

Unit 7: Structural and Computational Biology

Short Questions

SHORT ANSWER QUESTIONS EXERCISE

Q1. Define domains of the protein.

Ans. The distinct structural units with independent functions, within a protein are called its domains.

Example: HIV-1 reverse transcriptase have its polymerase domain which synthesizes DNA, this domain has been identified by structural biologists.

Q2. How corona virus enters the host cells?

Ans. The corona virus (SARS-CoV-2) possess spikes in its envelope. These spikes are made up of proteins known as spike proteins, which stick out from the surface of the virus. There is a protein on human cells that acts as receptor (Cd-4) of virus spike protein. When spike protein attach with receptor protein of human cells (helper T-lymphocytes), it gets entry into host cells.

Q3. Define genomics.

Ans. Genomics is one of the major areas in computational biology, which involves the study of genomes, which are the complete set of DNA within a single cell of an organism. It also involves sequencing, all assembling and analyzing the function and structure of genome.

Q4. Differentiate between genomics and proteomics.

Ans.

Genomics	Proteomics
- Genomics i.e., the study of genomes, which are the complete set of DNA within a single cell of an organism. - Genomics involves sequencing, assembling, and analyzing the function and structure of genomes. It helps in understanding genetic variations, gene function, and evolutionary relationships.	- Proteomics i.e., the large-scale study of proteins, including their structures and functions. - Proteomics aims to map the entire set of proteins (the proteome) produced by an organism and understand their interactions and roles in cellular processes.

Q5. What is GenBank. Describe it briefly.

Ans. It is a comprehensive public **database of nucleotide sequences** and supporting bibliographic and biological annotations. It provides access to a vast repository of DNA sequences from various organisms, facilitating genetic research and comparative genomics.

Q6. Write a short note on Protein Data Bank.

Ans. Protein Data Bank (PDB)

This database provides 3D structural data of large biological molecules, such as proteins and nucleic acids. It is important for studying the structures of macromolecules, understanding their functions, and designing drugs that target specific protein structures.

SLO BASED SHORT QUESTION ANSWERS

Application of Structural Biology

Q7. Define structural biology. What information it provides about?

Ans. Structural biology deals with the study of **three dimensional (3D) structures of macromolecules** (including proteins and nucleic acids) at atomic levels.

- It provides the detailed information about the structure of biomolecule, its functions, dynamics and interaction with ligands and other macromolecules.

Q8. Enlist the applications of structural biology in the field of medical research.

Ans. Applications in Medical Research

- Determining the **active site** and domains of macromolecules.
- Identifying **drug targets** for developing vaccines or medicines.

3. Identification of **Host-pathogen** interactions.

4. Identification of **protein misfolding** to diagnose diseases.

Q9. What is the significance of knowing the location and structure of domains of macromolecules?

Ans. Knowing the location and structure of domains of macromolecules helps in designing of antiviral drugs that can specifically target them.

Q10. What are drug targets? Give example satisfying the role of structural biology in designing vaccines and medicines.

Ans. Drug Targets

Structural biology helps scientists find the right place on a disease-causing molecule where a drug can work. These places are usually proteins and are called drug targets.

- By studying the 3D shape of these proteins, scientists can find specific spots where a drug can attach and stop the protein from working.

For example, in COVID-19, scientists used structural biology to study the spike protein of the coronavirus (SARS-CoV-2). This protein helps the virus to enter human cells.

- By knowing its 3D structure, scientists identified it as a drug target. Thus, they designed vaccines and medicines that block the spike protein, preventing the virus from infecting more cells.

Q11. How studying 3D-structures of host-pathogen/s helps in the development of drugs to prevent this interaction?

Ans. Host-Pathogen Interactions

- Structural biology helps in understanding how **pathogens** (like viruses or bacteria) **interact** with the host's body cells. This is called host-pathogen interaction.

Significance: By studying the 3D structures of both the pathogen and the host cell proteins, scientists can see how the pathogen attaches to and enters the host cell, and which molecules are involved in the process.

For example, structural biologists studied the spike protein of coronavirus, which sticks out from the surface of the virus. They also looked at a protein on human cells that acts as receptor of virus spike protein.

- So, the scientists discovered exactly how the virus enters human cells. This information was vital in developing the drugs that can bind with receptor proteins. Such drug inhibits the interaction of the virus with the receptor and consequently blocks the entry of virus into the host cells.

Q12. Describe the role of structural biology in identifying diseases due to protein misfolding?

Ans. The functionality of proteins depends on the correct folding into three dimensional shapes.

