Kód předmětu: Bi8980

MASARYKOVA UNIVERZITA

Protein expression and purification

- III. The important properties of proteins and how to explore them

Lubomír Janda, Blanka Pekárová and Radka Dopitová

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

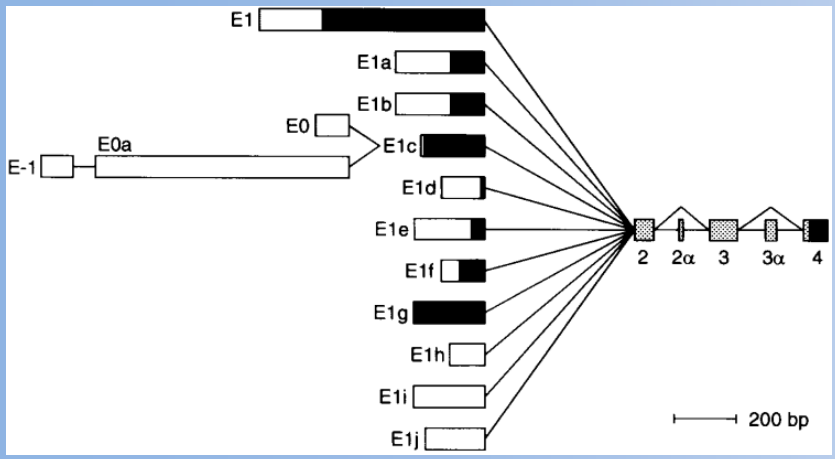
3.1. Introduction: the context for studies and data analysis

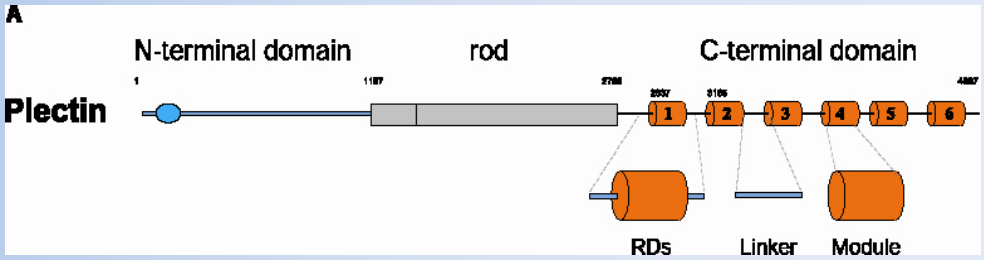
Key concept: • Appreciating the complexity of cellular systems in terms of the numbers of distinct protein species present

Yeast *Saccharomyces cerevisiae*	6,000 genes
Nematode worms *Caenorhabditis elegans*	12,000 genes
Vertebrate *Homo sapiens*	23,000 genes

Probably over 10 times as many distinct proteins:

- differential RNA processing
- post-translational modification





3.2. The key questions about a protein

Key concepts:

- Being aware of the range of functions of proteins
- Understanding the levels of protein structure
- Identifying appropriate methods to explore protein structure and protein interactions

Proteins are composed of amino acids which are linked by peptide bonds; however, it is the **sequence** of amino acids within a given protein that dictates its unique function and structure.

- Biological systems are not static entities; they respond to:
  - Environmental signals
  - Developmental signals
  - Metabolic signals
- Proteins almost always occur within complex cellular enenvironments, interacting with other proteins, metabolites and cellular structures.

3.2. The key questions about a protein

3.2.1. What are the functions of proteins?

I. **1. Catalytic proteins**
II. **2. Transport proteins**
3. Signalling proteins
III. **4. Structural proteins**
5. Motor proteins
6. Binding proteins
IV. **7. Storage proteins**

I. Enzymes
II. Proteins binding small molecules
III. Proteins binding macromolecules (protein-protein or protein-DNA interaction)
IV. Storage proteins

3.2. The key questions about a protein

3.2.1. What are the functions of proteins?

3.2.1.1. Protein classification

Enzymes are classified on the basis of their composition (60% of all proteins are enzymes). Activity of enzymes is also affected by changes in **pH**, **temperature** and **substrate concentration**.

Simple enzymes	enzymes composed entirely of proteins
Complex enzymes (**Holoenzymes**)	protein + small organic molecule (ligand)
Apoenzymes	protein component
Coenzyme	non-protein component (**not covalently bound**)
Prosthetic group	small organic molecule (**covalently bound**)

3.2. The key questions about a protein.

3.2.1. What are the functions of proteins?

3.2.1.2. Enzyme classification – EC number

Number	Classification	Biochemical properties
1.	Oxidoreductases	Act on many chemical groupings to add or remove electron or oxygen
2.	Transferases	Transfer functional groups between donor and acceptor molecules
3.	Hydrolases	Add water across bond, hydrolyzing it
4.	Lyases	Add water, ammonia or carbon dioxide across double bonds, or remove these elements to produce double bonds
5.	Isomerases	Carry out many kinds of isomerization (L/D), mutase reaction
6.	Ligases	Catalyze reactions in which two chemical group are joined with use of energy from ATP

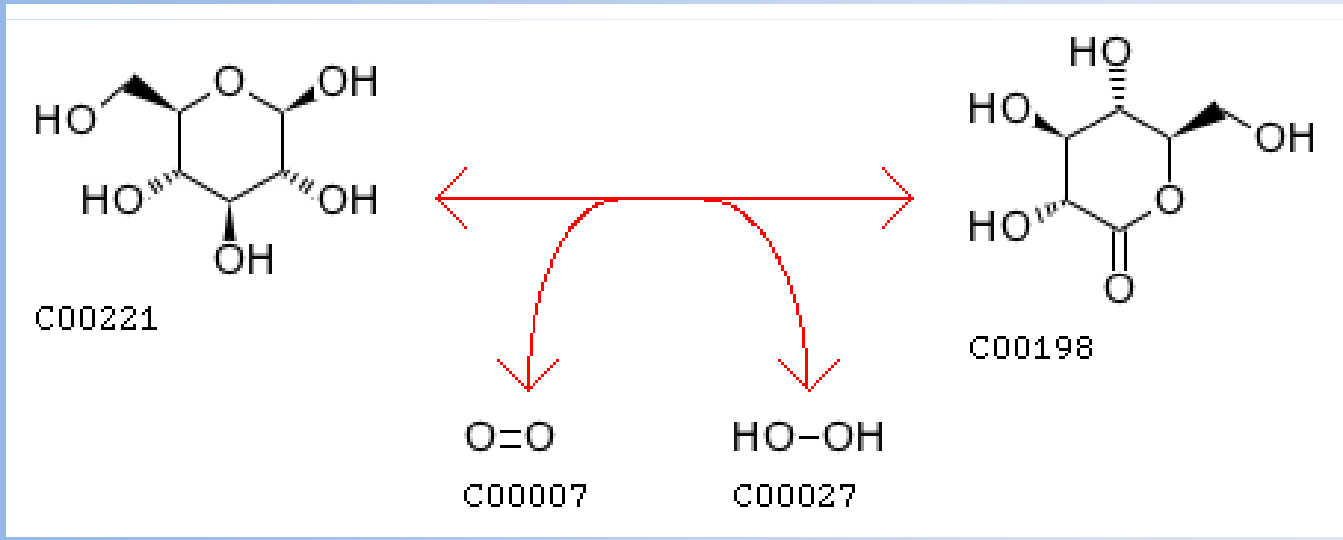
3.2.1. What are the functions of proteins?

3.2.1.2. Enzyme classification – EC number

3.2.1.2.1. Oxidoreductases

They are involved in redox reactions in which hydrogen or oxygen atoms or electrons are transferred between molecules.

Dehydrogenases	hydride transfer
Oxidase	electron transfer
Oxygenase	oxygen transfer
Peroxidase	electron transfer to peroxide



EC 1.1.3.4 beta-D-glucose oxidase

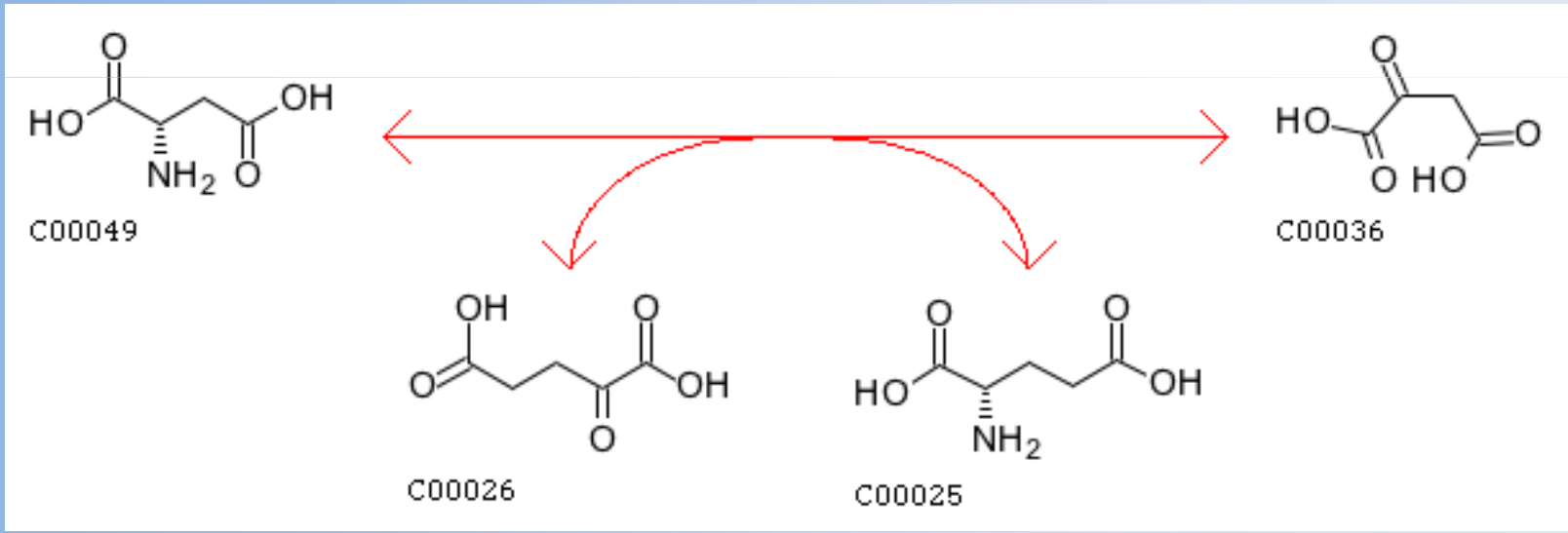
beta-D-Glucose + Oxygen ⇌ D-Glucono-1,5-lactone + H_2O_2

3.2.1. What are the functions of proteins?

3.2.1.2. Enzyme classification – EC number

3.2.1.2.2. Transferases

They catalyse the transfer of an atom or group of atoms (e.g. Acyl-, alkyl-, and glucosyl-) between two molecules, but excluding such transfers as are classified in the other groups.



2.6.1.1 aspartate aminotransferase

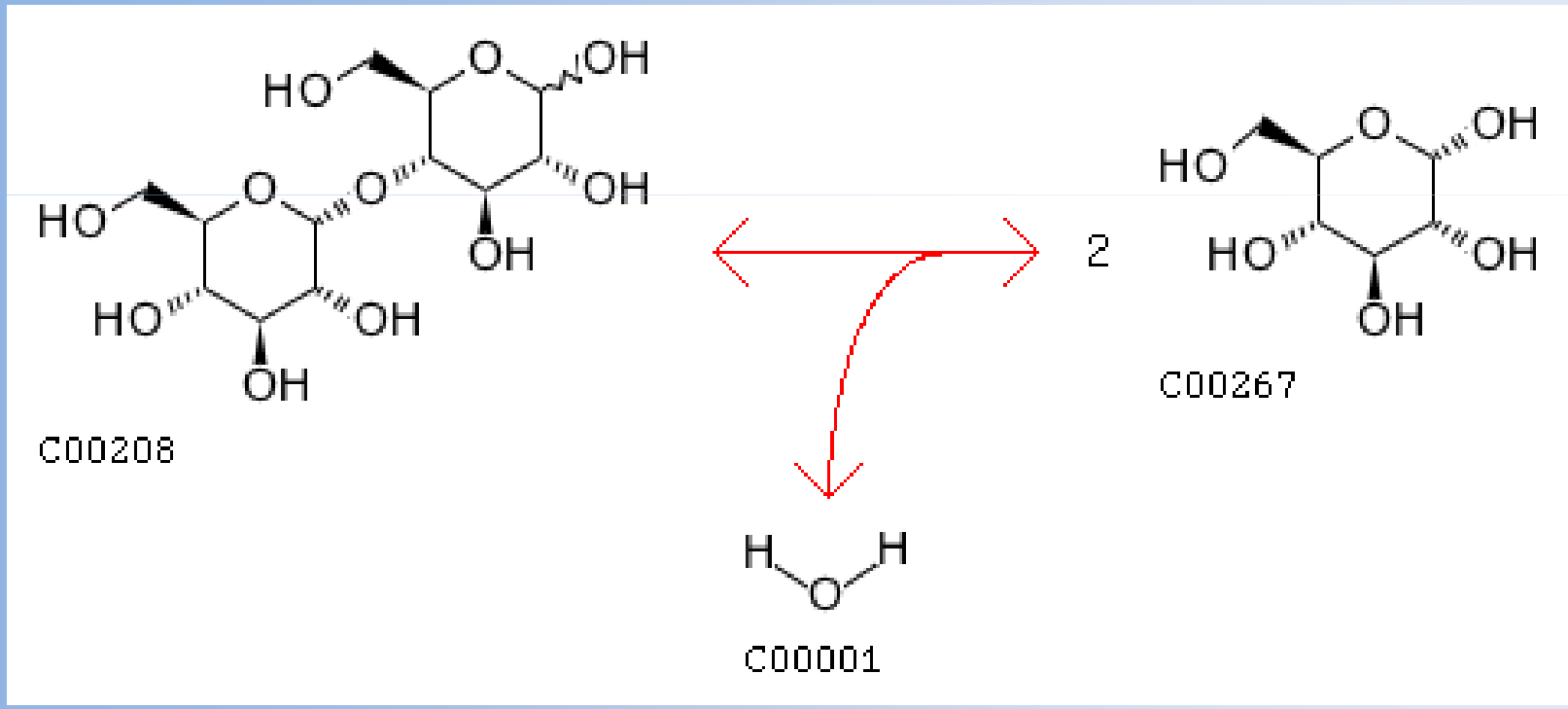
L-Aspartate + 2-Oxoglutarate ⇌ Oxaloacetate + L-Glutamate

3.2.1. What are the functions of proteins?

3.2.1.2. Enzyme classification – EC number

3.2.1.2.3. Hydrolases

They are involved in hydrolytic reactions and their reversal. This group of enzymes includes esterases, glycosidases, lipases and proteases.



