AIMS Biophysics

AIMS Biophysics, 13(2): 162−188.
DOI: 10.3934/biophy.2026010
Received: 14 January 2026
Revised: 14 March 2026
Accepted: 03 April 2026
Published: 16 April 2026

http://www.aimspress.com/journal/biophysics

Review

Biophysical techniques for protein secondary structure analysis: a comprehensive review

Adil Ahiri*

Polydisciplinary Faculty of Safi, Cadi Ayyad University, Safi, Morocco

* **Correspondence:** Email: ahiri.adil@gmail.com; Tel: +212666648554.

Abstract: The determination of protein secondary structure constitutes a fundamental step in understanding protein function, folding mechanisms, and disease pathology. This comprehensive review presents a systematic methodological comparison of the principal biophysical techniques employed for secondary structure elucidation, spanning seven decades of technological development from Pauling and Corey's foundational discoveries to contemporary high-resolution methods. We examine spectroscopic approaches, including circular dichroism (CD) and Fourier-transform infrared (FTIR) spectroscopy, which provide rapid global quantification of secondary structure content with minimal sample requirements, alongside high-resolution techniques, including nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography, and cryo-electron microscopy (cryo-EM), which offer atomic-level structural details. For each technique, we analyze theoretical foundations, experimental protocols, deconvolution algorithms, accuracy benchmarks, and optimal application domains. Comparative analysis reveals that no single technique achieves optimal performance across all secondary structure classes: FTIR generally provides improved accuracy for β-sheet-rich systems, particularly when sheet topology discrimination is required. NMR uniquely provides residue-level resolution with dynamics information across multiple timescales, whereas crystallography and cryo-EM define atomic-resolution standards for static structures. Complementary techniques, including small-angle X-ray scattering (SAXS), hydrogen-deuterium exchange mass spectrometry (HDX-MS), and Förster resonance energy transfer (FRET), provide essential orthogonal information on global shape, protection patterns, and distance constraints. We present detailed decision frameworks for technique selection based on sample characteristics and research objectives, propose integrative multi-technique workflows, and identify critical gaps in current methodology, including the need for comprehensive multi-method benchmarking and standardized protocols for challenging protein systems such as membrane proteins and intrinsically disordered proteins. This review provides

practical guidance for researchers designing experimental strategies for protein secondary structure characterization in the modern era of integrative structural biology.

Keywords: protein secondary structure; circular dichroism; FTIR spectroscopy; NMR spectroscopy; X-ray crystallography; cryo-electron microscopy; α-helix; β-sheet; protein folding

1. Introduction

Proteins, the molecular machines of life, derive their biological function from precise three-dimensional arrangements of their polypeptide chains. Within this hierarchy, secondary structure occupies a pivotal position: it represents the first level of spatial organization beyond the covalent backbone and serves as the fundamental building block from which larger structural domains are constructed [1,2].

This review provides a systematic, comparative analysis of the major biophysical techniques for protein secondary structure determination (Figure 1). We examine the theoretical foundations, experimental methodologies, quantitative accuracy, and practical considerations for each approach. Our analysis emphasizes the complementarity of methods and presents decision frameworks to guide technique selection. By synthesizing foundational work with recent advances, we aim to equip researchers with the knowledge to design optimal experimental strategies for their specific structural questions in the context of modern integrative structural biology.

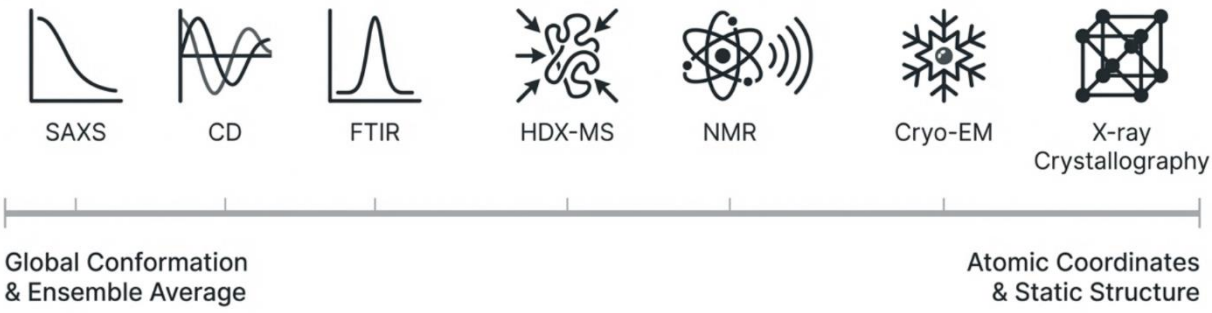

Figure 1. Spectrum of biophysical techniques for protein structure analysis. The techniques are arranged by type of structural information provided, ranging from methods yielding global conformation and ensemble averages (SAXS, CD, FTIR) to those determining atomic coordinates of static structures (NMR, cryo-EM, and X-ray crystallography).

It is important to note that these techniques operate at fundamentally different levels of structural resolution. Spectroscopic methods such as circular dichroism (CD) and Fourier-transform infrared (FTIR) spectroscopy provide global, ensemble-averaged fractions or ratios of secondary structure content for the entire protein, without localizing specific elements along the sequence. Nuclear magnetic resonance (NMR) spectroscopy, by contrast, provides residue-wise secondary structure assignments, mapping each amino acid to its local conformational state. X-ray crystallography and

cryo-electron microscopy (cryo-EM) yield complete three-dimensional structures from which secondary structure is determined at the atomic coordinate level. This hierarchy—from global content to residue-level mapping to full atomic structure—is a central organizing principle of this review and should guide the selection of techniques for specific research questions.

1.1. Definition and classification of secondary structure elements

Protein secondary structure describes the local conformation of the polypeptide backbone, determined by the patterns of hydrogen bonds between the backbone amide (N–H) and carbonyl (C=O) groups. The principal elements—α-helix, β-sheet, and various turn and loop configurations—are each characterized by specific backbone dihedral angles (φ and ψ) and hydrogen bonding patterns [3].

The α-helix is characterized by 3.6 residues per turn, a pitch of 5.4 Å, and intrachain hydrogen bonds between the carbonyl oxygen of residue i and the amide hydrogen of residue i + 4. The canonical α-helix occupies a well-defined region of Ramachandran space ($\varphi \approx -57°$, $\psi \approx -47°$), with side chains projecting outward from the helical axis in a characteristic pattern [4].

β-sheets are formed by the lateral association of extended β-strands through interchain hydrogen bonds. In parallel β-sheets, adjacent strands run in the same N–C direction, while antiparallel sheets feature alternating strand orientations. Antiparallel sheets exhibit more linear, and thus stronger, hydrogen bonds. The characteristic interstrand distance is approximately 4.7 Å, with side chains alternating above and below the sheet plane. β-sheets frequently display a right-handed twist that influences their overall topology and packing [5].

Turns and loops connect regular secondary structure elements and often participate directly in protein function, including enzyme active sites and binding interfaces. β-turns typically span four residues and involve hydrogen bonding between residue i and i + 3. While sometimes categorized as "irregular" structures, many loops adopt defined conformations stabilized by local interactions and are critical for molecular recognition [4].

1.2. Biological and biomedical significance

Secondary structure is intimately connected to protein function at multiple levels. The stability of the native state depends substantially on the integrity of secondary structure, as does the protein's susceptibility to misfolding and aggregation [6,7].

Beyond disease and therapeutics, secondary structure analysis supports protein engineering efforts, including directed evolution, rational design, and de novo protein design. Understanding how sequence changes affect secondary structure propensity informs stabilization strategies and enables the prediction of mutation effects. The recent success of computational protein design, exemplified by Rosetta and machine learning approaches, relies heavily on accurate secondary structure prediction and experimental validation [8,9].

2. Spectroscopic methods for global secondary structure quantification

Spectroscopic techniques provide a rapid assessment of overall secondary structure content in solution, making them invaluable for protein characterization, quality control, and studies of conformational change.

2.1. Circular dichroism spectroscopy

2.1.1. Physical principles and theoretical foundations

Circular dichroism (CD) arises from the differential absorption of left- and right-circularly polarized light by chiral molecules. Proteins are inherently chiral due to the L-configuration of their constituent amino acids; additionally, the spatial arrangement of peptide chromophores in secondary structure elements creates exciton coupling that produces characteristic CD signatures [10,11]. For proteins, the far-UV CD spectrum (190–260 nm) originates from electronic transitions of the peptide bond [12,13].

2.1.2. Characteristic spectral signatures

The α-helix produces the most distinctive and intense far-UV CD spectrum. The characteristic features include two negative bands at 222 nm (n→π* transition) and 208 nm (exciton-split π→π* transition), plus a strong positive band near 192 nm. The magnitude of $[\theta]_{222}$ is often used as a rapid indicator of helical content [14].

CD data are typically reported as mean residue ellipticity ([θ], in deg·cm²·dmol⁻¹), which normalizes the measured ellipticity by protein concentration, path length, and the number of residues. This normalization is essential for comparing spectra across different proteins and experimental conditions. The mean residue ellipticity is calculated as $[\theta] = \theta\_obs / (10 \times c \times l \times n)$, where θ_obs is the observed ellipticity in millidegrees, c is the molar concentration (mol/L), l is the path length in centimeters, and n is the number of residues. Accurate determination of protein concentration is therefore critical for reliable CD analysis.

β-sheets display a negative band between 215 and 218 nm and a positive band at 195–200 nm, though the precise positions and intensities vary considerably with β-sheet topology, twist, and strand length. The spectral features of β-sheets are generally less intense than those of α-helices, with $[\theta]_{216}$ typically ranging from −10,000 to −20,000 deg·cm²·dmol⁻¹ (Table 1). This reduced intensity and greater variability contribute to the somewhat lower accuracy of CD-based β-sheet quantification compared to α-helix estimation [15].

Disordered or "random coil" conformations produce a strong negative band below 200 nm (typically centered near 198 nm) with minimal ellipticity at higher wavelengths. It should be noted that the absolute values of ellipticity depend not only on the secondary structure content but also on instrumental factors (detector sensitivity, baseline calibration), sample concentration accuracy, and protein purity; aggregated or degraded samples may yield distorted spectra that do not reflect the true secondary structure composition.

