Circular dichroism and its uses in biomolecular research - A Review

Sangeeta 1 and Anu Radha Pathania 2

1,2 Department of Chemistry, University Institute of Sciences, Chandigarh University Gharuan-140413, Mohali, Punjab, India

Abstract: The higher-order structure of proteins as well as their thermal stability can be determined using the circular dichroism (CD). CD is a common approach for swiftly assessing binding, secondary structure, and folding properties of proteins. In a nutshell, circular dichroism is an absorption spectroscopy technique that employs circularly polarized light to explore structural properties of optically active chiral compounds. Biological molecules, as well as their interactions with metals and other compounds, are studied extensively. Circular dichroism is becoming more widely acknowledged as a useful technique for studying the various conformations taken by proteins and nucleic acids in solution. Because CD is a quantitative approach, it can be used to track protein denaturation and protein-ligand interaction. These CD measures will have two key advantages: they can be performed on small amounts of material in a physiological buffer, and they will provide one of the greatest methods for monitoring any structural changes that occur as a result of changes in environmental conditions. It has proven possible to generate proteins on a big scale for therapeutic reasons utilizing recombinant DNA technology. Circular dichroism is also well-known as a useful method which is used for determining the folding characteristics of proteins. CD is used to see if a purified, produced peptide is either bended or if it has a mutation that impacts its strength and confirmation. The basic steps in getting this CD data, as well as the methodologies for interpreting the spectra in order to predict the protein structure, are summarized in this article. However, many researchers' value is harmed when they use circular dichroism, either because of poor experimental design or because of insufficient data. The essential steps in getting this CD data, as well as the methodologies for interpreting the spectra in order to predict the protein structure, will be summarized in this article. However, the value of many investigations using circular dichroism is harmed due to insufficient attention to critical components of instrument calibration or sample characterization.

1. Introduction

The circular dichroism (CD) approach is a spectroscopic method used for the resolution of the secondary structure composition of the proteins present in a solution. Because as we know in case of the electronic transitions of the polypeptide backbone the peptide bonds that are present in different conformations will give various spectra of absorption in the vacuum UV and far UV wavelength regions, both of these are unique and independent linearly. As a result, CD data can be utilized to deduce the secondary structure components type that are found in a solution's net spectrum [1]. This spectroscopy is one of the common spectroscopic techniques which analyses the difference in absorption that is present between the circularly polarized light on right and left sides. The peptide bonds in proteins will mostly contribute between 170 and 250 nm to the far-UV CD spectrum. The chromophore at these wavelengths is a peptide bond, and the signal is generated when this is positioned in a folded, regular context. It was discovered almost 50 years earlier that these various proteins may have varied spectral profiles [2] based on their environment and conformation of the peptide bonds, as well as the pattern of their H-bonding, which is indicative of the proteins' secondary structure. The traditional benchtop CD instruments have become widely available as instrumentation has progressed [3]. The wavelength range of conventional CD data
Circular dichroism (cCD) recorded on the laboratory-based devices is typically between 190 and 300 nm [4]. The information that is contained in this CD spectra can be thought of as a collection of distinct individual spectra that is arising from each sort of the secondary structure seen in the sample of protein. The number of various secondary structural components that can be investigated will be determined by the wavelength range of spectra available [5].

Each of the alpha-helix, beta-sheet, and random coil configurations will produce a distinct amplitude and form of the CD spectrum. The signals recorded in the near-UV area will be almost negligible if the protein preserves its secondary structure but lacks a three-dimensional well-defined structure. Significant near-UV signals, while on the other hand, are solid indications that the protein has been folded into a three-dimensional structure. Because the n to π* and n to π transitions of the amide groups dominate the spectra of such molecules in regions like far ultraviolet (UV), and are impacted also by geometries of the polypeptide backbones, and their spectra are indicative of the different forms of secondary structure present [6]. As a result, we now know that circular dichroism has far more applications other than only simply calculating the structure of proteins.

For example; It is also known as a great tool for determination of the kinetics and thermodynamics of protein folding, as well as its denaturation and the mutations having effects on protein stability and folding properties. We are unable to provide a comprehensive assessment of all contemporary CD spectroscopy applications.

For example, over 1500 research papers have been written in Journal Biological Chemistry in the last ten years, in which the CD has been utilised to examine various areas of nucleic acid and protein chemistry [7]. As a result, this evaluation will concentrate on a few specific examples that will demonstrate the CD's broad utility as an analytical tool.

