Purification and Structural Characterization of N-Terminal 190 Amino Acid Deleted Essential Mammalian Protein; Transcription Termination Factor 1

Kumud Tiwari, Gajender Singh, and Samarendra Kumar Singh*

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: The mammalian transcription termination factor 1 (TTF1) is an essential protein that plays diverse cellular physiological functions like transcription regulation (both initiation and termination), replication fork blockage, chromatin remodeling, and DNA damage repair. Hence, understanding the structure and mechanism conferred by its variable conformations is important. However, so far, almost nothing is known about the structure of either the full-length protein or any of its domains in isolation. Since the full-length protein even after multiple attempts could not be purified in soluble form, we have codon optimized, expressed, and purified the N-terminal 190 amino acid deleted TTF1 (ΔN190TTF1) protein. In this study, we characterized this essential protein by studying its homogeneity, molecular size, and secondary structure using tools like dynamic light scattering (DLS), circular dichroism (CD) spectroscopy, Raman spectroscopy, and atomic force microscopy (AFM). By CD spectroscopy and DLS, we confirmed that the purified protein is homogeneous and soluble. CD spectroscopy also revealed that ΔN190TTF1 is a helical protein, which was further established by analysis of Raman spectra and amide I region deconvolution studies. The DLS study estimated the size of a single protein molecule to be 17.2 nm (in aqueous solution). Our structural and biophysical characterization of this essential protein will open avenues toward solving the structure to atomic resolution and will also encourage researchers to investigate the mechanism behind its diverse functions attributed to its various domains.
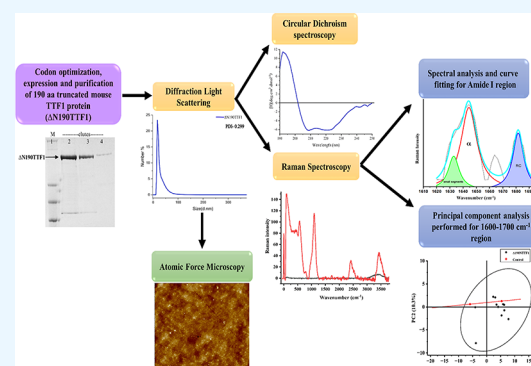
INTRODUCTION

Eukaryotic DNA replication and transcription are vital processes occurring in three prime stages: initiation, elongation, and termination. Regulation of these stages is critically important specially during the S phase of the cell cycle when certain regions of chromosomal DNA (like ribosomal DNA) are often transcribed and replicated simultaneously. Therefore, the process of termination (polar fork arrest) is physiologically necessary to avoid collision between replication and transcription machinery, which could lead to genomic instability and cell transformation. Polar fork arrest in mammalian cells takes place at certain defined regions of the chromosomes, and among such regions are non-transcribed spacer regions of ribosomal DNA (rDNA) flanking the area that codes for ribosomal protein. These particular sequences, situated 170 base pairs upstream and downstream of the rDNA gene repeats in mammalian cells, are known as Sal box elements (Figure 1a). In mammalian cells, these regulatory elements are responsible for site-specific Pol1 transcription termination and DNA replication fork arrest in a polar fashion, which is mediated by the multifunctional protein transcription termination factor 1 (TTF1) upon binding with Sal box sequences.[1] Orthologues of TTF1 proteins have been found in

a variety of organisms, including Rib2 of *Xenopus*, RNA polymerase I enhancer binding protein (Reb1) of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, and the mitochondrial transcription termination factor (mTERF) in human mitochondria.[2,3] Being a multifunctional protein, TTF1 is engaged in various cellular functions like chromatin remodeling, cell cycle regulation, and DNA damage sensing.[4] To execute these processes, it also interacts with numerous regulatory proteins, including TTF1 interacting protein (TIP 5), Pol I and transcript release factor (PTRF), mouse double minute 2 (MDM2), Cockayne syndrome B (CSB), and alternative reading frame (ARF), and catalyzes critical cellular functions in humans.[4−9] These diverse activities make TTF1 a truly multifunctional and essential protein for the survival of the cell, the mis-regulation of which has been linked with various cancers.[10−12]
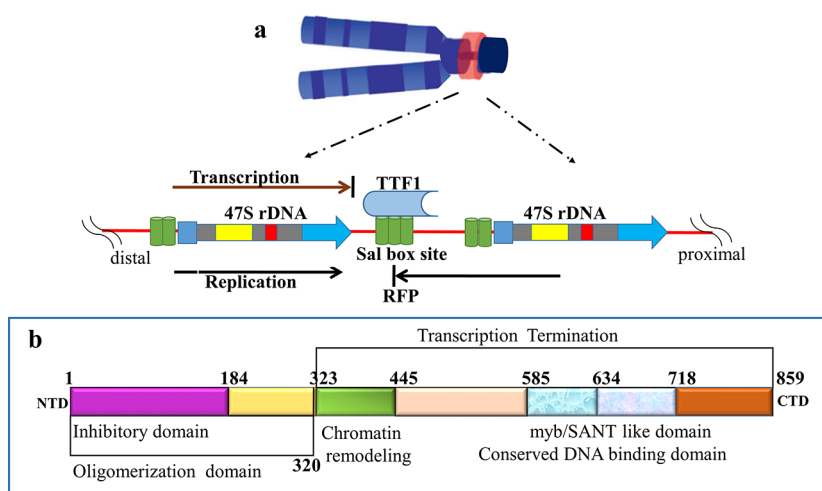
Figure 1. (a) Diagram showing a mouse chromosome de-condensed rDNA region (nucleolar organizing region (NOR), red); schematic representation of the non-transcribed spacer region of rDNA showing the rDNA genic region (47S),[14] Sal box site, and fork pausing site RFP. The TTF1 protein binds to the Sal box and arrests replication from one direction (black arrow) and RNA pol1 mediated transcription from the opposite direction (brown arrow).[1] (b) Domain architecture of mouse TTF1 protein showing the N-terminal inhibitory domain (1–184), oligomerization domain (1–320), chromatin remodeling domain (323–445), conserved DNA-binding domain (myb/SANT-like domain 585–718), and transcription termination domain (323–859).[13,15]

The domain architecture of mouse TTF1 comprises a putative N-terminal regulatory domain, a DNA-binding trans-activation domain (myb/SANT domain), and a C-terminal transcription termination domain. The N-terminal region of the protein, working as a flap in oligomeric state, represses its DNA-binding activity (Figure 1b).[13]