EC 3.2.1.20 alpha-1,4-glucosidase

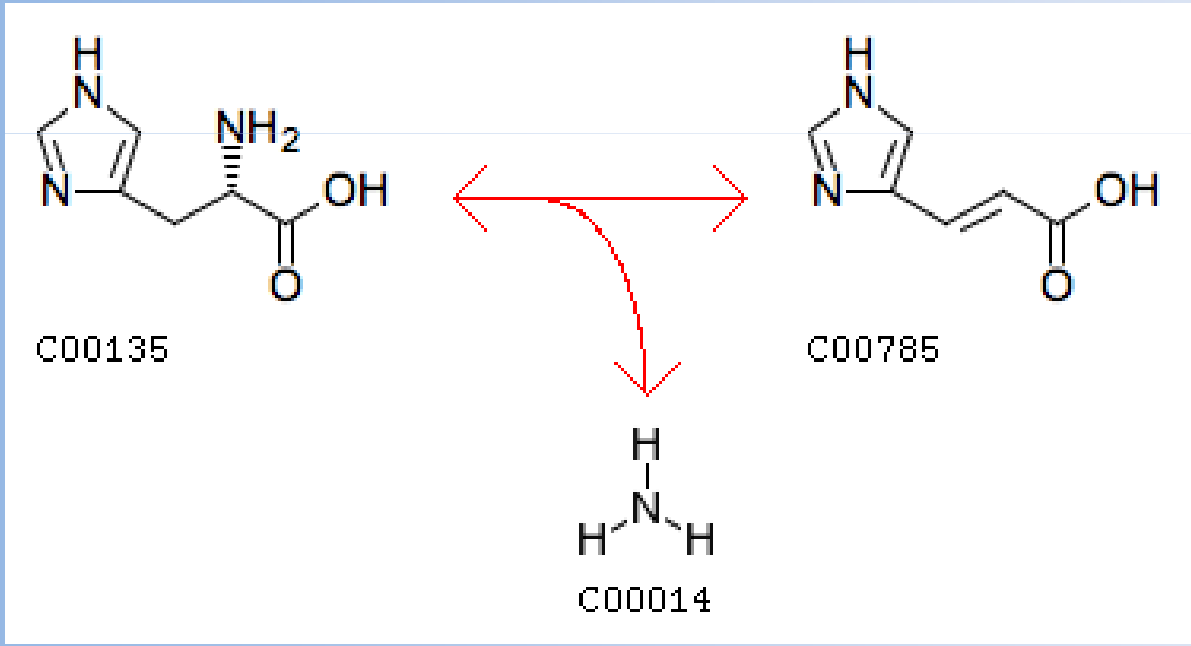
Maltose + H_2O ⇌ 2 alpha-D-Glucose

3.2.1. What are the functions of proteins?

3.2.1.2. Enzyme classification – EC number

3.2.1.2.4. Lyases

They are involved in elimination reactions in which a group of atoms is removed from the substrate. Lyases include aldolases, decarboxylases, dehydratases and some pectinases.



EC 4.3.1.3. L-histidine ammonia-lyase

L-Histidine ⇌ Urocanate + NH_3

3.2.1. What are the functions of proteins?

3.2.1.2. Enzyme classification – EC number

3.2.1.2.5. Isomerases

They catalyse molecular isomerizations and include epimerases, racemases and intramolecular transferases.

EC 5.3.1.5 D-glucose aldose-ketose-isomerase

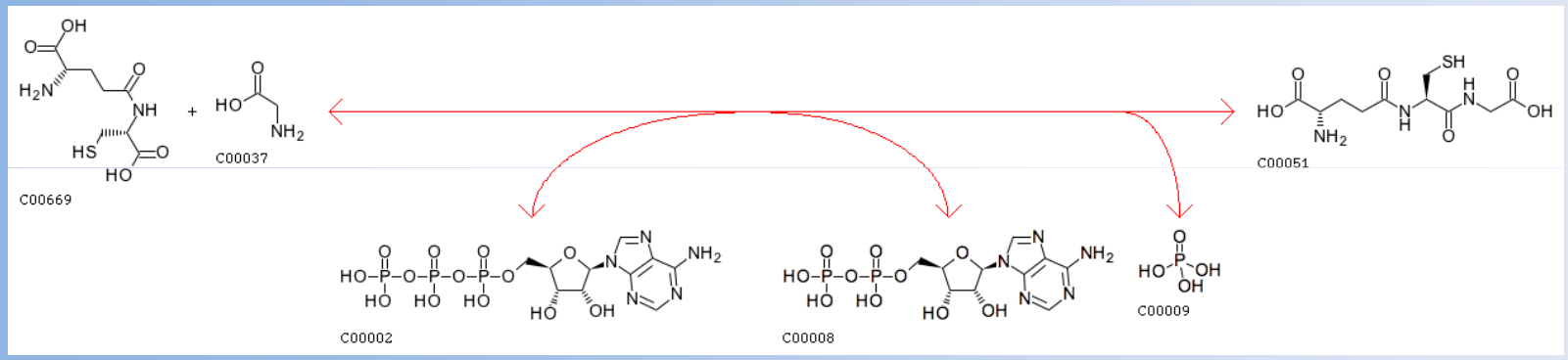
D-Glucose ⇌ D-Fructose

3.2.1. What are the functions of proteins?

3.2.1.2. Enzyme classification – EC number

3.2.1.2.6. Ligases

They are known as synthetases involved in the formation of a covalent bond joining two molecules together, coupled with the hydrolysis of a nucleoside triphosphate.

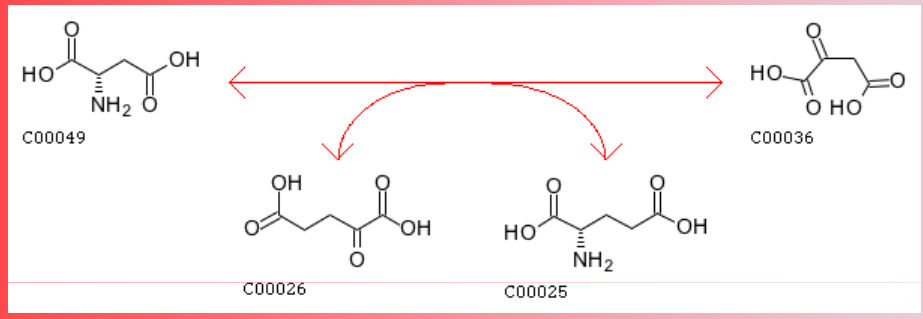


6.3.2.3 glutathione synthase

ATP + gamma-L-Glutamyl-L-cysteine + Glycine ⇌ ADP + Orthophosphate + Glutathione

Please solve the problem.

Question 1: To which enzyme group (EC number according to enzyme classification) does it belong?



5 points

L-Aspartate + 2-Oxoglutarate ⇌ Oxaloacetate + L-Glutamate 3 points

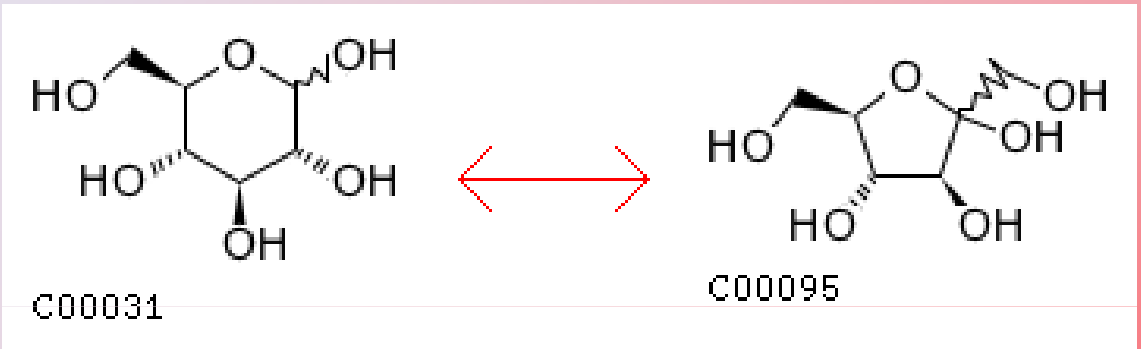
Aspartate aminotransferase catalyses this reaction. 2 points

This enzyme group can transfer functional groups between donor and acceptor molecules. 1 point

Second group EC 2.-.-.-.

Please solve the problem.

Question 2: To which enzyme group (EC number according to enzyme classification) does it belong?



5 points

D-Glucose ⇌ D-Fructose

3 points

D-glucose aldose-ketose-isomerase catalyses this reaction

2 points

This enzyme group catalyses molecular isomerizations and includes epimerases, racemases and intramolecular transferases.

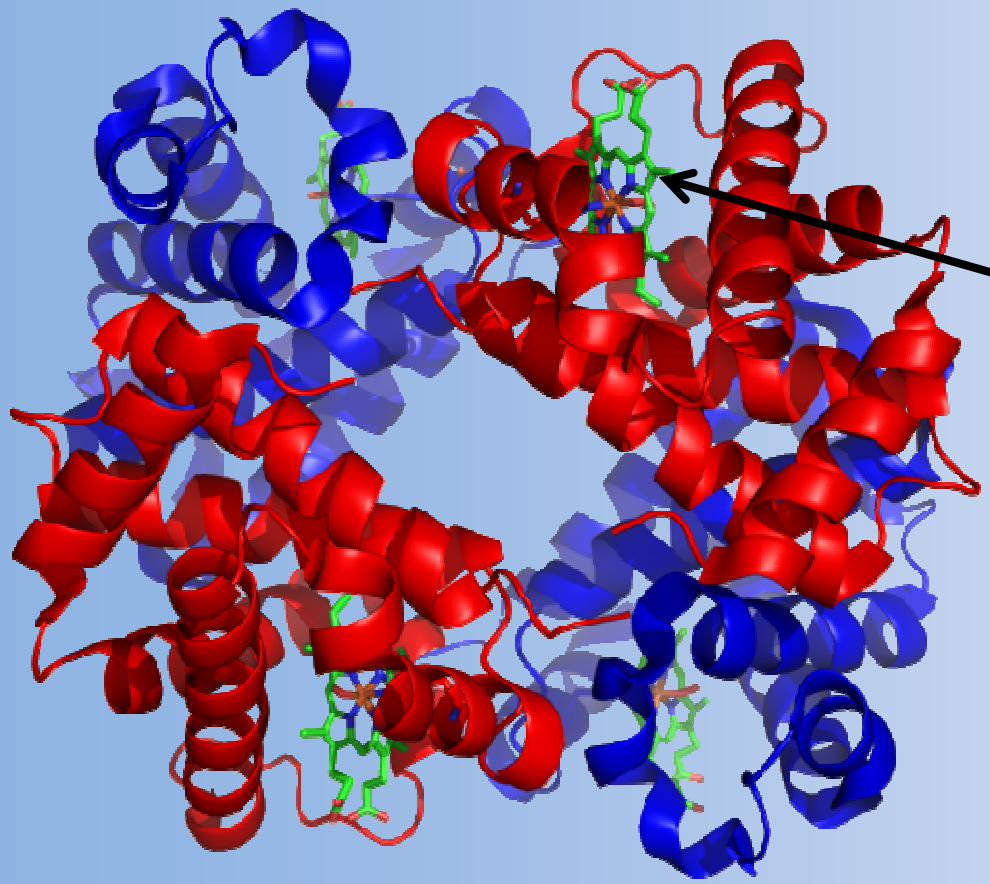
1 point

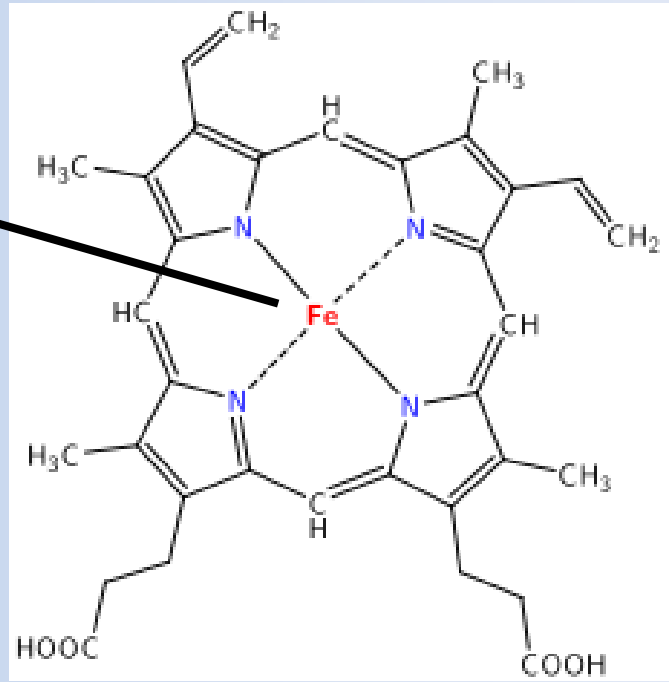
Fifth group EC 5.-.-.-.