Table 1. Characteristic far-UV circular dichroism spectral signatures of protein secondary structure elements. Spectral assignments are based on reference data by Greenfield [10], Kelly et al. [11], and Chen et al. [16].

Secondary structure	Negative band(s) (nm)	Positive band (nm)	[θ] range (deg·cm²·dmol⁻¹)
α-helix	222, 208	192	−30,000 to −36,000 at 222 nm
β-sheet	215–218	195–200	−10,000 to −20,000 at 216 nm
Disordered	< 200 (intense)	Minimal	−35,000 to −45,000 at 198 nm

An idealized α-helical CD spectrum is characterized by two well-resolved negative minima of approximately equal magnitude at 222 and 208 nm, and a large positive maximum near 192 nm, with the ratio $[\theta]_{222}/[\theta]_{208}$ approaching 1.0 for a fully helical polypeptide. A pure β-sheet spectrum shows a single broad negative minimum centered near 216–218 nm and a positive maximum at 195–200 nm, with markedly lower overall intensity compared to the α-helix. A fully disordered or random coil spectrum is dominated by a strong negative band below 200 nm (typically near 197–198 nm) with very low ellipticity above 210 nm. In practice, most protein CD spectra represent a superposition of these idealized components weighted by the fractional content of each secondary structure class (Figure 2).

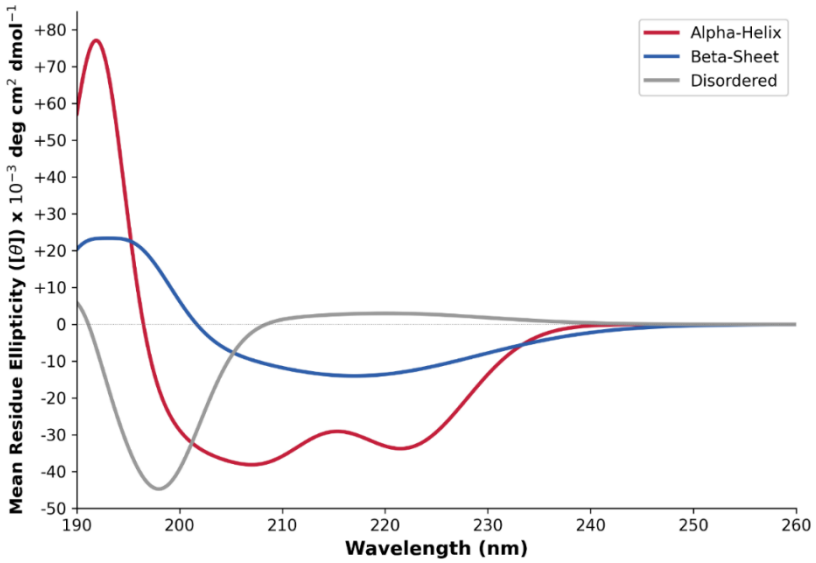

Figure 2. Idealized far-UV circular dichroism spectra of protein secondary structure elements. The α-helix (red) displays two negative minima at 222 and 208 nm and a large positive maximum near 192 nm. The β-sheet (blue) shows a negative band at ~216 nm and a positive band at ~195 nm. The disordered conformation (gray) is dominated by a strong negative band below 200 nm. Spectral patterns are based on reference data from Greenfield [10] and Kelly et al. [11].

2.1.3. Deconvolution algorithms and computational analysis

Quantitative estimation of secondary structure content from CD spectra requires deconvolution using algorithms that compare experimental spectra to reference datasets of proteins with known structures from crystallography. Several mathematical approaches have been developed, differing in their treatment of the basis set and fitting methodology [14,17].

The underlying principle of all CD deconvolution approaches is that the observed CD spectrum of a protein can be expressed as a linear combination of basis spectra corresponding to different secondary structure types, weighted by their fractional contributions to the total structure. The accuracy of deconvolution depends critically on the quality and diversity of the reference protein database, the mathematical method used for fitting, and the wavelength range of the experimental data (extending below 190 nm significantly improves accuracy). The most widely used algorithms are available through the DichroWeb online server, which provides standardized access to multiple methods and

reference datasets.

This approach treats the experimental spectrum as a linear combination of reference protein spectra, with regularization preventing overfitting. CONTIN has been widely used and provides robust results for many protein types, though it can struggle with unusual compositions outside the reference database range.

SELCON3, developed by Sreerama and Woody [14], employs singular value decomposition (SVD) with self-consistent selection of reference proteins. The algorithm iteratively selects a subset of reference proteins most similar to the target, improving accuracy for proteins that differ significantly from the average database composition.

CDSSTR, also by Sreerama and Woody [18], uses variable selection methods to identify optimal reference subsets for each analysis. By systematically testing combinations of reference proteins, CDSSTR can achieve improved accuracy, particularly for proteins with unusual secondary structure compositions. The computational cost is higher than simpler methods but remains practical for routine analysis.

The BeStSel (Beta Structure Selection) algorithm represents a significant advance in addressing the spectral heterogeneity of β-sheet structures [19]. Unlike earlier methods that treat β-sheets as a single structural class, BeStSel distinguishes parallel β-sheets, antiparallel β-sheets (further subdivided by twist angle), and mixed sheets. This differentiation is essential because the CD spectra of different β-sheet types vary substantially. BeStSel achieves Pearson correlation coefficients of 0.94 for total β-sheet content and 0.80 for α-helix, with root mean square deviations (RMSD) of approximately 6%–8% [20].

2.1.4. Practical experimental considerations

Successful CD experiments require attention to several practical factors. The product of concentration and path length should be adjusted to maintain absorbance below approximately 1.0–1.5 to avoid detector saturation and maintain linearity.

Temperature control is essential, as secondary structure content is temperature-dependent. Standard measurements are typically performed at 20 or 25 °C. Thermal denaturation experiments, monitoring $[\theta]_{222}$ as a function of temperature, provide information on protein stability and can reveal folding intermediates [21,22].

Several additional limitations should be considered when interpreting CD data. Sample turbidity and light scattering, arising from protein aggregation, large complexes, or membrane preparations, can severely distort CD spectra by introducing differential scattering artifacts that mimic or mask genuine CD signals. This represents a significant experimental failure mode, particularly for aggregation-prone proteins. Furthermore, CD deconvolution algorithms are inherently constrained by the composition of their reference databases; secondary structure types or topologies not represented in the reference set cannot be accurately predicted. For example, proteins with unusual β-sheet architectures or high turn/loop content may yield unreliable estimates if the reference database lacks structurally similar proteins. Users should therefore select reference datasets appropriate for their protein class and critically evaluate goodness-of-fit parameters.

2.2. Fourier-transform infrared spectroscopy

2.2.1. Theoretical foundations and vibrational modes

Infrared spectroscopy probes vibrational transitions of molecular bonds, with the amide bands of the protein backbone providing the primary structural information. Of these, the amide I band (1600–1700 cm^{-1}, predominantly C=O stretching) and amide II band (1500–1600 cm^{-1}, primarily C–N stretching coupled with N–H bending) are most informative for secondary structure analysis [23,24].

The sensitivity of amide I frequencies to secondary structure arises from two distinct physical effects. First, hydrogen bonding to the carbonyl oxygen affects C=O bond strength and hence vibrational frequency: stronger hydrogen bonds weaken the C=O bond, lowering the stretching frequency. Second, transition dipole coupling (TDC) between adjacent peptide groups in regular structures produces collective vibrational modes with characteristic frequencies. In β-sheets, the two-dimensional arrangement creates more complex coupling patterns [25,26].

The amide II band is less sensitive to secondary structure than amide I but provides complementary information. The combination of amide I and II analysis can improve quantification accuracy [27].

2.2.2. Structure-frequency correlations

The amide I band positions for different secondary structures have been established through extensive empirical correlation with crystallographic data. These assignments, while showing some variation among studies, provide reliable fingerprints for structural identification [28].

α-helices absorb at 1650–1658 cm^{-1}, reflecting moderately strong intrahelical hydrogen bonds and characteristic transition dipole coupling [29]. β-sheets produce the most distinctive FTIR signatures due to the two-dimensional transition dipole coupling arrangement. Antiparallel β-sheets exhibit two characteristic bands: a strong component at 1620–1640 cm^{-1} arising from the in-phase coupling of strands, and a weaker band at 1680–1695 cm^{-1} from out-of-phase coupling. Parallel β-sheets typically show a single band at 1630–1640 cm^{-1} without the high-frequency component, reflecting their different coupling geometry [30,31].

Intermolecular β-sheets in amyloid fibrils produce particularly intense and distinctive signals at 1615–1630 cm^{-1}, lower in frequency than intramolecular β-sheets (1630–1640 cm^{-1}) (Table 2). The distinctive amyloid signature enables sensitive detection and monitoring of protein aggregation even at early stages [32,33].

Table 2. Amide I band assignments for protein secondary structure elements in FTIR spectroscopy. Band assignments follow Jackson and Mantsch [28], Surewicz et al. [29], and Kong and Yu [31].

Secondary structure	Wavenumber (cm^{-1})	Intensity	Diagnostic features
α-helix	1650–1658	Strong	Single band
Antiparallel β-sheet	1620–1640, 1680–1695	Strong, weak	Characteristic doublet
Parallel β-sheet	1630–1640	Strong	Single band, no doublet
Disordered	1640–1650	Broad	Overlaps with other structures

It is important to note that the wavenumber ranges listed in Table 2 are approximate guidelines rather than strict boundaries. The precise frequencies of amide I components are influenced by several factors, including hydrogen-deuterium (H/D) exchange state, temperature, ionic strength, pH, and the specific local environment of the peptide groups. In D_2O solutions, amide I bands shift to lower wavenumbers (amide I') due to the different hydrogen bonding properties of deuterium. Temperature changes can alter hydrogen bond strengths and population distributions, while ionic strength affects electrostatic interactions that modulate vibrational coupling. These factors should be considered when assigning spectral components to specific secondary structure types.

