



# Extracting accurate infrared lineshapes from weak vibrational probes at low concentrations<sup>☆</sup>



Raiza N.A. Maia, Sunayana Mitra, Carlos R. Baiz\*

Department of Chemistry, University of Texas at Austin, Austin, TX 78712-1224, USA

## ARTICLE INFO

### Method name:

Analysis of weak vibrational probes

### Keywords:

Spectroscopy  
FTIR  
Thiocyanate  
Data processing  
Protein structure  
Vibrational modes

## ABSTRACT

Fourier-transform infrared (FTIR) spectroscopy using vibrational probes is an ideal tool to detect changes in structure and local environments within biological molecules. However, challenges arise when dealing with weak infrared probes, such as thiocyanates, due to their inherent low signal strengths and overlap with solvent bands. In this protocol we demonstrate:

- A streamlined approach for the precise extraction of weak infrared absorption lineshapes from a strong solvent background.
- A protocol combining a spectral filter, background modeling, and subtraction.
- Our methodology successfully extracts the CN stretching mode peak from methyl thiocyanate at remarkably low concentrations (0.25 mM) in water, previously a challenge for FTIR spectroscopy.

This approach offers valuable insights and tools for more accurate FTIR measurements using weak vibrational probes. This enhanced precision can potentially enable new approaches to enhance our understanding of protein structure and dynamics in solution.

## Specifications table

Subject area:	Chemistry
More specific subject area:	Biochemistry
Name of your method:	Analysis of weak vibrational probes
Name and reference of original method:	Not applicable
Resource availability:	Not applicable

## Method details

### Introduction

Proteins undergo often structural changes as part of their function [1]. Understanding these conformational dynamics is essential for arriving at structural hypotheses of protein function. Measuring protein conformations and ensembles often relies on spectroscopic methods such as electron paramagnetic resonance (EPR), nuclear magnetic resonance (NMR), circular dichroism (CD), and Fourier-transform infrared (FTIR) spectroscopies [2–4]. FTIR spectroscopy reports on angstrom-level changes in the structure

<sup>☆</sup> Related research article: Not applicable.

\* Corresponding author.

E-mail address: [cbaiz@cm.utexas.edu](mailto:cbaiz@cm.utexas.edu) (C.R. Baiz).

<https://doi.org/10.1016/j.mex.2023.102309>

Received 1 May 2023; Accepted 31 July 2023

Available online 1 August 2023

2215-0161/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

and environment around biomolecules. In proteins, coupled vibrations play a major role in determining the local environment [5–7]. Motions of the amino acid functional units are sensitive to changes in the local environment [8–10].

Vibrational probes such as the  $C\equiv N$  or  $N\equiv N$  stretching mode, which includes nitriles thiocyanates (SCN), or azides, provide an appealing route to probing local dynamics because these vibrational modes produce narrow lineshapes that appear in a region of the spectrum that avoids direct overlap with the solvent OH stretching or H–O–H bending modes [11–16]. While these probes are ideal for measuring local dynamics, their incorporation into proteins can be demanding, particularly for large proteins [12,17–19]. Thiocyanates, on the other hand, can be efficiently incorporated into proteins by cyanylation of cysteine amino acids [20,21]. In the absence of cysteines in a native protein, a point mutation to a cysteine residue can be chemically transformed into SCN through site-specific cyanylation, thus providing a relatively streamlined route to introduce local probes at multiple selected sites in a protein. Nitriles, while certainly useful, produce weak spectroscopic signals as a result of their low oscillator strengths compared to for example carbonyls or other “strong” vibrations. In addition, for biomolecules concentration must often be kept low to prevent aggregation or crowding effects which could affect the protein structure of its environment. Together, the combination of weak oscillator strength and dilute concentrations makes it challenging to measure accurate and reproducible lineshapes using nitrile probes [20–22].

Several approaches have been developed to measure weak absorption lineshapes of nitriles in solution. These include the use of sensitive MCT detectors; bandpass filters; and long acquisition times to measure high-quality absorption spectra [6,23,24]. Whereas weak IR peaks are extracted from the background baseline correction methods, fitting functions, smoothing, and Fourier filtering [15,16]. Among the different peak fitting methods, the polynomial fitting function can be used to remove the broad solvent background and extract the narrow vibrational probe lineshapes. However, such an approach is subject to specific assumptions and is particularly unreliable when the initial signal amplitude is not distinct from the broad solvent background. This is particularly problematic for an SCN probe in  $H_2O$ , where the CN stretching mode overlaps with the broad  $H_2O$  combination band, [25] making the spectral measurement challenging. This work presents an effective method to measure the absorption spectra of weak infrared probes and extract accurate lineshapes from a broad water background. We expect that the specific procedure presented here will be useful to researchers working with weak vibrational probes and particularly in the context of biological samples.

