细胞生物学中的多维蛋白质组

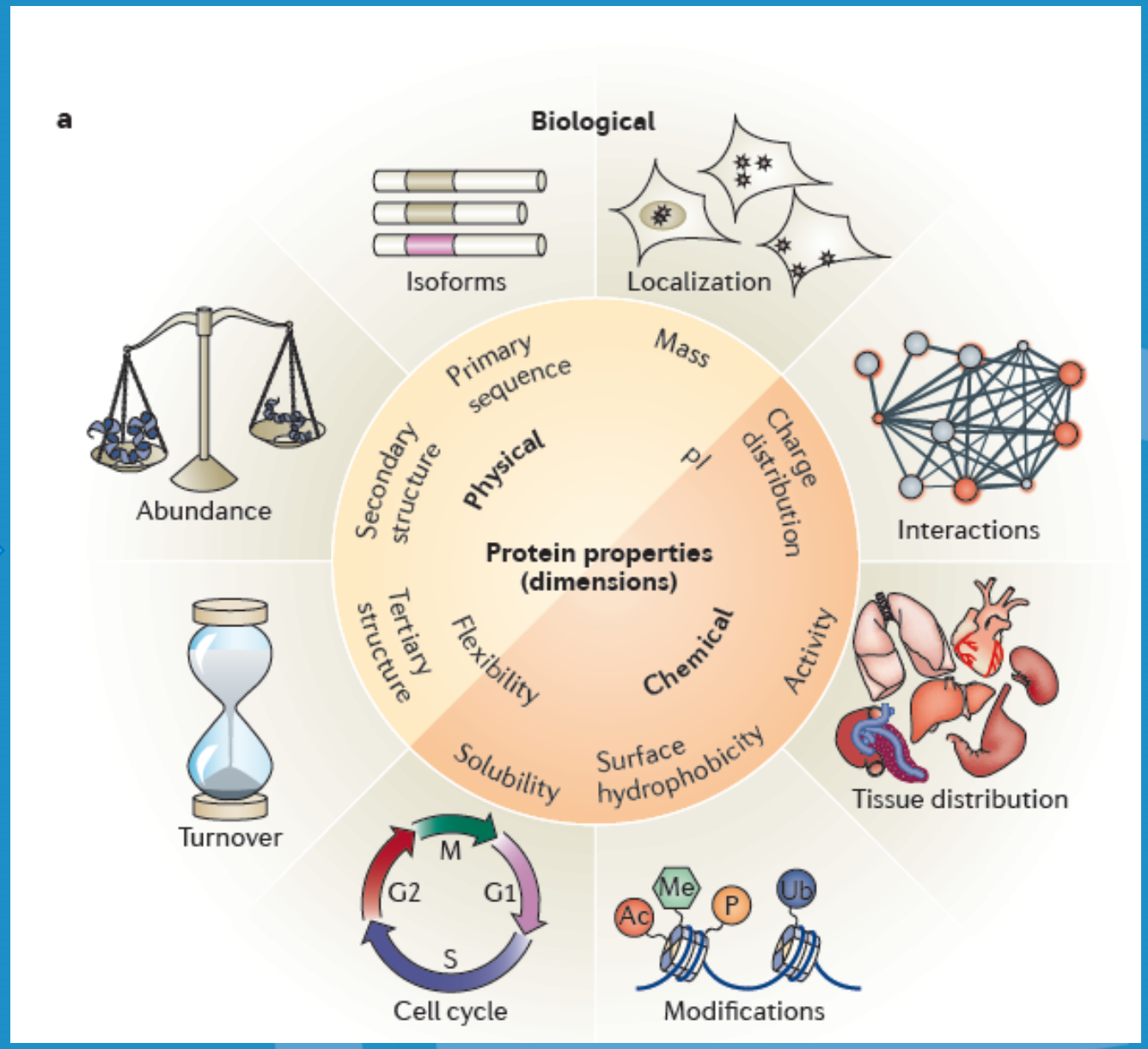
主讲人：姜军欢

翻译文献：何寿鹏
ppt 制 作：陈念航
有关资料的查询：何，姜，陈

Time：2016-3-31

蛋白组是个动态的系统，蛋白质之间的关联（维度），形成了细胞的表型

由于基于光谱测定法的发展，发现了蛋白质的不同性质。

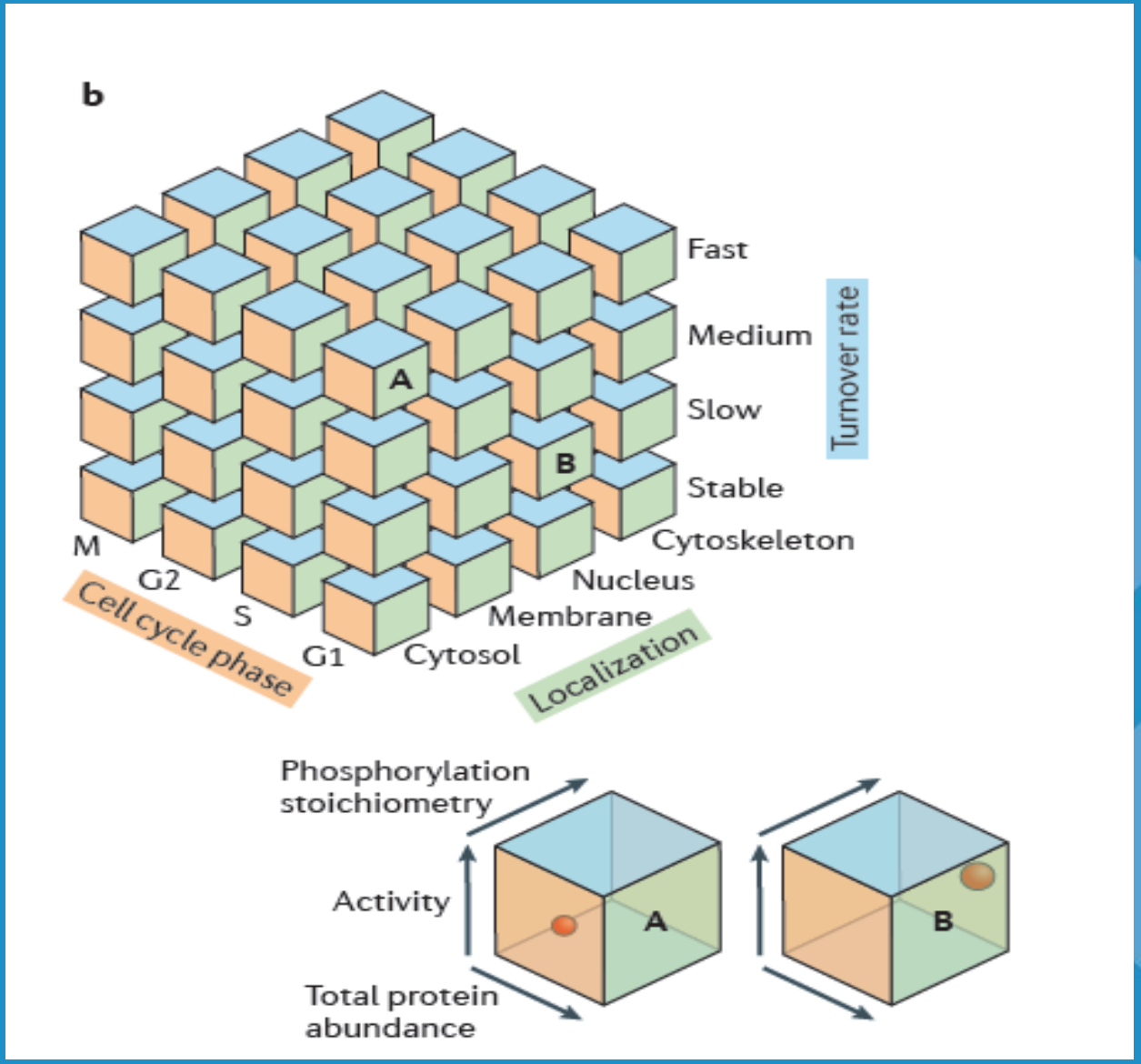


单维度的蛋白质分析会导致生物信息分析不准确，因此我们用到Spectrometry-based(基于光谱测定法)。

该技术结合改进的工具、分析方法、实验技术注释蛋白质的多维信息（例如蛋白质降解、蛋白质的合成与转化率）

通过以测量的信息，我们还能得到有关蛋白质的某些属性：如亚细胞定位、蛋白质丰富度、组织分布，以及蛋白质亚型或者突变型。

不同的亚细胞内或者细胞周期阶段



样本相对量化（丰富度）

- SILAC：用细胞培养中稳定同位素标记氨基酸的方法该方法是在细胞培养的过程中加入稳定同位素标记的必需氨基酸如赖氨酸(K) 和精氨酸(R)，使每条肽链相差质量数恒定。
- TMT:串联质量标签
- iTRAQ：等重同位素标记相对与绝对定量方法

iTRAQ和TMT是最常用的基于报告离子的蛋白质组相对定量方法，通过方法改进或者与其他方法结合，可实现更多重的标记，进而提高蛋白质组相对定量的通量。

相对定量是研究不同情况下同一蛋白质组样品组成含量的变化，如细胞受到刺激前后蛋白质表达的差异。

Table 1 | **Analysing the dimensions of the proteome**

Dimension	Examples of techniques used	Refs
Abundance (absolute and relative)	Label-free quantitation	5,16–18