3.2.1. What are the functions of proteins?

3.2.1.3. Transport proteins

Haemoglobin

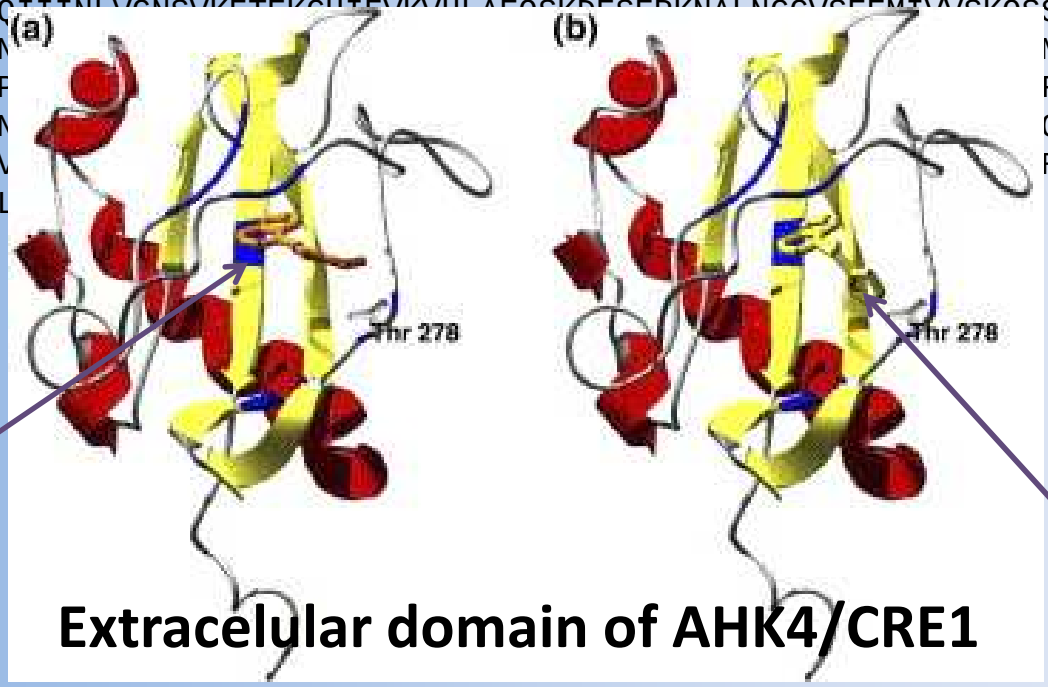


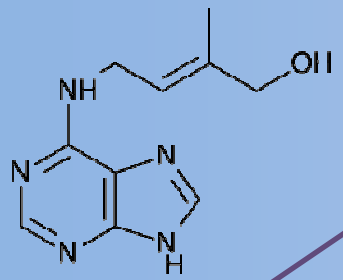


3.2.1. What are the functions of proteins?

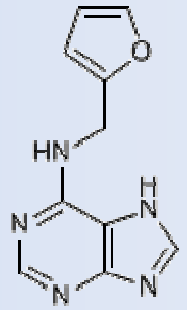
3.2.1.4. Signalling proteins – small molecules

MNWALNNHQEEEEEPRRIEISDSESLENLKSSDFYQLGGGGALNSSEKPRKIDFWRSGLMGFAKMQQQQQLQHSVAVKMNNNNNNDLMGNK
KGSTFIQEHRALLPKALILWIIIVGFISSGIYQWMDDANKIRREEVLVSMCDQRARMLQDQFSVSVNHVHALAILVSTFHYHKNPSAIDQE
TFAEYTARTAFERPLLSGVAYAEKVVNFEREMFERQHNWVIKTMDRGEPSPVRDEYAPVIFSQDSVSYLESLDMMSGEEDRENILRARETG
KAVLTSPFRLLETHHLGVVLTFPVYKSSLPENPTVEERIAATAGYLGGAFDVESLVENLLGQLAGNQAIVVHVYDITNASDPLVMYGNQDE
EADRSLSHESKLDFGDPFRKHKMICRYHQKAPIPLNVLTTVPLFFAIGFLVGYILYGAAMHIVKVEDDFHEMQELKVRAEAADVAKSQFLA
TVSHEIRTPMNGILGMLAMLLDTELSSTQRDYAQTAQVCGKALIALINEVLDRAKIEAGKLELESVPFDIRSILDDVLSLFSEESRNKGIE
LAVFVSDKVPEIVKGDSGRFRQIIINLVGNSVKFTEKGHIFVKVHLAEQSKDESERKNALNGGVSEEMIVVSKQSSYNTLSGYEAADGRNS
WDSFKHLVSEEQSLSEFDISSN MRGQINFISRPHIGST
FWFTAVLEKCDKCSAINHMKKP FERNGSPLPTKPQLDM
ILVEKDSWISTEDNDSEIRLLN QVLELRKTRQQHPEGS
SPATLKSLLTGKKILVVDDNIV RQIRMMEKEAKEKTNL
EWHLPILAMTADVIHATYEECL





trans-zeatin

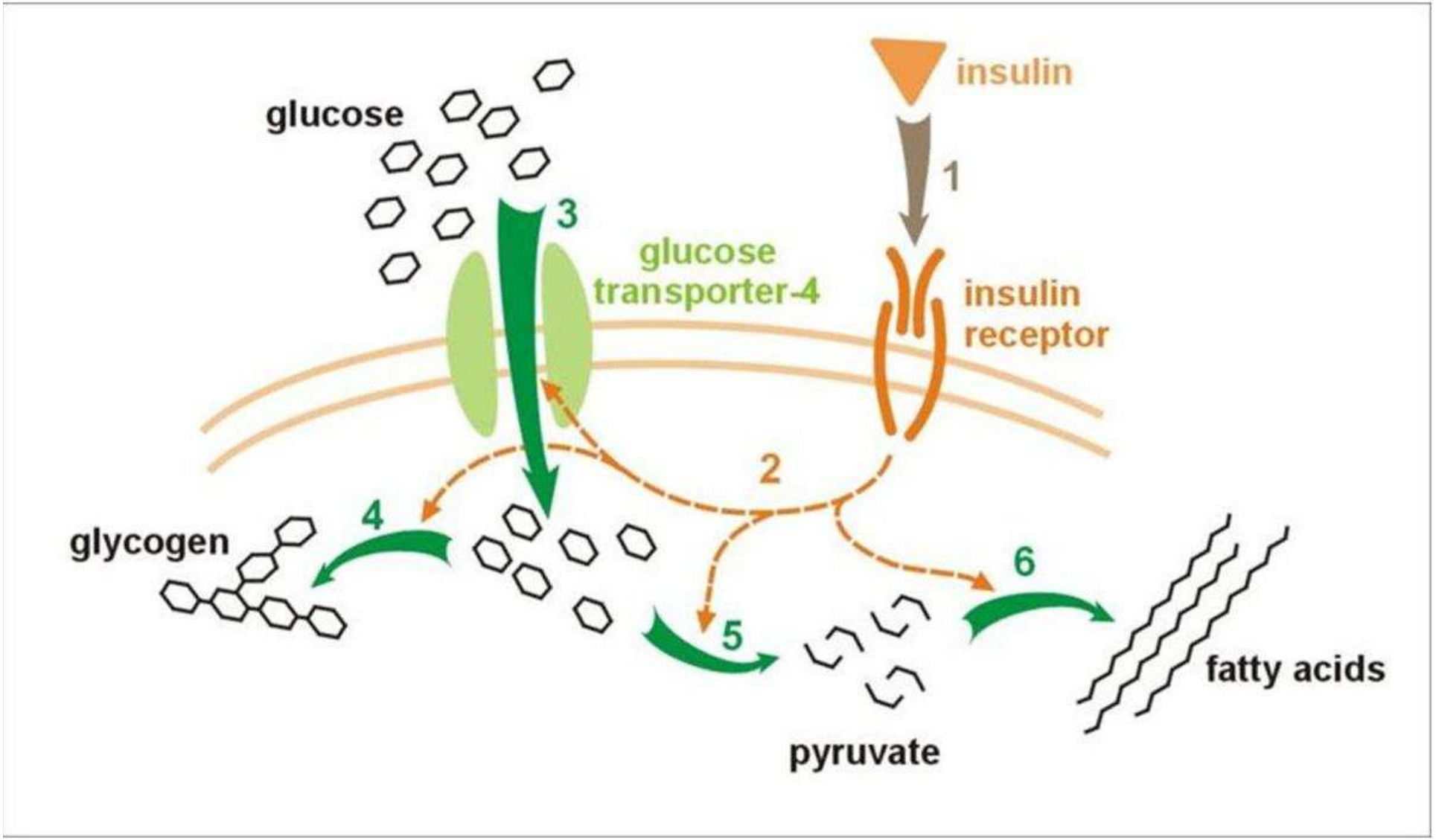


kinetin

Extracelular domain of AHK4/CRE1

3.2.1. What are the functions of proteins?

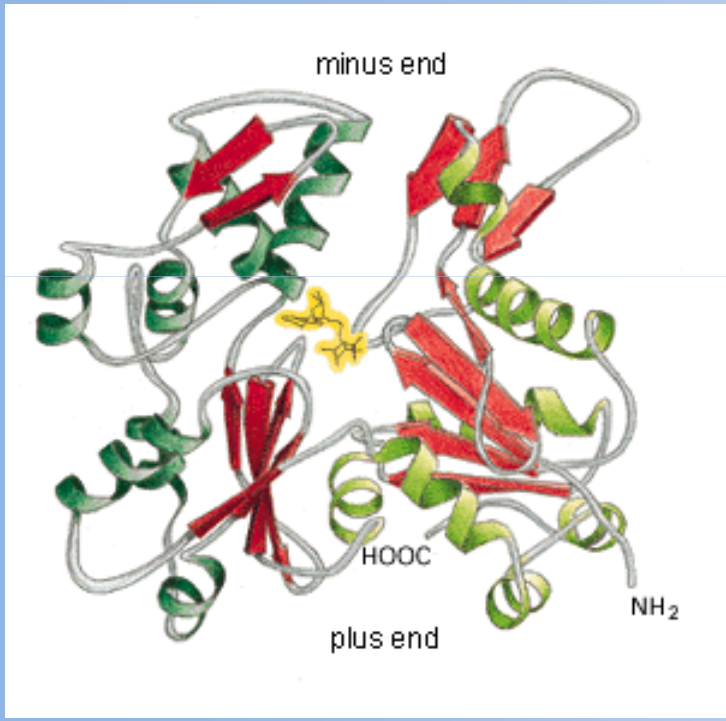
3.2.1.4. Signalling proteins – macromolecules



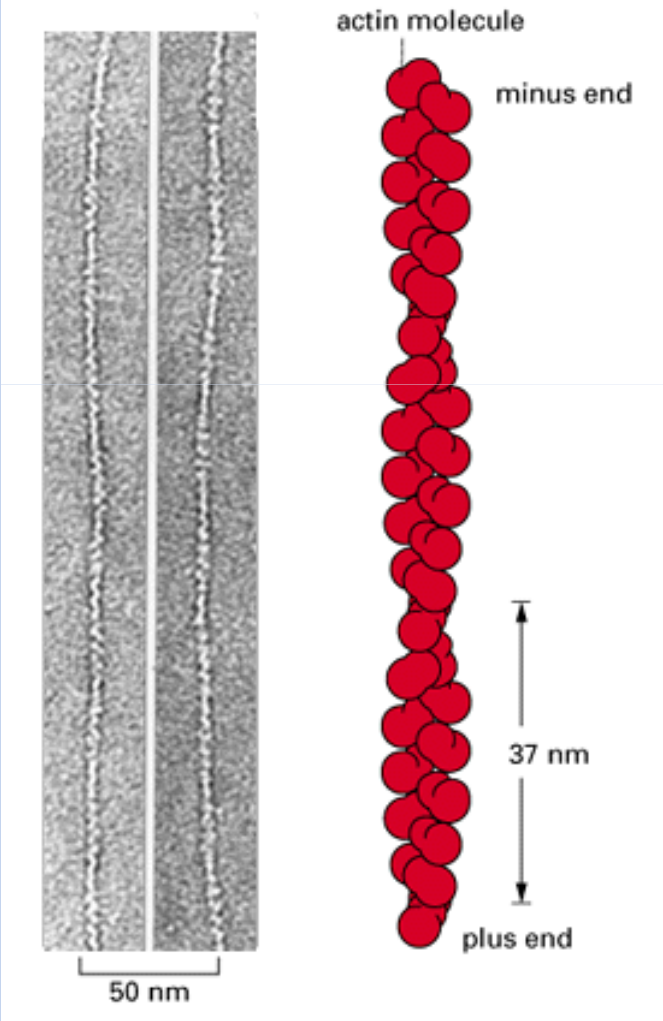
3.2.1. What are the functions of proteins?

