Biochimica et Biophysica Acta xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230 nm

Yang Wei, Aby A. Thyparambil, Robert A. Latour *

Department of Bioengineering, 501 Rhodes Engineering Research Center, Clemson University, Clemson, SC 29634, USA

ARTICLE INFO

Article history:
Received 3 July 2014
Received in revised form 27 September 2014
Accepted 1 October 2014
Available online xxxx

Keywords:
Circular dichroism spectroscopy
Empirical algorithm
Protein helical structure
Denaturant

ABSTRACT

Conventional empirical methods for the quantification of the helical content of proteins in solution using circular dichroism (CD) primarily rely on spectral data acquired between wavelengths of 190 and 230 nm. The presence of chemical species in a protein solution with strong absorbance within this range can interfere with the ability to use these methods for the determination of the protein's helical structure. The objective of this research was to overcome this problem by developing a method for CD spectral analysis that relies on spectral features above this wavelength range. In this study, we determined that the slopes of CD spectra acquired over the 230 to 240 nm region strongly correlate with the helix contents including α-helix and 3_{10}-helix of protein as determined using conventional CD algorithms that rely on wavelengths between 190 and 230 nm. This approach (i.e., the 230–240 nm slope method) is proposed as an effective method to determine the helix content within proteins in the presence of additives such as detergents or denaturants with high absorbance of wavelengths up to 230 nm.

1. Introduction

Circular dichroism (CD) has been extensively used to spectroscopically study the structure of biomolecules in solution and when absorbed to surfaces due to its characteristics of being non-destructive, relatively easy to perform, requiring small sample volume, and providing fast, reliable data analyses [1,2]. In particular, CD provides a very convenient experimental method for the determination of the secondary structure and environmentally induced structural changes in proteins since the different forms of the primary secondary structural elements found in proteins (e.g., α-helix, β-sheet, and random loop) exhibit distinctly different CD spectrum [3].

Most algorithms that have been developed for secondary structure determination of proteins by CD depend on the analysis of spectral features in the far UV range, primarily from 190 to 230 nm. Over this spectral range, the amides within the secondary structural components constituting a protein strongly absorb circularly polarized light and undergo varying extents of $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions for a given wavelength [4]. The CD spectrum (Fig. 1) for a pure α-helical structure acquired between 190 and 250 nm exhibits a characteristic double minimum at 208 nm ($\pi \rightarrow \pi^*$) and 222 nm ($n \rightarrow \pi^*$), and a stronger maximum at 191–193 nm ($\pi \rightarrow \pi^*$) [4]. Similarly, β-sheet structure exhibits a characteristic minimum at 215 nm ($n \rightarrow \pi^*$) and a maximum at 198 nm ($\pi \rightarrow \pi^*$) [4]. In contrast to these spectral features, random coil segments of protein tend to exhibit a maximum and a minimum that is essentially opposite from the minimum and maximum of the α-helical and β-sheet structures [5].

To quantify the relative proportion of each associated secondary structure contained in a protein sample, the resulting CD spectrum acquired between wavelengths of 190 and 240 nm is typically empirically interpreted as a sum of fractional multiples of reference spectra for each type of secondary structure [3]. This process is conducted using a variety of mathematical tools [6] along with reference datasets of highly resolved protein structures (i.e., protein structures from X-ray crystallography and NMR spectroscopy) [7]. As a result, these conventional algorithms cannot be used if the protein solution to be analyzed contains chemical species that strongly absorb at wavelengths below 230 nm, as commonly seen with various detergents and denaturants [8–11]. For example, Fig. 2 presents plots of the background absorbance for solutions containing various chemical additives using a 0.1 cm cuvette, which show saturating levels of absorbance (see for wavelengths below 225 nm). When this situation occurs, other methods are required for protein structural analysis [1,10,12]. This type of absorbance problem becomes increasingly problematic as the pathlength of the cuvette is increased, with pathlengths up to 1.0 cm being commonly used for temperature and titration experiments [13–23] and for the analysis of adsorbed proteins on nano-particles or flat material surfaces in order to provide sufficient signal strength for analysis [24–32].

Subsequently, conventional algorithms that rely on CD spectra over the range of 190–230 nm can only be used when the additives are present under extremely dilute conditions, thus greatly limiting the ability to investigate the influence of such additives on a protein's secondary structure.

* Corresponding author.
E-mail address: LatourR@clemson.edu (R.A. Latour).

http://dx.doi.org/10.1016/j.bbapap.2014.10.001
1570-9639/

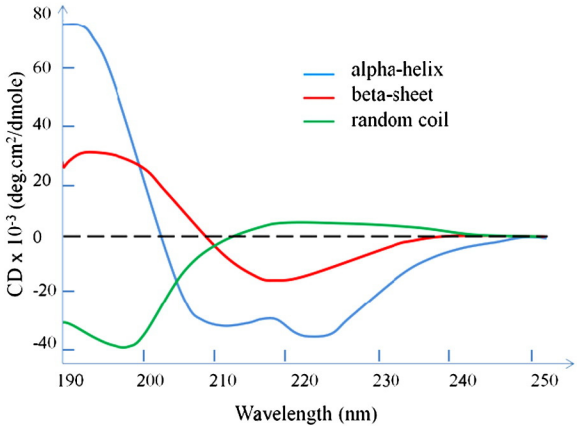

Fig. 1. Standard CD spectra redrawn from Corrêa et al. [5]. Each of the three basic secondary structures of a polypeptide chain (α-helix, β-sheet and random coil) shows a distinctly different characteristic CD spectrum [5].