## Procedure

### Sample preparation

Methyl thiocyanate (MeSCN, > 99.0% purity) was purchased from Tokyo Chemical Industry (T.C.I.). Ultrapure 18.2 m $\Omega$   $H_2O$  was dispensed from a Millipore milli-Q sigma water purification system. A 50 mM MeSCN in water stock solution was prepared from which a series of 25 mM, 10 mM, 5 mM, 1 mM, 0.5 mM, and 0.25 mM were prepared by serial dilution.

### Fourier transform infrared (FTIR) measurements

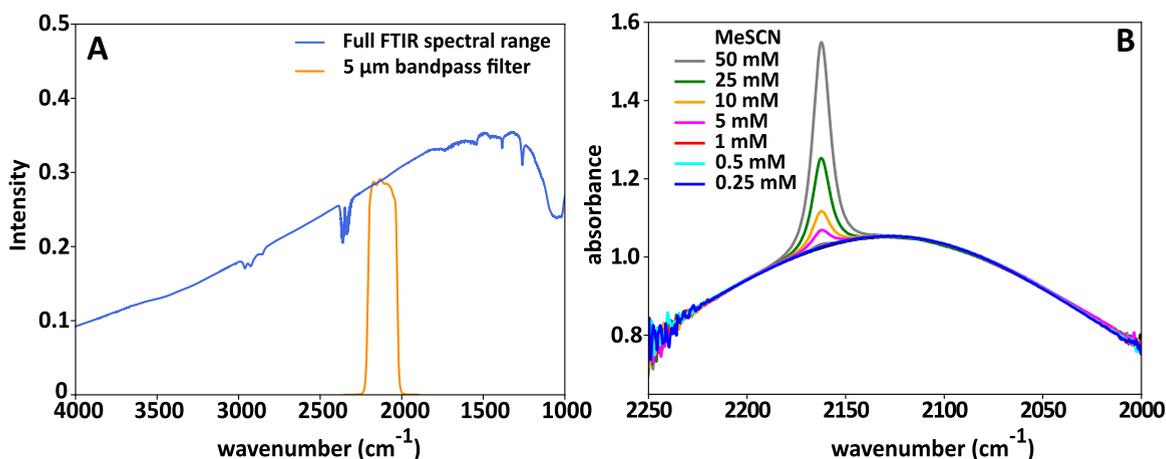
FTIR spectra were measured using a Bruker Invenio-S instrument equipped with a liquid-nitrogen-cooled mercury cadmium telluride (MCT) detector. The spectral resolution was set to 0.5  $cm^{-1}$  and 60 kHz scanner velocity. The sample enclosure was purged with dry air for 15 min before data acquisition. The sample was held between a pair of 1-inch diameter by 1-mm thick  $CaF_2$  windows, separated by a 50  $\mu m$  spacer. Measurements were performed at room temperature. All the MeSCN samples were prepared using 28  $\mu L$  of the solution. The analysis below focuses on polynomial fitting for background correction. There are, however, many additional fitting functions that can be used, including smoothing splines, oscillatory functions, Gaussian functions, Fourier series functions, and many more. We tested a selection and found that other functions have resulted in similar outputs or produced non-ideal fitting outcomes. Thus, in the following descriptions, only polynomial functions are discussed.