3.2.1.5. Structural proteins

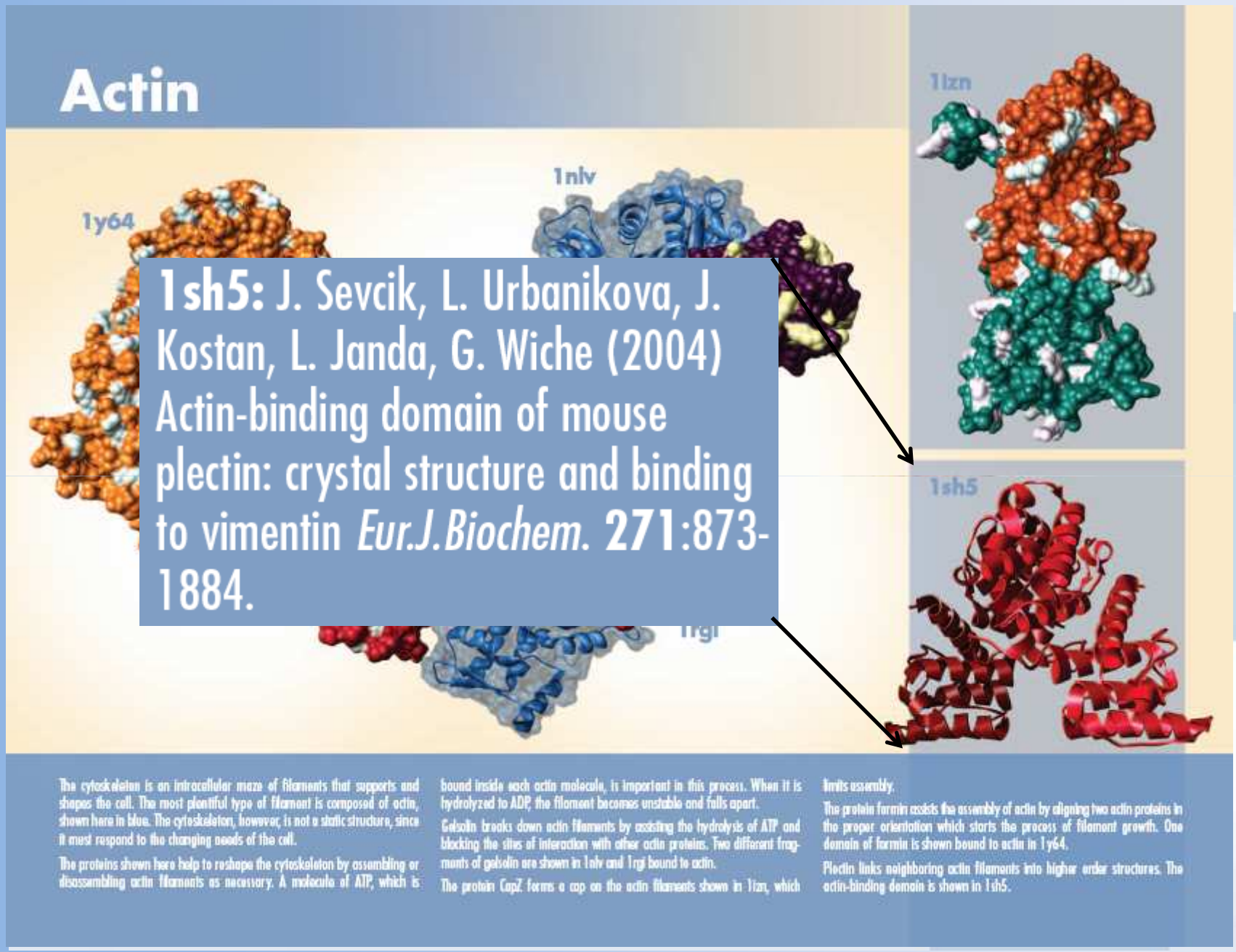
G-actin

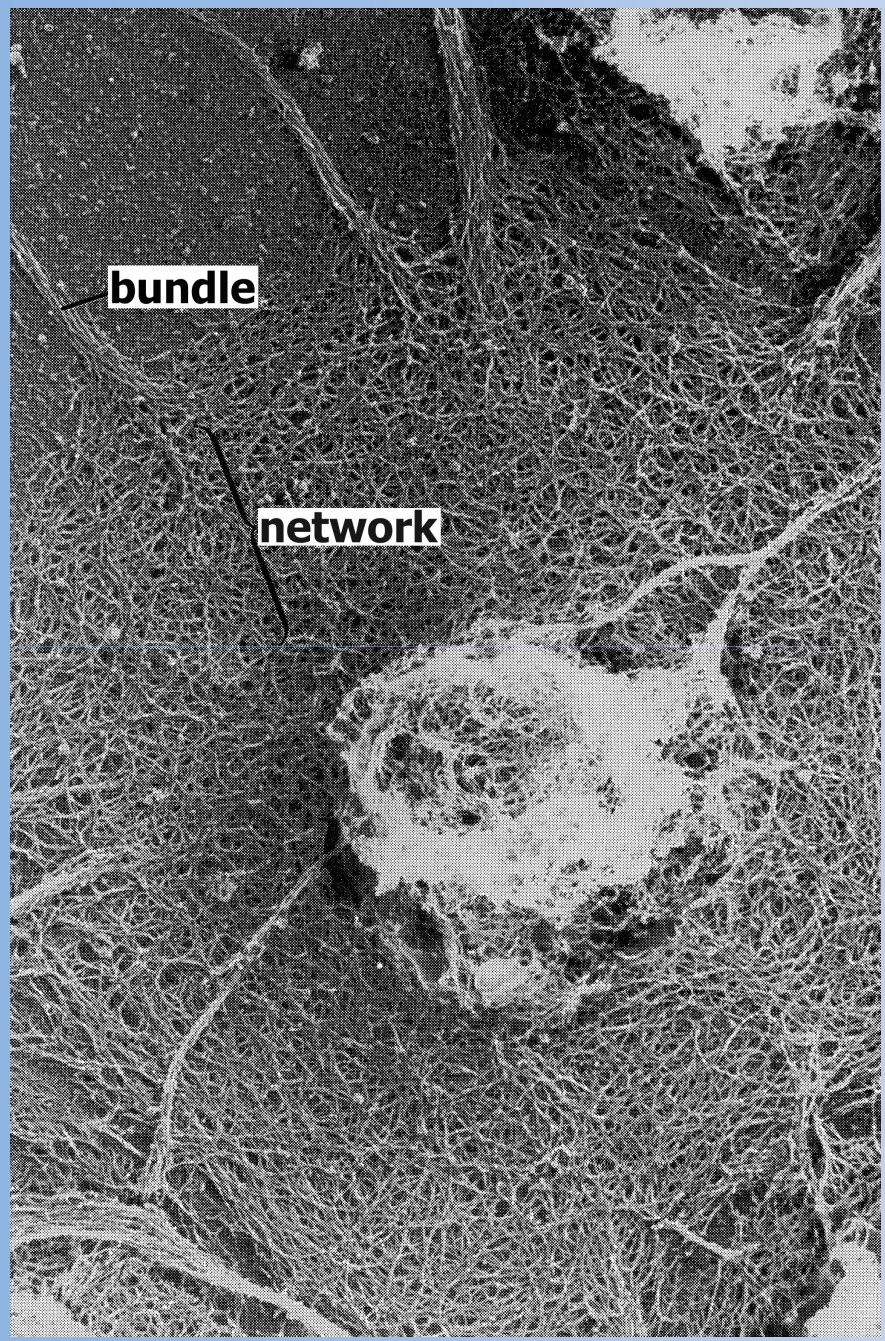


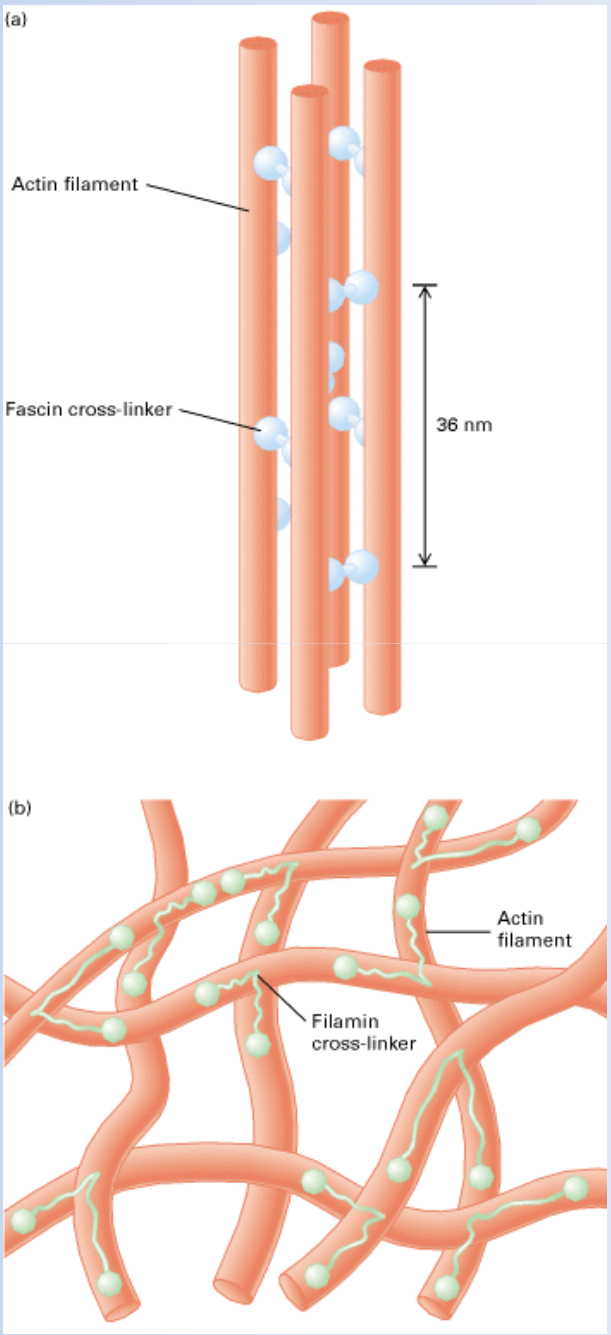
F-actin



Structural proteins – most abundant proteins within biological systems and include such proteins as collagen, keratin, actin, tubulin, and spectrin


Actin
1y64
1niv
1izn
1sh5
1sh5: J. Sevcik, L. Urbanikova, J. Kostan, L. Janda, G. Wiche (2004) Actin-binding domain of mouse plectin: crystal structure and binding to vimentin *Eur.J.Biochem.* **271**:873-1884.
The cytoskeleton is an intracellular maze of filaments that supports and shapes the cell. The most plentiful type of filament is composed of actin, shown here in blue. The cytoskeleton, however, is not a static structure, since it must respond to the changing needs of the cell.
The proteins shown here help to reshape the cytoskeleton by assembling or disassembling actin filaments as necessary. A molecule of ATP, which is bound inside each actin molecule, is important in this process. When it is hydrolyzed to ADP, the filament becomes unstable and falls apart.
Gelsolin breaks down actin filaments by assisting the hydrolysis of ATP and blocking the sites of interaction with other actin proteins. Two different fragments of gelsolin are shown in 1niv and 1rgi bound to actin.
The protein CapZ forms a cap on the actin filaments shown in 1izn, which limits assembly.
The protein formin assists the assembly of actin by aligning two actin proteins in the proper orientation which starts the process of filament growth. One domain of formin is shown bound to actin in 1y64.
Plectin links neighboring actin filaments into higher order structures. The actin-binding domain is shown in 1sh5.

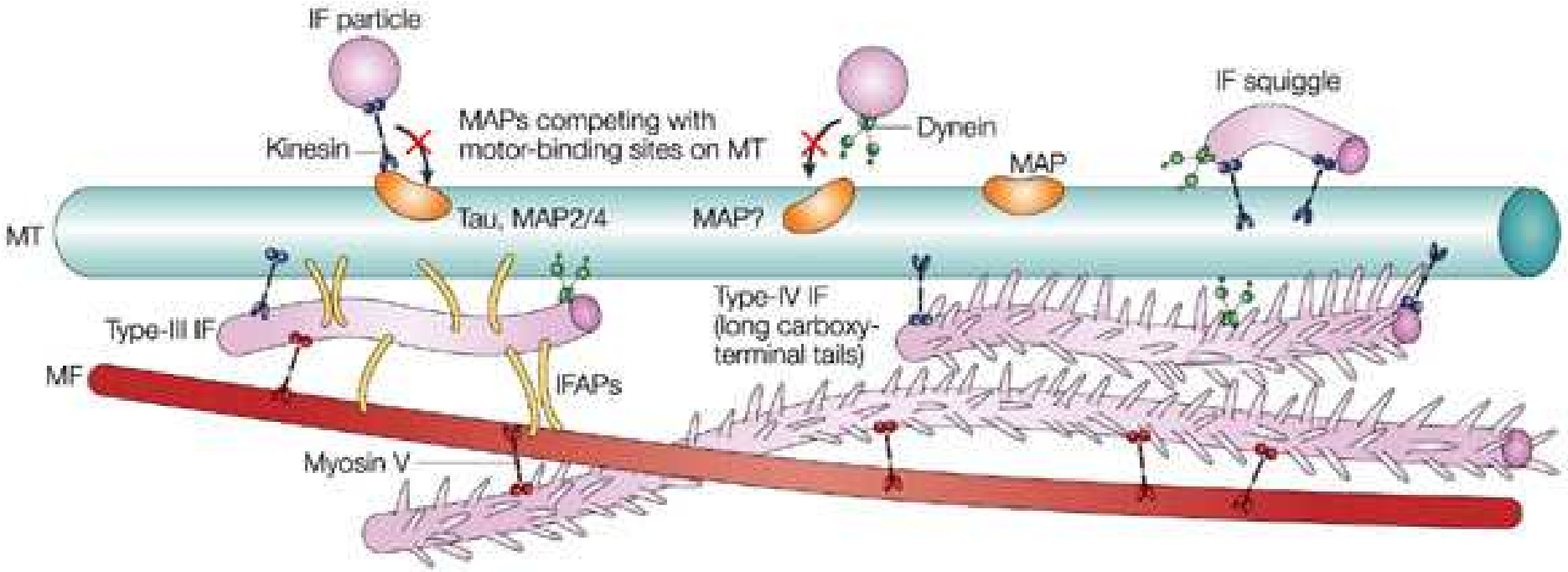

bundle
network


(a)
Actin filament
Fascin cross-linker
36 nm
(b)
Actin filament
Filamin cross-linker