	SILAC	19
	^{15}N-labelling	20
	NeuCode SILAC	21
	Dimethyl-labelling	22,23
	TMT	24
	iTRAQ	25
Cell cycle regulation	Centrifugal elutriation	124
	Chemical inhibitors of cell cycle regulators	125
	FACS (for DNA content or phase-specific markers)	126
Tissue distribution	Dissection	95,127
	FACS (for cell-type-specific markers)	126
Interactions	Affinity-enrichment (endogenous immuno-precipitation or tagged fusion protein pull-down)	63–67
	Protein correlation profiling	9,70,71
	Proximity-labelling	39,68
Post-translational modifications	Affinity enrichment: TiO_2	128,129
	Affinity enrichment: IMAC	128,130
	Modification-specific antibodies	90,131–133
	Chromatography: IEX	87
	Chromatography: HILIC	94
	Chromatography: ERLIC	134
Localization	Centrifugation	3,43,135
	Protein correlation profiling	38,44
	Proximity-labelling	39
	Detergent solubility	4
Turnover	Metabolic pulse-labelling	3,5,6,55
	Cycloheximide treatment	4
Isoform expression	• High sequence coverage to identify isoform-specific peptides • Targeted mass spectrometry analysis may be used to detect isoform-specific peptides	136,137
Solubility	Thermal denaturation followed by differential centrifugation	138
Activity	Analogue-sensitive kinases	139
	Activity-dependent binding domains	140
Tertiary Structure	Protease sensitivity	141
	Crosslinking	77,78

结合色谱法和离心分离可以更容易地组合多个分离的多维蛋白质

生物方面

亚细胞定位

融合报告基因的方法

- 亚细胞定位对于蛋白质功能是非常重要的!蛋白质必须处于合适的亚细胞定位才能行使其功能
- GFP 能自我催化形成发色结构并在蓝光激发下发出绿色荧光,所以可以与目标蛋白融合, 作为荧光标记分子, 特异性地进行蛋白质的亚细胞定位。
- GFP 能在蛋白质的N端或C 端融合而保持其天然蛋白的特性,而且灵敏度高、对活细胞无毒害作用,
- GUS 基因克隆后, 很快就发展起以GUS作为基因标记的系统 。GUS 基因表达产物具有检测容易、灵敏度高、易于定量及定性分析的优点。由于在绝大多数植物的细胞内不存在内源的GUS活性, 所以检测弱启动子驱动的GUS 活性则更容易、更精确.

