Unit 07

Structural and Computational Biology

STUDENTS LEARNING OUTCOMES (SLO's)

After studying this unit, the students will be able to

- Define structural biology.
- Explain that structure determination of biomolecules are important.
- Describe how X-ray crystallography works.
- Outline the online databases where biomolecule structures are available.
- Describe computational biology.
- Define sequence homology.
- Define structural homology.

7.1 APPLICATIONS OF STRUCTURAL BIOLOGY

Describe the applications of structural biology. (Question from chapter Exercise and is explained in Long Question number 1 to 5).

1. What is Structural Biology and what kind of information does it provide about macromolecules?

Ans. Structural Biology

Definition: Structural biology is a specialized branch of biology that deals with the study of three-dimensional (3D) structures of macromolecules, including proteins and nucleic acids, at the atomic level. This discipline plays a critical role in understanding the fundamental architecture and behavior of biological molecules.

Information Provided by Structural Biology

Structural biology provides detailed and in-depth information about various aspects of biomolecules:

- **Structure:** It reveals the precise 3D arrangement of atoms in a macromolecule.
- **Function:** It helps in understanding how the shape and structure of a molecule relate to its specific biological function.
- **Dynamics:** It explains the movements and conformational changes of macromolecules under different conditions.
- **Interaction:** It sheds light on how biomolecules interact with ligands (small molecules) and other macromolecules, which is essential for processes such as enzyme activity, DNA replication, and signal transduction.

2. How does structural biology help in determining active sites and domains of macromolecules?

Ans. Role of 3D Structures in Identifying Functional Regions

One of the primary applications of structural biology is determining the active sites and domains of macromolecules such as proteins and nucleic acids. Structural biologists use advanced techniques to visualize these structures at the atomic level.

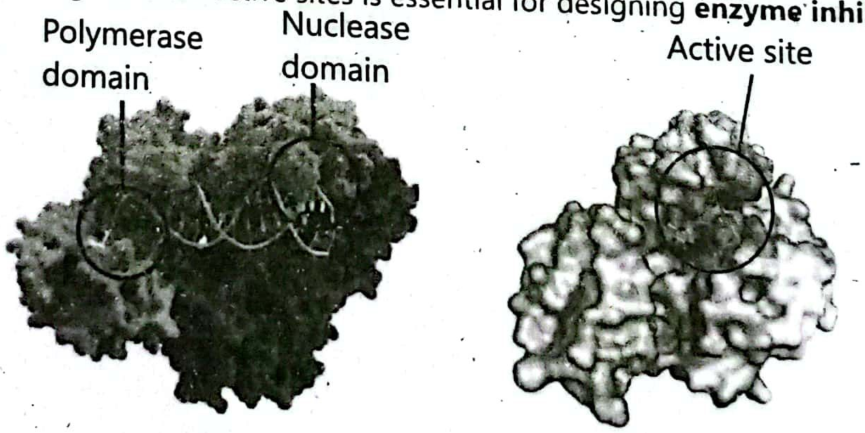
Active Sites: These are specific regions on a protein where substrate molecules bind and undergo a chemical reaction.

Domains: These are distinct structural units within a protein, each having independent and specific functions.

Examples of Functional Identification

- In the enzyme **HIV-1 reverse transcriptase**, structural studies have helped in identifying:
  - The **polymerase domain**, responsible for synthesizing DNA.
  - The **RNase H domain**, which breaks down the RNA strand of RNA-DNA hybrids.

  Understanding the precise location and structure of these domains has significantly contributed to the development of **antiviral drugs** that specifically target them.
- In **serine proteases**, the 3D structure reveals a **well-defined active site** that is responsible for breaking peptide bonds. Knowledge of such active sites is essential for designing **enzyme inhibitors**.



(a) (b)

Fig. 7.1: (a) 3D structure of HIV-1 reverse transcriptase: (b) 3D structure of serine protease

3. How does structural biology assist in identifying drug targets for disease treatment?

Ans. Understanding Drug Targets

Structural biology plays a vital role in the field of medical research, particularly in **drug discovery and development**. It enables scientists to identify potential **drug targets**—typically proteins involved in disease processes—by studying their 3D structures.

A **drug target** is a specific molecule in the body, often a protein, to which a drug binds in order to bring about a therapeutic effect.

The detailed structural information allows researchers to identify **binding sites** where drugs can attach and inhibit the function of the disease-causing molecule.

Application in COVID-19 Research

A notable example is the study of **SARS-CoV-2**, the virus responsible for COVID-19.

Scientists used structural biology to examine the **spike protein** of the virus, which plays a crucial role in helping the virus **enter human cells**.

Knowing the 3D structure of the spike protein enabled scientists to identify it as a key **drug target**.

This information facilitated the **design of vaccines and antiviral drugs** that can block the spike protein, thereby preventing the virus from infecting more cells.

4. How does structural biology help in understanding host–pathogen interactions?

Ans. Insights into Host–Pathogen Interactions

Structural biology is an essential tool for understanding how pathogens such as viruses and bacteria interact with the host organism. These interactions are central to infection and disease progression.

- By analyzing the 3D structures of both **pathogen molecules** and **host cell receptors**, scientists can observe how a pathogen **attaches to** and **enters** host cells.
- This process is known as **host–pathogen interaction**, and understanding it is crucial for developing therapeutic interventions.

Example of Coronavirus

A practical example is the interaction between the **spike protein** of the coronavirus and a **receptor protein** on human cells:

- The spike protein on the virus binds to a specific receptor on the surface of human cells, facilitating viral entry.
- Structural biology studies have revealed the exact mode of this binding interaction.

Based on this information, scientists can design **drugs that bind to the receptor proteins** and block the virus from entering the cell, thus **inhibiting infection**.

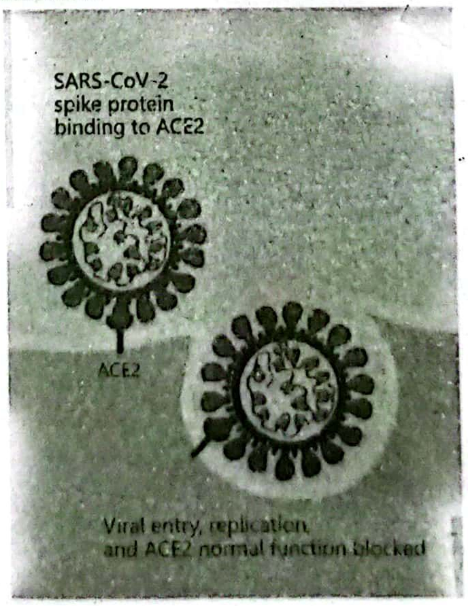


Fig. 7.2: Mechanism of corona virus binding with receptor of human cell.

5. What role does structural biology play in identifying protein misfolding and its relation to diseases?

Ans. Protein Folding and Its Importance

Proteins must fold into specific three-dimensional shapes to perform their biological functions. Incorrect or abnormal folding of proteins can result in **loss of function** or **toxic gain of function**, which is associated with several diseases.

Understanding Misfolding Through Structural Biology

Structural biology provides essential insights into:

- **Folding pathways** of proteins.
- **Mechanisms of misfolding** and how these lead to the development of diseases.