However, to date, no research on its structural, computational, or physical characterization has been reported yet. The very first article that reports ab initio modeling of this essential protein has been communicated from our lab.[16] Considering the multifunctional and essential nature of TTF1, it becomes important to elucidate the structural knowledge in order to understand its mechanisms of action and use the same toward its therapeutic applications. Because multiple attempts to purify soluble full-length TTF1 had been unsuccessful, we cloned the codon-optimized N-terminal 190 amino acid truncated TTF1 (ΔN190TTF1; to get rid of the inhibitory flap) mouse gene and expressed it in bacterial cells. In this article, we present biophysical and structural characterization of this essential protein using dynamic light scattering (DLS), circular dichroism (CD) spectroscopy, Raman spectroscopy, and atomic force microscopy (AFM), which will open up the field for further structural characterization of the protein to the atomic level and make it available for understanding its mechanistic function in detail. This work would be the first ever report of biophysical and structural characterization of the essential mammalian protein TTF1.

■ MATERIALS AND METHODS

Sequence-Based Analyses. The amino acid sequence of the mouse (*Mus musculus*) TTF1 protein was retrieved from UniProtKB (UniProtKB accession number: Q62187), and the physicochemical properties of TTF1 were analyzed using the ProtParam computational tool.[17] The Gravy index score of proteins was measured by the Kyte–Doolittle formula.[18]

Expression and Purification of Recombinant TTF1. The codon-optimized N-terminal 190 amino acid (aa) truncated TTF1 (ΔN190TTF1) mouse gene was cloned in a pMAL-c2X vector with a N-terminal His tag and expressed in the BL21-λDE3 bacterial strain. The tagged protein was then purified using a Ni-NTA column, the further details of which are presented in the extended experimental procedures in the Supporting Information.

Dynamic Light Scattering. DLS analysis was performed with a Zetasizer Nano ZS (Malvern, Worcestershire, UK) equipped with 4 mW He–Ne 633 nm laser with a detection angle of 173° backscatter at 25 °C, in a low-volume (45 μL) quartz cuvette. Measurements of intensity and volume in multiple narrow mode were used to derive the size distribution. Measurements were performed at least in triplicate with 15 μg/μL protein in 25 mM Tris-Cl at pH 7.5. The dispersity (polydispersity index (PDI)) and the particle size were determined at 25 °C performing four runs (five measurements each).

CD Spectroscopy. CD spectra were recorded using a Jasco J-1500 spectropolarimeter equipped with a Peltier thermo-stated cell holder (Jasco, Easton, MD, United States). Purified ΔN190TTF1 protein was diluted to 15 ng/μL in buffer A (25 mM Tris-Cl at pH 7.5 and 100 mM KCl). Data collection and analysis were carried in triplicate as described by Colarusso et al.[19] Secondary structure evaluation was performed using the Jasco[20] and Origin software (version 2022, Origin Lab Corp., Northampton, MA, United States),[21] following the standard procedure.[22]

Raman Spectroscopy. The Raman spectra were recorded using an alpha300 (WITec, Germany), which has a liquid-nitrogen-cooled charge-coupled device (CCD) detector and a spectrograph with a 600 g/mm grating with a resolution of 1 cm⁻¹. Backscattering geometry for spectrum collection was applied with a notch filter to reject the elastic contribution. A laser emitted at 532 nm with a power of 44 mW was employed as the excitation source (as described in the extended experimental procedures in the Supporting Information).

Raman Spectrum Recording. The procedure used here was modification of that by Signorelli et al.,[23] which is described in the extended experimental procedures in the Supporting Information.

Curve-Fitting Procedure. Curve fitting of the amide I band and its assignment to the structural components were performed essentially as described by Maiti et al.[24] Curve fitting was performed via the Origin software (version 2022).[21]

Principal Component Analysis. Principal component analysis (PCA) was applied on Raman spectroscopy data with 10 spectra, being satisfactorily described by a principal component number of 21. The correlation matrix was analyzed with the Origin software (version 2022),[25] and the number of components to be considered was defined as the number required to explain at least 70% of the total variance.

Atomic Force Microscopy. For AFM studies, purified ΔN190TTF1 protein was diluted to 11 ng/μL in 25 mM Tris-Cl buffer at pH 7.5 (filtered) and 10 μL of this diluted protein sample was deposited on a fresh silicon wafer and dried by desiccation. AFM measurements were performed with the P9 SPM digital control platform (Solver Next, NT-MDT, Russia). Images were acquired using a silicon tip (NHG-0.1 model) in tapping mode or semi-contact mode. Image processing and analysis were performed by NT-MDT Nova, which was provided by the manufacturer of the machine.

■ RESULTS AND DISCUSSION

Analysis with ProtParam. Analysis with ProtParam (amino acid based) showed that the full-length and 190 aa deleted mouse TTF1 are hydrophilic proteins (instability indices of 57.28 and 46.44) and have grand averages of hydrophobicity and hydrophilicity (Gravy) of −1.009 and −0.811, respectively. After deleting 190 aa, theoretically, we can see the instability index decrease from 57.28 to 46.66, which indicates that the truncated protein is much more stable than the full-length one (Table 1).[26] Proteins with hydro-

Table 1. Physicochemical Properties of Mouse TTF1, as Determined with ProtParam

physicochemical properties	values	
	full-length mouse TTF1	190 aa deleted mouse TTF1
number of residues	859	669
molecular formula	$C_{4262}H_{6916}N_{1256}O_{1333}S_{20}$	$C_{3317}H_{5323}N_{951}O_{1043}S_{17}$
molecular weight	97,722.61 Da	75,758.52 Da
theoretical pI	9.41	7.93
instability index	57.28	46.66
aliphatic index	65.42	73.06
total number of negatively charged residues (D + E)	133	112
total number of positively charged residues (R + K)	166	114
grand average of hydrophilicity (Gravy)	−1.009	−0.811
estimated half-life (mammalian reticulocyte, in vitro)	30 h	30 h

phobicity scores below 0 are more likely to be globular (hydrophilic), while those with scores above 0 are more likely to be membranous (hydrophobic) in nature;[27] therefore, TTF1 is supposed to be globular and hydrophilic. Various predicted physicochemical properties of the protein are listed in Table 1.