3.2.1. What are the functions of proteins?

3.2.1.6. Motor proteins

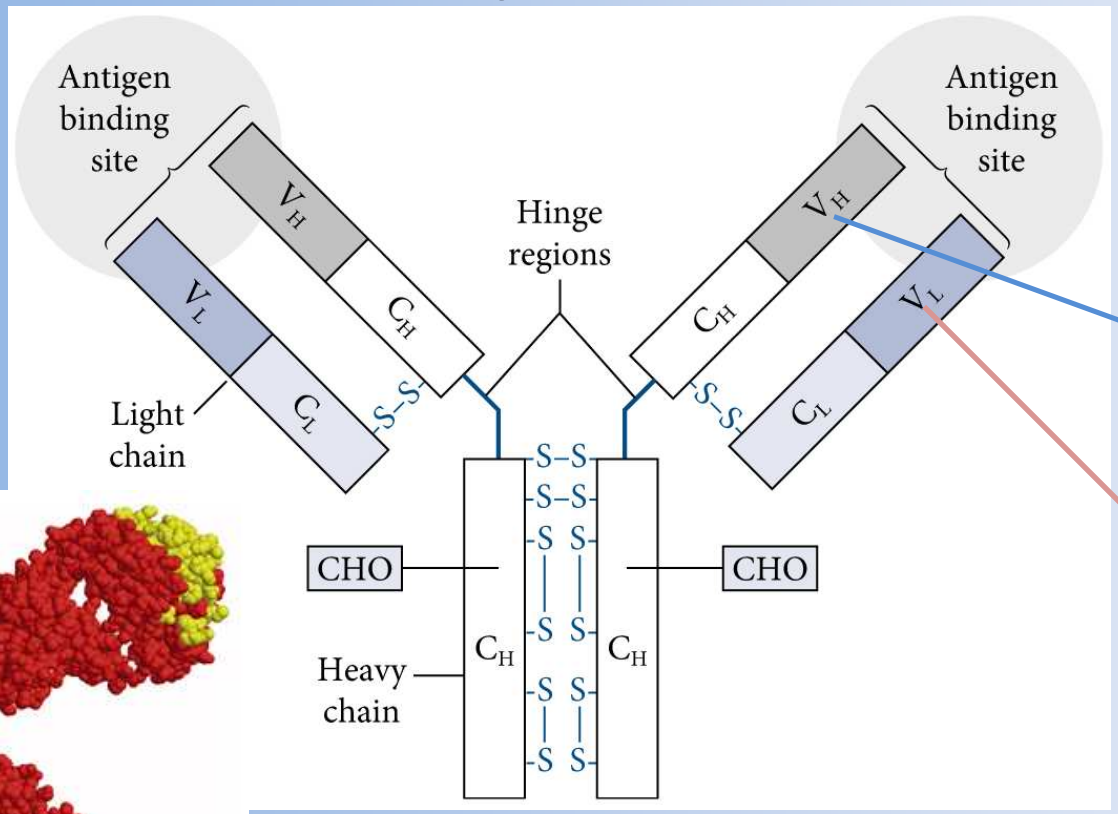
Actin and myosin – muscle contraction (60% efficiency)
Microtubules – kinesin
dynein

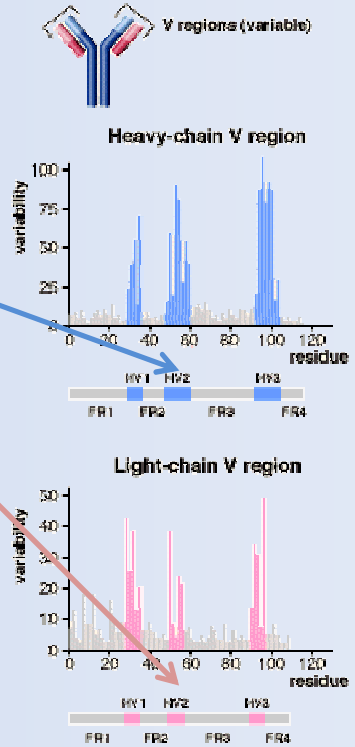


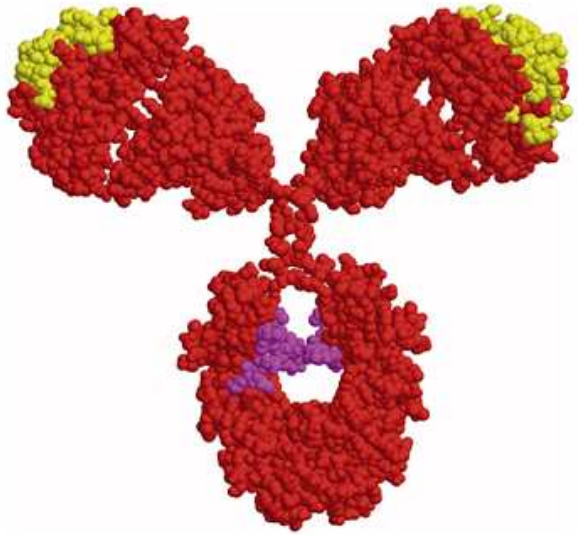
3.2.1. What are the functions of proteins?

3.2.1.7. Binding proteins

Antibody domain structure





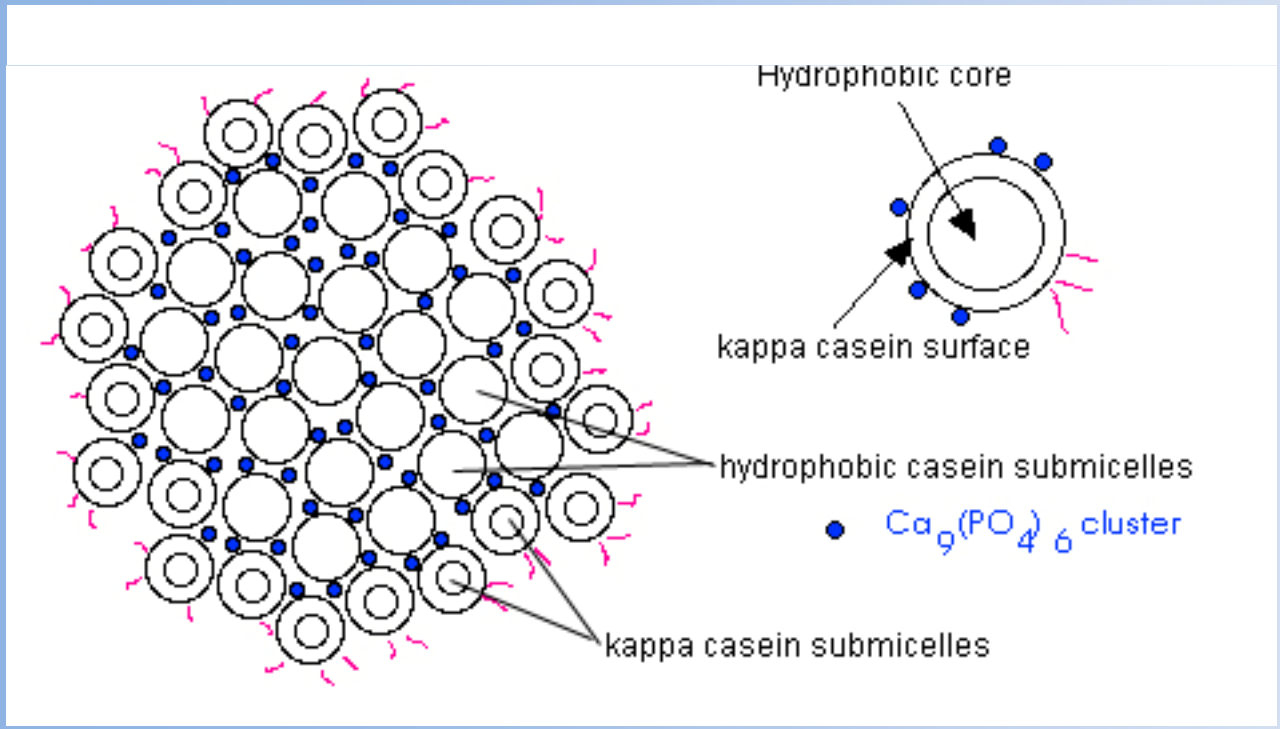


Antibodies play a central role in the adaptive immune response of vertebrates by recognizing and binding antigens present on the surface of viruses, bacteria and other infectious agents.

3.2.1. What are the functions of proteins?

3.2.1.8. Storage proteins

- The casein content of milk represents about 80% of milk proteins.
- Calcium binding by the individual caseins is proportional to the phosphate content.
- The high number of **proline residues** in caseins causes particular bending of the protein chain and inhibits the formation of close-packed, ordered secondary structures.
- Caseins contain no disulfide bonds. Furthermore, the lack of tertiary structure accounts for the stability of caseins against heat denaturation as there is very little structure to unfold.



Casein or **ferritin** sequester ligands (Ca^{2+} or Fe^{3+}), which may prove toxic within a biological system.

Please solve the problem.

Question 3: To which group does this protein belong?

The protein is dynein.	5 points
This protein does not interact with small molecules.	3 points
It is a motor protein.	2 points
This protein interacts with microtubules.	1 point

Third protein group (macromolecules binding proteins).

Please solve the problem.

Question 4: To which group does this protein belong?

The protein is AHK4.	5 points
This protein is a receptor histidine kinase.	3 points
This protein interacts with small molecules and macromolecules and has enzymatic activity.	2 points
This protein contains three domains.	1 point

Complex protein – Group 1, 2 and 3.

44,749 structures in the PDB have a structure factor file.

3.2. The key questions about a protein

3.2.2. What is the structure of the protein?

3.2.2.1. Determination of the amino acid sequence

The amino acid sequence of a protein can be used to generate a wealth of structural information:

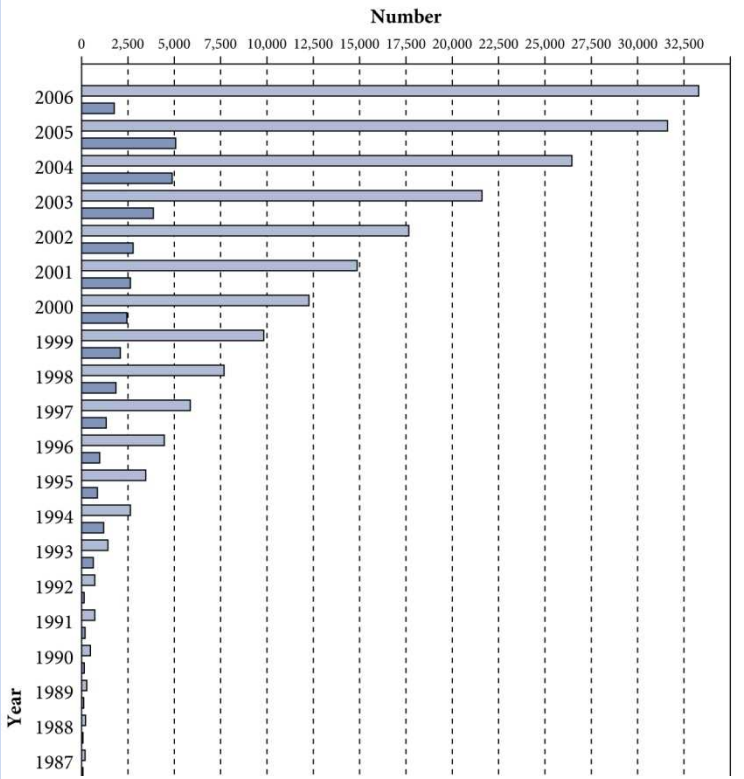
- theoretical mass,
- potential post-translational modification sites,
- structural and functional motifs, and
- X-ray crystallographic and NMR spec.

structures.