Following themes have been reviewed:

- Interconnection between proteins and their ligands.
- Transitions in conformation.
- Folding of proteins thermodynamics.
- Intermediates are folded.
- In a model peptide, the helices unfurl.
- Proteins in their 'molten globule' form.
- Folding of proteins is a complex process.
- DNA conformational polymorphism and circular dichroism.

2. Literature review

The rapid expansion of the area, including greater knowledge of the principles and development of novel applications, has been seen in the four years following publication of the first edition of the circular dichroism [8]. For many years, circular dichroism has been recognized as a valuable approach for analyzing the secondary structure of proteins. CD spectroscopy is a well-accepted method used for evaluating the dynamics and folding of the proteins. CD spectroscopy offers a lot of potential for drug discovery because of its ability to detect conformational changes very sensitively [9]. CD has the benefit of being able to perform assays on several samples comprising 20 μg or less of proteins present in physiological buffers even in a matter of hours. This will not, however, be able to provide residue-specific details that may be gathered via X-ray crystallography or nuclear magnetic resonance.

With contributions from dozens of the world's finest practitioners and researchers in chirality, biochemistry, chirality, biochemistry, analytical chemistry, luminescence, and vibrational spectroscopy, the second edition will maintain pace with this extraordinary progress. The second edition of Circular Dichroism provides an important insight into the vast potential of the CD and bridges the gap between theory and practice with the addition of nine fully new chapters and significant changes to the current material. The paper will begin with a review of historical developments before immediately moving on to exciting reporting on contemporary advances and new developing topics in CD. The CD's coverage has been revised;

- Theory of the VOA
- CD applications in solid state
- Polymer CD that has been induced
- CD measurements with a high time resolution
- CD containing nucleic acids
- Interactions between DNA and drugs and between DNA and proteins
- The CD has been applied to a number of significant medicinal molecules [10].

2.1 Circular dichroism's applications in protein and peptide analysis include

2.1.1 Protein - ligand interaction

Circular dichroism is a quantitative approach that can be practiced to track the interconnection between proteins and their ligands as well as their denaturation. Since the late 1960s, when the
instrumentation became available, it has been employed to examine this relationship [11]. CD can be used to track the ligands' interactions with the peptides, proteins, and nucleic acids that have been delivered. This ligand binding must be dissymmetric, taking into account the extrinsic optical activity of the attached ligands, particularly in chromophores [12].

**Types of the protein-ligand interaction:**

There are numerous ways for a protein to bind to a ligand or another protein. Here are some of the methods described below:

2.1.1.1 Protein-ligand binding that is reversible.

2.1.1.2 Protein-ligand interaction that is complementary.

2.1.1.3 Protein-ligand interaction that is cooperative.

2.1.1.4 Protein with an allosteric function [13].

**Extrinsic effects:** Extrinsic cotton effects occur when a highly chromatic ligand binds to proteins in a dissymmetric manner in the region of their absorbance spectrum, and these effects are referred to as extrinsic cotton effects [14]. In most varied systems, this effect is employed to research protein-ligand interactions. Example: It is used to research cofactor binding to proteins, such as nucleotide binding to the lactic dehydrogenase [15], ATP binding to GroEL [16], pyridoxal phosphate binding to the tryptophan synthase, and retinol tie up to the interphotoreceptor retinoid (binding protein) [17].

**Intrinsic effects:** Ligand binding to proteins can frequently result in conformational changes. The conformational change results in a difference in the inherent ellipticity of amides backbone, which will be utilised to calculate the binding constant of ligands. For example; Divalent cations will act as effectors in most of the biological systems, which frequently include substantial conformational changes in target proteins. (In the same way that calcium ions attach to muscle regulating proteins) [18].

2.1.2 **Conformational transitions**

While circular dichroism is typically employed to examine transitions from a coiled to an uncoiled state of proteins, it is recently discovered that some systems can undergo transformation from an alpha-helical to the beta-pleated sheet arrangement. Protein oligomerization, precipitation, and aggregation may occur as a result. These structural alterations will have significant clinical implications. Some of these transitions have been discovered to be implicated in the pathophysiology of a few disorders, including Alzheimer's disease and 'mad cow disorder. CDs have been utilized to track conformational transformation in prions [19], amylin, and the Alzheimer's beta-amyloid peptide [20]. This amyloid peptide has also been found in the pancreatic islets of diabetic patients [21].

