プロテオーム解析における 質量分析技術と応用研究例

Mass Spectrometry-based Proteomics and its Applications

大阪大学大学院医学系研究科
薬理学講座 生体システム薬理学
岡西 広樹

Graduate School of Medicine, Osaka University
Department of Bio-system Pharmacology
Hiroki Okanishi

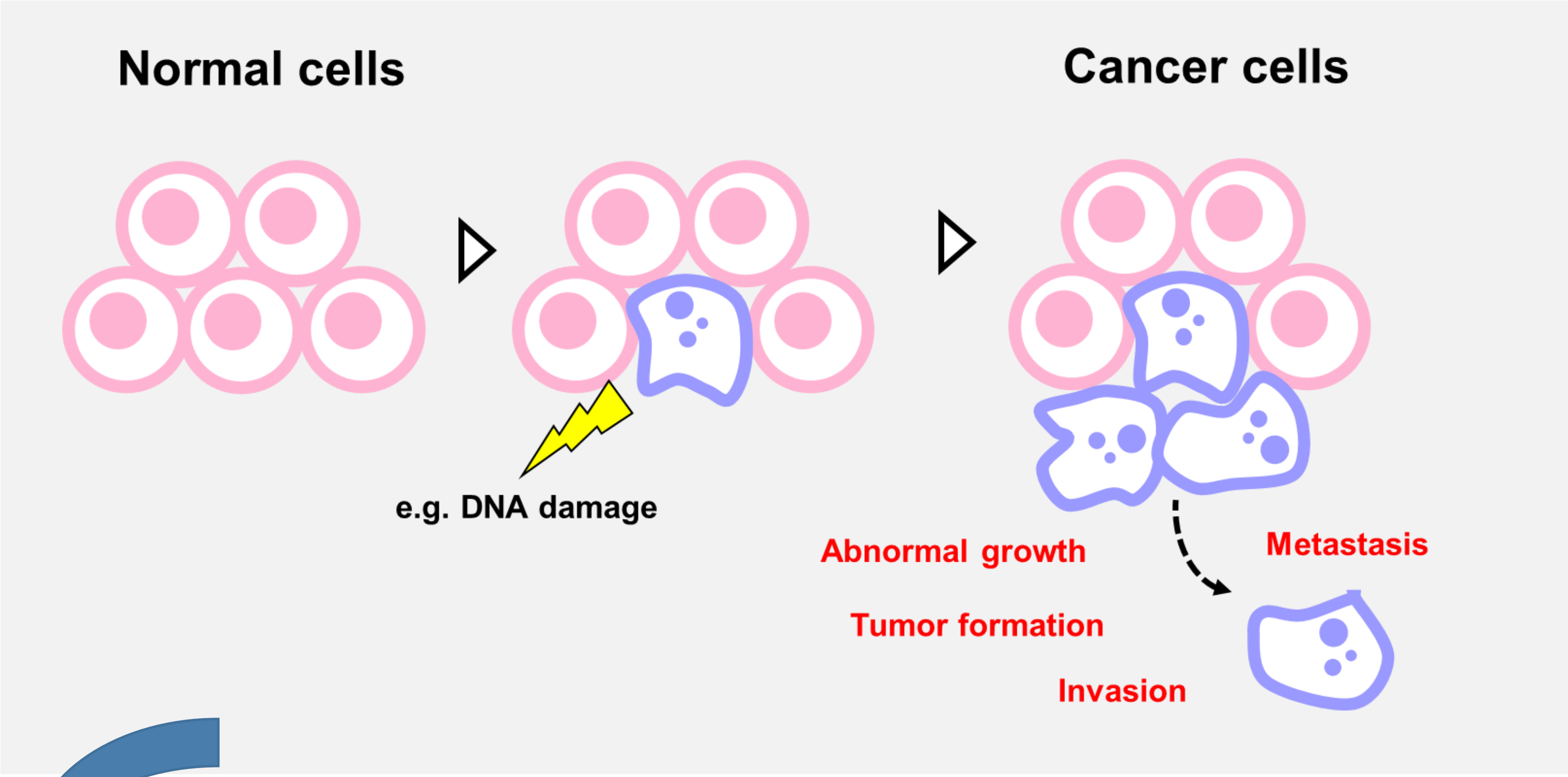
Today’s topics

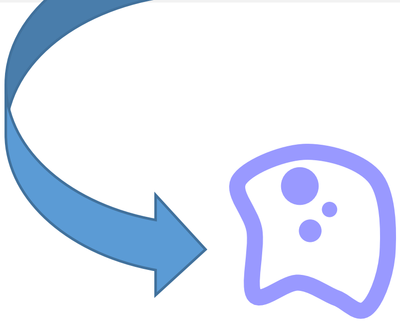
1. Mass spectrometry-based proteomics
2. Phosphoproteomics study on cancer cells treated with molecular targeted drug
3. LC-MS in CoMIT Omics Center (COC)

Today’s topics

1. Mass spectrometry-based proteomics
2. Phosphoproteomics study on cancer cells treated with molecular targeted drug
3. LC-MS in CoMIT Omics Center (COC)

Normal cells vs. cancer cells





The expression of proteins which contribute to growth, survival and function is dysregulated in cancer cells.

Analyze expressions of proteins associated to diseases

How to analyze protein expression

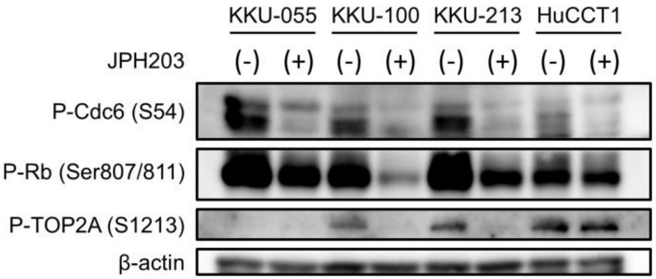
Western blotting, immunocytochemistry Immunohistochemistry, ELISA

- Detect target proteins using antibodies
- Can quantify expression of targeted proteins and reveal their localization
- Need specific antibodies

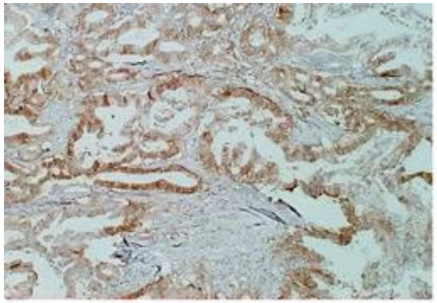
Mass spectrometry-based proteomics (Shotgun proteomics)

- Identify proteins using mass spectrometer
- Can quantify expression of thousands of proteins
- Don't need antibody (need protein sequence database)

Western blotting

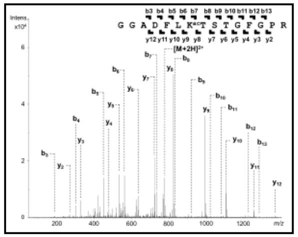


Immunohistochemistry



Kaira *et al.* (2012)

Mass spectrometry-based proteomics



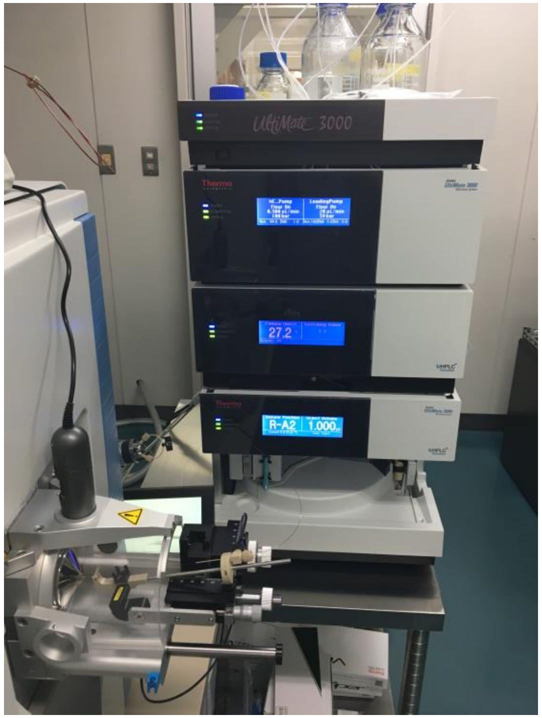
Genome = Gene (遺伝子) + Ome (全体)
Proteome = Protein (タンパク質) + Ome (全体)

Mass spectrometer

Mass spectrometer combined with nano liquid chromatography system (nLC-MS)

Mass spectrometer
(Q Exactive^{TM} mass spectrometer)

Nano liquid chromatography system
Ultimate^{TM} 3000 HPLC system

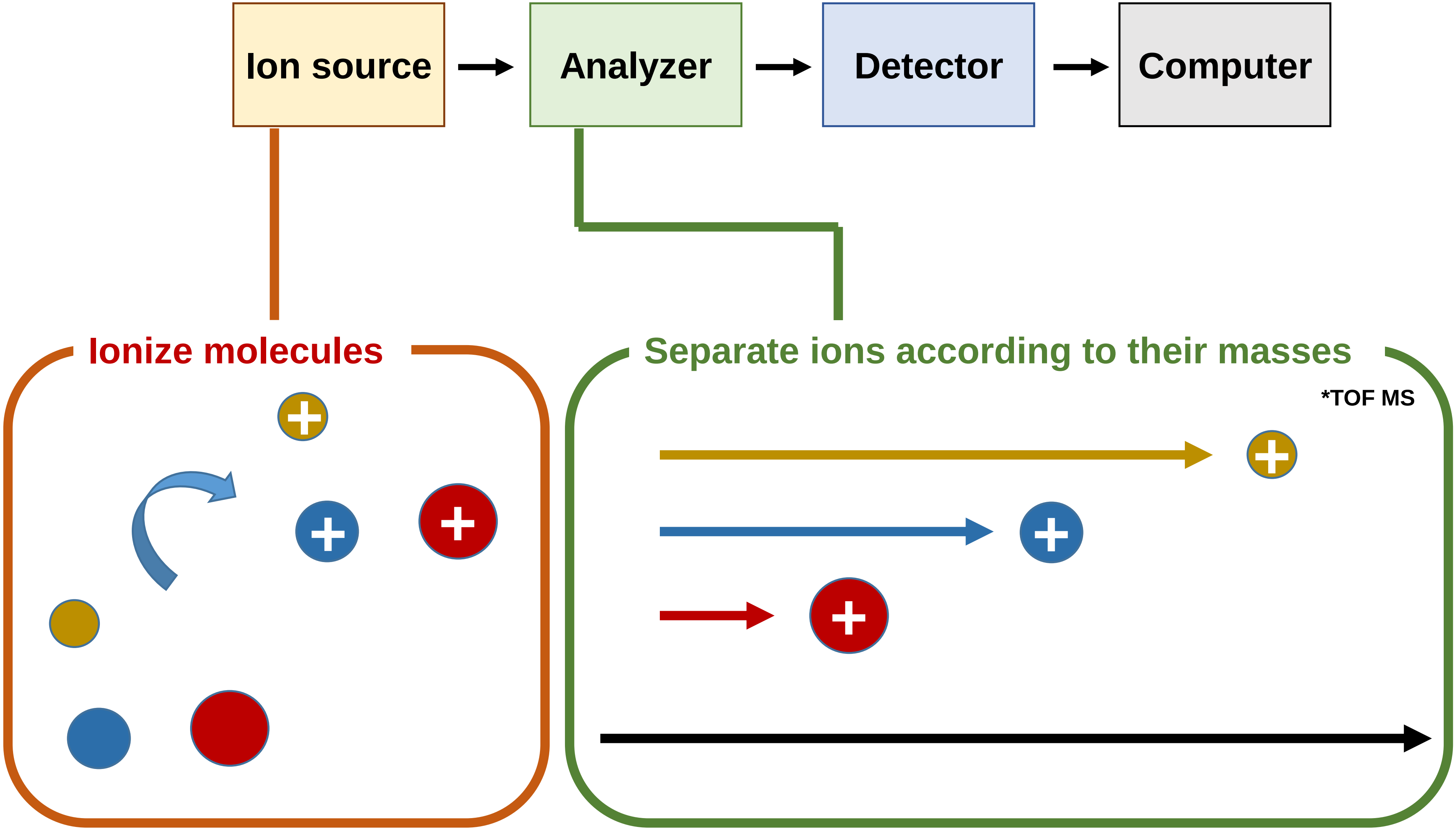


Mass spectrometer can measure the mass of molecule
(Resolution: 140,000 [maximum] Q-Exactive)

Mass spectrometer
(Exploris^{TM} 480)

Mass spectrometer can determine mass of molecule

Construction of mass spectrometer



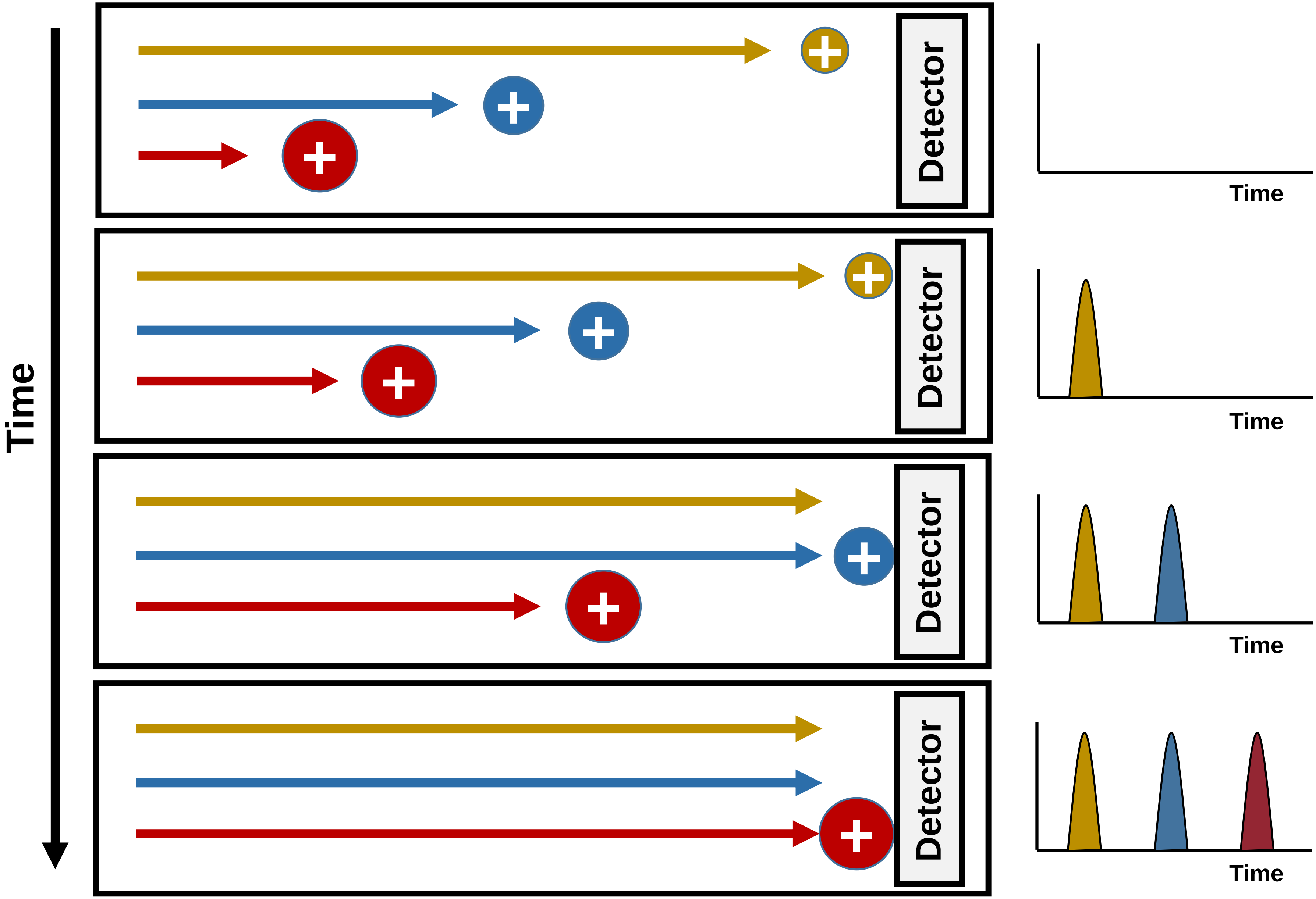
Ion source → Analyzer → Detector → Computer