- Several diseases originate due to incorrect folding of proteins, including cystic fibrosis, Parkinson's, Alzheimer's.
- Structural biology provides understanding of intricate folding pathways and how misfolding leads to the diseases.

X-ray crystallography

Q13. Who discovered x-ray crystallography? Describe its purpose of discovery.

Ans. X-Ray Crystallography

- X-ray crystallography was developed in 1912 by **William Henry Bragg** and **William Lawrence Bragg**.
- They were awarded **1915 Nobel Prize** in Physics for their work.
- Since then it has been used to **analyze the diverse substances including minerals, salts, metals, proteins, carbohydrates, nucleic acids and vitamins.**

Q14. Describe the working principle of x-ray crystallography.

Ans. In this technique, x-rays beam strikes crystal atoms and molecules in the crystals diffract the x-rays beam in specific directions. From the angles and intensities of diffracted beams, a 3D picture of electron density within the crystals are produced. The electron density is exploited to create 3D structure of the molecule.

Q15. Enlist the steps involved in working of x-ray crystallography.

Ans. In order to understand the working of x-ray crystallography its method is divided into following steps.

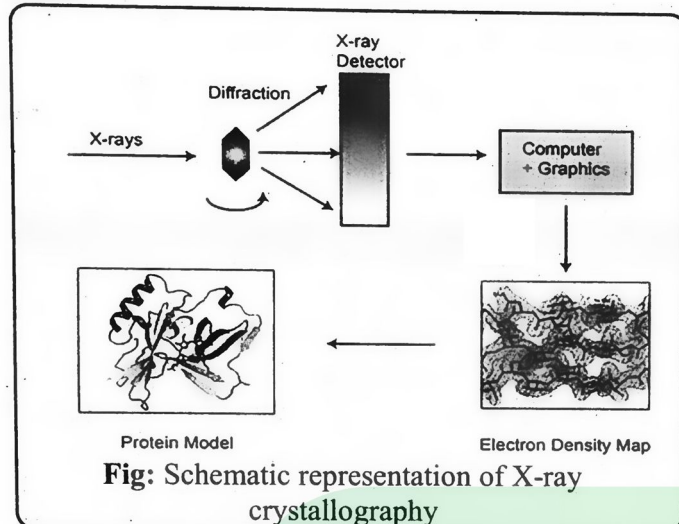
1. Protein crystallization
2. Production of diffraction pattern

3. Creating density map.

4. Determination of protein structure.

Q16. Draw the schematic representation of x-ray crystallography.

Ans.



Q17. Describe how proteins are crystallized in x-ray crystallography?

Ans. Protein Crystallization

- Protein crystallization means turning a purified protein into a solid crystal form. Crystals are needed because **they arrange protein molecules in a regular, repeating pattern**, which is important for getting a clear image during the x-ray process.
- To make crystals, scientists slowly mix the protein with special solutions that cause the protein molecules to stick together in an orderly way.
- This process can take hours, days, or even weeks. It often requires careful control of temperature, pH and salt concentration. Once a clear and stable protein crystal is formed, it can be used in the next steps.

Q18. How a diffraction pattern is produced in x-ray crystallography?

Ans. Production of a diffraction pattern

Once a good quality crystal is formed, it is mounted on the x-ray machine. The x-rays beam is bombarded at the crystal at various

angles. The atoms in the crystal diffract the x-rays beam and a diffraction pattern (which is a series of spots) is created on the detectors.

Computational Biology

Q19. Define computational biology describe its importance.

Ans. Computational Biology

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. 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.

Q20. Name major areas in computational biology, describe bioinformatics.

Ans. Major areas of computational biology include: Proteins, genomics, bio-informatics.

Bioinformatics: Bioinformatics i.e., the application of computer technology to manage and analyze biological data. Bioinformatics tools and techniques are used to store, retrieve and analyze DNA, RNA and protein sequences.

Q21. Write two applications of computational biology.

Ans. Applications of Computational Biology

Computational biology has vast applications, some of these are discussed here.

(i) Drug Discovery: Computational biology helps in **identifying potential drug targets** and simulating the effects of drugs on biological systems. It accelerates the drug

discovery process by predicting how drugs interact with proteins and other molecules.

(ii) **Genetic Research:** By analyzing DNA sequences, computational biology helps to **identify genetic variations** associated with diseases. It aids in understanding the genetic basis of diseases and can lead to the development of personalized medicine.

Q22. Name key databases used in computational biology.

Ans.

1. GenBank
2. Protein Data Bank (PDB)
3. Ensembl

Q23. Differentiate between BLAST and FASTA.

Ans.