蛋白质组学定位技术

1.细胞器蛋白质组学：主要用双向电泳进行蛋白质的分离，然后用质谱技术鉴定蛋白质，如能预先对细胞组分进行筛分，则可以了解不同细胞组分中的蛋白质成分及丰度。

2.用细胞器特异的的实验方法能够鉴定特定位置的蛋白质，这样就可以建立已知或未知蛋白质定位的目录。这种方法通常采用分离细胞组分、离心纯化细胞器或者细胞组分，然后用质谱技术鉴定多肽。

蛋白质周转

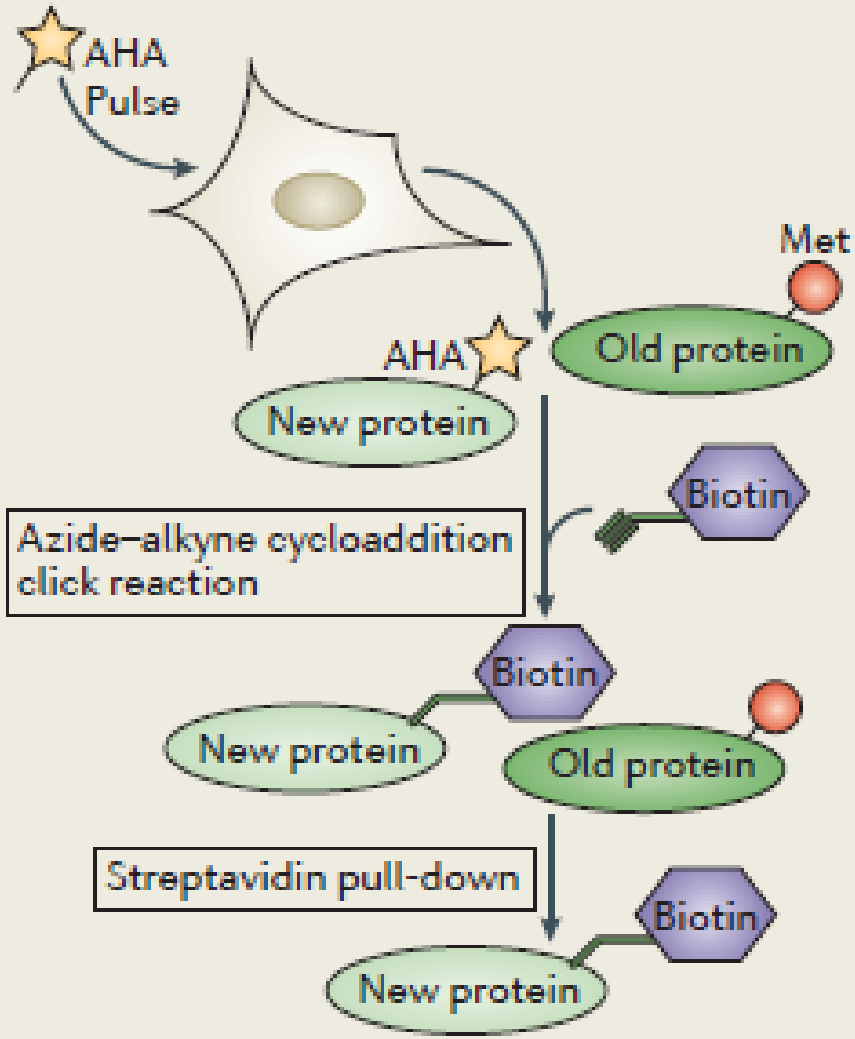
——蛋白质合成与降解的结合

蛋白质周转是蛋白质合成与降解的双向调节过程，这个循环中合成与降解的互相协调对维持细胞内酶和结构蛋白的稳定状态、细胞内环境的相对稳定及调节蛋白质在组织中的沉积都十分重要。