Diseases Caused by Protein Misfolding

Several human diseases are caused by **misfolded proteins**, including:

- **Cystic fibrosis**
- **Parkinson's disease**
- **Alzheimer's disease**

By understanding how misfolding occurs at the molecular level, structural biology aids in the development of **therapeutic strategies** to correct or prevent such misfolding, offering hope for treatment and prevention of these disorders.

mCQs ✓

1. **What does structural biology primarily study?**
A) Chemical reactions
B) 3D structures of macromolecules ✓
C) Blood cell formation D) Genetic mutations

2. **Which of the following is NOT typically studied in structural biology?**
A) Protein structures B) Nucleic acid structures
C) Cellular respiration ✓ D) Ligand interactions

3. **Structural biology provides insights at which level?**
A) Molecular B) Cellular
C) Atomic ✓ D) Organism

4. **What can be determined by studying the 3D structure of a protein?**
A) Temperature tolerance B) pH balance
C) Active sites and domains ✓
D) Color of molecule

5. **What are domains in proteins?**
A) Genetic mutations B) Regions of lipids
C) Structural units with independent functions ✓
D) pH-sensitive areas

6. **Which enzyme's domains were identified using structural biology in HIV research?**
A) Hemoglobin
B) HIV-1 reverse transcriptase ✓
C) DNA polymerase D) Pepsin

7. **The polymerase domain of HIV-1 reverse transcriptase helps in:**
A) Breaking RNA B) Synthesizing DNA ✓
C) Forming lipids D) Protein degradation

8. **The RNase H domain is responsible for:**
A) Cutting proteins
B) RNA strand degradation ✓
C) DNA synthesis D) Lipid metabolism

9. **What does the active site of serine proteases do?**
A) Join proteins
B) Break peptide bonds ✓
C) Create glucose D) Transport oxygen

10. **Which of the following is a major application of structural biology in medicine?**
A) Photosynthesis B) Protein synthesis
C) Drug design ✓ D) Cell division

11. **A drug target is usually:**
A) A DNA molecule B) A protein ✓
C) A carbohydrate D) A lipid

12. **How does structural biology help in drug design?**
A) By modifying hormones
B) By identifying the chemical bonds
C) By showing 3D structures of drug targets ✓
D) By increasing blood flow

13. **What protein was studied in COVID-19 to develop drugs?**
A) Hemoglobin B) Actin
C) Spike protein ✓ D) Myosin

14. **What does the spike protein of SARS-CoV-2 do?**
A) Helps virus replication
B) Prevents immune response
C) Enables viral entry into human cells ✓
D) Breaks down proteins

15. **What is meant by host–pathogen interaction?**
A) Virus replication B) Drug response
C) Interaction between pathogen and host cells ✓
D) Cell signaling

16. **Which tools help visualize host–pathogen interactions at atomic level?**
A) Light microscopy

B) Structural biology techniques ✓
C) Blood tests
D) Genetic engineering

17. Which viral component binds to the receptor protein on human cells?
A) Viral DNA
B) Spike protein ✓
C) Lipid envelope
D) Capsid

18. Blocking a receptor can:
A) Increase infection
B) Cause protein synthesis
C) Stop viral entry ✓
D) Alter genetic code

19. What is essential for protein functionality?
A) Temperature
B) Water availability
C) Correct folding ✓
D) Size of protein

20. Misfolded proteins can cause:
A) Immunity
B) Diseases ✓
C) Photosynthesis
D) Energy production

21. What field helps understand protein misfolding?
A) Biophysics
B) Structural biology ✓
C) Ecology
D) Biotechnology

22. Which of the following diseases is associated with protein misfolding?
A) Malaria
B) Parkinson's ✓
C) Influenza
D) Diabetes Type 1

23. Alzheimer's disease is linked to:
A) Viral infection
B) Nutritional deficiency
C) Misfolded proteins ✓
D) Overhydration

24. What does structural biology reveal about misfolding?
A) Folding temperature
B) Folding pathways and errors ✓
C) Protein color
D) DNA sequences

25. Which technique is essential in structural biology
A) MRI
B) CT scan
C) X-ray crystallography ✓
D) Ultrasound

26. Ligands in structural biology are:
A) Membranes
B) Hormones
C) Molecules that bind to macromolecules ✓
D) Organs

27. Why is 3D structure important for understandin function?
A) It shows chemical reactions
B) It reveals movement
C) Shape determines how a molecule works ✓
D) It gives color to molecules

28. Structural biology contributes most to:
A) Environmental science
B) Disease diagnosis
C) Molecular drug development ✓
D) Agricultural planning

29. Protein misfolding often results in:
A) Higher energy levels
B) Inactive or harmful proteins ✓
C) Rapid healing
D) Enzyme activation

30. Structural biology allows scientists to:
A) Predict weather
B) Grow crops
C) Visualize molecules in 3D ✓
D) Study population growth

SQ

1. What does structural biology study?

Ans. Structural biology studies the three-dimensional (3D) structures of macromolecules such as proteins and nucleic acids. It investigates these structures at the atomic level to understand their function and interactions.

2. Why is the 3D structure of macromolecules important?

Ans. The 3D structure reveals the precise shape, environment, and arrangement of atoms in a biomolecule. This helps in understanding how it functions and interacts with other molecules.

3. What kind of information does structural biology provide about biomolecules?

Ans. Structural biology provides detailed insights into the structure, function, dynamics, and interactions of biomolecules. This information is crucial for understanding biological processes at the molecular level.

4. What are active sites in proteins?

Ans. Active sites are specific regions in proteins where substrate molecules bind to carry out a chemical reaction. These sites are critical for the protein's biological function.

5. What are domains in macromolecules?

Ans. Domains are distinct structural units within macromolecules that often perform independent functions. Each domain typically has a specific role in the overall activity of the molecule.

6. How does structural biology help identify active sites and domains?

Ans. By analyzing the 3D structures of proteins, scientists can locate the exact position and shape of active sites and domains. This knowledge aids in understanding how these regions contribute to function.

7. How did structural biology help in studying HIV-1 reverse transcriptase?

Ans. It revealed the locations of the polymerase domain and RNase H domain within the enzyme. This helped scientists design antiviral drugs that specifically target those domains.

8. **What is the role of the active site in serine proteases?**

Ans. The active site in serine proteases is responsible for breaking peptide bonds. Structural studies show this site is highly specific and essential for the enzyme's function.

9. **What is a drug target in medical research?**

Ans. A drug target is a specific molecule, usually a protein, involved in a disease process that a drug can bind to and inhibit. Identifying drug targets is crucial for designing effective treatments.

10. **How does structural biology help in identifying drug targets?**

Ans. It provides detailed 3D images of disease-related proteins, helping scientists identify binding sites for potential drugs. This allows for the rational design of molecules that can inhibit disease progression.

11. **How was structural biology used during the COVID-19 pandemic?**

Ans. Scientists studied the spike protein of SARS-CoV-2 to understand how it enters human cells. The structural data helped identify it as a drug target for vaccines and therapeutic drugs.

12. **What is the function of the spike protein in SARS-CoV-2?**

Ans. The spike protein enables the virus to attach to and enter human cells. Blocking this protein can stop the virus from spreading in the body.

13. **What are host–pathogen interactions?**

Ans. Host–pathogen interactions refer to how disease-causing organisms, like viruses or bacteria, attach to and affect the host's cells. Understanding these interactions helps in developing targeted treatments.

14. **How does structural biology contribute to understanding host–pathogen interactions?**

Ans. It allows scientists to visualize how pathogen proteins interact with host cell receptors at the molecular level. This helps in designing drugs that can block these interactions.

15. **What role do human cell receptors play in viral infections?**

Ans. Receptors on human cells act as entry points for viruses like SARS-CoV-2. Structural biology helps identify these receptors and how they interact with viral proteins.

16. **How can blocking a receptor prevent viral infection?**

Ans. If a drug binds to the receptor and blocks it, the virus cannot attach or enter the host cell. This interrupts the infection cycle and reduces the spread of the virus.

17. **Why is protein folding important?**

Ans. Correct folding is essential for a protein's functionality. Misfolded proteins can lose their function or become toxic, leading to various diseases.

18. **How does structural biology help in understanding protein misfolding?**

Ans. It provides insights into the folding pathways and mechanisms that lead to misfolding. This understanding is important for finding ways to prevent or correct these errors.

19. **Name some diseases caused by protein misfolding.**

Ans. Diseases like cystic fibrosis, Parkinson's disease, and Alzheimer's disease are linked to protein misfolding. These conditions result from proteins not achieving their proper shapes.

20. **How can knowledge of protein misfolding help in medicine?**

Ans. Understanding misfolding helps researchers develop therapies to correct or prevent it. This can lead to treatments that address the root cause of certain diseases.

7.2 X-RAY CRYSTALLOGRAPHY

6. What is X-ray Crystallography and What is Its Historical Significance and Applications?

Ans. Introduction to X-ray Crystallography

X-ray crystallography is a scientific technique developed in 1912 by William Henry Bragg and his son William Lawrence Bragg. For their groundbreaking contribution, they were awarded the Nobel Prize in Physics in 1915. The technique has since revolutionized the field of structural analysis of matter.