Ionize molecules

Separate ions according to their masses

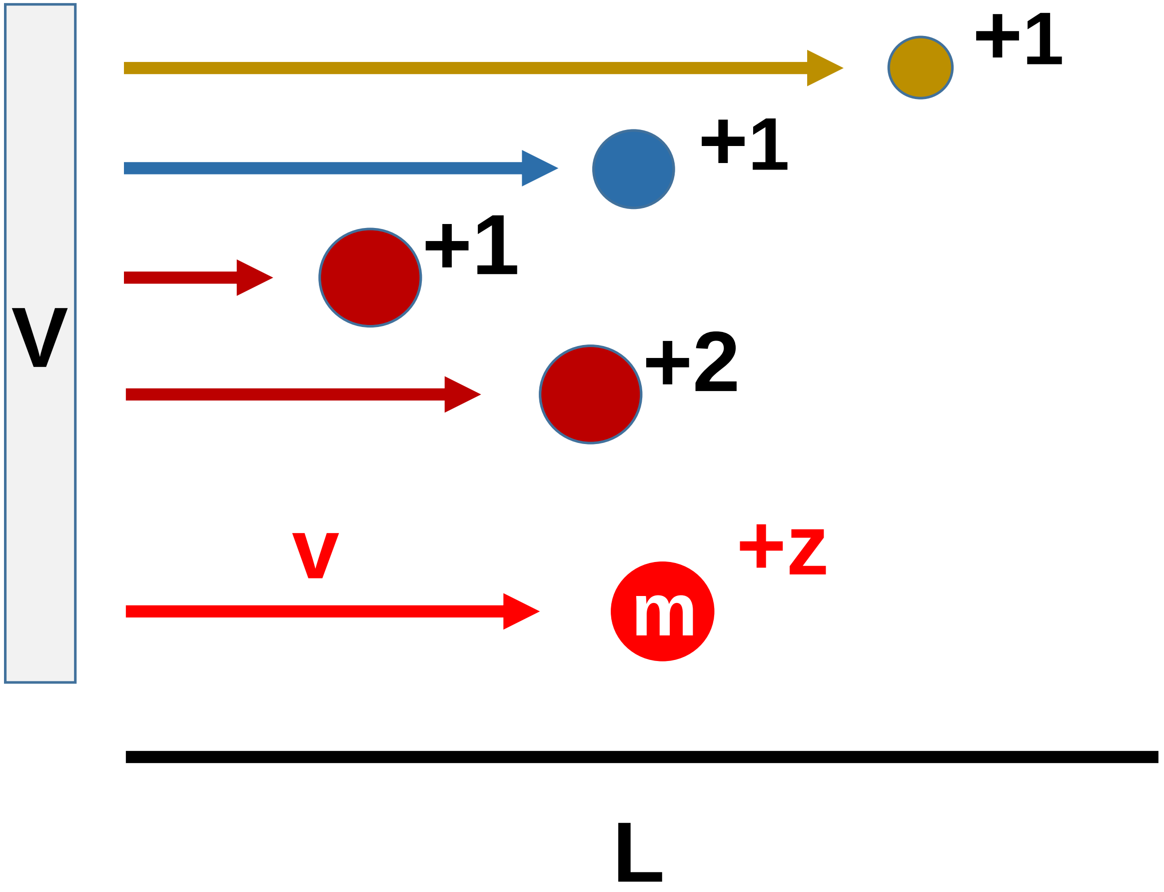
*TOF MS

Time-of-flight mass spectrometer



m/z

m/z: mass/charge number



$$zeV = \frac{1}{2}mv^2$$

$$\updownarrow \quad L = vt$$

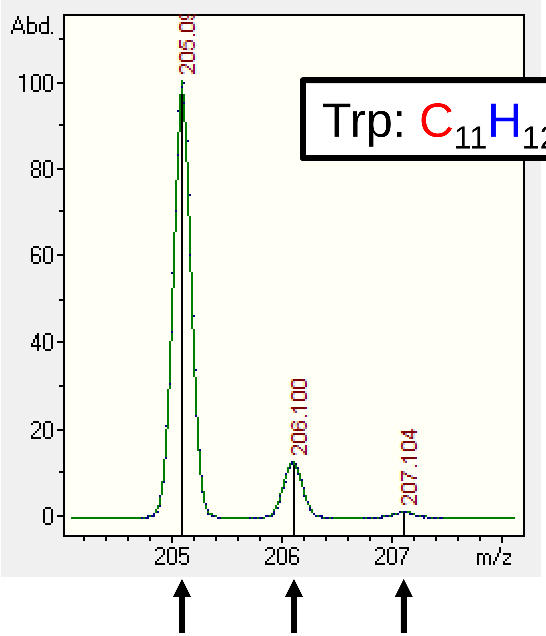
$$zeV = \frac{1}{2}m\left(\frac{L}{t}\right)^2$$

$$\updownarrow$$

$$t = L\sqrt{m/z \times \frac{1}{2eV}}$$

Average and monoisotopic mass

C, H, O, and N have isotopes
$[^{13}\mathrm{C}, ^{2}\mathrm{H}\ (\mathrm{D}), ^{15}\mathrm{N}, ^{17}\mathrm{O}, ^{18}\mathrm{O}]$



Trp: $C_{11}H_{12}N_2O_2$

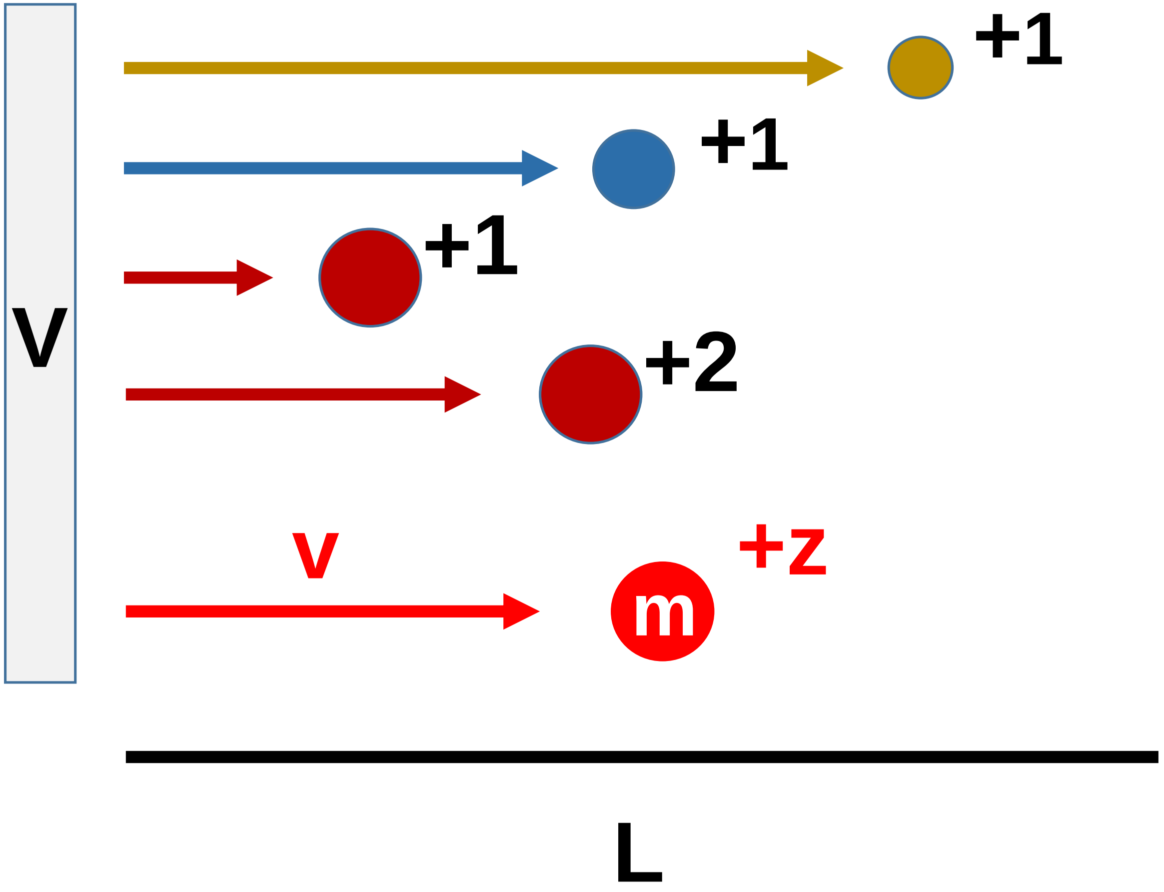
Isotopic peaks

Monoisotopic peak

	Mass of amino acid residue	
3-letter code	Average	Monoisotopic
Ala	71.079	71.037
Arg	156.188	156.101
Asn	114.104	114.043
Asp	115.089	115.027
Cys	103.139	103.009
Glu	129.116	129.043
Gln	128.131	128.059
Gly	57.052	57.021
His	137.141	137.059
Ile	113.159	113.084
Leu	113.159	113.084
Lys	128.174	128.095
Met	131.193	131.040
Phe	147.177	147.068
Pro	97.117	97.053
Ser	87.078	87.032
Thr	101.105	101.048
Trp	186.213	186.079
Tyr	163.176	163.063
Val	99.133	99.068

m/z

***m/z*: mass/charge number**



$$zeV = \frac{1}{2}mv^2$$

$$\updownarrow \quad L = vt$$

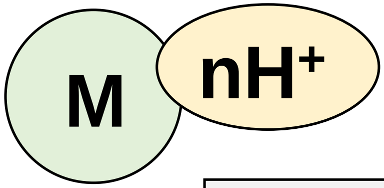
$$zeV = \frac{1}{2}m\left(\frac{L}{t}\right)^2$$

$$\updownarrow$$

$$t = L\sqrt{m/z \times \frac{1}{2eV}}$$

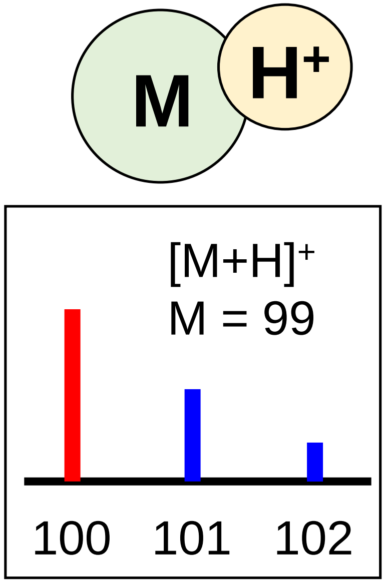
The mass can be calculated by *m/z* and isotopic peak pattern

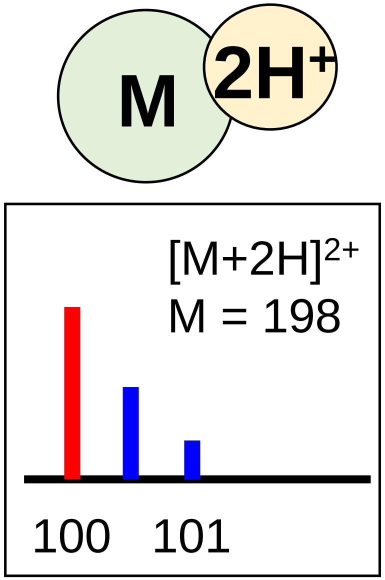
Molecules ionized by proton



m = M + n
z = n

$$m/z = \frac{M+n}{n}$$





Monoisotopic peak
→ m/z
Isotopic peaks
→ z

Type of mass spectrometer

Construction of mass spectrometer

Ion source → Analyzer → Detector → Computer

Ion source

- MALDI (Matrix Assisted Laser desorption/Ionization)
- ESI (Electrospray ionization)
- Electron ionization
- FAB (Fast atom bombardment)
- Chemical ionization

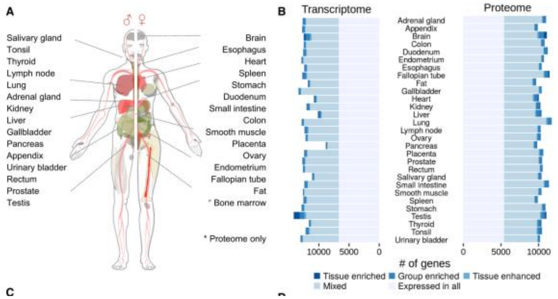
Analyzer

- TOF (Time-of-flight)
- Quadrupole
- Ion trap
- FT-ICR (Fourier-transform ion cyclotron resonance)
- Orbitrap
- Magnetic sector

Mass spectrometry-based protein analysis

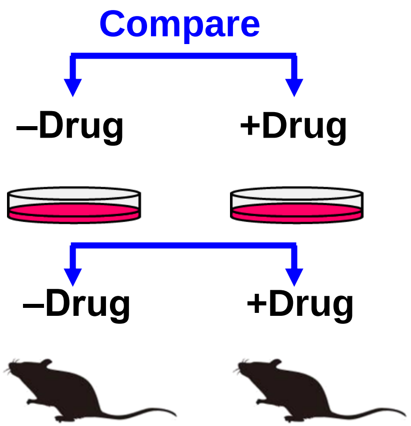
(1) Identification of unknown protein

(2) Profiling of proteome in cells or tissues

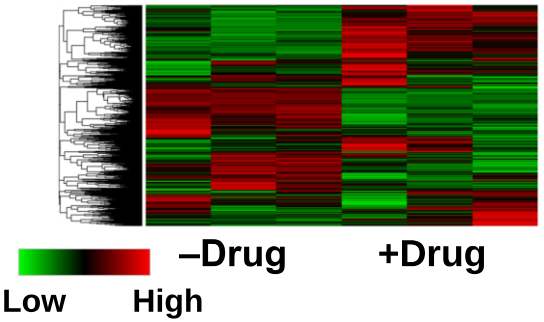


Wang *et al.* (2019) *Mol. Syst. Biol.*

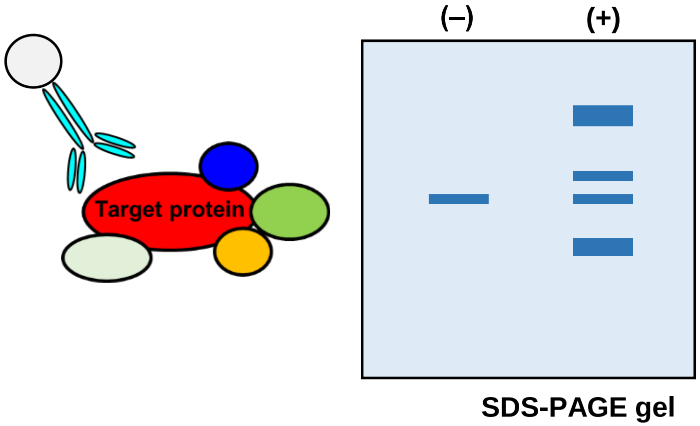
(4) Comparative quantification of protein expression