Purification of ΔN190TTF1. ΔN190TTF1 was cloned in a pMAL-c2X vector with a N-terminal His tag. It was expressed in the BL21-λDE3 bacterial strain and induced with 0.9 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Figure 2a). The induced TTF1 was extracted from lysed bacterial cells and bound with the Ni-NTA column. After a 30 column volume (CV) wash (as mentioned in the Materials and Methods section), the protein was eluted in a elution buffer (wash buffer plus 400 mM imidazole, Figure 2b). The protein at this stage was 1 mg/mL and more than 95% homogeneously pure (lanes 2−4, Figure 2b).

DLS Study Reveals That ΔN190TTF1 Is Monodispersed. In order to make sure that the purified protein from the Ni-NTA column is soluble and homogeneous, we next performed DLS to know its dispersity. DLS is used for measuring particle sizes ranging from nanometers to micrometers. In this procedure, by processing the time series, we could find the diffusion coefficient of the molecule.[28] For this process, the concentrated protein was diluted to 15 μg/μL in 25 mM Tris-Cl at pH 7.5. The analysis of the results obtained from size distribution by number (which is more reliable; Figure 3a)[29] showed that the solution contained a single particle size with a mean hydrodynamic radius (R_H) of 17.2 nm, confirming the mono-dispersive nature of the solution with a PDI of 0.299 (Figure 3a,b and Table 2).[29] This evaluation allowed us to move ahead and do structural characterization of the purified protein.

CD Spectroscopy Confirms That ΔN190TTF1 Is a Soluble and Helical Protein. After confirming that the purified protein is monodispersed and homogeneous, we wanted to know the secondary structure of the protein by performing CD spectroscopy (as described in the Materials and Methods section). CD spectroscopy is a versatile tool in structural biology, with an increasingly wide range of applications such as rapid determination of the secondary structure and folding properties of proteins (obtained using recombinant techniques or purified from cells or tissues). In the far-UV region (240−180 nm), which corresponds to peptide bond absorption, the CD spectrum can be analyzed to give the content of regular secondary structural features such as α helices and β sheets. CD spectroscopy, being a faster technique requiring less sample, plays a significant role in structural characterization and is often used as a confirmatory method for secondary structure determination.[22,30−32] The mean residue molar ellipticity, $[\theta]$, in degrees cm^2 dmol^{-1}, has been calculated from

$$[\theta] = [\theta]_{obs} \times MRW/(10cl)$$

where $[\theta]_{obs}$ is the measured molar ellipticity in degrees, MRW is the mean residue molecular weight, c is the protein concentration in grams per milliliter, and l is the optical path length in centimeters.[19] The far-UV CD spectrum of the ΔN190TTF1 protein (Figure 4), collected in buffer A at 20 °C, is characterized by a sharp maximum at 192 nm and two slightly pronounced minima at 209 and 221 nm. The secondary structure estimation after spectral deconvolution (as mentioned in the Materials and Methods section) gave estimated percentages of α helices (53.46%), β sheets (9.7%), random coils (23.78%), and turns (12.88%) (average of five runs, Table 3), confirming the predominantly helical nature of the protein.[33] The above finding is very much in agreement with our computational ab initio study and simulation data on human TTF1 protein.[16]

Raman Spectroscopy of ΔN190TTF1. Raman spectroscopy is an important tool for studying the secondary structures
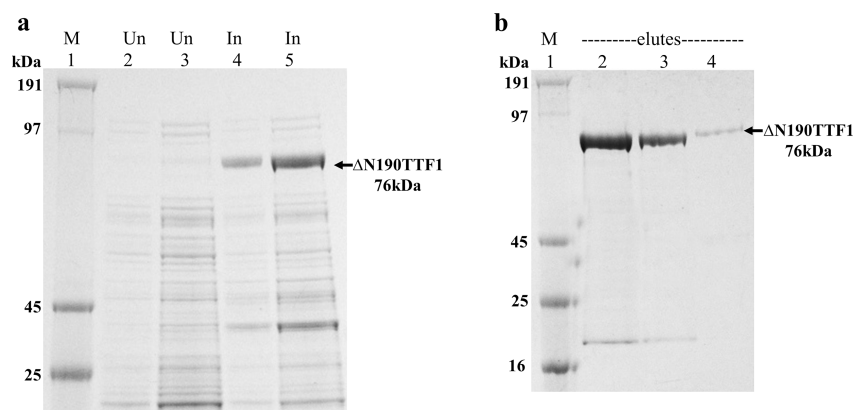
Figure 2. (a) Induction profile of the ΔN190TTF1 protein; "In" represents an induced expression profile of the protein (lanes 4 and 5), and "Un" represents an un-induced expression profile of the same (lanes 2 and 3). (b) Purification profile of the ΔN190TTF1 protein: lanes 2−4 show eluted TTF1. M represents the molecular weight marker.
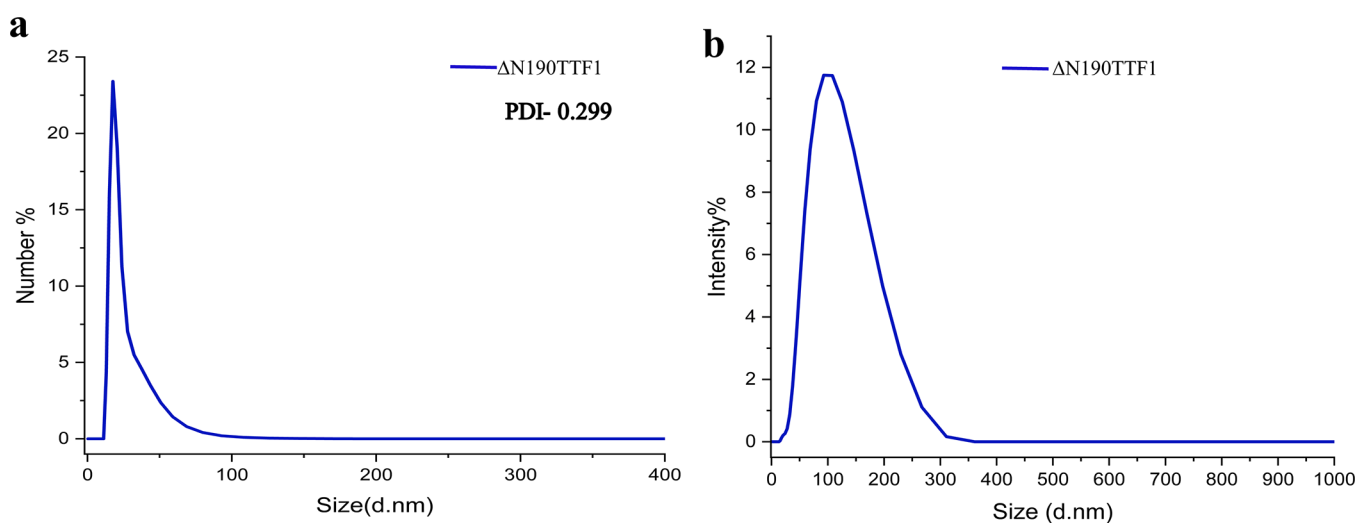

Figure 3. DLS analysis plots of purified ΔN190TTF1 protein: (a) size distribution by number plot; (b) size distribution by intensity plot.

