ENCYCLOPAEDIA OF VITAL PROTEINS

William Dean Howells
Dr. Sangeeta Kapoor

Encyclopaedia of Vital Proteins

William Dean Howells
Dr. Sangeeta Kapoor

Encyclopaedia of Vital Proteins

William Dean Howells
Dr. Sangeeta Kapoor

Dominant
Publishers & Distributors Pvt Ltd
New Delhi, INDIA

Knowledge is Our Business

ENCYCLOPAEDIA OF VITAL PROTEINS
By William Dean Howells, Dr. Sangeeta Kapoor

This edition published by Dominant Publishers And Distributors (P) Ltd
4378/4-B, Murarilal Street, Ansari Road, Daryaganj,
New Delhi-110002.

ISBN: 978-81-78886-52-7

Edition: 2022 (Revised)

Dominant
Publishers & Distributors Pvt Ltd

Registered Office: 4378/4-B, Murari Lal Street, Ansari Road,
Daryaganj, New Delhi - 110002.
Ph. +91-11-23281685, 41043100, Fax: +91-11-23270680

Production Office: "Dominant House", G - 316, Sector - 63, Noida,
National Capital Region - 201301.
Ph. 0120-4270027, 4273334

e-mail: dominantbooks@gmail.com
info@dominantbooks.com

w w w . d o m i n a n t b o o k s . c o m

CONTENTS

CHAPTER 1

EXPLORING THE DIVERSITY AND FUNCTIONALITY OF BIOLOGICAL POLYMERS: THE WORLD OF POLYPEPTIDES AND PROTEINS

Dr. Sangeeta Kapoor, Professor & HOD
Department of Biochemistry, TMMC&RC, TeerthankerMahaveer University, Moradabad, Uttar Pradesh India,
Email Id- drsangeeta2000@gmail.com

ABSTRACT:

Within biological systems, polymers are a basic type of macromolecules that are made up of complicated linear chains that are assembled from smaller building components. Polysaccharides, polynucleotides, and polypeptides are three main biological polymers that serve crucial functions. This chapter focuses on polypeptides, which are collections of units made of amino acids connected by peptide bonds. The length of these peptides varies, with proteins serving as notable examples. Proteins have amazing chemical and physical variety due to the twenty different amino acids they contain, each of which has properties that are both unique and common. This diversity results from the proteins' many linear configurations. Importantly, proteins are more than just strands of unstructured material; they fold into complex three-dimensional structures that support their many activities. Peptide bonds, which form when an amino group and a carboxyl group combine without a water molecule present, bind amino acids together in proteins. Polypeptide chains have two unique chemical endings, each with an amino or carboxy terminus, which are essential for defining the chain's linear directionality and have a considerable impact on the final structure and function of the protein. This chapter goes further into the intriguing world of proteins, examining their many roles, intricate structural details, and the crucial function of amino acid sequences in determining their three-dimensional shapes. Untangling the intricacy of biological systems and recognising how crucially important proteins are to maintaining life requires an understanding of these basic features of proteins.

KEYWORDS:

Biological Polymer, Diversity, DNA, Polypeptides, Proteins, RNA.

INTRODUCTION

Polymers make up many of the most significant macromolecules in biological systems. These polymers are made up of tiny building pieces connected by extensive linear chains. Polysaccharides, polynucleotides, and polypeptides are three of the most significant biological polymers. While polynucleotides like DNA and RNA are constructed from nucleotides, polysaccharides like starch are made up of sugar subunits. We concentrate on polypeptides in this chapter and the one after. Polypeptides are chains of amino acid-based building blocks connected by peptide bonds. Peptides are short polypeptides, whereas proteins are generally large polypeptides. Twenty different types of amino acids, which are both similar and different, make up proteins. They display individual chemical characteristics as well as similar traits that enable them to create peptide connections with one another. Proteins may have a wide range of chemical and physical characteristics due to the variety of amino acids and the enormous number of combinations that can be made when they are arranged in a linear fashion. In addition, proteins are not just unorganised chains. Instead, they fold in on themselves to create three-dimensional structures with distinctive characteristics.

Proteins serve a variety of purposes.

Cells use proteins for a variety of functions due to their tremendous diversity. Some proteins serve as enzymes that reduce G to catalyze chemical processes. The majority of enzymes are proteins, and as we saw in the last chapter, cells utilise enzymes to regulate reaction rates and link advantageous processes with unfavorable ones in order to carry out controlled chemical reactions. Enzymes really mediate almost all of the changes that take place inside the cell; without them, biological systems would do very little chemistry. Enzymes are often categorized based on the chemistry they carry out and catalyze a broad range of processes. The majority of enzymatic processes include one of the following: the transfer of functional groups, the creation or breakage of bonds, the rearrangement of bonds within specific molecules, or the use of ATP to covalently link molecules. In later chapters, we'll have more to say about how enzymes diminish G [1], [2].

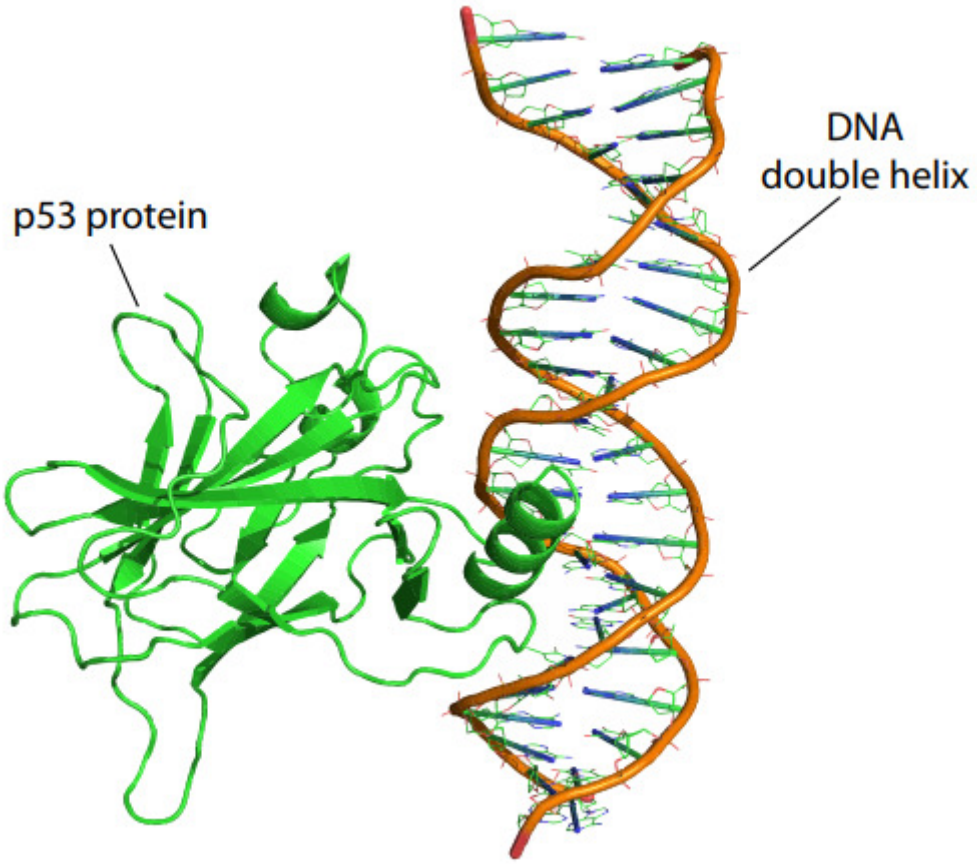

Figure 1: Proteins fold into a variety of different three-dimensional configurations, but the structure of DNA is predictable [RCSB.Org].

Even while proteins as a whole are more diverse, enzymes are more varied in terms of their forms and sizes and the wide range of particular chemical processes they may catalyze. Some proteins provide structural functions; not all proteins are enzymes. These proteins are also present in fingernails, skin's outer layers, and hair. A lot of common fabrics like wool, silk, and leather are formed of protein as well. These structural proteins have developed throughout time in order to protect the organisms that make them from certain mechanical pressures. On a cellular level, structural proteins support the cell's structural integrity and are in charge of a large portion of the compartmentalization and organisation seen in living systems. For instance, actin and tubulin, two cytoskeletal proteins, are often utilised as scaffolding to place and regulate the movement and localization of other cellular components.

As a result, even within the ostensibly limited category of structural proteins, we find a diverse variety of structures, functions, and behaviours. Numerous biological activities are regulated by proteins as well. Life relies on chemical reactions that take place quickly, as we've previously seen, but it also requires that these reactions be properly timed. Thus, other proteins that coordinate cellular processes by reacting to external variables control the activities of specific enzymes. We will investigate a particular regulatory protein, Abl, whose dysfunction leads to cancer by deregulating cell division in subsequent chapters. These chapters will also look at how proteins govern when and where RNA is created from DNA. Several proteins serve as mollusk carriers. One example is hemoglobin, which transports oxygen through the circulatory system by carrying oxygen gas. Protein carriers also transport lipids and cholesterol throughout the circulation. Some proteins are harder to categorise than others. Ion channels, for instance, can be viewed as either enzymes or carriers because they transport ions from one side of the membrane to the other while allowing ions to pass through cell membranes, which is essential for muscle contraction and neurotransmission.

The structure of a protein influences how it works. The X-ray crystal structure of DNA with a protein attached is shown in Figure 1. Each atom in a molecule is described in almost perfect detail by the X-ray crystal structure of the molecule in three dimensions. A crystal of the target molecule and X-rays, a kind of electromagnetic energy, are used in the procedure to generate such a structure. The crystal lattice is broken by an X-ray beam. As the X-rays come into contact with the individual atoms that make up the molecules in the crystal lattice, some of the X-rays are diffracted. The arrangement of atoms may be determined from the distinctive pattern of reflections that the diffraction produces and collects on an X-ray film. You'll find several examples of this methodology used by scientists to ascertain the structures of proteins and other molecules essential to life throughout this book [3], [4].The well-known DNA double-helical structure is seen in Figure 4. Regardless of the nucleotide sequence, DNA usually always has the same three-dimensional form, which is a double helix. The protein in Figure 1 appears quite differently from the proteins in Figures 2 and 3, in comparison. This is due to the wide variety of three-dimensional protein architectures. The reason why proteins are so adaptable is in large part due to their capacity to adopt a variety of unique three-dimensional conformations; a protein's shape dictates its function. Contrarily, DNA only serves one purposethe storing of genetic informationand adopts a specific structure for that purpose.

DISCUSSION

The folded structure of a protein is determined by the amino acid sequence. A protein's fundamental structure is defined as the particular arrangement of amino acids in the protein. A protein's three-dimensional architecture is determined by this sequence. In the next chapter, a well-known experiment that demonstrates how all the data required for a protein's appropriate folding is found in its fundamental structure is discussed. Amino acids have similar structural characteristics. An amino group (-NH2), a carboxylic acid group (-COOH), and an additional carbon atom that connects these two groups make up every amino acid. Alpha carbon is the carbon atom that comes next. An extra chemical group, known as a R group or side chain, is joined to the carbon atom in 19 of the 20 amino acids. Instead of a R group and a single hydrogen atom, the 20th amino acid, glycine, contains two hydrogen atoms attached to the carbon. The nature of the R group determines the specific chemical and structural characteristics of each amino acid.

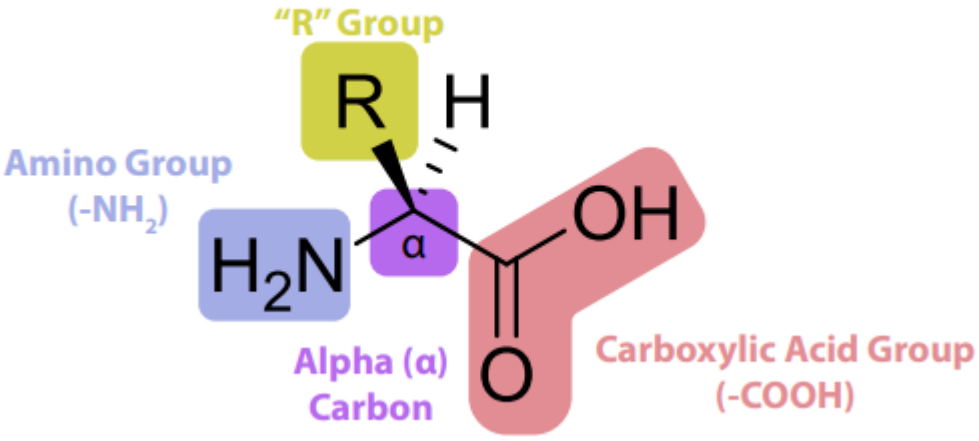

Figure 2: Shows the four constituents of an amino acid [Africagay.Org].

Chiral amino acids

Amino acids are distinguished by the chiral nature of their carbon atom. A chiral centre, often referred to as a stereocenter, is produced when a carbon atom is joined to four different groups. A chiral molecule is one that, to put it simply, can't be overlaid with its mirror copy. This is shown in Figure 3 by a cartoon whereby four distinct coloured spheres stand in for the four different substituents of the carbon atom of one of the 19 amino acids that has a R substituent (molecule A). Because molecule B can't be rotated to match molecule A, it is obvious that molecule A's mirror copy is chiral [5], [6]. A cartoon showing the substituents connected to the carbon of glycine is shown in Figure 4. As seen by the green spheres in molecule C, glycine is an achiral molecule in which two of the substituents (hydrogen atoms) are the same. You can observe that by rotating the mirrored molecule (molecule D), molecule C and molecule D become interchangeable.

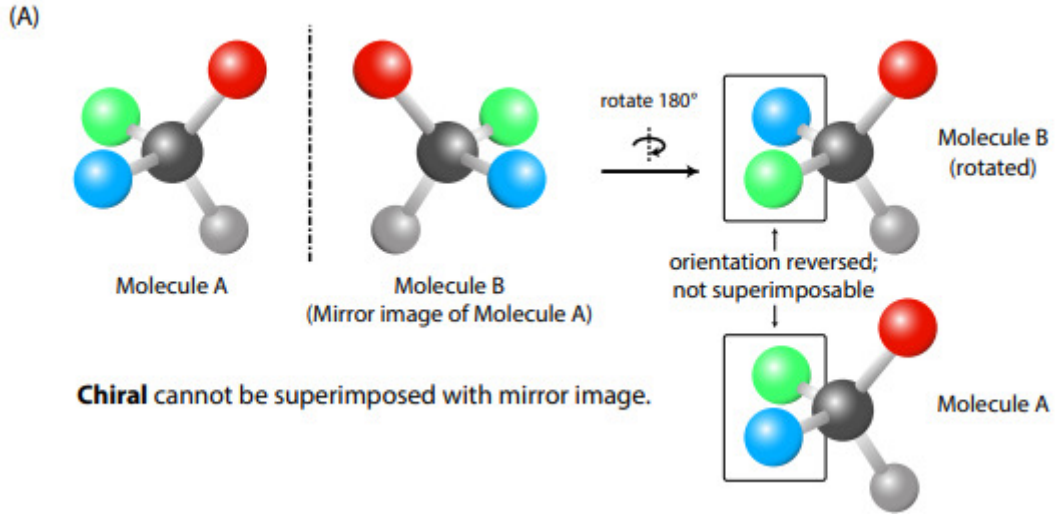

Figure 3: Atoms that are attached to four different substituents are chiral because their mirror images cannot be placed on them [Science Direct.Com].

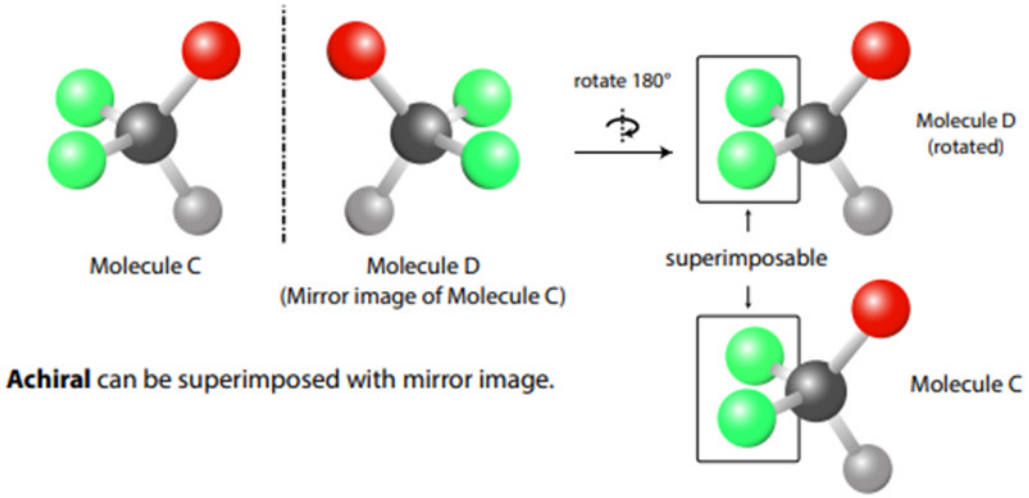

Figure 4: Atoms may be overlaid with their mirror copies if they are bonded to fewer than four distinct substituents, making them achiral [Science Direct.Com].

A chiral center's stereochemistry, also known as configuration, is binary; it may either have one configuration or the opposite of that configuration. Two extra kinds of lines are used to illustrate stereochemistry using normal line drawings: a solid, wedge-shaped line to indicate bonds that extend from the page, and a dashed, wedge-shaped line to represent bonds that extend into the page. Standard line drawings may be given a third dimension called depth, which allows us to express the stereochemistry of any molecule.Chirality is not an abstract idea, despite appearances to the contrary. You are surrounded by chiral items; the majority of your body's big and tiny molecules as well as many macroscopic structures are chiral. You cannot, for instance, put your left shoe on your right foot because of the chirality of your feet. Your hands are chiral, which is why it might be difficult for left-handed individuals to use most scissors since they are made for right-handed people, who also happen to be the majority.

As with hands or feet, chirality in molecules may have a significant impact on how they operate. Consider the two tiny molecules in Figure 5 as an illustration. Both molecules are known as carvones and have the same number, kind, and bond connections of atoms. The two molecules are, nevertheless, distinct because carvone only contains one chiral centre. The major odorant in spearmint is represented by the molecule on the left, while the primary odorant in caraway, the seed used in rye bread and Swedish cookies, is represented by the molecule on the right [7], [8]. The scents of spearmint and caraway cannot be mixed together. These compounds' distinctive aromas result from the chirality of the receptors to which they bind in the nose. The stereoisomers of carvone connect to the chiral receptors in the nose differently, much as your left shoe fits your right foot differently than your left foot. In other words, even if they may seem comparable on paper, their forms in three dimensions are quite different.

Carvone has two stereoisomers, each of which has unique characteristics. The perfume of spearmint comes from the stereoisomer of carvone on the left, while the aroma of caraway comes from the stereoisomer on the right. Although these molecules are mirror copies of one another reflected across the dashed line, with just one stereocenter's arrangement altered, our noses see them quite differently.The 19 chiral amino acids each have two stereoisomers, which are identified as D and L in Figure 6. In nature, only the L-stereoisomer is utilised to build proteins. While it can appear random, keep in mind that stereoisomers are unique entities with different shapes. Consider a protein with 100 amino acids, where each amino

acid may either be D or L. If so, the same protein may exist in 2100 different forms. These proteins would all have drastically different structures and fold in unforeseen ways. Because only the L-stereoisomer is used in live systems, there is no such issue, and proteins always fold into the same predicted shapes.

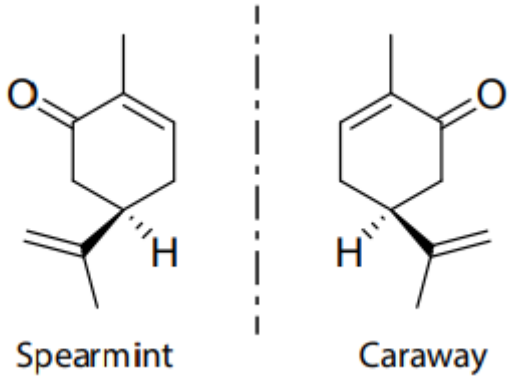

Figure 5: Chirality may significantly alter the characteristics of molecules [Lab X Changer].

A single stereoisomer may be used in applications other than proteins. All biological compounds having four distinct groups linked to the carbon atom, such as sugars and nucleotides, are only found in one stereoisomer.

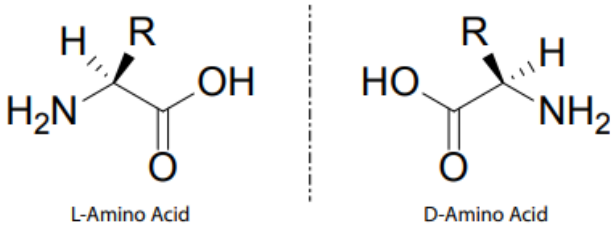

Figure 6: Each amino acid has two stereoisomers, but only the L-stereoisomer is employed by nature to make proteins [Research Gate. Net].

Peptide bonds hold amino acids together in proteins. As we now know, proteins are made up of lengthy chains of amino acids joined by covalent bonds. The term peptide bond refers to the covalent connection between two amino acids, therefore the word polypeptide. Since the peptide bond is an illustration of the amide functional group that we studied, it could appear familiar. Peptide bonds link the carbon atom of one amino acids carboxylic acid to the nitrogen atom of another amino acids amino group. This is analogous to removing a water molecule from each of the two amino acids, namely a hydrogen atom from the amino group and a hydroxyl group from the carboxylic acid. The immediate removal of water molecules would be energetically unfavorable, so instead peptide bonds in living systems are created via a sequence of intermediate biochemical processes [9], [10].Polypeptide chains have two distinct chemical groups at each end, and they are directed. A free carboxylic acid group is located at one end and a free amino group is located at the other. When the amino acid sequence of a protein is specified, it must indicate the N-terminal-to-C-terminal directionality; reversing the directionality indicates a very different protein sequence. The end of the polypeptide chain with a free amino group is the amino terminus, which is frequently written as N-terminus or NH3+-terminus. The end with a free carboxylic acid group is the

carboxy terminus, which is frequently written as C-terminus Amino acid sequences are normally written left to right, starting with the amino and ending with the carboxy terminus.

CONCLUSION

In conclusion, research on polymers in biological systems shows an amazing mosaic of intricacy and variety. Three of the most important biological polymers are polysaccharides, polynucleotides, and polypeptides, each of which plays a distinct part in the complex machinery of life. This chapter has illuminated the intriguing world of polypeptides and proteins, whereas DNA and RNA, built from nucleotides, contain the blueprint of genetic information. We have seen firsthand the enormous variety and complexity that underpin the operation of biological systems throughout our voyage into the realm of polymers. The study of proteins in particular holds the possibility of revealing a wealth of information about the mechanisms that keep life functioning, from catalyzing chemical reactions to controlling cellular functions and guaranteeing structural integrity. We will continue to be in awe of the beauty and intricacy of the biological world as we dig more into the realm of molecules and their intricate functions in the next chapters.

REFERENCES:

[1] C. Heinzmann, C. Weder, and L. M. De Espinosa, Supramolecular polymer adhesives: Advanced materials inspired by nature, *Chemical Society Reviews*. 2016.

[2] J. J. Green and J. H. Elisseeff, Mimicking biological functionality with polymers for biomedical applications, *Nature*. 2016.

[3] E. Yashima, N. Ousaka, D. Taura, K. Shimomura, T. Ikai, and K. Maeda, Supramolecular Helical Systems: Helical Assemblies of Small Molecules, Foldamers, and Polymers with Chiral Amplification and Their Functions, *Chemical Reviews*. 2016.

[4] J. M. Priegue, D. N. Crisan, J. Martínez-Costas, J. R. Granja, F. Fernandez-Trillo, and J. Montenegro, In Situ Functionalized Polymers for siRNA Delivery, *Angew. Chemie - Int. Ed.*, 2016.

[5] R. J. Wojtecki and A. Nelson, Small changes with big effects: Tuning polymer properties with supramolecular interactions, *J. Polym. Sci. Part A Polym. Chem.*, 2016.

[6] K. Ishihara, W. Chen, Y. Liu, Y. Tsukamoto, and Y. Inoue, Cytocompatible and multifunctional polymeric nanoparticles for transportation of bioactive molecules into and within cells, *Science and Technology of Advanced Materials*. 2016.

[7] S. Mohanty *et al.*, 3D Printed Silicone-Hydrogel Scaffold with Enhanced Physicochemical Properties, *Biomacromolecules*, 2016.

[8] C. Wang, J. Zhou, P. Wang, W. He, and H. Duan, Robust Nanoparticle-DNA Conjugates Based on Mussel-Inspired Polydopamine Coating for Cell Imaging and Tailored Self-Assembly, *Bioconjug. Chem.*, 2016.

[9] Y. Hou, J. Yuan, Y. Zhou, J. Yu, and H. Lu, A Concise Approach to Site-Specific Topological Protein-Poly(amino acid) Conjugates Enabled by in Situ-Generated Functionalities, *J. Am. Chem. Soc.*, 2016.

[10] J. M. Priegue, D. N. Crisan, J. Martínez-Costas, J. R. Granja, F. Fernandez-Trillo, and J. Montenegro, In Situ Functionalized Polymers for siRNA Delivery, *Angew. Chemie*, 2016.

CHAPTER 2

PEPTIDE BOND CHEMISTRY: PROFOUND INFLUENCE ON POLYPEPTIDE STRUCTURES AND FUNCTION

Dr. Sushil Yadav, Assistant Professor
Department of Biochemistry, TMMC&RC, TeerthankerMahaveer University, Moradabad, Uttar Pradesh India,
Email Id- yadav.sushil4644@gmail.com

ABSTRACT:

The chemical properties of the peptide bond have a significant impact on the shape of polypeptide chains. There are certain bonds within the peptide chain that are nonetheless freely rotatable even when the peptide bond itself only displays limited conformational flexibility. The orbital interactions that control covalent bonding are at the heart of this phenomena. Atomic orbitals mix and overlap as a consequence of covalent interactions between atoms, forming bonding orbitals that control how shared electrons are distributed. Single bonds are represented by the sigma bond, which is created by cylindrically symmetric orbitals joining two nuclei, while double bonds are made up of the sigma bond and the pi, which is created by overlapping p orbitals. One, two, and three bonding orbitals, respectively, are corresponding to single, double, and triple bonds. Because single bonds' sigma bonding orbitals are cylindrical and rotation does not affect orbital overlap, single bonds may freely spin. Contrarily, double bonds including pi bonds, such as those in ethylene (C2H4), are incapable of rotating because pi orbitals need parallel alignment, and doing so would cause the bond to be broken. It's interesting that the peptide bond, which at first glance could seem like a single bond between carbon and nitrogen, really has double-bond properties that inhibit free rotation. The coplanar alignment of the atoms that make up this bond is what causes its unique behaviour, which includes resonance stabilization and partial positive and negative charges on nitrogen and oxygen, respectively. These characteristics, as well as steric factors, contribute to the stiffness of the peptide bond. Additionally, peptide bonds may take on cis or trans forms, with the former being preferred since it prevents steric conflicts. Protein folding is substantially influenced by the arrangement of peptide links. Proteins are made up of amino acids, which have a variety of properties based on their side chains (R groups), which may be classified as polar or nonpolar. The side chains of polar amino acids are hydrophilic and polar, whereas those of nonpolar amino acids are hydrophobic and nonpolar.

KEYWORDS:

Amino Acid, Chemical, Chemistry, Peptide Bond, Polypeptide.

INTRODUCTION

Some amino acids contain side chains that may ionize; the pKa values of these side chains define the charge states of these amino acids at physiological pH. Given their critical role in protein folding and function, understanding the molecular characteristics of these amino acids is essential for understanding protein behaviour. Additionally, polypeptide chains include ionizable functional groups at their N- and C-termini, which affects how charged they are in aquatic conditions. The chemical properties of peptide bonds and amino acids as a whole contribute to the structural and functional variety of proteins. The coplanar orientation of the atoms attached to the peptide bond causes it to behave like a double bond, contrary to early predictions that it would be free rotating. This behaviour is strengthened even more by resonance stabilization, which leads to a hybrid state that favours the carbon-nitrogen bond's single-bond-like characteristics. This flat, stationary peptide link also causes polarization,

which affects how proteins fold by producing partial positive and negative charges on nitrogen and oxygen, respectively.

The form of the polypeptide chain is significantly influenced by the peptide bond's chemical characteristics. While the peptide bond itself can only have particular conformations, some of the bonds in the peptide chain are freely rotatable. We go back to the subject of the orbitals that encircle atoms to see why. When two atoms are joined together in a covalent bond, their orbitals mix and overlap, forming a bonding orbital that houses the shared pair of electrons. Atomic orbitals from each atom are combined to form bonding orbitals, which, like atomic orbitals, define where the shared electrons are most likely to be located. A cylindrically symmetrical orbital that connects two nuclei makes up a single link, often known as a sigma bond. One bond and a second bond known as a pi bond make up a double bond. The p orbitals of two nuclei collide to produce the connection [1], [2].

Only two electrons can fit in bonding orbitals. As a result, single, double, and triple bonds must, respectively, include one, two, and three bonding orbitals. In actuality, a single bond always consists of a single -orbital, a double bond always consists of a single -orbital and a triple bond always consists of a single -orbital and a double -orbital. This has a significant repercussion; bonds containing -orbitals cannot spin freely. Because the sigma bonding orbitals are cylindrical, single bonds may spin without affecting the orbital overlap that causes the bond to form. On the other hand, for p orbitals that make up -orbitals to overlap, they must be in parallel. The connection would rupture if the -bond rotated. The rotation of a covalent bond has a high barrier since it takes energy to break one. Double bonds are not free to rotate since they include -orbitals. The example of ethylene (C2H4), which has a carbon-carbon double bond, is seen in Figure 1. This double bond has a -bonding orbital, which prevents the bond from rotating. The p orbitals may overlap and create the -bonding orbital in ethylene because the hydrogen and carbon atoms are positioned in the same plane.

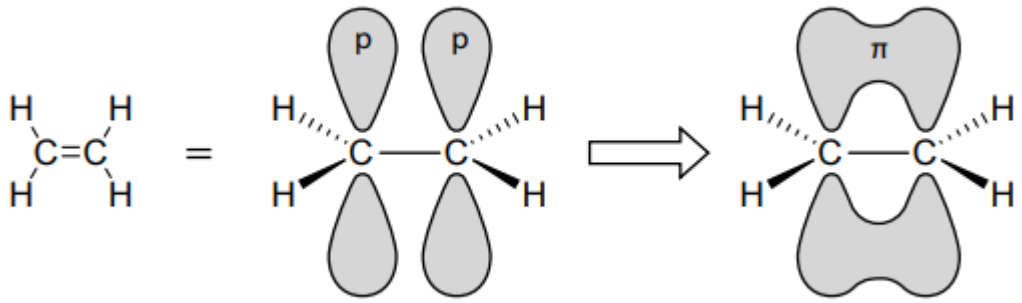

Figure 1: Double bonds cannot rotate because they contain π-orbitals [Research Gate. Net].

Why the peptide bond does not spin is still a mystery to us. The peptide bond should spin easily since we have shown it as a carbon-nitrogen single bond. However, it turns out that the peptide bond has certain double-bond characteristics and does not behave like a single bond. The typical line depiction of the peptide bond that we have shown is a little deceiving since it leaves out some important details. The atoms connected to nitrogen and carbon are in the same plane and rotation of the peptide bond is constrained, similar to how ethylene operates. Consider the bonding between nitrogen and the carbonyl carbon of the peptide bond to further understand this behaviour. A lower-energy state is produced when the nitrogen tends to share its single pair of electrons with the carbonyl carbon, delocalizing electrons among

the nitrogen, carbon, and oxygen atoms (Figure 2). Resonance stabilization is the term used to describe bonding in which electrons are spread across many atoms and delocalized. The peptide bond does not fluctuate between the two forms shown in Figure 2 as resonance stabilization implies; rather, it exists as a middle ground between these two extremes. However, this hybrid state's electron dispersion is not evenly distributed. Instead, it is biased such that the carbon-nitrogen bond's single bond-like nature is somewhat preferred (60%) over its double bond-like nature. The peptide bond is flat, which means that both the nitrogen and carbonyl carbon atoms have trigonal planar geometries. This is because the C-N bond exhibits certain double-bond characteristics.

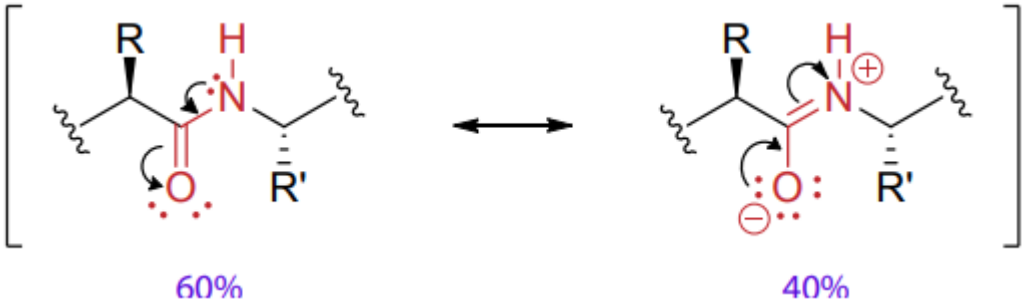

Figure 2: Resonance stabilization causes the peptide bond to have double-bond character [Lab X changer].

Another effect of resonance stabilization is the polarization of the peptide bond. In this hybrid form, the nitrogen has a partial positive charge while the oxygen has a partial negative charge due to resonance. The peptide bond is polar because of the charge difference between oxygen and nitrogen. The partial-positive charges of nitrogen and hydrogen, as well as the partial-negative charges of oxygen, are significant factors that affect how proteins fold, as we shall see later. There is more to polypeptide chains than just a never-ending string of peptide links. Instead, it also consists of bonds between carbonyl carbons and amide nitrogen's as well as between carbons and carbons. We will now explain why the other two bonds in this framework are allowed to spin whereas the peptide bonds are not[3], [4].

DISCUSSION

The peptide bond is flat because of its double-bond nature, much as in the ethylene example from Figure 1 did. The peptide bond forces all of the atoms that are immediately attached to it to be coplanar. The carbonyl carbon, carbonyl oxygen, peptide link nitrogen, hydrogen attached to that nitrogen, and the two neighbouring carbons are the six coplanar atoms. Figure 3 shows coloured rectangles as highlights for each group of six coplanar atoms. With regard to the neighbouring coplanar sets of atoms, each rectangle is free to spin. The carbon is coplanar with two distinct and dissimilarly coloured neighbouring blocks of atoms, as can be seen. The polypeptide backbone is stiffened and constrained into a certain set of conformations by the requirement that sets of six atoms be coplanar, which affects how the protein will fold as a whole. Generic polypeptide chain diagram in (A). Due to the peptide bond's double-bond nature, coloured rectangles show groups of six coplanar atoms. The peptide backbone's free-rotating bonds are denoted by green checks, whereas the non-rotating bonds are denoted by red Xs. Just the peptide backbone bonds are marked, take note. (B) On the left is a computer-generated model of a short peptide, and on the right is a typical line drawing of the same structure. Be aware that sets of coplanar atoms might spin in relation to one another due to the rotation of peptide backbone bonds other than the peptide bond.

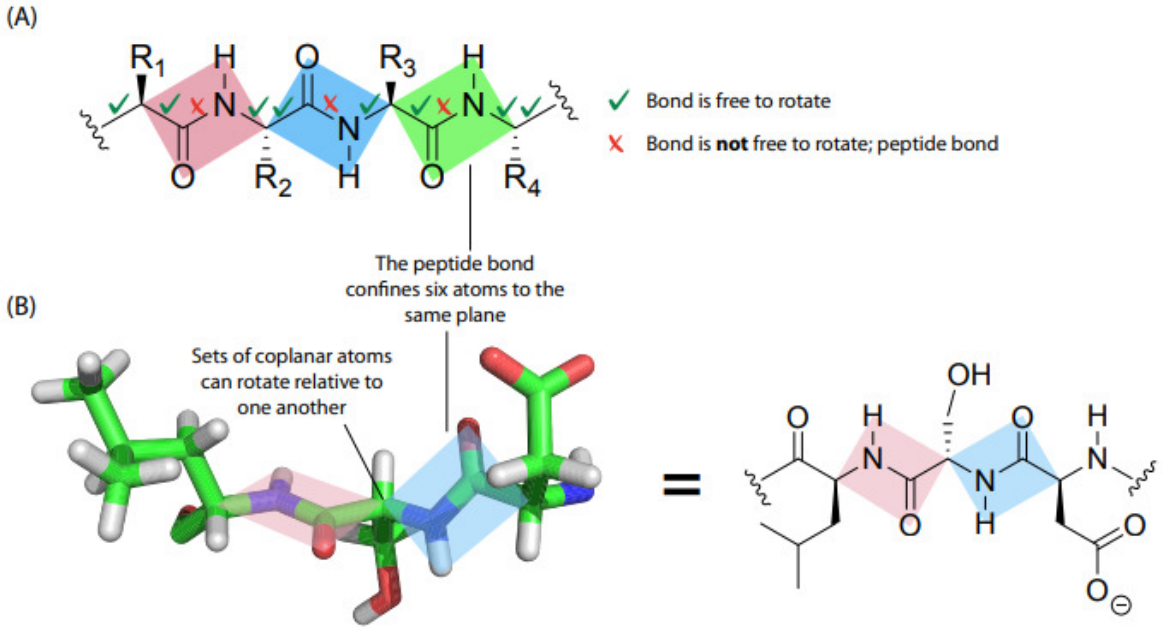

Figure 3: Shows a peptide link that is flat and immobile [Science Direct.Com].

The trans conformation is favoured by peptide bonds. The peptide bond may exist in two different configurations known as geometric isomers since it cannot spin freely. Let's first take the straightforward situation of the hypothetical molecule shown in Figure 3A in order to comprehend the idea of geometric isomers. The two geometric isomers cis and trans, which define how two atoms or groups of atoms are orientated in relation to one another on opposing sides of a double bond, may occur for this compound (Figure 3B). When two atoms or groups of atoms are on the same side of a hypothetical line drawn along the length of the double bond, they are said to be cis to one another. They are trans to one another if they are on opposite sides of that fictitious line. According on whether the carbons of neighbouring amino acids are on the same side of the peptide bond (Figure 4, right panel) or are on different sides (Figure 4, left panel), peptide bonds may also exist as cis or trans isomers. The trans isomer is often preferred over the cis isomer. The cis isomer brings the R groups of the two linked amino acids very close together, which accounts for this.

As a consequence, there is a steric clash, a non-bonded interaction in which electrons in one bond approach electrons in a neighbouring bond too closely. This idea may be easily seen as two chemical groups' electron clouds colliding as they attempt to occupy the same place at the same time. Steric conflicts are undesirable, as you would expect. The R groups are kept far apart and the steric conflict is avoided in the trans isomer of the peptide bond. As we shall see in a moment, the chemistry of the R groups determines how much the trans configuration is preferred over the cis form [5], [6]. The 20 amino acids have a range of characteristics. According to their R groups or lack thereof in the case of glycine, each of the 20 amino acids has unique chemical and physical characteristics. Despite their wide range of properties, amino acids may be categorized based on their size, charge, polarity, and, in certain circumstances, the peculiar structural characteristics they confer on the polypeptide backbone. You should get familiar with all 20 amino acids' structures, three-letter acronyms, and one-letter codes since they are crucial to life. They are a component of life's language.