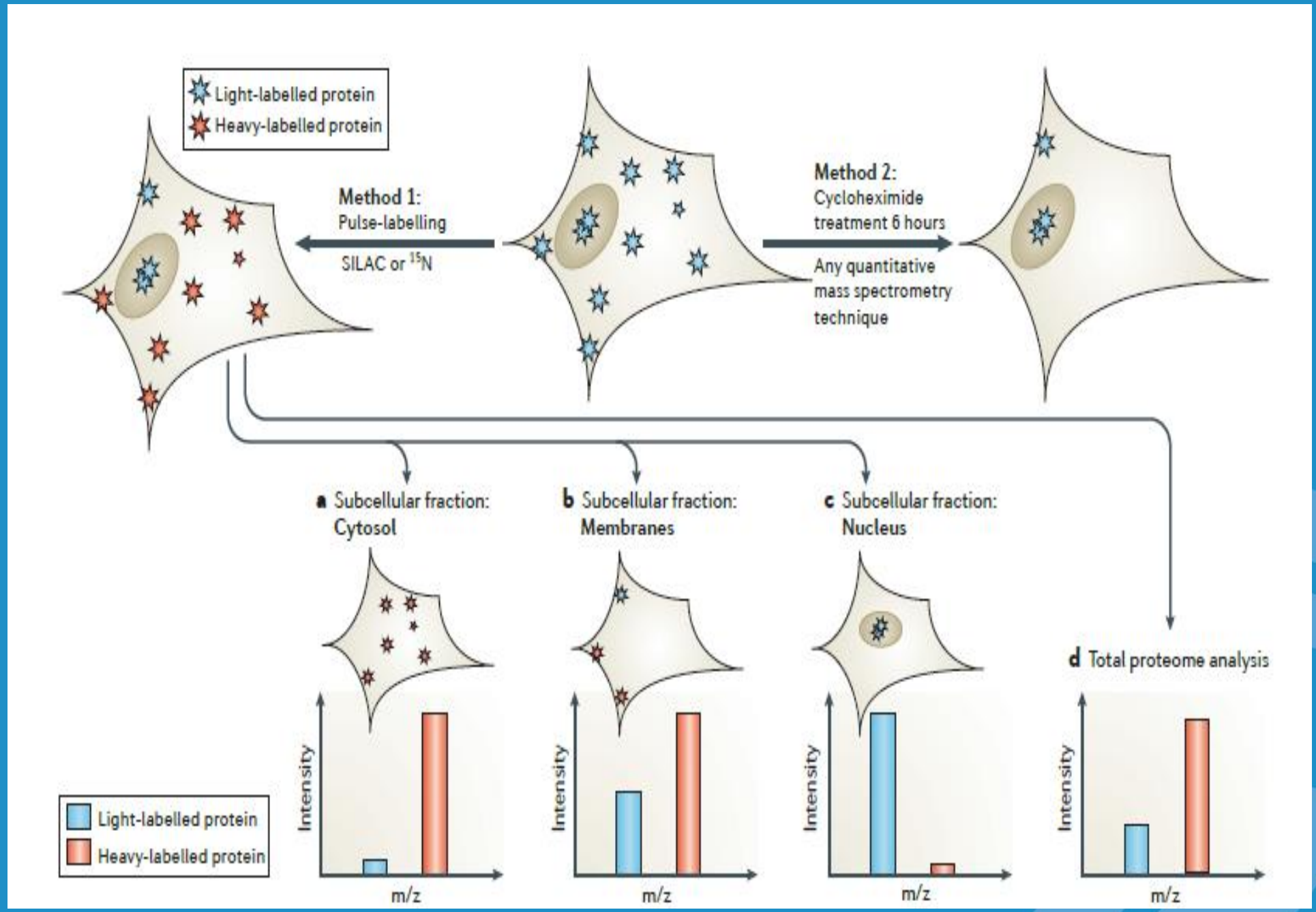
在满足机体能量需要的前提下，增加蛋白质水平则明显增加蛋白质周转代谢速率，使代谢表现出高周转和沉积的特点。