2.1.3 **Protein folding thermodynamics**

Obtaining the thermodynamics of folding by measuring the differences in the optical efficiency of protein molecules and polypeptides as a function of temperature. Indeed, one most common use of the CD currently is to determine thermodynamics of protein binding and effects of mutations on folding of protein and its stability [22]. Kelly and the Prince [23] recently evaluated the CD's applicability in protein folding research.

Freire has studied yet another equation that will characterize the cooperative folding behaviour of proteins as well as the thermodynamics of structural stability.

2.1.4 **Folding intermediates**

If peptides or proteins coil or uncoil, there is never always a one-to-one transformation between the native and completely disarranged states; often intermediate states are there. When a protein is subjected to denaturing conditions, like it is given high temperature, and is exposed to urea either, or detergents, such states are observed during kinetic studies of unfolding and folding and are temporary or may be stable intermediates [24]. These states are partially folded with molten globules, which are designated tertiary structures. Ptitsyn and his collaborators [25] coined the word. The CD method is frequently used to quantify and detect intermediate stages. Some molten globule folding stages, for example, have strong CD signals into the far UV regions due to structural components which are secondary such as alpha helices. While the near UV CD strands generated by structural interactions of the tertiary aromatic groups, which are frequently detected in the native protein spectra, are essentially absent in these molten globule folding stages. Alpha-lactalbumin is the most well-known example of a partially folded molten globule protein. The contribution of different states to protein unfolding is frequently useful to quantify [26].

CD's role in protein unfolding:

Proteins are most commonly unfolded by adding denaturants like urea or guanidine chloride, or by heat. In some processes, unfolding of protein can be
achieved by dilution simply. Differences in CD strands into the far-UV and the near-UV may also be utilised to alter loss of structure during unbinding; UV absorbance and fluorescence are further spectroscopic metrics that can be measured [27]. The available information from investigations of protein unfolding can include things like the denaturant content or the temperature at which 50% of the unfolding has happened. These are some metrics that will be used to compare the strength of different tertiary structures of proteins (or mutations of the same protein) [28]. The unfolding of the protein is thought to occur in a series of reversible, sequential unimolecular processes in which the kinds of distinct structures are lost. More recent research has found that equilibrium unfolding intermediates do not preserve an entire domain that is N-terminal, but have features like the molten globule rather [29].

2.1.5 Helix unfolding in model peptides

CD measurements will provide crucial information on the unfolding mechanism. The peptides investigated may range in length almost from 14 to 50 residues and it was based on the reshowing units of the succession Ala-Glu-Ala-Ala-Ala-Lys-Ala. These characteristics have been linked to the formation of monomeric helices in mixture, absence of long –range tertiary contacts, and hydrophobic cores.

2.1.6 Proteins having ‘molten globule’ state

Few convincing examples of protein structural state transitional in between unfolded and aboriginal form will come from CD investigations. Some other features of this kind of structure that was described include absence of dispersion of the chemical shift in case of NMR and increased tie up of the ANS fluorescent dye. This state could act as critical steps in folding pathways of protein [28].

2.1.7 Folding of proteins

Considerations on protein folding in general; Understanding the procedure of folding of the protein is one of the greatest issues in molecular biology, as we already know. Since Levinthal pointed out about 30 years ago that folding a protein required sampling of all conformations possible which are available to it, this approach may take an inordinate amount of time, this has been evident that folding of peptides must involve specialized passageways [2].

A) This is crucial to demonstrate that the peptide is thoroughly uncoiled prior to dilution.

B) Denaturant's remaining quantity after the dilution has no outcome on the protein.

It's crucial to investigate the time scale of rebinding from its denatured state, as well as to match up the data with rate of resumption of biological activities, to build a rebinding mechanism. The re-establishment of the secondary structure and protection in opposition to amide hydrogen exchange are frequently checked by far-UV CD, whereas the re-establishment of the structure which is tertiary are monitored by fluorescence and by near UV CD. The fast cross-linking with glutaraldehyde will be used to track the quaternary structure re-establishment [25].

2.1.8 DNA conformational polymorphism and circular dichroism

Double-stranded DNA molecules hold genetic information in all familiar autonomous structures of life, like bacteria to humans. This is a large fragment in humans, measuring hundreds of millions of base pairs in length. Massive DNA fragments are crammed into the comparably confined area of the cell kernel in eukaryotes. The B-type right-handed double helix was identified in 1953 as the basic configuration of DNA. Later, it was discovered that DNA may adopt alternative structures, such as the right-handed A-type of double coil representative of RNA, left-handed Z-DNA, other types of slipping structures and hairpins, triplexes, guanine quadruplexes, and other less well-characterized structures [29].