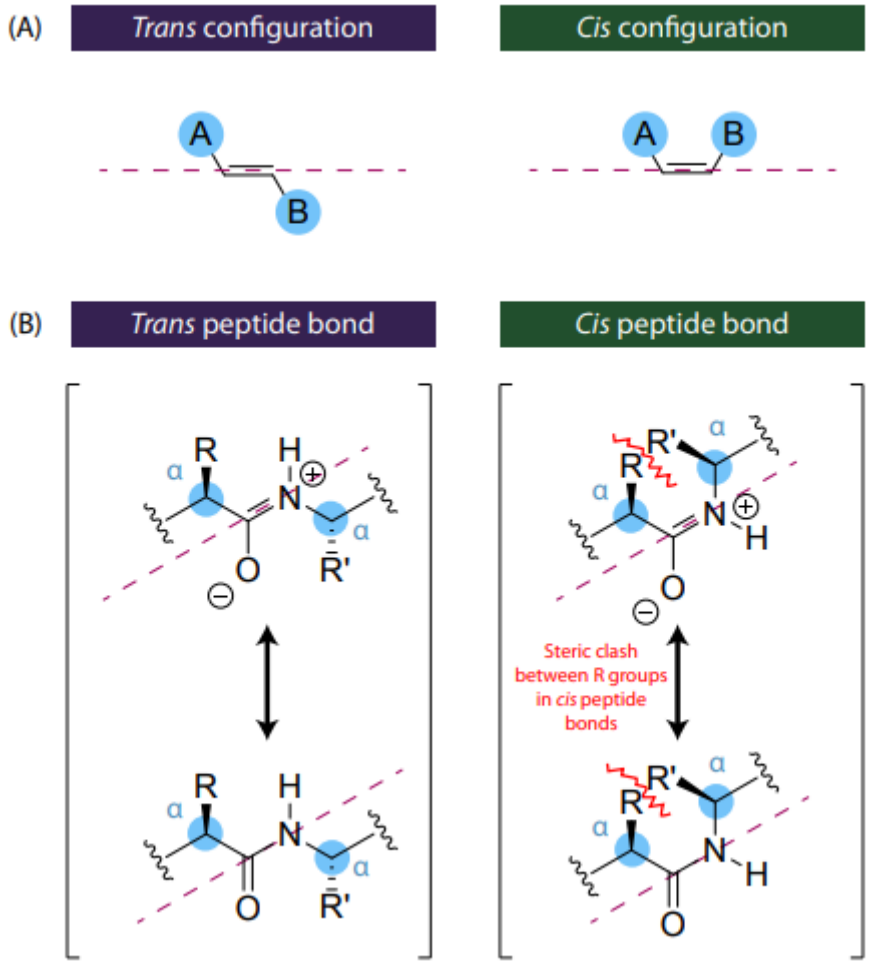

Figure 1 shows that the peptide bond's trans form is preferred and in (A) peptide bond Shows, (B). The cis and trans configurations of the peptide bond's extreme resonance states are shown in B [Research Gate. Net].

There are two types of amino acids: polar and nonpolar. The side chains of nonpolar amino acids often have no polar linkages. Nonpolar amino acids often contain hydrocarbon side chains made of carbon and hydrogen, which have a little electronegative difference (for example, isoleucine). These hydrocarbon side chains' bonds are not polarised as a result. Methionine, a nonpolar amino acid, also includes sulphur, which shares carbon's electronegativity. Because they tend to separate from water for reasons that we shall discuss later, nonpolar amino acids are hydrophobic. The polar amino acids, on the other hand, are those that have polar side chains. Because of the strong interactions between their side chains and water, polar amino acids are hydrophilic. Cysteine is an intriguing example, even though you may believe it belongs in the nonpolar category. However, as we shall see below, the S-H bond in cysteine has a propensity to ionise at neutral pH and hence acts something like a polar side chain. Both tryptophan and tyrosine have long side chains with polar and nonpolar characteristics, making it difficult to classify them as hydrophobic or hydrophilic [7], [8].

Ionizable side chains are found in several amino acids. At certain pH levels, some amino acids ionise (that is, take on a complete charge) (Figure 5). Acids include glutamic acid and aspartic acid. Both amino acids mostly reside in their deprotonated, negatively charged forms under physiological settings because their side chains have pKa values that are substantially below physiological pH. When aspartic acid and glutamic acid are deprotonated, they take on

distinct identities and are referred to as aspartate and glutamate, respectively. These names are often used interchangeably. Cysteine's pKa ranges from 8.0 to 9.0, therefore it may also be deprotonated.

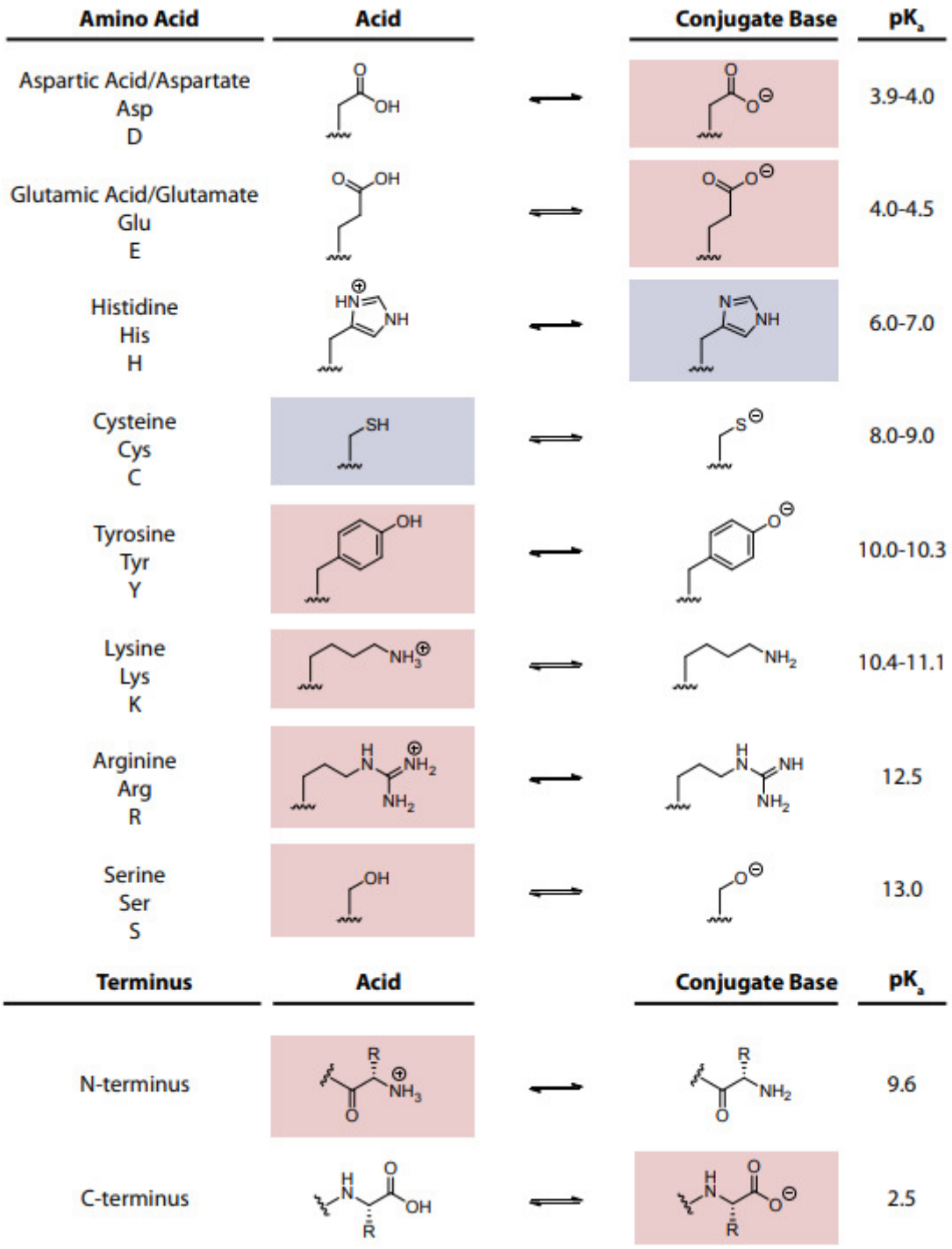

Figure 5: Ionizable functional groups are present in amino acids [Research Gate. Net].

At physiological pH, the neutrally charged, protonated form of cysteine is preferred; but, due to its low pKa, the deprotonated species is also present in large amounts. Other amino acids, like arginine and lysine, are basic and nearly exclusively reside in their protonated, positively charged forms at physiological pH. As a consequence, histidine exists at physiological pH as a combination of its positively charged, protonated form and its neutrally charged,

deprotonated form. Histidine is also basic, but the pKa of its conjugate acid is near to physiological pH. Since histidine's pKa ranges from 6.0 to 7.0, physiological pH favours the neutrally charged state somewhat more. Histidine is often employed by enzymes to transport protons during chemical reactions because its pKa is so near to physiological pH. Tyrosine and serine are two more amino acids that may be deprotonated at high pH levels, although they mostly reside in their protonated, neutrally charged forms at physiological pH levels.

At physiological pH, the polypeptide chain's N- and C-termini are ionised in addition to its side chains. The conjugate acid of the basic amino group at the N-terminus has a pKa of around 9.6. As a consequence, at physiological pH, the amino terminus is mostly protonated. With a pKa of around 2.5, the carboxyl group at the C-terminus is acidic. As a consequence, at physiological pH, the carboxy terminus is mostly deprotonated. Due to the delocalization of its lone pair, the amide nitrogen in the peptide bond cannot be protonated in water.The amino acids having ionizable side chains. It is underlined that the major ionization state occurs at physiological pH. If the specified ionization form is at least 100 times more abundant than the other form, the highlighted structures are highlighted in red; otherwise, they are highlighted in purple [9], [10]. The ionization of the free amino and carboxy groups at the termini of the polypeptide chain.

CONCLUSION

In conclusion, the shape and behaviour of polypeptide chains are significantly influenced by the structural and chemical properties of the peptide bond. The peptide bond's unusual structure, which includes both sigma andpi bonding orbitals, places restrictions on its ability to spin, resulting in a degree of stiffness uncommon for single bonds. Since pi orbitals must align parallel to one another in order to form double bonds, as observed in molecules like ethylene, this stiffness results from this necessity. Additionally, reducing steric conflicts between amino acid side chains via the preference for the trans configuration in peptide bonds increases the stability of protein structures. Understanding the 20 amino acids' chemical characteristics is crucial for understanding how proteins behave, especially their ionizable side chains.

These characteristics raise the complexity of protein structure and function coupled with ionization at the N- and C-termini of polypeptide chains. In conclusion, the study of protein structure and function is greatly aided by the intricate details of peptide bond behaviour, geometric isomerism, and amino acid properties, which provide light on the amazing details of the molecular world that underpin life's essential activities.

REFERENCES:

[1] M. Crisma, F. Formaggio, and C. Toniolo, Insight into peptide bond formation from 3D-structural chemistry, *Amin. Acids, Pept. Proteins*, 2016.

[2] N. E. Wezynfeld, T. Frączyk, and W. Bal, Metal assisted peptide bond hydrolysis: Chemistry, biotechnology and toxicological implications, *Coordination Chemistry Reviews*. 2016.

[3] K. Ulbrich, K. Holá, V. Šubr, A. Bakandritsos, J. Tuček, and R. Zbořil, Targeted Drug Delivery with Polymers and Magnetic Nanoparticles: Covalent and Noncovalent Approaches, Release Control, and Clinical Studies, *Chemical Reviews*. 2016.

[4] B. Albada and N. Metzler-Nolte, Organometallic-Peptide Bioconjugates: Synthetic Strategies and Medicinal Applications, *Chemical Reviews*. 2016.

[5] A. L. Pilo and S. A. McLuckey, Selective Gas-Phase Ion/Ion Reactions: Enabling Disulfide Mapping via Oxidation and Cleavage of Disulfide Bonds in Intermolecularly-Linked Polypeptide Ions, *Anal. Chem.*, 2016.

[6] F. Wu, J. P. Mayer, A. N. Zaykov, F. Zhang, F. Liu, and R. D. DiMarchi, Chemical Synthesis of Human Insulin-Like Peptide-6, *Chem. - A Eur. J.*, 2016.

[7] M. Luisa Di Gioia, A. Leggio, F. Malagrinò, E. Romio, C. Siciliano, and A. Liguori, N-Methylated α-Amino Acids And Peptides: Synthesis And Biological Activity, *Mini-Reviews Med. Chem.*, 2016.

[8] T. Lühmann, S. K. Mong, M. D. Simon, L. Meinel, and B. L. Pentelute, A perfluoroaromatic abiotic analog of H2 relaxin enabled by rapid flow-based peptide synthesis, *Org. Biomol. Chem.*, 2016.

[9] R. Gianatassio *et al.*, Organic chemistry: Strain-release amination, *Science (80-.).*, 2016.

[10] D. Sengupta, B. Roy, and B. Basu, Microwave-assisted Formation of Organic Disulfides of Biochem ical Significance, *Curr. Med. Chem.*, 2016.

CHAPTER 3

AROMATIC AMINO ACIDS, PEPTIDE CONFORMATION, AND PROTEIN FUNCTION: EXPLORING THE CHEMISTRY OF BIOMOLECULES

Dr. Jyoti Trivedi, Assistant Professor
Department of Biochemistry, TMMC&RC, TeerthankerMahaveer University, Moradabad, Uttar Pradesh India,
Email Id- jyoti.trivedi01@gmail.com

ABSTRACT:

Tryptophan, tyrosine, and phenylalanine are three aromatic amino acids that have distinct characteristics and functions. These amino acids have aromatic rings with delocalized electrons as a result of resonance. These aromatic rings are essential to the functioning of enzymes and UV absorbance for protein measurement in biological systems. The conformation of the peptide backbone is also covered, emphasising how the size and flexibility of the side chains of the amino acids affect how proteins fold. Due to their different characteristics, glycine and proline are demonstrated to have a considerable influence on the shape of the peptide backbone. The article also explores how cysteine side chains create disulfide bonds, which aid in the stabilisation and stiffness of proteins. With reference to the chemistry of hair, a practical example of disulfide bonding is shown. It is also investigated how amino acids work together in proteins, particularly in enzymes. An example of how amino acids cooperate in enzymatic activities is the catalytic triad of proteases, which is made up of aspartate, histidine, and serine. The article concludes by highlighting the basic significance of amino acid sequences in defining the structure and function of proteins, since even little changes may have a profound impact on how proteins fold and behave. We present the varied chemical properties of amino acid side chains and provide the groundwork for a more in-depth investigation of their effects on protein structure and function.

KEYWORDS:

Aromatic Amino Acids, Biomolecule, Chemistry, Peptide Conformation, Protein.

INTRODUCTION

One of the core principles of protein folding is the impact of amino acid side chains on the shape of the peptide backbone. These side chains, which differ in size and flexibility, influence the overall three-dimensional structure of proteins by steering the peptide backbone towards certain conformations. Glycine's flexibility and proline's distinctive side chain both bring unique obstacles and possibilities in protein folding. Amino acids like glycine and proline have unusual impacts on backbone conformation. In addition, the creation of disulfide bonds through cysteine side chains is essential for protein stability, increasing the stiffness of proteins and enabling certain conformations. The practical importance of disulfide bonding is shown through instances from the real world, such as the chemistry of hair treatments. Furthermore, the complex cooperation that underpins biological activities is shown by the cooperative interactions of amino acids inside proteins, particularly in enzymes. The catalytic trio in proteases is an illustration of how certain amino acid configurations optimise chemical processes inside the folded protein structure. The essay also emphasises how important the amino acid sequence is in defining the structure and function of proteins. The example of sickle-cell disease serves as a reminder that even little changes to this region may have a significant impact on how proteins fold. Proteins exhibit a wide range of behaviours and

functions as a result of the different chemical properties of their side chains, including polarity, charge, and ring formations [1], [2].

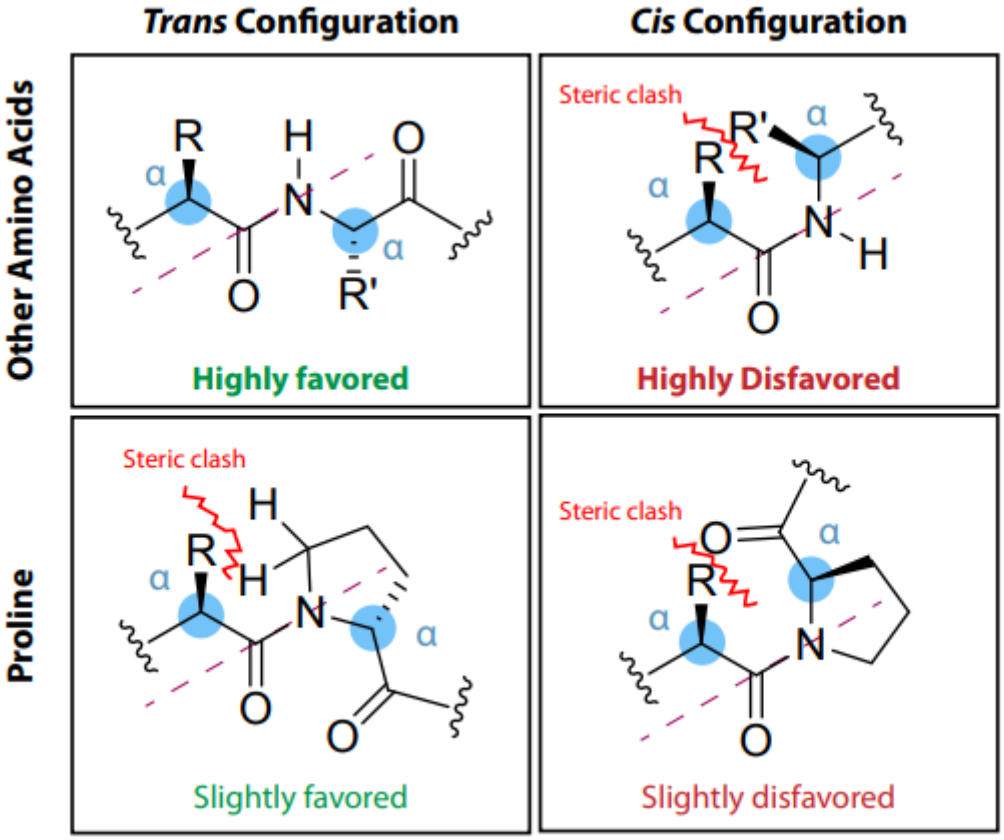

Figure 1: Peptide linkages created by proline may take on cis or trans configurations [Research Gate. Net].

Tryptophan, tyrosine, and phenylalanine all have aromatic rings. Phenylalanine, tyrosine, and tryptophan are three aromatic amino acids that are special in that they have rings with alternating double bonds. In reality, these rings' double bonds are not fixed. Instead, resonance causes the electrons to become delocalized among all the carbon atoms. Properties that are imparted by aromaticity are used in both the lab and in nature. For instance, they play a crucial role in the operation of several enzymes because they act as pathways for the transfer of electrons from one aromatic side chain to another. They are also helpful in the lab because they quickly absorb UV light, making it easier to quantify the amounts of proteins in solution.The conformation of the peptide backbone is affected by the side chains of amino acids. A polypeptide chain's side chains, as we previously said, control how it folds into a three-dimensional shape. This happens in part because the size and flexibility of the amino acid side chains vary, which in turn causes the peptide backbone to lean towards certain conformations. Alanine and glycine are two amino acids that have substantially smaller side chains than tryptophan and phenylalanine, respectively. Some amino acids also have more flexible side chains than others. In contrast to isoleucine and valine, which both contain a methyl (-CH3) group on the side chain near to the peptide backbone on the beta carbon, the second carbon from the carbonyl carbon, methionine, for example, may adopt many more conformations. Because it causes steric conflicts when certain conformations are adopted, this methyl group limits the conformations that the peptide backbone may take.

Even more dramatic impacts on the peptide backbone are those of glycine and proline. The peptide backbone may adopt a wider range of conformations around the two non-peptide bonds in the repeating unit of the peptide backbone because glycine basically lacks a side chain. Rotation around the connection forming the amine and carbon atom is prevented by the side chain of proline, which is joined to its amino group (Figure 1). This is unusual since in every other amino acid, the amine-carbon bond is free to spin. The only amino acid for which the trans peptide bond structure has an energy that is comparable to the cis form is proline. All other amino acids have a hydrogen substituent on the nitrogen of the amino group and a side chain-bound carbon atom in their peptide backbones. Because the hydrogen atom is substantially less bulky than the carbon atom, the trans form is favoured (Figure 1). For proline, however, there is always a steric conflict between a carbon atom immediately linked to nitrogen and the carbon atom of the next amino acid since its side chain is directly bound to its amino group.Due to steric incompatibilities in the cis configuration, the majority of amino acids form peptide bonds that greatly favour the trans configuration. Proline is an outlier due to steric conflicts that are brought on by its side chain in both the trans and cis forms. As a result, proline may occasionally serve as the nitrogen source for peptide bonds, giving them a cis structure [3], [4].

DISCUSSION

In proteins, cysteine side chains may create non-peptidic connections. The side chain of the amino acid cysteine has a sulfhydryl (-SH) group. Sulfhydryls are capable of being oxidised to create disulfide bonds, in which two cysteine side chains that are often located at different points in the main sequence join together to form a covalent sulfur-sulfur link. Such disulfide link production in proteins makes the protein stiff and may stabilise conformations that are not generally preferred.

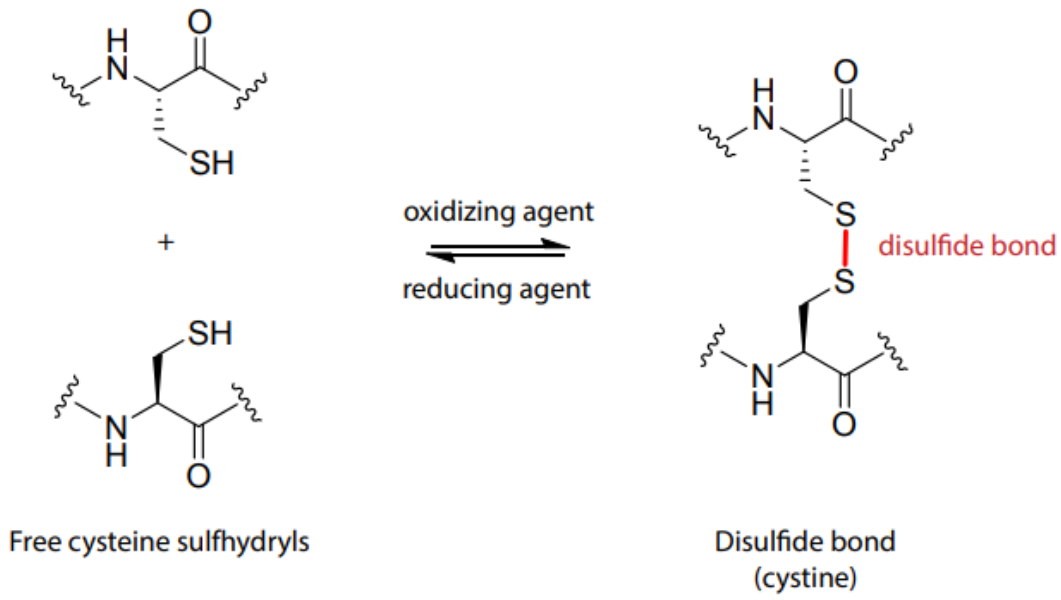

Figure 2: The reaction formula for creating a disulfide bond between two cysteine side chains is shown. Red is used to indicate the disulfide bond [Research Gate. Net].

Cysteine is the collective name for a pair of cysteine residues that are disulfide-bonded. Your hair serves as a well-known illustration of disulfide bonding. These connections are prevalent

in hair because they are crucial to its strength. Beauty parlors benefit from your hair's disulfide connections. If you were born with straight hair but now desire curly hair, you may visit the salon and obtain a perm. Your hairdresser treats your hair with a reducing agent, such thioacetal acid, which causes the disulfide bonds to dissolve and chemically change back into free cysteines. In order to create new disulfide bonds, your hair is then treated with an oxidizing substance, often hydrogen peroxide, after being curled around rollers to bring various sulfhydryl groups together. You now have curly hair instead of straight hair. Disulfide bonds are covalent connections that develop between the side chains of cysteine. Figure 2 shows the chemical equation that results in the production of a disulfide bond between two cysteine side chains is shown. Red represents the disulfide bond. The figure 3 depicts the formation of a disulfide bond (red) between two different parts of a polypeptide chain (black). Protein folding is encouraged by the tight conformations that the disulfide bonds force the polypeptide chain into.

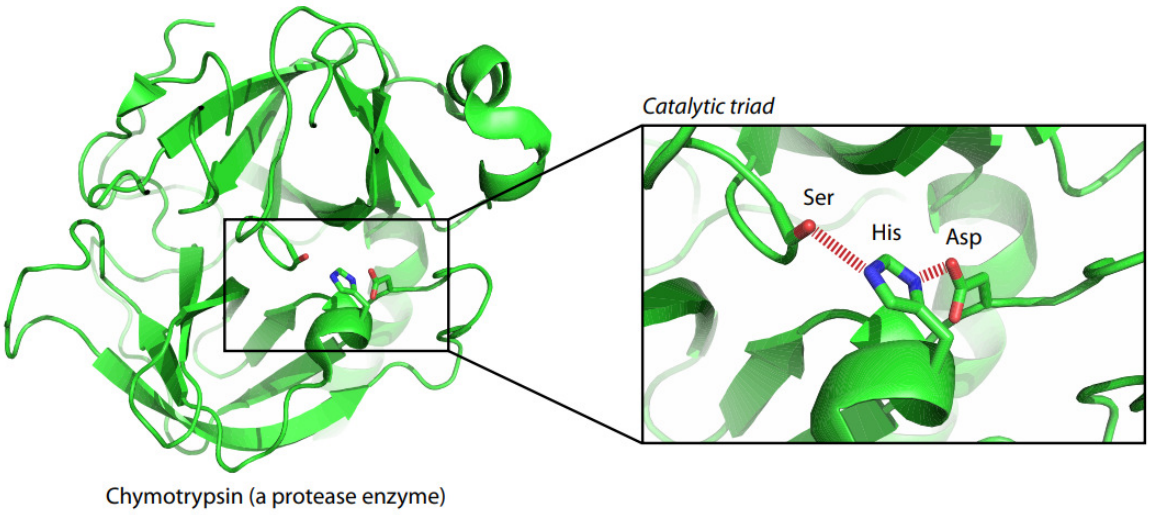

Figure 3: The functional elements of proteins are their side chains of amino acids [Springer].

Especially in enzymes, amino acids often cooperate with one another. Chymotrypsin is an enzyme that breaks peptide bonds as an example. Three amino acids aspartate, histidine, and serinethat work together to create a catalytic triad are used to speed up this chemical process. Figure 3 shows the side chains of amino acids in proteins are what make them function.In proteins, amino acids often cooperate with one another. The side chains of amino acids that we have been discussing are essential parts of how proteins operate. In circumstances when side chains are directly engaged in the processes the enzymes catalyze, many proteins also serve as enzymes. Proteases are a family of enzymes that have evolved to facilitate the dissolution of peptide bonds in proteins. In many proteases, a catalytic triadthree amino acids arranged in a certain wayworks in tandem to hydrolyze peptide bonds. Aspartate, histidine, and serine make up the catalytic triad. This catalytic triad is used by the enzyme to promote peptide bond hydrolysis. Of course, this is only possible because each of these amino acids is exactly positioned in three dimensions by the folded protein, allowing for optimal interactions between the amino acids and the substrate. Later chapters will cover enzyme catalysis in more detail, but first we must look at the folded structures of proteins and the thermodynamic principles behind their creation.

The most varied and adaptable macromolecules discovered in biological systems are proteins. The polymeric chains of amino acid monomers that make up proteins are joined by

covalent peptide bonds. Proteins, unlike the majority of other biological polymers, fold into specific structures with unique physical and chemical characteristics. Proteins are therefore used by cells for a variety of structural, catalytic, regulatory, and transport processes. A protein's fundamental structure, or amino acid sequence, alone determines how it is folded. In fact, as we saw in the case of sickle-cell disease, even minor alterations in the amino acid sequence have the capacity to modify the folded structure of a protein. All of the amino acids that make up proteins have a core carbon atom that is joined to a R group, a hydrogen atom, an amino (-NH2) group, a carboxylic acid (-COOH) group, and a carbon atom. The identification of the amino acid as well as its chemical characteristics are determined by the side chain's chemical makeup. The carbon in every amino acid is chiral, which means that it cannot be superimposed with its mirror image, with the exception of glycine, which has a hydrogen atom as its side chain. The stereochemistry, or unique arrangement at a chiral centre, may have a significant impact on the molecular characteristics. Enantiomers are molecules that are mirror copies of one another; one enantiomer is referred to as D and the other as L. Only the L-enantiomers of amino acids are used by living systems to build proteins[5], [6].

Peptide bonds join amino acids to create polypeptide chains. The N-terminus, which has a free amino group, and the C-terminus, which has a free carboxylic acid group, are the two ends of polypeptides. Both of these termini are ionized at physiological pH. But water cannot protonate the amide in the peptide backbone. Peptide bonds are flat, polar, and rigid; they cannot easily spin.

These characteristics affect the peptide backbone's potential conformations, which in turn affect how proteins fold. Additionally, amino acid side chains have an impact on the conformations of the peptide backbone by causing steric conflicts that direct the peptide backbone towards certain configurations. Backbone conformation is significantly influenced by glycine and proline. Due to glycine's absence of a side chain, it may take on a variety of conformations, making glycine-rich polypeptides more flexible. However, proline limits the backbone's conformational options to a smaller set. Proline has a special side chain that allows it to sometimes participate in cis peptide bonds, which are normally undesirable for other amino acids.Proteins include amino acids, which have a wide variety of characteristics that make them difficult to classify. There are polar side chains, nonpolar side chains, and slightly polar side chains. At physiological pH, some are positively or negatively charged. Others have ring formations that combine polar and nonpolar characteristics. We'll see how this range of amino acid characteristics affects the structure, folding, and other behaviors of proteins in the next chapter [7], [8].

CONCLUSION

In conclusion, this thorough investigation of amino acids, with a special emphasis on aromatic amino acids such as tryptophan, tyrosine, and phenylalanine and their function in proteins, offers insightful information into the complex field of biochemistry. With their delocalized electron systems, aromatic amino acids not only contribute to the variety of chemical characteristics in proteins but also play a crucial role in enzyme action and UV light absorption, assisting in the measurement of proteins. In essence, this investigation lays the groundwork for a fuller comprehension of the complicated world of proteins and their crucial functions in biological systems. By bridging the gap between chemical structures and biological activities, amino acids' special qualities and interactions are crucial for understanding the molecular secrets of life.

REFERENCES:

[1] O. Van Der Poorten, A. Knuhtsen, D. Sejer Pedersen, S. Ballet, and D. Tourwé, Side Chain Cyclized Aromatic Amino Acids: Great Tools as Local Constraints in Peptide and Peptidomimetic Design, *Journal of Medicinal Chemistry*. 2016.

[2] C. R. Forbes, A. K. Pandey, H. K. Ganguly, G. P. A. Yap, and N. J. Zondlo, 4R- and 4S-iodophenyl hydroxyproline, 4R-pentynoyl hydroxyproline, and S-propargyl-4-thiolphenylalanine: Conformationally biased and tunable amino acids for bioorthogonal reactions, *Org. Biomol. Chem.*, 2016.

[3] S. S. Bag, S. Jana, and M. K. Pradhan, Synthesis, photophysical properties of triazolyl-donor/acceptor chromophores decorated unnatural amino acids: Incorporation of a pair into Leu-enkephalin peptide and application of triazolylperylene amino acid in sensing BSA, *Bioorganic Med. Chem.*, 2016.

[4] A. Mollica *et al.*, Preparation of Constrained Unnatural Aromatic Amino Acids via Unsaturated Diketopiperazine Intermediate, *J. Heterocycl. Chem.*, 2016.

[5] N. DeGraan-Weber, D. C. Ashley, K. Keijzer, M. H. Baik, and J. P. Reilly, Factors Affecting the Production of Aromatic Immonium Ions in MALDI 157 nm Photodissociation Studies, *J. Am. Soc. Mass Spectrom.*, 2016.

[6] F. Merlino *et al.*, Structure–Activity Study of the Peptides P5U and Urantide by the Development of Analogues Containing Uncoded Amino Acids at Position 9, *ChemMedChem*, 2016.

[7] K. P. Chua, L. Y. Chew, and Y. Mu, Replica exchange molecular dynamics simulation of cross-fibrillation of IAPP and PrP106-126, *Proteins Struct. Funct. Bioinforma.*, 2016.

[8] S. S. Bag, S. Jana, M. K. Pradhan, and S. Pal, Trichromophoric pentapeptide: Impact of β-sheet conformation on dual path to excimer emission and sensing of BSA, *RSC Adv.*, 2016.

CHAPTER 4

UNRAVELING THE COMPLEXITY OF PROTEASES AND AMINO ACIDS: FROM DIVERSITY TO CHIRALITY

Ravi Kumar, Assistant Professor
College of Paramedical Sciences, TeerthankerMahaveer University, Moradabad, Uttar Pradesh, India Email Id- forensicravi001@gmail.com

ABSTRACT:

The extraordinary variety of proteases, a varied class of macromolecules, is due to the 20 amino acid monomers' combinatorial potential. The size and structure of every protein molecule may theoretically be altered by combining amino acids. The small number of protein sequences found in all living things stands in stark contrast to the enormous universe of possible protein sequences. The complex interaction between structural and functional characteristics of naturally occurring proteins, which have developed over billions of years in response to selection forces, is the cause of this difference. Specific binding sites for molecules, a careful balance between structural flexibility and stiffness, surface qualities appropriate for their immediate surroundings, and vulnerability to environmental variables are important characteristics that contribute to this variety. Proteins may be further distinguished by the amount, general makeup, and sequence differences of amino acids. The underlying structural and chemical properties of amino acids are examined in this chapter. The structural and functional properties of peptides and proteins, as well as the subtleties of protein folding, are then covered. The crucial relationship between polypeptide structure and function is often emphasised. Amino acids have other biological functions besides serving as the building blocks of proteins. Others aid in the formation of intricate nitrogen-containing substances including nucleotides, nucleic acids, heme, and chlorophyll. Some act as chemical messengers. The changed amino acids produced by post-translational modifications, which are present in many proteins, affect the activity and function of those proteins.

KEYWORDS:

Amino Acids, Chirality, Proteases, Peptide, Protein.

INTRODUCTION

The group of macromolecules known as proteases is varied (Figure 1). The 20 amino acid monomers' combinatorial potential has a direct bearing on this variety. Theoretically, any size or configuration of protein molecules may be created by joining amino acids. Think of a hypothetical protein with 100 amino acids, for instance. An enormous 20100 different combinations are possible for this kind of molecule. However, only a tiny portion (probably no more than 2 million) of the billions of potential protein sequences are actually produced in all living things. The complex combination of structural and functional characteristics of naturally occurring proteins, which have developed over billions of years in response to selection pressure, serve as an essential illustration of the cause of this astonishing difference. The presence of binding sites that are specific for one or a small group of molecules, an appropriate balance of structural flexibility and rigidity so that function is maintained, a surface structure that is appropriate for a protein's immediate environment (i.e., hydrophobic in membranes and hydrophilic in cytoplasm), and vulnerability to environmental factors are among these. The quantity of amino acids (also known as amino acid residues), the overall amino acyl composition, and the amino acid sequence may all be used to differentiate between different proteins.

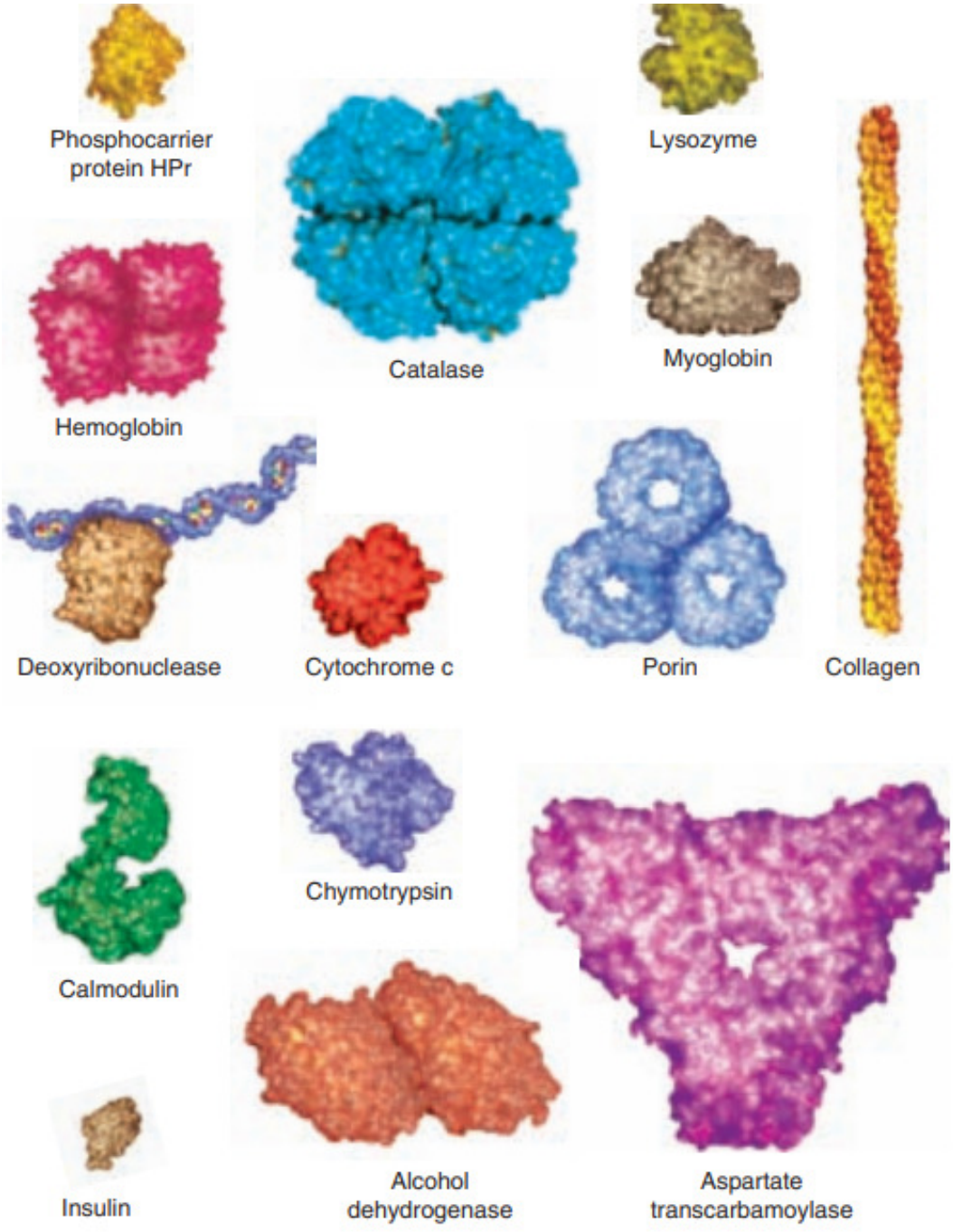

Figure 1: Presents a few instances of the variety of proteins [Slide Share].

Polypeptides are molecules with molecular weights between a few thousand and a few million daltons; peptides are molecules with low molecular weights, often with less than 50 amino acids. Molecules having more than 50 amino acids are known as proteins. A polypeptide chain or chains make up each protein [1], [2]. Reviewing the structures and chemical characteristics of the amino acids is the first section of this chapter. The structural and functional characteristics of peptides and proteins, as well as the method of protein folding, are discussed after this. The close connection between polypeptide structure and function is continually emphasised. The role of enzymes, a particularly significant class of proteins, is covered in Chapter 6. Chapter 19 discusses protein synthesis.

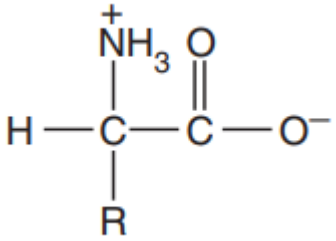

Figure 2: General Structure of the alpha-Amino Acids [Research Gate. Net].

The term standard amino acids refers to these amino acids. Figure 2 shows that 19 of the typical amino acids have the same basic structure. These compounds have a R (side chain) group, an amino group, a carboxylate group, a hydrogen atom, and a central carbon atom. Proline is an exception to the rule because it has a secondary amino group that is generated by ring closure between the R group and the amino nitrogen. The peptide chain gains stiffness from proline because rotation around the -carbon is impossible. This structural characteristic has important effects on the function and structure of proteins with a high proline concentration.Nonstandard amino acids are those that are present in live organisms but are not present in proteins, or amino acid residues that have undergone chemical modification after being incorporated into a polypeptide [3], [4]. Proteins often include unusual amino acids as a consequence of posttranslational modifications, which are chemical processes that occur after protein production. A carboxyl group is in its conjugate base form (—COO-) while an amino group is in its corresponding acid form (—NH3) at a pH of 7. As a result, each amino acid has the ability to act as an acid or a base. This characteristic is referred to as amphoteric. Zwitterions are molecules that have both positive and negative charges. Each amino acid receives its own characteristics from the R group.