Box 1 | **AHA and click chemistry**

Feeding azido-homoalanine (AHA) to cells in short pulses results in the incorporation of this amino acid into proteins, replacing Met, and thus enables the labelling of newly synthesized proteins (see the figure). The presence of the azide group in the side chain of this amino acid facilitates the covalent modification of AHA-containing proteins *in vitro* with affinity reagents that contain groups such as alkynes through cycloaddition reactions (click reactions). Commonly, an affinity group such as biotin is covalently attached to the AHA, resulting in tagged proteins that can subsequently be enriched by streptavidin-mediated pull-down.



蛋白质周转分析的方法

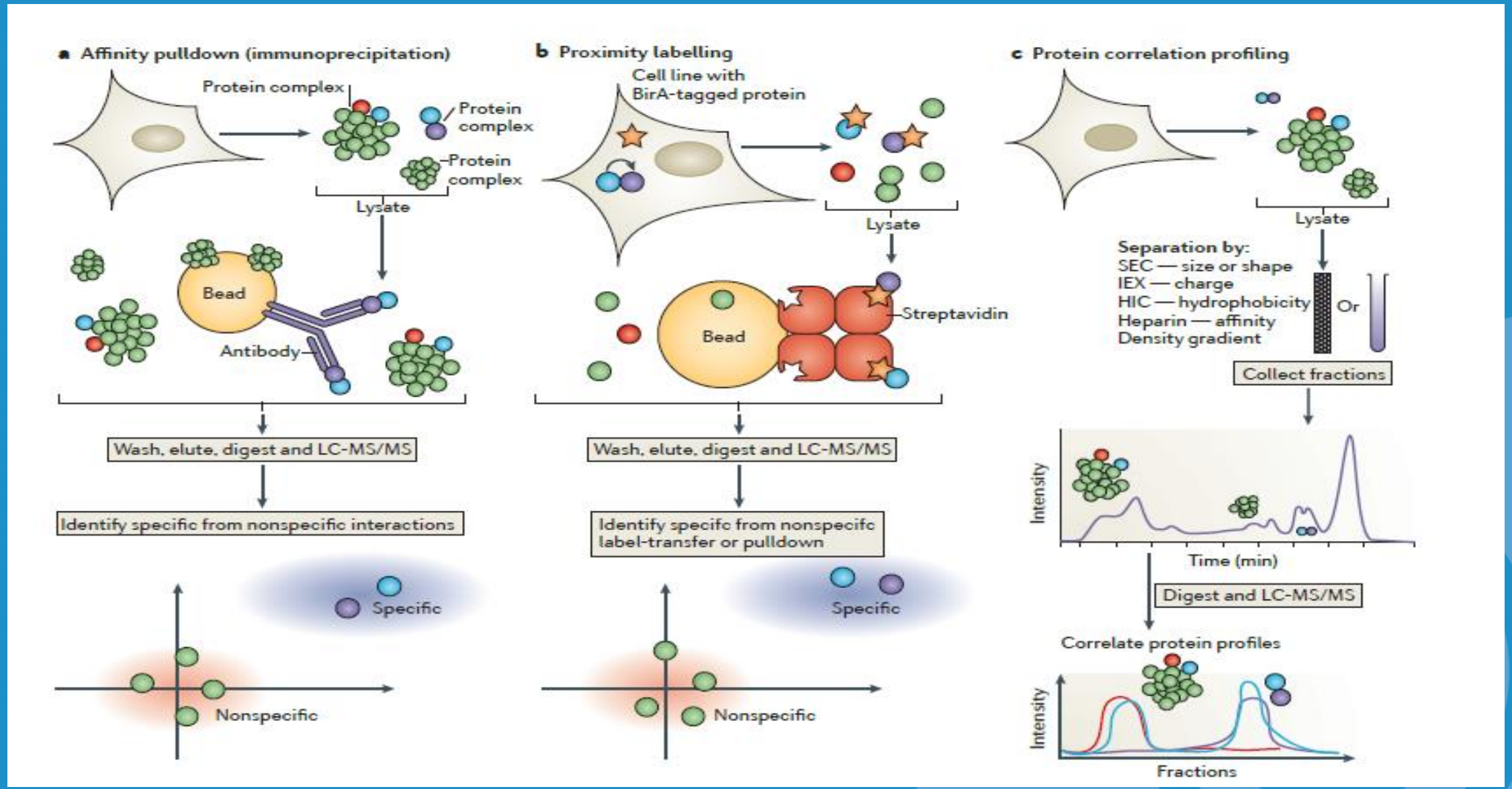

Light-labelled protein
Heavy-labelled protein
Method 1:
Pulse-labelling
SILAC or ^{15}N
Method 2:
Cycloheximide
treatment 6 hours
Any quantitative
mass spectrometry
technique
a Subcellular fraction:
Cytosol
b Subcellular fraction:
Membranes
c Subcellular fraction:
Nucleus
d Total proteome analysis
Intensity
m/z
Light-labelled protein
Heavy-labelled protein

蛋白质的互作

功能相关或组织相关的蛋白质在一个三维的空间里，通过动态的相互作用形成一个有秩序、有功能的整体，这些蛋白质互作随着时间的演变形成一个个四维的蛋白质互作网络(protein-proteininteraction network，PIN)。所有的蛋白质互作构成蛋白质互作组或互作蛋白质组(protein-protein interactome)