2.2.3. ATR-FTIR methodology

Attenuated total reflectance (ATR) methodology has become the preferred approach for protein FTIR measurements, eliminating many practical difficulties of transmission measurements.

ATR offers several practical advantages: minimal sample volume requirements (2–5 μL), elimination of problematic thin-cell path length control, and tolerance of higher protein concentrations (5–50 mg/mL). Diamond ATR crystals provide exceptional durability and chemical resistance, enabling measurements in aggressive solvents or at extreme pH. The shallow penetration depth reduces water absorption interference, though D_2O exchange remains beneficial for quantitative work [34].

For accurate secondary structure quantification, spectra should be acquired under controlled conditions: stable temperature (typically 25 °C) and consistent sample deposition. Background spectra of the buffer alone enable subtraction of solvent contributions [33].

2.2.4. Spectral analysis and deconvolution

The highly overlapping nature of amide I components requires mathematical resolution enhancement and fitting procedures for quantitative analysis. Several complementary approaches are employed [35–37]. Curve fitting involves decomposition of the amide I envelope into constituent Gaussian or Lorentzian-Gaussian bands. Component positions are typically assigned based on the established structure-frequency correlations, with areas proportional to secondary structure content. This approach requires assumptions about the component's number and shape, which can affect results. Partial least squares (PLS) and other multivariate methods avoid explicit peak assignment by correlating entire spectral patterns with known structures from reference proteins [37,38].

2.3. Nanoscale infrared spectroscopy for protein secondary structure analysis

Recent advances in nanoscale infrared (nano-IR) spectroscopy have opened new frontiers in protein secondary structure analysis by overcoming the diffraction limit that constrains conventional FTIR measurements. These techniques combine the chemical specificity of infrared spectroscopy with the nanometer-scale spatial resolution of scanning probe microscopy, enabling secondary structure characterization at the single-molecule and subcellular level. Three principal modalities have emerged: photothermal infrared (PTIR/AFM-IR) spectroscopy, photoinduced force microscopy (PiFM), and scattering-type scanning near-field optical microscopy (s-SNOM). For a comprehensive overview of nanoscale spectroscopic imaging technologies and their recent applications, the reader is referred to Xiao and Schultz [39]. A comparative discussion of the different modes of AFM-IR and their

implications for soft matter surfaces has been provided by Xie and Xu [40].

AFM-IR spectroscopy, pioneered by Dazzi and colleagues, detects the local photothermal expansion of a sample illuminated by a tunable IR laser through the deflection of an atomic force microscope (AFM) cantilever [41]. When the IR laser is tuned to an absorption band (e.g., amide I), the resulting thermal expansion is proportional to the local absorption coefficient, providing chemical contrast at spatial resolutions of 20–50 nm. This technique has been successfully applied to map secondary structure distributions in individual amyloid fibrils, revealing the spatial heterogeneity of β-sheet content along single fibrils that is invisible to ensemble-averaging methods such as conventional FTIR [42]. AFM-IR has also enabled the characterization of protein secondary structure in intact biological membranes and thin films at the nanoscale.

Photoinduced force microscopy (PiFM) provides an alternative detection mechanism based on the measurement of optical gradient forces between the AFM tip and the sample upon IR illumination [43]. Notably, IR-PiFM offers superior spatial resolution compared with other AFM-IR methods, achieving less than 5 nm resolution, as recently demonstrated by Ali et al. in the nanochemical cell-surface evaluation of antimicrobial interactions in *Bacillus subtilis* and vancomycin [44]. PiFM offers high sensitivity to surface and near-surface vibrational modes. Recent work has demonstrated the application of PiFM to map amide I and amide II band distributions in protein aggregates and biological specimens, providing nanoscale secondary structure contrast that complements topographic imaging [45]. Furthermore, Joseph et al. demonstrated the potential of PiF-IR for nanoscale evaluation of the secondary protein structure of F-actin, confirming the capability of this technique for biomaterial characterization at the nanoscale [46].

Scattering-type scanning near-field optical microscopy (s-SNOM) employs a metalized AFM tip to scatter evanescent near-field signals generated by broadband or tunable IR illumination [47]. By demodulating the scattered signal at higher harmonics of the tip oscillation frequency, s-SNOM achieves spatial resolution of ~10–20 nm and can record full nano-FTIR spectra at each pixel. This approach has been used to image protein secondary structure in individual amyloid fibrils and protein crystals, resolving β-sheet and α-helical domains at unprecedented spatial resolution [47]. Furthermore, Bakir et al. employed polarization-dependent IR spectroscopy of collagen from intact tendons down to the single fibril level [48], and Nishida et al. demonstrated sub-tip-radius near-field interactions in nano-FTIR vibrational spectroscopy on single proteins, achieving remarkable spectral sensitivity [49].

These nanoscale IR techniques share several advantages for protein secondary structure analysis: they require minimal sample preparation (typically air-dried thin films or cryosections), operate without labels or stains, and provide simultaneous topographic and chemical information. However, current limitations include the need for thin, flat samples; potential tip-induced artifacts; relatively long acquisition times for full spectral mapping; and challenges in quantitative analysis due to tip–sample interaction effects. Despite these constraints, nanoscale IR methods are rapidly maturing and offer unique capabilities for studying protein structure and misfolding at length scales inaccessible to conventional vibrational spectroscopy [50,51]. Recent developments have further expanded these capabilities: Waeytens et al. demonstrated the application of optical photothermal infrared spectroscopy for label-free secondary structure analysis of individual protein aggregates [52], and Kanevche et al. applied AFM-IR to map protein secondary structure changes in intact single cells [53]. Similarly, Quaroni et al. used AFM-IR to probe mid-infrared spectroscopy and microscopy of subcellular structures in eukaryotic cells [54]. In a complementary approach, Dos Santos et al. combined nanoscale infrared spectroscopy with chemometrics to enable detection of intracellular

protein distribution [55]. The continued advancement of the nanoscale IR field will benefit from further development of chemometrics approaches, including open-source tools such as the hyPIRana cluster analysis software for PiF-IR data analysis [56], as well as from theoretical modeling of the underlying physics, including the description of infrared near-field spectroscopy of molecular vibrations in thin layers [57] and the photothermal expansion of nanostructures in photoinduced force microscopy [58].

2.4. Comparative analysis: CD versus FTIR

Circular dichroism (CD) and Fourier-transform infrared (FTIR) spectroscopy represent the two most widely applied spectroscopic approaches for global secondary structure quantification in proteins. Although both techniques are routinely used for similar purposes, their underlying physical principles confer distinct strengths, limitations, and optimal application domains that warrant careful consideration.

CD spectroscopy probes electronic transitions of the peptide chromophore and is therefore particularly sensitive to the regular exciton coupling characteristic of α-helical structures. This physical basis explains the consistently high accuracy of CD for α-helix quantification (typically 3%–8% RMSD under optimized conditions) and the prominence of $[\theta]_{222}$ as a robust empirical marker of helical content. In contrast, β-sheet CD signatures are weaker, more heterogeneous, and strongly dependent on strand orientation, twist, and length, leading to greater uncertainty in β-sheet estimation, especially for proteins with mixed or atypical sheet topologies.

Table 3. Comparative characteristics of circular dichroism and FTIR spectroscopy for protein secondary structure analysis. Values compiled from Oberg et al. [59], Greenfield [22], and Goormaghtigh et al. [32].

Parameter	Circular dichroism	FTIR spectroscopy
Physical basis	Electronic transitions	Vibrational transitions
Sample quantity	20–100 μg	50–500 μg
Concentration range	0.1–1.0 mg/mL	5–50 mg/mL
Measurement time	5–15 min	10–30 min
α-helix accuracy (RMSD)	3%–8%	8%–12%
β-sheet accuracy (RMSD)	5%–15%	5%–8%
β-sheet type discrimination	Limited	Excellent (parallel vs. antiparallel)
Amyloid detection	Indirect	Direct, quantitative (1615–1630 cm^{-1})
Turbid/aggregated samples	Problematic (scattering)	Suitable (ATR)
Membrane proteins	Challenging (detergents)	Well-suited (ATR)
Temperature dependence	Thermal denaturation curves	Less commonly used
Buffer restrictions	Significant (< 200 nm)	H_2O interference (D_2O preferred)

FTIR spectroscopy, by probing vibrational transitions of the amide backbone, is governed primarily by hydrogen bonding strength and transition dipole coupling. These interactions are particularly diagnostic for β-sheet structures, which generate highly characteristic amide I band

patterns that enable reliable discrimination between parallel, antiparallel, and intermolecular β-sheets. As a result, FTIR routinely outperforms CD for β-sheet quantification (RMSD ~5%–8%) and uniquely enables direct detection of amyloid and aggregation-associated β-structures through their low-frequency signatures (1615–1630 cm^{-1}).

Importantly, the strengths of CD and FTIR are complementary rather than redundant. CD provides rapid, low-sample assessment of overall fold state and helical content in optically clear solutions, while FTIR excels in β-rich systems, turbid or aggregated samples, membrane proteins, and high-concentration regimes inaccessible to CD. Comparative benchmarking studies demonstrate that combined CD–FTIR analysis consistently improves prediction accuracy across all secondary structure classes relative to either method alone, underscoring the value of integrated spectroscopic strategies (Table 3). Consequently, CD and FTIR should not be viewed as competing techniques, but rather as synergistic tools whose deliberate combination yields the most reliable and information-rich spectroscopic characterization of protein secondary structure.

3. High-resolution techniques for atomic-level secondary structure determination

While spectroscopic methods provide valuable global assessments of secondary structure content, they lack the ability to localize structural elements along the polypeptide sequence or to visualize atomic details.

3.1. Nuclear magnetic resonance spectroscopy

3.1.1. Chemical shift-based secondary structure identification

The Chemical Shift Index (CSI), developed by Wishart, Sykes, and Richards [60–62], revolutionized NMR-based secondary structure determination by establishing systematic relationships between backbone chemical shifts and local conformation. The CSI compares observed chemical shifts with sequence-specific random coil reference values calculated from short unstructured peptides; deviations from these reference values indicate secondary structure (Figure 3).