DISCUSSION

Classes of Amino Acids

The ultimate three-dimensional configuration of each protein is determined by the amino acid sequence;thus, the structures of each protein are thoroughly explored in the next four subsections. The ability of an amino acid to interact with water determines how it is categorized. Four classifications may be identified using this criterion:

1. Nonpolar.
2. Polar.
3. Acidic.
4. Basic.

Nuclear Amino Acids

The majority of the R groups in nonpolar amino acids are hydrocarbons, which do not have positive or negative charges. Because they interact poorly with water, nonpolar amino acids are crucial in preserving the three-dimensional structures of proteins. This category contains both aliphatic and aromatic hydrocarbon side chains.

Cyclic structures seen in aromatic hydrocarbons make up a type of unsaturated hydrocarbons with planar conjugated electron clouds. One of the most basic aromatic hydrocarbons is benzene (Figure 3). Aliphatic describes nonaromatic hydrocarbons like cyclohexane and methane. Tryptophan and phenylalanine both have aromatic ring structures. Aliphatic R groups are present in glycine, alanine, valine, leucine, isoleucine, and proline. Methionine

and cysteine's aliphatic side chains both contain a sulphur atom. The side chain of methionine has a thioether group (—S—CH3).

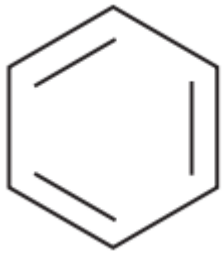

Figure 3: Representing the structure of the Benzene [Mol port].

Its product S-adenosyl methionine (SAM) is a significant metabolite that participates in a variety of metabolic processes as a methyl donor. Polar amino acids readily interact with water because they contain functional groups capable of hydrogen bonding. Polar amino acids are said to be hydrophilic, or water-loving.This group includes the amino acids serine, threonine, tyrosine, asparagine, and glutamine. Because they have polar hydroxyl groups, the amino acids serine, threonine, and tyrosine may engage in hydrogen bonding, a crucial aspect of protein structure. In proteins, the hydroxyl groups have additional uses. One such regulating mechanism is the production of the phosphate ester of tyrosine. Serine and threonine's OH groups are other sites where carbs might be attached. The acidic amino acids aspartic acid and glutamic acid, respectively, are the sources of the amide derivatives asparagine and glutamine. The capacity of asparagine and glutamine to form hydrogen bonds with the amide functional group, which is extremely polar, has a considerable impact on protein stability. The cysteine sulfhydryl group (—SH), which is extremely reactive and a crucial part of many enzymes. Additionally, it binds metal ions in proteins, such as iron and copper ions. Two cysteine molecules' sulfhydryl groups may also readily oxidise in the extracellular space to create the disulfide complex cystine [5], [6].

Amino Acids in Acid

There are carboxylate groups on the side chains of two common amino acids. Aspartate and glutamate are common names for the side chains of aspartic acid and glutamic acid because they are negatively charged at physiological pH.

Primary Amino Acids

At physiological pH, basic amino acids have a positive charge. Consequently, they are able to create ionic connections with acidic amino acids. Lysine, which possesses an amino group on its side chain, combines with water to generate the conjugate acid (—NH3). Strong intramolecular and intermolecular cross-linkages are created when lysine side chains in collagen fibrils, a crucial structural element of ligaments and tendons, are oxidised and then condensed.Proteins' guanidino group, which has a pKa range of 11.5 to 12.5 and is continuously protonated at physiological pH, is unable to participate in acid-base processes because of this. On the other hand, the imidazole side chain histidine is a weak base as its pKa is around 6, and it is only slightly ionised at pH 7. Numerous enzymes' catalytic activity is greatly influenced by their ability to receive or donate protons in response to minute pH changes when they occur under physiological settings.

Amino Acids That Are Active in Biology

Aside from serving as the building blocks of proteins, amino acids also have a number of other biological activities.

1. A number of -amino acids and their derivatives function as chemical messengers (Figure 4). For instance, neurotransmitters are chemicals produced from one nerve cell that affect the operation of another nerve cell or a muscle cell. Examples of neurotransmitters include glycine, glutamate, -amino butyric acid (GABA, a derivative of glutamate), and serotonin and melatonin.

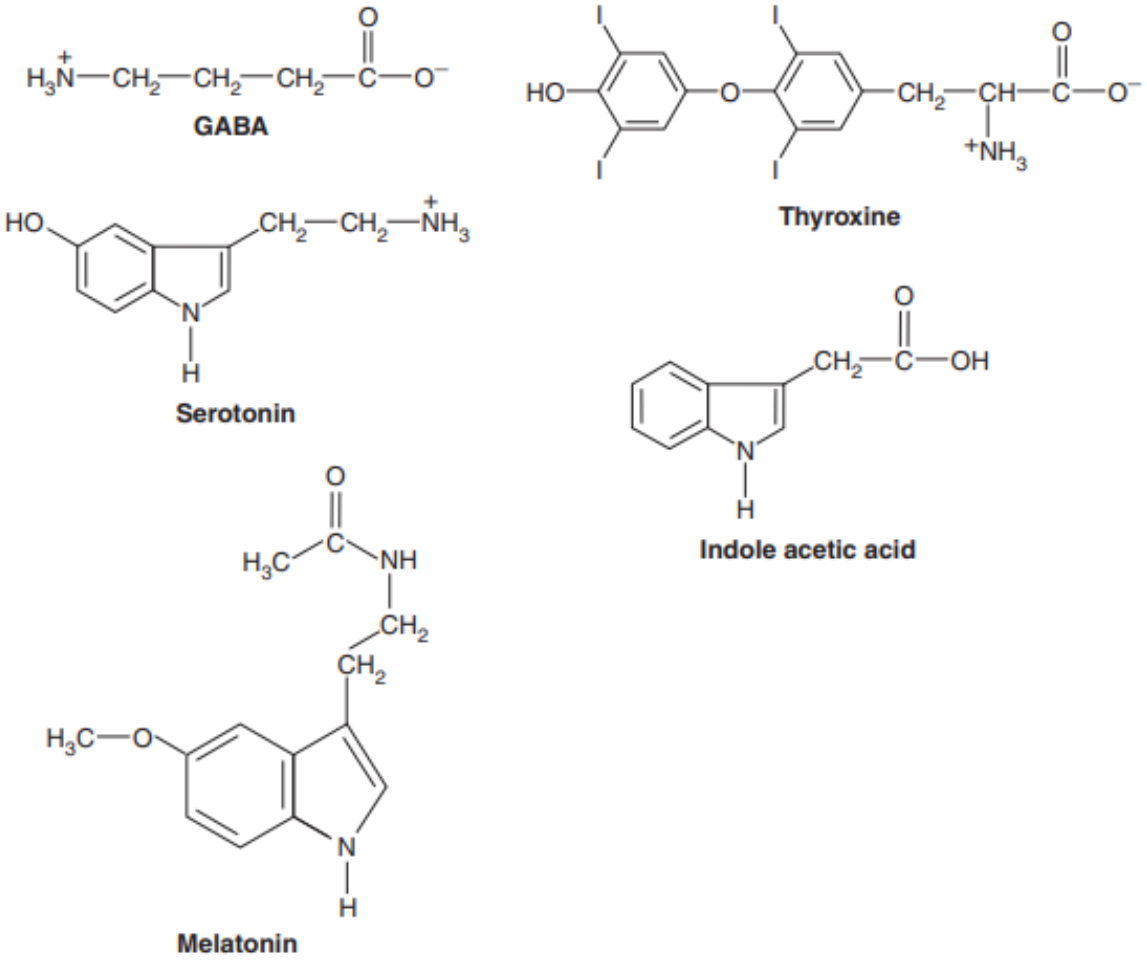

Figure 4: Representing the Some Derivatives of Amino Acids [Research Gate. Net].

Hormones are chemical signal molecules generated in one cell that control the activity of other cells, such as thyroxine a tyrosine derivative produced in the thyroid gland of mammals and indole acetic acid a tryptophan derivative found in plants.A variety of complicated nitrogen-containing compounds may be generated from amino acids. Examples include the nitrogenous bases found in nucleotides and nucleic acids, hemean organic group containing iron that is necessary for the biological function of a number of key proteinsand chlorophylla pigment that is essential for photosynthesis. A number of common and uncommon amino acids serve as metabolic intermediaries. The urea cycle, for instance, is made up of amino acids like arginine (Figure 2), citrulline, and ornithine (Figure 5). The main method for getting rid of nitrogenous waste is the synthesis of urea, a chemical created in the livers of vertebrates [7], [8].

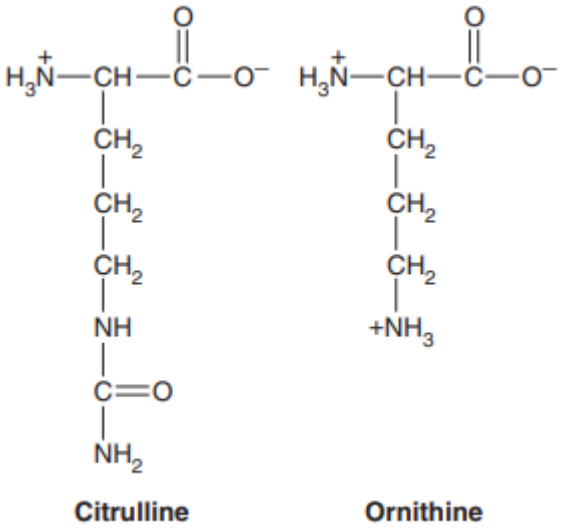

Figure 5: Representing the structure of Citrulline and Ornithine [Research Gate. Net].

Proteins with Modified Amino Acids

After a polypeptide chain has been created, amino acid derivatives are generated, and many proteins include these derivatives. Carboxylation acid, a calcium-binding amino acid residue present in the blood-clotting protein prothrombin, is one of these modified amino acids (Figure 6). Collagen, the most prevalent protein in connective tissue, has both 4-hydroxyproline and 5-hydroxylysine as significant structural constituents. Protein activity is often controlled by phosphorylation of the hydroxyl-containing amino acids serine, threonine, and tyrosine. For instance, when the enzyme glycogen synthase is phosphorylated, the production of glycogen is greatly reduced.

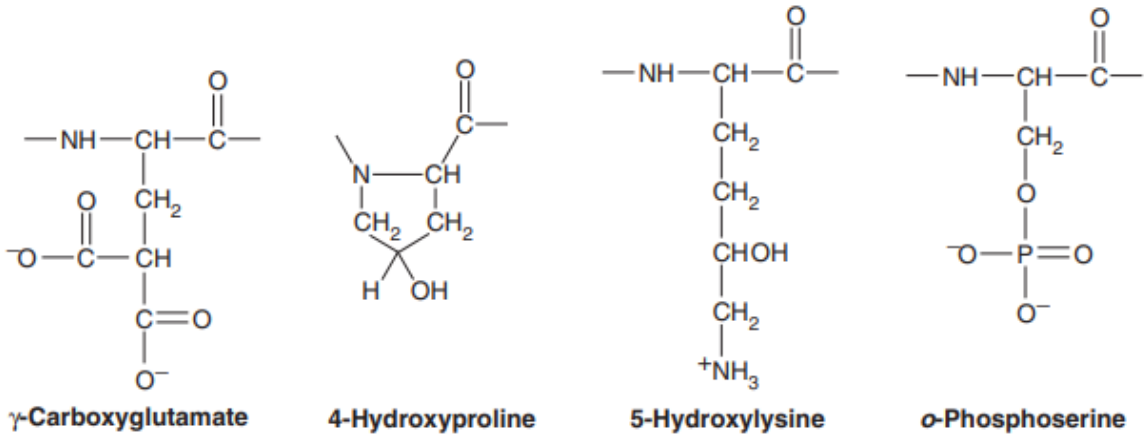

Figure 6: Some Modified Amino Acid Residues Found in Polypeptides [Science Direct. Com].

Animal Protein Stereoisomers

The -carbons of 19 of the 20 standard amino acids are referred to as asymmetric, or chiral, carbons because they are joined to four distinct groups namely, a hydrogen, a carboxyl group, an amino group, and a R group. Because the -carbon of the molecule glycine is joined to two hydrogen atoms, it is symmetrical. Chiral carbon compounds may exist as stereoisomers,

which are identical molecules that vary only in their atoms' spatial arrangements. Figure 7 shows stereoisomer representations of amino acids in three dimensions. With the exception of where the ammonium group and the hydrogen atom are located, the two isomers' atoms are bound to one another in the identical arrangement in the illustration.

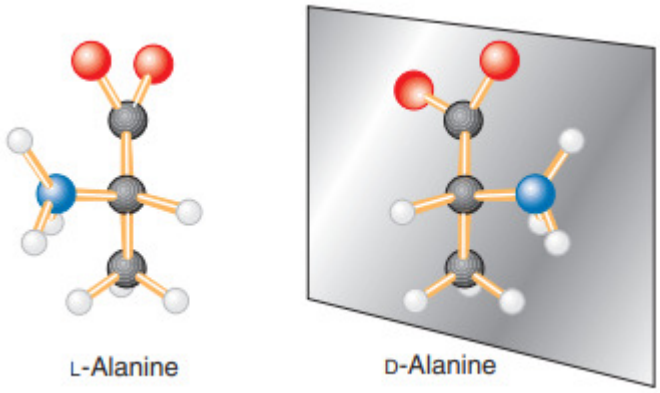

Figure 7: Two Enantiomers L-Alanine and D-alanine are mirror images of each other. (Nitrogen = large red ball; Hydrogen = small red ball; Carbon = black ball; Oxygen = blue balls [Quizlet].

The mirror images of these two isomers are one another. Enantiomers of this kind of molecule can't be overlaid on one another. Except for the direction in which they turn plane-polarized light, enantiomers are physically identical. Unpolarized light is transformed into plane-polarized light by passing it through a unique filter; the light waves only vibrate in one plane. Molecular entities with this characteristic are referred to be optical isomers. The reference substance for optical isomers is glyceraldehyde (Figure 8). One glyceraldehyde isomer, known as dextrorotatory (marked by), spins the light beam in a clockwise manner[9], [10].

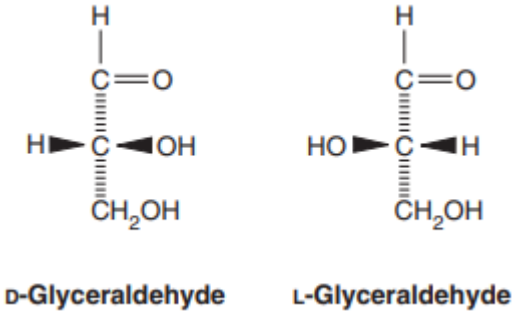

Figure 8: Representing the structure of D- and L-Glyceraldehyde [Springer].

The levorotatory isomer of glyceraldehyde spins the beam to the same extent but in the opposite direction. When the arrangement of atoms around an asymmetric carbon in a molecule resembles the arrangement of atoms around the asymmetric carbon in either of the glyceraldehyde isomers, optical isomers are often denoted as D or L. The majority of biomolecules include several chiral carbons. As a consequence, the letters D and L do not denote the direction in which a molecule spins plane-polarized light; rather, they denote the structural connection of a molecule to either of the glyceraldehyde isomers.The majority of asymmetries discovered in living things only exist in one stereoisomeric form, either D or L. For instance, proteins include solely L-amino acids, with a few exceptions.Chirality has a

significant impact on the morphological and biological functions of molecules. For instance, only L-amino acids are present in proteins and cause the right-handed helices that are seen in them. D- and L-amino acid mixtures used to create polypeptides in the lab don't result in helices. Furthermore, most substrate compounds are bound by enzymes in just one enantiomeric form since enzymes are chiral molecules. Artificial polypeptides made of D-amino acids cannot be broken down by proteases, enzymes that break down proteins by hydrolyzing peptide bonds.

CONCLUSION

In conclusion, the protease family of macromolecules, a varied group, is a living example of the astounding diversity and complexity of biological molecules. The enormous combinatorial potential of the 20 amino acid monomers, which theoretically permits the production of protein molecules of any size or structure, is the source of this variety. However, only a tiny portion of the trillions of possible protein sequences are actually realized in living things. This striking distinction is highlighted by the complex interplay of structural and functional traits present in naturally occurring proteins, which have been fine-tuned through billions of years of evolution in response to selective forces. These proteins have unique molecular binding sites, retain their function by balancing structural flexibility and stiffness, adjust their surface structure to their immediate surroundings, and are sensitive to environmental variables. Variations in the number of amino acids present, the overall amino acyl composition, and the individual amino acid sequence all contribute to the differences across proteins. In essence, research on proteases and amino acids provides an intriguing look into the intricate workings of life's molecular system and demonstrates the wonders of nature's design and adaptability. Understanding these concepts enhances our grasp of biochemistry and holds the potential to opening up new directions in biotechnology, medicine, and our investigation of nature.

REFERENCES:

[1] A. Rasoolizadeh, M. C. Goulet, F. Sainsbury, C. Cloutier, and D. Michaud, Single substitutions to closely related amino acids contribute to the functional diversification of an insect-inducible, positively selected plant cystatin, *FEBS J.*, 2016.

[2] J. Ando *et al.*, Alkyne-Tag SERS Screening and Identification of Small-Molecule-Binding Sites in Protein, *J. Am. Chem. Soc.*, 2016.

[3] N. Y. Siddiqui and L. G. DuBois, Optimizing Urine Processing Protocols for Protein and Metabolite Detection, *J. Proteomics Bioinform.*, 2016.

[4] Z. Li, C. Wang, F. Jiang, P. Huan, and B. Liu, Characterization and expression of a novel caspase gene: Evidence of the expansion of caspases in Crassostrea gigas, *Comp. Biochem. Physiol. Part - B Biochem. Mol. Biol.*, 2016.

[5] E. L. H. Tang, C. H. Tan, S. Y. Fung, and N. H. Tan, Venomics of Calloselasma rhodostoma, the Malayan pit viper: A complex toxin arsenal unraveled, *J. Proteomics*, 2016.

[6] R. Abalo, Drugs and other strategies for management of neuropathic pain. Where are we now? Where are we going?, *J. Physiol. Biochem.*, 2016.

[7] D. D'Aliberti *et al.*, NS3 genetic variability in HCV genotype 1a isolates from liver tissues and serum samples of treatment-naïve patients with chronic hepatitis C, *Dig. Liver Dis.*, 2016.

[8] A. Bah and J. D. Forman-Kay, Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications From the ‡ Program in Molecular Downloaded from, *J. Biol. Chem.*, 2016.

[9] A. F. ZEINA *et al.*, Enfoque De Los Trastornos Del Equilibrio Acido-Base, *Biol. Eng. Trans.*, 2016.

[10] J. De Smet, Impact of phages on Pseudomonas aeruginosa physiology, *PhD thesis*, 2016.

CHAPTER 5

EXPLORING THE COMPLEXITIES OF BIOLOGICAL MEMBRANES: LIPIDS, PROTEINS, AND FUNCTIONS

Chintakayal Purnima, Assistant Professor
College of Paramedical Sciences, TeerthankerMahaveer University, Moradabad, Uttar Pradesh, India Email Id- chintakayal.purnima3@gmail.com

ABSTRACT:

The dynamic border that separates the intracellular cytoplasm from the external environment is provided by the plasma membrane, an essential part of cellular design. Due to the stark chemical contrasts between these two worlds, its development, around 3.2 to 3.8 billion years ago, was a significant turning point in the evolution of life. All biological organisms have plasma membranes, with some differences like the outer membranes of gram-negative bacteria and the interior membranes of eukaryotic organelles. The fundamental function of the plasma membrane is to preserve the particular chemical environment of the cell, which includes ions, minute metabolites, and macromolecules. Ion concentrations and electric potentials, which are crucial for cellular operations and have an impact on activities like cellular transport and muscle contraction, are maintained by the plasma membrane. Due to the membrane's lipid-based structure, which limits permeability, translocation across the polymer matrix is required for polar molecules. Transporters and channels found in membrane proteins enable the import and export of essential molecules, preserving metabolic stability and limiting unwelcome material influxes. In addition to transport, membrane proteins are essential for signalling, cell recognition, energy generation, pathogen defence, and cellular motility. They make up a significant fraction of the genome and proteome, highlighting their importance in health and illness and having effects on a number of diseases. The foundation for in-depth research into the structure, organisation, and functions of biological membranes is laid by this thorough review, which emphasises the complex and dynamic character of biological membranes.

KEYWORDS:

Biological Membranes, Cell, Lipids, Polymer, Proteins.

INTRODUCTION

Understanding the structure of biological membranes, which is chemically and physically more complex than the cytoplasmic environment, is crucial to understanding the behaviour of membrane proteins. Singer and Nicolson's fluid mosaic model defines membranes as lipid bilayers dotted with proteins. The distinct characteristics of the membrane are influenced by membrane proteins, which may be peripheral or integral. Each of the several lipid molecules found in lipid bilayers, such as glycerophospholipids, sphingolipids, and sterols, has unique properties. For instance, glycosphingolipids have a role in the identification of molecules. Sterols regulate membrane fluidity and provide stiffness, as cholesterol does in mammals. Archaeans' ethers improve stability in harsh conditions. Varied species, organ systems, cellular compartments, leaflets, and microdomains have membranes with varied lipid compositions, reflecting distinct functional needs. The charge and functioning of the membrane are influenced by the asymmetrical distribution of lipids between the cytoplasmic and exoplasmic leaflets. Lipid rafts are one kind of microdomain that further diversity the characteristics and uses of membranes. Uncovering the varied functions of membrane proteins and their contributions to cellular physiology and disease requires an understanding

of the intricacy of biological membranes. Given that many medications target membrane proteins, this information has significant implications for drug development.

The plasma membrane is a lipid structure made up of both protein and carbohydrate building blocks. It defines the structure of the cell and functions as a physical partition separating the cytoplasm from the exoplasm, or outside environment. The creation of the first membranes, which occurred between 3.2 and 3.8 billion years ago, marked the evolutionary emergence of the first living beings since the chemical compositions of these two environments are substantially different. All organisms' cells are surrounded by plasma membranes, and some cells also have additional membranes. For instance, gram-negative bacteria have an outer membrane that is loaded with polysaccharides and contains a unique collection of identifying proteins. Inner membranes in eukaryotic cells, which first appeared roughly 1.5 billion years ago, distinguish the different organelles [1], [2]. The plasma membrane's major function is to preserve the special chemical environment of the cell, which includes the following:

$$\Delta G_{\text{transfer}} = G_m - G_w = 166\left(\frac{q}{r}\right)\left(\frac{1}{\varepsilon_m} - \frac{1}{\varepsilon_w}\right)$$

Chemical diversity exists among intracellular ions. Na+, K+, Cl-, Mg2+, Mn2+, Cu2+, Zn+, Co2+, and Fe2+/3+ are examples of elements, while PO3 is an example of a molecule with a range of sizes. The regular execution of several cellular and physiological processes depends on maintaining a steady concentration of these ions within the cell. First, ions make it possible for normal metabolic steps to occur by stabilizing charged groups in the active sites of enzymes or taking part in the binding of atoms and molecules. Second, the plasma membrane's electric potential, which is employed to power procedures including cellular transport, neuronal transmission, and muscular contraction, is caused by the concentration gradient of Na+, K+, and Cl- across the membrane. Finally, the control of bodily hydration depends on the physiological ionic equilibrium.Metabolic processes include and produce small organic molecules such ATP, amino acids, monosaccharides and disaccharides, nucleotides, pyruvate, and others.

Since many of these substances are polar, evolution has opted to maintain them within the cell using a barrier based on a lipid structure, which has a very low permeability to such molecules. A polar chemical must migrate from the aqueous environment into the membrane itself before it may cross the membrane. Due to the polymer compound's desolvation, this procedure is very rare. The Born equation of solvation may be used to calculate the transfer energy as shown below. The transfer free energy is: assuming that electrostatic contributions govern ion transport into the membrane. (Where m and w are the membrane and aqueous environments, respectively; q is the net charge of the molecule; r is its effective radius; and is the dielectric.

Consider a straightforward spherical ion with a radius of 1 and a charge of +1. We get to Gtransfer = +80 kcal/mol by assuming that the membrane's and the aqueous environment's dielectrics are, respectively, 2 and 80. As a result, P e(G/RT) = e(80/0.6) 1 1058 is the proportion of ions that transfer into the membrane (Equation (4.3b)). In essence, no cation will partition into the membrane at a physio- logical ion concentration of roughly 150 mM. This is also true, according to computational andexperimental research, for big ions such the side chains of charged amino acids.It is necessary to continually import certain compounds and export others in order to maintain stable quantities in cells since the majority of the

substances they contain are constantly created and used by metabolic processes. Furthermore, every step of the process has to be closely monitored to prevent the loss of crucial metabolites or the internalization of wastes and/or harmful compounds. In this instance as in the others we have seen, proteins have been given the task by evolution. Channels and transporters are the two main categories of transport proteins. Polar chemicals may travel across the membrane along their electrochemical gradient thanks to a channel that runs the full length of the membrane and includes a water annulus.

Polar molecules are bound by a transporter on one side of the membrane and released on the opposite. Others must diffuse from one side to the other in order to discharge the substrate into the appropriate compartment since they are not able to cover the whole width of the membrane. While some transporters move molecules along their electrochemical gradient, othersknown as pumpsmove them in the opposite direction while drawing energy from an external source. This energy source may be direct (ATP), indirect another molecule's electrochemical gradient, or both. Both channels and transporters have mechanisms for controlling the transport process [3], [4]. For example, channels can open or close in response to a variety of signals, including ligand binding, changes in the cross-membrane voltage, or the application of mechanical pressure. Transporters, on the other hand, only allow molecules large enough to pass through the water annulus to enter. When their ‘substrates' are present, transporters have the ability to bind or release them, but often they do so while waiting for the correct regulatory ligand to attach. Because of their fundamental nature, transport proteins account for 40% to 50% of all membrane proteins in bacteria.

DISCUSSION

The fundamental roles of the plasma membrane include the control of ions, tiny metabolites, and macromolecules. It carefully regulates the flow of these substances into and out of the cell, preserving the ion gradients necessary for cellular functions and making sure that metabolites are available for metabolic pathways. To enable this movement, specialised transport proteins, like as channels, transporters, and pumps, are required due to the membrane's poor permeability to polar molecules. The plasma membrane is essential for transport as well as signal transduction, cellular recognition, energy production, pathogen defence, and cellular motility. These tasks depend on a variety of membrane-associated enzymes and transporters, membrane-bound receptors, and cellular recognition proteins. Particularly important are membrane proteins, which make up a significant component of the genome and proteome in diverse species. A variety of diseases may result from membrane protein dysfunction, underlining the importance of these proteins in both health and illness. The lipid bilayer's fluidity and the wide range of proteins incorporated into or loosely linked with it are highlighted by the Singer and Nicolson's fluid mosaic model, which highlights the dynamic character of biological membranes. Membrane carbohydrates that make up the glycocalyx have a role in molecular recognition in addition to providing physical protection.

Glycerophospholipids, sphingolipids, sterols, and ethers make up the lipid bilayer itself. Even within the same membrane, various species, organ systems, and other variations in its exact makeup are possible. Cells are able to adapt to certain activities and environmental factors thanks to the variety in their lipid content.Signal transduction and communication. Membrane-bound receptors are present in many biological proteins. These serve as antennas and transmit communication signals entering the cell from the outside world. The majority of these receptors react to chemical messengers, including hormones, neurotransmitters, pheromones, odorants, and local mediators, although some also react to other signals, including electromagnetic radiation (light) and mechanical pressure. Thus, these proteins' ability to operate properly is essential for not just individual cells but also for several

physiological systems, including the neurological, endocrine, and immunological systems. The whole width of the membrane is covered with receptors, which have an extracellular side intended to bind or react to the external messenger and an intracellular side meant to interact with various cytoplasmic proteins. The latter do this by often catalyzing enzymatic processes to transfer signals into the cell. For instance, growth factor receptors encourage the phosphorylation of several cellular constituents to send the message grow or divide into the cell. Despite the fact that most of the cytoplasmic proteins that are receiving this signal are water-soluble, some of them, including G-proteins and the enzymes protein kinase A and C, are membrane-bound at least sometimes.Recognition of cells and cells and ECM. The extracellular matrix (ECM) or other cells' surface proteins or other components may be bound by certain membrane-binding proteins, such as integrins or cadherins. For the cell to recognised its neighbours or get fixed to its biological tissue, for example, these interactions are crucial.

A variety of proteins found within the inner mitochondrial membrane or the bacterial plasma membrane take part in the process of obtaining chemical energy from food and storing it in the form of the readily available compound ATP. The majority of these proteins function as proton pumps and electron transporters, while others build the ATP-producing portion of this system. In the assimilation of solar energy, or photosynthesis, a similar mechanism operates in the thylakoid membrane of plants and algae or the plasma membrane of photosynthetic bacteria.Certain membrane-bound proteins help protect cells and the body as a whole against pathogens like bacteria, viruses, and parasites that are invading. These proteins may perform a variety of tasks, the majority of which involve the identification of chemicals associated with pathogens. The T-cell receptor, which spans a T lymphocytes membrane and recognises peptides obtained from pathogens that have been destroyed, is a well-known example [5], [6].Membrane proteins often act as sites where other proteins may bind. Cells may concentrate metabolic enzymes or signal-transduction proteins in specific areas because to this capability. Some membrane-bound proteins also influence the movement of vesicles that deliver lipids and proteins between cellular compartments.

The importance of membrane proteins is shown by their prevalence as well as by the fact that they are thought to make up 20% to 30% of any genome and 23% of the human proteome, respectively, according to a recent comprehensive proteome study. Pathologies caused by membrane protein defects include cancer, depression, obesity, and cardiovascular or brain problems. Therefore, it is thought that around 60% of pharmaceuticals that have been authorized act on membrane proteins, the majority of which are G protein-coupled receptors. We previously spoke about how proteins' surroundings have an impact on their structural, thermodynamic, and functional properties. Therefore, it seems sense that one must first understand the nature of biological membranes in order to understand the behaviour of membrane proteins. The fact that membranes are inherently far more complicated than the aqueous solution that makes up the cytoplasm, both chemically and physically, supports this view.

Biological systems: structure and organisation membranes

Basic structure and characteristics

Singer and Nicolson presented their renowned fluid mosaic (FM) model in 1972 as a way to describe the composition and properties of biological membranes. The membrane is shown in the model as a structure comprised of two lipid bilayers, each of which contains a variety of proteins. These proteins are divided into two categories in general:

1. Integral proteins are found within the lipid bilayer, with one or more segments of their polypeptide chains spanning the whole width of the bilayer. Detergents must dislodge the bilayer structure in order to isolate these proteins.
2. A lipid monolayer or an integral protein is when peripheral proteins are loosely bound. It is not necessary to break the membrane in order to isolate these proteins; a simple procedure, such increasing the salt content, would do.

Membranes also include several forms of carbohydrates, often in the form of lengthy, branched chains, in addition to lipids and proteins. On the extracellular side of the membrane, they are joined to molecules of lipid and protein. 'Glycocalyx' is the term used to describe the membrane's complete carbohydrate coat. The glycocalyx is much larger than it is shown in traditional biochemistry and cell biology textbooks, and it can be seen with an electron microscope. The chains of carbohydrates not only provide the membrane physical defence but also take part in molecular recognition procedures. These may happen inside the body's cells or between the cell and a water-soluble chemical.Numerous lipid molecules are densely packed together to form the bilayer structure, which serves as the foundation for the whole membrane. However, the FM model postulates that the bilayer is only moderately viscous since each lipid molecule is essentially dynamic thus the word fluid in the model's name. Early research focused on the protein-to-lipid ratio in various cellular membranes in order to verify this hypothesis. They came to the conclusion that certain membranes varied significantly in this characteristic, despite the fact that most biological membranes have a weight ratio of 0.5. For instance, the ratio is 0.2 in the myelin membrane, which encircles the nerve cells' axons, and 0.8 in the inner mitochondrial membrane.

A somewhat different technique was used in research on red blood cells by taking into account a different metric, the percentage of the membrane surface that is occupied by the protein component. The findings showed that proteins occupied at least 23% of the membrane surface, which is a much larger percentage than would be predicted based on the protein-to-lipid ratio. This high value was ascribed to the extensive extramembrane domains seen in many integral membrane proteins as well as the tendency of these proteins to form large oligomers. According to these findings, biological membranes are rather stiff, in contrast to the FM model's fluid representation and research done in pure lipid bilayers, which had previously dominated the scientific viewpoint. The membrane is now thought to have qualities that are halfway between fluid and gel, allowing it to impede the free flow of polar solutes while yet maintaining its flexibility, which is crucial to its function. For the production of transport vesicles, which transfer cargo such as proteins and lipids between internal membranes and the plasma membrane, flexibility is crucial. Exocytosis is a mechanism through which the cargo molecules are either secreted or remain in the membrane. Polar solutes may also be internalized by the cell via endocytosis thanks to transport vesicles [7], [8].

Makeup of the lipid bilayer

Diverse lipid molecules of diverse sorts make up the cellular membrane. Even red blood cells, which are thought to be quite basic, have about 200 distinct kinds of lipids in their plasma membranes. The structures and characteristics of the major categories are outlined in the next subsections.

Glycerophospholipids

The most prevalent form of lipids in biological membranes are glycerophospholipids. The name of these compounds relates to their chemistry, which consists of two fatty acids esterified to a glycerol backbone with a third glycerol carbon connected to a negatively

charged phos- pate group in each case. The alcohol group, which may be serine, choline, ethanolamine, inositol 4,5-bisphosphate (IP2), or even glycerol, is usually connected to the phosphate group on the opposite side of most phospholipids. Each phospholipid is given a name that begins with the prefix phosphatidyl and the name of the alcohol group it belongs to. It is the alcohol group that defines the entire physicochemical uniqueness of each phospholipid, including its size and electric charge, as the numerous phospholipids only vary in the identity of their distinct alcohol groups. The form of cardiolipin sets it apart significantly from other phospholipids. It includes two phosphate groups, two glycerol groups, and four acyl chains since its alcohol group is a complete phosphatidylglycerol group. While the acyl chain in the second place (sn-2) likely to be poly- unsaturated4, the acyl chain in the first glycerol position (sn-1) tends to be either saturated or monounsaturated. The majority of biological membrane acyl chains have 18 carbons resulting in an average hydrophobic width of 30.

A single acyl chain, known as a ceramide, is present and connected to the sphingosine backbone. However, the overall shape of the lipid molecule is still comparable to that of glycerol phospholipids because the structure of sphingosine has a long hydrocarbon chain that mimics an acyl chain. Even though the third carbon of sphingosine is often joined to a phosphocholine group, there are certain instances when the phosphate group may be swapped out for a large carbohydrate group. 'Glycosphingolipids' (GSLs) are the name given to such intricate molecules. N-acetylneuraminic acid, a sialic acid group, is covalently linked to the sugar moiety in certain GSLs. Gangliosides, as these GSLs are sometimes known, are especially common in neuronal membranes, where they account for 2% to 10% of the total lipid content.Animal cell membranes often include GSLs, which make up a particularly fascinating subclass of sphingolipids. These molecules are ideal for molecular recognition processes because of the intricate carbohydrate patterns in GSLs and the fact that the majority of GSLs are found on the outer leaflet of the lipid bilayer. In fact, a wide variety of extracellular ligands, including lectins, toxins, hormones, and viruses, are known to interact with GSLs. The development of the cell is known to have an impact on how carefully GSLs manage their membrane composition. Additionally, it has been shown that some aberrant processes, such as neurological disorders and the trans-formation of malignant cells, cause its composition to shift substantially.

Sterols

A third kind of lipid, called a sterol, is present in eukaryotes. Four fused rings with a hydroxyl group at one end and a lipid tail at the other make up the distinctive structure of sterols. The precise kind of sterol found in the membrane varies depending on the type of organism: stigmasterol, ergosterol, and cholesterol are found in plants, fungi, and mammals, respectively. Sterols are hefty and unyielding in comparison to the thin-flexible phospholipids. As will be discussed below, these two sterol characteristics have a significant impact on the characteristics of the whole membrane. The restricted concentration range of cholesterol in the membrane reflects the overall significance of cholesterol in the mammalian membrane. The cell keeps a close eye on this area.

Ethers

The oldest living things on Earth are archaea’s. It should come as no surprise that they often inhabit niches like the hydrothermal vents at the ocean's bottom, where harsh environmental conditions are similar to those that were prevalent on our planet 3.5 billion years ago. Archaean’s are prokaryotes, although they differ from eukaryotes and eubacteria in a number of ways, while being classified as prokaryotes. Their membrane lipids' chemistry is one of

these variations. While the majority of membrane lipids in eukaryotes and eubacteria are made up of fatty acids esterified to glycerol backbones, in archaea's the lipid chain in the first position is joined to the glycerol through an ether bond. Given the severe environments these species live in, it's probable that the ether bond, which is more stable than an ester, offers a significant benefit [9], [10].

Variability

Different methods that cell have at their disposal allow them to regulate the lipid content of their membranes. Even though phospholipids predominate in most membrane compositions, there is still significant variation amongst membranes from various sources, as seen in the following examples:

1. Distinct biological species. In eukaryotes, phosphatidylcholine (PC) is the predominant phospholipid, but in most bacteria, it is either phosphatidylethanolamine (PE) or phosphatidylglycerol (PG). In contrast, cardiolipin (CL) is the predominant phospholipid in the mycobacterium M. tuberculosis.
2. Distinct organ systems within a same organism. For instance, membranes in the nervous system that are rich in cholinergic receptors contain no phosphatidylinositol (PI), very little sphingomyelin (SM), and no intestinal brush border membrane (CL).
3. In eukaryotes, plasma versus inner membranes. For instance, although CL is essentially missing from the ER and plasma membranes, it makes up around 20% of the lipids in the mitochondrial inner membrane. Furthermore, animal cholesterol is mostly found in the plasma membrane and only in trace levels in the ER membrane. Finally, the plasma membrane and the lysosomal membrane are where the majority of the SM in the cell is concentrated.
4. The plasma membrane's cytoplasmic vs exoplasmic leaflets. The exoplasmic leaflet of eukaryotic membranes is mostly composed of choline phospholipids (PC and SM), whereas the cytoplasmic leaflet is primarily composed of amino phospholipids (PS and PE), as well as PI in much lower amounts. The lipid asymmetry causes a charge differential between the two leaflets because PS and PI are negatively charged whereas phosphatidylserine (PS) and PC are both electrically neutral. In other words, as compared to the exoplasmic leaflet, the cytoplasmic leaflet is negative. The cytoplasmic leaflet of the bacterial inner membrane is abundant in PE and PI, whereas the exoplasmic leaflet is high with PG.
5. Various areas of the same membrane. Certain lipid molecules with identical characteristics prefer to congregate in certain membrane areas known as rafts or microdomains. One outcome of lipid-protein interactions is the creation of microdomains, which often has functional ramifications. PIP2 microdomains, as an instance, are crucial for certain signal transduction procedures.

CONCLUSION

In summary, the plasma membrane is a fascinating and crucial part of cells, acting as a dynamic border to segregate the inside of the cell from the outside world. This complex structure, which is made up of lipids, proteins, and carbohydrates, is essential for preserving the particular chemical environment required for cellular activity. To maintain the integrity and functioning of diverse living forms, evolution has fine-tuned the composition and characteristics of membranes over billions of years. These structures have evolved to meet the demands presented by various surroundings and cellular functions, from the appearance of the earliest membranes through the development of inner membranes in eukaryotic cells. In conclusion, the plasma membrane is a flexible and intricate structure that is necessary for

life. Its structure, composition, and capabilities have changed to accommodate the many demands placed on cells and organisms. Understanding the intricate workings of biological membranes is essential for bettering our understanding of the behaviour of membrane proteins and their function in diverse cellular processes.

REFERENCES:

[1] C. G. Palivan, R. Goers, A. Najer, X. Zhang, A. Car, and W. Meier, Bioinspired polymer vesicles and membranes for biological and medical applications, *Chemical Society Reviews*. 2016.

[2] K. R. Pothula, C. J. F. Solano, and U. Kleinekathöfer, Simulations of outer membrane channels and their permeability, *Biochim. Biophys. Acta - Biomembr.*, 2016.

[3] R. Ernst, C. S. Ejsing, and B. Antonny, Homeoviscous Adaptation and the Regulation of Membrane Lipids, *Journal of Molecular Biology*. 2016.

[4] G. Rossi and L. Monticelli, Simulating the interaction of lipid membranes with polymer and ligand-coated nanoparticles, *Advances in Physics: X*. 2016.

[5] M. G. Burton *et al.*, Direct Measurement of Pore Dynamics and Leakage Induced by a Model Antimicrobial Peptide in Single Vesicles and Cells, *Langmuir*, 2016.

[6] S. Eriksson, N. Eremina, A. Barth, J. Danielsson, and P. Harryson, Membrane-induced folding of the plant stress dehydrin Lti30, *Plant Physiol.*, 2016.