Thousands of proteins are quantified

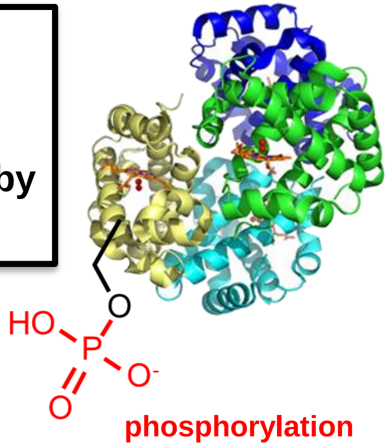


(3) Identification of proteins interacted with target protein

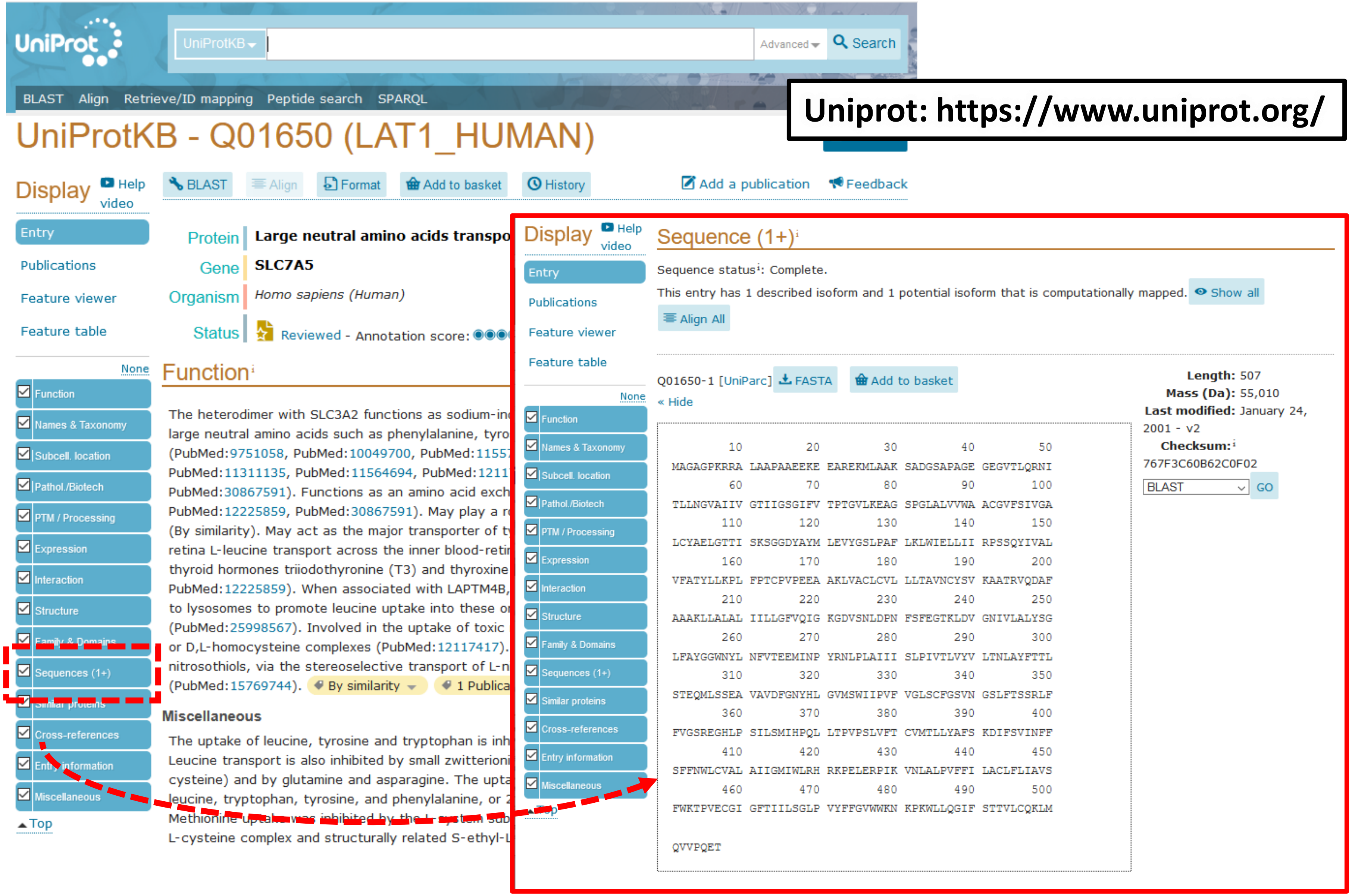


(5) Identification of phosphorylation sites (& PTM sites)

What type PTM?
Which sites?
When?
Up/down-regulated by drug?



Protein sequence database is available



Uniprot: https://www.uniprot.org/

UniProtKB - Q01650 (LAT1_HUMAN)

Protein: **Large neutral amino acids transpo**

Gene: **SLC7A5**

Organism: *Homo sapiens (Human)*

Status: Reviewed - Annotation score:

Function

The heterodimer with SLC3A2 functions as sodium-in
large neutral amino acids such as phenylalanine, tyro
(PubMed:9751058, PubMed:10049700, PubMed:1155
PubMed:11311135, PubMed:11564694, PubMed:1211
PubMed:30867591). Functions as an amino acid exch
PubMed:12225859, PubMed:30867591). May play a r
(By similarity). May act as the major transporter of t
retina L-leucine transport across the inner blood-reti
thyroid hormones triiodothyronine (T3) and thyroxine
PubMed:12225859). When associated with LAPTM4B,
to lysosomes to promote leucine uptake into these o
(PubMed:25998567). Involved in the uptake of toxic
or D,L-homocysteine complexes (PubMed:12117417).
nitrosothiols, via the stereoselective transport of L-n
(PubMed:15769744). By similarity 1 Publica

Miscellaneous

The uptake of leucine, tyrosine and tryptophan is inh
Leucine transport is also inhibited by small zwitterioni
cysteine) and by glutamine and asparagine. The upta
leucine, tryptophan, tyrosine, and phenylalanine, or 2
Methionine uptake was inhibited by the L-system sub
L-cysteine complex and structurally related S-ethyl-L

Sequence (1+)

Sequence status: Complete.

This entry has 1 described isoform and 1 potential isoform that is computationally mapped.

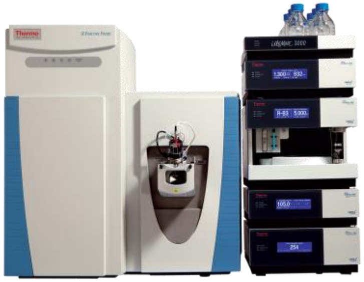
Q01650-1 [UniParc]

Length: 507
Mass (Da): 55,010
Last modified: January 24, 2001 - v2
Checksum:
767F3C60B62C0F02

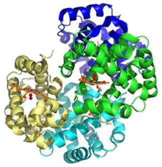
```
        10         20         30         40         50
MAGAGPKRRA LAAPAAEEKE EAREKMLAAK SADGSAPAGE GEGVTLQRNI
        60         70         80         90        100
TLLNGVAIIV GTIIGSGIFV TPTGVLKEAG SPGLALVVWA ACGVFSIVGA
       110        120        130        140        150
LCYAELGTTI SKSGGDYAYM LEVYGSLPAF LKLWIELLII RPSSQYIVAL
       160        170        180        190        200
VFATYLLKPL FPTCPVPEEA AKLVACLCVL LLTAVNCYSV KAATRVQDAF
       210        220        230        240        250
AAAKLLALAL IILLGFVQIG KGDVSNLDPN FSFEGTKLDV GNIVLALYSG
       260        270        280        290        300
LFAYGGWNYL NFVTEEMINP YRNLPLAIII SLPIVTLVYV LTNLAYFTTL
       310        320        330        340        350
STEQMLSSEA VAVDFGNYHL GVMSWIIPVF VGLSCFGSVN GSLFTSSRLF
       360        370        380        390        400
FVGSREGHLP SILSMIHPQL LTPVPSLVFT CVMTLLYAFS KDIFSVINFF
       410        420        430        440        450
SFFNWLCVAL AIIGMIWLRH RKPELERPIK VNLALPVFFI LACLFLIAVS
       460        470        480        490        500
FWKTPVECGI GFTIILSGLP VYFFGVWWKN KPKWLLQGIF STTVLCQKLM


QVVPQET
```

How does mass spectrometry identify protein? (1)

Mass spectrometer (MS)



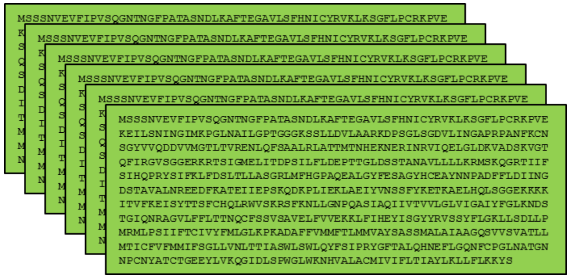
Protein



Measured
mass of protein

Protein sequences

~20,000 proteins



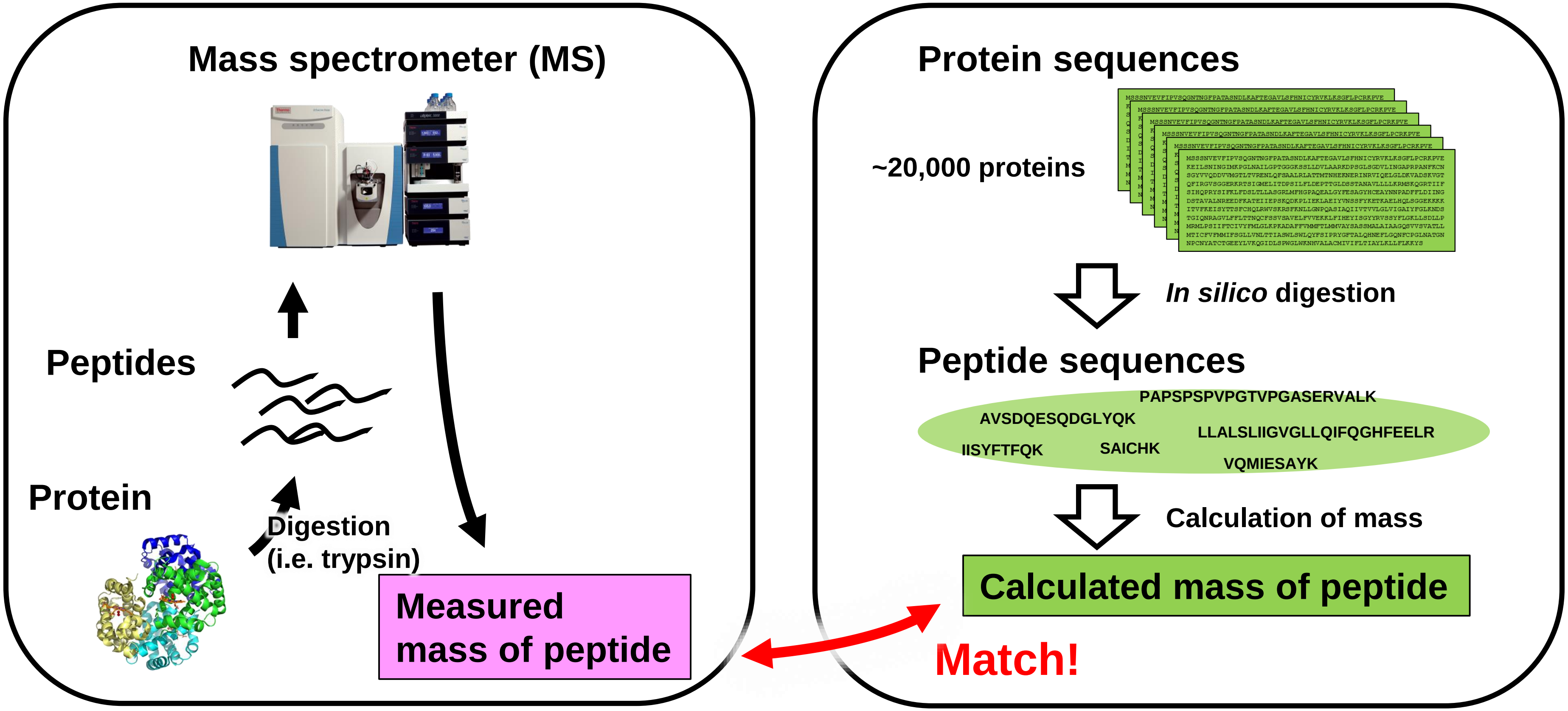
```
MSSSNVEVFIPVSQGNTNGFPATASNDLKAFTEGAVLSFHNICYRVKLKSGFLPCRKPVE
KEILSNINGIMKPGLNAILGPTGGGKSSLLDVLAARKDPSGLSGDVLINGAPRPANFKCN
SGYVVQDDVVMGTLTVRENLQFSAALRLATTMTNHEKNERINRVIQELGLDKVADSKVGT
QFIRGVSGGERKRTSIGMELITDPSILFLDEPTTGLDSSTANAVLLLLKRMSKQGRTIIF
SIHQPRYSIFKLFDSLTLLASGRLMFHGPAQEALGYFESAGYHCEAYNNPADFFLDIING
DSTAVALNREEDFKATEIIEPSKQDKPLIEKLAEIYVNSSFYKETKAELHQLSGGEKKKK
ITVFKEISYTTSFCHQLRWVSKRSFKNLLGNPQASIAQIIVTVVLGLVIGAIYFGLKNDS
TGIQNRAGVLFFLTTNQCFSSVSAVELFVVEKKLFIHEYISGYYRVSSYFLGKLLSDLLP
MRMLPSIIFTCIVYFMLGLKPKADAFFVMMFTLMMVAYSASSMALAIAAGQSVVSVATLL
MTICFVFMMIFSGLLVNLTTIASWLSWLQYFSIPRYGFTALQHNEFLGQNFCPGLNATGN
NPCNYATCTGEEYLVKQGIDLSPWGLWKNHVALACMIVIFLTIAYLKLLFLKKYS
```

Calculation of mass

Calculated mass of protein

Match!

How does mass spectrometry identify protein? (2)



Mass spectrometer (MS)

Peptides

Protein

Digestion (i.e. trypsin)

Measured mass of peptide

Protein sequences

~20,000 proteins

***In silico* digestion**

Peptide sequences

PAPSPSPVPGTVPGASERVALK

AVSDQESQDGLYQK

LLALSLIIGVGLLQIFQGHFEELR

IISYFTFQK

SAICHK

VQMIESAYK

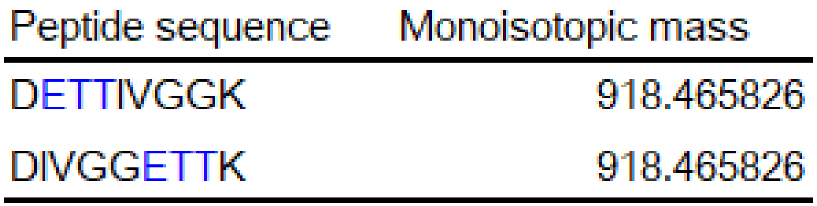
Calculation of mass

Calculated mass of peptide

Match!

Peptide mass fingerprinting (PMF) & MS/MS

MS level identification (PMF) has a problem

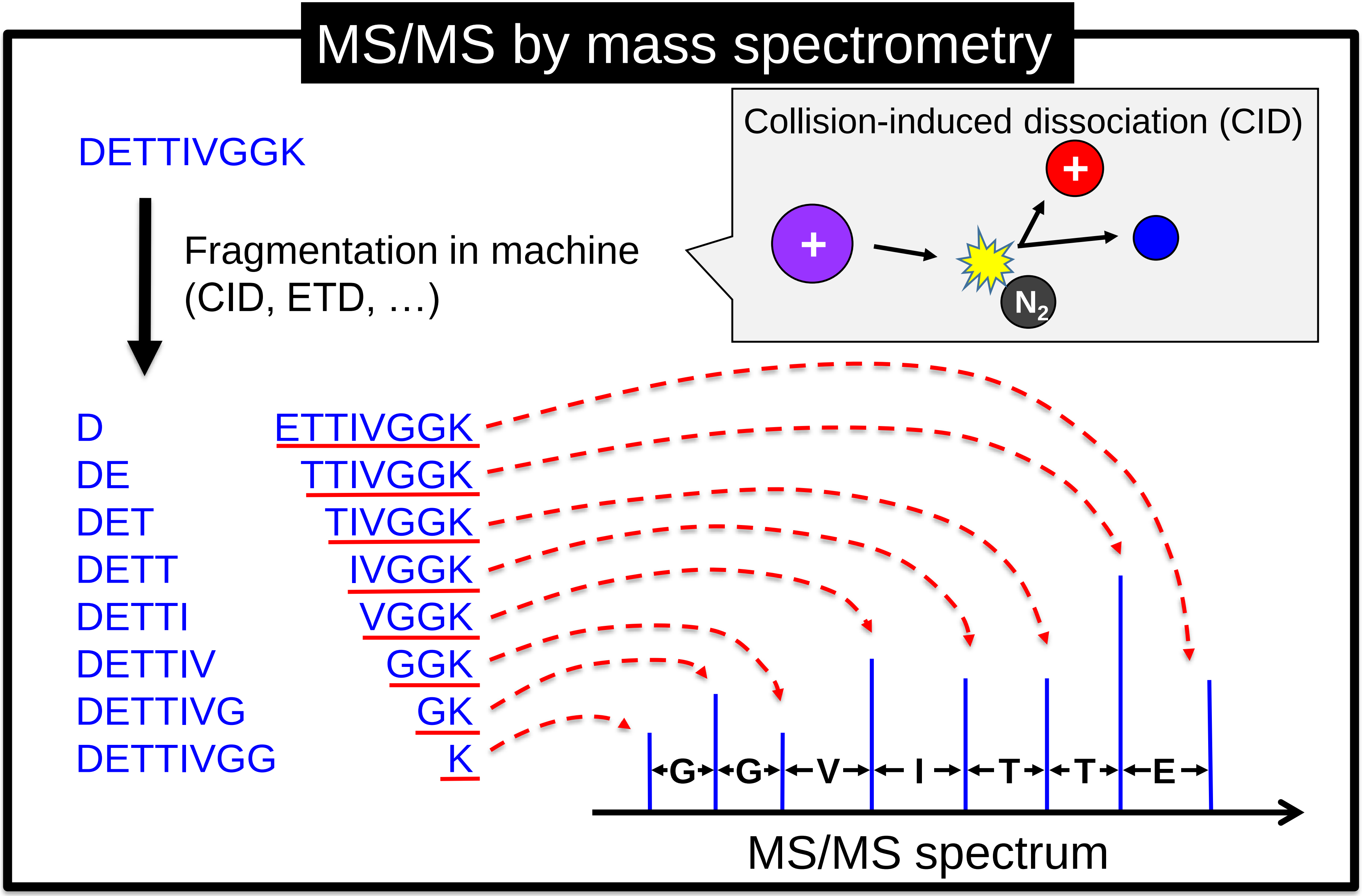


Peptide sequence	Monoisotopic mass
DETTIVGGK	918.465826
DIVGGETTK	918.465826

These different peptides (different sequence) have same mass

In order to distinguish these two peptides, peptide sequence information is needed.