Table 2. Properties of the ΔN190TTF1 Protein in 25 mM Tris Buffer at pH 7.5 after DLS Analysis

SN	properties	mean
1	polydispersity index (PDI)	0.299
2	peak one mean by intensity (nm)	17.2

of proteins and peptides. The Raman spectral profile of a protein in solution is a collection of signals arising from each conformation, thus providing an "instantaneous snapshot" of the population.[24,34] The majority of Raman spectral signals are produced by out-of-phase vibration due to C−N stretching, C−C−N deformation, and N−H in-plane bending. The spectrum presents a complex set of peaks and overlapping bands, among which the main spectra due to aromatic amino acids and polypeptide chain can be recognized. Hence, we performed Raman spectroscopy to characterize the secondary structure and also to understand the population heterogeneity of the purified TTF1. The Raman spectra of ΔN190TTF1 are shown in red color and that of the buffer (without protein, control) in black in Figure 5. In the spectra, the peaks associated with the amino acid Phe are located at 1000 cm^{-1}. The peaks around 1341 cm^{-1} are related to Trp vibrations. At about 1250 cm^{-1}, the amide III band could be observed, which occurs due to vibrations produced by in-phase combination of
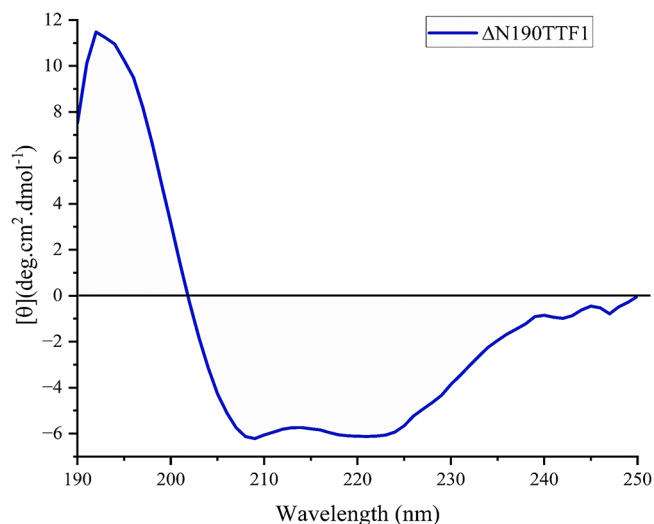
Figure 4. CD spectrum of the ΔN190TTF1 protein. The X axis represents the wavelength of the spectrum, and the Y axis represents the molar ellipticity in degrees.

N−H bending and C−N stretching and contributes to a mixture of secondary structures in the protein solution. The

Table 3. Secondary Structure Percentage Estimated after Deconvolution of CD Spectra[20]

SN	sec structure	percentage (%)
1	α helix	53.46
2	β sheet	9.7
3	random coil	23.78
4	turns	12.88

peak around 1550 cm^{-1} is the amide II band. Finally, the amide I band, which is found within the 1620−1725 cm^{-1} spectral region, could be observed, which is caused by vibration produced mainly due to C1/4 O stretching with minor contributions. The spectrum arising from amide I modes comes from the vibrational coupling between the motions of the peptide carbonyl groups organized in an ordered secondary structure through hydrogen bonds. An analysis of this band can provide detailed information on the structural conformation of the proteins in physiological conditions (Figure 5b). Since the amide I spectrum could significantly characterize the secondary structure of a protein, we described and analyzed the same in detail for ΔN190TTF1 (Table 4 and Figures 5 and 6).[23,24,34,35]

Curve Fitting for the Amide I Region and Deconvolution. The amide I band has been used to validate the atomic structures of several proteins and also to study intrinsically disordered proteins and peptides. For these reasons, we too performed a curve-fitting procedure for the amide I spectrum of our protein.[23,24] The aim of spectral curve fitting is to observe and quantify the general trends in the measured data by effectively removing noise. It is carried out by fitting the measured data into an analytical equation (to extract meaningful parameters). The data estimation is carried out via interpolation between discrete values in order to obtain a maximum or a minimum value, thereby deriving finite-difference approximation.[36] In particular, the band was deconvoluted in terms of three major components separately, which are associated with the α-helix, β-sheet, and random coil structures. In particular, the amide I band was fitted by considering a curve centered at 1650−1656 cm^{-1}, which could

Table 4. Raman Bands of ΔN190TTF1 Protein Spectra with Their Assignments[34,35]

source	frequency (Raman band position, cm^{-1})	assignment
amide I	1650−1657	α helices
	1680−1690	turn
amide II	1550	parallel/anti parallel β sheet
amide III	1243−1253	random coil
	1229−1235	β sheet
tryptophan	1340−1355	relative ratio is indicative of hydrophobicity
phenylalanine	1000	CN stretching
disulfide	500−550	S−S bond stretching
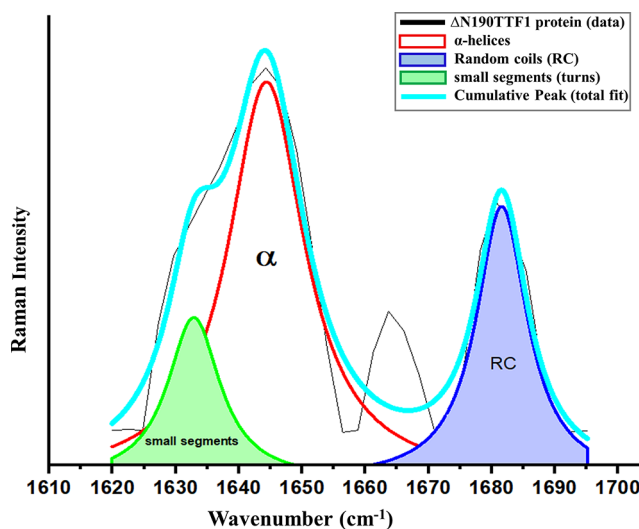
Figure 6. Raman spectra after curve fitting for the experimental amide I data.