Working Principle

In X-ray crystallography, a beam of X-rays is directed at a crystal. The atoms and molecules within the crystal scatter or diffract the X-rays in specific, predictable directions. This diffraction results in a pattern from which scientists can determine the arrangement of atoms. By analyzing the angles and intensities of these diffracted beams, a three-dimensional (3D) image of the electron density within the crystal is produced. The resulting electron density map is then used to construct the 3D structure of the molecule.

Applications of X-ray Crystallography

X-ray crystallography has a broad range of applications in science and industry. It is used to study:

- **Minerals and salts**
- **Metals**
- **Biomolecules like proteins, carbohydrates, and nucleic acids**
- **Vitamins and other small molecules**

Its primary significance lies in revealing atomic-level details about the structure of substances, which is essential in fields like biochemistry, pharmacology, and materials science.

7. How is X-ray Crystallography Used to Determine Protein Structure?

Ans. X-ray crystallography is one of the most reliable techniques for determining the 3D structure of proteins at atomic resolution. The process involves several sequential steps, each of which plays a crucial role in deriving structural information.

Step 1: Protein Crystallization

The first step is **protein crystallization**, where a purified protein is converted into a solid crystal form. Crystals are necessary because they organize protein molecules into a regular and repeating pattern, which is essential for producing a clear and interpretable diffraction pattern during the X-ray analysis.

To achieve crystallization:

- The protein is slowly mixed with specialized solutions.
- These solutions cause the protein molecules to aggregate in an orderly fashion.
- The crystallization process can take from a few hours to several days or weeks.
- Critical parameters like **temperature, pH, and salt concentration** must be carefully controlled.
- Once a stable and clear crystal is formed, it is ready for the next stage of analysis.

Step 2: Production of Diffraction Pattern

The next step is to mount the crystal on an X-ray machine. When the X-ray beam strikes the crystal at various angles:

- The atoms within the crystal scatter the X-rays.
- This scattering results in a **diffraction pattern**, which appears as a series of spots on a detector.
- The positions and intensities of these spots provide key information about the internal structure of the crystal.

Step 3: Creating the Density Map

Using the data from the diffraction pattern, scientists construct a **density map**:

- The diffraction data, particularly the angles and intensities of the scattered rays, is mathematically analyzed.
- This data helps to map the distribution of electrons within the crystal, forming what is called an **electron density map**.

Step 4: Determination of Protein Structure

Finally, advanced computational software is used to:

- Analyze the electron density map.
- Transform the map into a 3D atomic model of the protein.
- Reveal detailed information about the shape, folding, and atomic interactions of the protein.

This structural knowledge is vital for understanding protein function, designing drugs, and advancing biotechnology.

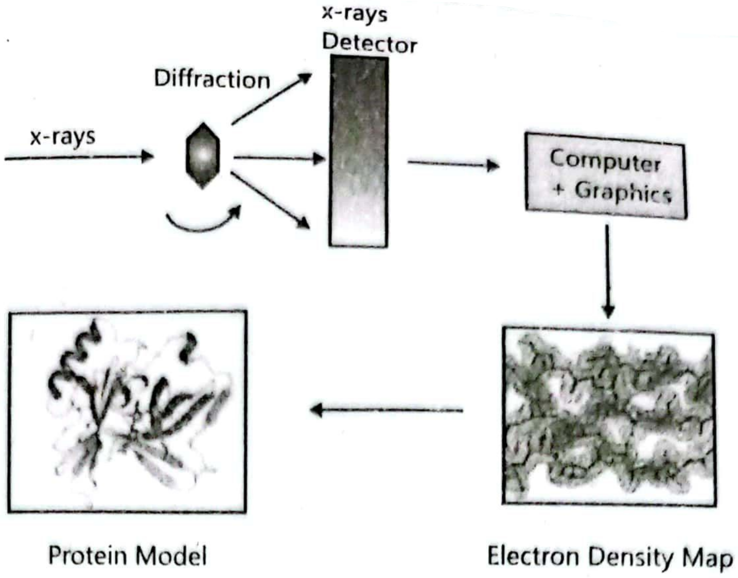


Fig. 7.3: Schematic representation of X-ray crystallography

mCQ's ✓

1. **Who developed X-ray crystallography?**
A) Watson and Crick B) Robert Hooke
C) William Henry Bragg and William Lawrence Bragg ✓
D) Alexander Fleming

2. **In which year was the Nobel Prize awarded to the Braggs for X-ray crystallography?**
A) 1905 B) 1920
C) 1915 ✓ D) 1932

3. **What does X-ray crystallography primarily help to determine?**
A) Chemical reactions
B) 3D structure of molecules ✓
C) Genetic mutations D) Energy levels of atoms

4. **What causes diffraction of X-rays in a crystal?**
A) Electrons in a vacuum B) Magnetic fields
C) Atoms in the crystal ✓ D) Water molecules

5. **What kind of image is created from X-ray diffraction?**
A) Heat map B) DNA sequence
C) Electron density map ✓
D) Ultrasound scan

6. **Which of the following substances can be studied using X-ray crystallography?**
A) DNA only B) Metals only
C) Proteins, salts, metals, vitamins ✓
D) Liquids only

7. **Why are crystals essential for X-ray crystallography?**
A) They change X-ray wavelength
B) They organize molecules in a repeating pattern ✓
C) They amplify signals
D) They produce magnetic fields

8. **What is the first step in determining protein structure using X-ray crystallography?**
A) Protein crystallization ✓
B) Data analysis
C) Electron microscopy D) DNA sequencing

9. **What is protein crystallization?**
A) Breaking protein into fragments
B) Converting purified protein into crystal form ✓
C) Melting proteins
D) Coloring proteins with dye

10. **What conditions must be controlled during protein crystallization?**
A) Light intensity only B) Voltage and pressure
C) Temperature, pH, and salt concentration ✓
D) Oxygen level

11. **What is a diffraction pattern?**
A) A light reflection B) A sound wave
C) A series of spots formed by X-ray scattering ✓
D) A UV absorption graph

12. **How is the diffraction pattern used?**
A) To purify the protein
B) To create an electron density map ✓
C) To test chemical bonds D) To separate crystals

13. **What information do the diffraction spots contain?**
A) Sound wave frequencies
B) Color spectrum data
C) Angles and intensities of scattered X-rays ✓
D) Heat levels

14. **What is done after generating the density map?**
A) The crystal is melted
B) Data is analyzed with computational tools ✓
C) The protein is frozen D) The protein is dyed

15. **What is the final product of X-ray crystallography analysis?**
A) DNA fingerprint
B) 3D atomic model of the molecule ✓
C) X-ray photograph D) Protein color image

16. **What does the electron density map show?**
A) Pressure distribution B) Light absorption
C) Likely positions of electrons ✓
D) Sound waves

17. **What role does computation play in X-ray crystallography?**
A) Produces radiation
B) Analyzes diffraction data ✓
C) Grows crystals D) Filters proteins

18. **What is the purpose of forming a high-quality crystal?**
A) To speed up reactions
B) To ensure clear diffraction patterns ✓
C) To improve protein flavor
D) To sterilize the protein

19. **Which biomolecules have been studied using X-ray crystallography?**
A) Only carbohydrates
B) Proteins, nucleic acids, carbohydrates ✓
C) Only vitamins D) Only salts

20. **Why is the 3D structure of a protein important?**
A) It determines its color
B) It reveals its function and interactions ✓
C) It tells the age of the protein
D) It reduces its size

SQ

1. **What is X-ray crystallography?**

Ans. X-ray crystallography is a scientific technique used to determine the 3D structure of molecules by analyzing how X-rays are diffracted by a crystal. It helps scientists visualize the arrangement of atoms within the crystal.

2. **Who developed X-ray crystallography and when?**

Ans. X-ray crystallography was developed in 1912 by William Henry Bragg and his son William Lawrence Bragg. They received the Nobel Prize in Physics in 1915 for their pioneering work.

3. **What happens when X-rays strike a crystal in crystallography?**

Ans. When X-rays strike a crystal, the atoms inside the crystal diffract the X-ray beams. This diffraction produces a pattern that can be analyzed to reveal the molecule's structure.

4. **What kind of image is created from the diffraction of X-rays?**

Ans. The diffraction of X-rays creates a 3D electron density map. This map shows the likely locations of electrons and helps build the 3D structure of the molecule.