### Background subtraction and analysis

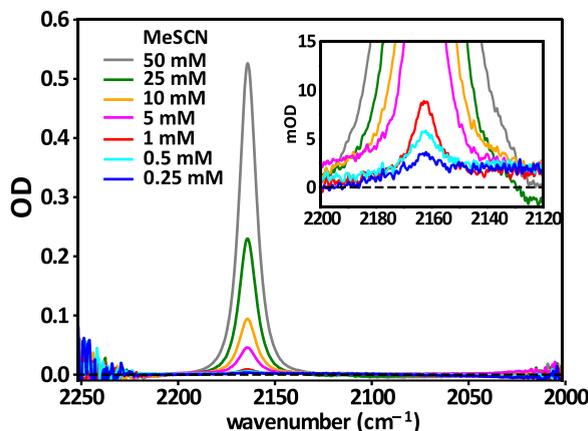
#### Bandpass filter

In FTIR spectroscopy all the wavelengths generated by the source simultaneously impinge on the detector which can limit the dynamic range of the measurement. In addition, FTIR spectroscopy in aqueous samples can bring few challenges to the data acquisition, since the IR spectrum will be dominated by the intense absorption of O–H stretching bands centered around 3450  $cm^{-1}$ , 3615  $cm^{-1}$  and H–O–H bend around 1640  $cm^{-1}$  [26]. Overcoming these issues and improving data quality in this region can be achieved using a bandpass filter, which serves to limit the wavelength range that impinges upon the detector, making better use of the detector's dynamic range. Essentially, the filter selectively blocks out unwanted frequencies while retaining only the desired range. This allows for an increase in the gain on the detector and enables the opening of the source aperture to enhance the amount of light captured. Therefore, when using the bandpass filter, the oscillations in the interferogram for the desired frequency range can exhibit a lower signal-to-noise ratio compared to not using the filter.

This step allows us a wider source aperture and consequently obtain higher infrared intensities at the detector for the wavelengths of interest, increasing sensitivity. Here we introduced an IR bandpass filter with a spectral window centered at 4.75  $\mu m$  and 50 nm of bandwidth (FB4750–500, Thorlabs Inc) as shown in Fig. 1A. Which in the conditions presented in this work, increased the dynamic range  $\sim 13\times$  by removing unwanted wavelength regions and thus reducing the zero-delay “spike” in the interferogram. The results described here are focus on the 2000–2400  $cm^{-1}$ , however the method describe in this paper can be applied to any frequency range. The data collected to different concentrations of MeSCN can be seen in Fig. 1B.



**Fig. 1.** Measurement of SCN peak for different MeSCN concentrations. (A) Raw detector intensity spectra shown with and without the implementation of the bandpass filter. In the orange spectrum wavelengths outside the bandpass filter window are blocked. (B) Spectra of MeSCN of concentrations 5, 10, 25, and 50 mM show thiocyanate peak sitting on top of the H<sub>2</sub>O band. The MeSCN CN stretching signal appears in ~2162 cm<sup>-1</sup> region of the spectrum. Within lower concentrations of MeSCN (0.25 mM, 0.5 mM and 1 mM) only the large background absorption of the water combination band mode is visible, and the SCN absorbance is too weak to observe without the background subtraction methods described here.



**Fig. 2.** Solvent subtraction leads to accurate lineshapes at low concentrations. Spectra following solvent subtraction from samples with different MeSCN concentrations. Inset shows zoomed in frequency region for better visualization. All samples show a clear SCN peak centered at 2162 cm<sup>-1</sup>.

#### Solvent background subtraction

The spectra shown in Fig. 2 are obtained standard difference ( $\Delta A$ ) between the sample ( $A_{\text{MeSCN}}$ ) and the solvent ( $A_{\text{H}_2\text{O}}$ ). Spectra are corrected using a linear baseline correction. Finally, the solvent band is subtracted from the individual samples to get the consistent lineshapes for the different MeSCN concentrations from 50 mM to 0.25 mM as indicated in the figure.

Following the above-described steps, we can successfully extract the SCN vibrational mode from the solvent background. The accuracy of the analysis is further established by observing the SCN peak position centered at 2162 cm<sup>-1</sup>. Here the lowest concentration (0.25 mM) exhibits a small SCN peak above the baseline within the current extraction method.

#### Polynomial background subtraction

To account for the background in the 2250–2080 cm<sup>-1</sup> range (Fig. 1) and to extract the SCN peak without the specific need to measure a H<sub>2</sub>O background, a polynomial regression can be used. Numerous methods have been introduced to perform baseline subtraction and background correction [27–29]. However, the issue if signal-to-noise adds complications and the performance of most of these methods for weak vibrational signals remain untested. Hence, in this study, we demonstrate a straightforward and effective method that can be applied for these cases involving weak vibrational modes. Other, more sophisticated approaches that could yield comparable baseline results [27–29]. In the samples with high MeSCN concentrations (50 mM, 25 mM, 10 mM, and 5 mM), the frequency range containing the SCN signal (2190–2134 cm<sup>-1</sup>) must be excluded from the polynomial fitting procedure. In the low MeSCN concentration samples (1 mM, 0.5 mM, and 0.25 mM), it was not necessarily excluded as the feature is sufficiently weak.