For Cα carbons, positive secondary chemical shifts ($\Delta\delta > +0.7$ ppm for three or more consecutive residues) indicate α-helical conformation, reflecting the deshielding effect of the helical geometry. Negative shifts ($\Delta\delta < -0.7$ ppm) indicate β-strand conformation. The pattern is reversed for Cβ carbons (negative = helix, positive = strand) and Hα protons (negative = helix, positive = strand) [62].

CSI 2.0, developed by Hafsa and Wishart [62], represents a significant advance that integrates all six backbone chemical shifts (Cα, Cβ, C', N, Hα, and HN) using support vector machine classification. The algorithm achieves three-state prediction accuracy (helix/strand/coil) exceeding 90% on benchmark datasets, substantially better than single-shift CSI methods. CSI 2.0 also provides probability scores for each prediction.

TALOS-N (Torsion Angle Likelihood Obtained from Shift and sequence similarity-Next generation), developed by Shen and Bax [63,64], extends chemical shift analysis to predict backbone dihedral angles (φ and ψ) directly rather than discrete secondary structure classes. Using artificial neural network analysis of chemical shift patterns from a database of 200 + proteins with known structures, TALOS-N achieves angular predictions with RMSD of ± 12° for approximately 90% of residues. These predictions are sufficiently accurate to unambiguously define secondary structure

regions and provide restraints for structure calculation.

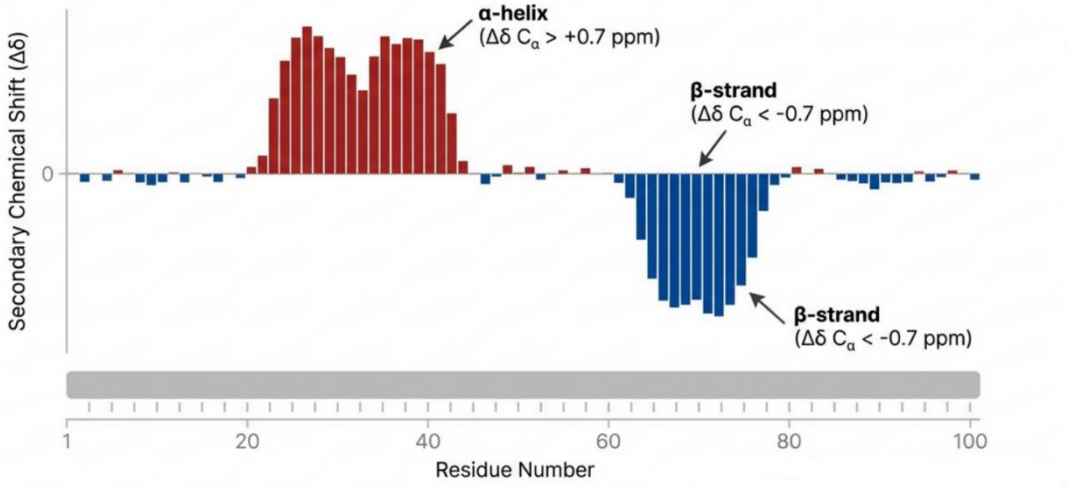

Figure 3. Secondary structure identification using the Chemical Shift Index (CSI). Secondary chemical shifts (Δδ) of Cα atoms report on local backbone conformation: positive values (Δδ > 0.7 ppm) are characteristic of α-helical regions (red), whereas negative values (Δδ < −0.7 ppm) indicate β-strand conformations (blue).

3.1.2. NOE-based structure determination

The nuclear Overhauser effect (NOE) provides distance information by reporting cross-relaxation between protons within approximately 5–6 Å. Characteristic NOE patterns define secondary structure types with high reliability [65].

α-helices show strong sequential dNN (i, i + 1) contacts (HN to HN of adjacent residues) and medium-range dαN (i, i + 3), dαβ (i, i + 3), and dαN (i, i + 4) NOEs that reflect the i → i + 4 hydrogen bonding pattern (Figure 4). β-sheets display strong sequential dαN (i, i + 1) contacts with weak dNN, reflecting the extended backbone conformation. Interstrand NOEs between adjacent strands provide direct evidence for sheet formation and define strand register [66].

3.1.3. Scalar coupling analysis

Three-bond scalar coupling constants (^3J) provide independent confirmation of secondary structure through their dependence on dihedral angles via the Karplus relationship: $^3J = A\cos^2\theta + B\cos\theta + C$, where coefficients A, B, and C are parameterized from experimental data [67].

The ^{3}JHN-Hα coupling reflects the φ angle: In Figure 4, values of 3–5 Hz indicate α-helical conformation (φ ≈ −60°), while 8–10 Hz indicates extended β-strand conformation (φ ≈ −120°). Intermediate values (6–7 Hz) suggest averaging between conformations or irregular structure. Coupling measurements require high-quality spectra but provide quantitative angle information complementary to NOE-based analysis [68].

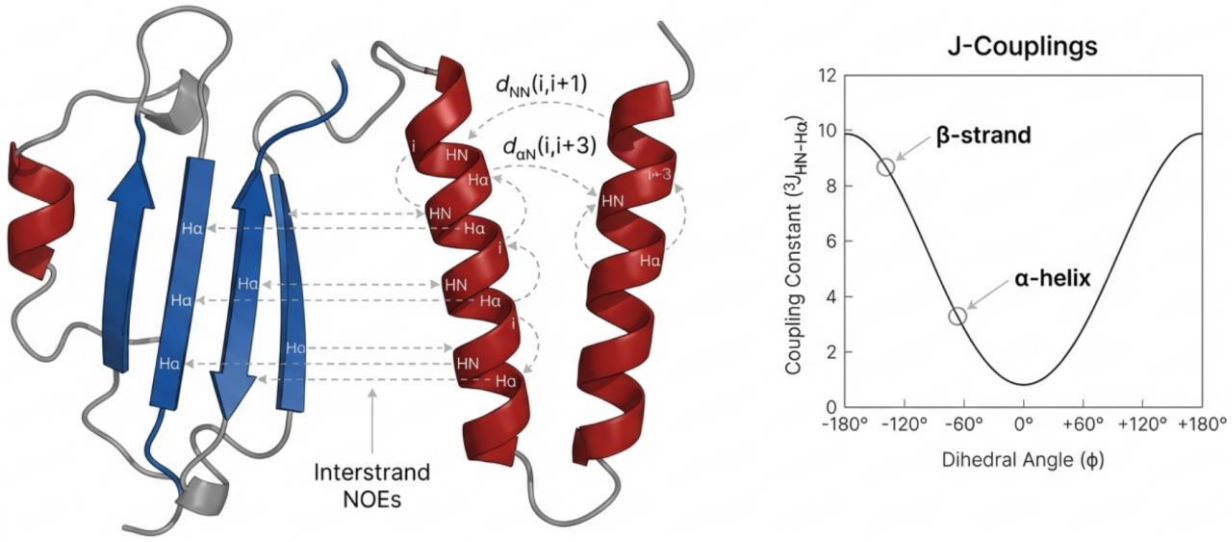

Figure 4. NMR structural restraints: NOEs and scalar couplings. (Left) Characteristic interstrand and sequential NOE patterns (dashed lines) define secondary structure topology and strand register. (Right) Dependence of the 3JHN−Hα coupling constant on the dihedral angle ϕ; extended β-strands typically show large couplings (8–10 Hz) while α-helices exhibit small couplings (3–5 Hz).

3.1.4. Size limitations and advanced NMR methods

Traditional NMR approaches face practical limitations as protein size increases, primarily due to faster transverse relaxation (shorter T_2) that broadens lines and reduces sensitivity [69].

TROSY (transverse relaxation-optimized spectroscopy), developed by Pervushin et al. [70], exploits interference between different relaxation mechanisms—dipole–dipole coupling and chemical shift anisotropy—that partially cancel at high magnetic fields. TROSY selects the slowly relaxing component of the multiplet, dramatically reducing linewidths. Combined with perdeuteration to eliminate relaxation from aliphatic protons, TROSY extends practical size limits to approximately 80–100 kDa for favorable cases.

3.2. X-ray crystallography

3.2.1. Principles and structure determination workflow

X-ray crystallography determines protein structure by analyzing the diffraction pattern produced when X-rays scatter from the regularly arranged molecules in a crystal (Figure 5). The electrons in the protein scatter X-rays coherently, and the regular crystal lattice creates constructive interference at specific angles described by Bragg's law: $n\lambda = 2d \sin\theta$. The resulting diffraction pattern contains information about electron density in the unit cell, encoded in the amplitudes and phases of the scattered waves [71].

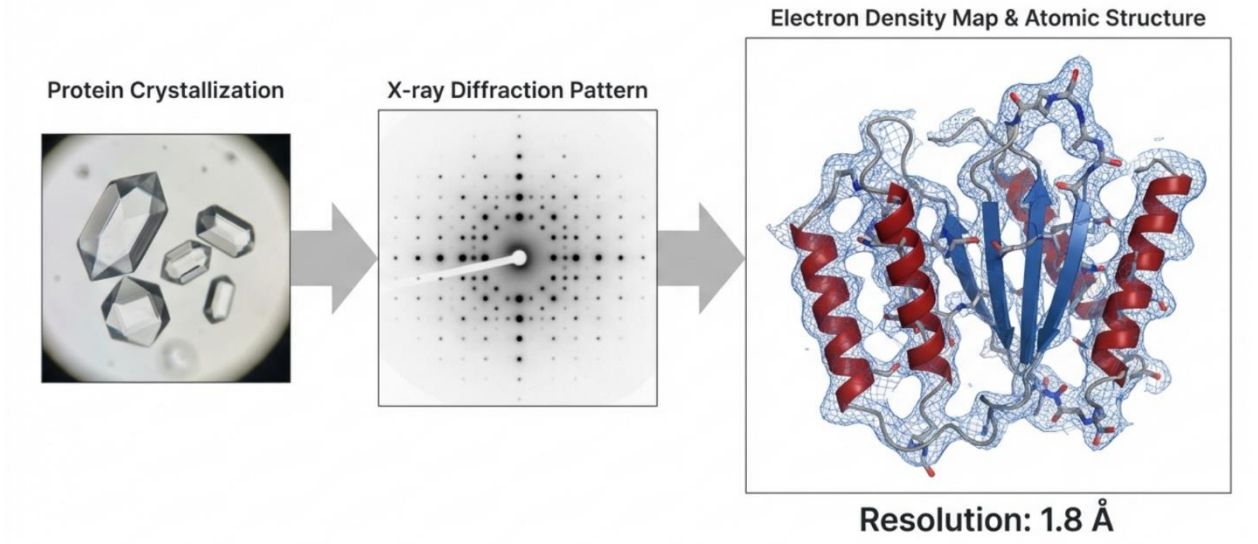

Figure 5. X-ray crystallography structure determination workflow. The process progresses from protein crystallization to the collection of an X-ray diffraction pattern. The resulting electron density map allows for the atomic-level modeling of the polypeptide backbone and side chains (right), shown here at a high resolution of 1.8 Å.