5. **What types of substances can be studied using X-ray crystallography?**

Ans. X-ray crystallography can be used to study a wide range of substances including minerals, salts, metals, proteins, carbohydrates, nucleic acids, and vitamins. It is highly versatile in application.

6. **Why are crystals necessary in X-ray crystallography?**

Ans. Crystals are necessary because they arrange molecules in a regular, repeating pattern. This orderliness is essential for producing a clear diffraction pattern.

7. **How is protein crystallization carried out?**

Ans. Protein crystallization involves mixing purified proteins with special solutions that encourage the molecules to stick together in an orderly way. This process often requires controlled temperature, pH, and salt conditions.

8. **How long can the crystallization process take?**

Ans. The crystallization process can take hours, days, or even weeks. It depends on various factors such as protein type and environmental conditions.

9. **What is a diffraction pattern in X-ray crystallography?**

Ans. A diffraction pattern is a series of spots formed when X-rays are scattered by the crystal. The pattern holds information about the internal atomic structure of the molecule.

10. **What is an electron density map?**

Ans. An electron density map is a 3D representation of where electrons are most likely located in a crystal. It i generated from the diffraction pattern data.

11. **What is the final step in protein structure determination using X-ray crystallography?**

Ans. The final step is mathematical analysis using computational tools to convert the electron density map into a detailed 3D model of the protein. This model reveals the atomic arrangement.

12. **Why is determining the 3D structure of proteins important?**

Ans. Determining the 3D structure of proteins is crucial for understanding their functions and interactions. It is also vital for drug design and medical research.

13. **What role does computational analysis play in X-ray crystallography?**

Ans. Computational analysis is used to interpret the diffraction data and create the final 3D structure of the molecule. It involves complex mathematical modeling.

14. **What does the diffraction pattern tell scientists?**

Ans. The diffraction pattern tells scientists about the angles and intensities at which X-rays are scattered. This information is key to building the electron density map.

15. **What scientific fields benefit from X-ray crystallography?**

Ans. Fields such as biochemistry, pharmacology, structural biology, and materials science benefit from X-ray crystallography. It helps them understand molecular structures and develop new technologies.

7.3 COMPUTATIONAL BIOLOGY

8. What is Computational Biology and why is it important in modern science?

Ans. **Definition and Scope**

Computational biology is an interdisciplinary field that uses computational techniques and tools to solve biological problems. It integrates knowledge from biology, computer science, mathematics, and statistics to analyze and interpret biological data.

Importance

The importance of computational biology lies in its ability to handle large datasets, uncover hidden patterns, and generate predictive models that can lead to new biological insights and applications. With the explosion of biological data from various fields such as genomics and proteomics, computational biology plays a crucial role in organizing, analyzing, and interpreting this information efficiently and effectively.

9. What are the major areas of study within Computational Biology?

Ans.

1. **Genomics**

Genomics is the study of genomes, which are the complete set of DNA within a single cell of an organism. It involves sequencing, assembling, and analyzing the function and structure of genomes. Genomics helps in understanding genetic variations, gene functions, and evolutionary relationships. It is foundational in uncovering the genetic blueprint of organisms.

2. **Proteomics**

Proteomics is the large-scale study of proteins, including their structures and functions. Since proteins are essential molecules responsible for numerous biological processes, proteomics aims to map the entire set of proteins (the proteome) produced by an organism. It also focuses on understanding protein interactions and their roles in cellular mechanisms.

3. **Bioinformatics**

Bioinformatics is the application of computer technology to manage and analyze biological data. It involves the use of specialized software tools to store, retrieve, and analyze DNA, RNA, and protein sequences. Bioinformatics enables researchers to make sense of complex biological information and supports other areas like genomics and proteomics.

10. What are the main applications of Computational Biology?

Ans. Computational biology has vast and diverse applications across multiple domains of biology and medicine. Some key applications include:

1. **Drug Discovery**

Computational biology significantly contributes to drug discovery by identifying potential drug targets and simulating the effects of drugs on biological systems. It accelerates the drug development process by predicting how drugs interact with proteins and other molecules, thus reducing the time and cost involved in laboratory-based testing.

2. Genetic Research

Through the analysis of DNA sequences, computational biology helps identify genetic variations associated with diseases. It aids in understanding the genetic basis of various health conditions and supports the development personalized medicine by tailoring treatments based on an individual's genetic profile.

3. Evolutionary Biology

Computational tools are widely used in evolutionary biology to compare genetic information across different species. These tools help reconstruct evolutionary relationships and provide insights into the processes that drive evolution. This contributes to our understanding of biodiversity and the genetic linkage between species.

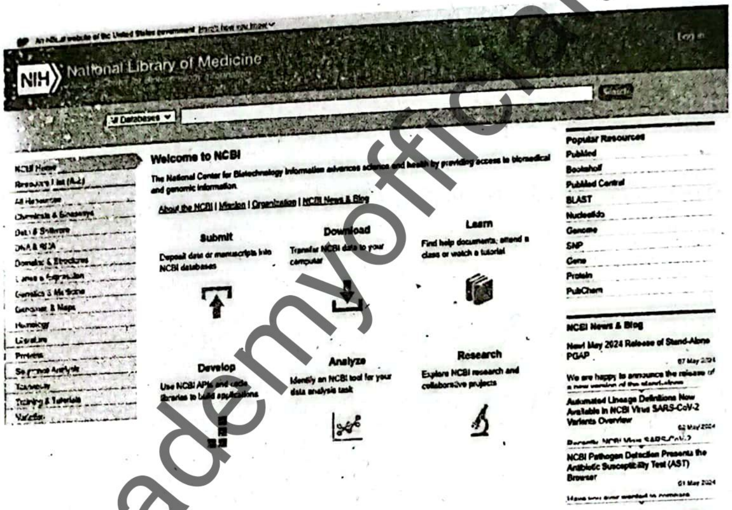
11. What are the key databases used in Computational Biology and what are their roles?

Ans. In computational biology, several major databases are used to store and access biological data, particularly related to nucleic acids and proteins. These include:

1. **GenBank**

   Website: https://www.ncbi.nlm.nih.gov/nuccore/

   GenBank is a comprehensive public database of nucleotide sequences along with supporting bibliographic and biological annotations. It offers access to a vast repository of DNA sequences from various organisms, thereby facilitating genetic research and comparative genomics.



Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

Deposit data or manuscripts into NCBI databases

Transfer NCBI data to your computer

Find help documents, attend a class or watch a tutorial

Use NCBI APIs and code libraries to build applications

Identify an NCBI tool for your data analysis task

Explore NCBI research and collaborative projects

New! May 2024 Release of Stand-Alone PGAP
07 May 2024

We are happy to announce the release of a new version of the stand-alone

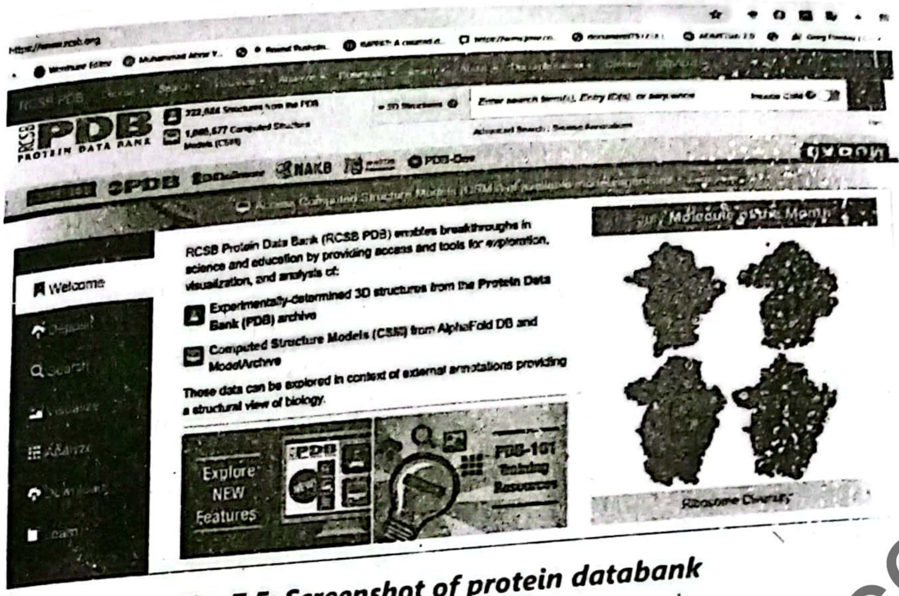
Automated Lineage Definitions Now Available in NCBI Virus SARS-CoV-2 Variants Overview
02 May 2024

NCBI Pathogen Detection Presents the Antibiotic Susceptibility Test (AST) Browser
01 May 2024

Fig. 7.4: Screenshot of GenBank database

2. **Protein Data Bank (PDB)**

   The PDB provides 3D structural data of large biological molecules, including proteins and nucleic acids. It is essential for studying the structures of macromolecules, understanding their functions, and aiding in drug design by targeting specific protein structures.



RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

Experimentally-determined 3D structures from the Protein Data Bank (PDB) archive

Computed Structure Models (CSM) from AlphaFold DB and ModelArchive

These data can be explored in context of external annotations providing a structural view of biology.

Fig. 7.5: Screenshot of protein databank

3. **Ensembl**

Ensembl is a genome browser that provides information on genome sequences, gene models, and comparative genomics for various species. It is a powerful resource for accessing and visualizing genomic data, supporting extensive research in genomics and evolutionary biology.

Q 12. Which key algorithms are used in Computational Biology for data analysis?

Ans. In addition to databases, computational biology relies on powerful algorithms for analyzing biological data. Two mportant algorithms include:

1. **BLAST (Basic Local Alignment Search Tool)**

BLAST is widely used for comparing primary biological sequence information, such as amino acid sequences of proteins or nucleotide sequences of DNA. It helps in identifying homologous sequences, predicting the functions of nknown genes, and studying evolutionary relationships between organisms.

2. **FASTA**

FASTA is another sequence alignment tool that compares a query sequence to a database of sequences to find egions of similarity. It is used in searching both protein and nucleotide databases, identifying sequence homology, nd analyzing sequence alignments. It plays a crucial role in understanding functional and structural similarities mong genes and proteins.

MCQs ✓

1. **What is computational biology?**
A) The study of computers B) The study of chemicals
C) The use of computational techniques to solve biological problems ✓
D) The study of astronomical systems

2. **Which fields are integrated in computational biology?**
A) Chemistry and Physics
B) Biology, Computer Science, Mathematics, and Statistics ✓
C) Physics and Engineering
D) Geography and Astronomy

3. **Which of the following is a major area of computational biology?**
A) Genomics ✓ B) Thermodynamics
C) Botany D) Geology

4. **What does genomics study?**
A) Protein folding
B) The complete set of DNA in a cell ✓
C) Enzyme reactions D) Cell membranes

5. **What is the aim of proteomics?**
A) Study of soil
B) Study of protein structures and functions ✓
C) Study of climates
D) Study of photosynthesis

6. **Bioinformatics is primarily used to manage and analyze:**
A) Chemical equations
B) Biological sequence data ✓
C) Mathematical problems
D) Environmental data

7. Which of the following is a key application of computational biology?
A) Painting B) Welding
C) Drug discovery✓ D) Agriculture only

8. Which area uses computational biology to understand disease causes?
A) Physics B) Genetic research✓
C) Cooking D) Botany

9. What is the purpose of evolutionary biology in this context?
A) Study of weather B) Study of plants
C) Comparing genes to understand evolution✓
D) Study of rocks

10. What is GenBank used for?
A) Selling genes B) Data storage
C) Public database for nucleotide sequences✓
D) Banking information

11. Which website hosts GenBank?
A) www.google.com B) www.geneticsbank.org
C) www.ncbi.nlm.nih.gov/nuccore/✓
D) www.genestudy.com

12. The Protein Data Bank provides data on:
A) RNA transcription B) Cell walls
C) 3D structures of proteins and nucleic acids✓
D) Fossil records

13. What is Ensembl primarily used for?
A) Genome browsing and gene models✓
B) Weather tracking
C) Plant growth analysis D) Bacterial growth rates

14. Which tool compares biological sequence information?
A) Word processor B) BLAST✓
C) Photoshop D) Calculator

15. What does BLAST stand for?
A) Biochemical Learning and Science Tool
B) Basic Local Alignment Search Tool✓
C) Biological Layout Analysis System Tool
D) Biological Link Algorithm Study Tool

16. FASTA is used for:
A) Food analysis
B) Sequence alignment✓
C) Making proteins
D) Writing chemical formulas

17. Which area studies all proteins in a cell or organism?
A) Genomics B) Proteomics✓
C) Anatomy D) Pharmacology

18. Which database includes bibliographic and biological annotations?
A) Ensembl B) GenBank✓
C) Excel D) PDB

19. The PDB is important for:
A) Climate predictions B) Tree classification
C) Drug design and structural analysis✓
D) Geological mapping

20. Ensembl helps in:
A) Translating languages
B) Accessing and visualizing genome data✓
C) Cooking D) Marketing

21. BLAST is helpful in identifying:
A) Homologous sequences✓
B) Food chains
C) Cell wall components D) Vitamins

22. FASTA is mainly used to search:
A) Internet history
B) Protein and nucleotide databases✓
C) Computer errors D) Weather records

23. Which ofhe following is not a key database i computational biology?
A) GenBank B) Britannica✓
C) PDB D) Ensembl

24. Which algorithm is used for fast sequence comparison?
A) Windows B) Paint
C) FASTA✓ D) Excel

25. Why is computational biology considered interdisciplinary?
A) It only uses biology
B) It integrates biology, computer science, mathematics, and statistics✓
C) It uses engineering and law
D) It combines art and science

26. How does computational biology suppor personalized medicine?
A) By analyzing individual genetic data✓
B) By making herbal medicine
C) By farming
D) By providing surgery kits

27. What is the full form of PDB?
A) Protein Data Bundle B) Protein Data Bank✓
C) Protein Digital Base D) Protein Design Bureau

28. Which is NOT an application of computational biology?
A) Drug discovery B) Genetic research
C) Road construction✓ D) Evolutionary studies

29. 1Which tool is best for understanding gen models?
A) FASTA B) BLAST
C) Ensembl✓ D) PowerPoint

30. What is a key advantage of computational biology?
A) It grows plants B) It fixes computers
C) It processes large biological datasets efficiently✓
D) It treats animals directly

1. **What is computational biology?**

Ans. Computational biology is an interdisciplinary field that applies computational techniques to solve biological problems. It combines biology, computer science, mathematics, and statistics to analyze biological data.

2. **Why is computational biology important?**

Ans. It is important because it can handle large biological datasets, identify hidden patterns, and generate predictive models. These capabilities lead to new insights and applications in biology and medicine.

3. **What is genomics?**

Ans. Genomics is the study of genomes—the complete set of DNA in a cell. It involves sequencing, assembling, and analyzing DNA to understand gene functions and variations.

4. **What does genomics help us understand?**

Ans. Genomics helps in understanding genetic variations, gene functions, and evolutionary relationships. It is crucial for disease research and genetic mapping.

5. **What is proteomics?**

Ans. Proteomics is the large-scale study of proteins, including their structures and functions. It aims to map the proteome of an organism and study protein interactions.

6. **Why are proteins important in biology?**

Ans. Proteins perform a wide range of functions in organisms, including catalyzing reactions and providing structure. Understanding them is essential for studying cell biology and disease.

7. **What is bioinformatics?**

Ans. Bioinformatics involves the application of computer technology to manage and analyze biological data. It is widely used to study DNA, RNA, and protein sequences.

8. **What are the applications of bioinformatics?**

Ans. Bioinformatics is used for storing, retrieving, and analyzing biological sequence data. It also supports drug discovery and genomics research.

9. **How does computational biology aid in drug discovery?**

Ans. It identifies potential drug targets and simulates how drugs interact with biological systems. This speeds up the drug discovery process and reduces costs.

10. **What is the role of computational biology in genetic research?**

Ans. It helps analyze DNA sequences to find genetic variations linked to diseases. This enables the development of personalized medicine.

11. **How does computational biology support evolutionary biology?**

Ans. It compares genetic information across species to understand evolutionary relationships. This helps reconstruct phylogenetic trees and study the process of evolution.

12. **What is GenBank?**

Ans. GenBank is a public database of nucleotide sequences. It includes biological annotations and supports genetic and comparative genomics research.

13. **What type of data is found in GenBank?**

Ans. It provides DNA sequences from various organisms along with related biological and bibliographic information. It is widely used in genomic studies.