To determine the polynomial order that best models the background without overfitting, we looked at the quality of the curve and the adjusted R<sup>2</sup> of each polynomial model from the 2nd to 4th order (Fig. 3). For our data, all the polynomial orders from 2nd

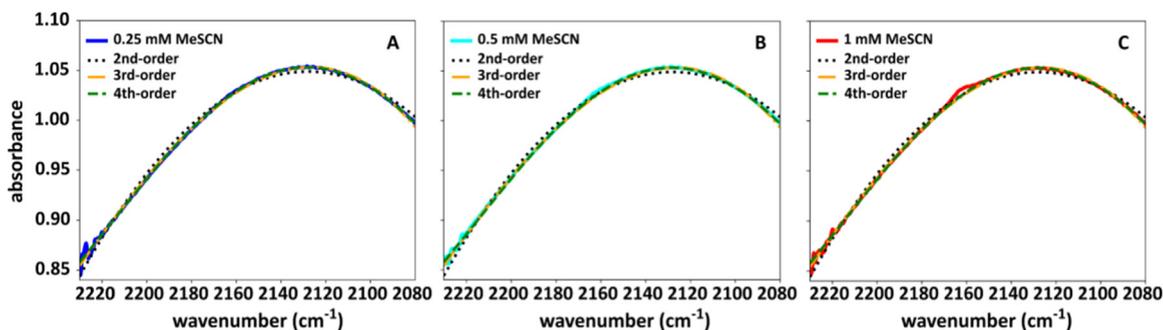


Fig. 3. Background modeling with 2th to 4th-order polynomials. The raw measured spectra for the three lower concentration of MeSCN are shown together with their polynomial fits as indicated in each figure. Here the 3rd order polynomial was chosen to correct the absorption background of H<sub>2</sub>O since it accurately models the baseline while avoiding overfitting.

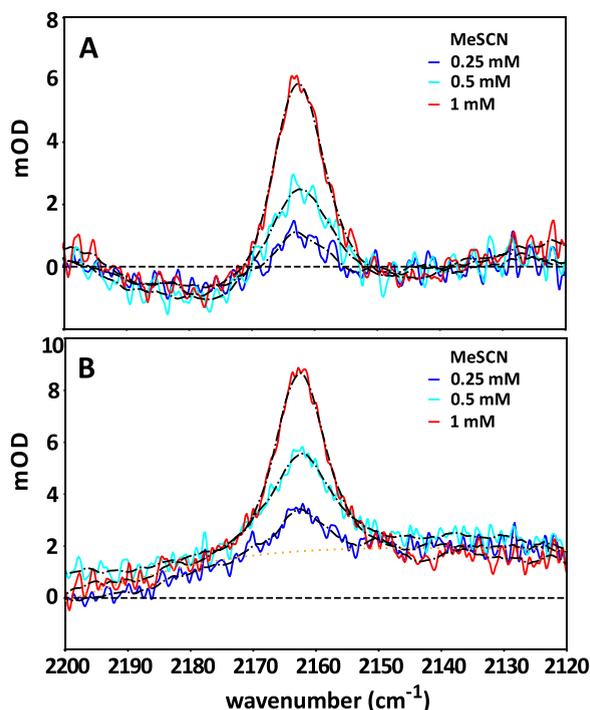


Fig. 4. Comparison of polynomial fit subtraction and baseline subtraction for SCN. (A) Spectra following 3rd order polynomial subtraction. The polynomial subtraction produces a relatively clean SCN peak with relatively clean baseline correction. (B) Spectra after the subtraction solvent subtraction. Similar to the spectra shown in Fig. 2, the SCN peak is centered at 2162 cm<sup>-1</sup> for all the three samples with lowest MeSCN concentration. For better visualization of the peaks a Savitzky–Golay smoothing filter was used (black dashlines) as described in the text.

to 4th order present the coefficient of variation is nearly identical  $R^2 \approx 0.99\%$ . Thus, to determine the best fit we decided to focus on the best-quality curve with lower polynomial order. Based on these parameters we choose to work with a 3rd order polynomial for all the data presented in Fig. 3.