[7] N. Erwin, B. Sperlich, G. Garivet, H. Waldmann, K. Weise, and R. Winter, Lipoprotein insertion into membranes of various complexity: Lipid sorting, interfacial adsorption and protein clustering, *Phys. Chem. Chem. Phys.*, 2016.

[8] M. Palamini, A. Canciani, and F. Forneris, Identifying and visualizing macromolecular flexibility in structural biology, *Frontiers in Molecular Biosciences*. 2016.

[9] Y. Zhao, M. Zhang, and Z. Wang, Underwater Superoleophobic Membrane with Enhanced Oil–Water Separation, Antimicrobial, and Antifouling Activities, *Adv. Mater. Interfaces*, 2016.

[10] M. L. Huang and K. Godula, Nanoscale materials for probing the biological functions of the glycocalyx, *Glycobiology*. 2016.

CHAPTER 6

MEMBRANE LIPID COMPOSITION: IMPACT ON BIOLOGICAL FUNCTIONALITY

Himanshu Yadav, Assistant Professor
College of Paramedical Sciences, TeerthankerMahaveer University, Moradabad, Uttar Pradesh, India Email Id- forensic.holmes@gmail.com

ABSTRACT:

Thorough summary of several topics pertaining to how lipid characteristics affect cellular membranes. It discusses issues including membrane curvature, the structural principles of membrane proteins, lipid amphipath city, membrane lipid distribution and asymmetry, lipid bilayer characteristics, and lipid bilayer distribution and asymmetry. Different lipids inhabit the exoplasmic and cytoplasmic leaflets of the membrane in an asymmetrical distribution of lipids. Several enzymes that transfer lipids from one side to the other preferentially keep this asymmetry in place. Changes in lipid asymmetry may have important physiological effects, such as affecting how macrophages identify apoptotic cells and how the immune system reacts to an infection. Various elements, including lipid saturation levels and the presence of sterols, have an impact on the characteristics of the lipid bilayer, including its fluidity and thickness. The biological activities of the membrane, such as its function as a barrier and its participation in vesicular transport processes, are significantly impacted by these features. Whether peripheral or integral, membrane proteins are essential for a variety of cellular processes. Integral membrane proteins are contained inside the lipid bilayer, and the necessity to reduce the energy cost of burying polar peptide bonds in the hydrophobic core of the membrane dictates how they are structured. These proteins' transmembrane regions are often hydrophobic and favour certain amino acid residues, including Leu, Ile, Val, Phe, Trp, Tyr, Lys, and Arg. These residues engage with the lipid molecules and aid in holding the proteins to the bilayer.

KEYWORDS:

Biological membrane, Cell, Lipid, Molecules, Protein.

INTRODUCTION

Biological membrane lipid molecules have distinct characteristics from one another, yet they all possess the property of amphipath city. In other words, there are polar and nonpolar regions in every membrane lipid. For instance, the ester-glycerol-phosphate-alcohol groups are found in the polar area of glycerophospholipids, while the acyl chains are found in the nonpo- lar region. The terms polar head and nonpolar tails are commonly used to describe these two areas. The hydroxyl group makes up the polar area of cholesterol, whereas the fused ring structure and associated hydrophobic tail make up the nonpolar part. The hydrophobic effect and amphipathic nature of these lipids cause them to form bigger structures in the typically aqueous environment of biological systems, where the polar areas face the aqueous medium and the nonpolar regions face one another. The lipid bilayer is one such sturdy structure. As previously mentioned, this structure is set up such that all lipids' nonpolar tails form a 30-hydrophobic core and their head groups form two 10- to 15-polar layers that are exposed to the surrounding aqueous environment. The most significant characteristic of the lipid bilayer, i.e., impermeability to most polar solutes, depends on this structural organisation. Once again, due to the membrane's impermeability, the cell is able to

closely control the quantity of its metabolites by designating certain transport proteins as the only pathways into and out of the cytoplasm.

The distribution of lipids in the membrane is asymmetric. For instance, in eukaryotic membranes, the exoplasmic leaflet mostly consists of the glycolipids and choline-containing lipids PC and SM, whereas the cytoplasmic leaflet primarily consists of the amino lipids PS and PE.

Since phospholipids may switch sides in a couple of hours, an active mechanism is required to maintain the asymmetry, namely membrane-bound enzymes that move lipids from one side of the bilayer to the other while drawing energy from ATP [1], [2]. There are two enzymes in particular that operate in opposition to one another:

1. The amino peptides PS and PE are moved from the exoplasmic side of the bilayer to the cytoplasmic side by the enzyme flippase.
2. In certain tissues, floppase moves PC and cholesterol from the cytoplasmic side to the exoplasmic side.

Certain activities, such as malignant cell transformation and intended or unintentional cell death, reduce membrane lipid asymmetry. This decrease in asymmetry is caused by either a drop in flippase activity or a rise in scramblase activity, an enzyme that evenly distributes phos- pholipids to both sides of the bilayer. At least in the case of PS, the loss of asymmetry may impact the cell and tissue. Particularly, it has been shown that PS mediates a number of physiological processes involving cellular identification when it is present in the exoplasmic leaflet:

1. Apoptotic cells are recognised by macrophages. As was previously mentioned, macrophages are phagocytes with the capacity to engulf and internalize a wide range of objects, including individual proteins and whole cells. They fulfil two roles in doing so. They first eliminate invasive microorganisms that might damage the body. Second, they aid in the removal of tissue's dead cells. The latter function is crucial for cleansing, as well as for preventing the onset of a damaging inflammatory reaction in the tissue after cell demise. On the other hand, inflammation occurs quickly when cells die via necrosis, which is not planned by the body but rather brought on by some kind of damage. PS on the surface of apoptotic cells has been linked to macrophages' ability to identify these cells, which helps to reduce inflammation by allowing macrophages to recognised these cells.
2. T-lymphocytes' recognition of activated endothelium cells. The immune system's ability to identify tissues that have been invaded by infections and respond promptly to eradicate the invaders is one of its functions. The issue is that lymphocytes often move about in flowing blood and lymph and do not stay still for long periods of time. Therefore, it becomes imperative to stop lymphocytes from spreading outside of this tissue when infections are found there. This is accomplished by the endothelial cells that line the blood arteries close to the invaded tissue, or neighbouring endothelial cells. They go through a procedure that makes their PS visible to the extracellular environment, and neighbouring T-lymphocytes are able to recognised and bind it.
3. The complement system's ability to identify microorganisms. A number of typically inactive proteins are part of the complement system. However, they develop a complex that assaults the invasive cells when they become active during pathogenic invasion. The assault entails the destruction of those cells' cytoplasmic and membrane components. The complement system is often triggered in response to bacterial invaders that have previously been recognised by antibodies. Alternative activation

mechanisms do, however, exist, and they seem to include PS on the bacterial invader's exoplasmic leaflet [3], [4]. If this is the case, cancer cells may also be identified in this manner as they are known to have less lipid asymmetry in their membranes than healthy cells.

DISCUSSION

Although the fluid mosaic model often depicts the lipid bilayer as being generally fluid, its fluidity may vary within a given range. The degree of order of the individual lipids, and more especially, their hydrophobic tails, determines this variance. The bilayer structure is more viscous due to the dense packing of linear tails. 'Liquid ordered' (lo) is the term used to describe such a structure. On the other hand, bent tails create a structure known as liquid disordered (ld), which is less compact and more fluid. Under physiological circumstances, the lipid bilayer contains both types, each of which contributes to the bilayer's biological properties: the lo phase improves the bilayer's ability to act as a physical barrier for polar solutes, while the ld phase gives it some degree of dynamics. In fact, bilayer portions that enter the ld phase enable individual lipids to transversely diffuseor 'flip'from one leaflet to anotherin addition to diffusing along the bilayer's surface. Furthermore, during the late phase, the membrane is able to organise structurally, which is necessary for several biological functions including the creation, budding, and fusing of transport vesicles. Two primary elements impact how tightly individual lipids pack. Lipid saturation level. Fully saturated lipids contain hydrophobic tails that are linear and likely to form bilayers of the lo type.

The bent tails of lipids with one or more double bonds, on the other hand, pack in a ld-type structure. Sterols are present in the bilayer. The lipid bilayer's characteristics are affected by sterols in two different ways. One way the sterol functions is as a stopper that stops solutes from freely moving through the spaces between phospholipids. On the other hand, the sterol molecule's stiff and bending conformation causes a spatial disruption in the tightly packed phospholipids of the lipid bilayer, which inhibits the bilayer from solidifying. This is significant because the bilayer must continue to be dynamic for the membrane to function biologically. Sterols enable biological membranes to maintain their fundamental role as a physical barrier while yet being dynamic.To attain the ideal balance between various properties, the cell may vary the two elements. For instance, the ER membrane preserves flexibility despite having tiny amounts of sterols owing to the abundance of unsaturated phospholipids. Numerous regions of biological membranes are thought to be in the lo phase, according to studies on lipid composition, sensitivity to detergents, and biophysical measurements of lipid motions, while the ld phase is typically only found in areas that are actively engaged in dynamic activity, like the formation of transport vesicles. The lipid bilayer's thickness is a crucial characteristic that results from the degree of order in the layer. Pure lipid bilayer measurements reveal that the bilayer has an average thickness of 50–60, a 30- nonpolar core, and two 10–15 –sized polar lipid–water interfaces. These values vary among various

even across several compartments of the same cell in the same organism. For instance, in liver cells, the basolateral membrane, which confronts other cells, is 7 times thicker than the basolateral membrane and is 5 times thicker than the apical membrane, which faces the lumen. Because the length of the acyl chains shortens when a portion of the lipid bilayer transitions from the lo to the ld phase, it gets thinner. The interactions between the phospholipids and the nearby integral membrane proteins are impacted by this variation. Since these effects also have an impact on the stability of the latter, proteins have a tendency to concentrate in areas of the membrane where only one of the phases is present. As a consequence, membrane protein clusters have a tendency to be physically apart from one

another. Protein concentrations in certain membranous areas are often employed to improve signal transduction pathways, which need close proximity of its constituent parts. Last but not least, modifications to protein-lipid interactions brought on by phase shifts may have an immediate impact on the former's function [5], [6].

The lipid bilayer may momentarily curl in certain places, despite the fact that it is often shown as being flat. Several activities, including the vesicular transport of proteins and lipids between the ER, Golgi apparatus, and plasma membrane, are made easier by this phenomenon. Vesicular transport starts with the source membrane gradually bending till the vesicle is created, then it progresses with the vesicle budding from the membrane and its diffusion towards the destination membrane, and it concludes with the fusing of the two. This process is involved in exocytosis, endocytosis, and interorgan Elle exchange.Membrane lipid and protein composition has an impact on membrane curvature. The ratio between the effective cross-section area of the head groups of the lipids and that of their tails determines the action of lipids. When the ratio is less than one, the lipids create a bilayer structure that is generally planar and arranged parallel to one another. On the other hand, the lipids create a curved membrane when there is an imbalance between the areas of the head group region and the tail group region. There are two examples of this:

1. When the head group section is larger than the tail section, as is the case with choline-containing lipids (PC and SM) as well as with PG, positive membrane curvature results. These lipids naturally have a propensity to form a convex leaflet.
2. When the head group section is smaller than the tail section, as it is with PE, negative membrane curvature results. These lipids naturally prefer to generate a concave curvature in the leaflet they create. The stability of the bilayer membrane is decreased by lipids that cause negative curvature, which might eventually result in bilayer disintegration.

Proteins and a variety of lipids with varying curvature preferences make up biological membranes, and it may be challenging to anticipate the precise shape that would result from a given lipid composition. For instance, the bacterial plasma membrane still resembles a membrane in principle even if 70% of its lipids are made up of PE. This is so because PGs, which make up the remaining 30%, generate a compensatory positive curvature. In fact, studies reveal that even in the absence of compensating lipids, the membrane will stay planar and entire as long as the concentration of negative curvature-inducing lipids is less than 20%. Nevertheless, the membrane experiences mechanical frustration due to the mixture of lipids' various curvature-inducing characteristics. In circumstances when the membrane's biological activity necessitates frequent curvature changes, such as in intracellular transport, it has been proposed that cells employ this so-called curvature frustration to make the membrane metastable. Thus, albeit minimal, the influence of lipid shape on bilayer curvature is nevertheless evident. The ER and Golgi membranes, for instance, are somewhat curved due to lipid shape. Integral membrane proteins have a far bigger influence on membrane morphology; they are in charge of dramatic modifications including the creation of transport vesicles.

Membrane Protein Structure Principles

Proteins that are attached to membranes may be divided into two main categories: peripheral proteins and integral proteins. An integral protein's membrane-spanning region may have one of two shapes. The D and L amino acids in certain antibiotic peptides, like gramicidin, alternate, allowing them to form the -helix, a third kind of structure. The -helix may act as a channel to transport monovalent ions across the membrane since it is broader than the -helix.

Helix-shaped membrane proteins, which make up the great majority of integral membrane proteins, are the main topic of this section. The characteristics of -barrel membrane proteins are discussed. Depending on how many segments cross the membrane, different subgroups of helical membrane proteins may be identified. While polytopic membrane proteins have several transmembrane segments, biopic membrane proteins only have one. Comparing various organisms reveals a weak preference for membrane proteins with seven transmembrane segments each in higher organisms like Caenorhabditis elegans and Homo sapiens, but integral membrane proteins with 6 or 12 transmembrane segments are more common in unicellular organ- isms. A well-known example of the latter kind of protein is the GPCR, which is a prominent target for pharmaceutical medicines and plays a crucial function in animal physiology. This chapter's last portion focuses on GPCRs [7], [8].

The majority of membrane proteins are integral membrane proteins, which play a variety of activities. Monotopic and bitopic proteins often serve as receptors for growth-factor-like messengers as well as recognition and/or adhesion molecules. While their cytoplasmic area transmits the signal into the cell by attaching soluble components or cytoskeletal proteins, their extracellular region is in charge of binding the chemical messenger. Typically, polytopic proteins serve as receptors or transporters. For instance, the aforementioned GPCRs react to a wide range of messengers, such as hormones, neurotransmitters, odorants, pheromones, and even electromagnetic radiation. On each side of the membrane, integral proteins or membrane lipids serve as anchors for peripheral membrane proteins. Carbohydrate moieties may act as a direct or indirect mediator of lipid attachment.Integral membrane proteins are surrounded by a lipid bilayer, therefore the principles governing their structuremore especially, the structure of their transmembrane domainsare quite different from those that apply to proteins that are water-soluble. As a result, we will concentrate on the structure of integral membrane proteins when discussing the membrane anchoring of peripheral proteins, which are often surrounded by a water-based environment.

Integral membrane protein structures

According to this definition, integral membrane proteins are globular, much like their cytoplasmic counterparts. However, given that they exist in a setting that is so unlike from the watery cytoplasm, it is possible that the energy factors that determine their structural stability are distinct from those that influence the structure of proteins that are water-soluble. Analysing multiple structures is necessary to understand these determinants, much as it has been done for water-soluble proteins during the last several decades. It is challenging to overexpress, extract, and purify membrane proteins, as well as to crystallise them, as was discussed; this makes it difficult to determine the structure of the protein. Replacement of the surrounding lipids with detergent molecules often solves the crystallisation issue, albeit the protein structure may change due to the changing environment, rendering the results useless. Researchers have made significant strides in the experimental identification of membrane protein structure in recent years. This improvement includes the creation and refinement of techniques like electron cryomicroscopy (cryo-EM), circular dichroism (CD), and small-angle X-ray scattering (SAXS), which have given important insights into membrane proteins that are difficult to crystallise as well as the supramolecular assemblies they form.

This is due to the lipid bilayer's anisotropic and chemically complicated composition, which places restrictions on the shape of resident proteins. As a consequence, membrane proteins have a generally simple design, and fewer structures are required to comprehend the fundamental ideas governing that architecture.The energy cost of burying the protein's polar peptide bonds within the hydrophobic hydrocarbon core of the lipid bilayer is the primary determinant of integral membrane protein structure. The transmembrane segment sequences

are very hydrophobic and have a strong propensity to generate organised secondary structures in order to offset this cost. There are more factors that have a little but significant impact on the shape of membrane proteins. The principles governing membrane protein structure are reviewed in the subsections that follow using the structural hierarchy used for water-soluble proteins. We suggest reading the reviews by von Heijne, White, Engelman, and Bowie for further information.An integral membrane protein's polypeptide chain must pass through the lipid bilayer at least once. Since the bilayer's hydrocarbon core is so hydrophobic, the proteins' transmembrane domains must be similarly hydrophobic. In fact, the lowest polarity of integral membrane proteins, especially in their transmembrane portions, makes them stand out from water-soluble proteins. Leu, Ile, Val, and Phe are notably strongly abundant in integral membrane proteins compared to water-soluble proteins, despite the fact that all kinds of nonpolar residues are frequent in transmembrane regions.

Polar residues may also be found in transmembrane domains, although they are less prevalent there, particularly in single-pass proteins, where they make up only around 20% of the sequence overall. Polar residues in multi-pass membrane proteins are often buried in the core rather than facing the more hydrophobic membrane particularly if they are charged. The presence of polar residues in a highly hydrophobic environment serves a similar purpose as in the cores of water-soluble proteins here, explaining the inevitable structural destabilization. Where the buried polar residue is surrounded by water molecules, other polar residues, or both such as in the voltage-sensing K+ channel, the destabilization is somewhat reduced. Integral membrane proteins are introduced into the ER membrane through the translocon machinery during co-translation. The translocon scans the nested polypeptide chain to find transmembrane regions, but how does it do this? According to structural studies, the translocon structure has a side gate that opens to the lipid bilayer at a certain frequency, revealing the sequences inside of it and complementing the main channel hole that accommodates the nested polypeptide chain [9], [10].

The length of transmembrane segments is another distinguishing feature. Although helical transmembrane segments may have between 15 and 39 residues, statistical analyses show that the 'average' transmembrane helix has between 21 and 26 residues, and there is a considerable predilection for helices with more than 20 residues. The limitations imposed by the membrane environment in addition to the structural characteristics of -helices are the cause of this once again. In other words, a 20-residue long -helical transmembrane segment would equate to a length of 30, matching the typical thickness of the hydrocarbon core of the lipid bilayer. This is because of the typical 1.5 increase per residue along the helix axis. It goes without saying that the helix has to be hydrophobic enough to traverse the membrane. As we shall see in a later section, longer helices often tilt to maximise their nonpolar interactions with the membrane's core. Comparatively, helices tend to be shorter on average, with a wider length distribution, in water-soluble proteins, whose environment does not impose the limits seen in the bilayer.The fact that membrane proteins can be successfully detected automatically in whole genomes and their number of transmembrane segments can be predicted based solely on sequence highlights the importance of the low polarity and characteristic length of transmembrane segments. Pro and Gly, as well as branched residues, are often found in transmembrane segments organised as helices.

This comes as a huge shock considering how seldom such residues are found in globular proteins' -helices. Proline is highly prevalent and often found next to Ser or Thr in helical transmembrane regions. The importance of these residues in membrane proteins may be understood by structural study; transmembrane segments often contain aromatic belts around the edges of the hydrocarbon area of the lipid bilayer. The aromatic residues Trp and Tyr,

which are often uncommon in proteins, form part of such a belt. This is noteworthy, particularly in the case of Trp, which is three times more common in membrane proteins than in water-soluble proteins. The aromatic residues are situated at the lipid bilayer's interface between the nonpolar tails and the polar head group area, close to the termini of transmembrane helices. There, they may engage in intricate interactions with both lipid components. The established explanation states that these interactions serve to 'anchor' the transmembrane segments to the bilayer, preventing them from'sliding' into or out of the cell. Trp and Tyr's affinity for the interface region might be attributed to their ability to interact with the amphipathic membrane interface due to their amphipathic nature. In other words, their high surface area enables them to interact with the nonpolar lipid tails, whilst their polar NH and OH groups hydrogen-bond with the polar lipid head groups of the interface. Additionally, it is predicted that insertion into the highly disordered acyl chains of the membrane core will be difficult due to the inflexible, bulky side chains of Trp and Tyr.

In their cytoplasmic portions, transmembrane segments often include the fundamental residues Lys and Arg. Von Heijne identified this tendency and named it the positive inside rule. This preference is also shown for histidine;however, it is less common in these areas than Lys or Arg. Given that the His side chain approximately equally likely to be positively charged or electrically neutral at physiological pH, this makes sense. The positive inner tendency may be explained in a number of ways. It could be related, for instance, to the phospholipid asymmetry that lipid bilayers naturally possess. Electrically neutral phospholipids (PC and SM) are located at the exoplasmic leaflet of the bilayer in eukaryotic membranes, whereas negatively charged phospholipids (PS, and PI) are located at the cytoplasmic leaflet. The exoplasmic and cytoplasmic leaflets of inner bacterial membranes both contain negatively charged lipids (PG and PI, respectively). Transmembrane segments containing basic residues in their cytoplasmic sections might therefore create salt bridges with these negatively charged lipids in both prokaryotic and eukaryotic membranes, stabilising the protein-membrane system. The absence of positively charged lipids in the lipid bilayer prevents the converse from happening, which is the presence of acidic residues on the exoplasmic side of the bilayer [11], [12].

The membrane potential, or the electric potential across biological membranes, may also be another explanation for the positive inside rule. In this potential, the cytoplasm is more negatively charged than the periplasm. The transposon machine's bias might also be reflected in the positive inside rule. Finally, the inclusion of polar groups at the end of lengthy non-polar chains in each of their side chains is likely what makes Lys and Arg compatible with the membrane interface. As a result, even when the residue is positioned farther into the bilayer's hydrocarbon core, the polar group may still interact favourably with the head groups of phospholipids. 'Snorkelling' is the term used to describe this phenomenon. Interestingly, despite having a predilection for this area identical to Lys, recent research reveals that only Arg substantially stabilises the cytoplasmic side of the membrane. This could be due to the Arg side chain's electronic delocalization, which disperses the stabilising positive charge across a wider region. Additionally, the side chain of Arg may form more hydrogen bonds with phospholipid head groups than that of Lys.

The distribution of positive residues is different in the -barrel proteins of the Gram-negative bacterial membrane from other membrane proteins. at other words, basic residues are more prevalent at the protein's outermost loops, or positive outside. Because there are so many lipopolysaccharides (LPS), the bacterial outer membrane is strongly negatively charged on its exoplasmic side compared to other cellular membranes. Through ionic interactions, the basic residues of membrane proteins in this area stabilise the negatively charged LPS. A bacterial

outer-membrane protein has more negatively charged residues on the opposing side. These interact with periplasmic cationic chaperones like Skp to help insert and fold the -barrel proteins into the outer membrane. In transmembrane regions, Gly, Ala, and Ser are often found as little residues. These residues often occur in -helices and help neighbouring helices optimise their hydrogen bonds and van der Waals interactions.

CONCLUSION

In conclusion, the structure and operation of biological membranes are fundamentally influenced by the characteristics of lipid molecules. Amphipathic membrane lipids with polar head groups and nonpolar tails are a distinctive feature of these molecules. They can create the lipid bilayer, a crucial component of cell membrane structure, thanks to their amphipathicity. Untangling the intricate activities of biological membranes and their participation in many cellular processes requires an understanding of the principles underlying the lipid and protein characteristics of membranes. Despite ongoing experimental difficulties, improvements in methods like cryo-EM and X-ray crystallography have shed light on the structures and roles of membrane proteins. In conclusion, the interaction of lipids and proteins in biological membranes is an exciting field of research with extensive ramifications for cell biology and physiology.

REFERENCES:

[1] M. Li, C. Gan, W. Shao, C. Yu, X. Wang, and Y. Chen, Effects of membrane lipid composition and antibacterial drugs on the rigidity of Escherichia coli: Different contributions of various bacterial substructures, *Scanning*, 2016.

[2] L. Dominguez, L. Foster, J. E. Straub, and D. Thirumalai, Impact of membrane lipid composition on the structure and stability of the transmembrane domain of amyloid precursor protein, *Proc. Natl. Acad. Sci. U. S. A.*, 2016.

[3] S. Rodriguez-Cuenca, L. Whyte, R. Hagen, A. Vidal-Puig, and M. Fuller, Stearoyl-CoA desaturase 1 Is a key determinant of membrane lipid composition in 3T3-L1 adipocytes, *PLoS One*, 2016.

[4] Z. Chorner, P. A. Barbeau, L. C. Castellani, D. C. Chorner, A. Chabowski, and G. P. Holloway, Dietary α-linolenic acid supplementation alters skeletal muscle plasma membrane lipid composition, sarcolemmal fat/cd36 abundance, and palmitate transport rates, *Am. J. Physiol. - Regul. Integr. Comp. Physiol.*, 2016.

[5] I. Haralampiev *et al.*, The interaction of sorafenib and regorafenib with membranes is modulated by their lipid composition, *Biochim. Biophys. Acta - Biomembr.*, 2016.

[6] R. Ernst, C. S. Ejsing, and B. Antonny, Homeoviscous Adaptation and the Regulation of Membrane Lipids, *Journal of Molecular Biology*. 2016.

[7] M. A. Tahir, R. C. Van Lehn, S. H. Choi, and A. Alexander-Katz, Solvent-exposed lipid tail protrusions depend on lipid membrane composition and curvature, *Biochim. Biophys. Acta - Biomembr.*, 2016.

[8] J. Schumann, Does plasma membrane lipid composition impact the miRNA-mediated regulation of vascular inflammation?, *Med. Hypotheses*, 2016.

[9] J. P. X. Cheng and B. J. Nichols, Caveolae: One Function or Many?, *Trends in Cell Biology*. 2016.

[10] A. P. Muir, F. L. D. Nunes, S. F. Dubois, and F. Pernet, Lipid remodelling in the reef-building honeycomb worm, Sabellaria alveolata, reflects acclimation and local adaptation to temperature, *Sci. Rep.*, 2016.

[11] G. Li, J. Kim, Z. Huang, J. R. St Clair, D. A. Brown, and E. London, Efficient replacement of plasma membrane outer leaflet phospholipids and sphingolipids in cells with exogenous lipids, *Proc. Natl. Acad. Sci. U. S. A.*, 2016.

[12] B. Daum *et al.*, Supramolecular organization of the human N-BAR domain in shaping the sarcolemma membrane, *J. Struct. Biol.*, 2016.

CHAPTER 7

CHALLENGES AND STRATEGIES: STRUCTURE AND FOLDING OF INTEGRAL MEMBRANE PROTEINS

Akash Chauhan, Assistant Professor
College of Paramedical Sciences, TeerthankerMahaveer University, Moradabad, Uttar Pradesh, India Email Id-akashchauhan3596@gmail.com

ABSTRACT:

Understanding integral membrane proteins' structure and function is vital since they are involved in many different biological processes. These proteins' transmembrane segments are mostly made up of non-polar residues, and the quantity of these residues often varies with the hydrophobic core of the membrane's thickness. As a result, computer methods for predicting transmembrane regions have been created, beginning with the groundbreaking work of Kyte and Doolittle, who used a sliding window method based on hydrophobicity. There have been arguments about the selection of the hydrophobicity scale and the kind of material used to assess hydrophobicity. As alternatives, knowledge-based scales that are based on statistical information about the architectures of membrane protein have evolved. Taking into consideration functional restrictions for charged residues, these scales take into account the statistical probability of residues occurring in transmembrane domains. Another difficulty is the representation of residues in transmembrane segments. The influence of charge magnitude on hydrophobicity has been studied using host-guest peptides in current studies, as opposed to earlier scales that employed individual amino acids. Algorithms like TopPred and TMHMM, which depend on statistical information from well-known structures, have been used to address topology prediction, which determines the orientation of segments in the membrane. These forecasting techniques have improved in precision. Alpha-helical segments are often abundant in integral membrane proteins, which bring particular difficulties and possibilities. Transmembrane helices' proline residues may operate as motion hinges, enabling conformational changes that are necessary for protein activity. Weak turns and tight twists are examples of other structural oddities that are still poorly understood.

KEYWORDS:

Helical, Hydrophobicity, Membrane Proteins, Protein.

INTRODUCTION

An integral membrane protein's transmembrane segments are mostly made up of non-polar residues, the number of which tends to correspond approximately to the membrane's hydrophobic core's thickness. This quantity is 20 residues in -helical proteins, which make up the majority of integral membrane proteins. Using this knowledge, researchers have developed computer algorithms that can identify transmembrane regions inside genomes based on certain trends. Kyte and Doolittle carried out the first endeavor. Their overall plan was to find transmembrane segments along the protein sequence by using a sliding window that covered around 20 amino acid positions. The overall or average hydrophobicity of the 20-amino acid sequence was determined at each point of the window along the sequence. The segment covered by the window was taken into consideration to be a possible trans-membrane segment in those locations where the estimated value surpassed a certain threshold. A hydropathy map, which shows the likelihood that each subsequent 20-residue section in the sequence is a transmembrane segment, was used to display the findings of this approach.

Calculating each candidate sequence's hybrid city was a crucial step in this methodology. This was done by employing a hydrophobicity scale, in which a hydrophobicity value was given to each of the 20 different kinds of natural amino acids. Although the design of this scale may seem unimportant at first, it is still up for debate. The choice of the physical quantity that may be utilised to indicate hydrophobicity must come first. The polarity of the molecule is the first quantity that springs to mind. Although polarity influences hydrophobicity, it does not always completely account for it. This is because hydrophobicity represents all the characteristics influencing the molecule's propensity to favour a nonpolar media over a polar one, while polarity only emerges from the geometric distribution of electronegative atoms. Therefore, even though a big residue like tyrosine is thought to be polar owing to its side chain OH group, it may actually turn out to be hydrophobic if its massive phenyl group can generate sufficient nonpolar interactions. So, based on the distribution of amino acids between polar (water) and nonpolar (vacuum) environments, the Kyte-Doolittle (KD) scale was experimentally created [1], [2].

This brings us to the second issue with creating a hydrophobicity scale, which is the different medium types utilised to assess hydrophobicity. The polar and nonpolar media used to assess the hydrophobicity values must be as near as feasible to the cytoplasm and lipid bilayer, respectively, for such a measurement to be effective. While water has long been considered a good representation of the cytoplasm, there is debate about the best medium to use to portray the lipid bilayer. In spite of the fact that vacuum, the medium used by Kyte and Doolittle, is nonpolar with a value even lower than the hydrocarbon area of the membrane), it is not amphipathic, and as a result, it is unable to accurately mimic a biological membrane. Other scales have been created utilising octanol as the nonpolar medium or even actual lipid bilayers to account for this crucial characteristic of the membrane. In order to improve the accuracy of the hydrophobicity measurements, White, von Heijne, and colleagues used a reconstructed system that included all the biological components necessary for the insertion of transmembrane segments into the membrane in actual cells, including the ribosome-translocon complex. This method avoided measuring the spontaneous, but artificial, partitioning of residues between simple polar and nonpolar media.

They created a scale of hydrophobicity that may be more accurate than scales created by transferring amino acids and/or peptides across simple mediums. The amino acid transfer energies discovered by White, von Heijne, and colleagues were substantially less in size than those discovered in the aforementioned research, raising questions about their veracity. These low energies were thought to be the consequence of several approximations made throughout the research and/or interactions between the implanted peptides and other membrane elements. It's noteworthy that a recent investigation of amino acid transfer energetics, which, like the one by White, von Heijne, and collaborators, employed a true biological membrane, produced transfer energies that were comparable to those found in experiments that used simple media. These findings imply that simple organic solvents accurately capture the hydrophobicity at the centre of biological membranes and confirm the validity of the latter research.

The two issues raised abovechoosing the physical amount to be utilised in building hydropathy scales and the kind of medium to employ for amino acid transfer experimentsare avoided by the use of knowledge-based scales. These relatively recent scales substitute the non-physical chance of a residue appearing within a transmembrane segment for the physically relevant hydrophobicity. The likelihood is determined using statistical data gathered from membrane proteins with known three-dimensional structures. Since there are more membrane protein structures today than there were decades before, it is now possible to

extract the data that were previously unavailable owing to the scarcity of membrane protein structures. It should be highlighted that functional limitations skew knowledge-based scales. This bias is particularly noticeable in the case of charged residues, whose real transfer energies into nonpolar media are extremely unfavorable. However, their statistical likelihood to appear in transmembrane domains is higher than suggested by these energies because they are required for functional purposes, which makes them more likely to appear in these domains.Choosing how to represent the residues in the transmembrane segment is the third contentious topic related to the creation of hydrophobicity scales. Although in actuality amino acids are joined by peptide bonds, the KD and GES scales employ individual amino acids.

The linkages turn the amino acids -amino and -carboxyl groups' full electric charges into simple partial dipoles. It has been argued that using individual amino acids is methodologically problematic since the partitioning between polar and nonpolar media much relies on the magnitude of the charge on the molecule. Wimley and White employed entire peptides to create their own hydrophobicity scale, using the 'host-guest' method, to address this issue. Except for one place, which had a different residue in each instance, the peptides they utilised in the numerous assays were identical. Wimley and White also employed a lipid bilayer in place of a straightforward nonpolar mimic, like Goldman, Engelman, and Steitz. The host-guest peptides, however, could not traverse the complete length of the lipid bilayer owing to their short lengths (5 or 6 residues), and instead partitioned solely into the polar head group area. As a consequence, transmembrane segments could not be covered by the scale created by this approach. Kessel and Ben-Tal chose 20-residue host-guest peptides in their computer investigation because they could cover the whole length of the bilayer. The peptides also had a -helical shape, much like actual transmembrane segments, where the polar backbone groups were joined by hydrogen bonding [3], [4].

When just a few 3D structures of membrane proteins were known, the first transmembrane segment identification algorithms were created. The length and hydrophobicity values of several of the transmembrane segments varied from the fundamental trends expected by the prediction algorithms when structures, like those of the K+ channel, began to form. First, despite being too short to cover the whole thickness of the membrane, certain transmembrane segments were buried within the membrane's core. The unpaired polar groups at the termini of these segments were electrostatically covered by polar groups on neighbouring segments or by water molecules filling the intramembrane pore, preventing them from being electrostatically exposed to the hydrophobic core. It was discovered that some transmembrane segments had hydrophobic lengths greater than the membrane's hydrophobic thickness. To enable as many nonpolar residues to interact with the hydrocarbon section of the membrane as feasible, these segments often acquired a tilted orientation with respect to the membrane normal. Finally, it seemed that some of the segments had kinks that caused the helical shape to be deformed. These findings demonstrate that forecasts of trans-membrane segments need data beyond sequence tendencies in order to be effective.

The problem of topology, or anticipating whether portions of each segment face the cytoplasmic or exoplasmic sides of the membrane, has come up as a result of the quest for general guidelines for characterising transmembrane segments. This kind of prediction is an essential first step in predicting the overall structure of an integral membrane protein because it restricts the possible arrangements of the transmembrane segments in space relative to one another. The first algorithm created for this aim, TopPred, coupled the aforementioned positive inside criteria with the overall inclination of transmembrane segments towards hydrophobicity. Later-developed algorithms, such MEMSAT and TMHMM, heavily

depended on statistical information on the locations of transmembrane segments in known-structural proteins. Anyone may utilise the many prediction techniques and algorithms that are now widely available to the public to create a strong foundational model. For instance, it has been shown that the TMHMM method, which is now available through server, is 80% accurate in predicting the topology of transmembrane segments in bacterial membrane proteins.The MEMSAT-SVM, which incorporates predictions for both signal peptides and re-entrant helixes, may also attain an accuracy level of 89%. Through the PsiPred server, you may also access this technique for free. Again, short-buried helices half-helices, re-entrant helices, as well as those with kinks, often result in a significant reduction in the accuracy of these instruments.

DISCUSSION

The large degree to which -helical (and to a lesser extent, the -strand) conformations fill their transmembrane segments is one of the most notable properties of integral membrane proteins. of order to reduce the energy cost of the backbone polar groups' exposure to the nonpolar protein core during folding, the and conformations of water-soluble proteins couple these backbone polar groups in hydrogen bonds. The main driver for these conformations' high frequency in membrane proteins is roughly the same, and it's even more obvious: these proteins' transmembrane regions are exposed to the cores of both the protein and the lipid bilayer, the latter of which is very nonporous. Therefore, in membrane proteins, electrostatic masking of backbone polar groups is considerably more crucial than in their water-soluble counterparts. Algorithms for predicting the structure of membrane proteins have taken use of the propensity of transmembrane segments to have significant amounts of -helical and -strand/sheet composition. For GPCRs in particular, specialised algorithms have been created.Integral membrane proteins are organised as -helices for the majority of their transmembrane segments. Despite the 'helix-breaking' capabilities of this residue and its low prevalence in the cores of helical segments of water-soluble proteins, these helices contain a significant quantity of Pro residues, as was previously indicated.

Based on their MD simulations, Sansom and colleagues hypothesized that Pro residues serve as motion hinges in trans-membrane segments, facilitating both functionally significant conformational alterations and improved helical orientation adjustment. As we'll see in more detail below, conformational changes are crucial for the operation of integral membrane proteins. For instance, they help channels gate, transporters transition between inside- and outside-open states, and receptors and enzymes move between active and inactive states. These adjustments include a variety of movements in the protein, from positional alterations of whole domains and subunits to hinge bending or displacement of individual helices in screw or pivot motions. As was before discussed, Pro is known to kink helices and provide the polypeptide chain rigidity. It may thus come as a surprise at first because Pro is crucial for conformational changes in transmembrane proteins. However, Pro residues often exist in trans-membrane regions of sequences that also include Ser and Thr, such as (S/T)P, (S/T)AP, and PAA(S/T). Simulations show that these residues make up for the Pro's structural deformation and imply that these structural effects enable the helices to go through the necessary conformational alterations.

This idea is supported by experiments, which reveal that changing Ser and Thr within these motifs alters the protein's function. However, it is still puzzling why these phenomena an are almost exclusively seen in integral membrane proteins as opposed to all proteins. This could be related to the limitations placed on membrane protein movements by the structure of the lipid bilayer. It follows those special properties, such as the inclusion of Pro residues inside secondary structures, may have developed as a means of giving these proteins the same

degree of flexibility that typically exists in their water-soluble counterparts. This is because integral membrane proteins are more constrained by their environment than water-soluble proteins. Many of the benefits that Pro confers are thus likely to turn into disadvantages in the absence of the surrounding membrane. This idea presents the intriguing possibility that, via a process known as negative design, Pro may also be crucial for membrane proteins. In particular, it inhibits the folding of membrane proteins outside of the membrane, which would make them unstable and susceptible to destruction. According to a recent research, the mere presence of Pro residues in certain locations on transmembrane segments might prevent membrane proteins from misfolding by favoring native structures over non-native ones that include a variety of -strands [5], [6].

Other structural abnormalities, including narrow turns of -helices and tight twists of 310 helices, may also be present in transmembrane segments. It is currently unknown if the latter two anomalies, which have received less research than Pro-induced kinks, have any functional purpose. Such distortions have the benefit of allowing close contact between polar groups in neighbouring helices, which improves electrostatic masking compared to intrahelical interactions alone. In fact, compared to just 19% of the helices in water-soluble proteins, approximately 40% of the transmembrane helices in membrane proteins are deformed. Transmembrane helices may also be broken or only partially penetrate the membrane core, creating re-entrant loops. In channels like aquaporins and K+-channels, where the exposed residues in the re-entrant loops often function as binding sites for ions or other substrates, the latter are especially prevalent.Integral membrane proteins with extended conformations are thought to make up a very tiny percentage of the whole proteome since they are less frequent than those with -helical conformations. These objects, sometimes known as -barrels, typically have a single chain but may also have numerous chains.

The majority of -barrels are small molecule channels that are found in the outer membranes of mitochondria and chloroplasts, as well as the membranes of Gram-negative bacteria. When porins in bacteria are exposed to their surroundings, many of them become attachment sites for phages and bacterial toxins. In fact, several of the toxins cause the host membrane to take on a -barrel form. In the barrel structure, the strands are anti-parallel and are joined by large loops on the outside of the cell or organelle and small loops on the periplasmic side. It has been hypothesized that the -hairpin motif is the main evolutionary unit of all -barrel proteins based on its structure. The barrel structure is amphipathic, with a water-filled centre and a nonpolar outside, in line with the function of porins. Furthermore, poor selectivity is related to the barrel's wide breadth. Comparatively to channels found in the plasma membrane or inner mitochondrial membrane, both of which are constructed of -helical bundles, porins may transport a wider range of polar compounds. Finally, within the membrane, porins have a propensity to oligomerize. Aquaporins, which are -helical channels and members of the major intrinsic protein (MIP) superfamily, should not be confused with porins.

