR

Length: 12 aa | Mass: 1395.73 Da

Peptide 2:

LAK

Length: 3 aa | Mass: 345.43 Da

Peptide 3:

EYEATLEECCAK

Length: 12 aa | Mass: 1415.60 Da

Analysis of Cleavage Pattern:

Peptide 1 (VHPEYAVSVLLR):

• Good length for MS (12 aa) ✓  
• Contains basic residue at C-terminus (R) for ionization ✓  
• Mass in optimal range (800-2500 Da) ✓

Peptide 2 (LAK):

• Too short (3 aa) - may be lost X  
• Low mass (345 Da) - below MS detection X  
• Often excluded from analysis

Note: Peptide 2 illustrates a common issue - trypsin creates some peptides that are too short for effective MS analysis.

## 2.2 Peptide Coverage Map

### Proteome Coverage with Trypsin

Full Protein (500 amino acids)

Tryptic Peptides (Ideal Digestion):

Peptide 1

Peptide 2

Peptide 3

Peptide 4

Gap

Peptide 5

Peptide 6

Peptide 7

Typical Coverage Statistics:

Sequence Coverage

70-90%

of protein sequence  
identified by peptides

Peptides per Protein

15-40

observable peptides  
per average protein

Missed Regions

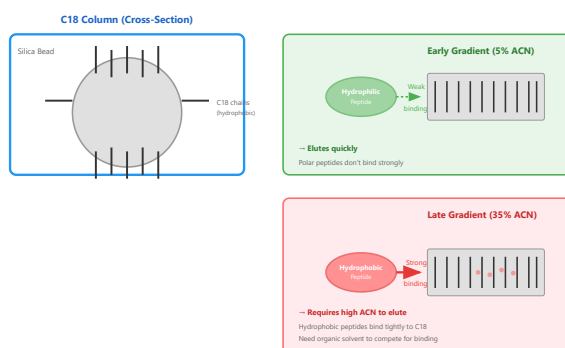
10-30%

Too short, hydrophobic,  
or modified peptides

### 3. Peptide Separation: Chromatography in Detail

#### 3.1 Reverse-Phase Mechanism

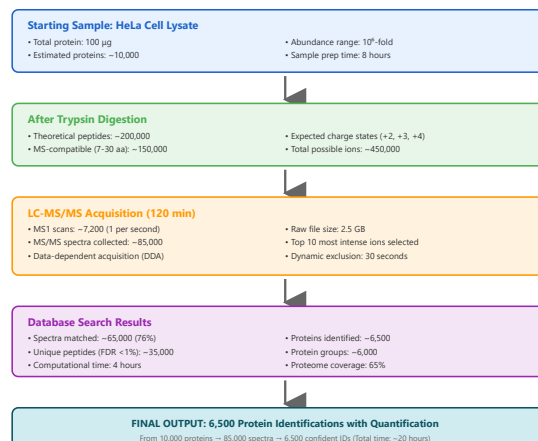
##### Molecular-Level View of Peptide Separation



### 4. Data Complexity: Real-World Example

#### 4.1 Case Study: HeLa Cell Proteome Analysis

##### From Sample to Identification: Numbers



**Key Insight:** Only about 65% of the proteome is typically identified in a single LC-MS run, even with modern instruments. The missing 35% includes low-abundance proteins, membrane proteins, very large/small proteins, and proteins with unfavorable chemical properties.

Fractionation or enrichment strategies are needed for deeper coverage.