The central challenge—the "phase problem"—arises because detectors measure only intensities (proportional to amplitude squared), not phases. Several methods have been developed to determine phases: molecular replacement (MR) uses a homologous known structure as a search model, isomorphous replacement (MIR/SIR) introduces heavy atoms that alter diffraction, and anomalous diffraction methods (SAD/MAD) exploit wavelength-dependent scattering by selenomethionine or other anomalous scatterers [72].

3.2.2. Secondary structure identification and DSSP algorithm

At sufficient resolution (better than approximately 3 Å), secondary structure elements are clearly visible in electron density maps. α-helices appear as tubes of helical density with regularly spaced side chain features projecting outward. β-sheets form planar or twisted arrangements with alternating density for residues above and below the sheet plane. The backbone can be traced unambiguously, and hydrogen bonding patterns can be inferred from geometry [73].

The DSSP (Dictionary of Secondary Structure of Proteins) algorithm, developed by Kabsch and Sander [73], provides the standard method for assigning secondary structure from atomic coordinates. DSSP identifies hydrogen bonds using an electrostatic energy criterion: $E = 0.084 \times (1/rON + 1/rCH - 1/rOH - 1/rCN) \times 332$ kcal/mol, with hydrogen bonds assigned where $E < -0.5$ kcal/mol.

Based on hydrogen bond patterns, DSSP classifies residues into eight categories: H (α-helix, minimum 4 residues with $i \rightarrow i + 4$ hydrogen bonds), G (3_{10}-helix, $i \rightarrow i + 3$ hydrogen bonds), I (π-helix, $i \rightarrow i + 5$ hydrogen bonds), E (extended β-strand, part of a β-ladder), B (isolated β-bridge), T (hydrogen-bonded turn), S (bend), and L or blank (loop/coil). This algorithm has been applied to nearly every structure in the PDB and provides the standard for secondary structure annotation [73,74].

3.2.3. Resolution requirements and structure validation

The resolution of a crystal structure determines the level of detail that can be reliably interpreted. Resolution is typically reported as the minimum d-spacing in the diffraction pattern, with lower numbers indicating higher resolution [75]. At resolutions worse than 4 Å, only the overall molecular envelope and major domain arrangement are visible. Secondary structure elements may be inferred from tube-like features, but cannot be reliably traced. At 3–4 Å resolution, secondary structure elements can be confidently identified and placed, though side chain positions remain uncertain. This resolution is typically adequate for secondary structure mapping. Resolutions of 2–3 Å permit confident placement of most side chains and accurate backbone tracing, and water molecules begin to appear in well-ordered regions [76].

Structure validation using tools, such as MolProbity [76], assess geometric quality and identify potential errors. The Ramachandran plot remains central to validation: high-quality structures typically show > 98% of residues in allowed regions and > 90% in favored regions.

3.3. Cryo-electron microscopy

3.3.1. The resolution revolution

Cryo-electron microscopy (cryo-EM) has undergone a transformative "resolution revolution" over the past decade, transitioning from a low-resolution technique limited to overall shape determination to a method capable of near-atomic resolution rivaling X-ray crystallography (Figure 6). This transformation resulted from the convergence of three technological advances: direct electron detectors with dramatically improved sensitivity and speed, motion correction algorithms that computationally correct for beam-induced specimen movement, and sophisticated image processing software for particle alignment and 3D reconstruction [77,78].

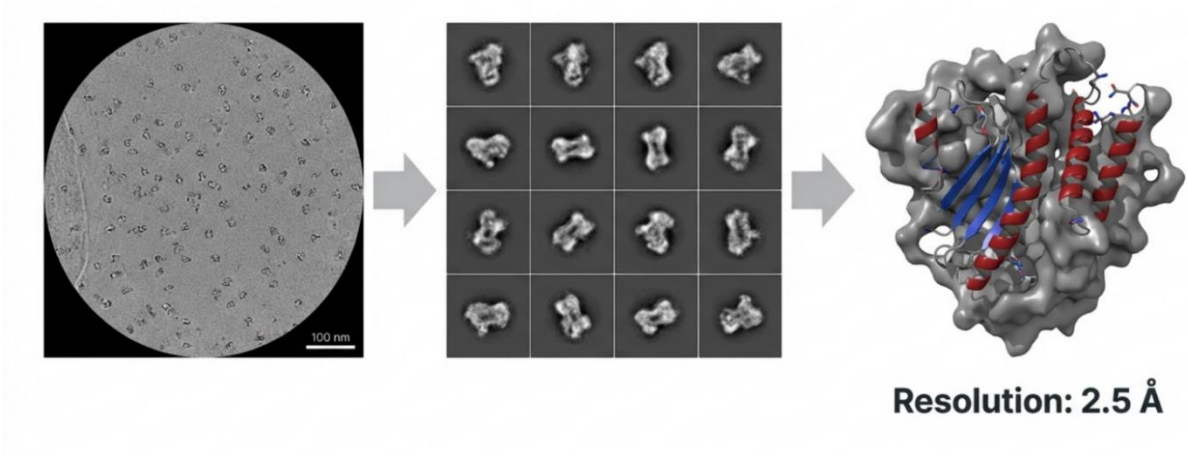

Figure 6. Single-particle cryo-electron microscopy (cryo-EM) reconstruction. From micrographs of vitrified protein solution (left), individual particles are computationally extracted and averaged into 2D classes (center). These are reconstructed into a high-resolution 3D density map (right, 2.5 Å), allowing direct visualization of secondary structure elements like α-helices and β-sheets.

Current state-of-the-art resolutions approach 1.2 Å for favorable specimens such as apoferritin and glutamate dehydrogenase, achieving detail comparable to high-resolution crystal structures [78].

3.3.2. Sample preparation and imaging

Cryo-EM sample preparation involves rapid vitrification of the protein solution to preserve native structure in amorphous ice. A small volume (3–5 μL) of sample at 0.5–5 mg/mL is applied to an EM grid, blotted to create a thin aqueous film, and plunge-frozen in liquid ethane [79].

Images are acquired with the sample maintained at cryogenic temperature (typically −170 to −180 °C) using low electron doses (20–50 $e^{-}/Å^{2}$) to minimize radiation damage. Direct electron detectors record movies that are subsequently aligned and averaged. Hundreds of thousands of individual particle images are computationally extracted, aligned, classified, and averaged to reconstruct the 3D structure [80].

3.3.3. Secondary structure visualization at different resolutions

At resolutions approaching 2–3 Å, cryo-EM density maps allow direct visualization of secondary structure elements with clarity comparable to crystallography. α-helices appear as characteristic helical tubes with resolved pitch, while β-sheets appear as extended planar features with visible strand separation; the polypeptide backbone can be traced unambiguously. Side chain density permits confident amino acid assignment for most residue types [81].

At intermediate resolutions (3–5 Å), secondary structure elements remain identifiable, though backbone tracing becomes less certain, and side chain positioning relies partially on modeling assumptions. The distribution of helices and sheets can be mapped, and large conformational differences between states can be detected.

At lower resolutions (5–10 Å), secondary structure assignment becomes challenging but not impossible. Machine learning approaches and molecular dynamics flexible fitting (MDFF) can help interpret density at these resolutions by incorporating prior structural knowledge [82].

3.3.4. Advantages and current limitations

Cryo-EM offers several key advantages: no crystallization requirement, which enables structure determination of samples that resist crystallization; capacity to capture multiple conformational states through computational classification; lower per-grid sample consumption, though total material requirements vary widely depending on optimization; and suitability for large complexes and membrane proteins in native-like environments [83]. Recent scaffold approaches have extended cryo-EM to small proteins previously considered too small for the technique [84].

Current limitations include challenges with flexible or highly heterogeneous samples where averaging breaks down, difficulties with proteins smaller than ~ 50 kDa without scaffolding (Table 4), the requirement for specialized and expensive equipment, and significant computational demands [85].

Table 4. Comparative characteristics of high-resolution techniques for protein secondary structure determination. Data compiled from Wider and Wüthrich [68], Rupp [72], and Nogales [82].

Parameter	NMR spectroscopy	X-ray crystallography	Cryo-EM
Optimal size range	< 40 kDa (solution)	No theoretical size limit, but strong practical constraints	> 100 kDa (optimal)
Sample state	Solution	Crystal	Vitrified solution
Sample quantity	0.5–5 mg	1–10 mg	50–200 μg
Isotope labeling	Required (> 10 kDa)	Not required	Not required
Typical resolution	Residue-level/NMR ensemble	1.5–3.0 Å	2–4 Å (improving)
Best reported resolution	~1 Å (small proteins)	< 1.0 Å	1.2 Å
Dynamics information	Excellent (ps to s)	Limited (B-factors)	Conformational heterogeneity
Disordered regions	Excellent characterization	Often invisible	Often invisible
Membrane proteins	Challenging (size, micelles)	Challenging (crystallization)	Well-suited (nanodiscs)
Time to structure	Weeks to months	Days to months	Days to weeks
Data collection	Days	Hours	Hours to days

4. Complementary biophysical techniques

While spectroscopic and high-resolution structural methods provide the primary means for secondary structure characterization, several complementary techniques offer orthogonal information that contextualizes and supports secondary structure analysis.