Integral membrane proteins are found in a lipid environment, which explains why many of their polar residues tend to face the inside of the protein and why they have nonpolar exteriors. These findings led to the first hypothesis that these proteins are 'inside-out' adaptations of water-soluble proteins. However, this assumption has shown to be oversimplified with the structural characterization of many membrane-bound proteins. Instead, there are some similarities between the structures of water-soluble proteins and integral membrane proteins. First off, the core of both protein types is more evolutionarily conserved than the protein's surface. It is also densely packed, mostly composed of nonpolar residues with a few numbers of functionally significant polar residues. The fact that residues in the core are more structurally restricted as a consequence of tighter inter-residue packing

than residue-lipid packing is likely what causes the remarkable conservation of the core. Second, both protein types' structural loops play comparable functions in ligand binding and signal transduction.

As was already established, integral membrane proteins often include minor residues like Ala, Ser, and Gly in their transmembrane portions that help pack helices tightly. In addition to polytopic proteins with helices, biopic proteins with a propensity to dimerize or oligomerize inside the membrane also need proper packing. Additionally, the arrangement of the smaller residues amid the larger residues results in ridges and grooves along the helix, which generates geometric complementarity between neighbouring helixes ('knobs-into-holes' packing). The optimal fit between neighbouringhelice requires them to lean across the membrane since these ridges and grooves are not mathematically parallel but rather twist around each helix. Although transmembrane helices in membrane proteins may be tilted by 5° to 35°, their average tilt is 20°, which seems to be the best tilt for interhelical packing.

Glycophorin A, which forms a -helical dimer when solubilized in detergent or lipid bilayers, is one of the most popular models used to research helix-helix interactions in membrane proteins. The helix-helix interface of the protein has the sequence motif LIxxGVxxGVxxT (x is any residue). The protein dimer divides into two monomers when these residues are swapped out for mutations. The pattern also contains the smaller motif GxxxG, which is overrepresented in transmembrane regions and is found in several interacting helices of water-soluble proteins. The two Gly residues enable the two interacting helices to be separated by just 6 since they are situated on the same face of the helix, a turn apart from one another. This close proximity enhances van der Waals interactions and permits a hydrogen bond between the C-H group of one helix and the backbone carbonyl group (C-O) of the other. However, it has been shown that the short residues Ser and Ala often take the place of Gly residues, expanding the GxxxG motif into a GxxxG-like (or GAS) motif. The leucine-isoleucine zipper, the heptad serine zipper, and the GxxxGxxxG glycine zipper motifs are additional motifs that have been proposed to mediate tertiary contacts in membrane proteins. The MeMotif database, which can also locate these motifs in particular sequences, contains these and additional linear motifs that mediate helix-helix and protein-protein interactions.

Finally, transmembrane helices' charged Glu, Asp, Lys, and Arg residues have been suggested to mediate helix-helix interactions. The next paragraph goes through how such interactions may have an impact on energy. Polar interactions inside the protein offer an extra advantage over nonpolar interactions, which only depend on steric complementarity, much as in water-soluble proteins. Integral membrane proteins have a tendency for certain structural configurations to recur. The -helical bundle, which recurs in numerous forms, is by far the most prevalent. For instance, GPCRs include a distinctive seven-helical bundle, while certain transporter families contain 12 or 14 helical bundles. The - barrel is a less frequent structural motif that describes channels with limited selectivity in bacterial outer membranes and eukaryotic organelles of bacterial origin. Integral membrane proteins go through a complicated process to achieve their ultimate active state in the membrane. The essential phases that each transmembrane segment goes through are as follows:

1. Ribosome-mediated translation.
2. The translocon complex inserting itself into the membrane.
3. Purchasing auxiliary structure.
4. Construction of the mature protein, often in the form of a -helical bundle, with the additional transmembrane segments. Steps 2 and 3 are often referred to as a single linked process since membrane insertion frequently hinges on the creation of secondary structures in transmembrane segments [7], [8].

The four-step process is complicated overall, and although recent structural studies have helped to clarify some of the complexity, it is still not fully understood. Using model systems based on whole proteins or isolated peptides and concentrating on the energetics of the four key steps rather than delving into the numerous kinetic barriers and minima that the complex translocon system involves is one alternative for studying this process, with all the biological components included therein. Estimates of the energy associated with the main phases, particularly the first three, have been provided by studies using such systems as well as more realistic settings (see and references therein). Generally speaking, the following factors predominate in the energetics of transmembrane segment transfer from the aqueous phase into the membrane:

1. The cost in free energy incurred when the (polar) peptide backbone is partitioned into the lipid environment. For C, O, and N backbone groups that are hydrogen-bonded to one another, the penalty is computed as +2.1 kcal/mol, and it is +6.4 kcal/mol when the hydrogen bond is not met.
2. The advantageous contribution of free energy brought on by the hydrophobic effect. This free energy value has been calculated to be between 25 and 30 cal/mol per mol of the protein engaged in nonpolar interactions on the basis of straightforward partition tests.

A lesser value of 10 cal/mol per 2 was obtained in experiments employing a more realistic system, in which peptides were introduced into the ER membrane via the Sec61 translocon. Therefore, for the net membrane-partitioning free energy to be advantageous, the favourable free energy must offset both the free energy penalty and the cost of polar side chain insertion.The nonpolar environment inside the membrane causes the inserted segments to develop secondary structures. This step's energetics are virtually identical to those of the folding process that induces secondary structure in water-soluble proteins. Bacteriorhodopsin observations using atomic force microscopy show that the free energy of a single -helical transmembrane segment during the linked insertion-folding process is 1.3 kcal/mol per residue. A scale may be used to estimate how deeply each transmembrane segment should be inserted into the membrane. Each naturally occurring amino acid is given a value on the scale that represents the free energy of its insertion into the hydrocarbon core or the polar head group area.Each folded transmembrane segment interacts with its neighbours within the membrane to generate the completely folded protein in the process of forming membrane proteins. Bacteriorhodopsin's net energy has recently been determined to be 11 kcal/mol, which is comparable to the energy required for water-soluble protein folding. With respect to its energy components and their size, this step is considerably more contentious than the previous phases of membrane insertion and acquisition of secondary structure of each transmembrane segment.

As we have seen, the hydrophobic effect is what propels the folding of water-soluble proteins. Since a lipid medium surrounds integral membrane proteins, a distinct driving factor must be present for their folding. According to studies, the amount of the protein's surface area that is buried during folding is correlated with the stability of membrane-bound proteins. When the hydrophobic effect is absent, this correlation should mostly represent van der Waals interactions. Indeed, compared to their water-soluble counterparts, van der Waals interactions are predicted to have a greater influence on the folding of membrane proteins. This anticipation is mostly based on the fact that water molecules are better at rearranging around the unfolded protein and producing a tight contact shell than lipid molecules because of their smaller size. Because of this, it seems sense that the van der Waals interactions between an unfolded protein and the molecules in its surroundings would be greater for

water-soluble proteins than membrane proteins. This indicates that in membrane proteins, the strengthening of these connections during folding is more significant. Additionally, the aforementioned patterns in membrane proteins allow transmembrane segments to be packed very tightly, which in turn enhances the van der Waals interactions between them.

The stability-surface area connection may potentially indicate an additional impact, which is entropic in nature and resembles the hydrophobic effect in water-soluble proteins, in addition to van der Waals interactions. When a protein is put into the membrane, the amount of mobility that the lipids that make up the membrane have reduced. Lipids in close proximity to protein residues are the main objects of this limitation of mobility. Some of those lipids are released into the bulk of the protein during folding, which improves the lipids' mobility. This means that folding raises the entropy of the lipid bilayer while decreasing the entropy of the polypeptide chain. This process may be seen as a folding driving force since it is advantageous. This phenomenon is comparable to the hydrophobic effect in proteins that are water-soluble because, in both situations, the release of 'solvent' molecules during protein folding, such as lipids or water, increases total entropy. Given that the change in the freedom of movement of lipids is anticipated to be less than in the case of the much smaller water molecules, this impact is most likely quantitatively weaker than the hydrophobic effect.

Like in the case of water-soluble proteins, it is unclear exactly how electrostatic interactions affect the folding of membrane proteins. As we previously observed, polar residues are often found in patterns that are known to favour helix-helix packing in membrane proteins, indicating that polar interactions are generally advantageous in this context. Studies that focused on hydrogen bonds between transmembrane segments did in fact imply that these connections are advantageous and are what causes the protein to assemble. DeGrado and colleagues, for instance, investigated this problem using dimerizing helical model peptides. Dimerization was prevented by mutagenesis when hydrogen bonds involving the amino acid residues Asn, Gln, Asp, and Glu were broken. Since the potential hydrogen bond donors and acceptors are surrounded by a less polar environment in their isolated state (within the lipid bilayer) than they are in the folded state, it makes thermodynamic sense that hydrogen bonds contribute favourably to the assembly of the helices into a folded protein. Again, this idea is supported by the fact that polar residues like Ser and Thr are more common in membrane-bound proteins than in water-soluble proteins [9], [10].

How substantial is the stabilisation, assuming transmembrane hydrogen bonding indeed stabilise the folded shape of membrane proteins? Bacteriorhodopsin, a bacterial light-activated proton pump with a well-known structure that is often used as a model for membrane proteins, was employed in research by Bowie and colleagues to examine this problem. They assessed the change in protein stability after replacing hydrogen-bonding pairs with non-hydrogen-bonding residues that were very identical to the original residues in all other respects.

Their findings indicated that the average hydrogen bond contribution to stability is rather little, on the order of 1kBT. What possible evolutionary justification exists for continuing to use such weak stabilizing forces? First off, numerous hydrogen bonds between transmembrane segments that include a number of polar residues may result in significant additional stabilization. Second, Bowie and colleagues hypothesised that, similar to Pro-related hinges, these weak interactions may be exactly what membrane proteins need to maintain their high degree of flexibility and generate the characteristic helical distortions that define their structure. Finally, some studies contend that hydrogen bonding weaken membrane proteins.

CONCLUSION

In conclusion, there have been important developments in molecular biology as a result of the research of integral membrane proteins and the creation of tools to comprehend their structure and function. These proteins are essential for many cellular activities, and the majority of their transmembrane regions are made up of non-polar residues. To find and describe these regions of genomes, researchers have created cutting-edge algorithms and hydrophobicity scales. In the area, there has been discussion on the best physical metric to use to describe hydrophobicity as well as the best measuring medium. In order to more reliably forecast transmembrane segments, knowledge-based scales have emerged as useful alternatives. These scales use statistical information from membrane protein structures. In order to forecast transmembrane segments well, more information is required than only sequence tendencies. The structural characteristics of transmembrane segments, such as the presence of Pro residues, kinks, and re-entrant loops, have been studied. Several methods have also been developed to predict the transmembrane segment topology, which has improved our comprehension of the structure of membrane proteins.

REFERENCES:

[1] S. H. DeLuca, S. L. DeLuca, A. Leaver-Fay, and J. Meiler, RosettaTMH: A method for membrane protein structure elucidation combining EPR distance restraints with assembly of transmembrane helices, *AIMS Biophys.*, 2016.

[2] N. Di Bartolo *et al.*, Complete reversible refolding of a G-protein coupled receptor on a solid support, *PLoS One*, 2016.

[3] D. M. Miller, H. E. Findlay, O. Ces, R. H. Templer, and P. J. Booth, Light-activated control of protein channel assembly mediated by membrane mechanics, *Nanotechnology*, 2016.

[4] H. T. Bergal, A. H. Hopkins, S. I. Metzner, and M. C. Sousa, The Structure of a BamA-BamD Fusion Illuminates the Architecture of the β-Barrel Assembly Machine Core, *Structure*, 2016.

[5] J. Stevenson, E. Y. Huang, and J. A. Olzmann, Endoplasmic Reticulum-Associated Degradation and Lipid Homeostasis, *Annual Review of Nutrition*. 2016.

[6] P. J. Focke *et al.*, Combining in Vitro Folding with Cell Free Protein Synthesis for Membrane Protein Expression, *Biochemistry*, 2016.

[7] J. L. Popot and D. M. Engelman, Membranes Do Not Tell Proteins How to Fold, *Biochemistry*, 2016.

[8] K. Claes *et al.*, Modular Integrated Secretory System Engineering in Pichia pastoris to Enhance G-Protein Coupled Receptor Expression, *ACS Synth. Biol.*, 2016.

[9] R. B. Thomson, C. L. Thomson, and P. S. Aronson, N-glycosylation critically regulates function of oxalate transporter SLC26A6, *Am. J. Physiol. - Cell Physiol.*, 2016.

[10] M. Palamini, A. Canciani, and F. Forneris, Identifying and visualizing macromolecular flexibility in structural biology, *Frontiers in Molecular Biosciences*. 2016.

CHAPTER 8

UNDERSTANDING THE COMPLEXITIES OF MEMBRANE PROTEIN STABILITY AND FUNCTION

Dr. Sangeeta Kapoor, Professor & HOD
Department of Biochemistry, TMMC&RC, TeerthankerMahaveer University, Moradabad, Uttar Pradesh India,
Email Id- drsangeeta2000@gmail.com

ABSTRACT:

Membrane proteins mediate a variety of functions, including as ion transport, signal transduction, and molecule absorption, which are crucial to cellular function. These proteins' intricate architectures and position inside lipid bilayers make them difficult to research. We examine the structural and functional characteristics of membrane proteins in this extensive overview, highlighting important examples like aquaporins and ion channels. The stability of membrane proteins is significantly influenced by the desolation cost related to exposing charged residues to the protein's core. Salt bridges in water-soluble proteins can be stabilising, but given that membrane proteins are exposed to lipids, their effects on membrane protein stability are more complex. But transmembrane salt bridges, particularly those arranged to promote helix-helix interactions, could collectively support membrane protein stability. Ancient water channels called aquaporins, which are present in many animals, make it easier for water to passively penetrate cell membranes. Small molecules like glycerol and ammonia are also transported by certain aquaporins, often referred to as aqua-glyceroporins. Selective water transport is made possible by their distinctive hourglass-shaped structure, which has two half-helices and six transmembrane segments. Asn-Pro-Ala and ar/R (aromatic-arginine), two important motifs, respectively create the small pore area and permit the transport of bigger solutes, which both contribute to selectivity. Asparagine residues in the NPA motif interact electrostatically with water molecules while excluding protons to achieve proton selectivity in aquaporins. Ion channels, which are highly selective for certain ions, are essential for many physiological functions. Similar to the ph-dependent bacterial K+ channel kcsa, the structure of K+ channels has an inverted teepee shape with a large water-filled hole crossing the lipid bilayer. The selectivity filter's coordination sites are where ions are selected, and molecular mimicry explains why K+ ions are preferred over Na+ ions of a smaller size.

KEYWORDS:

Glycerol, K+ Ion, Membrane Protein, Na+ Ions, Stability.

INTRODUCTION

There is a significant desolation cost associated with exposing charged residues to the protein's core, the effects of ionic interactions like salt bridges on membrane protein stability are more complex. Studies on salt bridges in the cores of water-soluble proteins show that these interactions may be generally stabilizing if they are optimised spatially and are a part of a larger network of electrostatic interactions that also includes hydrogen bonds. The salt bridges in membrane proteins may also be partly exposed to the lipid environment, which ought to increase the desolva- tion penalty compared to proteins that are water-soluble. However, the presence of salt bridges in transmembrane segments, where they are arranged as patterns that encourage helix-helix interactions, raises the possibility that these interactions function as a whole to stabilise membrane proteins.

As already indicated, a variety of structures with comparable properties have been produced as a result of the restrictions placed on membrane proteins by the chemically intricate and extremely anisotropic lipid bilayer. Transmembrane segments, for instance, are virtually invariably organised into -helical bundles, and to a lesser degree, -barrels. However, the basic design that all membrane proteins share may take on varied manifestations depending on the biological activities that each protein serves. Von Heijne, Gouaux, and mckinnon evaluate many significant instances of these architectural principles and their utilitarian value. Here, we concentrate on the channels and transporters—a class of proteins—that play a role in the movement of polar molecules across biological membranes. The structure-function link in the second main group of membrane proteins, ligand-activated receptors [1], [2].

Aquaporins

Aquaporins are ancient channels that are present in a variety of species and let water circulate passively through cells and tissues. For instance, they assist the kidneys of immatures in reabsorbing water from urine. Small substances like glycerol, ammonia, CO2, and oxygen may also be transported by some aquaporins, commonly known as aqua-glyceroporins. The primary intrinsic protein superfamily include aquaporins, which exist as tetramers in which each monomer functions as a separate pore. Aquaporins produce holes with an hourglass-shaped structure made up of two half-helices and six transmembrane segments. The narrow section (2.8) in the channel's middle is what gives it its selectivity against molecules bigger than water. In spite of the exceptionally fast transport rate (109 molecules/sec), even the water molecules themselves can only fit through the small tube in a single line. Side chains from two different kinds of motifs that face the inner side of the pore make up the narrow portion of this area.

1. There are two Asn-Pro-Ala motifs that are preserved.
2. An ar/R motif ('ar' for aromatic and 'R' for arginine), which is located on the extracellular side of the NPA motif at a distance of 8. Considered to be the main obstruction for big uncharged solutes, this pattern. Of fact, the ar/R residues of aqua-glyceroporins widen the aperture, enabling the transport of bigger solutes.

As was already noted, the NPA motif only has a little impact on the channel's ability to distinguish between charged and uncharged solutes. On the other hand, because protons are smaller than water molecules, it is essential for the selectivity of aquaporins against them. Because protons have a positive charge, the selection process is electrostatic as expected. Each water molecule that moves through the channel's middle is reoriented such that the NPA asparagine residues around it may form hydrogen bonds with the molecule's partly negative oxygen atom. Due to their positive charge, protons are unable to create these interactions. The dipoles of the half-helices that flank the NPA motif also help to exclude protons. As a result, protons are unable to go across the channel or hop between the single-file water molecules that fill the pore. Finally, it has also been proposed that the partial charges of NPA asparagine may aid in the water's quick flow down the channel.

Ion channels

Ion channels are engaged in a wide range of physiological processes, including muscular contraction, energy generation, chemical transport, neuronal transmission, and more. Defects in these proteins therefore cause a variety of diseases, including as cystic fibrosis, Bartter syndrome, and paralysis. An ion channel's structure is built around a central water-filled pore that spans the lipid bilayer and allows ions to dissolve without expelling energy into the hydrophobic environment of the bilayer core. The ph-dependent bacterial K+ channel (kcsa) is one of the most often used models for such channels. K+ channels have a distinctive

structure that consists of a lengthy (35 residue) cytoplasmic domain and a transmembrane domain, both of which are -helical. Each monomer contributes two trans-membrane helices (TM1 and TM2) and one cytoplasmic helix at the C-terminus to the channel, which is a tetramer. The transmembrane domain is shaped like an inverted teepee and features a large water-filled hole that spans more than two thirds of the lipid bilayer. As a result, most of the way across the membrane, the potassium ion may diffuse freely. The electrostatic repulsion between nearby K+ ions and this feature result in an extremely high K+ transit rate (. The four TM2 helices that make up the inner portion of the channel extend into the cytoplasmic domain. This domain stabilises the channel, enables channel interaction with regulatory components, and affects several functional aspects of the transport [3], [4].

The form, size, and width of channels vary widely, but almost all have some degree of selectivity towards the ions they convey. The specificity may be wide, such as allowing all elemental ions with a certain charge to pass, or tight, such as permitting just Ca2+ ions to pass. Most of the time, it seems that the 'candidate' ion must shed its solvation shell in order to attach to channel residues. The channel can evaluate the appropriateness of the ion thanks to this process. The selectivity filter, a loop-based substructure of the kcsa channel, houses the selectivity-related residues near the conclusion of the ion's passage. Four residues make up the TVGYG signature sequence, which forms the four equally spaced K+ binding sites known as S1 through S4 in each chain's loop that makes up the selectivity filter. Four backbone carbonyl groups and one side chain hydroxyl group help the sites coordinate the four K+ atoms. The desolvated ion's ability to continue travelling all the way to the opposite side depends on the filter. The biological significance of the filtration structure is shown, for instance, by the fact that scorpion poisons like kaliotoxin target it. While some ion channels choose the appropriate ion(s) based on a straightforward size limit, others may be more complex. This seems to apply to the kcsa channel, which prefers K+ over Na+ despite the latter's smaller size. Molecular mimicry is the term for the selection process.

DISCUSSION

This means that the arrangement of water-derived oxygen atoms surrounding each K+ ion in bulk solution is perfectly mimicked by the polar uncharged oxygen atoms that coordinate each K+ ion in the selectivity filter. When a result, almost no energy barriers are encountered when the K+ ion moves from the solution to the selectivity filter. Na+ ions may penetrate the selectivity filter because they are smaller, but they are too far away from the oxygen atoms for the latter to effectively electrostatically mask them. As a result, they often stay in bulk solution, or outside the channel. The ingenious molecular mimicry mechanism, put forward by mackinnon in his Nobel Prizewinning study, explains the kcsa channel's 1,000-fold selectivity for K+ over Na+. Data from subsequent investigations, however, indicate that the selection process is probably more intricate. In the selectivity filter area, for instance, NMR studies have shown greater protein flexibility, indicating that this region might be able to adjust to ions smaller than K+. Similar to this, variations of the selectivity filter throughout a range of ion-carbonyl distances, sufficient to coordinate either ion, as shown by molecular dynamics (MD) simulations of the kcsa channel in a lipid membrane.

According to the authors of the second research, selectivity is not determined by the pore's average size but rather by the inherent electrostatic properties of the coordinating carbonyl groups. Last but not least, a research integrating electrophysiology, X-ray crystallography, and MD suggested that smaller ions like Na+ and Li+ may attach to the selectivity filter readily but find it challenging to enter the filter from the intracellular side owing to a K+-dependent energy barrier. Once again, these data imply that ion channel selectivity is more complicated than previously thought. The K+, Na+, and Ca2+ channels have shown that,

despite the fact that minor elemental ions like K+ occupy relatively little space, their binding often includes noticeable changes in channel conformation.

The transport that ion channels carry out has to be controlled, or it needs to happen only in reaction to the proper signal, much like the situation with many biological proteins. The majority of channels feature a gate that serves to do this by blocking the flow of ions in their resting state. The signal that opens the channel might be natural in origin and either electric, chemical, or mechanical. An area of amino acids that serves as a sensor and reacts to an activating or deactivating signal is therefore constantly present in the gate. In voltage-gated K+ (Kv) channels, for instance, the positively charged amino acids Lys and Arg that are found on S4 helices are part of the sensor. The extracellular side of the membrane's high concentration of Na+ ions repel the positively charged S4 helices, keeping the channel closed in its resting state. However, the considerable drop in the amount of Na+ ions at the extracellular side during neu- ronalsignalling lessens the electric repulsion, allowing the S4 helices to be displaced and the channel to open. In this process, electrical energy is transformed into mechanical energy, which causes motion. The chemical energy stored in ATP or the electrochemical energy stored in an ion gradient is turned into motion in transporters that serve as ion pumps, as we shall see in more detail later. The intracellular side of the transmembrane domain and the start of the cytoplasmic domain are where the H+-activated gate, one of the two gates that make up the kcsa channel's complicated gating system, is located. When the ph drops, this gate opens, and it is closed when the ph is high.

An H+-unrelated gate that is located at the selectivity filter. Since the first gate is the one that reacts to a change in ph, it is understandable that the majority of investigations on kcsa gating have concentrated on it. In both the transmembrane and cytoplasmic domains of the TM2 helices, activation of this gate results in conformational changes. The helices move in a 15° hinge-bending motion around Gly-104 in the transmembrane domain. The conformational shift in the cytoplasmic domain is concentrated at Val-115, where the TM2 helices form a bulge. Each TM2 helix there moves 4 outwards upon activation, resulting in an overall 20-opening of the channel*1 at the narrowest location for K+ penetration. Three residues, Arg-117, Glu-118, and Glu-120, have been identified as the kcsa channel's ph sensors, and it seems that ph sensing occurs in the bulging area as well. It has been proposed that a ph reduction below the channel's pka causes the two glutamate residues to protonate, rendering them electrically neutral. However, the precise mechanism of this sensing is still unclear. These residues can no longer hide and stabilise the positive charges on Arg-117 as a consequence. The displacement of the helices and opening of the channel are caused by the electrostatic repulsion that results from the interaction of Arg-117 with nearby monomers. The channel also has a second gate, located at the selectivity filter, as was already indicated. The second gate is believed to be affected allosterically by the conformational shift described for the H+-activated gate, which causes the channel to enter a completely open state [5], [6].

Ion channels may also exist in an inactivated state in addition to the active and resting states. After a period of activity, channels often deactivate to reduce the cell's response. K+ channels may be inactivated in one of two ways. The first is seen in voltage-gated K channels and is referred to as N-type inactivation. An electrically charged N-terminal portion of the channel physically blocks the pore during this quick process. The kcsa channel exhibits the second form of inactivation, known as C-type inactivation, which is substantially slower than N-type inactivation. The H+-independent gate at the selectivity filter is hypothesised to be involved in C-type inactivation via the reorientation of the backbone carbonyls and consequent destabilisation of K+ ions within the filter. Intriguingly, the broad cytoplasmic domain seems

to delay the pace of channel inactivation even if it limits the degree of transmembrane domain opening during activation.

Some transporters enable the passive passage of ions along their electrochemical gradients, much like ion channels do. However, many other transporters move bigger solutes and some of them can even 'pump' or move solutes against their electrochemical gradients. In any event, it is evident that transporters need to make use of a more advanced transport mechanism than channels do. In fact, transporters exist in two distinct states, one of which permits the ligand to enter the transporter on one side of the membrane and the other of which releases it on the other side. This is contrary to the common misconception that transporters are only water-filled structures. 'The alternating access mechanism' is the name given to this paradigm, which Oleg Jardetzky presented in an abstract form. According to its structural requirements, each transporter must have at least two different conformations that correspond to the aforementioned states. The semisweet trans- porter, which promotes passive diffusion of sugars into bacteria, has such a mechanism.

This transporter performs a similar role to the human GLUT and SGLT transporters, which are essential for sugar metabolism. The inner side of the transporter is blocked by a collection of aromatic and nonpolar residues, while it is in the outward-open state. Several conformational changes are required to transition from the outward-open state to the inward-open state. The cytoplasmic half of TM1 (TM1b), which is hinged to the highly conserved Pro-21, first tilts 30° in relation to the exoplasmic half of TM1 (TM1a). Second, a rotational movement of the other monomer's TM1b, TM1a, TM2, and TM3 results in a movement known as a binder clip. In other words, the cytoplasmic portions of the transmembrane helices are pushed apart while their exoplasmic portions move towards one another. This motion causes the cytoplasmic gate to open and the extracellular gate to close simultaneously. Tyr-53, Arg-57, and Asp-59 make up the EC gate. These residues interact polarly with the corresponding residues in the neighbouring monomer when the gate is closed, stabilizing the closure.

An active transporter, often known as a pump, moves substances in opposition to their electrochemical gradient. It is necessary for this process to have an energy source, which may either be ATP or a common cellular ion's electrochemical gradient. Active transporters alternate between outward- and inward-facing conformations, much as passive transporters. However, in an active trans- porter the prevailing conformation at any given moment is decided by the binding of ATP/ADP/Pi or ions to a particular location in the protein. This is so that a unique conformation may be stabilized by the binding of each of these species to the protein. In an ATP-dependent transporter, switching the protein between several conformations that face various sides of the membrane occurs throughout a cycle that comprises (1) ATP binding, (2) ATP hydrolysis to ADP and Pi, and release of one or both of these. The bacterial multidrug transporter Sav1866, which belongs to the class of ATP-binding cassette (ABC) transporters, serves as an example of such a mechanism. Both prokaryotes and eukaryotes include these proteins, which pump tiny molecules out of the cell.

The way these transporters' function is also crucial for health; some ABC transporters in bacteria pump out antibiotics, which breeds resistance. Similar to this, tumour cells employ these transporters to release chemotherapy, which lowers the treatment's effectiveness. When Sav1866 is in a monomeric form and its transmembrane domain is exposed to the inside of the cell, the solute export cycle of the protein starts. The targeted solute may bind to the transmembrane domain thanks to this conformation. Two Sav1866 monomers are nucleotide-bound by ATP, which causes dimerization and an outward-facing orientation. It is believed that ATP hydrolysis causes further conformational changes that return the trans- porter to an

inward-facing configuration so it is prepared to bind a fresh solute molecule. It should be emphasised that the precise relationship between ATP binding or hydrolysis and solute transport in ABC transporters is not always as straightforward as shown above, and it may sometimes be quite weak. For each substrate, the bacterial vitamin B12 importer btuc2d2 hydrolyzes many ATP molecules, and this hydrolysis occurs even when the substrate is not present.

The alterations between inward- and outward-facing conformations in active transporters caused by nucleotide or ion binding do not adequately explain how these transporters may pull a solute from a compartment with low concentration and release it into a compartment with high concentration. This capacity is related to each conformation's innate affinity for the solute. Since the binding and release of the solute are controlled by its concentration in the two opposing compartments, the affinity of the inward-facing conformation in passive transporters may be identical to that of the outward-facing conformation. The conformation facing the compartment with low solute concentration in active transporters needs to have a sufficiently high affinity to the solute to bind it, whereas the conformation facing the compartment with high solute concentration needs to have a sufficiently low affinity to release the solute. Binding of nucleotides or ions indirectly regulates variations in affinity throughout the transport cycle by stabilising the two distinct conformation types. For instance, ATP binding stabilises the outward-facing conformation of the ABC transporters, which has a low affinity for the solute, allowing the solute to be discharged into the outside of the cell. Binding proteins that scavenge the solute and then attach to a specific domain in the transporter help certain transporters function [7], [8].

Therefore, rather of releasing energy directly, ATP hydrolysis and electrochemical gradient dissipation may drive active transport by sequentially binding nucleotides or ions that stabilise different conformations of the transporter. These conformations vary from one another depending on whether the transmembrane domains of the transporter are exposed to the inner or outer sides of the membrane and how receptive they are to the solute being transported.The ATP-dependent ion trans- porters use the active transport principles mentioned above. These proteins are grouped into the following categories since they are essential components of every organism:

E1E2-atpases, or P-atpases. These transporters are present in eukaryotic plasma membranes and organelles, as well as in bacteria, fungi, and other organisms. They move several ions across membranes, including as H+, Na+, K+, and Ca2+. Some of the most important main active transport proteins in the cell fall within this group, including the Na+/K+-, the H+/K+-atpase, and the Ca2+-atpase. Different uses are made of the electrochemical gradients that these proteins create. Bacteria, for instance, employ the proton gradient to power activities like chemotaxis and secondary active transport. Cytoplasmic and transmembrane domains are seen in P-atpases. Phosphorylation (P)*1, nucleotide-binding (N), and actuator (A) subdomains are all found in the cytoplasmic domain. Six helices make up the transmembrane domain, which is where membrane transport occurs. But most P-atpases also have extra transmembrane helices. For instance, the transmembrane domain of the Na+/K+-atpase and Ca2+-atpase has 10 helices. These helices also include cation-binding sites, which are expected to help the trans- porter be more ion-selective. The extracellular component of several P-atpases, including the Na+/K+-atpase and the H+/K+-atpase, is crucial for the correct trafficking of the transporter to the plasma membrane and influences other functional aspects of the transport.

F-atpases. These H+-atpases are found in bacterial plasma membranes, chloroplasts, and mitochondria. In contrast to P-atpases, F-atpases employ proton transport as the energy

source for ATP production rather than the other way around. To make ATP, F-atpases specifically exploit the H+ electrochemical gradient that is produced during cellular respiration in mitochondria or photosynthesis. Thus, rather of being called atpases, these proteins are more commonly referred to as 'ATP synthases'. The mitochondrial F-atpase's structure and mechanism are detailed. (V1VO-atpases) V-atpases. The membranes surrounding eukaryotic organelles, such as vacuoles and lysosomes, are where these H+-atpases are predominantly located. In the latter, by pumping protons into the organelle, V-atpases serve to acidify it. Additionally, they aid in the net excretion of acid into urine in some kidney segments. They are made up of more than 10 subunits, giving them a complicated structure. Archaea include A-atpases (A1AO-atpases), which work similarly to F-atpases. A variety of nucleoside triphosphates (ntps), including extracellular ATP, are hydrolyzed by E-atpases, which are cell-surface enzymes.

Ion pumps operate similarly to other transporters by cycling between various conformations that 'take up' the solute on one side of the membrane and release it on the other side. The three-dimensional architectures of the Na+/K+-atpase and the Ca2+-atpase provide evidence of such a cycle. Both structures show that the ions must 'hop' between amino acids inside the transporter, which act as temporary binding sites, in order to penetrate the membrane (see for the Na+/K+-atpase). the negatively charged side chains of glutamate and aspartate residues, as well as the partial charges of main chain or side chain carbonyl and hydroxyl groups, as well as individual molecules, are all present in the cation-binding sites in the Na+/K+-atpase and the Ca2+-atpase, respectively. Similar to what we have seen in the kcsa (potassium channel), partial charges that momentarily bind desolubilized ions are involved. The introduction of partial charges here also guarantees that ions may circulate across the protein rather than being immobilised in one location. Although the lack of a water-filled path in transporters may seem disadvantageous, it is actually one of the things that prevents the transported ion from being accidentally exposed to both sides of the membrane at once, which would allow it to return to the compartment from which it was taken [9], [10].

As we've seen before, the conformational changes that underpin the transport mechanism in active ATP-dependent transporters (such the ABC transporter Sav1866) are triggered by ATP binding and hydrolysis. All atpases, including the ion pumps described in the preceding chapters, have this property. The impacts of ATP and its hydrolysis products (ADP and Pi) on the transporter's energy landscape are the method by which atpases convert these events into conformational changes. That is, ATP, ADP, or Pi binding to the transporter causes a conformational change, a shift in energy, and a return to the energy minimum. Each binding results in a different conformation because ATP, ADP, and Pi have distinct impacts on the energy of the transporter. The ability of atpases to cycle between conformations in response to the binding of ATP, ADP, and Pi allows them to transport ions. As the A, N, and P subdomains shift in relation to one another within the cytoplasmic domain of P-atpases like the Na+/K+-, Ca2+-, and H+-atpases, the conformational changes spread to the transmembrane domain through linkers and tertiary contacts. We will examine the Na+/K+-atpase, which moves two K+ ions in the opposite direction and three Na+ ions from the cytoplasm to the cell's exterior, to show the whole transport cycle. The phases in the transport cycle are as follows:

1. The Na+-binding site on the transporter's ATP-bound form, known as E1, faces the cytoplasm and has a strong affinity for Na+ ions. Three cytosolic Na+ ions attach to the protein as a consequence of these characteristics, occluding the binding site.
2. The transporter has an intrinsic atpase activity, and the transfer of the -phosphate to a conserved aspartate in the protein's P domain occurs simultaneously with the

hydrolysis of ATP. Two effects result from this occurrence, which is followed by the release of ADP. First, it significantly lowers the transporter's affinity for Na+, and second, it encourages a conformational shift that exposes the Na+-binding site to the cytoplasm. These modifications cause the protein to protonate and the release of three Na+ ions to the membrane's extracellular side. They also cause the binding of two K+ ions.

3. The dephosphorylation of the protein's conserved aspartate and blockage of the binding site, as well as the release of Pi and the binding of fresh ATP, are all facilitated by the K+ ion binding. The ATP-bound protein changes back to the E1 form, causing the released K+ ions to be transported to the cytoplasm. The stages requiring binding and release of the second ion type are not included in the transport cycle in P-atpases that only transport one kind of ion.

Structure forecast

We have seen how scientists can recognised transmembrane segments in the amino acid sequence and forecast their structure in relation to other protein sections thanks to the strong sequence-related tendencies of these segments. The orientations of helices and the conformation of the protein backbone and side chains are the two last stages for a successful prediction of the whole structure. According to the qualities they predict, such as secondary structure, topology, tertiary structure, etc., structural prediction approaches of membrane proteins may therefore be categorized. Structure prediction techniques may also be categorized according to the data they utilize to make their predictions. Methods may be divided into three major categories based on the latter requirement. Ab initio techniques fall under the first category. These techniques only use physicochemical properties of membrane proteins, such as length, hydrophobicity, etc. Such a technique, which was created especially for gpcrs, is PREDICT. The main benefit of such approaches is that they just need the amino acid sequence as extra information about the protein. Their fundamental drawback is that they need large, computationally costly conformational sampling. The second category of techniques includes those that make use of statistical tendencies for residues to occur in certain protein locations. One such is the kprot scale, which measures how likely amino acids are to interact with the protein core or the lipid bilayer based on statistical information gleaned from bitopic and polytopic proteins, respectively.

Statistical propensities for evolutionary conservation, linked mutations, and tight packing against other residues are three further statistical tendencies that have been effectively used to structure prediction algorithms.

The third category of membrane protein structural prediction techniques comprises template-based techniques that compare the query protein's sequence to membrane proteins with known structures. The recommended technique is homology modelling when the query and template proteins are homologous. However, since there are so few homologues with known structures that can be used as templates, this technique, which is very effective for water-soluble proteins, is less effective for membrane proteins. Fold recognition techniques may be used in these circumstances. In these techniques, the tem- plate is selected based on similarities in the sequence of the term plate rather than the query protein's sequence. Others are physically significant, such as the propensity to form a specific secondary structure, to be exposed to the surrounding medium, to have specific dihedral angles, to interact with specific amino acids, etc. Some of these properties are statistically extracted from multiple sequence alignments.

CONCLUSION

In conclusion, research on membrane proteins demonstrates the astounding variety and complexity of these vital biological components. The specific difficulties presented by the hydrophobic lipid bilayer have given rise to the creation of various structural and functional adaptations in membrane proteins, despite the fact that water-soluble proteins have been the subject of substantial study. The considerable desolvation cost connected with exposing charged residues to the protein's core is one of the important results included in this review. This desolvation penalty emphasises how crucial it is to understand how ionic interactions, including salt bridges, contribute to membrane protein stability. Membrane proteins, in contrast to water-soluble proteins, also have to contend with the risk of exposure to the lipid environment, which raises the desolvation penalty. The positioning of salt bridges in transmembrane regions, which encourages helix-helix interactions, however, raises the possibility that they serve a practical purpose in stabilising membrane proteins. In conclusion, membrane proteins are vital parts of biological membranes that enable a variety of crucial processes. The rigorous lipid bilayer environment shapes their structures and modes of activity, leading to a wide variety of protein adaptations. Undoubtedly, increased understanding of the functions and complexity of membrane proteins in cellular biology will come from further study in this area.

REFERENCES:

[1] G. Hedger and M. S. P. Sansom, Lipid interaction sites on channels, transporters and receptors: Recent insights from molecular dynamics simulations, *Biochim. Biophys. Acta - Biomembr.*, 2016.

[2] H. Kowata, S. Tochigi, T. Kusano, and S. Kojima, Quantitative measurement of the outer membrane permeability in Escherichia coli lpp and tol-pal mutants defines the significance of Tol-Pal function for maintaining drug resistance, *J. Antibiot. (Tokyo).*, 2016.

[3] M. L. Ortiz-Suarez, F. Samsudin, T. J. Piggot, P. J. Bond, and S. Khalid, Full-Length OmpA: Structure, Function, and Membrane Interactions Predicted by Molecular Dynamics Simulations, *Biophys. J.*, 2016.

[4] P. Barth and A. Senes, Toward high-resolution computational design of the structure and function of helical membrane proteins, *Nature Structural and Molecular Biology*. 2016.

[5] Y. Lu, Identification and roles of photosystem II assembly, stability, and repair factors in Arabidopsis, *Frontiers in Plant Science*. 2016.

[6] J. S. Gunn, L. O. Bakaletz, and D. J. Wozniak, What's on the outside matters: The role of the extracellular polymeric substance of gram-negative biofilms in evading host immunity and as a target for therapeutic intervention, *Journal of Biological Chemistry*. 2016.

[7] C. Bergsdorf *et al.*, An Alternative Thiol-Reactive Dye to Analyze Ligand Interactions with the Chemokine Receptor CXCR2 Using a New Thermal Shift Assay Format, *J. Biomol. Screen.*, 2016.

[8] N. Di Bartolo *et al.*, Complete reversible refolding of a G-protein coupled receptor on a solid support, *PLoS One*, 2016..

[9] S. Di Lella, A. Herrmann, and C. M. Mair, Modulation of the pH Stability of Influenza Virus Hemagglutinin: A Host Cell Adaptation Strategy, *Biophysical Journal*. 2016.

[10] R. A. Black and M. C. Blosser, A self-assembled aggregate composed of a fatty acid membrane and the building blocks of biological polymers provides a first step in the emergence of protocells, *Life*. 2016.