研究方法

1. pull-down (immunoprecipitation) :免疫沉淀反应，一种特定的蛋白质及其结合物通过内源抗体标记，形成抗原抗体复合物。
2. proximity labelling：生物识别（意译）基于异位表达的蛋白质融合到细菌的生物素连接酶或过氧化物酶中，生物素迅速共价结合赖氨酸或酪氨酸残基。
3. Protein correlation profiling:蛋白的相关性分析,根据蛋白质复合物的大小、密度、形状、电荷使用色谱或密度梯度离心法分离，分析液相色谱串联质谱(LC - MS / MS)鉴定其组成蛋白质。


a Affinity pulldown (immunoprecipitation)
Protein complex
Protein complex
Protein complex
Lysate
Bead
Antibody
Wash, elute, digest and LC-MS/MS
Identify specific from nonspecific interactions
Specific
Nonspecific
b Proximity labelling
Cell line with BirA-tagged protein
Lysate
Streptavidin
Bead
Wash, elute, digest and LC-MS/MS
Identify specifc from nonspecifc label-transfer or pulldown
Specific
Nonspecific
c Protein correlation profiling
Lysate
Separation by:
SEC — size or shape
IEX — charge
HIC — hydrophobicity
Heparin — affinity
Density gradient
Or
Collect fractions
Intensity
Time (min)
Digest and LC-MS/MS
Correlate protein profiles
Intensity
Fractions

蛋白质翻译后的修饰（protein PTMs）

蛋白质翻译后修饰是蛋白质水平发生的一种重要的生物化学过程，可为蛋白质的特定氨基酸位点引入各种化学基团，例如磷酸根、糖基、甲基、乙酰基和泛素链等。

以蛋白质磷酸化为例，使底物蛋白质发生磷酸化修饰的酶是激酶，而使磷酸化蛋白质发生去磷酸化的酶是磷酸酶激酶和磷酸酶常有多种底物蛋白质，而每种底物蛋白质往往有多个磷酸化位点。这些被磷酸化的氨基酸序列常会特异性地被蛋白结构域所识别(例如SH2、PTB、BRCT 结构域等)，从而形成基于磷酸化的蛋白质间相互作用。