be assigned to an α helix; a second one at the 1664−1670 cm^{-1} region due to β sheets; a third one at about 1680 cm^{-1}, corresponding to random coils; and the fourth at 1640−1650 cm^{-1}, attributed to the turns that connects helices. Table 4 lists
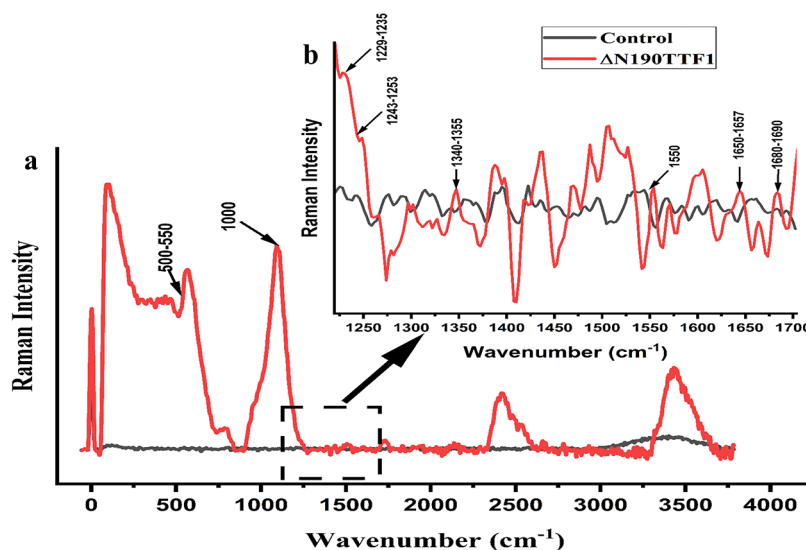
Figure 5. (a) Raman spectrum of the ΔN190TTF1 protein; (b) Raman spectra for the 1200−1700 cm^{-1} region (magnified). The TTF1 spectrum is shown in red, while that of the control (buffer without protein) is shown in black.

a summary of the corresponding fitting parameters. The contribution from the appropriate secondary structures was connected with the integrated curve intensities fitting the amide I band under the presumption that the Raman cross section is the same for each conformation. The fitting curves of the amide I Raman bands of the ΔN190TTF1 protein proclaim the major area (42.44%) of the curve for α helices (white area under the red curve), which clearly shows that TTF1 mostly contains helices. The rest curve area is divided into random coils (blue curve, 21.1%) and turns (green curve, 13.03%) (Figure 6). The amide I band full width at half-maximum (FWHM) and area percentage of each curve of secondary structures for the ΔN190TTF1 protein are listed in Table 5. Deconvolution of spectra collected from CD and

Table 5. Results from the Curve-Fitting Procedure of the Amide I Band of ΔN190TTF1

| secondary structure	frequency (cm^{-1})	area (%)	FWHM[a] (cm^{-1})
α helix	1646	42.44	15.95
small segments (turns)	1635	13.03	11.75
random coil	1681	21.34	11.44

[a]FWHM, full width at half maximum.

Raman spectroscopy shows that TTF1 is a helical protein, and this finding justifies its DNA-binding activity. The above finding is in agreement with our computational model for the full-length TTF1 as well.[16] Previous publications reporting protein DNA atomic structures have shown that the DNA is stabilized between the helices of those proteins.[2,37]

PCA of the Amide I Region. Additionally, the amide I region (1600−1700 cm^{-1}) was analyzed through the PCA technique, which is a multivariate statistical analysis method that allows one to reduce the multi-dimensionality of the data set by recollecting the characteristics of the original ensemble, primarily contributing to its quantitative variance. Concisely, a limited number of variables, known as principal components (PCs), contain most of the spectral information and they can be divided into patterns, scores, and loadings. Figure 7a shows the PCA scores of the PC1 vs the PC2 component for the amide I band of ΔN190TTF1 (experimental, black dots) and the buffer (control, red dots). The PC1 and PC2 components

of the amide I band of ΔN190TTF1 combined shows 81% of the total variance. The plot shown in Figure 7a represents the distribution of the spectra for a given PC and reveals the relationship existing among the samples. Meanwhile, the loading plot shown in Figure 7b represents how much an original variable, or range of variables, influences a given PC for a particular defined region. Hence, PCA is particularly effective for categorizing Raman spectra that are otherwise hardly distinguishable.[23,38] Along the PC1 axis, two separate clusters can be seen. Most of the scores of ΔN190TTF1 are located in the positive portion of the plot along PC1 with a spread along both axes, while the control scores also display both negative and positive PC1 values with a large variability along PC2. The PC2 loading plot, providing complementary information to PC1, shows a significant variation in the 1600−1700 cm^{-1} region. Variation specifically corresponding to the tyrosine peak (1613−1616 cm^{-1}) could be observed, which occurs mainly due to a change in the strength of the hydrogen bond between water and tyrosine molecules. Downshifting is observed in the band frequency in the 1670−1685 cm^{-1} region when helices form coils. A high spectral variance also suggests the presence of α helices (1646 cm^{-1}), turns (1635 cm^{-1}), and random coils (1681 cm^{-1}),[35] as previously observed (Figures 5b and 7b). PCA strengthens our Raman spectral data via introducing theoretical statistical calculation to show the quantitative variability in buffer components and the existence of the ΔN190TTF1 protein in the optimized buffer.