CHAPTER 9

HYBRID PREDICTION ALGORITHMS: LIPID-PROTEIN AND MEMBRANE PROTEINS

Dr. Sushil Yadav, Assistant Professor
Department of Biochemistry, TMMC&RC, TeerthankerMahaveer University, Moradabad, Uttar Pradesh India,
Email Id- yadav.sushil4644@gmail.com

ABSTRACT:

The prediction of protein structures remains a key difficulty in the field of structural biology, and modern prediction algorithms often use hybrid techniques that include components from many disciplines. Such hybrid models have been successfully used to predict protein structures with great accuracy using two well-known approaches, Rosetta and I-TASSER. Enhancing structure predictions has also been made possible by recent developments in the use of experimental data from techniques like cryo-electron microscopy (cryo-EM), small-angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR), circular dichroism (CD), and Förster resonance energy transfer (FRET). Particularly cryo-EM has become a useful technique for allowing structures with almost atomic resolution. Software tools like SAXTER, CHESHIRE, Rosetta, and CS23D have been modified to add experimental data-derived empirical limitations into the prediction process, therefore restricting the conformational space that algorithms may explore. It might be difficult to predict the architectures of proteins enmeshed in peripheral membranes. On the extracellular or cytoplasmic side of the lipid bilayer, these proteins mostly interact with the aqueous phases. Through both particular lipid-protein interactions and general bulk physical properties, the lipid bilayer has a variety of impacts on membrane proteins. These impacts have an impact on the stability, folding, assembly, and activity of proteins. The behaviour of membrane proteins may be affected by the lipid composition, degree of order, viscosity, and other characteristics of the bilayer. In addition, the hydrocarbon area of the bilayer's thickness and its compatibility with transmembrane protein segments are key factors in protein-lipid interactions. Additionally, mechanosensitive proteins aid in controlling water efflux in response to membrane stretching, and lipids' capacity to create positive or negative curvature in the bilayer may have an impact on protein function. There are three types of lipids that interact specifically with proteins inside the bilayer: bound lipids, annular lipids that create a contact layer surrounding the protein, and bulk lipids. These relationships are mediated by non-polar interactions with acyl chains and interactions between protein residues and lipid head groups. Near the polar-nonpolar barrier, aromatic residues like Tyr and Trp bind transmembrane protein segments and are crucial for the stability of membrane proteins.

KEYWORDS:

Hydrocarbon, Hybrid Prediction, Lipid-Protein, Membrane Proteins.

INTRODUCTION

Prediction algorithms used nowadays are often hybrid, i.e., they are built on a scoring function that incorporates expressions taken from many methodologies. For instance, although some of the expressions may depend significantly on statistically determined inclinations of certain amino acids to form particular structures or to be engaged in particular physical interactions, others may rely more substantially on physicochemical considerations. A unified modelling framework is provided by two well-known techniques, Rosetta *1 and I-TASSER *2, in which several methodologies are employed for various modelling issues.

Both approaches have been tried on several proteins and have produced accurate predictions.Last but not least, a relatively recent method has been implemented that employs experimental data in order to improve predictions. These results are generated from experimental techniques, primarily cryo-EM, SAXS, and NMR, but also CD and FRET, that give low-resolution structures of membrane proteins.

The fundamental concept is to utilise the retrieved data as spatial constraints to limit the conformational space that the prediction algorithms explore, greatly increasing the likelihood that they will provide the native protein structure. In this regard, cryo-EM is particularly interesting since standard EM techniques have been utilised to offer information on the quantity, tilt, and general positions of transmembrane helices in query structures. But, at least for certain proteins, recent advances in single-particle cryo-EM technology have made it easier to produce structures with close to atomic resolution. In fact, a number of software programmes have been developed or modified recently that allow the structure prediction and modelling process to include empirically obtained constraints. These applications include SAXTER, CHESHIRE, Rosetta, and CS23D. Proteolytic cleavage of the extramembrane regions of the proteins or their identification using antibodies, point mutations of various residues of the protein, and chemical crosslinking are examples of biochemical methods that provide experimental data that can be used to guide prediction tools. The fact that such approaches for structure prediction are used on samples with a lot of molecules presents a significant difficulty. As a consequence, they could provide information corresponding to several protein substates, complicating the prediction procedure [1], [2].

Proteins in peripheral membranes

The majority of the surface of peripheral membrane proteins, which are connected to either the exoplasmic or cytoplasmic side of the membrane, is in aqueous solution the extracellular matrix or cytoplasm, respectively. As a result, the rules governing their structures are comparable to those of proteins that are water-soluble. In light of this, the emphasis of our discussion of peripheral membrane proteins is on their interaction with the lipid bilayer. Three different processes exist that might lead to this attachment. A binding site in certain membrane proteins is specifically designed geometrically to bind a particular membrane phospholipid. The latter is often negatively charged (PS, PG, PIP2), and the binding site has basic residues that may interact with it well. Other proteins, like MARCKS, have patches of basic amino acids that make the whole area positively charged. With the help of this charge, the protein may indiscriminately bind to regions of the lipid bilayer that include microdomains of negatively charged phospholipids; in eukaryotes, these regions are located on the cytoplasmic side of the membrane. In these instances, the protein is placed at an electrostatically ideal distance of roughly 3 from the phospholipid head group. That is, since there is a layer of water between the protein and lipid membrane, there is a maximum Coulomb attraction between the two charged entities at this distance, which more than makes up for the unfavorable Born repulsion. In Subsection, which examines the interaction between proteins and membrane PIP2, one of the two types of electrostatic binding discussed here is shown using a well-known example.

Some proteins go through post-translational modifications that give them the ability to attach to membranes covalently. N′-myristylation, S-palmitoylation, and S-prenylation are some of the changes. These proteins, which attach to lipid chains from the cytoplasmic side of the membrane, include ras and src, which are important signal transduction proteins. Amphipathic helices are included into the binding. Amphipathic helices seen in certain membrane proteins allow them to partly pierce one of the bilayer leaflets. The hydrophobic face of the helix's nonpolar residues engages with the bilayer's hydrocarbon core in this

manner of protein-membrane interaction, whereas the polar face of the helix's residues interacts with the lipid head groups [3], [4]. Peptides exhibit this kind of binding as well; for instance, antimicrobial peptides may bind to bacterial membranes.

DISCUSSION

The last relationship is the most fascinating. First, both membrane proteins and peptides exhibit it. It also has an impact on the lipid bilayer's structure. As shown in the instance of the signal transduction protein MAR- CKS, none of the three modes of binding seem to be adequate on their own. The first two kinds of binding, as well as perhaps the third, are used by this protein to maintain its bond with the membrane. The electrostatic component of the interaction is neutralized by phosphorylation or binding to calmodulin, causing the release of the protein into the cytoplasm.Because the lipid bilayer is a chemically complex medium, it affects its resident proteins in a variety of ways. However, these effects, which come about as a consequence of lipid and protein physicochemical interactions, may be divided into two categories. The first kind derives from the bulk physical characteristics of the bilayer, which may be thought of as a complicated solvent. Studies have shown that general lipid bilayer characteristics including topology, degree of order, viscosity, hydrophobic thickness, curvature, degree of acyl chain packing, free volume, stiffness, and more have an impact on membrane proteins. Specific interactions between particular lipid molecules and proteins drive the second sort of bilayer impact. In the last decades, the two sorts of impacts have been investigated using various methodologies.

The following subsections examine the key results. To prevent the unfavorable exposure of their backbone polar groups to the hydrophobic core of the lipid bilayer, transmembrane portions of membrane proteins often develop an ordered secondary structure. When amphipathic peptides or protein fragments attach to the bilayer interface, a related event also takes place. The amphipathic peptides or segments, then, fold and form an orderly helical shape. However, in this instance, the folding is partially motivated by the need to maintain the peptide or segment's amphipathic character, which enables it to interface favorably with the lipid bilayer. So, the lipid bilayer's fundamental polar/nonpolar topology may have an impact on the structure of the proteins there.Different membrane proteins have different preferences for areas of the membrane with different levels of lipid organisation. Numerous integral proteins have been shown to favourld-phase areas, according to studies, whereas others have been found to prefer lo-phase regions. The various preferences cause lateral protein segregation within the membrane and the development of microdomains with particular lipid and protein compositions. When proteins are covalently connected to the bilayer through a fatty acid or other forms of hydrocarbon chains, the preference in some situations is the consequence of direct protein-lipid interactions.

Examples of these proteins include the case of the protein tyrosine kinase. The choice in these circumstances is determined by how saturated the acyl chain is. For instance, since the acyl chain of GPI (glycophosphatidylinositol) is saturated, proteins that are covalently bound to it prefer or- desired areas of the bilayer. This is likewise true for myristoylated or palmitoylatedsrc and other kinases in that family. In contrast, disordered areas of the bilayer are preferred by GTPases of the ras family, which are connected to the unsaturated prenyl chain [5], [6].By affecting the thickness of the membran, the degree of order in the bilayer indirectly influences proteins as well. Because they are longer than unsaturated acyl chains, saturated acyl chains tend to form more ordered bilayers. The thickness of the hydrocarbon area has a greater impact on the proteins than the total thickness of the bilayer. As previously mentioned, nonpolar residues often span a length that nearly corresponds to the hydrophobic

thickness of the lipid bilayer in transmembrane regions. Transmembrane segments may sometimes mismatch the bilayer core in terms of hydrophobicity.

Even so, the protein could also go through some modifications to reduce the mismatch. When the transmembrane segment's nonpolar stretch is longer than the thickness of the bilayer's core, positive hydrophobic mismatch results. As a result, the transmembrane segment's nonpolar residues are unable to properly connect with the bilayer's central region. The transmembrane segment may tilt in relation to the vertical axis of the bilayer to get around this issue. It is far less energy-efficient to have the opposite circumstance, or a negative hydrophobic mismatch, where the thickness of the bilayer's core is higher than the length of the nonpolar stretch in the trans-membrane section. The issue is the protein's polar amino acid residues being partitioned to the nonpolar bilayer core, which, as we've seen before, has the potential to seriously destabilise the system. There are many strategies to reduce detrimental mismatch or its effects:

The protein may travel along the lipid bilayer's plane in the direction of areas with a lesser degree of order and a thinner hydrophobic layer. The lipid bilayer becomes microdomain-like as a result. One example of a cellular activity that is heavily reliant on the existence of such microdomains is signal transduction. In particular, a concentration of certain protein and lipid components in a constrained area of the membrane is necessary for signal transmission. PIP2-dependent signal transduction pathways provide as an illustration of such a necessity. The trafficking of proteins within the cell is another biological function that microdomains have an impact on. The membrane of a certain cellular compartment must hydrophobically match the protein that is being supplied to that compartment. Large enough membrane proteins may go through conformational modifications to lessen hydrophobic mismatch. These modifications often include the screwing or sliding of domains or helices, which are structural protein components. Although they resolve the mismatch issue, such structural modifications may result in additional issues since they decrease the protein's activity. When the thickness of the sarcoplasmic reticulum membranes of muscle cells differs from that of the plasma membrane, for instance, the enzyme Ca2+-ATPase's activity is reduced. The effects of this decrease are not necessarily bad since it offers a way to control the enzyme's activity under various circumstances that alter the SR membrane's thickness.

The system may lessen the amount of unfavorable contacts by substituting some of them with beneficial protein-protein interactions when two distinct transmembrane segments interact unfavorably with the lipid bilayer, such as when there is a hydrophobic mismatch. In other words, the two transmembrane segments are linked together. In certain circumstances, activation may be aided by association. The antibacterial protein gramicidin exhibits this. In this instance, the connection enables the dimerization of two short segments into a transmembrane segment long enough to bridge the membrane and serve as an ion channel. This might represent an evolutionary process wherein particular proteins and peptides are activated by protein-membrane mismatches.The protein's activity may alter as a result of all the actions on the membrane outlined above. Numerous proteins, including Na/K-ATPase, cytochrome c oxidase, Ca2+-ATPase, melibiose permease, and diacylglycerol kinase, have previously been shown in studies to demonstrate this [7], [8].

Almost all proteins must dynamically switch between several conformations in order to function. Water molecules surround the majority of globular proteins and may quickly adapt to whatever new shape the protein may take on. Although it involves friction from water-water interactions, such adjustment facilitates structural changes in the protein. In contrast, lipids that surround membrane proteins have a restricted ability to reorganise in response to conformational changes because they are less mobile than water. Confor- mational

modifications of integral membrane proteins therefore entail considerable friction with surrounding lipids, especially with the acyl chains. Such friction resists the change on the one hand, but by preventing post-change vibrations, it may also make the shift smoother. Prokaryotes, which are subject to shifting environmental circumstances, demonstrate the significance of maintaining continuous viscosity in biological membranes. The lipid composition of the bilayer, which may alter in response to changes in the environment, is what maintains the constant membrane viscosity that is a homeostatic trait in these organisms. For instance, a bacteria will react to an increase in the ambient temperature by raising the proportion of long-saturated phospholipids, which counteract the increase in membrane dynamics brought on by heat.It has been discovered that the activity of integral proteins in certain locations is influenced by the lipid bilayer's ability to develop positive or negative curvature in such places. The mechanisms that cause negative curvature in the bilayer have received a lot of attention in studies looking at this problem.

This is because, under extreme circumstances, such activities may result in the bilayer losing its planarity and producing, for example, an inverted hexagonal phase. The findings show a complicated scenario in which enrichment of PE, a lipid known to cause negative curvature, increases certain proteins' activity while lowering that of others. Prokaryotes that are subjected to environmental changes experience significant water entry into the cell as a result of a rapid reduction in the solute concentration of the external environment. Aside from the decrease in intracellular solute concentration that results, this is a dangerous scenario because the stretching of the plasma membrane might cause it to rupture owing to the intense pressure being exerted to it from inside. The latter issue is resolved in bacteria by specific membrane proteins that serve as mechanosensitive sensors. These proteins enable an uncommon occurrence to take place when the membrane is stretched: a significant outflow of cellular solutes. Water efflux results from this, which fixes the issue. It is interesting to note that these proteins only react to membrane stretching when the increase in surface area approaches 4%, which also happens to be the threshold for membrane rupture.According to the preceding detailed explanation, lipids of various types are present in biological membranes. Regardless of kind, these lipids may be categorised into three main groups based on how they interact with proteins. All lipids that are not specifically interacting with proteins are considered bulk lipids, and their diffusion is consequently influenced by their interactions with nearby lipids.

All lipids that surround a protein in a contact layer yet continually migrate between the layer and the bulk are referred to as annular lipids. This movement is roughly an order of magnitude slower than the diffusion rate of bulk lipids, according to EPR studies. The protein-lipid interface is sealed by this contact layer, which also helps orient integral proteins vertically in the lipid bilayer. As is evident from the structure of the V-Type Na+-ATPase, annular lipids are not only located at the protein's edge but may also be found in significant amounts inside multimeric proteins.Lipid molecules that engage strongly with the protein are known as bound lipids. These lipids' movements are severely inhibited by the contact, which enables them to crystallize alongside the protein. Therefore, they are visible in X-ray diffraction-produced structures. Bound lipids may be either buried inside the protein or in clefts on the protein surface. 'Integral' refers to lipids that are firmly embedded into a protein or complex. Evolutionarily conserved residues are often found in the lipid-binding site, such as the cholesterol-binding motif in G-proteins. The attached lipid molecule has a propensity to choose a shape that offers the greatest possible interaction with the protein, which often results in lipid molecule deformation. It is known that such distortion occurs even when the lipid is saturated, and it may cause the lipid to translate inward, towards the bilayer core. The lipid distortion might even occur in situations where the polar head group of the lipid is

below where the phosphoresce groups of the bilayer are usually located, or where its acyl chains bend and wrap around protein -helices [9], [10].

There are certain protein-lipid interactions that are usually noncovalent and include the polar head group of the lipid as well as nonpolar and van der Waals interactions involving the acyl chains. As previously indicated, interactions between basic protein residues and acidic head groups on the electronegative side of the membrane (the mitochondrial matrix, the stroma of chloroplasts, or the cytoplasmic side of the plasma membrane) are some of the most advantageous interactions involving lipid head groups. While there are many interactions involving lipids' alcoholic groups, those involving the phosphodiester group, which is present in all phospholipids, are mediated by a number of two-residue combinations of basic and polar-neutral residues. These include KT, KW, KY, RS, RW, RY, RN, HS, HW, and HY.

The following residues, listed in order of frequency, are involved in the main interaction with the phosphodiester group:

Trp, Ser, Asn> Arg > Lys > Tyr > His

Thr and Gln have additional positive interactions. Recurring residues, the most of which are polar-neutral and very few of which are basic, are also involved in the interactions at the opposite, less electronegative side of the bilayer. Lys and Arg do not interact at all with PC, the predominant phospholipid on that side of the membrane, most likely due to the presence of the positively charged trimethylamino group. They do, however, include His and Ser in conjunction or His, Ser, and Thr individually. Cardiolipin (CL), which has two phosphodiester groups, is another unique example. These interact with a motif made up of three residues, the first two of which are basic and the third of which is polar-neutral.

The aromatic residues Tyr and Trp, which are found in transmembrane segments, often occur close to the polar-nonpolar boundary, where they serve as anchors to prevent the segment from sliding in or out. The complicated interactions between these residues and nearby phospholipids are largely responsible for the anchoring effect. Two characteristics of Tyr/Trp are crucial to these interactions: The form of a ring. This enables the residue to bind strongly with the acyl chains of the phospholipid via van der Waals and non-polar contacts. These interactions rely on the geometry because of the ring's planarity.This characteristic is a consequence of the residue's aromaticity as well as its chemical makeup (OH group in Tyr and NH in Trp). It enables electrostatic interactions between the residue and polar head groups of surrounding lipids. Since the aromatic residue is situated where it is, the interaction predominantly consists of hydrogen bonds with the carbonyl groups of phospholipids, even though a recent NMR research indicated that they are not crucial for membrane anchoring.

Integral membrane proteins' stability, folding, assembly, and activity may be impacted by certain lipid molecules in the bilayer. Most of the time, evolutionary forces rendered these effects advantageous, and it is already known that certain proteins can only function when specific lipids are present around them. For instance, cardiolipin, which is present in large amounts in the inner mitochondrial membrane, is necessary for the action of the metabolic proteins NADH dehydrogenase, ADP/ATP carriers, cytochrome c oxidase, ATP synthase, and cytochrome bc1. The lipid molecule is regarded as a cofactor in these circumstances. In certain instances, the molecular foundation for the functional reliance may be determined from the three-dimensional structure of the protein and attached lipid. The light-harvesting complex of the photosystem II (LHC-II) in plants exhibits this behaviour. The development of the LHC, and hence its activity, relies on PG. The phospholipid molecule is found near the subunit interface of the LHC structure, where one of its acyl chains is positioned within the trimer. As a result, PG helps to keep LHC's oligomeric structure stable.

Additionally, PG interacts with other lipids close to the LHC, such as chlorophyll and carotenoids, which aids in stabilising the LHC structure's loosely packed and moderately hydrophobic -helices. The dimeric structure of cytochrome c oxidase appears to be stabilised by two cardiolipin molecules that face the mitochondrial intermembrane side of the protein, whereas two additional cardiolipin molecules that face the matrix side appear to act as proton traps, facilitating proton translocation along the protein's surface. Additional examples of lipid molecules with identifiable functional activities are provided in. However, it is not always simple to infer the molecular cause of the protein's lipid dependence from its three-dimensional structure. The tetrameric KcsA channel, whose activity is dependent on PG, is an example of this. Each of the two Arg residues that the phospholipid molecule binds to comes from a separate subunit. This shows that PG's function is to support the protein's quaternary structure's stabilisation. However, according to other investigations, KcsA may settle with any negatively charged phospholipid and is not especially dependent on PG. Therefore, rather than forming a particular connection, it seems that PG's function in this system is to lessen the electrostatic repulsion between positive charges at the subunit interface [11], [12].

Peripheral membrane proteins interact with lipid bilayers by a variety of ways, such as particular geometric binding sites, post-translational changes, and the presence of amphipathic helices. several proteins display a variety of binding strategies, often combining several techniques to maintain their interaction with the membrane. Lipid bilayers are essential for determining how membrane protein structure and function are shaped. They influence proteins both via particular interactions between lipids and proteins as well as through the general physical characteristics of the bilayer, such as thickness, order, and curvature. Protein conformational changes, lateral protein segregation, and the creation of microdomains may all be caused by hydrophobic mismatch, which results from variations in hydrophobic thickness between the protein and bilayer. Proteins must constantly adapt to changes in their environment, and because of the fluidity of their surroundings, water-soluble proteins have an advantage in this respect. Contrarily, while undergoing conformational changes, integral membrane proteins come into contact with the surrounding lipids, which may both fight and ease the shift. In response to changing environmental circumstances, prokaryotes have developed ways to maintain membrane viscosity, emphasising the significance of lipid composition. The activity of integral proteins may be influenced by the lipid bilayer's capacity to produce positive or negative curvature. Certain lipids are known to serve as cofactors for certain proteins, enabling them to function. These lipid-protein interactions could be crucial for activity, stability, folding, and assembly.

CONCLUSION

In conclusion, there have been considerable improvements in the area of protein structure prediction, especially with regard to membrane proteins and their interactions with lipid bilayers. Modern prediction algorithms have developed into hybrid techniques that include many methodology, from physicochemical factors to the statistical propensities of amino acids. Rosetta and I-TASSER, two well-known methods, have shown to be effective in correctly predicting protein structures and are still useful tools in structural biology. Additionally, to improve prediction accuracy, subsequent advancements have used experimental data from cryo-EM, SAXS, NMR, CD, and FRET. These methods raise the possibility of discovering natural protein structures by offering useful spatial restrictions that direct prediction algorithms. Particularly Cryo-EM has made great progress in creating structures with almost atomic resolution, bringing up new opportunities for structure prediction. In conclusion, the complicated interaction between lipid bilayers and membrane proteins is a fascinating field of research with broad ramifications for structural biology and

cellular function. The development of therapeutic treatments that target membrane proteins and the advancement of our understanding of the biology of membrane proteins both depend on our ability to comprehend these interactions.

REFERENCES:

[1] H. Turabieh, A Hybrid ANN-GWO Algorithm for Prediction of Heart Disease, *Am. J. Oper. Res.*, 2016.

[2] W. Musa and S. Tansa, Ham: A hybrid algorithm prediction based model for electricity demand, *J. Theor. Appl. Inf. Technol.*, 2016.

[3] V. Verma, S. Bhardwaj, and H. Singh, A Hybrid K-Mean Clustering Algorithm for Prediction Analysis, *Indian J. Sci. Technol.*, 2016.

[4] L. Cornejo-Bueno, J. C. Nieto-Borge, P. García-Díaz, G. Rodríguez, and S. Salcedo-Sanz, Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm - Extreme Learning Machine approach, *Renew. Energy*, 2016.

[5] M. Qiu, Y. Song, and F. Akagi, Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market, *Chaos, Solitons and Fractals*, 2016.

[6] A. Swanepoel, S. Barnard, F. Recknagel, and H. Cao, Evaluation of models generated via hybrid evolutionary algorithms for the prediction of Microcystis concentrations in the Vaal Dam, South Africa, *Water SA*, 2016.

[7] T. Wang, Y. P. P. Chen, P. J. Bowman, M. E. Goddard, and B. J. Hayes, A hybrid expectation maximisation and MCMC sampling algorithm to implement Bayesian mixture model based genomic prediction and QTL mapping, *BMC Genomics*, 2016.

[8] M. Inthachot, V. Boonjing, and S. Intakosum, Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend, *Computational Intelligence and Neuroscience*. 2016.

[9] J. Ding, X. Li, and H. Hu, TarPmiR: A new approach for microRNA target site prediction, *Bioinformatics*, 2016.

[10] M. Mazari and D. D. Rodriguez, Prediction of pavement roughness using a hybrid gene expression programming-neural network technique, *J. Traffic Transp. Eng. (English Ed.*, 2016.

[11] M. A. Ahmadi, S. Zendehboudi, M. B. Dusseault, and I. Chatzis, Evolving simple-to-use method to determine water–oil relative permeability in petroleum reservoirs, *Petroleum*, 2016.

[12] S. Barak and S. S. Sadegh, Forecasting energy consumption using ensemble ARIMA-ANFIS hybrid algorithm, *Int. J. Electr. Power Energy Syst.*, 2016.

CHAPTER 10

EXPLORING THE INTRICATE WORLD OF PIP2: KEY ROLES IN CELLULAR SIGNALING AND MEMBRANE PROTEIN REGULATION

Dr. Jyoti Trivedi, Assistant Professor
Department of Biochemistry, TMMC&RC, TeerthankerMahaveer University, Moradabad, Uttar Pradesh India,
Email Id- jyoti.trivedi01@gmail.com

ABSTRACT:

The phospholipid phosphatidylinositol 4,5-bisphosphate (PIP2), which is mostly found on the cytoplasmic side of the plasma membrane, serves as a focal point in the complex interaction between protein and lipid in cellular membranes. PIP2 is not widely distributed, yet it has a significant impact on key cellular processes such vesicle trafficking, phagocytosis, exocytosis, and endocytosis. The crucial function of PIP2 as a substrate for the cytoplasmic enzyme phospholipase C (PLC), which converts PIP2 into the crucial second messenger's diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3) in signal transduction cascades, is at the core of its relevance. These messengers work together to activate protein kinase C (PKC), which phosphorylates a variety of cellular targets and triggers a variety of biological reactions. PIP2 has a significant impact on membrane proteins beyond its function as a source of second messengers, especially ion channels and transporters, which rely on PIP2 for activation and control. This reliance offers two important benefits. First off, these proteins function as cellular signalling hubs localised inside the plasma membrane, promoting connections and signal transmission since they remain inactive in the absence of PIP2. Second, PIP2-dependent proteins may respond to external stimuli by having their amounts of PIP2 altered by outside signals, typically through activating PLC. PIP2 interacts with proteins through non-specific electrostatic interactions and specific binding, respectively. Through geometrically and chemically appropriate binding sites, such as the pleckstrin homology (PH) domain and other domains including FYVE, PX, and ENTH, specific binding takes place. Non-specific binding is based on disordered protein segments with a high concentration of basic residues interacting with PIP2 clusters that are electrostatically charged within the membrane. PIP2-dependent signal transduction activities are regulated by competition between PIP2 binding and other interactions, such as calmodulin (Ca2+/CaM) binding to proteins like MARCKS.

KEYWORDS:

Cytoplasmic, Membrane Protein, Polymerization, Signal Transduction.

INTRODUCTION

PIP2, a phospholipid found in minute amounts on the cytoplasmic side of the plasma membrane, is one of the most studied instances of particular protein-lipid interactions. PIP2 is a rare protein, yet it plays a critical role in important cellular functions such vesicle trafficking, phagocytosis, exocytosis, and endocytosis. The discovery that this molecule serves as a substrate for the cytoplasmic enzyme phospholipase C (PLC) sparked interest in it. PLC divides PIP2 into diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP3), two crucial second messengers in signal transduction pathways. They work together to activate the enzyme protein kinase C (PKC), which phosphorylates a variety of cell targets and triggers significant biological reactions. Today, we know that PIP2 is the actual source of three second messengers and that these messengers are primarily responsible for PIP2's effect on the aforementioned processes. PIP2 seems to also have a direct impact on membrane

proteins, particularly ion channels and transporters. What benefit does these proteins' PIP2 dependence provide? There are at least two that spring to mind, both involving regulation. The first is straightforward: these proteins are inactive unless they are bound to the membrane because of their need on PIP2. The plasma membrane does in fact often serve as a meeting place for signalling proteins. This facilitates the interaction and signal propagation between the many elements of the route. The second benefit of proteins being dependent on PIP2 is indirect; since external signals may influence PIP2 levels (by activating PLC), dependent proteins are also susceptible to those external signals [1], [2].

This is accomplished through a bind- ing site that is geometrically and chemically suitable and recognises PIP2 utilising basic and other residues. There are multiple known binding sites of this kind; the pleckstrin homology (PH) domain is the best-known. Over 100 residues make up this domain, which has previously been identified in roughly 250 human proteins. It is made up of a sandwich of seven strands and one helix. Basic residues that form salt bridges with the phosphate groups of the lipid are used to bind to PIP2, as are hydrogen bonds to other forming residues. Different PH domains bind various phosphoinositides (PI(3,4)P2, PI(3,4,5)P3)), where the specificity is determined by the spatial arrangement of the residues in the bind-ing site. Other domains including FYVE, PX, and ENTH have additional PIP2 binding sites.This is accomplished by using disorganised protein segments that cling to electrostatically charged PIP2 clusters that are located in certain membrane areas in an unspecific manner. A number of basic residues in close proximity to one another in the sequence, which provide their corresponding segments a general positive charge, enable adhesion. The chaotic structure of such segments is also owing to the close closeness of the basic residues to one another. The 331-residue protein MARCKS, which is generally acidic except for a 24-residue stretch in which 13 residues are basic, is one often cited example of this sort of binding. This section may bind PIP2 molecules and group them into a microdomain because of its positively charged portion.

Additionally, PIP2 connections with other basic cytoplasmic proteins are in competition with MARCKS binding, and it seems that this serves as a regulating mechanism for PIP2-dependent signal transduction activities. In fact, the increase in cytoplasmic Ca2+ levels caused by IP3 results in the binding of calmodulin (Ca2+/CaM) to MARCKS, which causes the protein to lose its positive electrostatic potential and leave the PIP2-rich area of the membrane. As a result, additional cytoplasmic signalling proteins may access PIP2, enabling the signal to spread. It should be noted that PKC-induced phosphorylation of the protein also affects the electric field on MARCKS. As previously mentioned, the identical signal route that stimulates Ca2+/CaM also activates PKC.The two types of PIP2 binding discussed above vary from one another in terms of binding specificity and binding site structural requirements. Both kinds occur, although it seems that specific binding predominates. Even though several research have provided exciting new information, a mechanistic characterization of PIP2 reliance is often not simple. For instance, it has been proposed that PIP2 contributes to the gating mechanism of tetrameric K+ channel-like proteins by exerting an electrostatic stress on certain protein regions that causes the regions to move and open the channel. The description of further models that have been proposed is in.

Membrane proteins' effects on the characteristics of lipid bilayers

Proteins and other hydrophobic intrusions may disturb the lipid arrangement within a membrane. The lipids in the bilayer may respond in a variety of ways as a result of the disturbance, which operates on numerous levels. We list the primary impacts in the subsections that follow. Membrane lipids are dynamic, moving in a variety of ways, from little vibrations to significant lateral diffusion or flipping between bilayer leaflets, for

example. Simply having a hydrophobic inclusion causes the lipids flanking the stiff inclusion to move less quickly, which results in a loss of entropy. The introduction of even a single helix into the lipid bilayer decreases the entropy, with a corresponding free energy penalty of +2 kcal/mol, according to statistical-thermodynamic models.As we have already shown, the length of a specific transmembrane segment may differ from the thickness of the hydrocarbon portion of the lipid bilayer, leading to a hydrophobic mismatch. This mismatch may lead to alterations in the membrane lipids as well as the structure and/or orientation of the transmembrane segment. The lipid bi-layer's principal reaction to this mismatch is deformation, which involves stretching or compressing the acyl chains surrounding the transmembrane region, depending on whether the mismatch is positive or negative. At the protein-lipid contact, the deformation causes a local alteration in the lipid bilayer's curvature. This is made feasible by the lipid bilayer's flexible nature, which has a compressibility of 109 to 108 N/m2.

Recent measurements have shown that the influence of protein-induced deformation on membrane thickness is five times more important than the influence of cholesterol, which was previously thought to be the main factor affecting membrane thickness in higher eukaryotes. Certain signalling proteins seem to be concentrated in one area of the membrane through lipid-induced changes in curvature. For a 4-fold decrease in membrane thickness, the energy cost of deformation has been calculated at 0.4 kcal/mol. The segment may tilt at least 10° away from the membrane's normal to boost its entropy of procession, according to computational studies, even when the length of the transmembrane segment is equivalent to or less than the hydrophobic thickness of the membrane. The membrane must bend inward to do this, but the entropy gain more than offsets the deformation cost. In actuality, the entropy of the lipid chains is linked to the deformation penalty. Another instance of two entropy-based variables balancing one another may be seen in this calculation of the ideal tilt angle of the transmembrane segment in the lipid bilayer [3], [4].It has been discovered that protein shape influences membrane curvature. Therefore, integral proteins with asymmetric profilesthat is, proteins with extracellular regions that are narrower or wider than their intracellular counterpartscreate either a positive or negative curvature, especially when they oligomerize or aggregate. In addition to these basic characteristics of proteins, there are known instances when the actions of particular proteins have pronounced impacts on membrane curvature:

DISCUSSION

Polymerization of actin. variations in the plasma membrane, including variations in curvature, are closely related to the capacity of the cytoskeletal protein actin to polymerize in response to certain signals. In particular, the membrane is subjected to mechanical pressure during polymerization to the point of curvature. Numerous cellular activities, including the development of pseudopodia, phagocytic cups, endocytic invaginations, and even axonal growth cones produced during the development of the neural synapse, depend on this action.Generation of vesicles by coat proteins. Coat proteins that are peripherally linked to the membrane, such as clathrin, caveolin, COPI and COPII, are responsible for forming transport vesicles inside of cells. They exert mechanical pressure on the bilayer as a result of their activity, and this pressure progressively makes the bilayer more curved until the transport vesicle is created. It was originally believed that the pressure put on the membrane was caused by these proteins polymerizing. Recent research has shown that instead of making direct connections with the membrane, clathrin, COPI, and COPII utilise other proteins to exert mechanical pressure. The latter often include amphipathic -helices that are wedged between the lipid head groups and partly submerged within the bilayer with their polar side

towards the hydrophobic core. Positive curvature in the bilayer is produced when the lipids in contact with such -helices alter conformation to account for the membrane deformation. What function do clathrin, COPI, and COPII serve then? The pressure-producing proteins seem to be contained by these proteins in a specific area of the membrane, and when the vesicle has formed, they polymerize to create a framework around it. As an illustration of a dramatic but physiologically significant impact of proteins on membrane curvature, consider the formation of transport vesicles.

Protein-Coupled G Proteins

Despite the small number of membrane proteins with empirically defined structures, substantial study has produced a wealth of knowledge about the structural and sequence-related requirements for the function of these proteins. The G protein-coupled receptors (GPCRs), which are important in many physiological processes and diseases associated with them, are discussed in this subsection. The activity of GPCRs entails structural complexity, such as changing conformation and binding numerous partners.Highly complicated signal transduction systems are used by cells in multicellular animals as well as single cells to interact with their surroundings. The cell surface is only one of many parts in these systems, which also include internal proteins and tiny chemicals. The membrane-bound receptor is the first part of the cell to respond to an incoming message. Numerous receptors are present in cells, and they react to various messengers. These include internal ones that are absorbed from the environment of the organism. Cell-surface receptors that have been triggered may transmit the signal to a variety of enzymes, tiny compounds, or elemental ions within the cell. Since they transmit the information while amplifying it by acting on several cellular targets, the majority of these species perform the job of transducers-amplifiers [5], [6]. Others function as end effectors, which are proteins that, upon activation, produce an outcome. This outcome may take a variety of forms, from relatively minor cellular reactions like the creation and/or release of a chemical component to more significant reactions like cellular division and even suicide. Membrane-bound receptors may be categorized based on how they react when a ligand is bound:

1. Atomic channels.
2. The tyrosine kinases.
3. Threonine and serine kinases.
4. The guanylate cyclase's.
5. Defining cytokine receptors according to ligand type.
6. GPCRs are G protein-coupled receptors.

The biggest and most prevalent family of membrane receptors is without a doubt the GPCRs. They are extensively distributed throughout the majority of living form; in vertebrates, they make up 1% to 5% of the whole genome, and in the human genome, more than 800 genes are responsible for encoding them. The capacity of GPCRs to react to a wide range of external messengers, including as proteins, peptides, tiny chemical compounds, elemental ions, and even photons of light, is another amazing characteristic of these receptors. These messengers might be pheromones, neurotransmitters, hormones, local mediators, neurotransmitters, neurotransmitters, or environmental variables. Since they are either inactive or hyperactive in many illnesses and pathological syndromes, GPCRs are implicated in a number of physiological processes. Hypertension, congestive heart failure, stroke, cancer, thyroid dysfunction, congenital intestinal obstruction, improper bone growth, night blindness, and newborn hyperparathyroidism are some of these illnesses. GPCRs are undoubtedly attractive pharmacological targets; in fact, it is thought that between 30% and 50% of clinically prescribed medications work by attaching to GPCRs and altering their activity.

GPCRs, as their name suggests, typically use big GDP/GTP-binding proteins known as G-proteins to transmit signals into cells. Once activated, G-proteins may go on to activate several effector proteins in what seems to be a cascade. In other words, each molecule activates a certain number of effector proteins, and as the signal travels farther down the route, the number of active proteins increases. A huge range of effector proteins, including as enzymes, ion channels, proteins linked to transport vesicles, and others, are often activated or inhibited as a consequence of GPCR signalling. The messenger molecule, GPCR, and G-protein that are activated determine the kinds of proteins that are activated in a specific GPCR pathway. For instance, signalling through certain GPCR and G-protein types activates adenylyl cyclase (AC) and results in the production of the second messenger cAMP in the cAMP-PKA pathway. The latter causes a number of cytoplasmic proteins to be phosphorylated by protein kinase A (PKA), which is then activated. Some of the proteins are made active by phosphorylation, whereas others are made inactive. Instead of adenylyl cyclase, the G-protein activates phospholipase C (PLC) via a different universal signalling route. The second messenger's diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3) are produced when PLC hydrolyzes the membrane lipid phosphatidylinositol 4,5-bisphosphate. The combined action of both messengers activates protein kinase C (PKC), which phosphorylates different proteins than its PKA counterpart and has a distinct effect, via a significant but transient Ca2+ influx into the cytoplasm. GPCR signalling is inherently complicated since it includes several elements. The following characteristics add to this complexity:

1. Several G-proteins, as well as several non-G proteins, may be activated by a single GPCR. For instance, it is well known that the MAP kinase pathway is activated by the 2-adrenergic receptor.
2. The majority of GPCRs typically exhibit some level of baseline activity. In other words, even when they are not binding their activating ligands (agonists), they are still somewhat active.
3. A single GPCR may react to several ligands, each of which will have a distinct effect:

Full agonists increase receptor activity to its maximum potential by maintaining an active condition. By maintaining an inactive conformation, reverse agonists lower the receptor's basal activity. Due to their partial affinities for both active and inactive conformations, partial agonists generate partial activity. Antagonists stop other ligands from attaching to and activating the GPCR. These findings are consistent with two of the most widely used theories of protein dynamics.

The 'pre-existing equilibrium' hypothesis states that even in the absence of a binding agonist, the active conformation of the protein is sampled enough to produce some level of baseline activity. The second is the so-called conformational selection, which postulates that every ligand binds to and stabilizes a unique conformation of the GPCR, each of which has a unique intrinsic activity. The same ligand may have conflicting effects on two distinct routes that are controlled by the same GPCR by stabilising a conformation of the GPCR that is compatible with just one of the pathways since a single GPCR may modulate several pathways. Drugs that operate on GPCRs are often designed using this phenomenon. GPCR oligomerization localization to certain membrane compartments, or membrane lipid composition may all have an impact on the activity of GPCRs[7], [8].

Mechanisms and control of G-proteins

As previously mentioned, GPCRs use G-proteins to transmit outside signals into the cell. Each of these proteins is made up of three subunits called G, G, and G (or G1, G2, and G3,

respectively), which may take on many forms. To far, 23 genes have been identified as encoding G, G5, and G12, respectively. G-proteins have been divided into four families based on the G types, with each family having a propensity to either activate or inhibit a particular target:

1. Gs stimulates the cAMP-PKA-adenylyl cyclase signalling cascade. The cholera toxin also over activates this family by covalent alteration.
2. GIO covalently inhibited by the pertussis toxin; inhibits adenylyl cyclase and stimulates c-Src tyrosine kinases.
3. Gq/11 initiates the IP3-PKC signalling cascade in the PLC.
4. G12/G13 - Rho is activated as a result.
5. Cyclic GMP (cGMP) phosphodiesterase (transducin) is activated by G-transducin in the retina.