Identification of peptides by Tandem MS (MS/MS)



MS/MS by mass spectrometry

DETTIVGGK

Fragmentation in machine
(CID, ETD, …)

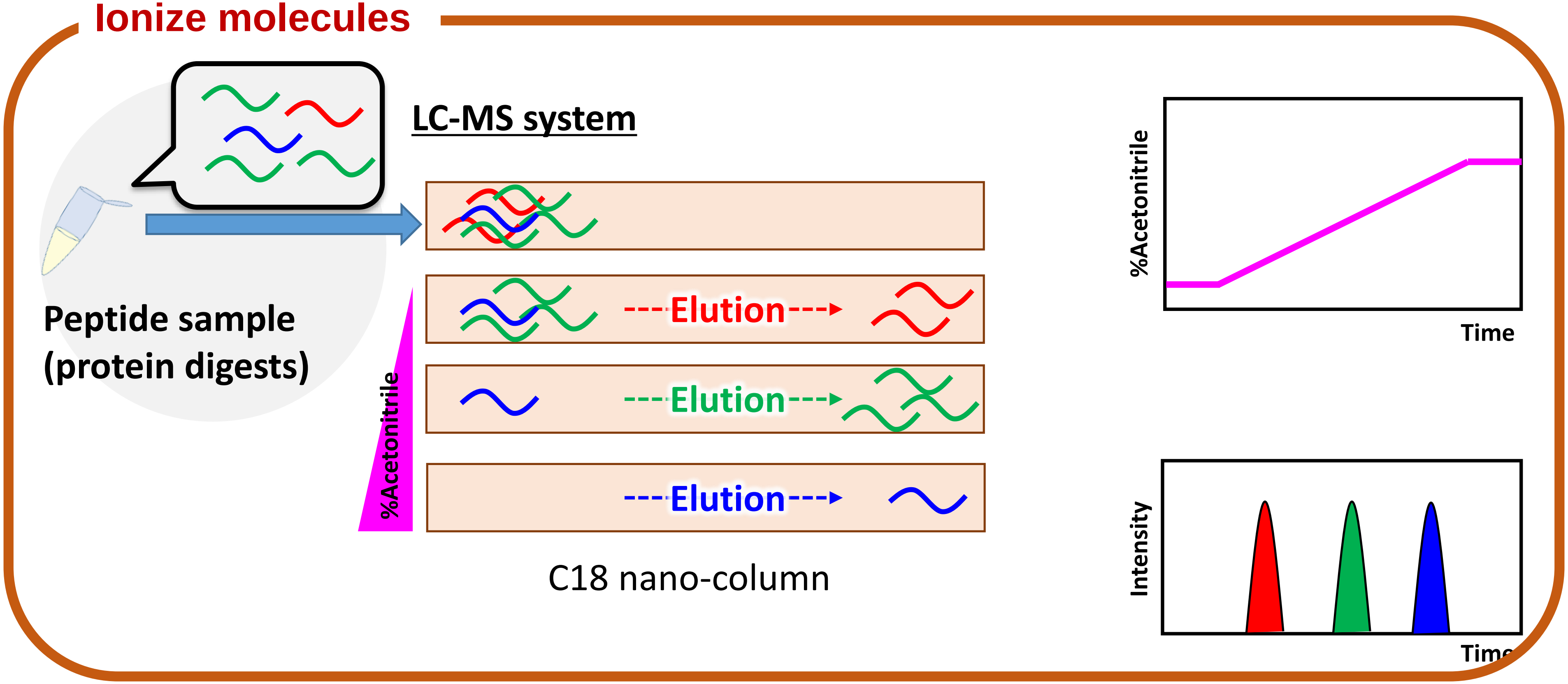
D	ETTIVGGK
DE	TTIVGGK
DET	TIVGGK
DETT	IVGGK
DETTI	VGGK
DETTIV	GGK
DETTIVG	GK
DETTIVGG	K

Liquid chromatography-mass spectrometry system (LC-MS system)

Construction of LC-MS

LC → Ion source → Analyzer → Detector → Computer

Ionize molecules



Today’s topics

1. Mass spectrometry-based proteomics

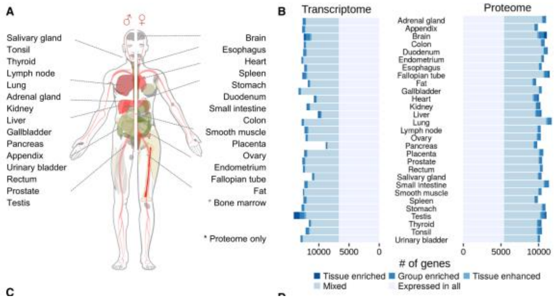
2. Phosphoproteomics study on cancer cells treated with molecular targeted drug

3. LC-MS in CoMIT Omics Center (COC)

Mass spectrometry-based protein analysis

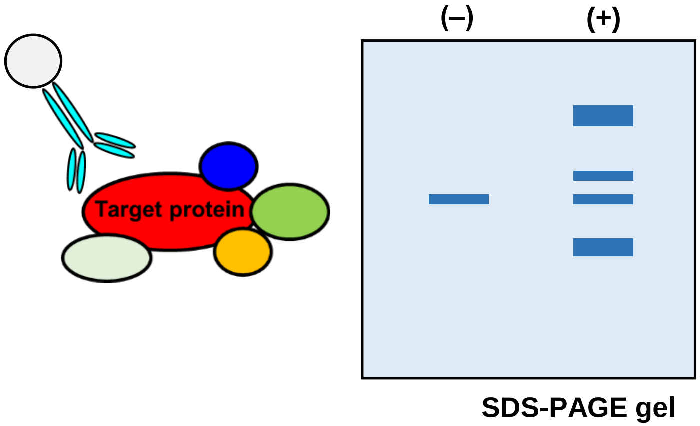
(1) Identification of unknown protein

(2) Profiling of proteome in cells or tissues

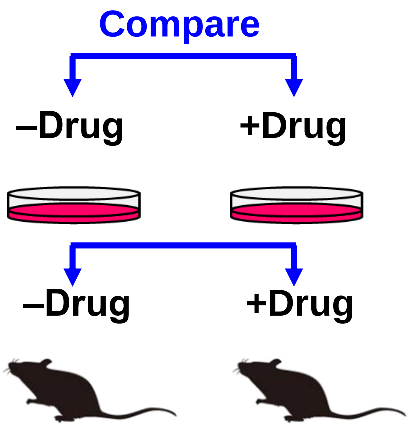


Wang *et al.* (2019) *Mol. Syst. Biol.*

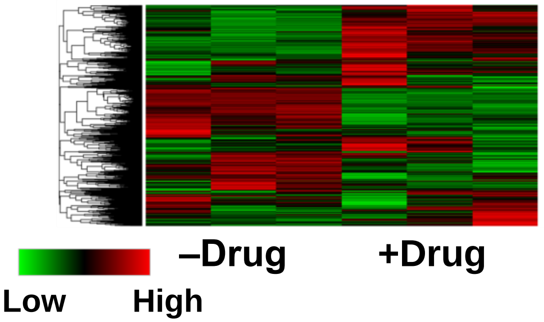
(3) Identification of proteins interacted with target protein



(4) Comparative quantification of protein expression

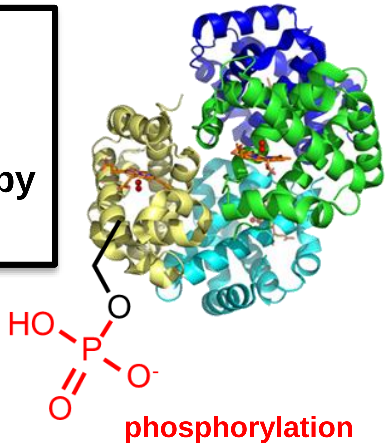


Thousands of proteins are quantified

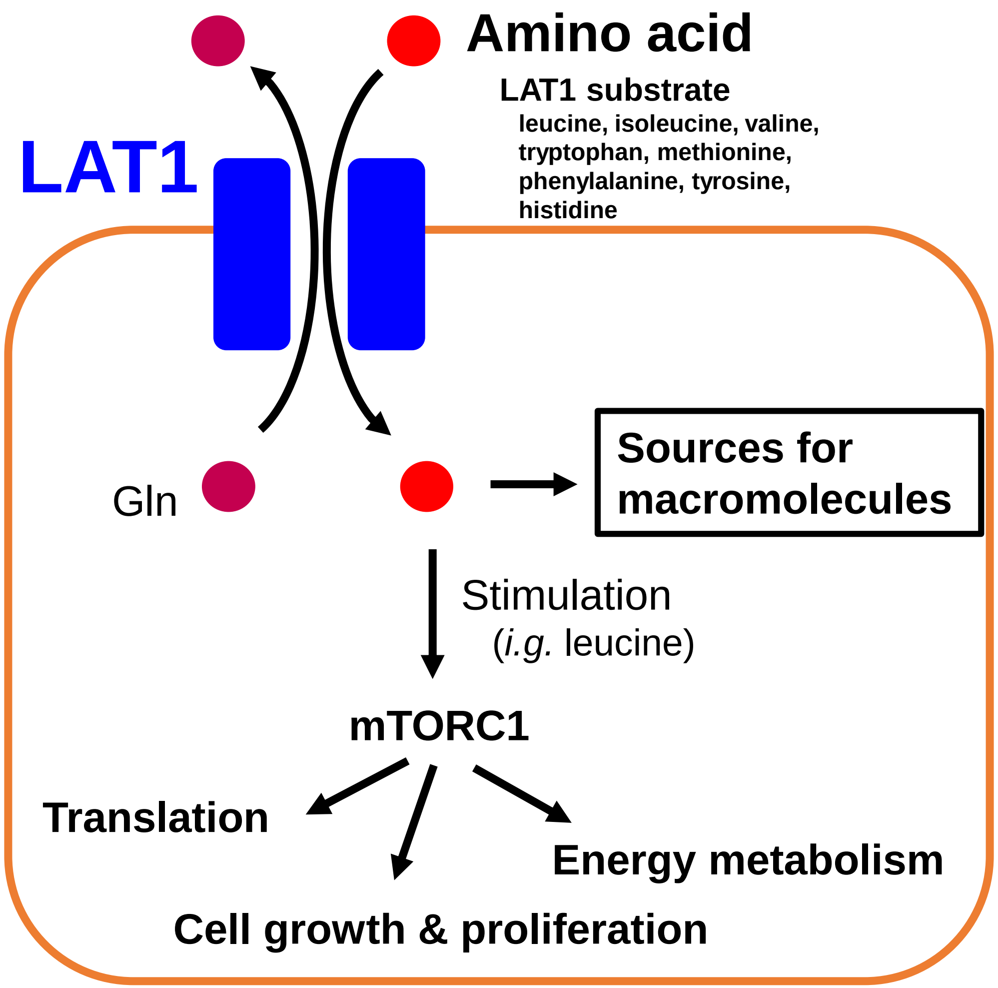


(5) Identification of phosphorylation sites (& PTM sites)

What type PTM?
Which sites?
When?
Up/down-regulated by drug?



LAT1: Amino acid transporter highly expressed in cancer cells



Amino acid

LAT1 substrate

leucine, isoleucine, valine,
tryptophan, methionine,
phenylalanine, tyrosine,
histidine

LAT1

Gln

Sources for macromolecules

Stimulation
(*i.g.* leucine)

mTORC1

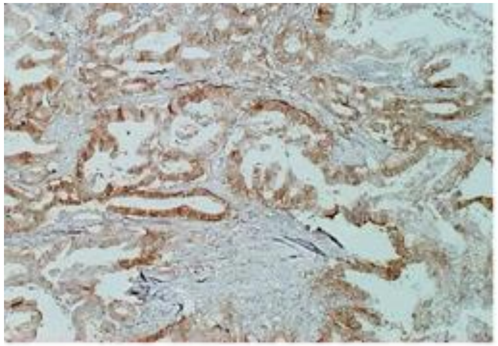
Translation

Energy metabolism

Cell growth & proliferation

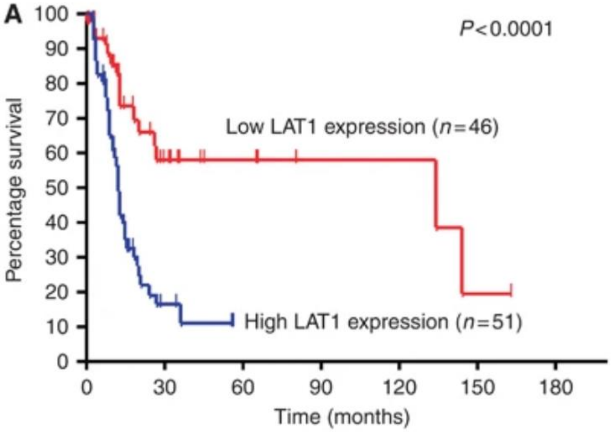
LAT1 substrates function (1) as sources of macromolecules needed for rapid growth, and (2) as signaling molecules

LAT1 is highly expressed in tumor



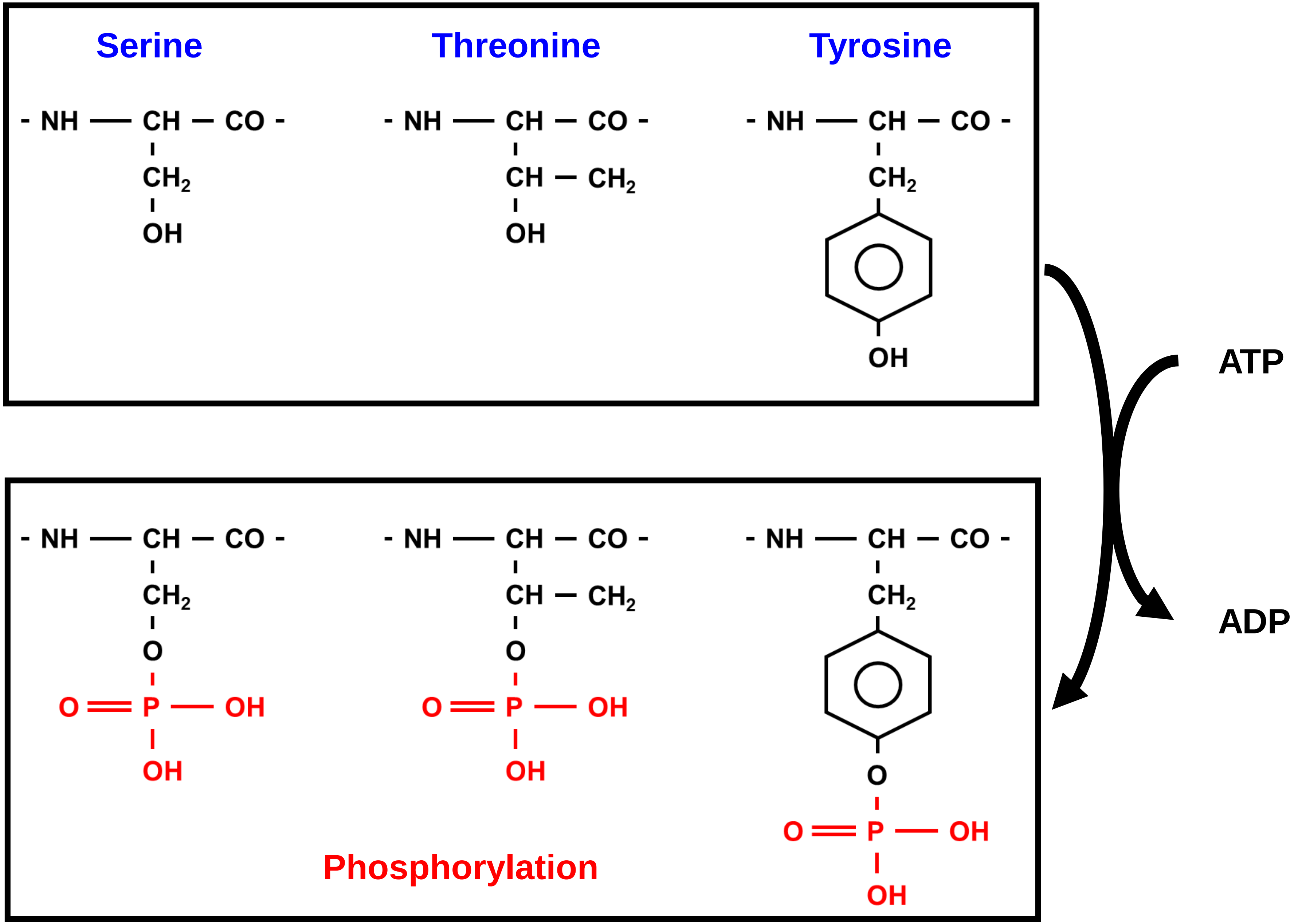
Kaira, K., et al. (2012). *British journal of cancer*, *107*, 632–638.