Following the polynomial fit subtraction, only the SCN vibrational mode should remain above the baseline in the 2250–2080 cm<sup>-1</sup> range (Fig. 4A). To removing high-frequency noise for better visualization of the lineshapes, a smoothing function has been applied to the spectra (Fig. 4A and 4B, respectively). Here, we use the Savitzky–Golay filter to reduce noise at the expense of slightly reduced spectral resolution, highlighting the relative peak amplitudes and weights [30]. Alternative smoothing filters can be explored and implemented to achieve similar results [31–33].

Fig. 4 and Table 1 shows that the polynomial subtraction method provides a better baseline correction than the solvent-subtraction method described above. Specifically, in Fig. 4A one the SCN feature centered at 2162 cm<sup>-1</sup> is clearly visible even within the lowest (0.25 mM) MeSCN concentration. A small negative feature in the 2190–2185 cm<sup>-1</sup> range, with similar shape and intensity for all the three different MeSCN concentrations also appears in Fig. 4A, suggesting that this could be an artifact derived from polynomial analysis. However, these negative features do not affect the SCN peak lineshape and should not complicate interpretation of data.

**Table 1**

Maximum amplitude of  $2162\text{ cm}^{-1}$  for the three lowest concentrations. For the solvent background subtraction data, the amplitude baseline offset with maximum amplitude of  $\sim 2.3\text{ mOD}$  (Fig. 4B – orange line), was subtracted from the final amplitude as shown in the table below. Here we show the intensity measurements of the MeSCN signal and comparing it with the standard deviation of background noise. To accurately assess the background noise, we calculated its standard deviation within a specific spectral range ( $2140 - 2050\text{ cm}^{-1}$ ) where no MeSCN signal is present. By performing such comparisons, we aim to evaluate the signal-to-noise ratio (SNR) for our measurements, a critical metric in assessing the quality of the data.

	Polynomial background fit	Solvent background subtraction
0.25 mM	$1.2 \pm 0.4\text{ mOD}$	$1.2 \pm 0.9\text{ mOD}$
0.5 mM	$2.5 \pm 0.4\text{ mOD}$	$3.3 \pm 0.7\text{ mOD}$
1 mM	$5.9 \pm 0.5\text{ mOD}$	$6.4 \pm 0.8\text{ mOD}$

Within the solvent-subtracted spectra (Fig. 4B), following the smoothing filter, the presence of the SCN peak around  $2162\text{ cm}^{-1}$  in the 0.25 mM MeSCN is also clearly visible. Here we highlight that the main advantage of a polynomial-subtraction method is to make a separate solvent background measurement unnecessary, thereby reducing the number of samples required and the overall experimental data acquisition effort.

The combination of the spectral filter, MCT detection, and, most importantly, the "backgroundless" polynomial subtraction approach enables the measurement of extremely weak thiocyanate signals even in the presence of a strong water background. In other words, this approach can be used to measure absorbances as low as  $\sim 2\text{ mOD}$  on a water background a 500-fold higher absorbance in the same region (Fig. 1B). Based on this, the method could also be valuable for various experimental approaches and samples, such as surface-enhanced infrared absorption spectroscopy (SEIRAS), which in which monolayer sensitivity requires accurate measurements of lineshapes in the mOD or sub-mOD range in many cases [34,35].

## Conclusions

This brief report demonstrates improvement in acquisition of FTIR spectra for weak vibrational probes using a spectral filter combined with direct polynomial subtraction. Using this combination we were able to extract the absorption peak of a weak oscillator, MeSCN at low concentrations. This is important, especially for vibrational labels in biological molecules such as proteins, where it may be impractical to collect data at high concentrations due to solubility or aggregation. Given common sample constraints it is important to perform careful data analysis to extract accurate lineshapes. Thus, to accurately interpret a weak IR band shape, post-processing mathematical tools are necessary. Here, we showed that implementing the direct polynomial subtraction of the baseline approach can be a successful method to extract IR peaks in low-concentration samples, and avoid measuring a separate solvent background.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Ethics statements

Not applicable

## CRediT author statement

**R.M.:** Methodology, Software, Analysis, Writing. **S.M.:** Methodology. **C.B.:** Conceptualization, Resources. All authors approved the final version of the manuscript.

## Acknowledgments

This work was supported by the National Institute of Health funding (R35GM133359) and the Welch Foundation (F-1891). We thank Prof. Jens Bredenbeck (University of Frankfurt) for insightful discussions.