2.2 The B-form, the hairpin, and DNA denaturation

This B-form is one of common DNA conformations. Sequentially conservative heterogeneous DNA's CD spectrum at long wavelengths. The pairs of bases are perpendicular in B-form to the axis of double-coil, giving the molecules a few chirality and resulting in low peak intensities. The negative band will become deeper as the content of (A+T) grows, and conformational variability will increase [30]. The CD spectra of synthetic poly deoxynucleotides' B-forms are determined by the main sequence.

The family of B-DNA forms will be made up of these structures. All of these are distinguished by a long wavelength band that is positive between 260 and 280 nm whereas negative between 245 and 265 nm [31]. The amplitude and locations of the CD bands, on the other hand, will vary significantly with the series, chromophores differ not only because of that, but also due to differing properties of conformations [32]. A hairpin is an intramolecular variation of the B-type of DNA. It has a single-stranded loop in addition to the double-helical stem. The recurrent expansions associated with a number of ailments, most notably neurodegenerative disorders, are
2.3 The B-A transformation and the A-form

The A-type of RNA is a fundamental structure. DNA will take on the A-type in watery circumstances like ethanol. All DNA molecules do not acquire this A-shaped conformation. The CD range of A-type DNA with a diverse primary structure is nearly identical to that of A-form RNA or a DNA/RNA blend with the same primary structure [33].

The B-A transition is a two-step, collaborative procedure. This causes isodichroic spots in CD range and an S-shaped CD dependency on inducing substances including methanol, ethanol, and trifluoroethanol [34].

2.4 The B-Z transformation and the Z-form of DNA

The Z-type of DNA has been studied extensively, and it, like the regular B- and A forms, has been included in textbooks [35]. In comparison to the B- and A-forms, the pairs of bases in the Z-DNA double coil will face the backbone in the opposite direction [36].

2.5 Triplexes

Unlike A-form, Z-form, and triplexes, DNA triplexes do not have a distinct CD spectrum. The diverse CD spectra will be provided by triplexes generated by varied sequences [37]. It will be quite straightforward to detect the characteristic CD spectrum of the investigated triplex of a given succession when used in conjunction with gel electrophoresis; this is a clearly unique spectrum from that of its individual strands. As a result, CD spectroscopy is well-known as a tool for identifying triplex-forming oligonucleotides and studying triplex formation. The duplex DNA will be bound by these oligonucleotides in a sequence-specific manner [38].

2.6 Quadruplexes of cytosine

The quadruplexes are formed by DNA strands that are high in cytosine. In an antiparallel orientation, the two duplexes are intercalated. Therefore, such structures are called the intercalated structures or i-tetrplex. The positive dominating band at 290 nm of these cytosine quadruplexes will produce a distinctive CD spectrum [39]. The pH, which is slightly acidic in nature, promotes the creation of these cytosine quadruplexes. Even at pH values greater than 7, series that are prone to folding into quadruplexes of cytosine will tie up to the proton in order to construct the structures. Slow kinetics, analogous to guanine quadruplexes, are used to create intermolecular cytosine quadruplexes [40].

2.7 Guanine quadruplexes

All guanine quadruplexes are based on guanine tetrads and come with a difference of sizes and shapes. Guanine quadruplexes are divided into two categories. Antiparallel quadruplexes' spectra will have a positive band at 290 nm and a negative band at 260 nm, whereas parallel quadruplexes' spectra will have a positive band at 290 nm and a negative band at 260 nm [41]. At a wavelength of 210 nm, both of these quadruplexes will have a positive peak and an additional feature. Antiparallel and parallel quadruplexes have different CDs due to the various stacking interactions of guanine residues that are positioned differently all over their glycosidic linkages. Transitions between different types of quadruplexes and intermolecular guanine quadruplexes will be gradual. Physiologically relevant cations, such as potassium, trigger the formation of quadruplexes, however ethanol is still a more potent inducer. At 210 nm, the quadruplexes, on the other hand, will show a positive band [42].