近年来，基于液相色谱与生物质谱联用技术(LC-MS)的发现，，为系统水平上的蛋白质翻译后修饰研究提供了强有力的研究工具。

离子交换色谱分离

概念：离子交换色谱用于多肽分离主要是基于磷酸化多肽与固定相表面的强阳离子交换基团或强阴离子交换基团(间的静电相互作用实现分离。另外，在高乙腈浓度流动相中，众多离子交换色谱填料可以展示出很好的亲水相互作用色谱(HILIC) 性质，从而可以基于亲水相互作用分离酶解多肽。

主要原理：在酸性条件下，磷酸化多肽的N 端氨基和C 端的赖氨酸或精氨酸残基都被质子化，磷酸基团的存在会降低多肽的价态。因此，磷酸化肽在阳离子交换基团色谱填料上的保留较常规多肽更弱，从而实现了磷酸化多肽的有效富集和分离。

数据分析与共享

高通量DNA,RNA测序的数据带来基因组，转录组数据的快速增长，目前蛋白质组研究数量的增长需要细胞生物学界来处理,分析和分享这些大型蛋白质组学数据集。

已被广泛应用于细胞生物学领域的软件有MaxQuant,Skyline,COMPASS,Census

Table 2 | **Data handling and sharing resources**

Resource	Key features	Strengths	Comments	Refs
Protein dimension annotation resources				
Encyclopedia of Proteome Dynamics	Graphical display of protein dimension data and diverse dimensions analysed	Diverse dimensions analysed, including multidimensional data sets	Single laboratory as source of data	4
Human Protein Atlas	Proteomic annotation of protein abundance and localization	High sensitivity, and localization is analysed within many tissues	Antibody-based, single laboratory as source of data and human data only	122
Human Proteome Map	Human proteome data annotated for tissue abundance	Whole pathway or protein family analysis	One protein dimension and human data only	7
The MaxQuant Database	Proteome data tabulated and graph for label-free abundance measurement	Quality control parameters for mass spectrometry acquisition are presented	Single laboratory as source of data	142
The Multi-Omics Profiling Expression Database	Enables search and visualizations of protein data derived from multiple species	Many different experiments can be visualized and compared	Chromosome-centric	143
The proteins across organisms database	Absolute protein abundance values determined across many organisms and tissues	Abundance histogram and STRING integration	Basic user interface	144
Phosphorylation site database	Provides data on phosphorylation, acetylation and *N*-glycosylation of proteins, including EGF-treated, cell cycle regulated, kinome-related data sets	Diverse dimensions analysed with high depth of coverage for phosphorylation data sets	Single laboratory as source of data	145
ProteomicsDB	Human proteome data annotated for tissue abundance and mass spectrometry spectra shown	Tissue protein abundance pattern and mass spectrometry spectral annotation and multi-protein analysis	Human data only	8
Mass spectrometry-based raw proteomics data repositories				
Chorus	Offers storage, search and visualization of mass spectrometry-based proteomics data files	Well-developed search and mass spectrometry data file visualization	Limited public mass spectrometry data included	–
Global Proteome Machine Database	Enables search and visualization of mass spectrometry data derived from many species	Rich graphical interface for mass spectrometry data visualization	MS2 spectral validation emphasized	146, 147
ProteomeXchange Consortium Includes the Proteomics Identifications, PeptideAtlas, PeptideAtlas SRM Experiment Library and Mass spectrometry	Enables centralized submission of mass spectrometry raw data and associated files for shotgun and targeted mass spectrometry analyses	New interface for submission and download of data; managed by the EBI	Requires visiting consortium member sites for the visualization of mass spectrometry data files	148

检索工具从蛋白质数据库中为蛋白质基因互作的数据库提供了大量信息

蛋白质组学的数据可以在细胞生物学界共享，这就意味着我们可以很容易的搜索其资源，在EBI可以找到蛋白质组的原始质谱文件 PRIDE ，虽然很方便，但是不利于作大批量蛋白质组的交互。

如今需要一致的、可搜索的格式，一致的实验，蛋白质组学数据的分析工具，实验室信息管理系统(LIMS)来收集，管理，使用大量的蛋白组数据。