To overcome this problem, a CD cell with shorter path-length is commonly used to minimize the background absorbance (See S.1. in the supporting information) [34]. Alternatively, if CD data from wavelengths above 220 nm are available, the helical content of protein can at least still be estimated by choosing a single wavelength above 220 nm where the difference in signal between a folded and unfolded protein is large (unfortunately there are no analogous methods for estimating β-sheet or random coil structure) [35]. For example, CD molar ellipticity data at 222 nm are commonly used for quantifying the helical content of protein (i.e., the 222 nm method), where helical structure exhibits a characteristic minimum in ellipticity (see Fig. 1) [5]. Even higher wavelengths than 222 nm, such as 225 nm [10] or 228 nm [35], have also been used for estimating helical content when the background absorbance has influenced the CD response at 222 nm. The situation, however, becomes particularly problematic when samples exhibit strong background absorbance all of the way up to 230 nm even when using CD cells of short path length (See S.1. in the supporting information). In this case, there are presently no existing methods that can be used for the quantitative analysis of even the helical structure of a protein when in solution.

While seeking for alternative methods for the analysis of CD spectra for the determination of protein helical structure in the presence of strongly absorbing additives [24,36], we observed that a linear region of varying slope typically occurs in CD spectrum between 230 nm and 240 nm. Based on this observation, we hypothesized that if the relative change in the molar ellipticity values between 230 nm and 240 nm was primarily caused by the helical structure of the protein, then the relative change in the slope, which can be simply derived from multiple CD points over this wavelength range, may provide a sensitive method of estimating the helical content of protein in solution when the background absorbance occurs for wavelengths all of the way up to 230 nm.

The purpose of current study was therefore to investigate if a linear correlation exists between the slope of CD spectra over the range of 230–240 nm and protein fractional helicity determined by existing methods for a range of proteins and their conformational states in aqueous solution. The specific objective of this research was then to use this correlation (if found) to provide a method (i.e., the 230–240 nm slope method) that could be used to reliably quantify the helical content in proteins in solution with backgrounds exhibiting strong absorbance up to 230 nm.

2. Analytical model

The 222 nm wavelength method for CD analysis of the helicity of protein structure uses the molar ellipticity CD data at 222 nm [7], which is the wavelength corresponding to the characteristic minimum of the CD spectrum of the helical structure of protein (see Fig. 1). Accordingly, the fractional helicity (*FH*) of a protein in solution can be estimated from the CD response at 222 nm, and similarly at other nearby wavelengths such as 225 and 228 nm [10,35], by Eq. (1): [5,37]

$$FH = \frac{\left(\theta_\lambda^{\exp} - \theta_\lambda^u\right)}{\theta_\lambda^h - \theta_\lambda^u} \quad (1)$$

where $\theta_\lambda^{\exp}$ is the experimentally observed mean residue ellipticity (usually given in deg cm^2 dmol^{-1}) for a given wavelength (λ), and θ_λ^u and θ_λ^h correspond to the ellipticity for a protein with 0% and 100% helical content at wavelength λ, which are typically experimentally or theoretically estimated to be −3000 and −39,500 deg cm^2 dmol^{-1}, respectively, for a λ of 222 nm [5,38,39].

Eq. (1) can be rearranged to generally express $\theta_\lambda^{\exp}$ as a function of *FH* as designated by Eq. (2):

$$\theta_\lambda^{\exp} = FH(\theta_\lambda^u - \theta_\lambda^u) + \theta_\lambda^u. \quad (2)$$

Accordingly, by extending this relationship over a linear region of a CD spectrum in the region of 230–240 nm [24,36], Eq. (2) can be

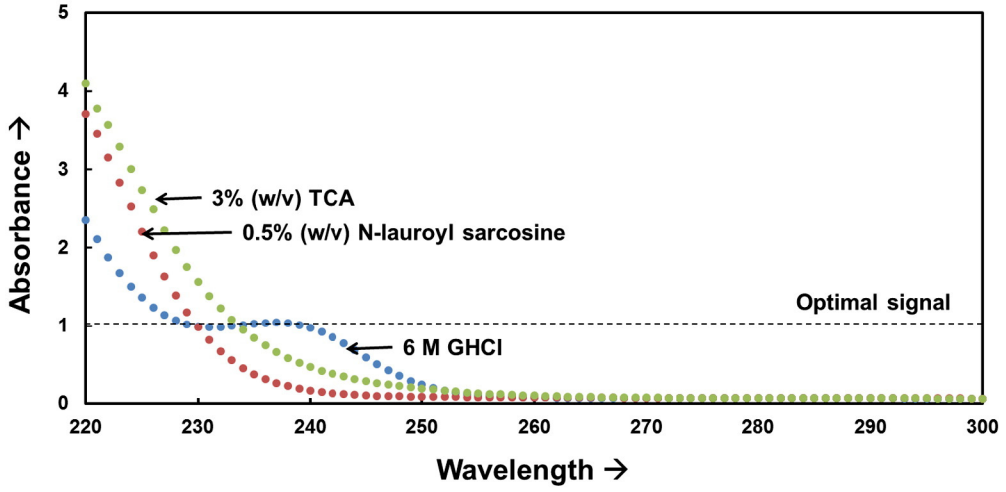

Fig. 2. Effective absorbance spectra for different chemical additives (3% (w/v) trichloroacetic acid (TCA, Sigma: T6399), 0.5% (w/v) N-lauroyl sarcosine solutions (Sigma: T7414), and 6 M Guanidine hydrochloride (GHCl, Sigma: G3272)) in deionized water in a 0.1 cm pathlength cuvette. Strong absorbance (>1) results in signal saturation at high wavelengths, which can prohibit the use of conventional structural analysis algorithms that require CD signal sensitivity over the range from 190 to 230 nm [1,12,33].