## References

- [1] L. Banci, *Metalloomics and the Cell*. *Metalloomics Cell* 2013.
- [2] M.C. Cavanagh, R.M. Young, B.J. Schwartz, The roles of the solute and solvent cavities in charge-transfer-to-solvent dynamics: ultrafast studies of potassium and sodium in diethyl ether, *J. Chem. Phys.* 129 (13) (2008) 134503.
- [3] C.K. Chan, Y. Hu, S. Takahashi, D.L. Rousseau, W.A. Eaton, J. Hofrichter, Submillisecond protein folding kinetics studied by ultrarapid mixing, *Proc. Natl. Acad. Sci. USA*. 94 (5) (1997) 1779–1784.
- [4] H.I. Okur, J. Kherb, P.S. Cremer, Cations bind only weakly to amides in aqueous solutions, *J. Am. Chem. Soc.* 135 (13) (2013) 5062–5067.
- [5] A. Barth, C. Zscherp, What vibrations tell us about proteins, *Q. Rev. Biophys.* 35 (4) (2002) 369–430.
- [6] A. Barth, Infrared spectroscopy of proteins, *Biochim. Biophys. Acta* 1767 (9) (2007) 1073–1101.
- [7] A. Barth, The infrared absorption of amino acid side chains, *Prog. Biophys. Mol. Biol.* 74 (3–5) (2000) 141–173.
- [8] S.C. Edington, S. Liu, C.R. Baiz, Infrared spectroscopy probes ion binding geometries, *Methods Enzymol.* 651 (2021) 157–191.
- [9] L.P. Deflores, Z. Ganim, R.A. Nicodemus, A. Tokmakoff, Amide I-II' 2D IR spectroscopy provides enhanced protein secondary structural sensitivity, *J. Am. Chem. Soc.* 131 (9) (2009) 3385–3391.
- [10] P. Stevenson, C. Gotz, C.R. Baiz, J. Akerboom, A. Tokmakoff, A. Vaziri, Visualizing KcsA conformational changes upon ion binding by infrared spectroscopy and atomistic modeling, *J. Phys. Chem. B* 119 (18) (2015) 5824–5831.
- [11] M.P. Wolfshorndl, R. Baskin, I. Dhawan, C.H. Londergan, Covalently bound azido groups are very specific water sensors, even in hydrogen-bonding environments, *J. Phys. Chem. B* 116 (3) (2012) 1172–1179.
- [12] S. Naganathan, S. Ray-Saha, M. Park, H. Tian, T.P. Sakmar, T. Huber, Multiplex detection of functional G protein-coupled receptors harboring site-specifically modified unnatural amino acids, *Biochemistry* 54 (3) (2015) 776–786.
- [13] M.J. Tucker, X.S. Gai, E.E. Fenlon, S.H. Brewer, R.M. Hochstrasser, 2D IR photon echo of azido-probes for biomolecular dynamics, *Phys. Chem. Chem. Phys.* 13 (6) (2011) 2237–2241.
- [14] B.A. Lindquist, K.E. Furse, S.A. Corcelli, Nitrile groups as vibrational probes of biomolecular structure and dynamics: an overview, *Phys. Chem. Chem. Phys.* 11 (37) (2009) 8119–8132.
- [15] J.M. Schmidt-Engler, L. Blankenburg, B. Blasiak, L. van Wilderen, M. Cho, J. Bredenbeck, Vibrational Lifetime of the SCN Protein Label in H(2)O and D(2)O Reports Site-Specific Solvation and Structure Changes During PYP's Photocycle, *Anal. Chem.* 92 (1) (2020) 1024–1032.
- [16] J.M. Schmidt-Engler, L. Blankenburg, R. Zangl, J. Hoffmann, N. Morgner, J. Bredenbeck, Local dynamics of the photo-switchable protein PYP in ground and signalling state probed by 2D-IR spectroscopy of -SCN labels, *Phys. Chem. Chem. Phys.* 22 (40) (2020) 22963–22972.
- [17] A.L. Le Sueur, R.E. Horness, M.C. Thielges, Applications of two-dimensional infrared spectroscopy, *Analyst* 140 (13) (2015) 4336–4349.
- [18] J. Zimmermann, K. Gundogdu, M.E. Cremeens, J.N. Bandaria, G.T. Hwang, M.C. Thielges, C.M. Cheatum, F.E. Romesberg, Efforts toward developing probes of protein dynamics: vibrational dephasing and relaxation of carbon-deuterium stretching modes in deuterated leucine, *J. Phys. Chem. B* 113 (23) (2009) 7991–7994.