3.2.2.2. Experimentally determined mass

The theoretical mass of a protein, calculated on the basis of the amino acid sequence, does not take into account any post-translational modifications:

- disulphide bond formation
- phosphorylation
- glycosylation
- proteolysis



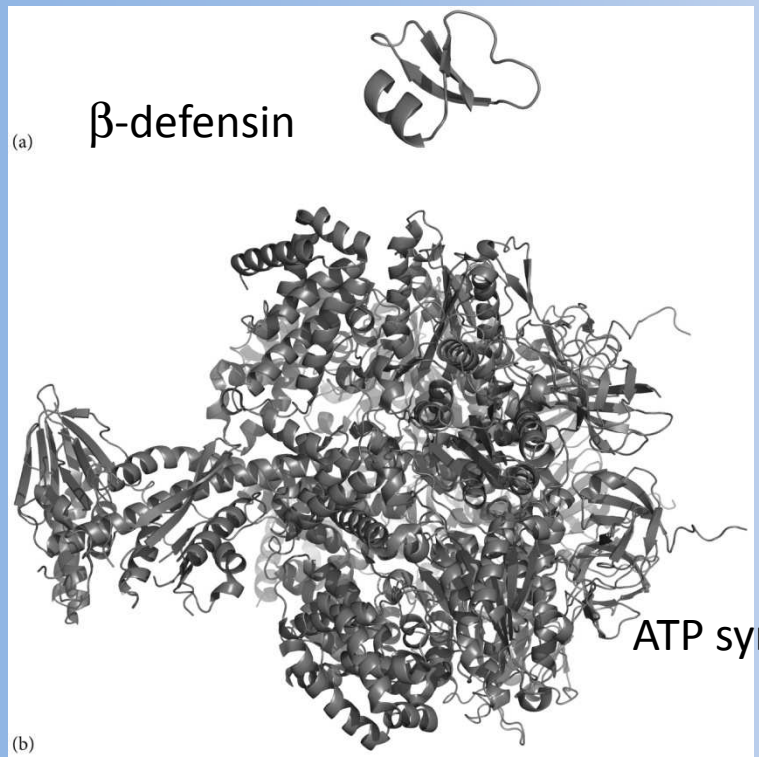
Experimental Method	Proteins	Nucleic Acids	Protein/Nucleic Acid complexes	Other	Total
X-ray diffraction	51756	1202	2386	17	55361
NMR	7240	895	154	7	8296
Electron microscopy	189	17	73	0	279
Hybrid	18	1	1	1	21
Other	120	4	4	13	141
Total:	59323	2119	2618	38	**64098**

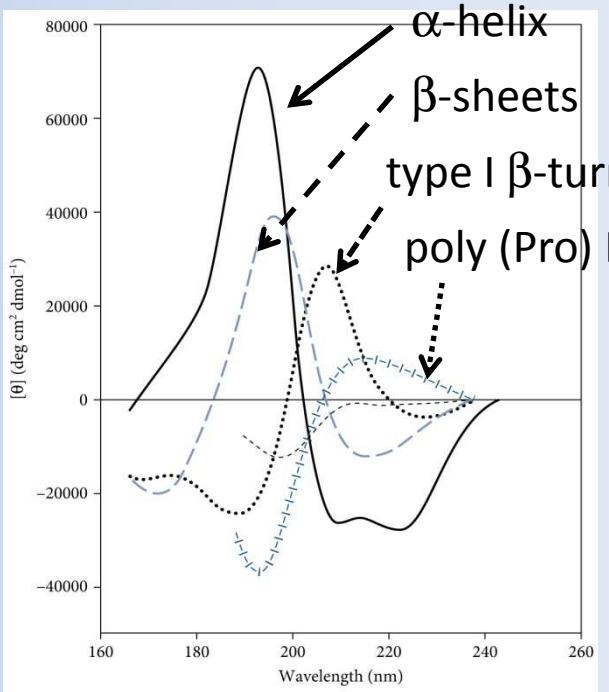
3.2. The key questions about a protein

3.2.2. What is the structure of the protein?

3.2.2.3. Characterization of secondary and tertiary structure

Proteins adopt three-dimensional structures that are dictated by the amino acid sequence of the protein. The predominant forms of secondary structure are α-helices and β-sheets. The secondary structure of a protein can be determined theoretically from its amino acid sequence or measured experimentally.





Far-UV CD spectra of various types of protein secondary structure

3.2. The key questions about a protein

3.2.2. What is the structure of the protein?

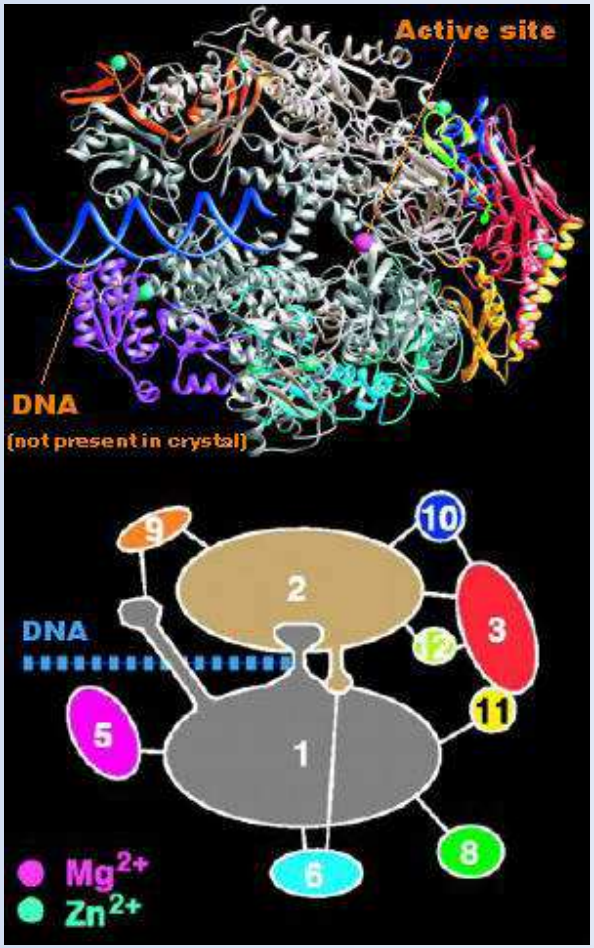
3.2.2.4. Characterization of quaternary structure

The quaternary structure can range in complexity from two identical subunits to multiple non-identical subunits.

- ribulose-biphosphate carboxylase EC 4.1.1.39 (photosynthetic bacteria) 2x55 kDa.
- RUBISCO (algae, plant) (8 x 55 kDa and 8 x 15 kDa)

High-resolution X-ray crystallography is required to provide the molecular details of subunit arrangements within multi-subunit proteins.

- proteasome
- RNA polymerase II



3.2. The key questions about a protein

3.2.3. What factors might affect the function of the protein?

Biological systems are dynamic and are capable of responding to environmental, developmental and metabolic changes.

- altering rate of gene expression
- altering rate of protein degradation
- altering the activity of the protein
  - non-covalent binding of other molecules
  - availability of ligands
  - reversible covalent modification
  - irreversible covalent changes

3.2. The key questions about a protein

3.2.4. How does the protein interact with other molecules?

Prokaryotic cell has a protein concentration: > 200 mg/ml.
In erythrocytes the concentration was found to be: > 300 mg/ml.
Mitochondrial matrix has a protein concentration: > 500 mg/ml.

To address how proteins interact with ligands, we should aim to determine:

- the three-dimensional structure of the protein-ligand complex,
- the site on the protein which interacts with the ligand and the molecular details of the interaction,
- the number of ligands interacting with the protein (stoichiometry), and
- the strength of the interaction and the rate constant involved.

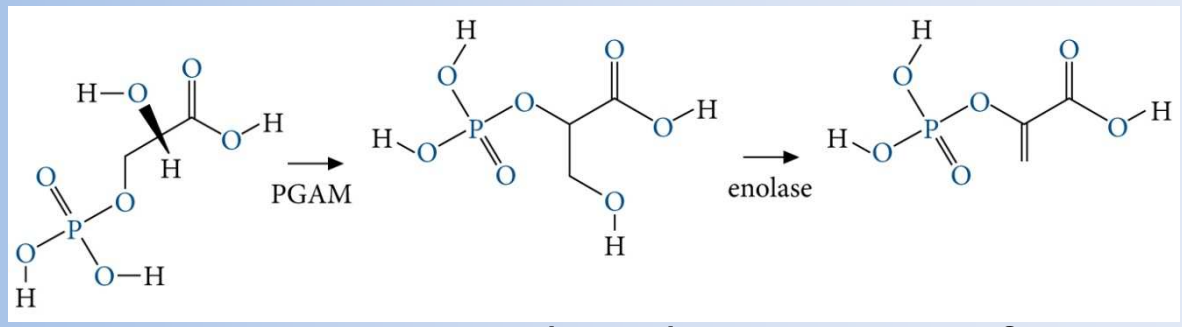
3.3. Assays for biological activity

Key concept: Appreciating the range of assays for biological activity measurements and their limitations

A specific assay is required for purification of a protein and to address key questions, relating its structure and function.

We must consider the care which must be exercised when interpreting assay data in general:

- Is the activity measurement due solely to the presence of the protein of interest?

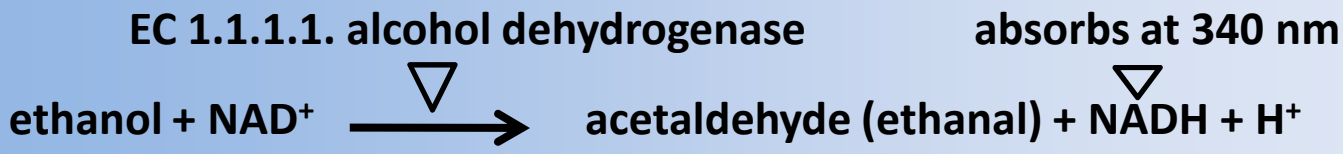


- Is the activity measurement proportional to the amount of protein present?
  - assay component limiting the activity measurement
  - continuous or discontinuous assay

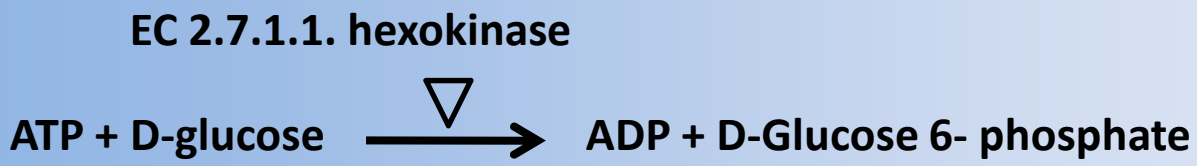
3.3. Assays for biological activity

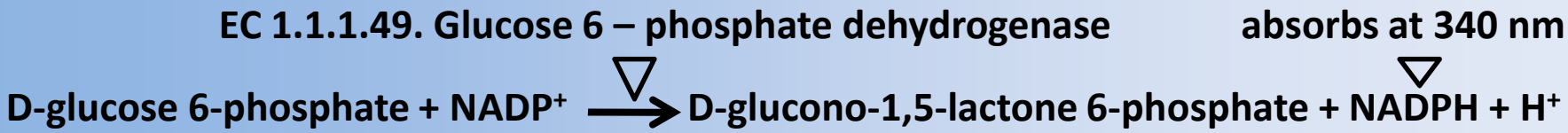
3.3.1. Catalytic proteins (enzymes)

Direct reaction:



Coupled reaction:



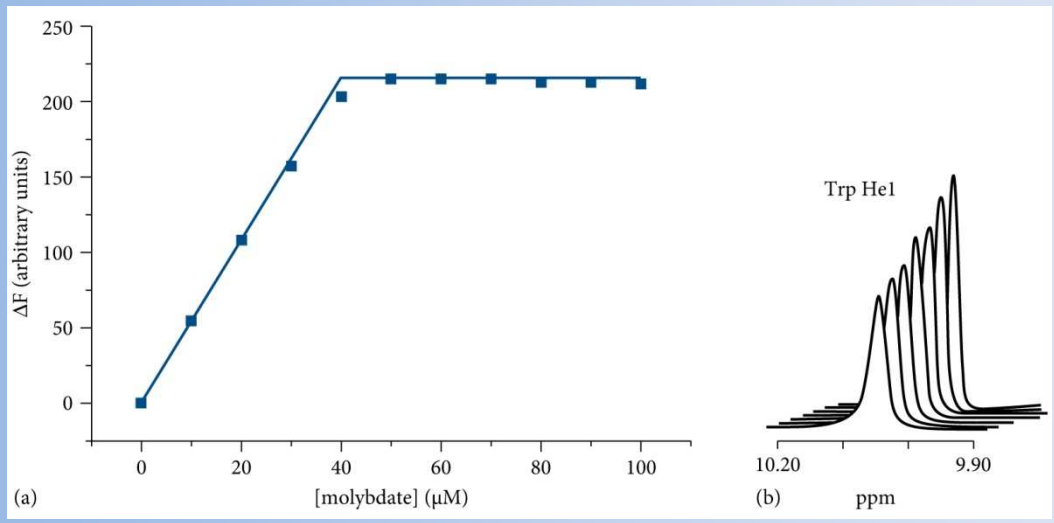


3.3. Assays for biological activity

3.3.2. Binding proteins

Protein binding assays can be used to both identify ligands and to characterize the nature of the protein–ligand interaction. Two categories:

- those based on biophysical changes in protein or ligand upon complex formation
- those that employ direct quantification of free and bound ligands.



- ❖Spectroscopic method – absorbance, fluorescence, CD spectroscopy, NMR analysis
- ❖Equilibrium dialysis and membrane filtration
- ❖Solid phase techniques – SPR (surface plasmon resonance), overlay assay, ELISA.

3.3. Assays for biological activity

3.3.3. Transport proteins

Transport protein assays are designed to measure the rate of transport of a ligand from one location to another.

➢In a typical assay, a known amount of isotopically labelled ligand is added to the system and incubated for a fixed period of time.

3.3.4. Other types of proteins

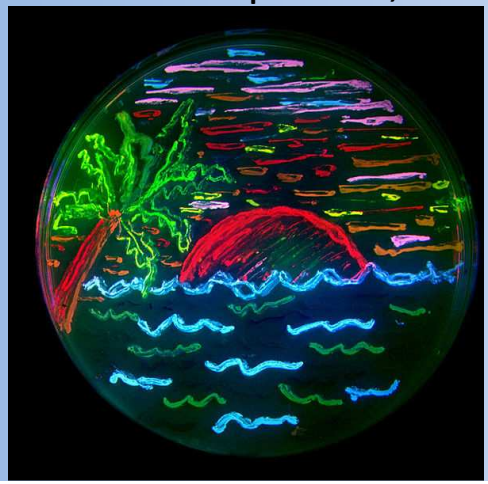
How to test the biological function of proteins which cannot be measured using the more conventional assays?

❖GFP – green fluorescent protein, which naturally occurs in the jellyfish *Aequorea victoria.*

❖*YFP*

❖*RFP*

❖*CFP*

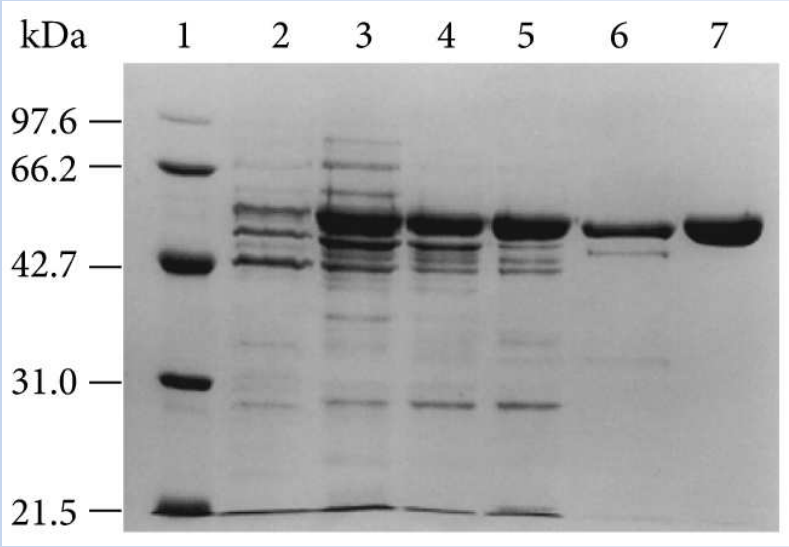


The diversity of genetic mutations is illustrated by this San Diego beach scene drawn with living bacteria expressing **8 different colors** of fluorescent proteins.