The G-protein's three subunits, G, G, and G, are securely connected to one another in the relaxed state, and G binds the guanine nucleotide GDP. A palmitoyl or myristoyl chain coupled to the N-terminus of the G-protein and an isoprenyl chain tied to the C-terminal CAAX motif of G are the two covalently linked lipid chains that hold the G-protein to the membrane (not illustrated). It is not necessary for G to be covalently bonded to the membrane since G is strongly connected to G through nonpolar interactions. The resting G-protein may wander in the membrane or bind to an active or inactive GPCR. It transforms from a resting state to an active one, however, when it binds to an active GPCR. Specifically, the G-protein's subunit (G) undergoes a conformational shift in response to the active, agonist-induced GPCR conformation, which results in the exchange of GDP for GTP. Thus, GPCRs' primary function is to function as GDP/GTP exchange factors (GEFs). When GTP binds to G, it causes the latter to break free from the G complex and attach to an effector protein. The effector protein is activated by this binding. The G complex contains its own effectors, such as PLC, which may be distinct from or similar to those of G. G has GTPase activity, which causes it to hydrolyze its bound GTP into GDP over a brief period of time. As a result, G assumes its initial shape and may reassociate with the subunits.

The system is prepared for a new activation cycle after being returned to its resting condition. It's interesting to note that despite being active, G's intrinsic GTPase activity is too sluggish for cellular needs. In order to increase GTP hydrolysis to the needed rate, G-proteins must work with proteins known as regulators of G protein signalling (RGSs). The GPCR cycle also contains regulatory stages that guarantee the GPCR's internalization and inactivation. Below is a discussion of the steps.The structures of GPCRs have also been delayed by technological issues, as has been the case with other membrane-bound proteins. However, the challenges have successfully been overcome in recent years, and a variety of GPCR architectures have evolved, albeit the majority of them belong to class. Brian Kobilka's group has solved a lot of these structures. For their research on GPCRs, which began with the discovery of the beta-adrenergic receptor and the identification of the gene encoding it, Kobilka and Robert Lefhowitz were awarded the 2012 Nobel Prize in Chemistry. 220 experimentally determined GPCR structures were included in the PDB as of the end of 2017. The receptors for adrenaline, adenosine, dopamine, histamine, acetylcholine, serotonin, glutamate, opioids, sphingosine, chemokines, neurotensin, purine nucleosides, and free fatty acids are among those that correspond to these.

According to these studies, GPCRs have a number of structural similarities, the most notable of which is a transmembrane core made up of seven -helical segments, designated TM1 through TM7, that are arranged. Most seven-transmembrane receptors are GPCRs, however not all of them are. The extracellular N′ comes before the seven helices, while the

intracellular C′ comes after. At both the extracellular and intracellular sides, loops of various lengths referred to as ECL1-3 and ICL1-3, respectively connect the transmembrane segments. Seven is the bare minimum number of transmembrane helices required for maintaining the environment of the ligand, so spectroscopic studies focusing on the refolding of bacteriorhodopsin a seven-transmembrane protein that does not interact with a G protein in lipid vesicles suggest that the common seven-transmembrane core of GPCRs is not accidental. White hypothesized that seven helices may provide ample space for ligands, through relatively minor helix distortions and reorientations, without the need to increase or decrease the number of transmembrane helices based on this finding. Most GPCRs share certain sequence motifs, as well as a disulfide link between a cysteine residue at the extracellular tip of TM3 and another cysteine residue in ECL2, in addition to the number of transmembrane helices and loops. The ligand-binding pocket's entrance is shaped in part by the disulfide bond. However, in many GPCRs, the ligand may also interact with portions of the transmembrane domain (TMD), as we will see below, the extracellular domain (ECD) of GPCRs is the primary ligand-binding site [9], [10].

Despite having a number of structural similarities, GPCRs vary in a number of ways, and this diversity is the result of these variances. While some variations are found in the transmembrane region of the protein, the majority are found in the intra- and extracellular regions of the protein. A GPCR, for instance, may have more helices than the seven that span the membrane. This is true of the -adrenergic receptors, which are located along the extracellular membrane plane and have an eighth helix. Additionally, reports of a subtype of the -opioid receptor with only six transmembrane helices and a GPCR-like sequence with a predicted transmembrane domain with just five helices have not been structurally validated. According to phylogenetic study, GPCRs may be divided into six groups depending on how related they are. Rhodopsin class A or one. The majority of GPCRs belong to this class, and members of this class react to both endogenous and external signals. The following subclasses may be created from this class of proteins:

Small ligands, like as neurotransmitters or even light photons, may activate receptors of subclass I, which have ligand-binding sites that are located in the transmembrane space. Inferred from the clustering of GPCRs in this subclass are the following subgroups: amine, opsin, melatonin, prostaglandin, and MECA (melatonin, EDG, cannabinoid, and adenosine). Receptors for peptides are a part of subclass II. Each member of this subclass has a multi-segmented ligand-binding site that is located on the protein's extracellular side. Glycoprotein hormone receptors are a part of subclass III. Each protein in this subclass has a binding site that is largely located in its extremely large extracellular domain.Despite having no sequence similarity, these proteins are related to class AIII. They react to a variety of big protein and peptide hormones, including secretin, glucagon, glucose-dependent insulinotropic polypeptide (GIP), vasoactive intestinal peptide (VIP), and glucagon-like peptide 1 (GLP-1). Other hormones and factors include calcitonin, pituitary adenylate cyclase activating polypeptide (PACAP), growth hormone-releasing factors (GRF), and parathyroid hormone (PTH). The poison produced by the black widow spider, -latrotoxin, likewise targets class B GPCRs.The glutamate class.

Metabotropic glutamate receptors (mGluRs), Ca2+-sensing receptors (CaSRs) of the parathyroid, kidney, and brain, GABAB receptors, pheromone receptors, sweet and amino acid taste receptors (TAS1R), and odorant receptors in fish are all members of this class. Class D or 4 contains receptors for fungi's mating pheromones. CAMP receptors in Dictyosteliumdiscoideum are classified as Class E or 5. They are engaged in the organism's developmental regulation. Class F or 6. Members of this class participate in a variety of

physiological and cellular processes, including embryonic development. They stimulate important signalling pathways including the Wnt pathway.It should be noted that all classes include contain orphan receptors, which are proteins that resemble recognised receptors structurally but are triggered by undiscovered natural ligands. We shall concentrate on class A GPCRs, for which we have a wealth of structural information, in the subsections that follow. We cover GPCRs from classes B, C, and F. In contrast to receptor tyrosine kinases, which are known to dimerize, GPCRs were first thought to function as single polypeptide chains. However, it is now known that many GPCRs dimerize or oligomerize inside biological membranes. These processes are thought to play a part in GPCR signal transduction and crosstalk across various signalling pathways.

Class C receptors, which are solely functional in their dimeric state, provide the majority of the information on GPCR dimerization. Some of them are homodimers, like mGluRs, whereas others, like the GABAB receptor, are heterodimers. One polypeptide chain binds the ligand in the case of the GABAB receptor, whereas the other binds the G-protein and is also required for bringing the whole receptor to the cell surface. Class A GPCRs have also been shown to dimerize and oligomerize, with a predilection for homodimers. It's interesting to note that in this class, the monomeric form is likewise functional, and dimerization functions primarily as a regulator. In other words, dimerization enables more effective GPCR activity control via cooperativity, crosstalk between several GPCRs within a single heterodimer, etc. The trans-membrane region of the GPCRs often mediates dimerization, although other areas may also be involved. For instance, the dimer in the 1-adrenergic receptor includes two interfaces that involve residues from extracellular and intracellular loops as well as trans-membrane helices. Additionally, membrane lipids could play a role in dimerization, as has been proposed for the mGlu, 2-adrenergic, and -opioid receptors.

CONCLUSION

In conclusion, PIP2 is an important component of cell signalling and membrane biology due to its diverse involvement in cellular processes, complex protein interactions, and control of membrane properties. Phosphatidylinositol 4,5-bisphosphate (PIP2) is a phospholipid that is present in minute quantities on the cytoplasmic side of the plasma membrane and is one of the most well researched instances of protein-lipid interactions. Despite its limited availability, PIP2 is essential for several cellular processes, such as vesicle trafficking, phagocytosis, exocytosis, and endocytosis. Significant interest in the actions of PIP2 was ignited by the finding that it is a substrate for the cytoplasmic enzyme phospholipase C (PLC). Diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3), both of which are crucial second messengers in several signal transduction pathways, are produced as a result of PIP2's enzymatic cleavage by PLC. These second messengers cause protein kinase C (PKC) to phosphorylate proteins, which sets off a series of events that control a variety of cellular activities.

Therefore, PIP2 acts as the source of these second messengers and largely uses them to affect the cellular processes described. As a result of acting as a precursor for essential second messengers and having a direct impact on membrane proteins, PIP2 is a tiny but significant phospholipid that has a variety of functions in cellular processes. Its importance in cell regulation is increased by its capacity to interact with a variety of proteins via both selective and non-specific binding methods. With vast ramifications for cell biology and signal transduction pathways, research into the complexity of PIP2's relationships and activities are still ongoing.

REFERENCES:

[1] K. Lahiri-Dutt, The diverse worlds of coal in India: Energising the nation, energising livelihoods, *Energy Policy*, 2016.

[2] R. Blaustein, New Atlas Illuminates the Intricate World of Soil Biodiversity, *BioScience*. 2016.

[3] P. Whitridge, Fractal worlds: An archaeology of nested spatial scales, *Arctic*, 2016.

[4] M. Pazzaglia and M. Molinari, The re-embodiment of bodies, tools, and worlds after spinal cord injury: An intricate picture. Reply to comments on 'The embodiment of assistive devices-From wheelchair to exoskeleton'., *Physics of Life Reviews*. 2016.

[5] Rachiyata Mahesh Kumar, Globalization and Human Resource, *Golden Res. Thoughts*, 2016.

[6] H. Latief, Philanthropy and 'muslim citizenship' in post-suharto Indonesia, *Southeast Asian Stud.*, 2016.

[7] E. Nakouzi and O. Steinbock, Self-organization in precipitation reactions far from the equilibrium, *Sci. Adv.*, 2016.

[8] M. Ali, Islam and Local Tradition: A Comparative Perspective of Java and Sulawesi, *JICSA (Journal Islam. Civiliz. Southeast Asia)*, 2016.

[9] J. Uptin, J. Wright, and V. Harwood, Finding education: Stories of how young former refugees constituted strategic identities in order to access school, *Race Ethn. Educ.*, 2016.

[10] S. Lakshmi, M. Harshitha, G. Vaishalia, S. Keerthana, and R. Muthappa, Studies on different methods for removal of phenol in waste water - review, *Int. J. Sci. Eng. Technol. Res.*, 2016.

CHAPTER 11

STRUCTURAL INSIGHTS INTO GPCR ACTIVATION: FROM RHODOPSIN TO CLASS A GPCRS

Ravi Kumar, Assistant Professor
College of Paramedical Sciences, TeerthankerMahaveer University, Moradabad, Uttar Pradesh, India Email Id- forensicravi001@gmail.com

ABSTRACT:

Bovine rhodopsin has served as the main resource for detailed structural understandings into the realm of G Protein-Coupled Receptors (GPCRs) since the year 2000. Rhodopsin, which is found in retinal rod cells, is essential for transforming visual information into neural impulses that can be sent to the brain. It is the perfect model for class A GPCRs because of its abundance and structural stability. Rhodopsin, unlike many GPCRs, retains a single, inactive conformation in complete darkness, making structural research easier using a variety of techniques. In this study, we concentrate on the extracellular, transmembrane, and intracellular domains to examine the structural features of class A GPCRs. Notably, the ligand-binding site and surrounding loops of GPCRs' extracellular domains show remarkable structural variability. We contrast class A GPCR structural variants, highlighting the variety of extracellular conformations. Despite low sequence commonality, GPCRs' ligand-binding pocket and transmembrane region exhibit striking structural similarities. At the intracellular terminus of TM7, the conserved NPxxY motif is a key mediator of conformational changes that occur during GPCR activation. We draw attention to the proline, which causes helical deformation in this motif and is stabilised by electrostatic interactions. Further investigation is done into how lipids like cholesterol interact with GPCRs. GPCR function and possible allosteric regulation are affected by these interactions. With rhodopsin as its focal point, the review explores the structural alterations that GPCRs go through during activation. The retinal-binding pocket, the dissolution of the ionic lock between TM3 and TM7, and the critical function of the NPxxY motif in stabilising the active conformation are among the important modifications. We also go through the structural intricacies of GPCR-G protein interactions revealed by the fully active states of the 2-adrenergic receptor and the muscarinic (M2) receptor.

KEYWORDS:

GPCR Activation, Muscarinic, Protein, Rhodopsin.

INTRODUCTION

Since 2000, the sole source for accurate structural characterization of GPCRs has been inactive bovine rhodopsin. To transform visual information into neural signals that may be sent to the brain (through the optic nerves) for processing, retinal rod cells contain a protein called rhodopsin. Rhodopsin, while having a very specific function, has historically served as a model for the structure of class A GPCRs, mostly due to its extreme abundance and stability; in total darkness, it only adopts one (inactive) conformation. These characteristics made it possible for researchers to thoroughly analyse the protein using a variety of methods and to characterise the structure of rhodopsin far sooner than with other GPCRs. A common model for class A GPCRs has been the catecholamine-responsive -adrenergic receptor. The -adrenergic receptors were the first GPCRs to be sequenced and cloned after rhodopsin, and their three-dimensional structures were also solved relatively early. Like rhodopsin, their prominence is based on historical factors. Adrenergic receptors and many other GPCRs are,

however, far less stable than rhodopsin, necessitating the application of stabilisation techniques in order to crystallise them.

These often include the attachment of these portions to an antibody fragment or to another protein such as T4-lysozyme and the mutation or truncation of unstable regions such as the third intracellular loop, or ICL3. Lower stability is often correlated with greater flexibility, indicating that the GPCR may sample different conformations, including the active one, even when it is agonist-free and inactive. As was already mentioned, most GPCRs exhibit baseline activity even in the absence of an activating stimulus, while rhodopsin does not.As previously mentioned, GPCRs exhibit strong structural homology particularly within each class, which is noteworthy given the minimal sequence identity among group members. For instance, just 21% of the amino acids in rhodopsin and the amino acids in the 2-adrenergic receptor are similar, however the r.m.s.d. between their structures is only 2.3, and in the transmembrane region it is 1.6. However, there are some variations among GPCRs, and this need examination. We summarise the key findings for class A GPCRs in the paragraphs that follow, focusing on each of the three GPCR subtypes (extracellular, trans-membrane, and intracellular) individually [1], [2].

Extracellular (EC) area

As might be predicted, the majority of variations among GPCRs are found in the extracellular domain (ECD), notably in the ligand-binding site and surrounding loops. Since ECL2 is the biggest loop and may take on several conformations, it is especially crucial for differentiating amongst GPCR architectures. ECLs 1 and 3 on the other hand, are substantially shorter and often lack a clear secondary structure. Rhodopsin has a tertiary structure that is compact and stiff, which considerably limits the ability of solvent and other molecules to enter the ligand-binding pocket. This is not unexpected, since light (photons) is the 'ligand' of rhodopsin and does not need a large entry. Second, solvent exposure to the retinal cofactor would cause the bond tying the retinal to the polypeptide chain to undergo hydrolysis. ECL2, which forms a brief -sheet, is the primary structural component preventing the solvent from reaching the retina. Additionally, it stops the transmembrane helices from moving and blocks the major entry to the ligand-binding pocket. These illustrate the adverse scenario, when the EC area is open. In particular, the N′ is completely disordered, ECL3 does not engage with any of the other loops, and ECL2 forms a very short helix.

Adenosine: Disulfide bridges that stabilise it as well as the whole extracellular area of the receptor allow ECL2 to maintain its rigidity in the absence of secondary structure. Stabilisation is further aided by polar and van der Waals interactions involving the three loops. Disulfide bridges may act as structural restraints on EC areas that are involved in ligand binding, according to certain theories. The neurotensin receptor, and other peptide-binding GPCRs ECL2 produces a hairpin shape in the extracellular region, which is more open than that of non-peptide-binding GPCRs.Other lipid-activated GPCRs include the sphingosine 1-phosphate receptor: The N′ (arranged as a helix) and ECL1, which cap the extracellular area, prevent the ligand from entering its binding pocket. It is believed that these GPCRs' very hydrophobic ligands pass through the membrane to reach their binding pockets.

The ligand-binding pocket and transmembrane region classes. Even yet, there are still certain areas here that are fundamentally more similar than others. This is shown by a structural alignment of the several structures, which reveals a core of 97 residues with a C-r.m.s.d. of only 1.3. The other, less structurally related residues are anticipated to be involved in the ligand- and G-protein binding processes that distinguish one GPCR from another. The chemical environment of the NPxxY motif, which is found at the intracellular end of TM7, is

one of the shared structural features of the GPCRs mentioned above. Important conformational changes during GPCR activation are mediated by this region [3], [4]. It's interesting to note that the presence of Pro in the sequence motif causes the helix in all of the aforementioned GPCR structures to be bent in this area. However, hydrogen connections to neighbouring residues or nearby water molecules stabilise the helical deformation electrostatically.

DISCUSSION

The nature of the GPCR-lipid interactions is interestingly shown by the structure of the 2-adrenergic receptor. Between transmembrane helices 2 through 4, the structure features a cholesterol-binding site made up of evolutionarily conserved residues. These include Ser-74 on TM2 and Trp-158 and Ile-154 on TM4. High geometric conservation makes Trp-158 compatible with the ring(s) of the bound cholesterol. The nonpolar, van der Waals, and CH-interactions between Trp-158 and cholesterol are optimised by this compatibility. According to the postulated role of cholesterol as a modulator of the activity of membrane proteins in general and of GPCRs in particular, it has been proposed that the cholesterol-binding residues comprise an allosteric site of GPCRs. As previously mentioned, cholesterol and other membrane lipids may also aid in the dimerization of GPCRs, which is hypothesised to play a role in the control of GPCR signalling.Despite the fact that the ligands of all class A GPCRs interact with both the transmembrane and EC domains, the location of the binding pocket varies considerably, though not always, among the various receptors. For instance, the ligand extends between TMs 3 and 7 and the interface between TMs 5 and 6 in the pockets of rhodopsin and the -adrenergic receptors, respectively.

The bound retinal cofactor is physically close to Trp-265 on TM6, a component of the conserved CWxP motif, and extends farther in rhodopsin. This residue, along with Phe-208, is a component of the so-called transmission switch mechanism, which aids in the signal's propagation to the rhodopsin intracellular portion following activation. The inverse agonist and Trp-286 (the Trp-265) are separated by aromatic residues in the -adrenergic receptors. The adenosine (A2A) receptor's binding pocket is located considerably differently from those of rhodopsin and beta-adrenergic receptors. The second extracellular loop, ECL2, might interact with the ligand because it is situated closer to the junction between TM6 and TM7. Second, the ligand seems to be moved towards the membrane-solvent interface, where part of it is solvent-exposed, and extends perpendicular to the membrane plane.The chemokine, neurotensin, and opioid receptors are examples of peptide-binding GPCRs that must accept ligands that are bigger than biogenic amines like adrenaline or nucleotides like adenosine. Due to this, the peptide ligands bind to these GPCRs at locations that are often closer to the extracellular domain than those of the other class A GPCRs. There is still variation in ligand sites and interactions inside the peptide-binding GPCRs, just as there is in GPCRs that bind small molecules.

Due to the limited range of possible binding partners, which include G-proteins, arrestins, and G protein-coupled receptor kinases, the structure of the intracellular area of GPCRs is largely constant. One of the distinctive structural characteristics shared by members of the class A GPCRs is the conserved sequence motif D/ERY at the intracellular side of TM3. The first GPCR to have its three-dimensional structure identified, rhodopsin, showed an electrostatic connection between the motif's Arg-135 and the nearby Glu-247 on TM6 that is known as a ionic lock. The ionic lock's sequence was preserved, indicating that the interaction was significant. Additionally, it was discovered that rhodopsin activation causes the ionic lock to be broken. These findings, together with information from mutational research, led to the hypothesis that the ionic lock plays a crucial role in both the stabilisation of the GPCRs'

dormant status and, potentially, their activation or signalling. The ionic lock was shown to be absent in the three-dimensional structures of additional dormant GPCRs, such as the -adrenergic receptors, the A2A receptor, and the muscarinic (M2) receptor. The ionic lock is really in continual equilibrium between two conformations, one in which it is created and the other in which it is broken, according to earlier investigations of -adrenergic receptors. These receptors show observable baseline activity in their inactive states, but rhodopsin does not, which may be explained by the partial or entire lack of the ionic lock in the inactive forms of and A2A receptors [5], [6].The comparison of the invertebrate and vertebrate rhodopsin structures, as illustrated by the rhodopsins of the squid and the bovine, raises still another intriguing topic. The two proteins simply vary in their G-protein specificity and have the identical functional response to an agonist. It's interesting to note that this change seems to be enough to cause noticeable topological variations in the intracellular areas of the two structures. The third intracellular loop, or ICL3, exhibits the biggest difference, which is ascribed to the longer sequence of squid rhodopsin in that region.

In fact, ICL3 is thought to impart specificity to the intracellular binding partner, and altering this component across GPCRs changes the selectivity of the receptors' G-proteins. The heart and lungs must provide enough oxygen to the skeletal muscles, and the liver and muscles must swiftly break down the body's sugar reserves so that the muscles have access to fuel. This reaction requires the activities of multiple organs. Additionally, body processes that ordinarily occur when at rest and depend on the operation of certain organs, such as digesting by the gastrointestinal (GI) system, must be blocked to enable large amounts of peripheral blood to reach the muscles.It is obvious that some of these reactions are neurological in origin such as the activation of skeletal muscles and others are metabolic in form. While noradrenaline is more focused on the former, adrenaline participates in both kinds of reactions. The two catecholamines work via a group of adrenergic receptors, which are found in different organs. The precise outcomes of activating these receptors. These effects primarily include the contraction of certain muscles and the relaxation of others, as well as the constriction or dilatation of blood vessels depending on the organ. The physiological results of these interactions are as follows:

1. An increase in stroke volume and heart rate.
2. The transfer of blood to the heart, lungs, brain, and skeletal muscles from the skin and GI tract.
3. An increase in blood pressure.
4. Widening of the pupils.
5. Windpipe enlargement.
6. GI smooth muscle relaxation.
7. Blood clotting rate is increased, reducing the risk of bleeding after injuries.
8. Increased perspiration helps the exhausted body calm down.
9. A rise in metabolic rate brought on by the breakdown of muscle glycogen and liver lipid reserves.

The adrenal glands respond to both immediate dangers like fight-or-flight, which may be fatal, and more indirect threats like stress, which last for a longer period of time. The hypothalamus, the brain's hormonal control centre, stimulates the pituitary adrenocorticotropic hormone (ACTH) in the latter scenario, when the adrenal glands are hormonally stimulated by the brain. The chief metabolic controller of the brain, the hypothalamus, commands the anterior pituitary gland to release this physiologically active peptide into circulation. The hypothalamic-pituitary-adrenal (HPA) axis is the term for this sort of stimulation, which also causes the cortex of this gland to produce enormous quantities

of glucocorticoid hormones, particularly cortisol, into the bloodstream. The stress reaction, which is less important than the fight-or-flight response, is mediated by cortisol. Cortisol does indeed function more gradually than adrenaline and noradrenaline, but it still has a wide range of physiological consequences. For instance, its metabolic actions operate to raise blood glucose levels and are anabolic in the liver and catabolic in muscle and fat cells. The cardiovascular system, the central nervous system, the immune system, and the kidneys are just a few of the additional systems that cortisol may impact. The inflammatory reaction, which may be potentially dangerous after damage, is most critically inhibited by cortisol [7], [8].

Numerous pharmaceuticals are targeting the adrenaline-noradrenaline system because of the impact it has on various organs. These may be divided into the following categories. Adrenergic agonists, to start. In the trachea and bronchi, 2 receptors may bind. As a result, 2 agonists like albuterol are used in the treatment of asthma. The smooth muscles that regulate blood vessel diameter are known to have the receptors. Vasoconstriction results from the contraction of smooth muscle caused by the activation of 1 receptor, while vasodilation results from the activation of 2 receptors, which regulate their 1 counterparts. In order to manage high blood pressure, 2 agonists like clonidine are utilised as antihypertensive medications. Because of its catastrophic consequences on those who go untreated and the relative lack of symptoms, high blood pressure has been dubbed the silent killer.

The majority of 1 receptors are found in the heart, and medications that block these receptors, including atenolol, are used to treat hypertension, angina pectoris, and certain arrhythmias.Reuptake inhibitors, third. The main method that adrenaline and noradrenaline are inhibited from acting in the nervous system is by their absorption or reuptake into their secreting cells and away from the synapse. 'Reuptake inhibitors' are medications that target trans- porters, which are responsible for carrying out these activities. These medications inhibit transporters to increase the concentrations of these neurotransmitters in the synapses, therefore enhancing their effects. Adrenaline and noradrenaline reuptake inhibitors are mostly used in medicine to treat depression. The reuptake of both catecholamines and indoleamines is inhibited by several antidepressants of an earlier generation, such as the tricyclics (such as desipramine). Modern antidepressants (SSRIs) are more targeted and are solely intended to increase serotonin levels. However, a relatively recent family of antidepressants called SNRIs raises serotonin and noradrenaline levels.

Certain addictive substances, such as amphetamines (cocaine, MDMA), target the catecholamine reuptake system. These medications have a significantly stronger effect than the adrenaline-noradrenaline reuptake inhibitors previously stated. Amphetamines are very harmful to the cardiovascular system and are addictive.Inhibitors of monoamine oxidase (MAO). Catalysing the oxidative deamination of catecholamines and indoleamines is the enzyme MAO. As a result, its inhibition enhances serotonergic and adrenergic actions. The main purpose of MAO inhibitors, such as selegiline, is as an antidepressant.In addition to the medications mentioned above, there are additional medications that work on the parasympathetic branch of the autonomic nervous system to produce effects that are comparable to those of the aforementioned medications. For instance, atropine, an antagonist of the acetylcholine receptor, generates some of the physiological effects of agonists of adrenergic receptors acetylcholine being the primary neurotransmitter in the parasympathetic nervous system.

Activation of G-proteins and GPCR

Our current understanding of the alterations that GPCRs go through after activation is mostly derived from in-depth biochemical and biophysical research conducted on class A receptors. GPCRs must be understood in both their inactive and active states in order to study the activation process. There are several GPCR structures that have been crystallized in complex with an agonist at this time. Such configurations, however, only partially activate; investigations reveal that the GPCR also has to bind a G- protein or a protein that mimics it for example, a portion of an antibody on its intracellular side in order to take a completely active shape. Only three GPCRs, all of which are class A members, have so far been crystallized in fully active states:

1. Rhopadin. Two active structures have been identified, one coupled to an antibody and the other to an 11-amino acid fragment that represents the C-terminus of G.
2. The 2-adrenoceptor. Two active structures have been identified in association with an agonist, one attached to a nanobody the heavy chain of an antibody and the other to a complete G-protein molecule on its intracellular side.
3. The (M2) muscarinic receptor. Each of the two identified active structures is connected to a nanobody on its intracellular side and in complex with an agonist. An allosteric activator is also connected to one of these structures.

We concentrate on these three GPCRs in the explanation of GPCR activation that follows, but we also make mention of a few more GPCRs, such the A2A receptor, for which a partially active structure is known. We exclusively discuss the active structure of the G-protein-bound 2-adrenergic receptor since it has been discovered to be quite similar to the structure coupled to the nanobody. Please refer to the review for a more thorough explanation of the data derived from recognised GPCR structures.

GPCR structural modifications upon activation

Rhodopsin was the first GPCR whose active and inactive states' 3D structures could be acquired, making it the initial source of information on the activation process. A photoactivated protein called rhodopsin, which is found in the membranes of retinal rod cells, allows these cells to transmit visual information to the brain. Rhodopsin employs an organic cofactor termed 11-cis-retinal as the photo- reactive element since polypeptide chains are not well adapted for reacting to electromagnetic photons. Through a Schiff base to Lys-296, this molecule is covalently but reversibly connected to the polypeptide chain. The protonated Schiff base is stabilised in the inactive state by Glu-113 on TM3 . The retina reacts to light by going through isomerization, or switching from an 11-cis to an all-trans arrangement. Consequently, the protein undergoes conformational modifications that provide room for the corresponding G-proteins of rhodopsin to bind.Rhodopsin's intracellular side is bound by the G-protein's subunit, which stabilises the protein's active shape.

Rhodopsin is bleached after it has been activated. That is, the retinal-binding Schiff base is hydrolyzed, enabling the retinal to escape and inactivate the receptor for around 30 minutes. Opsin is the name for the retinal-free polypeptide chain that is still present and serves as the active form of rhodopsin.In the presence of all-trans retinal, two structures of completely active rhodopsin were identified in 2011. One of the structures was additionally associated with an intracellular fragment of 11 amino acids produced from Gs. Rhodopsin's stacked structures in the inactive and active states at first inspection show relatively minor conformational changes, particularly involving helices 5, 6, and 7 . The retinal pocket is examined more closely, and rather slight alterations are seen there, which appears to support the first conclusion. For instance [9], [10].

Certain side chain movements free up space and encourage the retinal cofactor's conversion from cis to trans. A gap is established between TM3 and TM7, mostly because of a shift in TM7. These movements include the side chain of Phe-208 on TM5 and that of Trp-265 on TM6, which moves into a region formerly occupied by the -ionone ring of the retinal cofactor. This increases the separation between Glu-113 and the Schiff base from 3.5 to 5.3 metres, breaking the salt bridge between them. The conformational alteration also enhances the salt bridge that now connects the Schiff base with Glu-181, which takes the position of Glu-113 as the primary stabiliser of the Schiff base.Additionally, according to NMR studies, the retinal cofactor dissociates just before the hydrogen bonds between ECL2 and the extracellular portions of helices 4, 5, and 6 are broken.

However, examination of the TMD of rhodopsin's intracellular side indicates more significant modifications in TMs 3, 5, 6, and 7. A region for transducin binding is created by the displacement of TMs 3 and 6 from the protein's centre and from each other . The local motions of Trp-265 and Phe-208 in the retinal-binding pocket, also known as the transmission switch, cause these structural alterations. The 'ionic lock' between Glu-247 and Arg-135 of the D/ERY motif is broken by the gap created between TMs 3 and 6 . The novel contacts between Arg-135 and some of the TM5 residues, such as Tyr-223 and TM6, that are created as a result of the shifts somewhat make up for the loss of this interaction. It has been suggested that the absence of the ionic lock, in at least some instances, in the other structurally defined GPCRs, accounts for the variations in baseline activity between the GPCRs. Rhodopsin, in contrast, is entirely inactive in the dark, whilst the other three GPCRs, like many others, continue to function even when they are not bound by their agonists. Inverse agonists that act on the 2-adrenergic receptor, such carazolol, may decrease this activity.

In contrast to the D/ERY motif, which stabilises the inactive state of GPCRs, the NPxxY motif in TM7 is involved in activating GPCRs. This is how the two motif types vary. In fact, the motif's Tyr-306 moves into TM6's former location,stabilising rhodopsin's active conformation. When these GPCRs are activated, similar movements are also seen in the equivalent places of Tyr-306 in the muscarinic receptor, the A2A receptor, and the 1 and 2 adrenergic receptors. It is thought that the movements similarly contribute to the activation process in these GPCRs. Stabilisation of the active site, provided by the Tyr-306-equivalent position in the 2-adrenergic receptor, depends in part on a water-mediated hydrogen bond with the Tyr-223-equivalent location on TM5. Although water molecules are absent from the M2 receptor's active structure, the similar water-mediated contact between the two tyrosine residues is expected to occur there as well. The conservation of this contact and the two tyrosine residues' identical locations in the three active GPCRs imply that the interaction is a distinguishing feature of GPCR activation. Conclusion: Despite affecting several rhodopsin components, the alterations outlined here are generally minor, measuring between 2 and 6. Nevertheless, they accomplish their goal of providing transducin with a binding site by separating transmembrane helices 3, 5, 6, and 7. We shall show that during activation, the M2 receptor and the 2-adrenergic receptor exhibit comparable changes in their overall conformations.

The rhodopsin-transducin complex's experimentally determined structures reveal several contacts, both polar and nonpolar, between the residues of transducin and those of rhodopsin that stabilise the latter's active conformation. Given the great degree of conservation of Arg-135, the direct contact between Arg-135 of the D/ERY motif and a backbone group in transducin would appear significant. The 2-adrenergic receptor, which has crystallised in association with a whole G-protein molecule, does not, however, exhibit this interaction. The

orientation of the transducin-derived peptide seems to be specifically determined by hydrogen bonds, which appear to induce rhodopsin-transducin binding.We discussed the conformational changes that GPCRs go through when activated in the preceding subsection. However, the ultimate objective of GPCR research is to comprehend how these modifications are brought about by ligand binding and how they result in the activation of the G-protein. This feature also has significant pharmacological ramifications since various ligands designed to bind to the same pocket in the GPCR may have various effects; some may operate as agonists while others may serve as antagonists or inverse agonists. Given that this GPCR possesses a non-diffusible agonist and that the matching structure of the active receptor is coupled to a tiny segment of the Gs-protein, the aforementioned rhodopsin structures are unable to provide a response to these queries. Luckily, the fully active structures of the muscarinic receptor and the 2-adrenergic receptor have both been identified [11], [12].

CONCLUSION

In conclusion, inactive bovine rhodopsin, a protein present in retinal rod cells and critical in the conversion of visual information into brain impulses, has been the main source for precise structural characterization of GPCRs (G Protein-Coupled Receptors) since the year 2000. Because of its stability and abundance, rhodopsin has been used as a useful model to comprehend the structure of class A GPCRs. This has allowed researchers to extensively study it. Rhodopsin is very stable and mostly inactive in the absence of light, but it's crucial to remember that many other GPCRs have basal activity even in the absence of an activating stimulus. Despite low sequence commonality across group members, GPCRs structurally show high homology within their respective classes. The ligand-binding site and surrounding loops of the extracellular domains (ECDs) of GPCRs differ significantly amongst receptors. For instance, the tertiary structure of rhodopsin is compact and stiff, restricting solvent and molecule access to the ligand-binding pocket, which is necessary for the specialised function of the protein as a photon receptor. Adrenergic receptors, like the -adrenergic receptor, on the other hand, have a larger open EC region, making them more approachable to ligands.

REFERENCES:

[1] S. S. Dong, W. A. Goddard, and R. Abrol, Conformational and Thermodynamic Landscape of GPCR Activation from Theory and Computation, *Biophys. J.*, 2016.

[2] M. P. Wescott *et al.*, Signal transmission through the CXC chemokine receptor 4 (CXCR4) transmembrane helices, *Proc. Natl. Acad. Sci. U. S. A.*, 2016.

[3] A. L. Isola and S. Chen, Exosomes: The link between GPCR activation and metastatic potential?, *Frontiers in Genetics*. 2016.

[4] Y. Wang, X. Xu, M. Pan, and T. Jin, ELMO1 directly interacts with Gßγ subunit to transduce GPCR signaling to Rac1 activation in chemotaxis, *J. Cancer*, 2016.

[5] A. J. Venkatakrishnan *et al.*, Diverse activation pathways in class A GPCRs converge near the G-protein-coupling region, *Nature*, 2016.

[6] S. Kerridge *et al.*, Modular activation of Rho1 by GPCR signalling imparts polarized myosin II activation during morphogenesis, *Nat. Cell Biol.*, 2016.

[7] R. Schrage, A. De Min, K. Hochheiser, E. Kostenis, and K. Mohr, Superagonism at G protein-coupled receptors and beyond, *Br. J. Pharmacol.*, 2016.

[8] H. Block *et al.*, Gnb isoforms control a signaling pathway comprising Rac1, Plcb2, and Plcb3 leading to LFA-1 activation and neutrophil arrest in vivo, *Blood*, 2016.

[9] I. Rodríguez-Espigares, A. A. Kaczor, and J. Selent, In silico Exploration of the Conformational Universe of GPCRs, *Molecular Informatics*. 2016.

[10] T. Kawano, A. Baki, G. Xiang, and D. E. Logothetis, Construction of G Alpha-16 Chimeras for Detection of GPCR Activation, *Biophys. J.*, 2016.

[11] A. Cannaert, J. Storme, F. Franz, V. Auwärter, and C. P. Stove, Detection and Activity Profiling of Synthetic Cannabinoids and Their Metabolites with a Newly Developed Bioassay, *Anal. Chem.*, 2016.

[12] Y. Ye, X. Tang, Z. Sun, and S. Chen, Upregulated WDR26 serves as a scaffold to coordinate PI3K/AKT pathway-driven breast cancer cell growth, migration, and invasion, *Oncotarget*, 2016.

CHAPTER 12

STRUCTURAL INSIGHTS INTO THE ACTIVATION MECHANISMS: DIVERSE G PROTEIN-COUPLED RECEPTORS

Chintakayal Purnima, Assistant Professor
College of Paramedical Sciences, TeerthankerMahaveer University, Moradabad, Uttar Pradesh, India Email Id- chintakayal.purnima3@gmail.com

ABSTRACT:

This abstract summarizes research on G-protein-coupled receptors (GPCRs) spanning different classes, including class A, B, C, and F receptors. It highlights the structural details and activation mechanisms of these receptors, emphasizing the significance of ligand binding and intracellular interactions. Notable findings include the role of hydrogen bonds in triggering conformational changes, the impact of receptor-ligand interactions, and insights into allosteric modulation. The abstract also touches on desensitization processes in GPCRs, involving phosphorylation, arrestin binding, and internalization. Overall, this research enhances our understanding of GPCR activation mechanisms and holds potential implications for drug development in various therapeutic areas. Additionally, the discovery of allosteric sites and modulators in GPCRs offers new possibilities for developing subtype-specific medications with potential applications in various disorders. Despite these significant findings, there are still unanswered questions and areas of ongoing research in the field of GPCR structural biology. These include understanding the full activation process, the GDP-GTP exchange mechanism, ligand selectivity, and the functional roles of allosteric modulators. Further research will continue to unveil the intricacies of GPCR signaling and its therapeutic implications.

KEYWORDS:

Adrenergic Receptor's, Diverse, Intracellular, Protein.

INTRODUCTION

The 2-adrenergic receptor's active structure was crystallized, coupled to a high-affinity agonist and an entire Gs-protein . Since the G-protein in the crystallized form lacks nucleotides, it was stopped in its tracks during the GDP-GTP exchange. The 2-adrenergic receptor's active structure may be compared to its inactive form when it is coupled to the inverse agonist carazolol. Similar to rhodopsin, activation of the 2-adrenergic receptor results in more pronounced intracellular modifications than extracellular changes to the protein's structure. Three hydrogen bonds are formed between the bound agonist and TM5 residues: two with Ser-203 and one with Ser-207. TM5 seems to be somewhat pulled inward by these interactions. The hydrophobic interaction network between Phe-208, Pro-211, Ile-221, and Phe-282 is rearranged as a result of this little displacement. To push its intracellular tip outward as a consequence, TM6 engages in a hinge movement, which is followed by considerably smaller inward motions of TMs 3 and 7. The carboxyl terminus of the G-protein's helix, which is partially forced into the receptor's transmembrane core, may fit comfortably thanks to the motions of TMs 5 and 6. In the inactive G-protein, this mechanism results in a significant displacement of one of G's domains with respect to the other; this displacement is thought to facilitate the exchange of GDP for GTP upon activation. Overall conformational change of the 2-adrenergic receptor is comparable to the alterations seen in rhodopsin, the 1-adrenergic receptor, and the muscarinic receptor. It's interesting to note that the displacement of TM6 at the 1-adrenergic receptor is much lower than that at the 2-

adrenergic receptor. This discrepancy might be attributed to the fact that only the structure of the 2-adrenergic receptor was crystallised when attached to a G-protein, which was probably responsible for inducing more significant conformational changes. The adenosine receptor, whose active structure was crystallised in the absence of a G-protein like the active structure of the 1-adrenergic receptor, similarly exhibits minor displacements of helices. NMR and molecular dynamics investigations have shown that G binding is necessary for properly stabilising the conformational changes brought on by the agonist [1], [2].

We have shown that the hydrogen bonds generated between the agonist and serines 203 and 207 in TM5 are the first in a series of events that cause significant conformational changes on the 2-adrenergic receptor's intracellular side. These interactions shift additional, more downstream interactions in addition to moving TM6 as they pull TM5 slightly inward. Interestingly, Trp-286, the other residue present in the transmission switch, does not seem to be involved in the activation of the -adrenergic receptor, but Phe-208 is implicated in this rearrangement. Serines 203 and 207 are highly conserved in aminergic receptors, therefore it is not unexpected that they are involved in the activity. Additionally, TM5 interacts with carazolol, which is bonded to the dormant structure of the 2-adrenergic receptor. Although this contact is weaker than that of BI-167107 in the active structure, it only creates one hydrogen bond with Ser-203, and it seems that this interaction is inadequate to cause the aforementioned structural modifications. This might explain why BI-167107 functions as an agonist while carazolol works as an inverse agonist. The 1-adrenergic receptor exhibits the same differential interactions with TM5; however, the resulting conformational changes seen in this GPCR are considerably less.