14. **What is the Protein Data Bank (PDB)?**

Ans. PDB is a database of 3D structural data for biological molecules like proteins and nucleic acids. It helps scientists study molecular structure and function.

15. **Why is the Protein Data Bank useful in drug design?**

Ans. It provides accurate 3D models of protein structures, allowing researchers to design drugs that target specific molecules. This is essential in precision medicine.

16. **What is Ensembl?**

Ans. Ensembl is a genome browser that provides data on genome sequences, gene models, and comparative genomics. It supports researchers in visualizing and interpreting genomic information.

17. What kind of information does Ensembl offer?

Ans. Ensembl provides access to genome sequences from different species. It also includes gene predictions and evolutionary analysis tools.

18. What is BLAST?

Ans. BLAST (Basic Local Alignment Search Tool) is used to compare biological sequences such as DNA or protein. It finds similarities between sequences and helps in gene function prediction.

19. How does BLAST help in evolutionary biology?

Ans. BLAST identifies homologous sequences across species. This allows scientists to trace gene evolution and relationships between organisms.

20. What is the FASTA algorithm?

Ans. FASTA is a sequence alignment tool that compares a query sequence to a database. It finds regions of similarity to identify gene or protein homology.

21. Why are algorithms like BLAST and FASTA important?

Ans. They automate the comparison of genetic sequences and save time in data analysis. These tools are essential in genomics and proteomics research.

22. What does computational biology reveal through data analysis?

Ans. It reveals patterns, relationships, and functional predictions from biological data. This leads to discoveries in genetics, medicine, and evolutionary biology.

23. How is computational biology interdisciplinary?

Ans. It integrates biology, computer science, mathematics, and statistics. This allows researchers to tackle complex biological questions using diverse methods.

24. How does computational biology contribute to personalized medicine?

Ans. By analyzing genetic data, it identifies individual genetic variations. This helps in designing targeted treatments specific to a patient's genetic profile.

25. What kind of problems can computational biology solve?

Ans. It solves problems related to genome sequencing, protein structure prediction, and disease gene identification. It also supports drug target discovery.

26. How are biological datasets handled in computational biology?

Ans. They are processed using algorithms and computational tools. This helps in storing, retrieving, and analyzing vast amounts of biological information.

27. What is the proteome?

Ans. The proteome is the complete set of proteins produced by an organism. Studying the proteome helps in understanding how proteins function and interact.

28. Why is data visualization important in computational biology?

Ans. Data visualization helps in understanding complex biological information. Tools like Ensembl allow researchers to explore genomic data easily.

29. What is sequence alignment?

Ans. Sequence alignment is the process of arranging sequences to identify regions of similarity. It is used to detect evolutionary and functional relationships.

30. How does computational biology speed up research?

Ans. It allows for the rapid analysis of large datasets using automated tools. This accelerates discoveries in genetics, medicine, and biology.

7.4 SEQUENCE HOMOLOGY

13. What is sequence homology and what are its main types? Explain with examples.

Ans. Definition of Sequence Homology

Sequence homology refers to the similarity between DNA, RNA, or protein sequences due to shared ancestry. Homologous sequences have evolved from a common ancestral sequence.

Types of Sequence Homology

Sequence homology is mainly categorized into two types:

- **Orthologs**: These are sequences found in different species that originated from a common ancestral gene during the process of speciation. Orthologs often retain the same function across different species. For example, a gene in humans and a similar gene in mice performing the same function could be orthologs.
- **Paralogs**: These are sequences found within the same species that originated due to gene duplication events. Although they arise from the same ancestral gene, paralogs can evolve new or diversified functions over time.

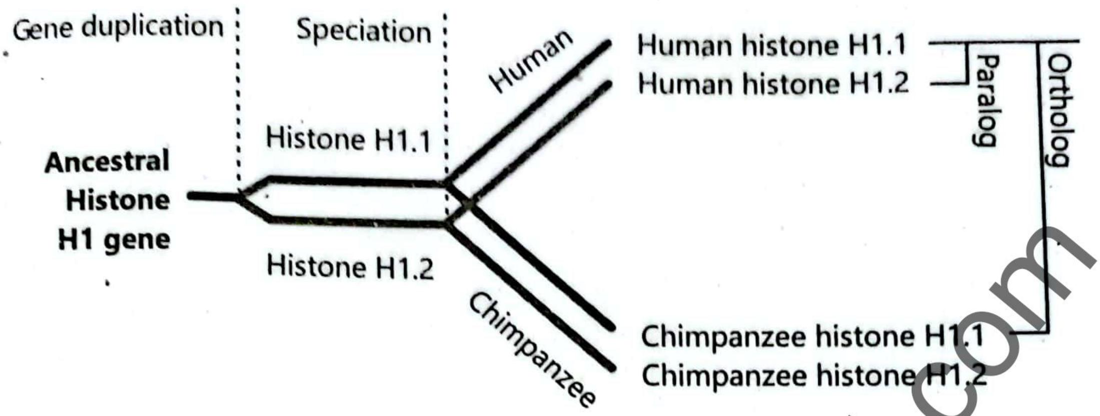


Fig. 7.6. Types of Homologous Sequences

14. How does sequence homology help in understanding evolutionary relationships and gene function?

Ans. Understanding Evolution through Homology

Sequence homology provides deep insights into the evolutionary relationships between organisms. By comparing homologous sequences across species, scientists can infer the evolutionary history and divergence of species over time.

Predicting Gene or Protein Function

Sequence homology also aids in predicting the function of unknown genes or proteins. If an unknown gene or protein is homologous to one with a known function, it is likely that the unknown entity performs a similar function. This functional prediction is fundamental in genomic studies, where many newly identified genes still have unknown roles.

Application in Disease Research

Identifying homologous genes involved in diseases across different species helps scientists understand the underlying mechanisms of diseases. It also plays a vital role in developing medical treatments. For example, homologous genes in model organisms like mice or zebrafish can be studied in laboratory settings to gain insights into human diseases. This comparative approach has greatly advanced genetic research and therapeutic developments.

15. What is structural homology and what is its importance in biological research and drug design?

Ans. Definition of Structural Homology

Structural homology refers to the similarity in the three-dimensional structures of proteins or other macromolecules that arises due to shared ancestry. While two proteins may not share high sequence similarity, their structures may still be alike due to evolutionary conservation.

Function Prediction through Structure

Proteins with similar three-dimensional structures often perform similar functions, even if their sequences differ. The structure of a protein is crucial for understanding how it works, as the shape determines how the protein interacts with other molecules.

Role in Drug Design

Structural homology is especially important in drug design. Many drugs are designed to interact specifically with the 3D structure of proteins. By understanding the structural relationships between proteins, scientists can develop drugs that bind effectively to target proteins, enhancing their therapeutic potential and specificity.

Evolutionary Insights

Studying structural homology also provides insights into the evolutionary processes that shape protein functions and molecular interactions. Proteins that retain similar structures over time offer clues about conserved biological mechanisms and can be key targets in biomedical research.