4.1. Small-angle X-ray scattering (SAXS)

Small-angle X-ray scattering (SAXS) provides low-resolution information on overall protein size, shape, and folding state in solution, complementing secondary structure measurements by reporting on global compaction and quaternary organization [86].

4.2. Hydrogen–deuterium exchange mass spectrometry (HDX-MS)

Hydrogen–deuterium exchange mass spectrometry (HDX-MS) probes backbone amide protection and solvent accessibility, indirectly reporting on secondary structure stability. Regions with persistent hydrogen bonding exhibit reduced exchange rates, making HDX-MS useful for comparing conformational states and mapping structural changes upon ligand binding or mutation [87,88].

4.3. Fluorescence spectroscopy and FRET

Fluorescence-based techniques, including Förster resonance energy transfer (FRET), report on intramolecular distances and conformational changes. While not directly measuring secondary structure, FRET provides complementary distance constraints when combined with other structural approaches [89].

5. Comparative analysis and decision frameworks

5.1. Quantitative accuracy comparison across techniques

Systematic benchmarking studies enable quantitative comparison of secondary structure estimation accuracy across techniques. For CD spectroscopy, meta-analyses indicate root mean square deviations (RMSD) of 3%–8% for α-helix content estimation and 5%–15% for β-sheet content when using optimized deconvolution methods with appropriate reference databases [32]. The greater uncertainty in β-sheet estimation reflects spectral variability among different sheet types.

FTIR-based quantification typically achieves an RMSD of 8%–12% for α-helix and 5%–8% for β-sheet when amide I band analysis is performed with appropriate curve-fitting or multivariate methods [90]. The superior β-sheet accuracy compared to CD reflects the distinctive spectral signatures of different sheet types in the infrared.

Combined CD and FTIR analysis consistently outperforms either technique alone. The optimization study by Oberg et al. [59] achieved standard errors of prediction as low as 5.5% for α-helix, 6.6% for β-sheet, and 3.4% for β-turns using combined spectroscopic data with optimized reference databases. These results demonstrate the value of spectroscopic integration.

NMR chemical shift-based predictions achieve a backbone dihedral angle RMSD of ± 12° for approximately 90% of residues, corresponding to highly accurate secondary structure identification at the residue level. The CSI 2.0 three-state accuracy exceeds 90% for helix/strand/coil classification [62]. High-resolution structures from crystallography and cryo-EM provide the ultimate reference standards against which other methods are validated (Table 5).

Table 5. Sample requirements and practical considerations for major biophysical techniques. Requirements are based on established protocols from Greenfield [10], Yang et al. [34], Shen and Bax [64], and Nogales [82].

Technique	Quantity	Concentration	Time	Special requirements
CD	20–100 μg	0.1–1 mg/mL	10–30 min	Clear solutions, suitable buffers
SRCD	10–50 μg	0.1–5 mg/mL	30–60 min	Synchrotron access
FTIR (ATR)	50–500 μg	5–50 mg/mL	15–45 min	D_2O exchange preferred
NMR	0.5–5 mg	0.1–1 mM	Days–weeks	$^{13}C/^{15}N$ labeling (> 10 kDa)
X-ray	1–10 mg	5–20 mg/mL	Days–months	Diffracting crystals
Cryo-EM	50–200 μg	0.5–5 mg/mL	Days–weeks	Homogeneous sample, good ice
SAXS	50–200 μg	1–10 mg/mL	Hours	Monodisperse sample
HDX-MS	10–100 μg	1–20 μM	Hours–day	MS-compatible buffers

Table 5 summarizes the practical requirements and constraints for each major technique. Several important considerations merit emphasis. CD spectroscopy requires optically clear, non-scattering solutions in buffers that do not absorb in the far-UV (chloride-containing buffers and detergents can be problematic below 200 nm). FTIR measurements benefit from D_2O exchange to reduce water absorption interference in the amide I region, and higher protein concentrations are necessary to achieve adequate signal-to-noise. NMR spectroscopy requires isotopic labeling (^{13}C, ^{15}N) for proteins larger than approximately 10 kDa, and sample stability over extended acquisition periods is essential. X-ray crystallography demands well-diffracting crystals, which may require extensive screening and optimization. Cryo-EM requires homogeneous, monodisperse samples that behave well on EM grids, and specimen preparation remains a significant bottleneck. These practical constraints often dictate technique selection as much as the scientific question itself.

5.2. Decision framework for technique selection

The optimal choice of method for secondary structure analysis depends critically on the biological question being addressed, the physical properties of the sample, and the required level of structural resolution. No single technique is universally optimal; instead, informed selection requires balancing accuracy, resolution, throughput, and feasibility.

For rapid assessment of fold state, conformational changes, or sample quality during protein production and purification, far-UV CD spectroscopy provides the most efficient entry point. With minimal material and short acquisition times, CD can readily distinguish folded from unfolded states, estimate helical content, and monitor thermal or chemical denaturation. However, CD should be interpreted cautiously for β-rich proteins or samples prone to aggregation, where light scattering and spectral heterogeneity can compromise accuracy.

For systems with substantial β-sheet content, aggregation propensity, membrane association, or optical turbidity, FTIR spectroscopy—particularly in ATR mode—offers decisive advantages. FTIR not only provides superior β-sheet quantification but also uniquely distinguishes native intramolecular β-sheets from intermolecular amyloid structures. In studies of protein misfolding, fibrillation, or membrane protein topology, FTIR is often indispensable rather than optional.

When residue-level resolution or dynamic information is required, NMR spectroscopy remains unmatched for proteins amenable to solution studies. Chemical shift–based analyses provide highly reliable secondary structure mapping even in partially disordered systems, while NOEs and scalar couplings validate and refine these assignments. For intrinsically disordered proteins or proteins undergoing conformational exchange, NMR is frequently the only technique capable of capturing the relevant structural ensemble.

Atomic-resolution determination of secondary structure ultimately relies on X-ray crystallography or cryo-electron microscopy. Crystallography remains the method of choice when well-diffracting crystals can be obtained, offering the highest typical resolution and mature validation standards. Cryo-EM is preferred for large assemblies, membrane proteins, and conformationally heterogeneous systems that resist crystallization, and increasingly rivals crystallography in achievable resolution.

In practice, the most robust structural insights arise not from reliance on a single technique but from integrative workflows that combine rapid spectroscopic screening with high-resolution structural methods tailored to the system under investigation.

5.3. Integrative multi-technique workflows

The most comprehensive structural characterization results from combining multiple techniques that provide complementary information (Figure 7). A practical multi-technique workflow for thorough characterization might proceed as follows:

Phase 1 (initial characterization): CD spectroscopy for secondary structure estimation and thermal stability; DLS for size and polydispersity; mass spectrometry for mass confirmation and post-translational modification mapping.

Phase 2 (detailed spectroscopic analysis): FTIR for β-sheet characterization and aggregation monitoring; SEC-SAXS to confirm monodispersity, measure Rg and Dmax, and assess overall shape; additional CD experiments including wavelength scans under different conditions.

Phase 3 (high-resolution structure): NMR for residue-level structure and dynamics (if size permits); crystallization trials for X-ray structure determination; cryo-EM for large or difficult-to-crystallize targets.

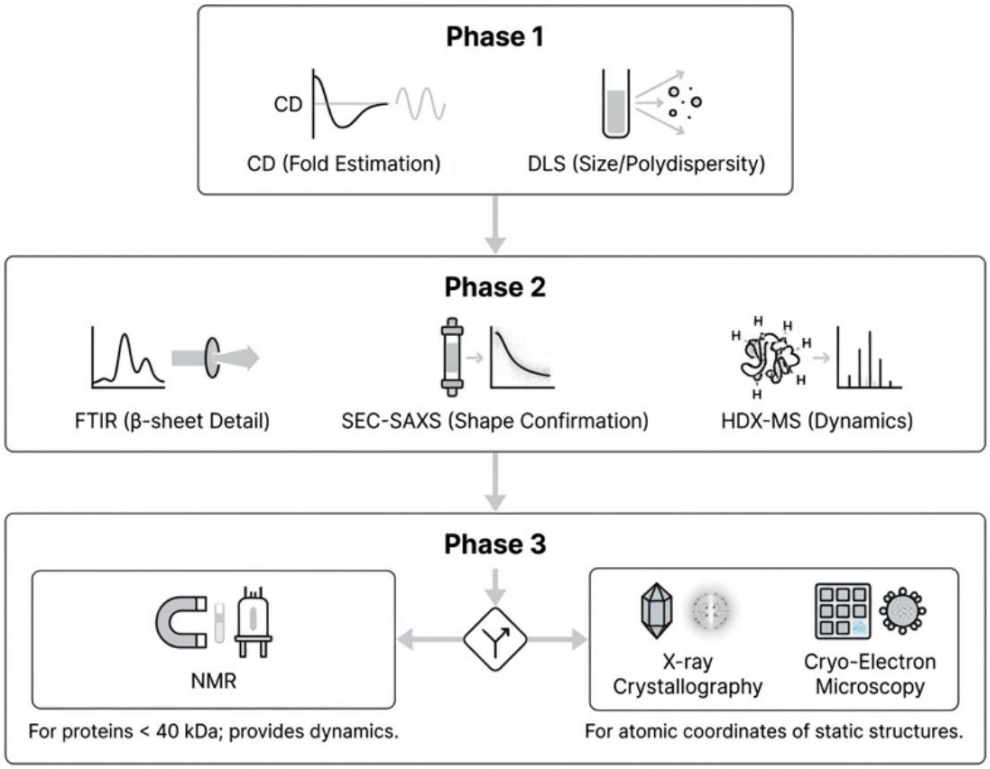

Figure 7. Integrative multi-technique workflow for protein characterization. A hierarchical approach proceeding from phase 1 (global assessment via CD and DLS) and phase 2 (detailed spectroscopic analysis via FTIR and shape confirmation by SEC-SAXS/HDX-MS) to phase 3 (high-resolution structural determination via NMR, X-ray crystallography, or cryo-EM).