Atomic Force Microscopy. AFM imaging of single protein molecules are used to quantitatively evaluate their dimensions and is quite useful to determine their geometrics per unit area of interface, which eventually facilitates the determination of the size of the molecule.[28] For this, an area of 10 μm × 10 μm was scanned and the topography of the surface was analyzed as shown in Figure 8. After the calculations (as shown in the micrograph of Figure 8a), the average size of the molecule was determined to be 0.094 μm (94 nm, Figure 8c). The difference in the size from that of DLS data could be due to surface accumulation of protein molecules (aggregation) during sample preparation (spreading and drying, which are very common).[39] In Figure 8b, the differences in the baselines of the top and bottom Z axes indicate the dimension of protein
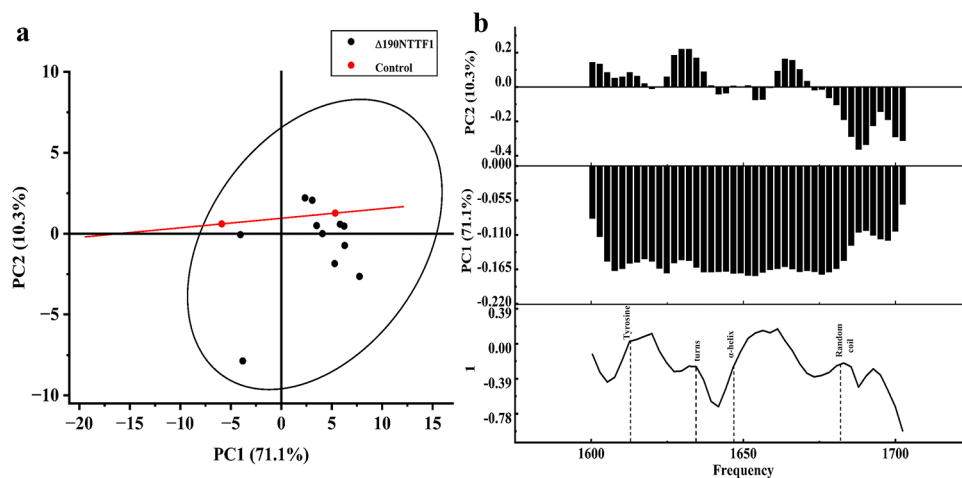
Figure 7. (a) Two-dimensional score plot of PC1 vs PC2 of the Raman spectra for ΔN190TTF1 (black dots) and the control (buffer; red dots) performed on the amide I band. The protein group is illustrated as an ellipse and the control group as a line. (b) One-dimensional loading plot versus frequency (PC2 10.3% and PC1 71.1% of total variance).
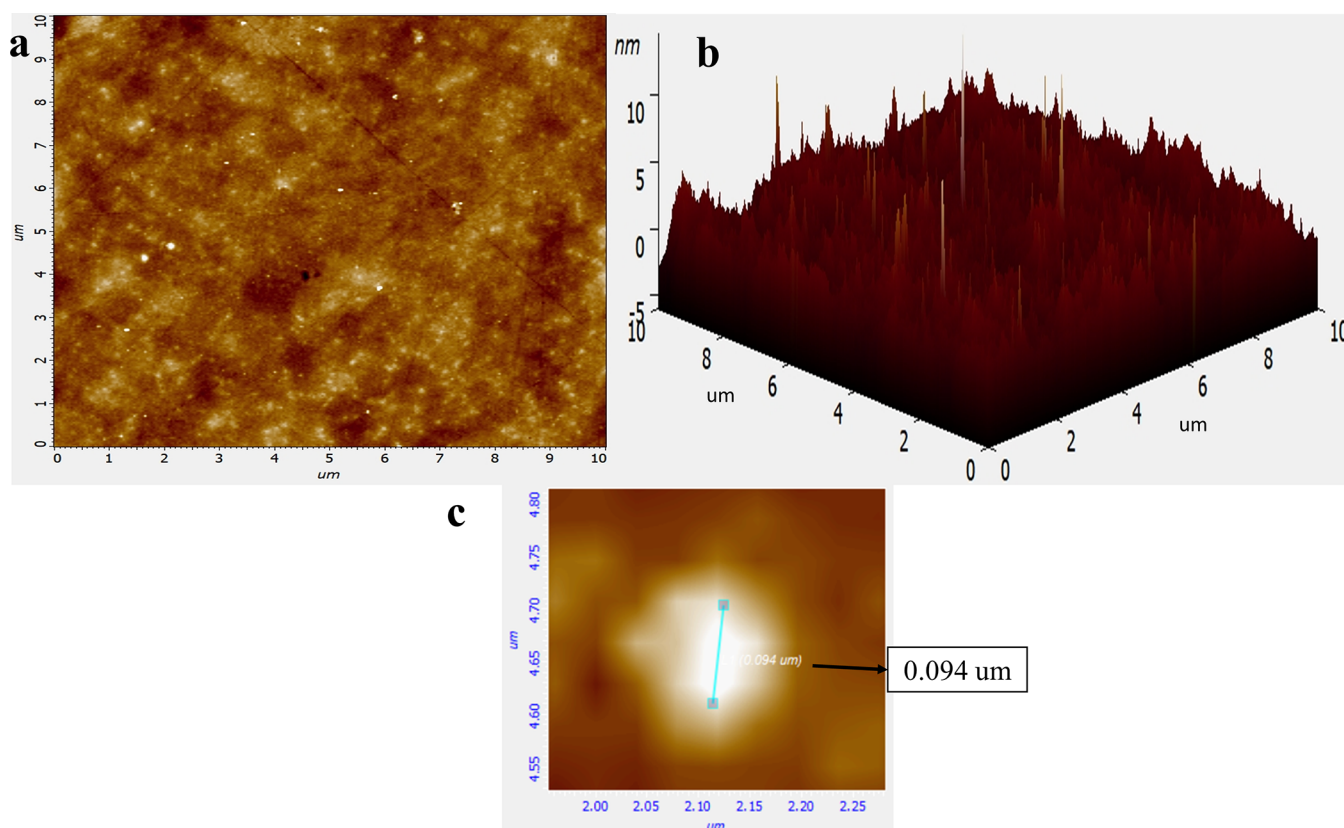
Figure 8. 10 × 10 μm AFM micrographs of the ΔN190TTF1 protein: (a) 2D ΔN190TTF1 image; (b) 3D ΔN190TTF1 image; (c) ΔN190TTF1 protein molecule under higher magnification.

molecules. Hence, the AFM study confirmed the homogeneity of the purification and revealed molecular assembly of the protein as well.