Poor prognosis of patient with highly expressed LAT1



Kaira, K., et al. (2012). *British journal of cancer*, *107*, 632–638.

Ser, Thr, Tyr phosphorylation



Serine **Threonine** **Tyrosine**

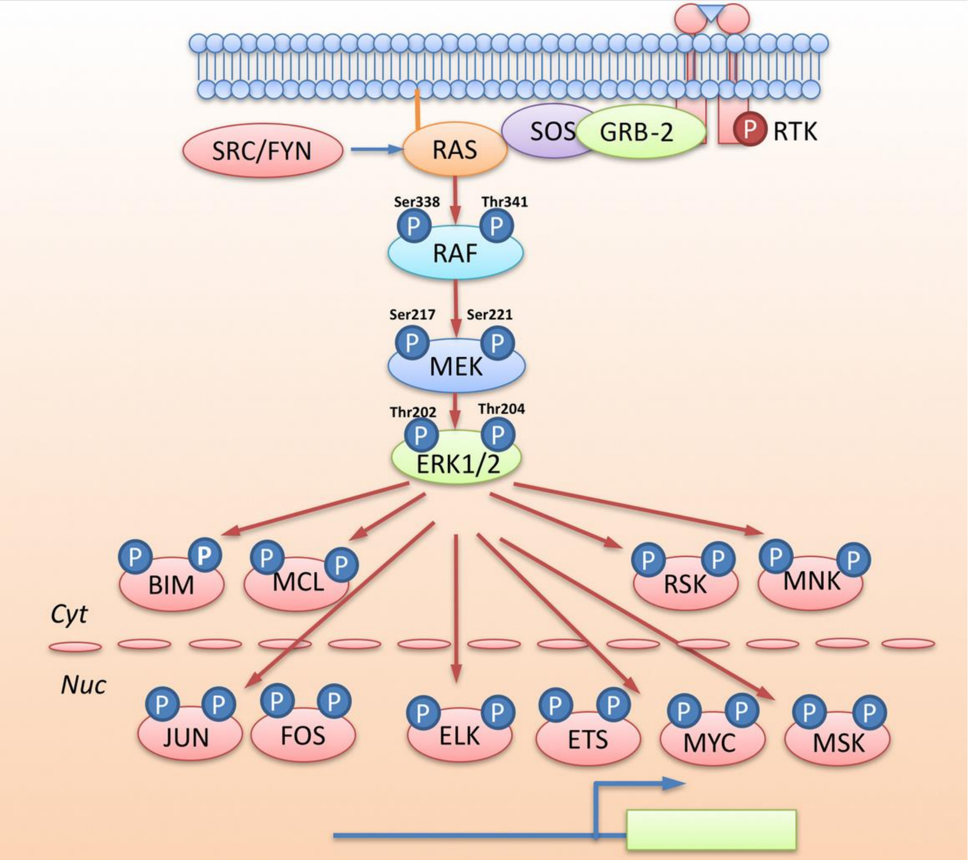
ATP

ADP

Phosphorylation

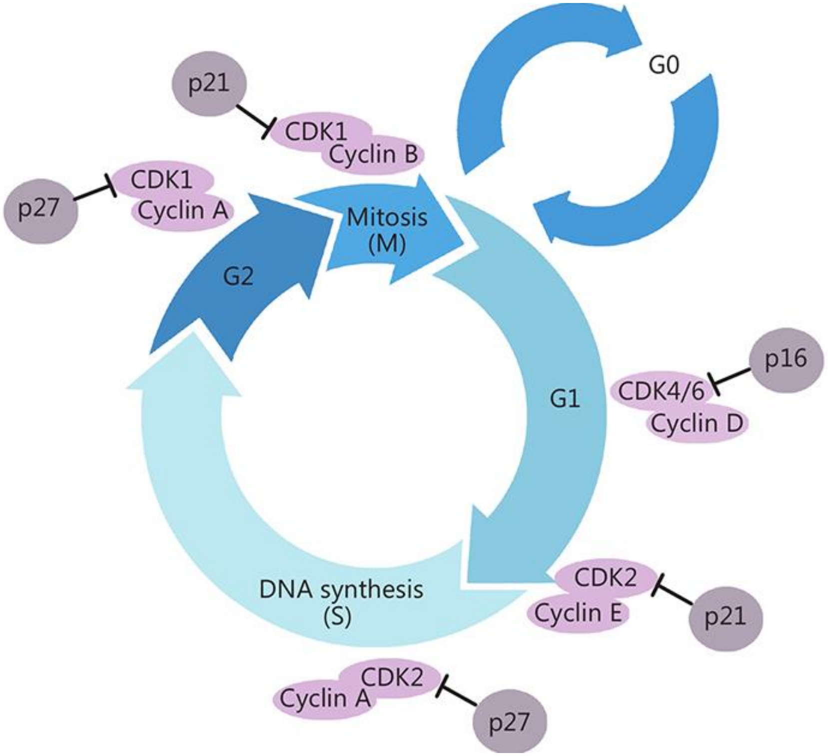
Phosphorylation signaling pathways

MAPK pathway



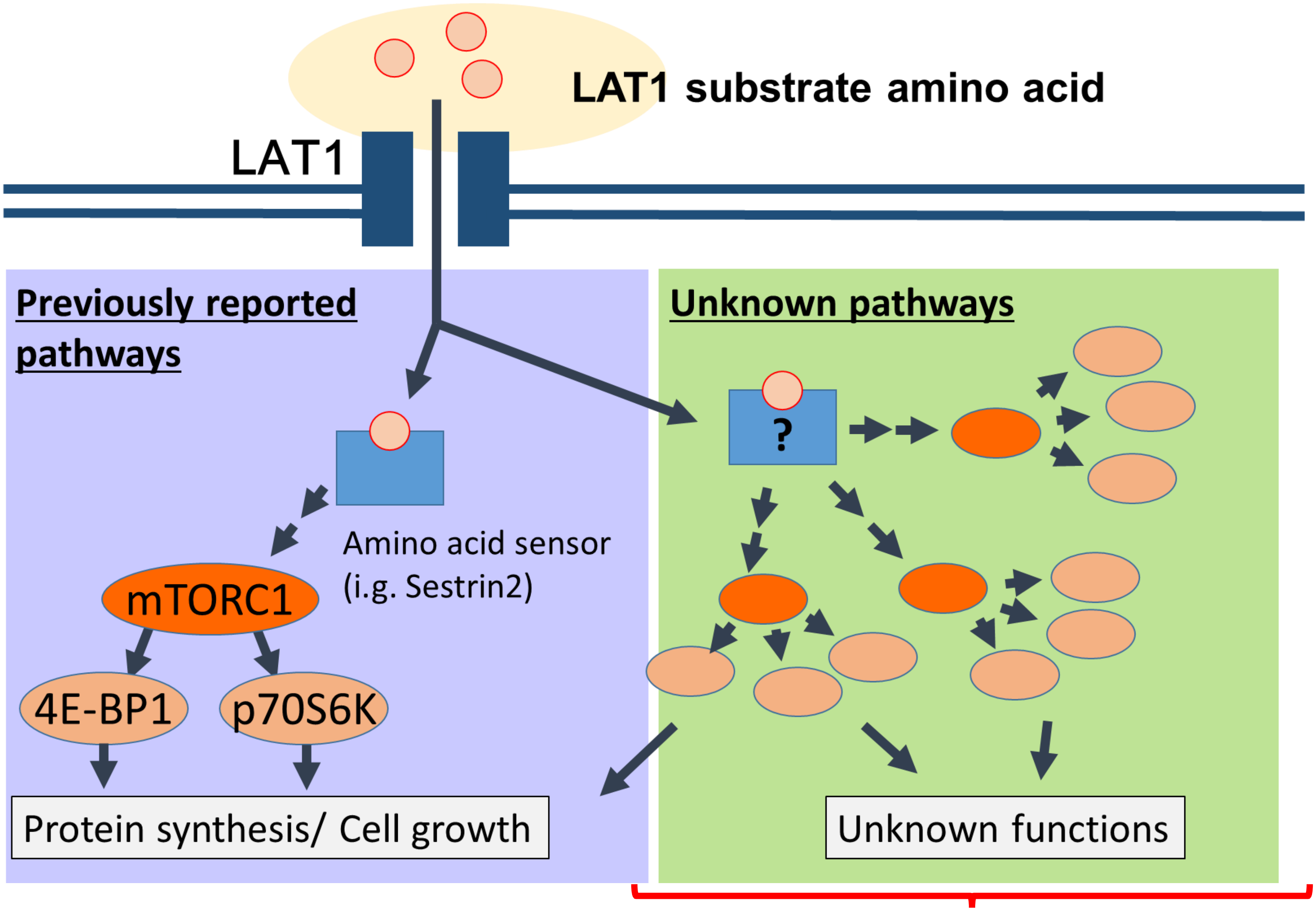
Liu, F. et al. (2018). *Acta pharmaceutica Sinica. B*, *8*(4), 552–562.

Cell cycle



Bai, J., Li, Y., & Zhang, G. (2017). *Cancer biology & medicine*, 14(4), 348–362.

LAT1 and phosphorylation signaling



LAT1 substrate amino acid

LAT1

Previously reported pathways

Amino acid sensor (i.g. Sestrin2)

Protein synthesis/ Cell growth

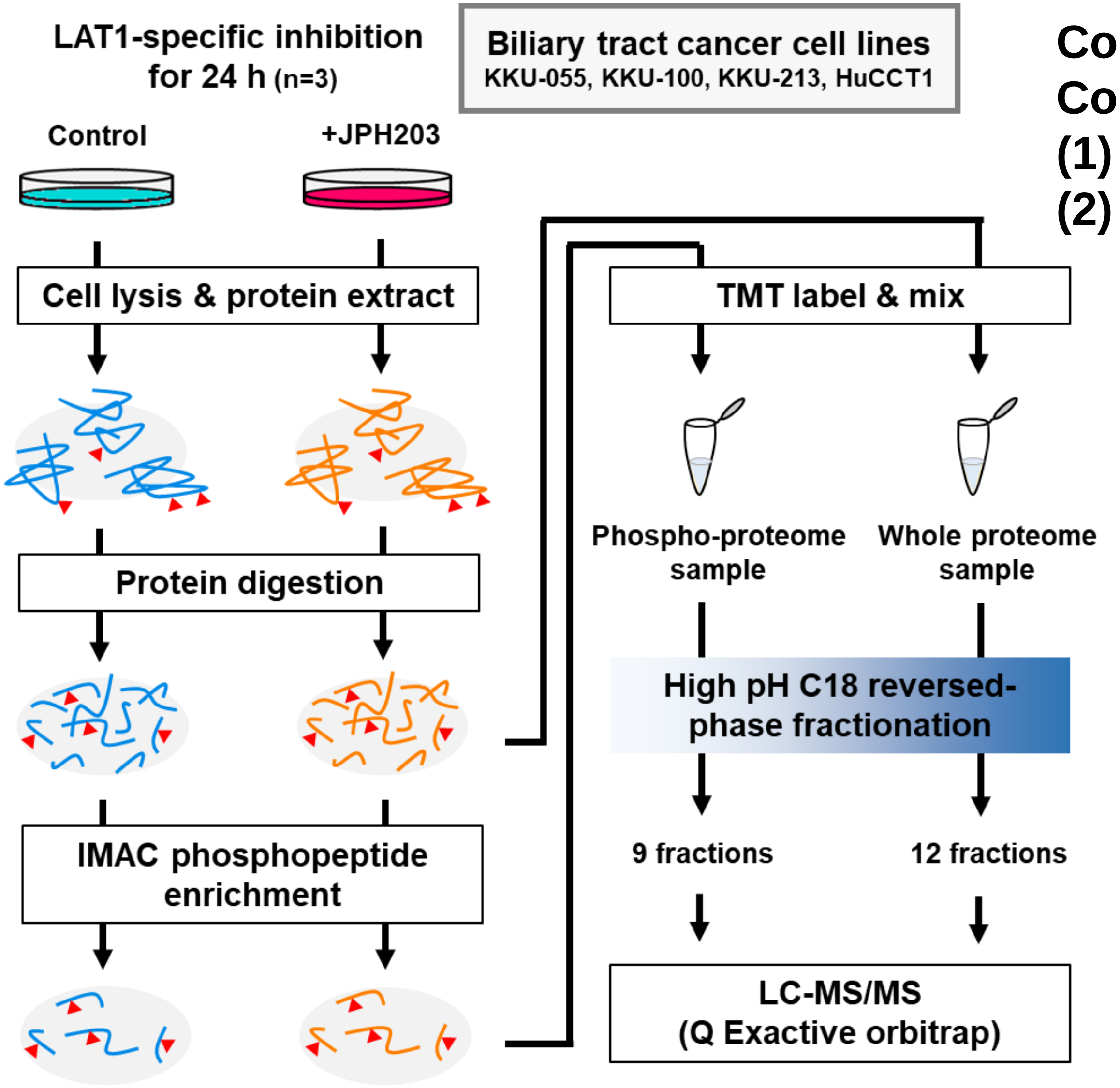
Unknown pathways

Unknown functions

Proteomics and phosphoproteomics study

Phosphoproteomics workflow

Comparative quantification
Control vs. Drug-treatment
(1) Protein expression
(2) Phosphorylation



LAT1-specific inhibition for 24 h (n=3)

Biliary tract cancer cell lines
KKU-055, KKU-100, KKU-213, HuCCT1

Cell lysis & protein extract

Protein digestion

IMAC phosphopeptide enrichment

TMT label & mix

High pH C18 reversed-phase fractionation

LC-MS/MS (Q Exactive orbitrap)

Okanishi, H., *et al.* (2021). Proteomics and phosphoproteomics reveal key regulators associated with cytostatic effect of amino acid transporter LAT1 inhibitor. *Cancer science*, *112*, 871–883.