6. Conclusions

Protein secondary structure analysis remains a foundational pillar of structural biology, linking amino acid sequence to three-dimensional architecture, stability, dynamics, and function. Over seven decades of methodological development have produced a diverse and powerful toolkit, spanning rapid spectroscopic techniques to atomic-resolution structural methods. This review demonstrates that the value of this toolkit lies not in any single "best" technique but in the strategic integration of complementary approaches.

Spectroscopic methods such as CD and FTIR provide rapid, accessible, and quantitatively meaningful assessments of global secondary structure content, each excelling for different structural classes. High-resolution techniques—NMR, X-ray crystallography, and cryo-EM—define the gold standard for secondary structure assignment at the residue and atomic level, while also revealing dynamics and conformational heterogeneity inaccessible to ensemble-averaging methods. Orthogonal techniques, including SAXS and HDX-MS, contextualize secondary structure within global shape and stability landscapes.

A central conclusion of this analysis is that methodological complementarity consistently outperforms methodological optimization in isolation. Combined CD–FTIR spectroscopy yields superior quantitative accuracy; spectroscopic data guide and validate high-resolution structure determination; and integrative workflows mitigate the limitations inherent to any individual technique. This paradigm is particularly critical for challenging systems such as membrane proteins, intrinsically disordered proteins, and aggregation-prone assemblies.

Looking forward, continued advances in instrumentation, data analysis, and machine learning will further enhance secondary structure characterization. However, experimental rigor, cross-validation, and thoughtful technique selection will remain essential. By articulating clear decision frameworks and highlighting integrative strategies, this review aims to support researchers in designing robust, efficient, and informative experimental approaches to protein secondary structure analysis in the era of integrative structural biology.

Use of generative-AI tools declaration

Language-support tools were used for editorial assistance.

Acknowledgments

The author gratefully acknowledges helpful discussions during the preparation of this manuscript.

Conflict of interest

The author declare no conflicts of interest.

Author contributions

Adil Ahiri - conceptualization, literature review, writing, editing.

References

1. Pauling L, Corey RB (1951) The pleated sheet, a new layer configuration of polypeptide chains. *Proc Natl Acad Sci USA* 37: 251–256. https://doi.org/10.1073/pnas.37.5.251
2. Pauling L, Corey RB, Branson HR (1951) The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* 37: 205–211. https://doi.org/10.1073/pnas.37.4.205
3. Richardson JS (1981) The anatomy and taxonomy of protein structure, In: Anfinsen C.B., Edsall J.T., Richards F.M., *Advances in Protein Chemistry*, New York: Academic Press, 167–339. https://doi.org/10.1016/S0065-3233(08)60520-3
4. Banci L, Cantini F, Cevec M, et al. (2012) Structure of biomolecules: fundamentals, In: Bertini I., McGreevy K.S., Parigi G., *NMR of Biomolecules*, Chichester: John Wiley & Sons, Ltd, 7–32. https://doi.org/10.1002/9783527644506.ch2
5. Pauling L, Corey RB (1953) Stable configurations of polypeptide chains. *Proc R Soc Lond B Biol Sci* 141: 21–33. https://doi.org/10.1098/rspb.1953.0012
6. Fersht AR (2008) From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat Rev Mol Cell Biol* 9: 650–654. https://doi.org/10.1038/nrm2446
7. Dill KA, MacCallum JL (2012) The protein-folding problem, 50 years on. *Science* 338: 1042–1046. https://doi.org/10.1126/science.1219021
8. Huang PS, Boyken SE, Baker D (2016) The coming of age of de novo protein design. *Nature* 537: 320–327. https://doi.org/10.1038/nature19946
9. Kuhlman B, Bradley P (2019) Advances in protein structure prediction and design. *Nat Rev Mol Cell Biol* 20: 681–697. https://doi.org/10.1038/s41580-019-0163-x
10. Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1: 2876–2890. https://doi.org/10.1038/nprot.2006.202
11. Kelly SM, Jess TJ, Price NC (2005) How to study proteins by circular dichroism. *BBA-Proteins Proteom* 1751: 119–139. https://doi.org/10.1016/j.bbapap.2005.06.005
12. Sreerama N, Woody RW (2004) Computation and analysis of protein circular dichroism spectra, In: Brand L., Johnson M.L., *Methods Enzymol*, New York: Academic Press, 318–351. https://doi.org/10.1016/S0076-6879(04)83013-1
13. Woody RW (1995) Circular dichroism. In: Sauer K., *Methods Enzymol*, New York: Academic Press, 34–71. https://doi.org/10.1016/0076-6879(95)46006-3
14. Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287: 252–260. https://doi.org/10.1006/abio.2000.4880
15. Lees JG, Smith BR, Wien F, et al. (2004) CDtool—an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal Biochem* 332: 285–289. https://doi.org/10.1016/j.ab.2004.06.002
16. Chen YH, Yang JT, Chau KH (1974) Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. *Biochemistry* 13: 3350–3359. https://doi.org/10.1021/bi00713a027
17. Johnson WC (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* 35: 307–312. https://doi.org/10.1002/(SICI)1097-0134(19990515)35:3<307::AID-PROT4>3.0.CO;2-3

18. Sreerama N, Woody RW (2004) On the analysis of membrane protein circular dichroism spectra. *Protein Sci* 13: 100–112. https://doi.org/10.1110/ps.03258404
19. Micsonai A, Wien F, Bulyáki É, et al. (2018) BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* 46: W315–W322. https://doi.org/10.1093/nar/gky497
20. Micsonai A, Moussong É, Wien F, et al. (2022) BeStSel: webserver for secondary structure and fold prediction for protein CD spectroscopy. *Nucleic Acids Res* 50: W90–W98. https://doi.org/10.1093/nar/gkac345
21. Whitmore L, Wallace BA (2004) Analysis of circular dichroism data, In: Brand L., Johnson M.L., *Methods Enzymol*, New York: Academic Press, 282–317. https://doi.org/10.1016/S0076-6879(04)83012-X
22. Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc* 1: 2527–2535. https://doi.org/10.1038/nprot.2006.204
23. Krimm S, Bandekar J (1986) Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Adv Protein Chem* 38: 181–364. https://doi.org/10.1016/S0065-3233(08)60528-8
24. Barth A, Zscherp C (2002) What vibrations tell about proteins. *Q Rev Biophys* 35: 369–430. https://doi.org/10.1017/S0033583502003815
25. Torii H, Tasumi M (1992) Model calculations on the amide-I infrared bands of globular proteins. *J Chem Phys* 96: 3379–3387. https://doi.org/10.1063/1.461939
26. Hamm P, Zanni M (2011) *Concepts and Methods of 2D Infrared Spectroscopy*, Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511675935
27. Miyazawa T, Blout ER (1961) The infrared spectra of polypeptides in various conformations: amide I and II bands. *J Am Chem Soc* 83: 712–719. https://doi.org/10.1021/ja01464a042
28. Jackson M, Mantsch HH (1995) The use and misuse of FTIR spectroscopy in the determination of protein structure. *Crit Rev Biochem Mol Biol* 30: 95–120. https://doi.org/10.3109/10409239509085140
29. Surewicz WK, Mantsch HH, Chapman D (1993) Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. *Biochemistry* 32: 389–394. https://doi.org/10.1021/bi00053a001
30. Arrondo JLR, Muga A, Castresana J, et al. (1993) Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy. *Prog Biophys Mol Biol* 59: 23–56. https://doi.org/10.1016/0079-6107(93)90006-6
31. Kong J, Yu S (2007) Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta Biochim Biophys Sin* 39: 549–559. https://doi.org/10.1111/j.1745-7270.2007.00320.x
32. Goormaghtigh E, Ruysschaert J-M, Raussens V (2006) Evaluation of the information content in infrared spectra for protein secondary structure determination. *Biophys J* 90: 2946–2957. https://doi.org/10.1529/biophysj.105.072017
33. Byler DM, Susi H (1986) Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* 25: 469–487. https://doi.org/10.1002/bip.360250307
34. Yang H, Yang S, Kong J, et al. (2015) Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nat Protoc* 10: 382–396. https://doi.org/10.1038/nprot.2015.024