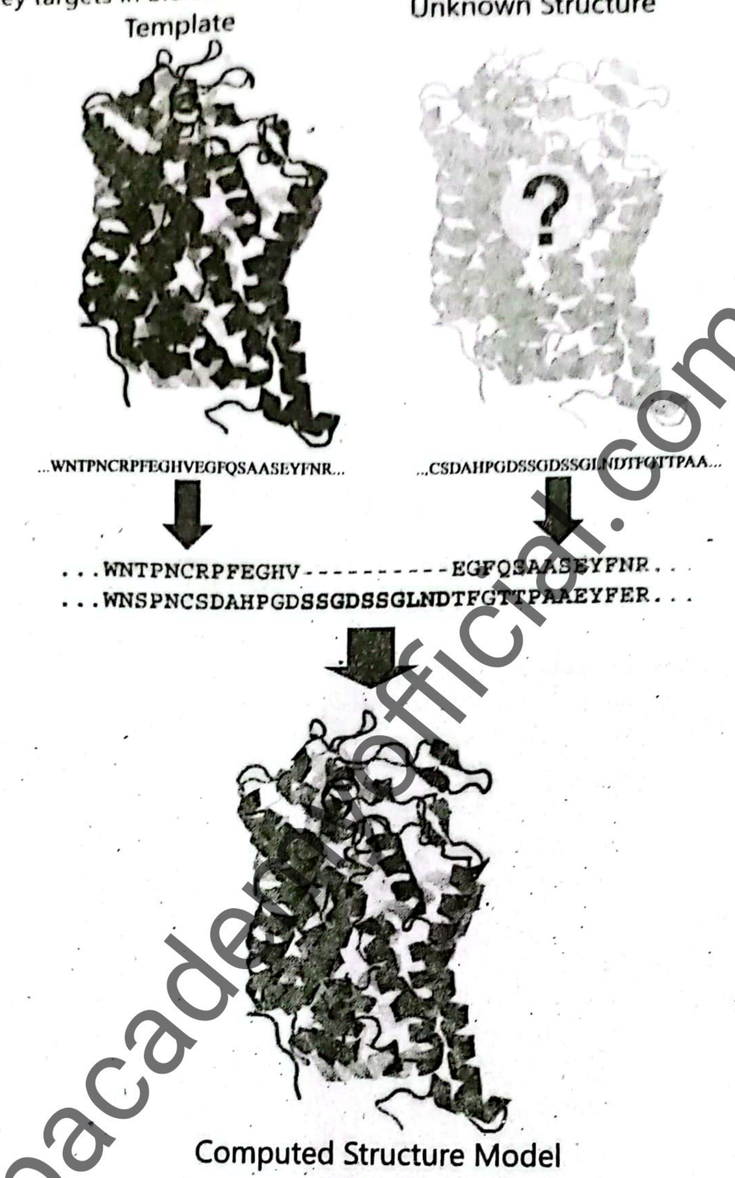


Fig. 7.7: Structural homology of protein

mQs ✓

1. **What does sequence homology refer to?**
   A) Similarity in cell shape
   B) Similarity between DNA, RNA, or protein sequences due to shared ancestry ✓
   C) Similarity in metabolic activity
   D) Similarity in organism size
2. **Orthologous sequences are found in:**
   A) The same species after gene duplication
   B) Different species originating from a common ancestral gene ✓
   C) Different species with unrelated functions
   D) The same species with unrelated genes
3. **Paralogous sequences arise due to:**
   A) Horizontal gene transfer
   B) Speciation
   C) Gene duplication within the same species ✓
   D) Random mutations
4. **Which of the following is true about orthologs?**
   A) They usually have different functions
   B) They often retain the same function across species ✓
   C) They only occur in bacteria
   D) They never have similar sequences
5. **What can sequence homology help infer?**

A) Evolutionary history and gene function✓
B) Protein folding speed
C) Cell membrane composition
D) Nutrient absorption rates

6. **How can homologous genes aid disease research?**
A) By increasing mutation rates
B) By helping understand disease mechanisms across species✓
C) By creating new pathogens
D) By decreasing gene expression

7. **Structural homology refers to similarity in:**
A) Genetic sequences only
B) Three-dimensional protein structures due to shared ancestry✓
C) Cellular organelles
D) Environmental adaptations

8. **Why is protein structure important?**
A) It determines the protein's function✓
B) It only affects color
C) It regulates DNA replication
D) It controls cell shape

9. **How does structural homology assist drug design?**
A) By predicting weather changes
B) By helping design drugs that target specific protein structures✓
C) By increasing mutation rates
D) By degrading drugs faster

10. **Proteins with low sequence similarity can still have:**
A) Identical sequences
B) Similar 3D structures due to structural homology✓
C) Different functions always
D) No relation

11. **Which type of homology involves gene duplication within the same species?**
A) Orthology B) Paralogy✓
C) Xenology D) Analogy

12. **Which of the following best describes orthologs?**
A) Genes duplicated within a species
B) Genes in different species with common ancestry✓
C) Random sequences D) Genes only in bacteria

13. **Sequence homology helps predict:**
A) Unknown gene or protein function✓
B) Cell size
C) Organism lifespan D) Atmospheric conditions

14. **Studying homologous genes in model organisms helps:**
A) Understand human diseases✓
B) Create new species
C) Reduce biodiversity
D) Change gene sequences randomly

15. **Which of the following statements is true?**
A) Paralogs always retain the same function
B) Paralogs can evolve new functions after duplication✓
C) Orthologs evolve new functions rapidly
D) Paralogs are found only in viruses

16. **Structural homology is important because:**
A) It determines protein color
B) It can indicate similar functions even with low sequence similarity✓
C) It prevents mutations D) It controls cell division

17. **How does structural homology help understand evolution?**
A) By studying protein shapes and their conservation✓
B) By comparing organism size
C) By analyzing metabolic rates
D) By sequencing RNA only

18. **A protein's 3D structure can help in:**
A) Predicting its function and interactions✓
B) Determining its color only
C) Changing its DNA sequence
D) Measuring cellular respiration

19. **Which term refers to genes with similar sequences due to common ancestry but in different species?**
A) Paralog B) Ortholog✓
C) Analog D) Isoform

20. **The study of sequence homology can contribute to:**
A) Personalized medicine development✓
B) Changing atmospheric gases
C) Altering soil composition
D) Increasing water evaporation

---

1. **What is sequence homology?**

Ans. Sequence homology refers to the similarity between DNA, RNA, or protein sequences due to shared ancestry. It indicates that the sequences have evolved from a common ancestral gene.

2. **What are orthologs?**

Ans. Orthologs are homologous sequences found in different species that originated from a common ancestral gene during speciation. They often retain similar functions across those species.

3. **What are paralogs?**

Ans. Paralogs are homologous sequences within the same species that arise due to gene duplication. They may evolve new functions even though they come from the same ancestral gene.

4. **How does sequence homology help in evolutionary studies?**

Ans. Sequence homology allows scientists to compare genetic sequences across species to infer evolutionary relationships. It helps reconstruct the evolutionary history and divergence of species.

5. **How can sequence homology assist in predicting gene function?**

Ans. If an unknown gene or protein shows homology to a gene or protein with a known function, it is likely to have a similar role. This helps researchers infer functions of newly discovered genes.

6. **Why is identifying homologous genes important in disease research?**

Ans. Homologous genes involved in diseases across species help scientists understand disease mechanisms. Studying these genes in model organisms can lead to insights relevant to human diseases.

7. **What is structural homology?**

Ans. Structural homology refers to the similarity in the three-dimensional structures of proteins or macromolecules due to shared ancestry. It can exist even when sequence similarity is low.

8. **Why is protein structure important?**

Ans. The three-dimensional structure of a protein determines its function. Understanding protein structure helps in predicting how proteins interact and perform biological roles.

9. **How is structural homology useful in drug design?**

Ans. Drugs often target specific protein structures, so understanding structural homology helps in designing drugs that bind effectively to target proteins. This enhances drug specificity and efficacy.

10. **Can proteins with different sequences have similar structures?**

Ans. Yes, proteins with different sequences can have similar three-dimensional structures due to structural homology. Such similarities often indicate related functions.

11. **What evolutionary insights can be gained from studying structural homology?**

Ans. Studying structural homology reveals how protein functions and interactions have been conserved or adapted through evolution. It helps understand the molecular basis of biological processes.

12. **How do orthologs differ from paralogs in terms of function?**

Ans. Orthologs usually retain the same function across species, while paralogs may develop new or specialized functions within the same species after gene duplication.

13. **What role does sequence homology play in personalized medicine?**

Ans. By analyzing genetic variations through homology, researchers can better understand the genetic basis of diseases. This knowledge aids in developing personalized treatments tailored to individual genetic profiles.

14. **Why are model organisms important in studying homologous genes?**

Ans. Model organisms share homologous genes with humans, making them useful for studying gene functions and disease mechanisms. Insights from these organisms often apply to human biology.

15. **What is the relationship between sequence homology and protein function?**

Ans. Sequence homology suggests that proteins may share similar functions if their sequences are related. This relationship is a fundamental tool for annotating unknown proteins in genomics.

SOLVED EXERCISE

MULTIPLE CHOICE QUESTIONS

Tick (✓) the correct answer.