Quantitative proteomics methods

Labeling quantification

Reporter ion-based

- iTRAQ (Isobaric Tag for Relative and Absolute Quantitation)
- TMT (Tandem Mass Tag)

Precursor ion-based

- SILAC (Stable Isotope Labeling)
- Dimethyl labeling
- ^{18}O labeling

Label-free quantification

Precursor ion-based

- Chromatogram peak area-based

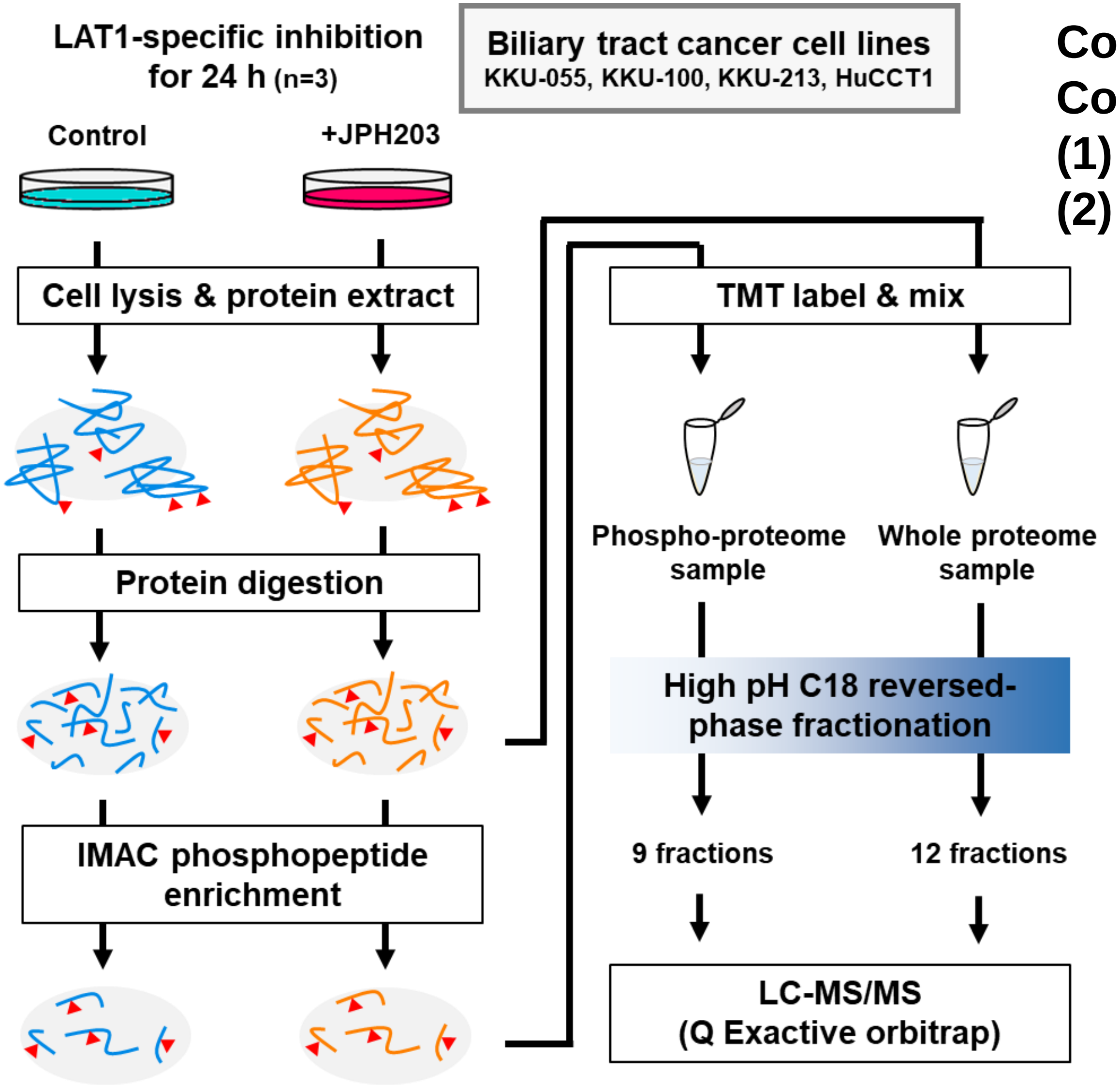
Data-independent

- SWATH-MS

Others

- Spectral counts
- emPAI

Phosphoproteomics workflow



LAT1-specific inhibition for 24 h (n=3)

Biliary tract cancer cell lines
KKU-055, KKU-100, KKU-213, HuCCT1

Control

+JPH203

Cell lysis & protein extract

Protein digestion

IMAC phosphopeptide enrichment

TMT label & mix

Phospho-proteome sample

Whole proteome sample

High pH C18 reversed-phase fractionation

9 fractions

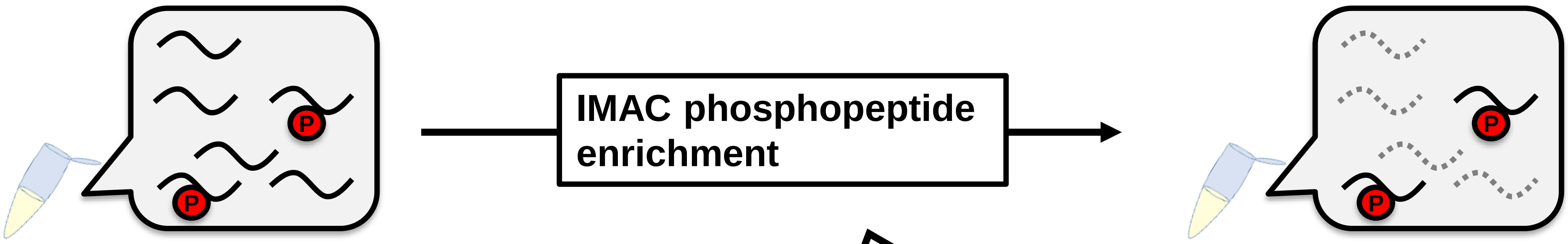
12 fractions

LC-MS/MS
(Q Exactive orbitrap)

Comparative quantification
Control vs. Drug-treatment
(1) Protein expression
(2) Phosphorylation

Okanishi, H., *et al.* (2021). Proteomics and phosphoproteomics reveal key regulators associated with cytostatic effect of amino acid transporter LAT1 inhibitor. *Cancer science*, *112*, 871–883.

Phosphopeptide enrichment technique: Immobilized metal affinity chromatography (IMAC)



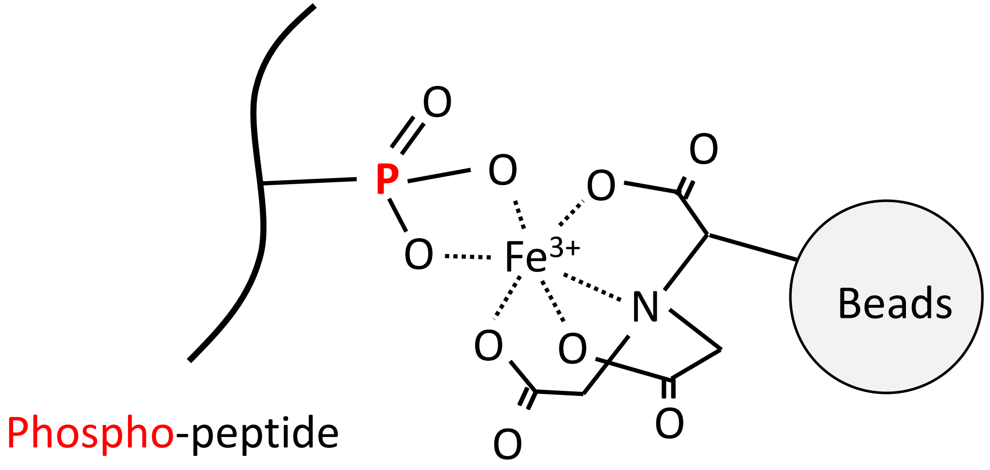
Too many peptides!

IMAC phosphopeptide enrichment

Peptide sample

Phospho-peptides

Immobilized metal affinity chromatography (IMAC)



Identification and quantification

Identification summary

Average identified peptides:

54106 peptides (q-value < 0.01) from 6431 proteins

Average identified phosphopeptides:

21340 phosphopeptides (q-value < 0.01)

Quantification summary

Average quantifiable proteins:

4983 proteins (#unique ≥ 2 with q-value < 0.01)

Average quantifiable phosphorylation sites:

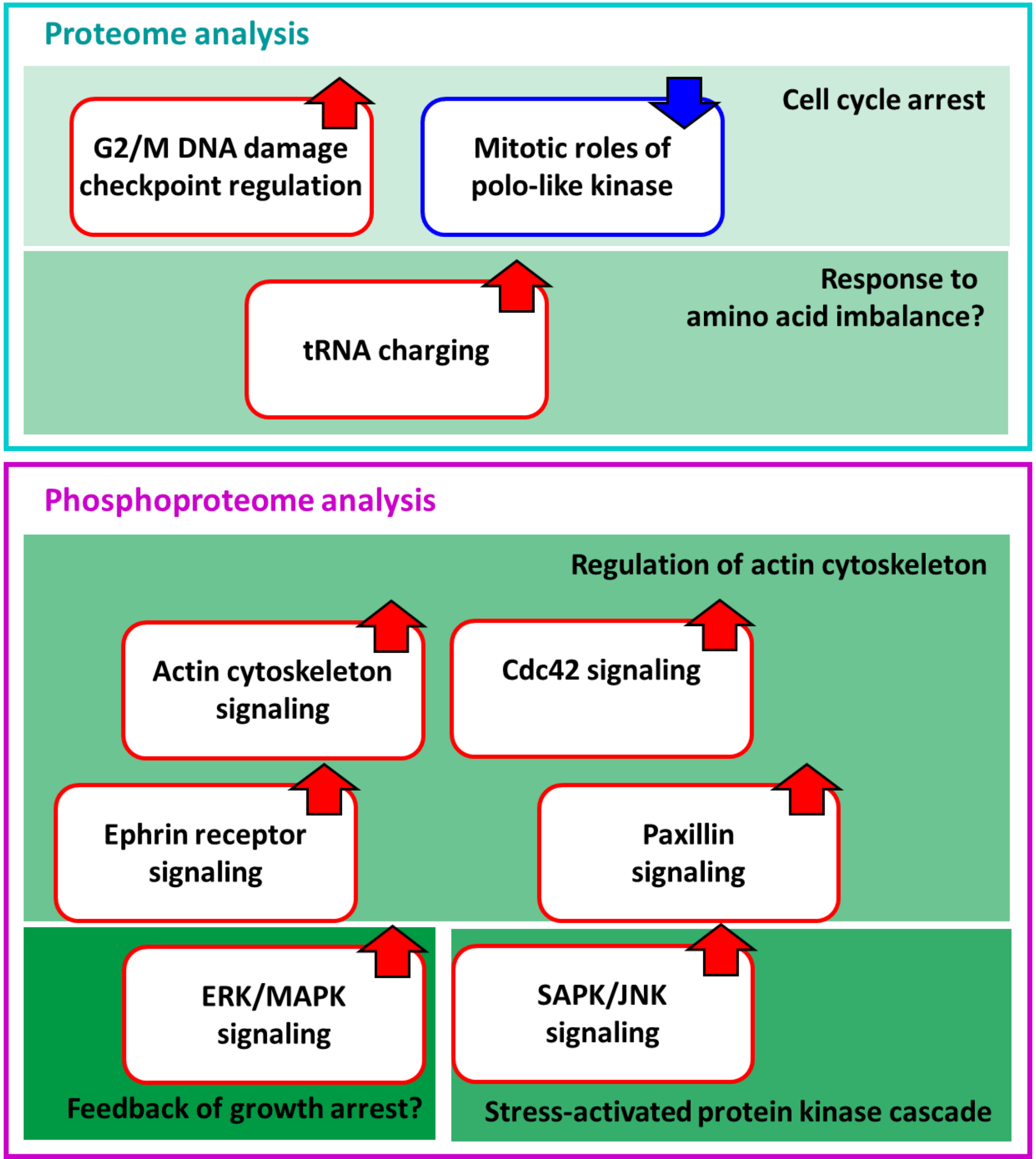
15019 phosphorylation sites (q-value < 0.01, Site probability > 0.75)

Differentially expressed proteins and phosphorylated sites

(p-value < 0.05 and fold change ≥1.2)

Proteomics	**Phosphoproteomics**
348 proteins in KKU-055	4777 sites in KKU-055
342 proteins in KKU-100	3943 sites in KKU-100
252 proteins in KKU-213	3990 sites in KKU-213
1315 proteins in HuCCT-1	5574 sites in HuCCT1

Commonly activate/inactivated pathways



Proteome analysis

Cell cycle arrest

G2/M DNA damage checkpoint regulation

Mitotic roles of polo-like kinase

Response to amino acid imbalance?

tRNA charging

Phosphoproteome analysis

Regulation of actin cytoskeleton

Actin cytoskeleton signaling

Cdc42 signaling

Ephrin receptor signaling

Paxillin signaling

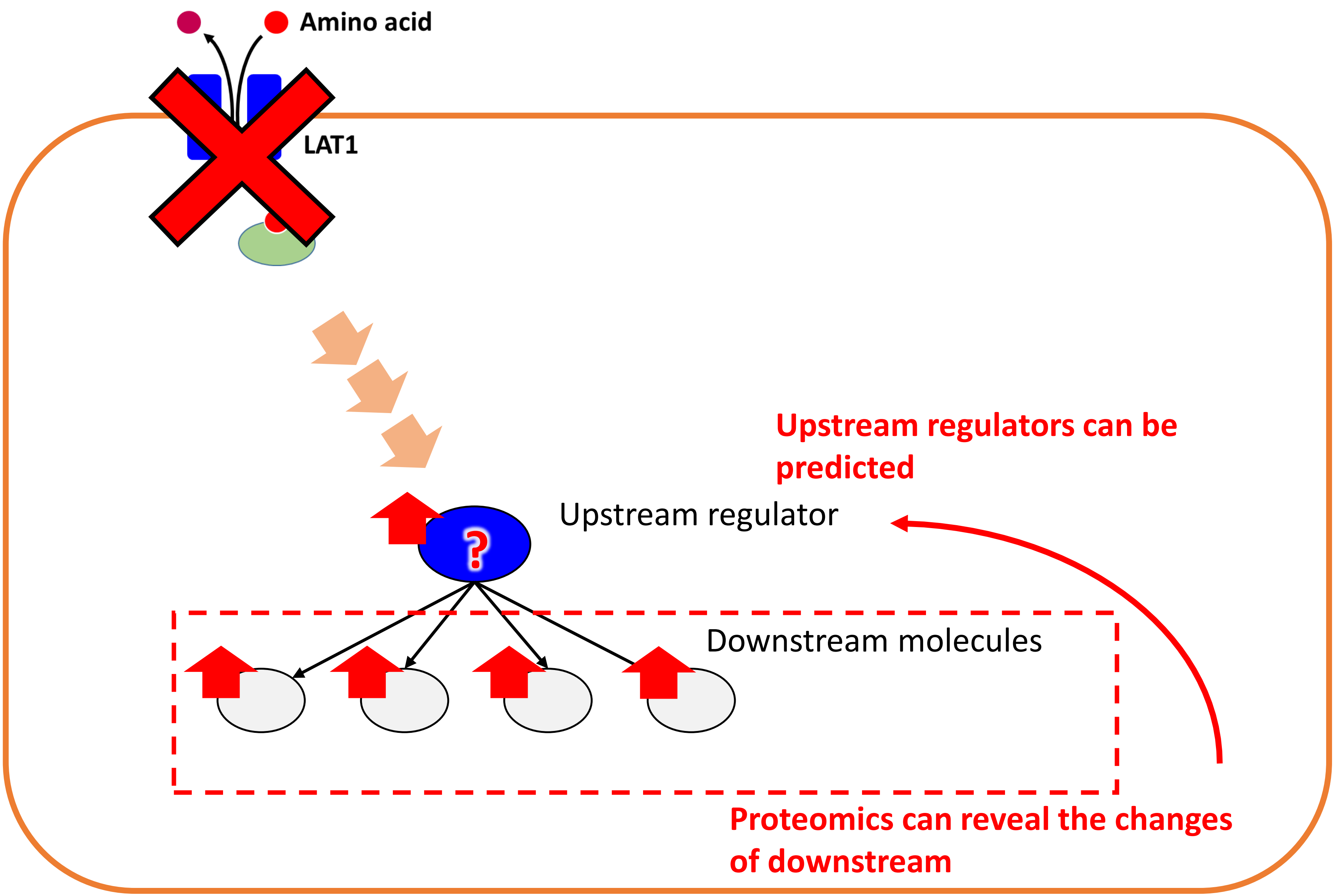
ERK/MAPK signaling

Feedback of growth arrest?