35. Goormaghtigh E, Cabiaux V, Ruysschaert J-M (1994) Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy, In: Hilderson H.J., Ralston G.B., *Physicochemical Methods in the Study of Biomembranes*, Boston: Springer US, 405–450. https://doi.org/10.1007/978-1-4615-1863-1_10
36. Arrondo JL, Goñi FM (1999) Structure and dynamics of membrane proteins as studied by infrared spectroscopy. *Prog Biophys Mol Biol* 72: 367–405. https://doi.org/10.1016/S0079-6107(99)00007-3
37. Lee DC, Haris PI, Chapman D, et al. (1990) Determination of protein secondary structure using factor analysis of infrared spectra. *Biochemistry* 29: 9185–9193. https://doi.org/10.1021/bi00491a012
38. Ganim Z, Chung HS, Smith AW, et al. (2008) Amide I two-dimensional infrared spectroscopy of proteins. *Acc Chem Res* 41: 432–441. https://doi.org/10.1021/ar700188n
39. Xiao L, Schultz ZD (2018) Spectroscopic imaging at the nanoscale: technologies and recent applications. *Anal Chem* 90: 440–458. https://doi.org/10.1021/acs.analchem.7b04151
40. Xie Q, Xu XG (2023) What do different modes of AFM-IR mean for measuring soft matter surfaces?. *Langmuir* 39: 17593–17599. https://doi.org/10.1021/acs.langmuir.3c02950
41. Dazzi A, Prater CB (2017) AFM-IR: technology and applications in nanoscale infrared spectroscopy and chemical imaging. *Chem Rev* 117: 5146–5173. https://doi.org/10.1021/acs.chemrev.6b00448
42. Ruggeri FS, Longo G, Faggiano S, et al. (2015) Infrared nanospectroscopy characterization of oligomeric and fibrillar aggregates during amyloid formation. *Nat Commun* 6: 7831. https://doi.org/10.1038/ncomms8831
43. Sifat AA, Jahng J, Potma EO (2022) Photo-induced force microscopy (PiFM) – principles and implementations. *Chem Soc Rev* 51: 4208–4222. https://doi.org/10.1039/D2CS00052K
44. Ali M, Schneider R, Strecker A, et al. (2025) Nanochemical cell-surface evaluation in photothermal spectroscopic imaging of antimicrobial interactions in the model system *Bacillus subtilis* and *Vancomycin*. *Anal Chem* 97: 23914–23926. https://doi.org/10.1021/acs.analchem.5c03502
45. Jahng J, Brocious J, Fishman DA, et al. (2014) Gradient and scattering forces in photoinduced force microscopy. *Phys Rev B* 90: 155417. https://doi.org/10.1103/PhysRevB.90.155417
46. Joseph J, Spantzel L, Ali M, et al. (2024) Nanoscale chemical characterization of secondary protein structure of F-Actin using mid-infrared photoinduced force microscopy (PiF-IR). *Spectrochim Acta A Mol Biomol Spectrosc* 306: 123612. https://doi.org/10.1016/j.saa.2023.123612
47. Amenabar I, Poly S, Nuansing W, et al. (2013) Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy. *Nat Commun* 4: 2890. https://doi.org/10.1038/ncomms3890
48. Bakir G, Girouard BE, Wiens R, et al. (2020) Orientation matters: polarization dependent IR spectroscopy of collagen from intact tendon down to the single fibril level. *Molecules* 25: 4295. https://doi.org/10.3390/molecules25184295
49. Nishida J, Otomo A, Koitaya T, et al. (2024) Sub-tip-radius near-field interactions in nano-FTIR vibrational spectroscopy on single proteins. *Nano Lett* 24: 836–843. https://doi.org/10.1021/acs.nanolett.3c03479

50. Centrone A (2015) Infrared imaging and spectroscopy beyond the diffraction limit. *Annu Rev Anal Chem* 8: 101–126. https://doi.org/10.1146/annurev-anchem-071114-040435
51. Ruggeri FS, Habchi J, Chia S, et al. (2021) Infrared nanospectroscopy reveals the molecular interaction fingerprint of an aggregation inhibitor with single Aβ42 oligomers. *Nat Commun* 12: 688. https://doi.org/10.1038/s41467-020-20782-0
52. Waeytens J, De Meutter J, Goormaghtigh E, et al. (2023) Determination of secondary structure of proteins by nanoinfrared spectroscopy. *Anal Chem* 95: 621–627. https://doi.org/10.1021/acs.analchem.2c01431
53. Kanevche K, Burr DJ, Nürnberg DJ, et al. (2021) Infrared nanoscopy and tomography of intracellular structures. *Commun Biol* 4: 1341. https://doi.org/10.1038/s42003-021-02876-7
54. Quaroni L, Pogoda K, Wiltowska-Zuber J, et al. (2018) Mid-infrared spectroscopy and microscopy of subcellular structures in eukaryotic cells with atomic force microscopy – infrared spectroscopy. *RSC Adv* 8: 2786–2794. https://doi.org/10.1039/C7RA10240B
55. Dos Santos ACV, Heydenreich R, Derntl C, et al. (2020) Nanoscale infrared spectroscopy and chemometrics enable detection of intracellular protein distribution. *Anal Chem* 92: 15719–15725. https://doi.org/10.1021/acs.analchem.0c02228
56. Maryam A, Sebastian U, Daniela T (2025) hyPIRana v2.0.0 - analysis code for combined analysis of hyperspectral data in mid-IR photo-induced force microscopy (PiF-IR). Available from: https://doi.org/10.5281/zenodo.15270457
57. Pascual Robledo I, Maciel-Escudero C, Schnell M, et al. (2025) Theoretical description of infrared near-field spectroscopy of in- and out-of-plane molecular vibrations in thin layers. *ACS Photonics* 12: 3782–3793. https://doi.org/10.1021/acsphotonics.5c00798
58. Anindo ST, Täuber D, David C (2025) Photothermal expansion of nanostructures in photoinduced force microscopy. *J Phys Chem C* 129: 4517–4529. https://doi.org/10.1021/acs.jpcc.4c08305
59. Oberg KA, Ruysschaert JM, Goormaghtigh E (2004) The optimization of protein secondary structure determination with infrared and circular dichroism spectra. *Eur J Biochem* 271: 2937–2948. https://doi.org/10.1111/j.1432-1033.2004.04220.x
60. Wishart DS, Sykes BD (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. *J Biomol NMR* 4: 171–180. https://doi.org/10.1007/BF00175245
61. Spera S, Bax A (1991) Empirical correlation between protein backbone conformation and Cα and Cβ 13C nuclear magnetic resonance chemical shifts. *J Am Chem Soc* 113: 5490–5492. https://doi.org/10.1021/ja00014a071
62. Hafsa NE, Wishart DS (2014) CSI 2.0: a significantly improved version of the chemical shift index. *J Biomol NMR* 60: 131–146. https://doi.org/10.1007/s10858-014-9863-x
63. Atkinson RA (2025) NMR of proteins and nucleic acids. *Nucl Magn Reson* 51: 196–224. https://doi.org/10.1039/9781837679966-00196
64. Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56: 227–241. https://doi.org/10.1007/s10858-013-9741-y
65. Wüthrich K (1986) NMR with proteins and nucleic acids. *Europhys News* 17: 11–13. https://doi.org/10.1051/epn/19861701011

66. Vuister GW, Bax A (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNHα) coupling constants in 15N-enriched proteins. *J Am Chem Soc* 115: 7772–7777. https://doi.org/10.1021/ja00070a024
67. Wang AC, Bax A (1996) Determination of the backbone dihedral angles φ in human ubiquitin from reparametrized empirical Karplus equations. *J Am Chem Soc* 118: 2483–2494. https://doi.org/10.1021/ja9535524
68. Wider G, Wüthrich K (1999) NMR spectroscopy of large molecules and multimolecular assemblies in solution. *Curr Opin Struct Biol* 9: 594–601. https://doi.org/10.1016/S0959-440X(99)00011-1
69. Pervushin K, Riek R, Wider G, et al. (1997) Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94: 12366–12371. https://doi.org/10.1073/pnas.94.23.12366
70. Tugarinov V, Hwang PM, Ollerenshaw JE, et al. (2003) Cross-correlated relaxation enhanced 1H–13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* 125: 10420–10428. https://doi.org/10.1021/ja030153x
71. Taylor GL (2010) Introduction to phasing. *Acta Crystallogr D* 66: 325–338. https://doi.org/10.1107/S0907444910006694
72. Rupp B (2009) Biomolecular crystallography: principles, practice, and application to structural biology, New York: Garland Science.
73. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22: 2577–2637. https://doi.org/10.1002/bip.360221211
74. Touw WG, Baakman C, Black J, et al. (2015) A series of PDB-related databanks for everyday needs. *Nucleic Acids Res* 43: D364–D368. https://doi.org/10.1093/nar/gku1028
75. Wlodawer A, Minor W, Dauter Z, et al. (2013) Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS J* 280: 5705–5736. https://doi.org/10.1111/febs.12495
76. Chen VB, Arendall WB, Headd JJ, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* 66: 12–21. https://doi.org/10.1107/S0907444909042073
77. Cheng Y (2018) Single-particle cryo-EM—how did it get here and where will it go. *Science* 361: 876–880. https://doi.org/10.1126/science.aat4346
78. Nakane T, Kotecha A, Sente A, et al. (2020) Single-particle cryo-EM at atomic resolution. *Nature* 587: 152–156. https://doi.org/10.1038/s41586-020-2829-0
79. Dubochet J, Adrian M, Chang JJ, et al. (1988) Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* 21: 129–228. https://doi.org/10.1017/S0033583500004297
80. Subramaniam S, Earl LA, Falconieri V, et al. (2016) Resolution advances in cryo-EM enable application to drug discovery. *Curr Opin Struct Biol* 41: 194–202. https://doi.org/10.1016/j.sbi.2016.07.009
81. Wang RY-R, Kudryashev M, Li X, et al. (2015) De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nat Methods* 12: 335–338. https://doi.org/10.1038/nmeth.3287

82. Nogales E (2016) The development of cryo-EM into a mainstream structural biology technique. *Nat Methods* 13: 24–27. https://doi.org/10.1038/nmeth.3694
83. Castells-Graells R, Meador K, Arbing MA, et al. (2023) Cryo-EM structure determination of small therapeutic protein targets at 3 Å-resolution using a rigid imaging scaffold. *Proc Natl Acad Sci USA* 120: e2305494120. https://doi.org/10.1073/pnas.2305494120
84. Danev R, Yanagisawa H, Kikkawa M (2019) Cryo-electron microscopy methodology: current aspects and future directions. *Trends Biochem Sci* 44: 837–848. https://doi.org/10.1016/j.tibs.2019.04.008
85. Putnam CD, Hammel M, Hura GL, et al. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* 40: 191–285. https://doi.org/10.1017/S0033583507004635
86. Kikhney AG, Svergun DI (2015) A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* 589: 2570–2577. https://doi.org/10.1016/j.febslet.2015.08.027
87. Englander SW, Kallenbach NR (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q Rev Biophys* 16: 521–655. https://doi.org/10.1017/S0033583500005217
88. Wales TE, Engen JR (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev* 25: 158–170. https://doi.org/10.1002/mas.20064
89. Stryer L (1978) Fluorescence energy transfer as a spectroscopic ruler. *Annu Rev Biochem* 47: 819–846. https://doi.org/10.1146/annurev.bi.47.070178.004131
90. Wien F, Miles AJ, Lees JG, et al. (2005) VUV irradiation effects on proteins in high-flux synchrotron radiation circular dichroism spectroscopy. *J Synchrotron Radiat* 12: 517–523. https://doi.org/10.1107/S0909049505006953

AIMS Press