BLAST	FASTA
- It is used for comparing primary biological sequence information, such as the amino-acid sequences of proteins or the nucleotides of DNA sequences. - It helps identify homologous sequences, predict functions of unknown genes and study evolutionary relationships.	- It is a sequence alignment tool that compares a query sequence to a database of sequences to find regions of similarity. - It is used for searching protein and nucleotide databases, identifying sequence homology and analyzing sequence alignments.

Q24. What do you know about Ensembl?

Ans. Ensembl

It is a genome browser providing information on genome sequences, gene

models and comparative genomics for various species. Ensembl helps to access and visualize genomic data, supporting studies in genomics and evolutionary biology.

Sequence Homology

Q25. Differentiate between sequence homology and structure homology.

Ans.

Sequence Homology	Structure 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.	Structural homology refers to the similarity in the three-dimensional structures of proteins or other macromolecules due to shared ancestry. Proteins with similar structures often perform similar functions, even if their sequences are not highly similar.

Q26. Categorize sequence homology.

Ans. Sequence homology is categorized into two main types.

Orthologs: Sequences in **different species** that **originated** from a common ancestral gene during speciation. Orthologs often retain the same function across species.

Example: Human histone H1.1, human histone H1.2 and chimpanzee histone H1.1.

Paralogs: Sequences within the **same species** that originated from gene duplication. Paralogs can evolve new functions even if they originally arise from the same ancestral gene e.g., Human histone H1.1 and human histone H1.2.

Q27. Describe applications of sequence homology study.

Ans. Sequence of homology study

- Sequence homology provides an insight into the **evolutionary relationships** between organisms.
- By comparing homologous sequences scientists can infer the **evolutionary history** and divergence of species.
- Sequence homology provides a clue about the **function of an unknown gene** or protein. If an unknown gene/protein is homologous to a gene/protein with a known function, it is likely to have a similar function.
- Identifying homologous genes involved in diseases across different species helps in **understanding disease mechanisms** and developing treatments.

- Homologous genes in model organisms can be studied to gain **insights into human diseases**.

Q28. Write some significances of studying structural homology.

Ans. Understanding structural homology helps in predicting the function of newly discovered proteins. Furthermore, structural homology is crucial in drug designing, as drugs are often designed to interact with specific protein structures. Understanding the structural relationships between proteins can help in designing more effective drugs. Also studying the structural homology of proteins help in understanding the evolutionary processes that shape protein functions and interactions.

LONG QUESTIONS (EXERCISE)

Q1. Describe the applications of structural biology.

Ans. See Q1. of theory.

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

Ans. See Q2. of theory.

Q3. Briefly describe key databases of computational biology.

Ans. See Q4. of theory.

INQUISITIVE QUESTIONS

Q1. Consider there is a pandemic of a new unknown disease, and the causative agent in a virus. You also know that virus belongs to X family. How structural biology can be helpful in preventing the disease?

Ans. Structural biology is pivotal in understanding viral out breaks. For a new virus from the X-family, following steps can be taken to prevent the disease from Novel virus;

1. Use X-ray crystallography to determine the 3D-structure of viral surface proteins (e.g spike proteins)
2. Identify how viral proteins interact with human cell receptor proteins. (e.g., ACE2 for SARS-CoV-2)
3. Locate conserved domains to design inhibitors.
4. Develop drugs that block the viral entry (e.g., by binding to spike proteins)
5. Use structural data to engineer mRNA vaccines.

This is how structural biology can be helpful in preventing that pandemic.

Q2. Suppose you find an unknown protein and determine amino acid sequence by Edman degradation/mass spectrometry. How can you exploit the computational biology to predict the structure and function of the protein?

Ans.

1. First of all, use BLAST to compare the amino acid sequence against GenBank or Protein Data Bank for homologs.
2. Then go for sequence homology, which will provide clue about the function of unknown protein.
3. If the unknown protein is homologous to a protein with known function, it is likely to have a similar function.
4. Finally use structural homology, the 3D structure of protein provides critical information about its function because proteins with similarity/similar in structures often perform similar functions.

Q3. Homology models of macromolecules differ from experimentally determined structures of the macromolecules. Please comment.

Ans. The homology models are basically approximations, while experimental structures reveal atomic levels, because homologous models are predicted from templates in database, while experimentally determined structures of macromolecules are derived from empirical data (e.g., x-ray crystallography) in which x-ray beam is bombarded in specific direction on crystals, from angles and intensities of diffracted beams, a 3D picture of electron density within crystals are produced which is exploited to create 3D-structure of molecule, that gives a real image rather than homology models.

Q4. Draw a flow chart to describe the steps involved in drug development till its prescription.

Ans.