- [19] M.C. Thielges, J.Y. Axup, D. Wong, H.S. Lee, J.K. Chung, P.G. Schultz, M.D. Fayer, Two-dimensional IR spectroscopy of protein dynamics using two vibrational labels: a site-specific genetically encoded unnatural amino acid and an active site ligand, *J. Phys. Chem. B* 115 (38) (2011) 11294–11304.
- [20] H.A. McMahon, K.N. Alfieri, K.A. Clark, C.H. Londergan, Cyanylated Cysteine: a Covalently Attached Vibrational Probe of Protein-Lipid Contacts, *J. Phys. Chem. Lett.* 1 (5) (2010) 850–855.
- [21] S.R. Dalton, M.S. Poretz, C.H. Londergan, S. Linse, Cyanylated Cysteine used to Characterize Binding Interfaces of Calcium Binding Proteins, *Biophys. J.* (2011) 100.
- [22] L.J. van Wilderen, D. Kern-Michler, H.M. Muller-Werkmeister, J. Bredenbeck, Vibrational dynamics and solvatochromism of the label SCN in various solvents and hemoglobin by time dependent IR and 2D-IR spectroscopy, *Phys. Chem. Chem. Phys.* 16 (36) (2014) 19643–19653.
- [23] M.G. Maienschein-Cline, C.H. Londergan, The CN stretching band of aliphatic thiocyanate is sensitive to solvent dynamics and specific solvation, *J. Phys. Chem. A* 111 (40) (2007) 10020–10025.
- [24] S.R.G. Naraharisetty, D.V. Kurochkin, I.V. Rubtsov, C–D Modes as structural reporters via dual-frequency 2DIR spectroscopy, *Chem. Phys. Lett.* 437 (4) (2007) 262–266.
- [25] P.K. Verma, A. Kundu, M.S. Poretz, C. Dhooonmoon, O.S. Chegwidden, C.H. Londergan, M. Cho, The Bend+Libration Combination Band Is an Intrinsic, Collective, and Strongly Solute-Dependent Reporter on the Hydrogen Bonding Network of Liquid Water, *J. Phys. Chem. B* 122 (9) (2018) 2587–2599.
- [26] Nakamoto, K. In *Infrared and Raman Spectra of Inorganic and Coordination Compounds*, 1978.
- [27] G. Schulze, A. Jirasek, M.M. Yu, A. Lim, R.F. Turner, M.W. Blades, Investigation of selected baseline removal techniques as candidates for automated implementation, *Appl. Spectrosc.* 59 (5) (2005) 545–574.
- [28] S. Guo, J. Popp, T. Bocklitz, Chemometric analysis in Raman spectroscopy from experimental design to machine learning-based modeling, *Nat. Protoc.* 16 (12) (2021) 5426–5459.
- [29] S. He, W. Zhang, L. Liu, Y. Huang, J. He, W. Xie, P. Wu, C. Du, Baseline correction for Raman spectra using an improved asymmetric least squares method, *Anal. Methods* 6 (12) (2014) 4402–4407.
- [30] W.H. Press, S.A. Teukolsky, Savitzky-Golay Smoothing Filters, *Computer in Physics* 4 (6) (1990) 669–672.
- [31] H.L. Weinert, Efficient computation for Whittaker-Henderson smoothing, *Comput. Stat. Data Anal.* 52 (2) (2007) 959–974.
- [32] M. Schmid, D. Rath, U. Diebold, Why and How Savitzky-Golay Filters Should Be Replaced, *ACS Measurement Science Au* 2 (2) (2022) 185–196.
- [33] P.H.C. Eilers, A Perfect Smoother, *Anal. Chem.* 75 (14) (2003) 3631–3636.
- [34] K. Ataka, T. Kottke, J. Heberle, Thinner, smaller, faster: IR techniques to probe the functionality of biological and biomimetic systems, *Angew. Chem. Int. Ed Engl.* 49 (32) (2010) 5416–5424.
- [35] O. Masatoshi, Dynamic Processes in Electrochemical Reactions Studied by Surface-Enhanced Infrared Absorption Spectroscopy (SEIRAS), *Bull. Chem. Soc. Jpn.* 70 (12) (1997) 2861–2880.