In this instance, they include the agonist isoprenaline and the antagonist cyanopindolol. In contrast, the A2A receptor's TM2 and TM7 helices interact differently with the agonist than they do with the inverse agonist , and TMs 3, 6, and 7rather than TM5transmit conformational changes to the intracellular side. The precise mechanisms by which these interactions inhibit or activate the GPCR may vary across GPCRs and involve different molecular switches and/or transmembrane segments, even though antagonist and inverse agonist interactions with their respective ligand-binding pockets generally differ from those of agonists.Rhodopsin and beta-adrenergic receptors are members of the same class as muscarinic receptors. There are five kinds of these receptors, which mediate cholinergic transmission. Drugs are designed to target them in order to cure various disorders. modulating a wide range of physiological activities. The completely active structure of the M2 receptor, attached to the agonist iperoxo and to a nanobody on its intracellular side, was found in 2013, a year after the antagonist-bound inactive structures of the M2 and M3 receptor subtypes were determined. The conformational changes in the ligand-binding site of the M2 receptor, which involve TMs 5, 6, and 7, are bigger than those seen in rhodopsin and the 2-adrenergic receptor, according to a comparison of the inactive and active structures of the receptor. These modifications cause the agonist to be completely buried within the pocket and cut off from the solvent.

An aromatic lid over the agonist, made up of the amino acids tyr-104, tyr-403, and tyr-426 , causes this blockage. The lid is partially open while the receptor is in its inactive state, but when it is activated, movements of these residues, particularly Tyr-403 and Tyr-426, enable them to form hydrogen bonds with one another and shut the lid. The conformational changes that occur after activation on the muscarinic receptor's intracellular side are often comparable to those seen in rhodopsin and the 2-adrenergic receptor . Particularly, the outward motions of TMs 6 and 2 and the corresponding inward movements of TMs 7 and 3 closely resemble those of the 2-adrenergic receptor and similarly produce the G-protein binding site. The

significant conformational shift of TM6 starts in the ligand-binding pocket and spreads to the intracellular side, much as in the other GPCRs.

Both the antagonist and the agonist hydrogen bond with Asn-404 in the ligand-binding pocket. The agonist, although being smaller than the antagonist, is situated nearer to TM3 and further from TM6. As a result, TM6 is drawn towards TM3 in this location by the interaction between the agonist and Asn-404 . Thr-399 causes TM6 to rotate, and the intracellular side of the protein moves farther away from TM3 as a result. Given the above and the findings of mutational investigations, it is believed that Asn-404 is a crucial residue in the activation of the M2 receptor. The modest displacement of TM3 caused by the creation of a stabilising hydrogen bond between Asp-120 of the D/ERY motif and Asn-58 is also responsible for the gap that is generated between TMs 3 and 6 following activation. This interaction has previously been seen in the active 2-adrenergic receptor, indicating that it functions generally in GPCR activation and disproving the notion that the D/ERY motif stabilises the inactive state. The M2 receptor seems to be activated by the second motif associated with GPCR activation, NPxxY. Particularly, the water-mediated interaction that is predicted to take place between Tyr-440 of the motif on TM7 and Tyr-206 on TM5 is likely to stabilise the receptor's active state.

DISCUSSION

The aforementioned research looked at allostery in the M2 receptor as well. The protein crystallised in conjunction with the positive allosteric modulator LY2119620 and the agonist iperoxo. In the allosteric activator binds right above the agonist. The two structures of the GPCRthose bound and unbound to the activatorare very identical, with the exception of a few minor modifications that mostly include residues that interact with the latter. This suggests that the allosteric site is pre-formed by the agonist. According to our present understanding of allostery, which holds that allosteric activators stabilise an active conformation of the protein while allosteric inhibitors stabilise an inactive conformation (, this theory is consistent with our understanding of allostery. Numerous class A GPCRs, including as the adenosine, dopamine, histamine, serotonin, and chemokine receptors, as well as class C GPCRs have been shown to include allosteric modulators. These GPCRs have allosteric sites that are located in their extracellular or transmembrane domains. Lipids, amino acids, ions, and numerous tiny molecules are among the chemically varied modulators. In reality, since each binds preferentially to an active or inactive conformation of the receptor and stabilises it, the G-proteins and other intracellular binding partners of GPCRs such as GRK and arrestins; see the next paragraph may also be thought of as allosteric modulators. Due to the fact that allosteric modulators are widely desired by the pharmaceutical industry. GPCRs have allosteric sites that may be found throughout the protein, even on the outside, membrane-facing surface. An allosteric site in this location is shown, for instance, in the P2Y1 receptor [3], [4].

To sum up, the aforementioned research as well as a large number of others conducted recently have revealed the following insights about the activation of class A GPCRs. The activation process typically entails only minor conformational changes in the extracellular and ligand-binding domains. The activation causes the transmission of small, local conformational changes from the ligand-binding site of the GPCR to its intracellular side by a variety of molecular triggers, which include ionic locks and transmission switches (. One exception is the P2Y12 receptor, where 5 to 10 shifts are observed on the extracellular sides of TM6 and TM7. Despite not acting the same way in all GPCRs, these triggers include physically comparable processes and structurally analogous locations that are found in highly conserved motifs like D/ERY and NPxxY.The rearrangements of helices on the intracellular

side are generally similar in both light-activated rhodopsin and agonist-activated GPCRs. G undergoes large conformational changes upon binding to its rhodopsin.But there are still some unresolved details of the activation procedure.

The wide range of structural modifications seen in rhodopsin and beta-adrenergic receptors. Finding more fully-active GPCR structures, or those that are concurrently bound to an agonist and an intracellular binding partner, is necessary for evaluating generality. Gaining such knowledge is crucial for GPCRs belonging to classes other than A, for which there aren't any completely active structures.The GDP-GTP trade-off. The complex of the 2-adrenergic receptor and nucleotide-free Gs was identified. Additional research on the GTP-bound form of GPCRs and the possible intermediates between the two states are required to provide a comprehensive picture of the activation mechanism. It should be emphasised that while GTP promotes the G-protein's separation from the receptor, crystallizing the active state of the GPCR with G-protein attached to GTP is not simple.Selectivity of ligands. Additional research is needed to comprehend the mechanistic effects of biassed ligands, or ligands that bind to GPCRs and activate particular signalling pathways, as well as the effects of ligands with different functionalities and receptor subtype specificity. Allosteric ligands may potentially have an impact on GPCRs, and it will be fascinating to characterise the activation or inhibition mechanisms used by such ligands.It is fascinating that GPCRs have an order of magnitude higher diversity than their G-protein counterparts at the system level. One can wonder what advantage there is to having such a varied sensing end (GPCRs) that ultimately reduces to the constrained G-protein repertoire from the perspective of signal processing [5], [6].

Desensitisation of GPCR

In order to avoid overstimulating the signalling pathway, GPCR function is regulated in part by GPCR inhibition after brief or prolonged activation. Desensitisation is the term used to describe the first form of downregulation, which takes place immediately after the GPCR is activated. Desensitisation comes in two forms. When Ser and Thr residues in ICL3 or the C′ of the GPCR are phosphorylated, homologous desensitisation affects the activated receptor. A particular class of Ser/Thr kinases known as GRKs, which exclusively operate on the receptor's agonist-bound conformation, catalyses the phosphorylation. The phosphorylation increases the GPCR's affinity for proteins in the arrestin family and makes it easier for those proteins to bind to the receptor. The GPCR is prevented from connecting with its corresponding G-protein, which halts signalling, as a result of the binding to arrestin. The GPCR is then internalised into the cell through clathrin-coated vesicles, where it is either recycled or destroyed. Next, it attracts clathrin and its adaptor protein AP-2.

It's interesting to note that after internalisation, class A GPCRs lose their clathrin coat and become dephosphorylated; as a consequence, it's possible that they may still signal when within the endosome. However, in this scenario, the GPCR and its effectors are situated nearer to the cell's nucleus than they were in the plasma membrane, which may increase the efficiency of the transcriptional pathway's activation. Class B GPCRs, on the other hand, continue to be coupled to arrestin after internalisation, which causes their ubiquitylation and degradation. Recent investigations have shown that arrestins have a far more complex role in cellular signalling than was previously thought. Arrestins seem to have a specific role in biassedagonism because when they bind to GPCRs they stabilise particular conditions that inhibit certain signalling pathways and support others. Therefore, rather from being thought of as sig- nal terminators, arrestins should be seen as multifunctional adaptor proteins.Other receptors are affected by heterologous desensitisation, which is carried out by second

messenger-activated kinases such PKA or PKC. Following sustained receptor stimulation, further downregulation processes take effect on active GPCRs.

These may influence gene transcription and translation among other processes. Visual arrestin and active rhodopsin's structural interaction was investigated in 2015. Intriguingly, the structure revealed that arrestin binds to rhodopsin asymmetrically, which should enable arrestin's conserved hydrophobic residuesPhe-197, Phe-198, Met-199, Phe-339, and Leu-343to touch or perhaps intrude into the nonpolar part of the membrane. Arrestins do not have hydrophobic chains that bind to the membrane, unlike G-proteins and GRKs. As a result, arrestins may only be able to attach to the membrane through the conserved hydrophobic patch, and this anchoring may then stabilisearrestin's connection with the GPCR. In fact, changing any of these residues to alanine has an impact on how arrestin binds to rhodopsin. It has also been proposed that the extremely asymmetric structure of the rhodopsin-arrestin complex influences the curvature of the membrane, perhaps as a result of arrestin's function in triggering rhodopsin's endocytosis [7], [8].A number of components facilitate binding to rhodopsin from the arrestin end. Similar to how transducin, the G subunit of rhodopsin's G-protein, inserts, a brief helix-like segment of arrestin inserts into the intracellular side of rhodopsin.

Interactions between the arrestin helical segment and helix 8's amino terminus and TM7's carboxy-terminus occur. In fact, earlier investigations have shown that both of these components are crucial for arrestin binding. Therefore, arrestin and transducin are in direct competition. The second intracellular loop of rhodopsin and arrestin interact in another intriguing way. Arrestin adopts a tight shape in its apo form, which is independent of rhodopsin, making the interaction area inaccessible. When it binds, it expands to provide room for the ICL2 of rhodopsin. ICL2 assumes a helical shape in arrestin-bound rhodopsin, whereas it is organised as a loop in active rhodopsin that is arrestin-free.As previously noted, active rhodopsin can only bind to arrestin with high affinity after being phosphorylated by a GRK. Rhodopsin is not phosphorylated in the rhodopsin-arrestin complex as previously reported, and the binding was enabled by the introduction of mutations into the two binding partners. As a result, arrestin's structure in such scenario depicts a protein in a pre-activated state. Understanding the structure of phosphorylated rhodopsin linked to active arrestin will help us better understand how arrestin is activated and, hopefully, how arrestin binding triggers signalling pathways.

Other types of GPCRs

Compared to class A structures, the number of structures identified for GPCRs outside of class A is much less. The following examples of non-class-A structures are listed:

1. Class B –glucagon and corticotropin-releasing factor (CRF) receptors.
2. Class C – mGlu1, mGlu5, and GABAB receptors.
3. Class F – Smoothened protein (SMO, a GPCR-like receptor).

Class B GPCRs

Large peptide hormones are bound by class B GPCRs, making them appealing targets for therapeutic medicines used to treat conditions related to glucose metabolism such as diabetes, the stress response, cardiovascular control, etc. Unfortunately, there are only two members of this groupthe corticotropin-releasing factor (CRF) and glucagon receptorswhose structural characteristics have been characterised. Additionally, in order to stabilise the structures for crystallisation, the N- and C-termini must be removed in considerable parts. As a result, the structures mostly reflect the receptors' trans-membrane domains. Combining the non-peptide

antagonist CP-376395 with the CRF receptor subtype 1 (CRFR1) allowed researchers to establish its structure. Between the structure of CRFR1 and the class A GPCRs mentioned above, there are three striking differences:

The binding location for the antagonist is close to the intracellular side. However, it should be noted that class B GPCRs' natural ligands bind to both the extracellular and trans-membrane domains of the receptor. Around Gly-356, TM7 exhibits a severe kink just above the transmembrane domain's midpoint. This residue is a component of the class B GPCRs' conserved QGxxV motif. Similar to the NPxxYproline in class A GPCRs, which likewise causes a distortion in TM7, gly-356 enables TM7 to acquire the kink. Contrarily, the intracellular side of the transmembrane domain resembles the form of class A GPCRs despite the absence of ICL2, suggesting that this structure is probably capable of interacting with the homologous G-protein.The glucagon receptor (GCGR), which crystallised in the same year as CRFR1, was similarly shown to include a sizable cavity, but the precise binding site is unclear since the ligand could not be resolved. The glucagon receptor does not have the same prominent V shape as the CRFR1 receptor. Furthermore, in terms of the locations and orientations of the transmembrane helices, GCGR seems to have a structure that is more comparable to class A GPCRs than CRFR1. TM1, which is substantially longer in the GCGR structure than in the structures of class A GPCRs, is one obvious distinction between the two. The placement of the ECD in relation to the TMD and the binding of glucagon are both believed to occur in this area [9], [10].

Parts of the ECD and a bound peptide are absent from the structures mentioned above. This prevents us from comprehending how the natural peptide agonist binds to the receptor and activates it, as well as the spatial relationship between the ECD and the TMD. However, biochemical evidence links the GWGxP motif to a network of inter-actions that is crucial for certain functions. Although these connections can be detected in the CRFR1 structure, it is unclear exactly what function they provide since we don't know the receptor's active structure. The natural peptide ligands of class B GPCRs bind to both the ECD and the TMD, as was already noted. Particularly, the peptide agonist's N-terminus binds to the TMD and the C-terminus predominantly to the ECD. Studies of the isolated ECD of class B GPCRs in combination with a peptide agonist have been motivated by this trend (see review by Parthier and collaborators). According to the study, this group's members have big ECDs that are all comparable in size. This is in line with the big class B GPCR ligands, which are peptides of around 30 amino acids. Three disulfide bonds serve to stabilise the two tiny -sheets, a neighbouring -helix, and the peptide-binding site. As with inherently unstructured proteins the binding of the peptide ligands seems to be connected to their folding. When the two -sheets of the ECD are pressed together by the pep- tide ligand's C′, it makes mostly nonpolar but also polar interactions with the corresponding ECD residues of loops 2 and 4, as well as with the C′ of the GPCR.

The peptide ligand's N′ is retained near to the TMD on the opposite side. The N′ of the ligand may interact with the GPCR's extracellular loops and transmembrane helices in this configuration. Additionally, the binding causes conformational changes in ECD loops that might be communicated to the TMD and are thus a component of the activation process.It is hardly unexpected that so much attention is being paid to mGlu receptors given its significance in biology, medicine, and therapy. The primary excitatory neurotransmitter in the central and peripheral nervous systems is glutamate, which is the natural agonist of these receptors. It is involved in a wide range of neurological processes, including memory and learning, sensory and motor functions, emotions, etc. Therefore, disorders including epilepsy, neurodegeneration, chronic pain, schizophrenia, anxiety, and autism are caused by mGluR

dysfunction. In light of this, mGlu receptors represent attractive targets for several neurological and psychiatric medications.

A Venus fly trap (VFT) domain, which binds the glutamate agonist, and a cysteine-rich (CR) domain, which connects the VFT domain to the TMD, make up the structure of the ECD of mGlu receptors. Consequently, class C GPCRs exclusively bind their ligands via the ECD, in contrast to GPCRs of classes A and B. In addition to mediating mGluR*1 dimerization, the VFT domain has significant conformational flexibility in its crystallised forms, which is maintained by agonist binding. In 2014, it was discovered how the negative allosteric modulator mavoglurant interacted with the TMD of mGluR5. This building was very significant for two reasons. First, it was proposed that allosteric modulation of mGluR5 might be used to treat schizophrenia, cognitive function abnormalities, and anxiety disorders (positive modulation) as well as anxiety disorders. Second, because allosteric sites are less conserved than orthostatic sites, medicines acting on allosteric sites are more likely to be subtype-specific, as was previously stated. Large flexible parts of the extracellular and intracellular domains of mGluR5 had to be excised for crystallisation, similar to the scenario with the class B receptors mentioned above. As a result, the precise location of the ECD with regard to the TMD is still uncertain. In the same year, the structure of the mGluR1 TMD coupled to the antagonistic allosteric modulator FITM was established.

However, in the instance of mGluR1, the linker region to the ECD was identified, enabling us to hypothesise the monomeric receptor's complete structure. Six cholesterol molecules were found to be located at the interface between the monomers of the structure after it was solved as a dimer (not illustrated). The prior hypothesis that cholesterol molecules inside the membrane are involved in GPCR dimerization is supported by these data. Particularly on the intracellular side, the TMDs of mGluR1 and mGluR5 are generally comparable to the TMDs of class A and class B GPCRs. This makes sense, as previously stated, given that all GPCRs only have a small number of intracellular binding partners, necessitating a minimum amount of structural diversity in the intracellular area. The majority of the changes seem to be in TMs 5 through 7 when the different classes are examined in terms of the conformation of each helix in the TMD (see more below). The allosteric binding site in both mGluR structures is deeper than the typical class A binding site, but it is not as deep as the binding site in the (class B) CRF receptor for the antagonist CP-376395. Even yet, FITM extends farther towards the transmembrane domain's extracellular side. The two allosteric modulators interact differently with the TMD, highlighting the possibility that such modulators might work as subtype-specific medications. ECL2 closes up the allosteric site entrance in both structures, and it is relatively narrow.

This is mostly because TMs 5 and 7 are oriented more inwards than their equivalents in class A and class B GPCRs. There is no need for a large entry to the TMD since, unlike in class A GPCRs, the natural ligands of mGlu receptors bind to the ECD, which is compatible with the capping of the TMD by ECL2. As we've seen with class A GPCRs, the mGlu receptors also feature a complex of ionic locks and other interaction, which are crucial for structural stabilisation and the activation process. Although the GPCR classes' locations in these interactions vary, their functional processes are the same. The functional motifs of class A GPCRs and class C (mGlu5R) GPCRs are compared in further depth in.A GPCR-like protein called Smoothened (SMO) is a component of the hedgehog (Hh) signalling system, which controls the development of animal embryos. This pathway's dysfunction causes abnormalities in the developing embryo and, sometimes, cancer in adults. Since 2013, the structures of SMO in association with several ligands have been identified. Despite the low sequence similarity (10%) between the two categories of proteins, the design of the TMD in

the initial structure of human SMO in association with the antagonist LY2940680 was identical to that of class A GPCRs. Similar to the glutamate (mGlu) receptors, TMs 5 through 7 make up the majority of the variations in this domain. The ligand interacts mostly with ECL2 and ECL3 at the interface between the complex extracellular domains and the transmembrane. Similar to mGlu receptors, SMO has a limited ligand-binding cavity, in part due to the inward placement of TM5. ECL2, however, is found within the TMD in the case of SMO.

The structures of SMO that are agonist- and antagonist-bound were determined in 2014. An ionic lock between the amino acids Arg-400 (TM5) and Asp-473 (TM6) at the ligand-binding region is one of the most obvious distinctions between the two structures. Due to a conformational shift in Arg-400, this contact is present in the antagonist-bound structure but is removed following antagonist binding. Arg-400 forms a hydrogen connection with Asn-477 at its new location, which is also a component of TM6. Although the existing structures do not show the underlying mechanism, this 'remodeling' of contacts is probably what acts as a molecular switch in the activation of SMO. The orientation of TM5 seems to be the sole substantial variation between the two structures on the intracellular side of the transmembrane domain, in contrast to class A GPCRs, whose activation is predominantly mediated by alterations in TM6. The agonist-bound SMO lacks an intracellular binding partner, which implies it is not completely activated; nevertheless, this discrepancy does not necessarily represent a fundamental differentiation between class A and class F GPCRs. Complete activation of the receptor might result in modifications that are more like to those we saw in rhodopsin and the 2-adrenergic receptor.

CONCLUSION

In conclusion, the structural studies of G-protein-coupled receptors (GPCRs) have provided valuable insights into the mechanisms underlying their activation and regulation. These studies have revealed that GPCRs, despite their diversity, share common features in their activation processes. The activation of class A GPCRs typically involves minor conformational changes in their extracellular and ligand-binding domains. These changes trigger small, local conformational alterations in the intracellular side, mediated by various molecular triggers like ionic locks and transmission switches. This leads to the formation of the G-protein binding site, allowing downstream signaling. The intracellular conformational changes in GPCRs are generally similar to those observed in rhodopsin and the β2-adrenergic receptor. The activation of class B GPCRs, such as the corticotropin-releasing factor (CRF) and glucagon receptors, relies on large peptide hormones. The ligands for these receptors bind to both the extracellular and transmembrane domains. Structural studies have revealed the presence of ECD and CR domains in the ECD of mGlu receptors, providing valuable information about ligand binding and potential allosteric modulation sites. Class C GPCRs, like metabotropic glutamate receptors (mGluRs) and Smoothened protein (SMO), have unique structural features. mGluRs utilize a Venus fly trap domain for glutamate binding and display distinctive structural elements. SMO, a member of class F GPCRs, shares structural similarities with class A GPCRs but exhibits variations in the intracellular side, particularly in TM5.

REFERENCES:

[1] J. S. Smith and S. Rajagopal, The β-Arrestins: Multifunctional regulators of G protein-coupled receptors, *J. Biol. Chem.*, 2016.

[2] D. P. Staus *et al.*, Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation, *Nature*, 2016.

[3] C. Oswald *et al.*, Intracellular allosteric antagonism of the CCR9 receptor, *Nature*, 2016.

[4] K. N. Woods, J. Pfeffer, A. Dutta, and J. Klein-Seetharaman, Vibrational resonance, allostery, and activation in rhodopsin-like G protein-coupled receptors, *Sci. Rep.*, 2016.

[5] Y. Wang, X. Xu, M. Pan, and T. Jin, ELMO1 directly interacts with Gßγ subunit to transduce GPCR signaling to Rac1 activation in chemotaxis, *J. Cancer*, 2016.

[6] V. Mahadevan and M. A. Woodin, Regulation of neuronal chloride homeostasis by neuromodulators, *Journal of Physiology*. 2016.

[7] A. W. Kahsai *et al.*, Conformationally selective RNA aptamers allosterically modulate the β 2-Adrenoceptor, *Nat. Chem. Biol.*, 2016.

[8] E. Schauberger, M. Peinhaupt, T. Cazares, and A. W. Lindsley, Lipid Mediators of Allergic Disease: Pathways, Treatments, and Emerging Therapeutic Targets, *Current Allergy and Asthma Reports*. 2016.

[9] S. E. Mansoor, W. Lü, W. Oosterheert, M. Shekhar, E. Tajkhorshid, and E. Gouaux, X-ray structures define human P2X 3 receptor gating cycle and antagonist action Crystallization and structure determination, *Nature*, 2016.

[10] D. P. Poole and N. W. Bunnett, G protein-coupled receptor trafficking and signalling in the enteric nervous system: The past, present and future, in *Advances in Experimental Medicine and Biology*, 2016.

CHAPTER 13

COMPLEXITY OF PROTEIN STRUCTURE: FROM PRIMARY TO TERTIARY LEVELS

Dr. Jyoti Trivedi, Assistant Professor
Department of Biochemistry, TMMC&RC, TeerthankerMahaveer University, Moradabad, Uttar Pradesh India,
Email Id- jyoti.trivedi01@gmail.com

ABSTRACT:

This gives a general overview of the complex nature of proteins and how their basic, secondary, and tertiary structures affect both their structure and function. It opens by discussing the difficulty of depicting intricate protein structures and the value of concise examples. The explanation of proteins' primary, secondary, tertiary, and quaternary structures emphasises the importance of genetic information and interactions between amino acid residues. The abstract goes on to cover the evolutionary features of protein structure, including how mutations affect the evolution of amino acid sequences through time and how conservative and variable substitutions affect protein function. Using cytochrome C as an example, it discusses how sequence homologies may be used to identify genetic connections across different species. The abstract switches the subject to molecular illnesses, focusing on sickle-cell anaemia in particular and describing the genetic mutation that causes this ailment and how it affects both red blood cells and general health. In certain communities, the coexistence of the sickle-cell gene with malaria resistance is also considered. The abstract then delves into proteins' intricate structural components, highlighting the existence of secondary structures like alpha-helices and beta-pleated sheets. It introduces the idea of super secondary motifs or structures and goes into detail about the variables affecting these structures. Following that, the abstract discusses tertiary structure and explains how globular proteins take on distinct three-dimensional conformations. It discusses the compactness of these structures, the function of domains in big proteins, and the importance of electrostatic contacts, hydrogen bonds, covalent bonds, hydrophobic interactions, and hydrophobic interactions in stabilising protein structures. The detailed description of the complexities of proteins is concluded by the abstract, which also discusses the role of water in protein structure and the dynamic hydration shell surrounding proteins.

KEYWORDS:

Cytochrome, Genetic, Protein, Polypeptide, Water.

INTRODUCTION

Proteins are very intricate molecules. Detailed representations of even the shortest polypeptide chains are almost challenging to grasp. Simpler illustrations that emphasize certain facets of a molecule are helpful. Two different ways to communicate structural information about proteins. The ball-and-stick model, a different structural illustration, is shown subsequently. There are many layers to the structural organisation of proteins, according to biochemists. Genetic information determines the amino acid sequence, which is the primary structure. Secondary structure is created when the polypeptide chain folds and takes the shape of certain localized arrangements of neighbouring but not necessarily contiguous amino acids. The tertiary structure of a polypeptide refers to its overall three-dimensional form [1], [2]. A protein is considered to have a quaternary structure if it has two or more polypeptide chains or subunits.the enzyme adenylate kinase. This space-filling model depicts the overall shape and volume of the molecular parts. Flat arrows in a ribbon model

depict pleated parts. The -helices resemble ribbon-like spirals. Each polypeptide contains a unique sequence of amino acids. The three-dimensional structure, function, and relationships to other proteins of a protein are all governed by interactions between amino acid residues [3], [4].

DISCUSSION

Homologous polypeptides are those that share an ancestor gene and have comparable amino acid sequences. The genetic links of several animals have been traced using comparisons of the sequences of homologous polypeptides. For instance, the sequence homologies of cytochrome C, a mitochondrial redox protein, have been widely exploited in the study of species evolution. Cytochrome c, a crucial molecule in the creation of energy, has a high degree of sequence conservation across a wide range of species, according on comparisons of its amino acid sequence. All homologues of a protein that have the same invariant amino acid residues are thought to be necessary for the protein's function. The invariant residues in cytochrome c engage with heme, a prosthetic group, or certain other proteins involved in energy production.

Molecular Diseases, Evolution, And Primary Structure

The amino acid sequences of polypeptides evolve throughout time as a consequence of evolutionary processes. These variations are brought about by mutations, which are spontaneous and random changes to DNA sequences. The function of a polypeptide is unaffected by a large number of fundamental sequence alterations. Because an amino acid with a side chain that is chemically identical is replaced, some of these replacements are referred to be conservative substitutions. Leucine and isoleucine, which both have hydrophobic side chains, may be swapped out for one another at certain sequence places, for instance, without impacting function. Some points in a sequence have much less restrictions. These residues, also known as variable residues, seem to have a variety of functions in how the polypeptide works.

It has been used to determine the evolutionary links between substitutions at conservative and variable locations. These studies make the assumption that the number of structural changes in a particular polypeptide will increase with the length of time since the divergence of two species. For instance, it is thought that the evolutionary history of humans and chimpanzees separated just recently. This assumption, which is backed by the main sequence data for cytochrome c, which shows that the protein is the same in both species, is mostly based on fossil and anatomical evidence. It is thought that kangaroos, whales, and sheep all descended from a common ancestor more than 50 million years ago since each of their cytochrome c molecules differs from the human protein by 10 residues. It is noteworthy to observe that despite several changes in the amino acid sequence, the overall three-dimensional structure often remains the same. There may be a striking similarity in the shapes of proteins that are coded for by genes that split millions of years ago.

The hydrophobic side chain of valine, the substituted amino acid in the -chain, interacts with a hydrophobic pocket in a second haemoglobin molecule to induce HbS molecules to group together form rod-like filaments. However, mutations may also be harmful. These arbitrary alterations in gene sequence might be mild or severe. For instance, it is unable to sustain individual organisms with nonconservative, variable amino acid changes at the conservative, unchanging residues of cytochrome. However, mutations don't always result in death; they might still have a significant impact. One of the most well-known examples of a class of illnesses that Linus Pauling and his colleagues referred to as molecular disorders is sickle-cell anaemia, which is brought on by mutant haemoglobin. Dr. Pauling used electrophoresis to

show for the first time that sickle-cell patients had a mutant haemoglobin. Two identical -chains and two identical -chains make up adult human haemoglobin (HbA). A single amino acid alteration in the HbA -chain causes sickle-cell anaemia. The sole variation between HbA and sickle-cell haemoglobin (HbS), according to analysis of the haemoglobin molecules from sickle-cell patients, is at amino acid residue 6 in the -chain. In the oxygen-free state, HbS molecules assemble to form stiff rod-like structures because a hydrophobic valine replaces a negatively charged glutamic acid. Severe anaemia develops as a consequence of the patient's red blood cells taking on the sickle shape and being vulnerable to hemolysis. The ability of these red blood cells to bind oxygen is extremely poor [5], [6].

Tissues lose oxygen as a result of sickled cells intermittently blocking capillaries. The symptoms of sickle-cell anaemia include extreme agony, eventual organ damage, and a premature death. Due to the crippling nature of sickle-cell disease, afflicted people almost ever lived into infancy until recently. Therefore, it would seem to reason that the harmful mutational alteration that causes this ailment would be swiftly eradicated from human populations. The sickle-cell gene is not as uncommon as one would anticipate, however. Only those with two copies of the sickle-cell gene are susceptible to sickle-cell disease. These people, known as homozygotes, get one faulty gene copy from each parent. The sickle-cell trait is thought to run in both parents. These individuals, known as heterozygotes because they have one healthy HbA gene and one damaged HbS gene, are comparatively symptom-free despite having roughly 40% HbS in their haemoglobin. In certain parts of Africa, the prevalence of sickle-cell trait is extremely high. Malaria, which is brought on by the parasite Plasmodium and spread by Anopheles mosquitoes, is a major public health issue in these regions. Because red blood cells with the sickle-cell trait provide a less favourable environment for the parasite to flourish than normal cells do, sickle-cell trait carriers are less likely to have malaria. The prevalence of the sickle-cell gene has stayed high because people with sickle-cell trait are more likely to survive malaria than healthy people.

Subtle Structure

Polypeptides have a number of repeating patterns that make up its secondary structure. The -helix and the -pleated sheet are the two secondary structure types that are most often seen. Localised hydrogen bonding between the carbonyl and N—H groups in the polypeptide's backbone stabilises both -helix and -pleated sheet patterns. The -carbons are swivel points for the polypeptide chain because peptide bonds are stiff. The and angles are influenced by a number of characteristics of the R groups linked to the -carbon, such as size and charge, if any. Specific secondary structural patterns may be supported or prevented by particular amino acids. Secondary structural patterns make up the majority of many fibrous proteins.When a polypeptide chain twists into a right-handed helical shape, the result is the -helix, a hard, rod-like structure. Each amino acids N—H group forms hydrogen bonds with the amino acid's carbonyl group, which is located four residues distant. The pitch the distance between equivalent spots per turn of the helix is 0.54 nm, and there are 3.6 amino acid residues per turn. From the helix, amino acid R groups protrude outward. Certain amino acids do not promote the development of -helices due to a number of structural restrictions namely, the stiffness of peptide bonds and the permitted limitations on the values of the and angles.

For instance, the R group in glycine is so tiny that it might cause the polypeptide chain to become overly flexible. The stiff ring of proline, on the other hand, prevents the N—C bond from spinning. Additionally, proline lacks the N—H group necessary to create the intrachain hydrogen bonding necessary for the -helix structure. When two or more polypeptide chain segments are arranged side by side, pleated sheets are formed. A -strand designates each separate section. Each strand is completely stretched rather than coiling. Hydrogen bonding

between the polypeptide backbone's N—H and the carbonyl groups of nearby chains help to stabilise pleated sheets. Sheets with pleats might be parallel or antiparallel. The polypeptide chains' hydrogen bonds are structured in parallel -pleated sheet structures in one direction, while they are arranged in antiparallel chains in the other way. Mixed parallel-and-antiparallel sheets have been seen on occasion [7], [8].

Multiple globular proteins combine helix and pleated sheet secondary structures. These patterns are referred to as super secondary motifs or structures. An -helix segment in the unit connects two parallelpleated sheets. Hydrophobic interactions between nonpolar side chains extending from the interacting surfaces of the -strands and the helixstabilize the structure of units. The structural components known as loops are responsible for a polypeptide's abrupt shifts in direction. The -turn, a frequently seen loop type, is a 180o turn with four residues. A hydrogen bond is created between the amide hydrogen of the fourth residue and the carbonyl oxygen of the first residue in the loop. Residues of glycine and proline often occur in turns. A tight turn may occur in a polypeptide strand because glycine lacks an organic side group, which allows a contiguous proline to adopt a cis orientation on the same side of the peptide plane.

Proline is a helix-breaking residue that modifies the polypeptide chain's orientation. Proteins with several -helical segments often include the b-turn. In the -meander pattern, polar amino acids and glycines link two antiparallel -sheets to create a more abrupt direction shift known as a reverse or hairpin bend. Two helical regions that are separated by a nonhelical loop align in a certain manner in units also known as helix-loop-helix units as a result of interacting side chains. Different sheet configurations fold back on themselves to generate various barrel formations. The Greek key motif is created when an antiparallel sheet folds back on itself in a way that mimics a typical Greek pottery pattern.

Tierre Facility

Although globular proteins often include a sizable number of secondary structural components, a number of additional elements also play a role in their structure. As globular proteins fold into their native structures and prosthetic groups, if any, are inserted, they adopt distinctive three-dimensional conformations known as tertiary structures. Protein folding, a process in which an unorganised, nascent molecule acquires a highly organised structure, occurs as a result of the interactions between the side chains in their primary structure. Several crucial characteristics of tertiary structure include:

1. A lot of polypeptides fold such that amino acid residues that were far apart in the initial structure are now near together.
2. Because the polypeptide folds efficiently, globular proteins are compact. The majority of water molecules are kept out of the core of the protein during this process, enabling interactions between polar and nonpolar groups.
3. Numerous compact structures known as domains are often found within large globular proteins, especially those that have more than 200 amino acid residues. In general, domains are structurally distinct portions with particular roles such as binding an ion or small molecule. A fold is a domain's primary three-dimensional structure. The Rossman fold, which binds nucleotides, and the globin fold are well-known examples of folds. Based on the structure of their primary motif, domains are categorized. Only -helices make up -domains, while -domains are made up of antiparallel strands. Different combinations of a -helix and -strands may be found in Domains.

Several eukaryotic proteins, known as modular or mosaic proteins, include several defective or duplicate copies of one or more connected domains. Three repeating domains make up fibronectin: Fl, F2, and F3. The extracellular matrix (ECM) proteins that include all three of these domains include collagen and heparan sulphate, as well as proteins that bind to certain cell surface receptors. Genetic sequences produced by gene duplications extra gene copies caused by mistakes in DNA replicationcode for domain modules. Living things employ these sequences to create new proteins. For instance, a number of cell surface proteins have the immunoglobulin structural domain in addition to antibodies [9], [10].

Hydrophobic interactions, to start. Hydrophobic R groups are excluded from water during polypeptide folding, bringing them close together. The highly ordered water molecules in the solvation shells are then freed from the interior, resulting in a rise in the water molecules' entropy. Protein folding is primarily driven by the favourable entropy change. The centre of folded proteins still contains a few water molecules, which each establish up to four hydrogen bonds with the polypeptide backbone. The polypeptide may be released from some of its internal interactions thanks to the stabilization provided by tiny structural water molecules. The polypeptide chain's consequent increase in flexibility is thought to be crucial in the binding of ligand molecules to certain locations. It is crucial for proteins to bind ligands.Between ionic groups with opposing charges, the highest electrostatic contact in proteins takes place. These noncovalent connections, also known as salt bridges, are important mainly in areas of the protein where water is not present due to the energy needed to remove water molecules from ionic groups close to the surface. Bridges made of salt have been seen to aid in the interactions of nearby subunits in complex proteins.

The weaker electrostatic interactions are the same way. They are important for interactions between subunits, inside the folded protein, and with ligands. Each polypeptide chain in proteins with several polypeptide chains is referred to as a subunit. Regions of the protein that lack water are known as ligand-binding pockets. A protein's inside and outside both create a considerable amount of hydrogen bonds. Polar amino acid side chains may interact with water or the polypeptide backbone in addition to creating hydrogen bonds with one another. The development of hydrogen bonds with other species is once again prohibited by the presence of water.Chemical processes that modify a polypeptide's structure before, during, or after synthesis result in covalent bonds. The disulfide bridges that are present in many extracellular proteins are the most noticeable covalent connections in tertiary structure. These robust connections shield protein structure from harmful pH or salt concentration variations in extracellular settings. Because there are substantial amounts of reducing agents in the cytoplasm, intracellular proteins do not have disulfide bridges. Water intake. Structured water is a significant stabilizing component of protein structure, as previously discussed. The adaptability needed for biological activity is further aided by the dynamic hydration shell that develops around proteins.

CONCLUSION

In conclusion, proteins are among the living world's most complex and important molecules. Their amazing structural complexity results from a range of organisational levels, from elementary to quaternary structures, each of which is essential to their operation. The visualisation of secondary structural components including -helices and -pleated sheets, as well as space-filling models, ribbon models, and other approaches have all been investigated as means to describe and comprehend protein structures. Protein research goes beyond only looking at their physical makeup. Homologous polypeptides, which have similar amino acid sequences and a shared heritage, provide important information on the evolution of species. Understanding the links between species has benefited by the extraordinary conservation of

certain proteins, such as cytochrome c, across a variety of creatures. Proteins' overall three-dimensional conformation is greatly influenced by their secondary structure, which is composed of -helices and -pleated sheets. These structural patterns contribute to the many forms and functionality of proteins in live organisms and provide the groundwork for understanding the roles of numerous fibrous proteins. Water molecules in the hydration shell play a key role in protein folding, a dynamic process that enables proteins to acquire their physiologically active conformations. The folding process is driven by the complex interplay of covalent bonds, electrostatic interactions, and hydrophobic interactions. In conclusion, research into proteins covers a wide range of topics, including their various structures, evolutionary importance, involvement in hereditary illnesses, and the complex folding mechanisms that underpin their activities. Scientists are still fascinated by these extraordinary biomolecules because they help them better grasp the basic functions of life.

REFERENCES:

[1] N. Wang, E. H. Xia, and L. Z. Gao, Genome-wide analysis of WRKY family of transcription factors in common bean, Phaseolus vulgaris: Chromosomal localization, structure, evolution and expression divergence, *Plant Gene*, 2016.

[2] M. M. Babu, The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease, *Biochemical Society Transactions*. 2016.

[3] D. Mrozek, P. Daniłowicz, and B. Małysiak-Mrozek, HDInsight4PSi: Boosting performance of 3D protein structure similarity searching with HDInsight clusters in Microsoft Azure cloud, *Inf. Sci. (Ny).*, 2016.

[4] A. Divan and J. A. Royds, 4. Proteins, *Mol. Biol. A Very Short Introd.*, 2016.

[5] T. Klose *et al.*, Structure of faustovirus, a large dsDNA virus, *Proc. Natl. Acad. Sci. U. S. A.*, 2016.

[6] B. Carpenter, R. Nehmé, T. Warne, A. G. W. Leslie, and C. G. Tate, Structure of the adenosine A2A receptor bound to an engineered G protein, *Nature*, 2016.

[7] Z. M. March, O. D. King, and J. Shorter, Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease, *Brain Research*. 2016.

[8] H. A. Zeinabad, A. Zarrabian, A. A. Saboury, A. M. O. Alizadeh, and M. Falahati, Interaction of single and multi wall carbon nanotubes with the biological systems: Tau protein and PC12 cells as targets, *Sci. Rep.*, 2016.

[9] Y. Zhang, J. K. Weber, and R. Zhou, Folding and Stabilization of Native-Sequence-Reversed Proteins, *Sci. Rep.*, 2016.

[10] D. T. Capraro and P. A. Jennings, Untangling the Influence of a Protein Knot on Folding, *Biophys. J.*, 2016.