CONCLUSIONS

Even though TTF1 is an essential multifunctional protein, almost nothing is known about its biophysical character or structure. The only study revealing its structural model (computationally) is reported from our lab.[16] Hence, it is important to understand the biophysical characters of this protein in order to engineer the same for structural and mechanistic studies. Toward achieving the above objectives, we first assayed the homogeneity of the purified protein and then characterized the secondary structure. The purified protein profile on gel and DLS and AFM data confirmed the homogeneity and stability (in the optimized buffer) of the preparation, which gave us confidence to proceed further with structural characterization. CD and Raman spectroscopy established that ΔN190TTF1 is a helical protein. The deconvolution of the amide I band, remarkably sensitive to the α-helix, β-sheet, and random coil structures, allowed us to assess the main structural motifs of these proteins. In particular, a careful fitting procedure of this band revealed that about ~50% of native ΔN190TTF1 displays α helices, while the rest is almost equally shared between random coil and small segment structures. Our lab is actively working to solve the physical structure of this protein by cryo-electron microscopy, and our current finding in this article infuses confidence in us that we are proceeding in the right direction. Knowing the biophysical and structural properties of this essential protein will significantly help researchers to know the

effect of mutation and its role in diseases like Cockayne syndrome B and various others to be discovered yet.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c05603.

Extended experimental procedures of expression and purification of recombinant TTF1 and Raman spectrum recording (PDF)

AUTHOR INFORMATION

Corresponding Author

Samarendra Kumar Singh − *Cell Cycle and Cancer Laboratory, School of Biotechnology, Institute of Science, Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India;* ⓘ orcid.org/0000-0002-2001-0025; Phone: +91-8009561887; Email: samarendra.singh@bhu.ac.in, biotech3@rediffmail.com

Authors

Kumud Tiwari − *Cell Cycle and Cancer Laboratory, School of Biotechnology, Institute of Science, Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India*
Gajender Singh − *Cell Cycle and Cancer Laboratory, School of Biotechnology, Institute of Science, Banaras Hindu University, Varanasi, Uttar Pradesh 221005, India*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.2c05603

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Akamatsu, Y.; Kobayashi, T. The Human RNA Polymerase I Transcription Terminator Complex Acts as a Replication Fork Barrier That Coordinates the Progress of Replication with RRNA Transcription Activity. *Mol. Cell. Biol.* 2015, 35, 1871−1881.

(2) Jaiswal, R.; Choudhury, M.; Zaman, S.; Singh, S.; Santosh, V.; Bastia, D.; Escalante, C. R. Functional Architecture of the Reb1-Ter Complex of Schizosaccharomyces Pombe. *Proc. Natl. Acad. Sci. U. S. A.* 2016, 113, E2267−E2276.

(3) Roberti, M.; Polosa, P. L.; Bruni, F.; Manzari, C.; Deceglie, S.; Gadaleta, M. N.; Cantatore, P. The MTERF Family Proteins: Mitochondrial Transcription Regulators and Beyond. *Biochim. Biophys. Acta, Bioenerg.* 2009, 1787, 303−311.

(4) Nemeth, A. The Chromatin Remodeling Complex NoRC and TTF-I Cooperate in the Regulation of the Mammalian RRNA Genes in Vivo. *Nucleic Acids Res.* 2004, 32, 4091−4099.

(5) Németh, A.; Perez-Fernandez, J.; Merkl, P.; Hamperl, S.; Gerber, J.; Griesenbeck, J.; Tschochner, H. RNA Polymerase I Termination: Where Is the End? *Biochim. Biophys. Acta, Gene Regul. Mech.* 2013, 1829, 306−317.

(6) Lessard, F.; Morin, F.; Ivanchuk, S.; Langlois, F.; Stefanovsky, V.; Rutka, J.; Moss, T. The ARF Tumor Suppressor Controls Ribosome Biogenesis by Regulating the RNA Polymerase I Transcription Factor TTF-I. *Mol. Cell* 2010, 38, 539−550.

(7) Lessard, F.; Stefanovsky, V.; Tremblay, M. G.; Moss, T. The Cellular Abundance of the Essential Transcription Termination Factor TTF-I Regulates Ribosome Biogenesis and Is Determined by MDM2 Ubiquitinylation. *Nucleic Acids Res.* 2012, 40, 5357−5367.

(8) Boutin, J.; Lessard, F.; Tremblay, M. G.; Moss, T. The Short N-Terminal Repeats of Transcription Termination Factor 1 Contain Semi-Redundant Nucleolar Localization Signals and P19-ARF Tumor Suppressor Binding Sites. *YALE J. Biol. Med.* 2019, 92, 385−396.

(9) Aamann, M. D.; Muftuoglu, M.; Bohr, V. A.; Stevnsner, T. Multiple Interaction Partners for Cockayne Syndrome Proteins: Implications for Genome and Transcriptome Maintenance. *Mech. Ageing Dev.* 2013, 134, 212−224.

(10) Stults, D. M.; Killen, M. W.; Williamson, E. P.; Hourigan, J. S.; Vargas, H. D.; Arnold, S. M.; Moscow, J. A.; Pierce, A. J. Human RRNA Gene Clusters Are Recombinational Hotspots in Cancer. *Cancer Res.* 2009, 69, 9096−9104.

(11) Ueda, M.; Iguchi, T.; Nambara, S.; Saito, T.; Komatsu, H.; Sakimura, S.; Hirata, H.; Uchi, R.; Takano, Y.; Shinden, Y.; Eguchi, H.; Masuda, T.; Sugimachi, K.; Yamamoto, H.; Doki, Y.; Mori, M.; Mimori, K. Overexpression of Transcription Termination Factor 1 Is Associated with a Poor Prognosis in Patients with Colorectal Cancer. *Ann. Surg. Oncol.* 2015, 22, 1490−1498.

(12) Komatsu, H.; Iguchi, T.; Ueda, M.; Nambara, S.; Saito, T.; Hirata, H.; Sakimura, S.; Takano, Y.; Uchi, R.; Shinden, Y.; Eguchi, H.; Masuda, T.; Sugimachi, K.; Doki, Y.; Mori, M.; Mimori, K. Clinical and Biological Significance of Transcription Termination Factor, RNA Polymerase i in Human Liver Hepatocellular Carcinoma. *Oncol. Rep.* 2016, 35, 2073−2080.

(13) Evers, R.; Smid, A.; Rudloff, U.; Lottspeich, F.; Grummt, I. Different Domains of the Murine RNA Polymerase I-Specific Termination Factor MTTF-I Serve Distinct Functions in Transcription Termination. *EMBO J.* 1995, 14, 1248−1256.

(14) McStay, B.; Grummt, I. The Epigenetics of RRNA Genes: From Molecular to Chromosome Biology. *Annu. Rev. Cell Dev. Biol.* 2008, 24, 131−157.