3.4. Purification of proteins

Key concepts: Knowing the objectives of protein purification
Being aware of the experimental considerations required to meet these objectives

- ❖ To retain maximum biological activity.
- ❖ To ensure the protein is indeed pure.
- ❖ To maximize the amount of protein recovered.



3.4.1. Wild-type proteins

The isolation of proteins from their natural source exploits their heterogeneous biochemical properties – molecular mass, pI, hydrophobicity, post-translational modifications, putative ligands.

3.4.2. Recombinant proteins

Protein purification has been simplified greatly by the advent of recombinant DNA technologies that enable high levels of protein expression – *E. coli, Aspergillus nidulans, S. cerevisiae, P. pastoris*, insect cells, mammalian cells, and transgenic plants and animals.

3.5. Structure determination

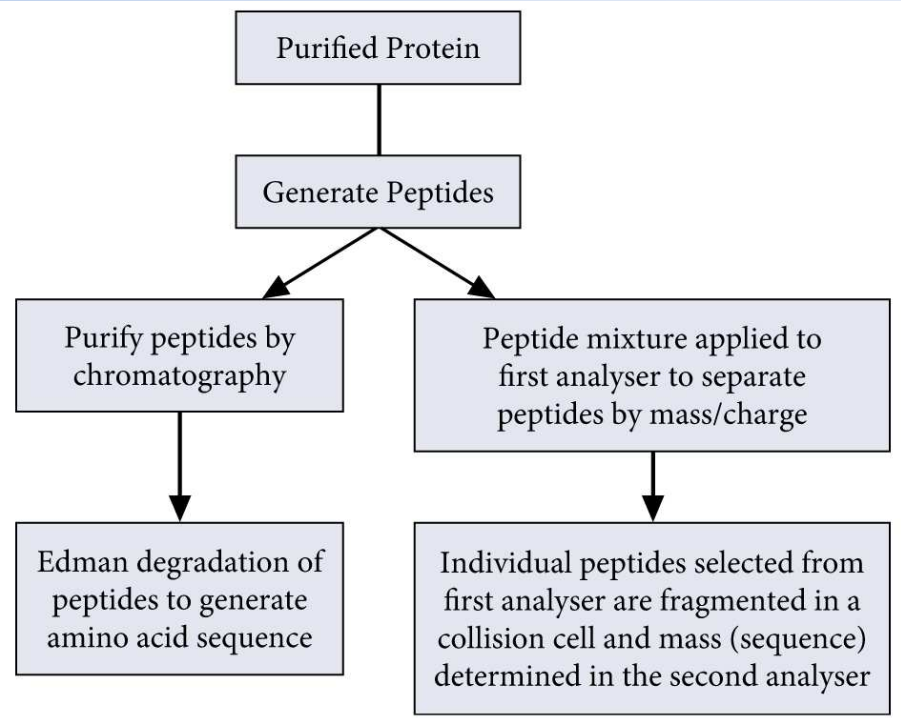
Key concepts:

- Understanding the tools required to explore the different levels of protein structure
- Appreciating the importance of an integrated experimental approach to provide a more complete picture of protein structure

The ultimate aim of protein structure determination is to gather and interpret data to reveal the three-dimensional structure of the protein at an atomic level.

- X-ray crystallography
- High-resolution NMR spectroscopy

The nature of the post-translational modification can be determined by SDS-PAGE combined with specific removal of modifications and mass spectrometry to analyze the mass of peptides with modifications.

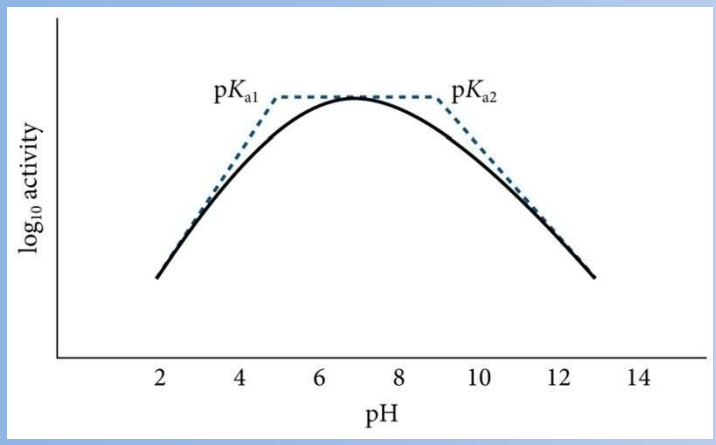


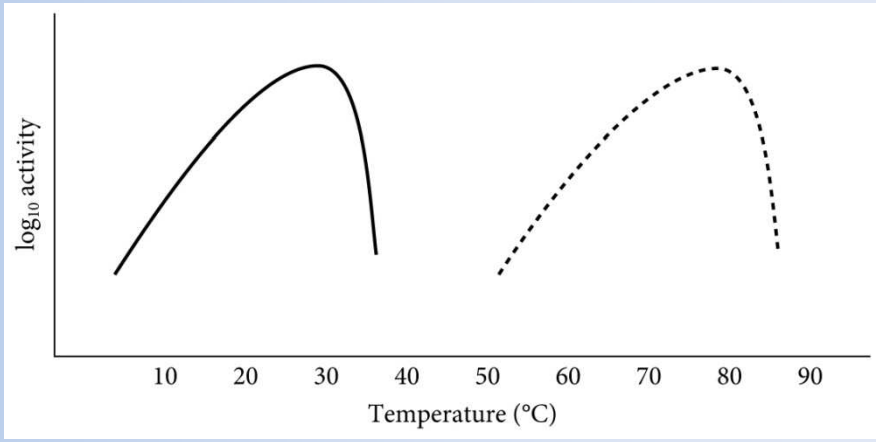
Direct determination of the structure.

3.6. Factor affecting the activity of proteins

Key concepts:
- Defining the major factors which influence the activity of proteins
- Understanding how to monitor their effects on protein structure and function

3.6.1. pH and temperature





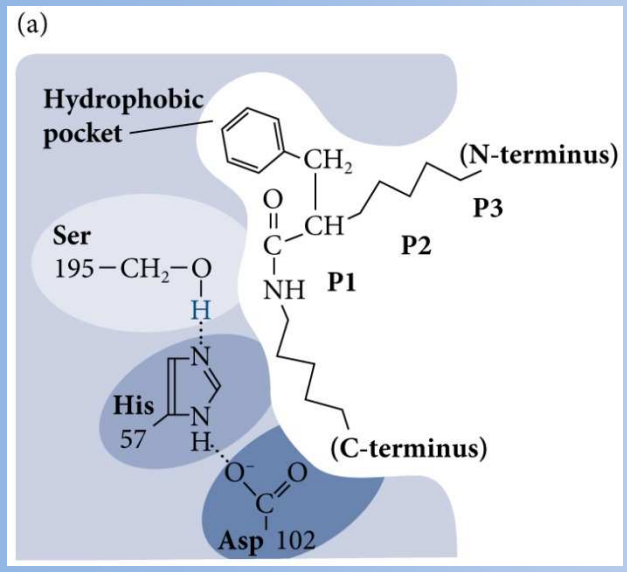
pH conditions that elicit the highest level of activity promote side chain side ionization states (pK_a values of 4.0 and 10.5 are indicative of glutamic acid and lysine, respectively).

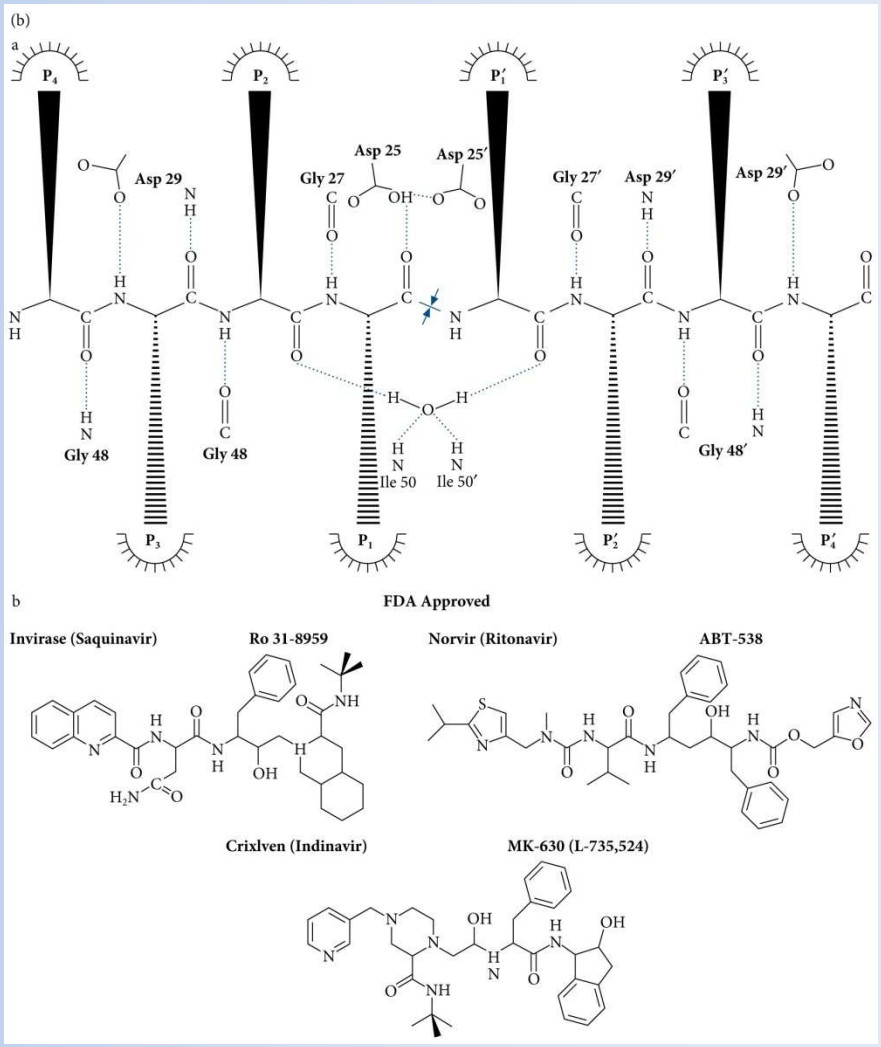
3.6. Factor affecting the activity of proteins

3.6.2. Inhibitor and activator molecules

Inhibitors and activators are important effectors of protein activity within the cell.

Biochemists usually quantify the effects of inhibitors in terms of an inhibitor constant, K_i, whereas pharmacologists quantify the effects of an inhibitor with the term IC_{50} (or $I_{0.5}$), which is the concentration of inhibitor required to reduce the protein activity by 50%.





Competitive inhibition and non-competitive inhibition

3.6. Factor affecting the activity of proteins

3.6.3. Post-translational modifications

Post-translational modification is one of the major effectors of protein activity within the cell.

The complexity of higher organisms is due to the differential RNA processing and post-translational modification of gene products.

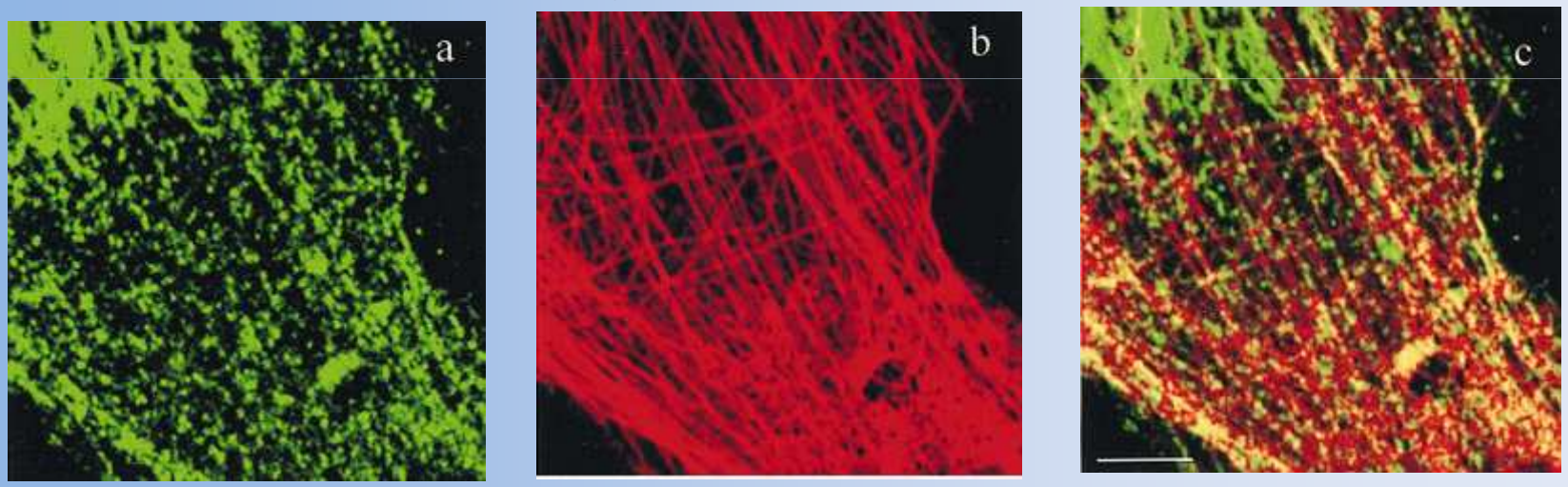
Examples of post-translational modifications	
Amino acid modifications	
Modifications	Example
Cysteine: disulphide bond formation	Lysozyme
Lysine biotinylation	Acetyl CoA carboxylase
Serine phosphorylation	Glycogen phosphorylase
Threonine phosphorylation	Cyclin-dependent kinase
Tyrosine phosphorylation	Cortactin
Addition of prosthetic group – thiamine diphosphate	Pyruvate dehydrogenase
Proteolytic processing	Chymotrypsin
Protein targetting (signal sequences)	Penicillin acylase

3.7. Interactions with other macromolecules

Key concepts:

- Appreciating the importance of protein–protein interactions *in vivo*
- Understanding the need to use appropriate experimental conditions to study these interactions

Whilst it is convenient to study isolated proteins *in vitro*, this does not reflect how they function *in vivo*.