2.8 Guanine and adenine-rich DNA fragments

Bits of DNA high in adenine and guanine will show cooperative melting of conformers, which is different from the traditional structures. Despite numerous attempts, these structures have yet to be solved [43]. The anti-parallel homoduplex with guanine adenine pairs is the first conformer. The zinc cations, which are divalent, are primarily responsible for stabilizing this duplex. The various conformers are induced by low zinc cation concentrations. This antiparallel homoduplex is most likely a component of the zinc-stabilized triplex [44]. The CD spectrum of the second conformer of the GA succession, which is a parallel homoduplex with the guanine guanine and adenine adenine pairings, will be identical to that of the parallel quadruplexes of guanine but with smaller positive amplitudes [45]. An orderly single strand containing protonated adenine is the third conformer. It has a CD spectrum that is nearly identical to homoduplex. At increased ionic strength, this single strand will become a homoduplex. The progressive deepening of the negative band reflects the uncooperative nature of this process. This means that guanine-guanine stacking will not alter considerably,
and duplex formation will be dominated by adenine-adenine interactions [46].

2.9 Condensation of DNA

Light above 300 nm will not be absorbed by DNA that includes natural bases. A signal greater than 300 nm suggests that the DNA has been condensed into light-scattering atoms. Because CD modifications can no longer be explained in terms of secondary DNA structure changes, this is the part of the spectrum that has to be investigated. The CD, on the other hand, can be used to investigate these species of condensed, regularly packed DNA [47]. High ionic stability in polyethylene glycol (PEG) and other alcohols, as well as basic peptides, can promote condensation. Such conditions will cause the production of small non-B DNA segments, which will propagate as a double helix perturbation, resulting in cooperative condensation. Proteins like histones and protamines, which are DNA binding proteins that regulate and manage the ionic stability and hydrophobic environment in close proximity to the DNA molecules, can cause condensation and disintegration in vivo [48].

2.10 Basic principle of the CD

The CD's basic principle is the differential absorption of the plane-polarized light radiations circularly polarized right and left components [49]. When a chromophore is placed in an asymmetric environment or when the chromophore is covalently connected to any of the chiral centres present in a molecule, this type of effect will occur due to the structure of the chromophore becoming chiral (optically active) [50].

Discussion

The major goal of this brief review was to outline the features of CD spectroscopy, which is a very helpful approach for studying a variety of DNA structural aspects. CD is an excellent supplement to NMR and X-ray crystallographic research because it provides features that these methods miss. A CD may be employed to investigate lengthy DNA fragments, which DNA in actuality, a CD can be employed. It can be used to quickly assess the impact of changing environmental variables on DNA behaviour and track conformational transformations in real time. The isomerization of DNA between the various conformers, as well as the factors that govern these isomerization’s, are important because they will very certainly be used to regulate gene expression.

● CD spectroscopy is a strong tool for researching DNA's structural characteristics. However, a comprehensive comprehension of the spectral shifts necessitates tracing the full spectral region. We must evaluate all CD variations caused by solution factors (such as pH, temperature, and salts) that lead to a given DNA layout, rather than making structural assumptions based on a single spectral value.

● Following only one band while disregarding the rest of the spectral regions will lead to inaccurate conclusions.

● The characterization of the structure will be aided by the course of denaturation and the appearance of the structure.

● These CD spectroscopic approaches are crucial for discovering the DNA's essential characteristics. CD has a long history of being a pioneering tool for studying new DNA structures.

● The CD spectroscopy, for example, is used to show the numerous sorts of arrangements present in DNA fibers (B, C, D, T), all of which belong to the broad B-family of the structures.

Z-DNA which is left-handed was discovered via CD spectroscopy, which was utilised to identify uncommon X-forms with a molecular structure that is still up for debate.

● This is a technique for analyzing and mapping abnormal configurations in specific genomic DNA regions that are known to be medically or biologically important.

● CD and other methods for protein structural examination show the secondary and tertiary forms of proteins. It also goes over the secondary and tertiary DNA forms found in the linear genomic primary succession, which have been shown to play a key role in both normal and abnormal DNA genome functioning.

● Recent DNA structure-function experiments have focused on quadruplexes.

These genomic areas, particularly gene promoters, have the potential to create quadruplexes, which are considered to occur in vivo and have an impact on cellular activities.

One of the most successful methods for analyzing the formation of quadruplexes in distinct DNA molecules is CD spectroscopy.

● In the case of telomeric sequence DNA, the development of quadruplexes is hypothesized to be linked to cancer and fluorescence.
Understanding the structure of quadruplexes and their capacity to monitor the dissociation and creation of such structure is critical for rational medication outline.

This CD range will also offer the support for a hypothesis that a lengthy series of telomeres has a structure similar to nucleosomes, with beads on strings.

References