(15) Sander, E. Oligomerization of the Transcription Termination Factor TTF-I: Implications for the Structural Organization of Ribosomal Transcription Units. *Nucleic Acids Res.* 1997, 25, 1142−1147.

(16) Tiwari, K.; Gangopadhyay, A.; Singh, G.; Singh, V. K.; Singh, S. K. Ab Initio Modelling of an Essential Mammalian Protein: Transcription Termination Factor 1 (TTF1). *J. Biomol. Struct. Dyn.* 2022, 1−10.

(17) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A.Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, 2005; pp. 571−607, DOI: 10.1385/1-59259-890-0:571.

(18) Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* 1982, 157, 105−132.

(19) Colarusso, A.; Caterino, M.; Fabbri, A.; Fiorentini, C.; Vergara, A.; Sica, F.; Parrilli, E.; Tutino, M. L. High Yield Purification and First Structural Characterization of the Full-Length Bacterial Toxin CNF1. *Biotechnol. Prog.* 2018, 34, 150−159.

(20) Yang, J. T.; Wu, C. S. C.; Martinez, H. M. Calculation of Protein Conformation from Circular Dichroism. *Methods Enzymol.* 1986, 130, 208−269.

(21) Ricciardi, A.; Piuri, G.; Porta, M. D.; Mazzucchelli, S.; Bonizzi, A.; Truffi, M.; Sevieri, M.; Allevi, R.; Corsi, F.; Cazzola, R.; Morasso, C. Raman Spectroscopy Characterization of the Major Classes of Plasma Lipoproteins. *Vib. Spectrosc.* 2020, 109, 103073.

(22) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* 2006, 2876−2890.

(23) Signorelli, S.; Cannistraro, S.; Bizzarri, A. R. Structural Characterization of the Intrinsically Disordered Protein P53 Using Raman Spectroscopy. *Appl. Spectrosc.* 2017, 71, 823−832.

(24) Maiti, N. C.; Apetri, M. M.; Zagorski, M. G.; Carey, P. R.; Anderson, V. E. Raman Spectroscopic Characterization of Secondary Structure in Natively Unfolded Proteins: α-Synuclein. *J. Am. Chem. Soc.* 2004, 126, 2399−2408.

(25) Brozek-Pluska, B. Statistics Assisted Analysis of Raman Spectra and Imaging of Human Colon Cell Lines − Label Free, Spectroscopic Diagnostics of Colorectal Cancer. *J. Mol. Struct.* 2020, 1218, 128524.

(26) Guruprasad, K.; Reddy, B. V. B.; Pandit, M. W. Correlation between Stability of a Protein and Its Dipeptide Composition: A Novel Approach for Predicting in Vivo Stability of a Protein from Its Primary Sequence. *Protein Eng., Des. Sel.* 1990, 4, 155−161.

(27) Magdeldin, S.; Yoshida, Y.; Li, H.; Maeda, Y.; Yokoyama, M.; Enany, S.; Zhang, Y.; Xu, B.; Fujinaka, H.; Yaoita, E.; Sasaki, S.; Yamamoto, T. Murine Colon Proteome and Characterization of the Protein Pathways. *BioData Min.* 2012, 5, 11.

(28) Adamczyk, Z.; Kujda, M.; Nattich-Rak, M.; Ludwiczak, M.; Jagura-Burdzy, G.; Adamczyk, M. Revealing Properties of the KfrA Plasmid Protein via Combined DLS, AFM and Electrokinetic Measurements. *Colloids Surf., B* 2013, 103, 635−641.

(29) Stetefeld, J.; McKenna, S. A.; Patel, T. R. Dynamic Light Scattering: A Practical Guide and Applications in Biomedical Sciences. *Biophys. Rev.* 2016, 8, 409−427.

(30) Kelly, S.; Price, N. The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Curr. Protein Pept. Sci.* **2000**, *1*, 349−384.

(31) ud din Parray, M.; AlOmar, S. Y.; Alkhuriji, A.; Wani, F. A.; Parray, Z. A.; Patel, R. Refolding of Guanidinium Hydrochloride Denatured Bovine Serum Albumin Using Pyridinium Based Ionic Liquids as Artificial Chaperons. *Colloids Surf., A* **2021**, *610*, 125737.

(32) Kumari, M.; Maurya, J. K.; Tasleem, M.; Singh, P.; Patel, R. Probing HSA-Ionic Liquid Interactions by Spectroscopic and Molecular Docking Methods. *J. Photochem. Photobiol., B* **2014**, *138*, 27−35.

(33) Wei, Y.; Thyparambil, A. A.; Latour, R. A. Protein Helical Structure Determination Using CD Spectroscopy for Solutions with Strong Background Absorbance from 190 to 230 Nm. *Biochim. Biophys. Acta* **2014**, *1844*, 2331.

(34) Rygula, A.; Majzner, K.; Marzec, K. M.; Kaczor, A.; Pilarczyk, M.; Baranska, M. Raman Spectroscopy of Proteins: A Review. *J. Raman Spectrosc.* **2013**, *44*, 1061−1076.

(35) Kuhar, N.; Sil, S.; Umapathy, S. Potential of Raman Spectroscopic Techniques to Study Proteins. *Spectrochim. Acta, Part A* **2021**, *258*, 119712.

(36) Bradley, M. Curve Fitting in Raman and IR Spectroscopy : Basic Theory of Line Shapes and Applications. *Thermo Fish. Sci. Appl. Note* **2007**, 50733.

(37) Jiménez-Menéndez, N.; Fernández-Millán, P.; Rubio-Cosials, A.; Arnan, C.; Montoya, J.; Jacobs, H. T.; Bernadó, P.; Coll, M.; Usón, I.; Solà, M. Human Mitochondrial MTERF Wraps around DNA through a Left-Handed Superhelical Tandem Repeat. *Nat. Struct. Mol. Biol.* **2010**, *17*, 891−893.

(38) He, X.; Liu, Y.; Huang, S.; Liu, Y.; Pu, X.; Xu, T. Raman Spectroscopy Coupled with Principal Component Analysis to Quantitatively Analyze Four Crystallographic Phases of Explosive CL-20. *RSC Adv.* **2018**, *8*, 23348−23352.

(39) Ruggeri, F. S.; Šneideris, T.; Vendruscolo, M.; Knowles, T. P. J. Atomic Force Microscopy for Single Molecule Characterisation of Protein Aggregation. *Arch. Biochem. Biophys.* **2019**, *664*, 134−148.