The colocalization of vimentin dots and microtubules at the edges of spreading cells. (*a) Confocal* immunofluorescence localization of vimentin at the edge of a fixed cell 30–45 minutes after replating. (*b) Microtubule* localization in the same cell. (*c) An overlay of images a and b. Yellow,* the regions of colocalization between vimentin dots and microtubules. Bar, 2.5 mm.

3.8. Use of bioinformatics

Key concepts:
- Appreciating the range of databases and bioinformatic tools available to assist protein characterization
- Understanding the theoretical basis of the tools used to calculate properties of a protein from its sequence

The field of bioinformatics has harnessed the exponential growth in the amount of information relating to nucleotide sequences, protein sequences, and biomolecular structures.

3.8.1. Web resources and databases

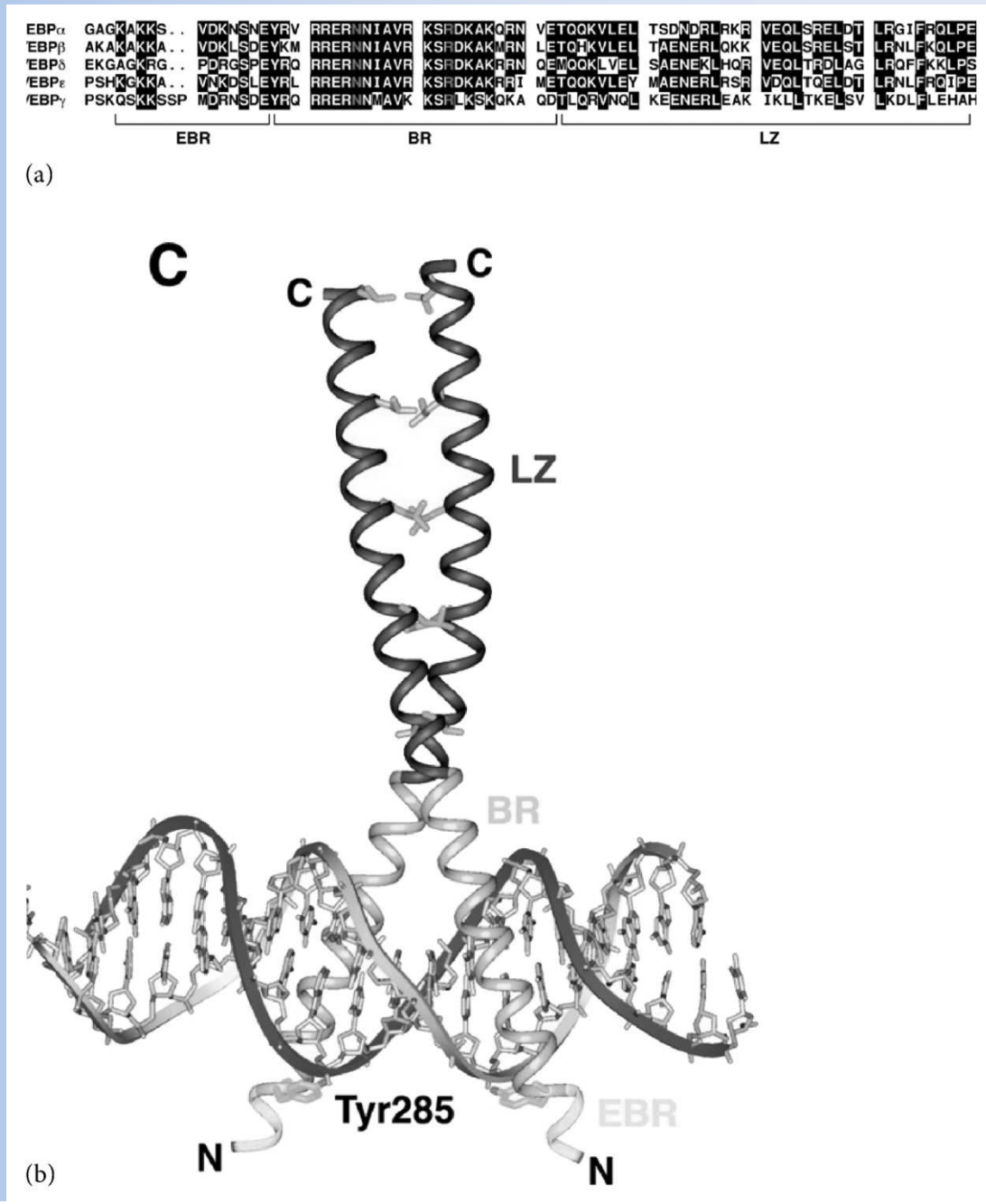
Nucleotide sequences, amino acid sequences, and protein structures are collected within a number of web-based databases.
More recently, there has been an effort to integrate data sets (e.g. linking individual **nucleotide/protein sequences** to related **three-dimensional structures**, **metabolic pathway** databases, **enzyme** databases, **disease** databases, **organism-specific** databases, **two-dimensional gel** databases, and **associated references**), allowing researchers to characterize more fully the structural and functional properties of proteins.

3.8. Use of bioinformatics

3.8.2. Sequence analysis

Protein sequences, derived from experimental data or database entries, can generate a wealth of information that can assist protein characterization.

A number of software packages which calculate the physico-chemical properties of proteins from their sequences are available on the web, e.g. ProtParam, available on the ExPASy server.



3.8. Use of bioinformatics

3.8.2. Sequence analysis

It is possible to compute the absorption coefficient of a protein knowing the amino acid sequence and the molar absorption coefficients (E) of tyrosine, tryptophan, and disulphide bonds at A_{280} using the following equation:

$$E_{Prot} = N_{Tyr} \times E_{Tyr} + N_{Trp} \times E_{Trp} + N_{Cys} \times E_{Cys}$$

where N_x is the number of amino acid X per polypeptide chain. The values of E_{Tyr}, E_{Trp} and E_{Cys} are 1,490, 5,500 and 62.5 $M^{-1}cm^{-1}$, respectively.

Sample calculations of the absorption coefficients for a range of proteins.

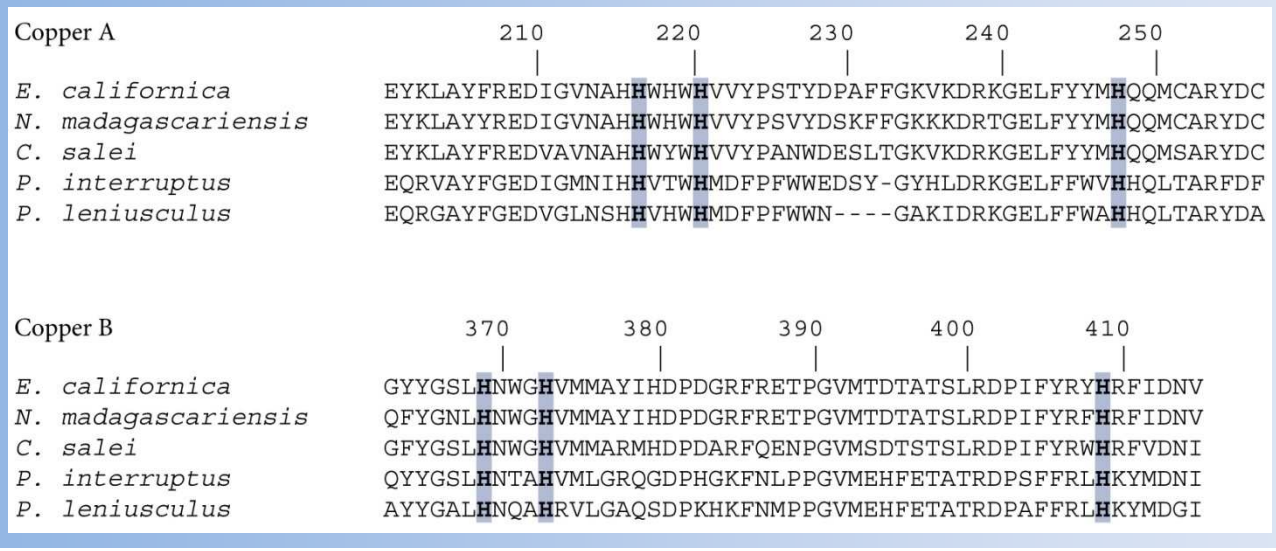
Protein	Molecular mass (Da)	N_{Tyr}	N_{Trp}	N_{Cys}	E_{Prot} (calc) (1M)	E_{Prot} (calc) (1 mg/ml)	E_{Prot} (exp) (1 mg/ml)
Insulin (bovine)	5,734	4	0	6	5,960 (6,335)	1.04 (1.10)	0.97
Lysozyme (hen)	14,314	3	6	8	37,470 (37,970)	2.62 (2.65)	2.63
Chymotrypsinogen (bovine)	25,666	4	8	10	49,960 (50,585)	1.95 (1.97)	1.98
Phosphoglycerate kinase (yeast)	44,607	7	2	1	21,430 (21,430)	0.48 (0.48)	0.49
Pyruvate kinase (rabbit muscle)	57,917	9	3	9	29,910 (30,410)	0.52 (0.53)	0.54
Serum albumin (bovine)	66,296	20	2	35	40,800 (42,925)	0.62 (0.65)	0.66

3.8. Use of bioinformatics

3.8.3. Sequence comparison

Whilst pairwise alignments provide a measure of the similarity of two sequences, the information arising from multiple sequence alignments, using more than two sequences from a protein family, can be used to:

- Suggest the function of unknown proteins: E-scores less than 0.2 are indicative of homology and values greater than 1.0 indicate that the similarities are just as likely to have arisen by chance.
- Identify functionally and/or structurally important residues.
- Improve structure prediction tools.
- Detect and characterize evolutionary relationships.



```
Copper A                210       220       230       240       250
                        |         |         |         |         |
E. californica          EYKLAYFREDIGVNAHHWHWHVVYPSTYDPAFFGKVKDRKGELFYYMHQQMCARYDC
N. madagascariensis     EYKLAYYREDIGVNAHHWHWHVVYPSVYDSKFFGKKKDRTGELFYYMHQQMCARYDC
C. salei                EYKLAYFREDVAVNAHHWYWHVVYPANWDESLTGKVKDRKGELFYYMHQQMSARYDC
P. interruptus          EQRVAYFGEDIGMNIHHVTWHMDFPFWWEDSY-GYHLDRKGELFFWVHHQLTARFDF
P. leniusculus          EQRGAYFGEDVGLNSHHVHWHMDFPFWWN----GAKIDRKGELFFWAHHQLTARYDA


Copper B                370       380       390       400       410
                        |         |         |         |         |
E. californica          GYYGSLHNWGHVMMAYIHDPDGRFRETPGVMTDTATSLRDPIFYRYHRFIDNV
N. madagascariensis     QFYGNLHNWGHVMMAYIHDPDGRFRETPGVMTDTATSLRDPIFYRFHRFIDNV
C. salei                GFYGSLHNWGHVMMARMHDPDARFQENPGVMSDTSTSLRDPIFYRWHRFVDNI
P. interruptus          QYYGSLHNTAHVMLGRQGDPHGKFNLPPGVMEHFETATRDPSFFRLHKYMDNI
P. leniusculus          AYYGALHNQAHRVLGAQSDPKHKFNMPPGVMEHFETATRDPAFFRLHKYMDGI
```

3.8. Use of bioinformatics

3.8.4. Structure comparisons

Protein structure comparisons are powerful tools in establishing structure, function, and evolutionary relationships, particularly when comparing distantly related proteins with low-level sequence identity.

- SCOP (Structural Classification of Proteins)
- CATH (Class Architecture Topology Homologous superfamily)

3.8.5. Predicting possible functions of proteins

The classical experimental approach to determining the function of a novel protein requires structure characterization, determination of factors influencing activity, and identification of ligands.

NetPhos http://www.cbs.dtu.dk/services/NetPhos
NetNGlyc http://www.cbs.dtu.dk.services/NetNGlyc
CASTp http://www.sts.bioengr.uic.edu/castp/
PONDR http://www.pondr.com

Please solve the problem.

Question 5: What is the name of this method?

I can measure protein-protein interaction, but predominantly I determine if a protein is folded. 5 points

I work in far UV. 3 points

I am able to determine secondary but not tertiary structure. 2 points

I am a spectroscopic method. 1 point

CD spectroscopy.

Please solve the problem.

Question 6: What is the name of this post-translational modification?

I modify arginine, lysine and five other amino acids. 5 points

I am detectable by MALDI and western blot (but not for all kinds of modified amino acids). 3 points

Very often I need ATP for modification. 2 points

Serine and threonine are the most often used amino acids for this modification. 1 point

Phosphorylation.