SAPK/JNK signaling

Stress-activated protein kinase cascade

Upstream regulator analysis



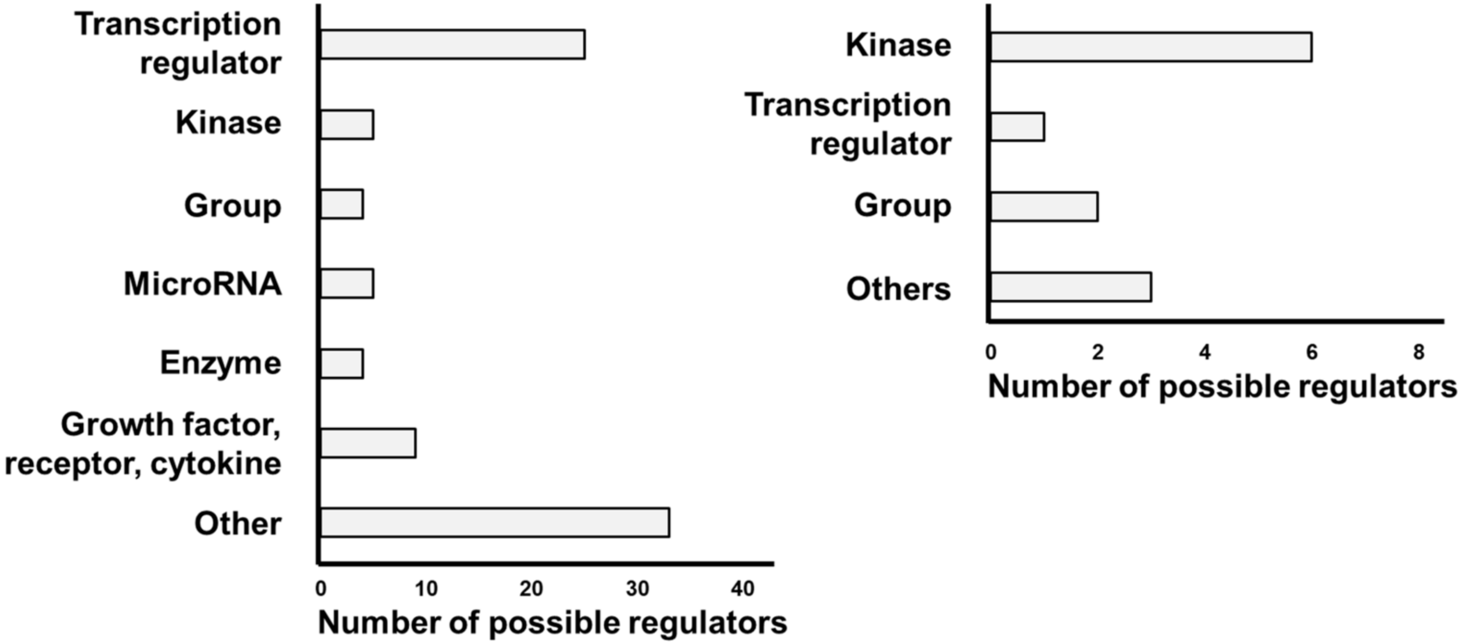
Upstream regulators can be predicted

Proteomics can reveal the changes of downstream

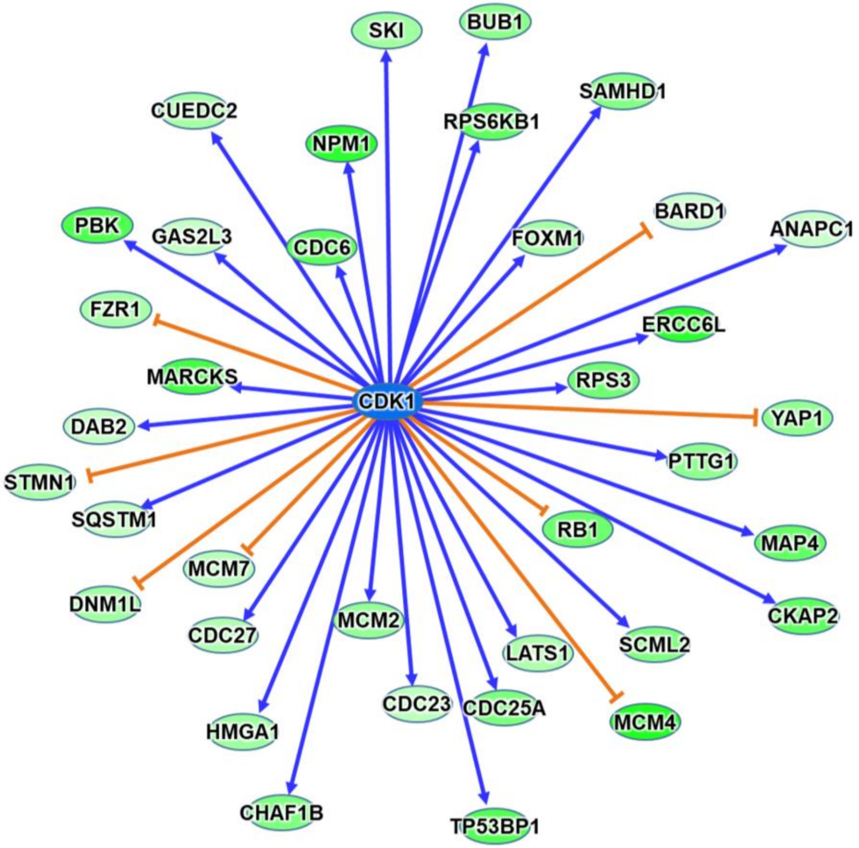
Upstream regulator analysis

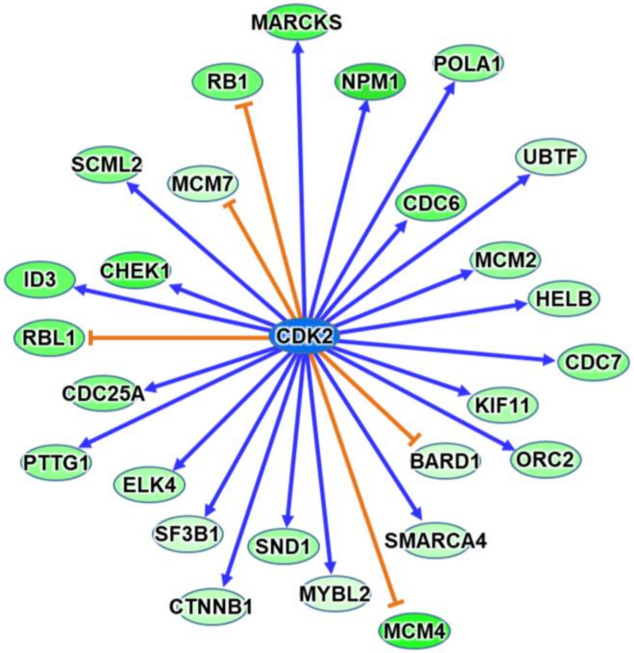
Proteomics

Phospho-proteomics



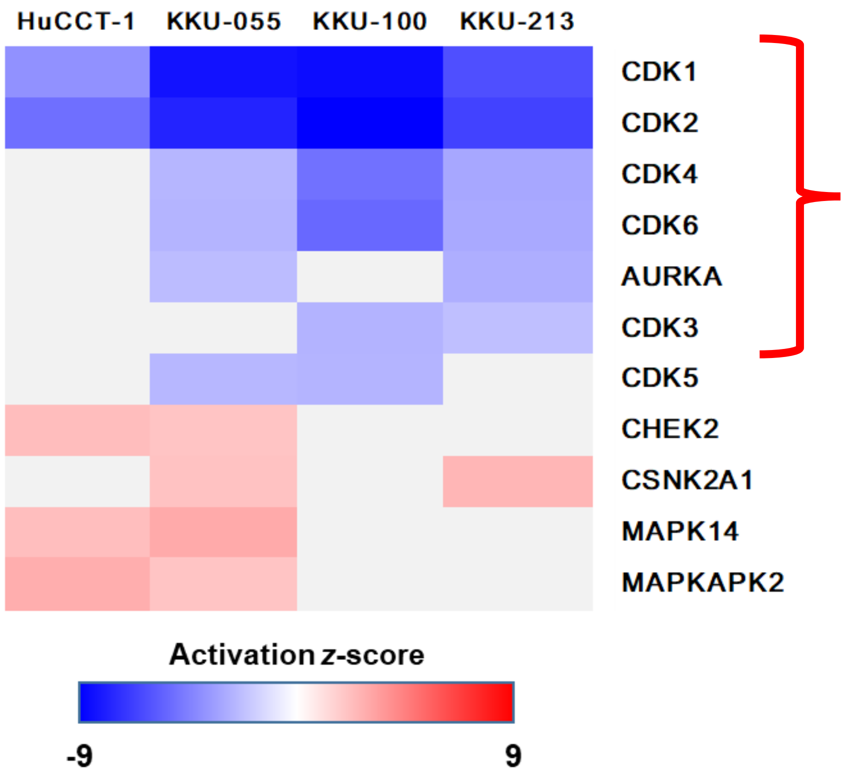
Upstream regulator analysis





Upstream regulator:
Inactivation of cell cycle proteins (CDK1 and CDK2)

Kinase substrate enrichment analysis (KSEA)

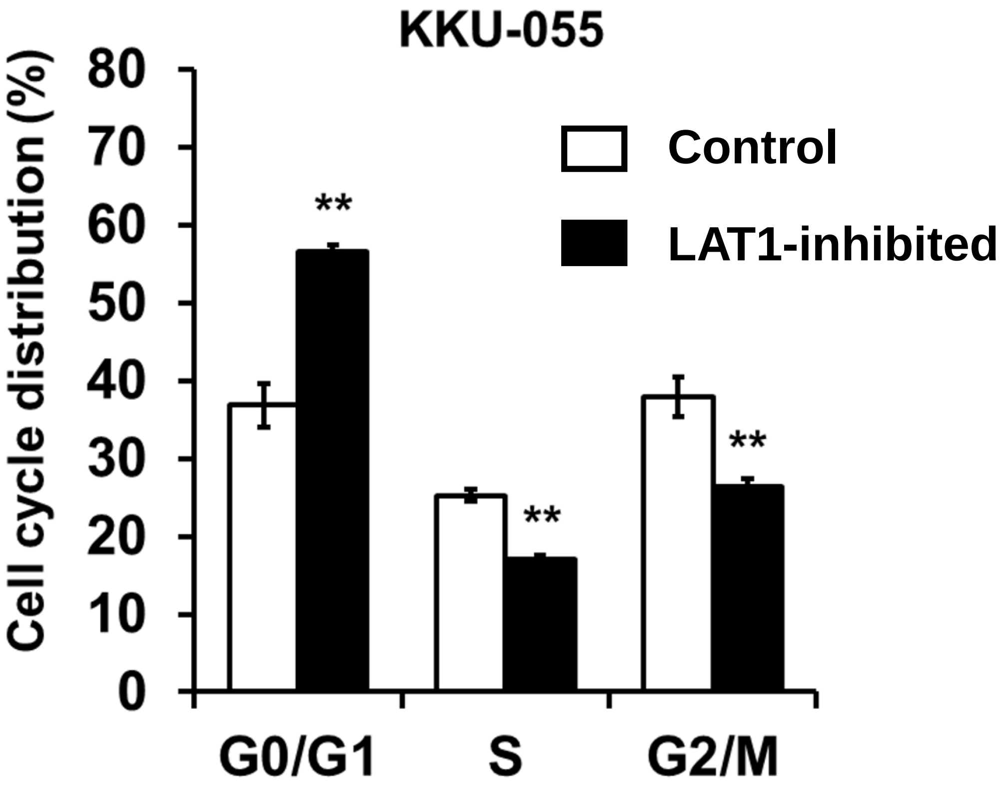


Inactivation of cell cycle-related kinases

CDK1, CDK2, CDK3, CDK4, CDK6
AURKA (aurora kinase A)

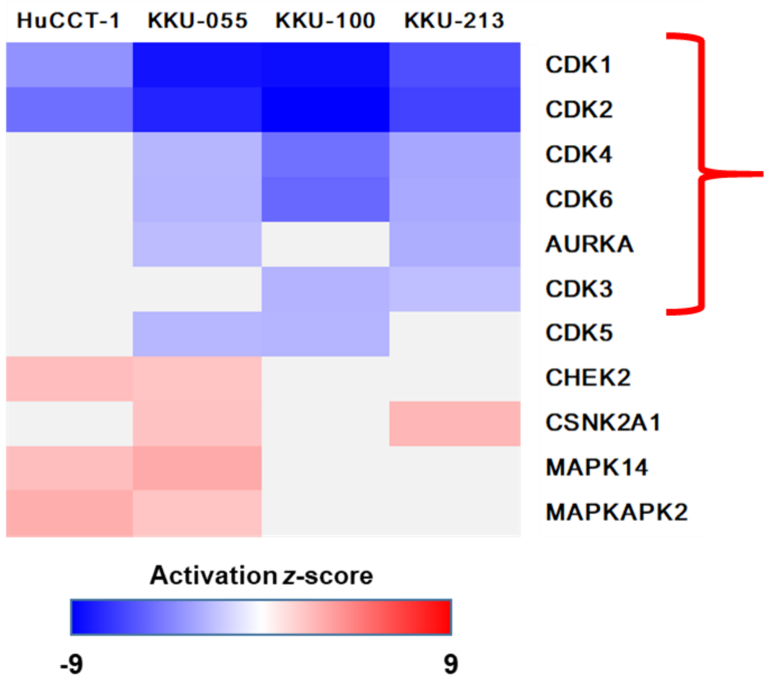
Especially, CDK1 and CDK2.

Cell cycle assay



Increase of G0/G1 phase cells by LAT1 inhibition

Combination of LAT1 inhibitor and CDKs inhibitors



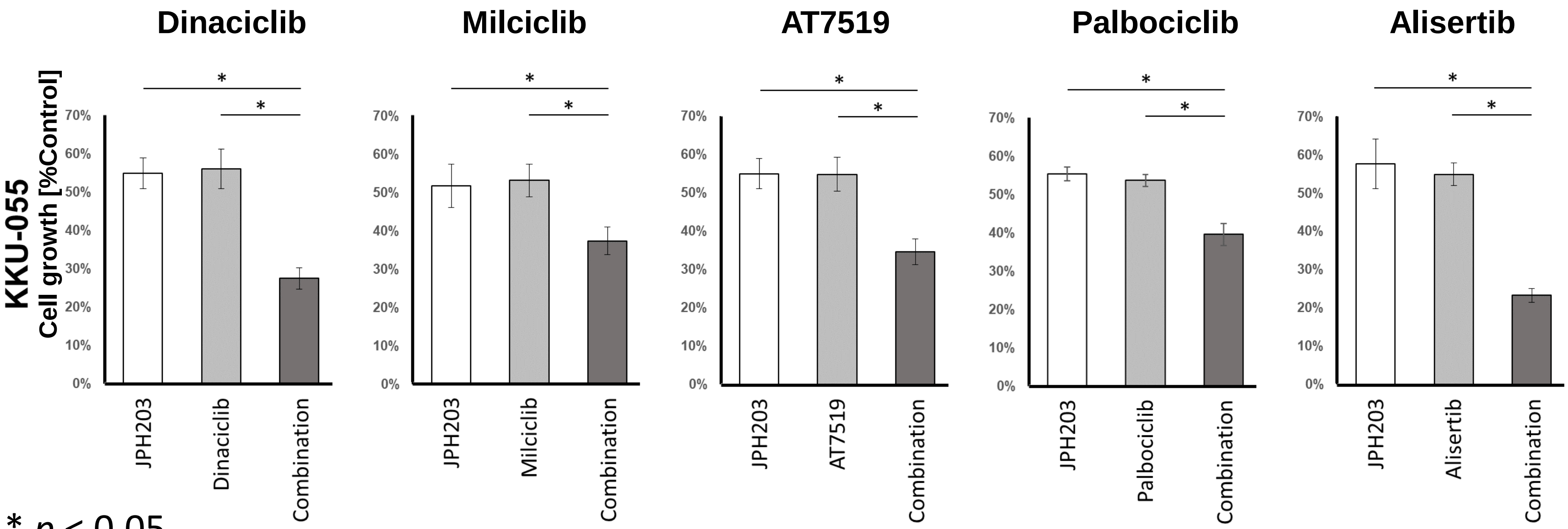
Inactivation of cell cycle-related kinases

CDK1, CDK2, CDK3, CDK4, CDK6
AURKA (aurora kinase A)

Especially, CDK1 and CDK2.