1 **Generally, the function of a protein depends on its:**

(a) One-dimensional structure
(b) Two-dimensional structure
(c) Three-dimensional structure✓
(d) Four-dimensional structure

2 **The protein domains are:**

(a) Functional and structural units within protein✓
(b) Secondary structural elements
(c) Linear sequences of amino acids
(d) Specific regions for post-translational modification

3 **The first step in x-ray crystallography experiment is:**

(a) Compute an electron density
(b) Build a model of your molecule
(c) Measure a diffraction pattern
(d) Grow a crystal✓

4 **What is the primary role of computational biology?**
(a) Using computer algorithms to analyze data✓
(b) Identifying genetic mutations
(c) Studying protein functions
(d) Analyzing the expression patterns

5 **Which computational approach is used to predict protein structure based on amino acid sequence?**
(a) Multiple sequence alignment (b) Homology modelling✓
(c) Clustering analysis (d) BLAST searches

SHORT ANSWER QUESTIONS

1. **Define Domains of the Protein**

Ans. Protein domains are distinct functional and structural units within a protein. Each domain can fold independently into a stable three-dimensional structure and often has a specific function. Domains can act as modular building blocks, and a single protein may contain multiple domains, each contributing to the protein's overall function.

2. **How Does Coronavirus Enter Host Cells?**

Ans. The coronavirus enters host cells through a process involving the spike protein on its surface. This spike protein binds to a specific receptor protein found on the surface of human cells, called the ACE2 receptor. The binding allows the virus to attach firmly to the host cell and facilitates its entry, enabling the virus to infect the cell and replicate.

3. **Define Genomics**

Ans. Genomics is the study of genomes, which encompasses the complete set of DNA within a single cell of an organism. It involves sequencing, assembling, and analyzing the structure, function, and evolution of the genome to understand genetic variations, gene function, and how genes contribute to an organism's traits and diseases.

4. **Differentiate Between Genomics and Proteomics**

Ans. **Genomics** is the study of the entire genome, focusing on DNA sequences, their structure, function, and variation. It helps understand genes and their hereditary information.

Proteomics is the large-scale study of proteins produced by an organism, focusing on their structures, functions, interactions, and roles within cellular processes. While genomics looks at genetic information, proteomics studies the functional molecules derived from genes.

5. **What is GenBank? Describe It Briefly**

Ans. GenBank is a comprehensive public database maintained by the National Center for Biotechnology Information (NCBI) that stores nucleotide sequences from various organisms. It provides access to a vast repository of DNA sequences along with supporting bibliographic and biological annotations, facilitating genetic research, comparative genomics, and evolutionary studies worldwide.

6. **Write a Short Note on Protein Data Bank**

Ans. The Protein Data Bank (PDB) is an essential database that provides three-dimensional structural data of large biological molecules such as proteins and nucleic acids. It allows researchers to study the detailed 3D structures of macromolecules, understand their functions, and use this knowledge to design drugs that specifically target protein structures for therapeutic purposes.

LONG QUESTIONS

Q1. **Describe the applications of structural biology.**

Ans. See section 7.1 long question number 1 to 5.

Q2. **Write a note on principle and working of x-ray crystallography.**

Ans. See section 7.2 long question number 6 and 7.

Q3. **Briefly describe key databases of computational biology.**

Ans. See Section 7.3 Computational Biology Long Question number11.

INQUISITIVE QUESTIONS

1. How Can Structural Biology Help in Preventing a New Viral Disease During a Pandemic?

Ans. Structural biology can be extremely valuable in preventing disease caused by a new virus. By determining the three-dimensional structure of viral proteins—especially those involved in infection such as spike proteins—scientists can identify active sites and domains crucial for the virus's ability to enter host cells. This detailed structural information enables researchers to design targeted antiviral drugs or vaccines that specifically block these functional sites, thus preventing the virus from infecting cells. Additionally, structural biology helps understand how the virus interacts with host molecules, guiding the development of therapies that disrupt these interactions and stop viral replication.

2. How Can Computational Biology Be Used to Predict Structure and Function of an Unknown Protein?

Ans. Once the amino acid sequence of an unknown protein is determined (using Edman degradation or mass spectrometry), computational biology tools can be employed to predict its structure and function. Techniques like homology modeling use known 3D structures of similar proteins (homologs) to build a predicted model of the unknown protein's structure. Bioinformatics databases and algorithms (such as BLAST) help identify similar sequences and possible functions by comparing the unknown protein to well-characterized proteins. Computational simulations can further analyze how the protein folds, interacts with other molecules, and performs biological functions, providing valuable insights without the need for immediate experimental structure determination.

3. How Do Homology Models Differ From Experimentally Determined Macromolecular Structures?

Ans. Homology models are computational predictions based on the known structures of related proteins, constructed by aligning the unknown protein's sequence to a similar template. While these models provide useful approximations of structure, they are limited by the quality of the template and the sequence similarity. In contrast, experimentally determined structures, such as those obtained via X-ray crystallography or cryo-electron microscopy, provide direct and highly accurate three-dimensional details of the molecule at atomic resolution. Homology models may lack precision in regions with low sequence conservation or flexible loops, whereas experimental methods reveal true structural conformations and dynamic aspects, making them more reliable for detailed functional and drug design studies.

4. Draw a Flow Chart Describing the Steps Involved in Drug Development Till Prescription

Ans. Here's a verbal description of the flow chart steps in drug development:

- **Identification of Drug Target:** Use structural biology and computational biology to identify and validate a target protein involved in disease.
- **Drug Design and Screening:** Design molecules that can interact with the target; screen libraries of compounds using computational docking and experimental assays.
- **Preclinical Testing:** Evaluate drug candidates for efficacy, toxicity, and pharmacokinetics in cell cultures and animal models.
- **Clinical Trials:** Conduct multi-phase human trials (Phase I, II, III) to assess safety, dosage, efficacy, and side effects.
- **Regulatory Approval:** Submit trial data to regulatory authorities (like FDA) for approval.
- **Manufacturing:** Scale up drug production under quality control.
- **Prescription and Post-Market Surveillance:** Drug is prescribed by doctors; ongoing monitoring for adverse effects and long-term efficacy.

Self-Assessment Unit 7

Max. Marks: 28 **Time allowed 60 Mins**

Q1. Each of the following question has four options. Select the correct answer. (10x1=10)

1. **The function of a protein primarily relies on its:**
   (a) Linear structure (b) Two-dimensional form
   (c) 3D conformation (d) Four-dimensional arrangement
2. **Protein domains are best described as:**
   (a) Independent functional and structural segments within a protein
   (b) Elements of secondary structure like alpha helices
   (c) Straight chains of amino acids
   (d) Areas targeted for post-translational changes
3. **The initial requirement in an X-ray crystallography procedure is to:**
   (a) Generate a diffraction image
   (b) Build an atomic model
   (c) Grow crystals of the molecule
   (d) Calculate electron density maps
4. **What is the primary objective of computational biology?**
   (a) Employ computer-based methods for biological data analysis
   (b) Detect gene defects
   (c) Investigate the roles of proteins
   (d) Track gene expression in real-time
5. **Which technique is commonly used to predict a protein's shape from its sequence?**
   (a) Sequence alignment tools (b) Homology-based modeling
   (c) Cluster analysis techniques (d) BLAST similarity searches
6. **Which level of protein structure refers to its linear amino acid sequence?**
   (a) Primary (b) Secondary (c) Tertiary (d) Quaternary
7. **Which method can help compare unknown DNA or protein sequences to known databases?**
   (a) PCR (b) ELISA (c) BLAST (d) X-ray diffraction
8. **The three-dimensional folding of a polypeptide chain is known as its:**
   (a) Primary structure (b) Secondary structure
   (c) Tertiary structure (d) Domain structure
9. **Which of the following best defines 'homology modeling' in protein bioinformatics?**
   (a) Predicting gene expression levels
   (b) Designing new proteins from scratch
   (c) Estimating protein structure from known related sequences
   (d) Mapping the chromosomal location of a gene
10. **A region of a protein that can evolve, function, and exist independently of the rest of the protein chain is called a:**
   (a) Motif (b) Domain (c) Helix (d) Sheet

Q2. Write short answers to the following questions. (5x2=10)

1. Define domains of the protein.
2. How corona virus enters the host cells?
3. Define genomics.
4. Differentiate between genomics and proteomics.
5. What is GenBank. Describe it briefly.

Q3. Write detailed answer to the following question. (4+4=8)

1. Describe the applications of structural biology.
2. Write a note on principle and working of x-ray crystallography.