Inhibitors of these CDKs promote anti-cancer effect of LAT1 inhibitor?
(Dinaciclib, Milciclib, AT7519, Palbociclib, …)

Evaluation of drug combinations



* $p < 0.05$

☐ **JPH203 (LAT1 inhibitor)**

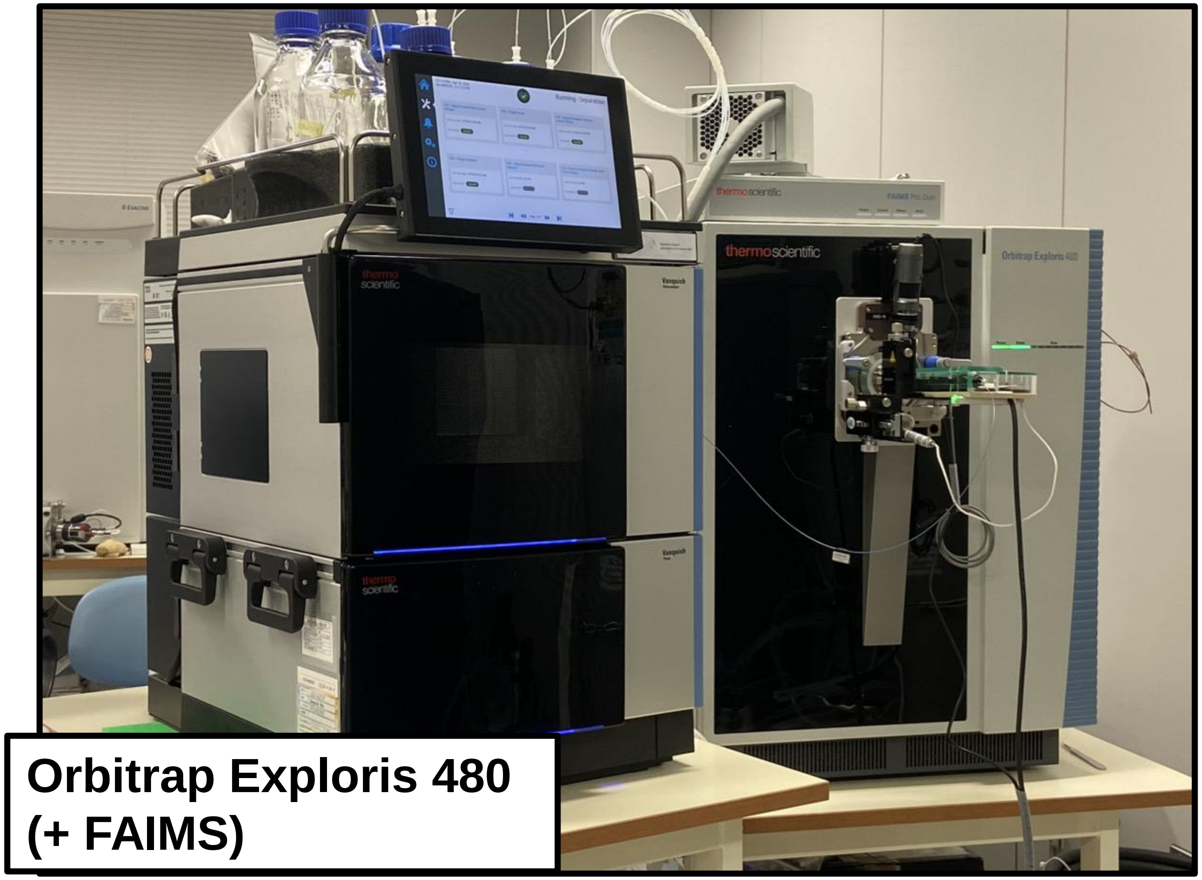
▒ **CDK inhibitor**

▓ **Combination**

Today's topics

1. **Mass spectrometry-based proteomics**
2. **Phosphoproteomics study on cancer cells treated with molecular targeted drug**
3. **LC-MS in CoMIT Omics Center (COC)**

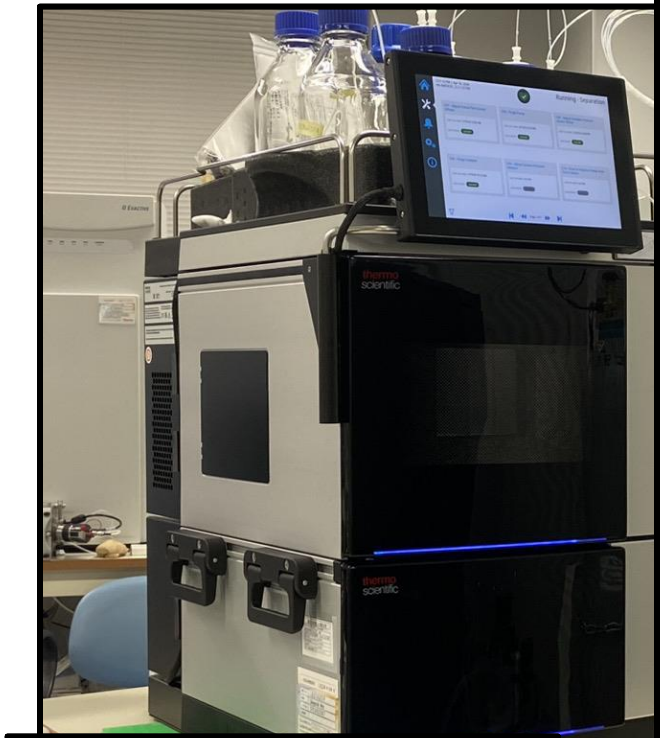
LC-MS in CoMIT Omics Center (COC)



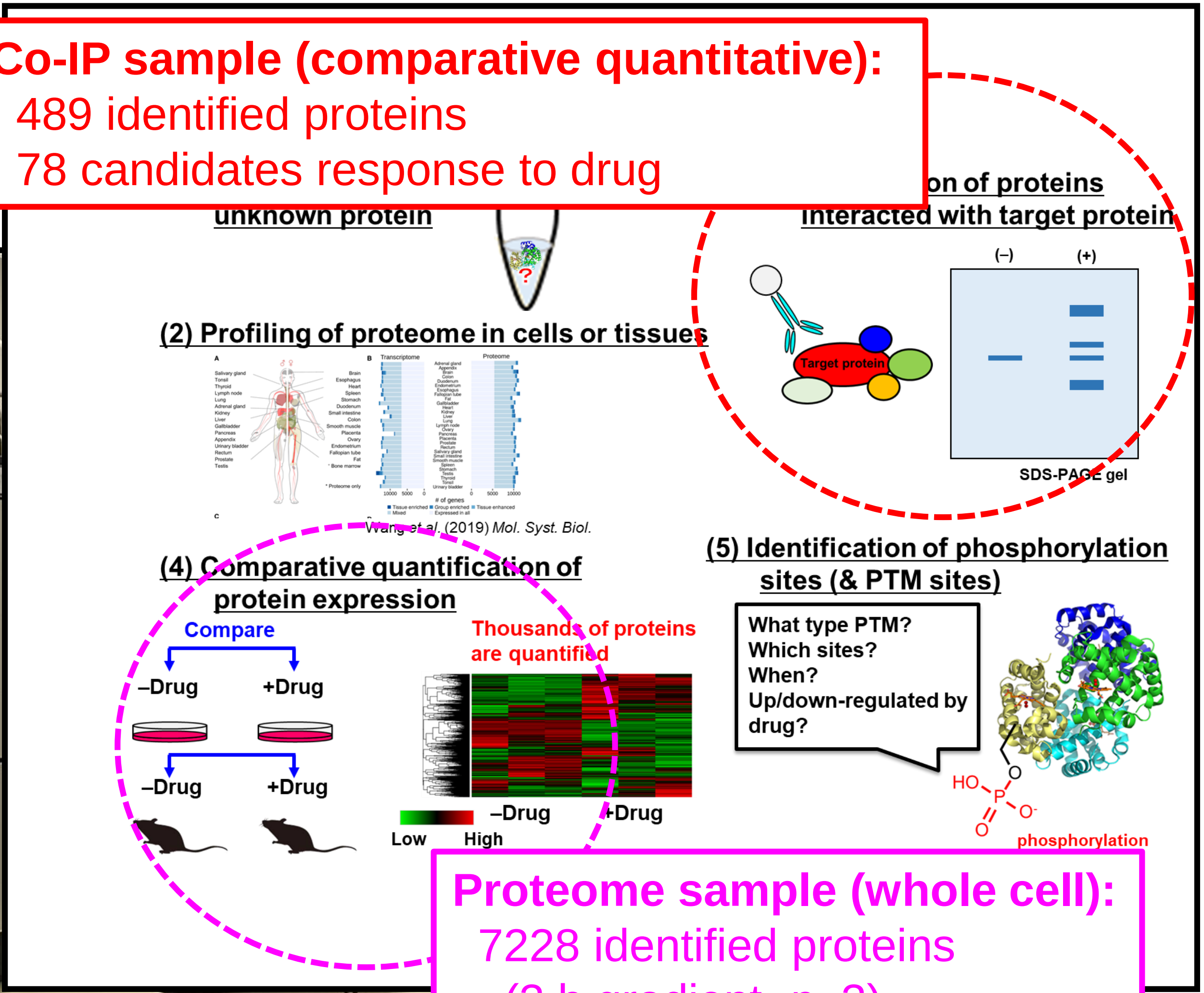
Orbitrap Exploris 480
(+ FAIMS)

Maximum resolution:
48,000
Max scan speed:
>40 Hz
Turbo TMT available
FAIMS Pro available

LC-MS in CoMIT Omics Center (COC)



Orbitrap Exploris 480 (+ FAIMS)



unknown protein

on of proteins interacted with target protein

(2) Profiling of proteome in cells or tissues

Wang *et al.* (2019) *Mol. Syst. Biol.*

(4) Comparative quantification of protein expression

(5) Identification of phosphorylation sites (& PTM sites)

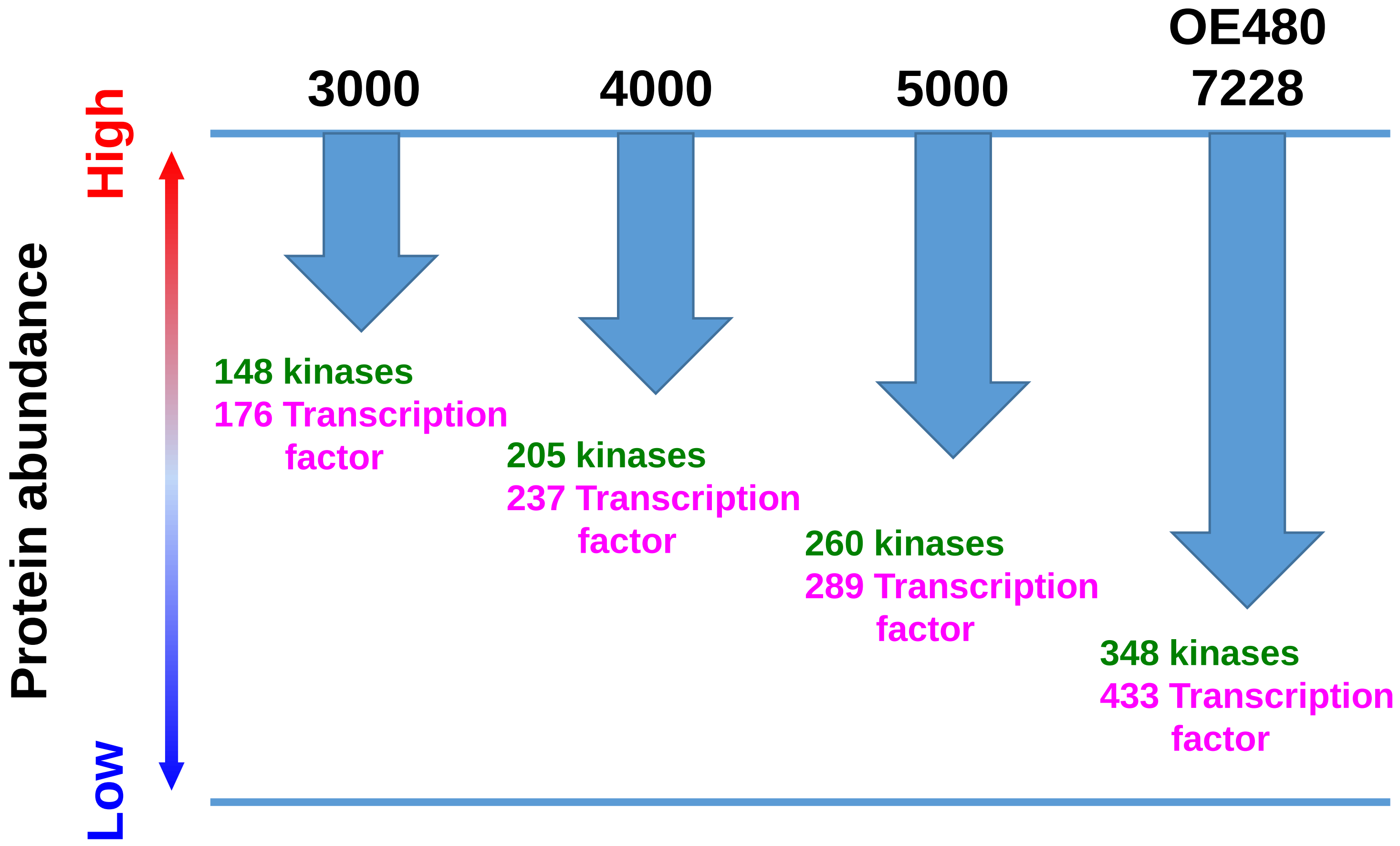
What type PTM?
Which sites?
When?
Up/down-regulated by drug?

Co-IP sample (comparative quantitative):
489 identified proteins
78 candidates response to drug

Proteome sample (whole cell):
7228 identified proteins
(3 h gradient, *n*=3)

LC-MS in CoMIT Omics Center (COC)

No. of identified proteins



OE480

3000

4000

5000

7228

High

Protein abundance

Low

148 kinases
176 Transcription factor

205 kinases
237 Transcription factor

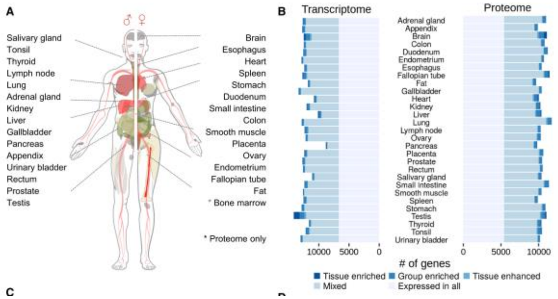
260 kinases
289 Transcription factor

348 kinases
433 Transcription factor

Mass spectrometry-based protein analysis

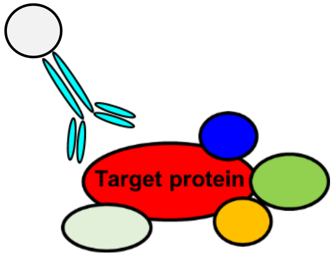
(1) Identification of unknown protein

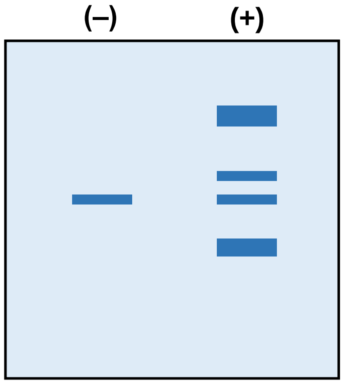
(2) Profiling of proteome in cells or tissues



Wang *et al.* (2019) *Mol. Syst. Biol.*

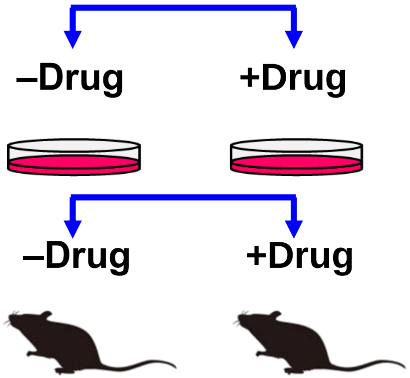
(3) Identification of proteins interacted with target protein





SDS-PAGE gel

(4) Comparative quantification of protein expression

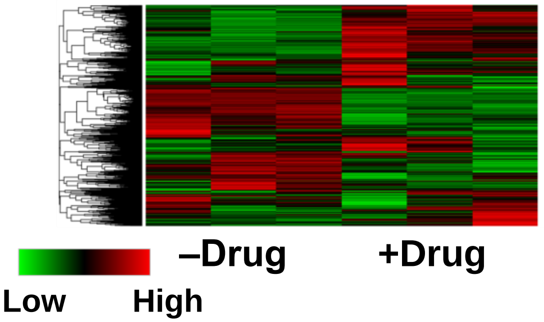


Compare

–Drug **+Drug**

–Drug **+Drug**

Thousands of proteins are quantified



(5) Identification of phosphorylation sites (& PTM sites)

What type PTM?
Which sites?
When?
Up/down-regulated by drug?