Compensation of Strong Water Absorption in Infrared Spectroscopy Reveals the Secondary Structure of Proteins in Dilute Solutions

Bonghwan Chon, Shuyu Xu, and Young Jong Lee\*

Biosystems and Biomaterials Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

**ABSTRACT**

Infrared (IR) absorption spectroscopy is a powerful tool that can quantify complex biomolecules and their structural conformations. However, conventional approaches to protein analysis in aqueous solutions have been significantly challenged because the strong IR absorption of water overwhelms the limited dynamic range of the detection system and thus allows only a very short path length and a limited concentration sensitivity. Here, we demonstrate an adaptive solvent absorption compensation (SAC) approach can improve the concentration sensitivity and extend the available path length by distinguishing the analyte signal over the full dynamic range at each wavelength. Absorption spectra without any post-processing show good linearity from 100 mg/mL to 0.1 mg/mL protein concentration, allowing a >100 times enhanced signal-to-noise ratio in the amide I band compared to the non-SAC results. We apply this method to *in-situ* investigate the isothermal kinetics of insulin fibrillation at two clinical concentrations at 74 °C for 18 hours. Simultaneous monitoring of both reactants (native forms) and products (fibrils) allows quantitative discussion of the detailed fibrillation mechanisms, which are not accessible with other single modality measurements. This simple optical technique can be applied to other absorption spectroscopies of analytes in strongly absorbing solvents, allowing for enhanced sensitivity without changing the detection system.

**INTRODUCTION**

Infrared (IR) absorption spectroscopy has been widely used to characterize the secondary structure of proteins and polypeptides.1–3 In particular, the amide I band, observed in the range of 1600 cm-1 and 1700 cm-1, is extensively studied because its absorption cross section is high compared to other functional groups and, more importantly, its absorption peak width and position are unique for different secondary structures, such as α-helices, β-sheets, and random coils.2 The distinct spectral features of the amide I band are so well investigated that IR spectroscopy can be used to measure semi-quantitatively the fractions of mixed secondary structures. In addition to the amide I band, analysis of broader IR absorption spectra of other molecular vibrations enables label-free chemical identification with the help of extensively accumulated spectral libraries. This non-emission-based approach can also provide absolute concentrations of subcomponents from the Beer–Lambert law. The SI (international system of units)-traceable measurement can improve data reproducibility and inter-laboratory comparability significantly for measurement assurance.

Despite these many advantages, the application of conventional IR spectroscopy, represented by Fourier-transform IR (FT-IR), is significantly limited when it is used to analyze proteins in aqueous solutions. The substantial light absorption by water reduces the transmitted signal intensity, dominates the dynamic range of a detection system, and overwhelms analyte absorption contributions. As a result, the optical path length is required to be short, typically 6 m or 8 m, and the analyte concentration needs to be high, typically >10 mg/mL for protein.3 The thin optical path length makes it extremely challenging to keep the same path length between the sample solution and the solvent reference measurements, which makes practical measurement reliability challenging. Also, an optical sample cell with a narrow space (< 10 m) is not compatible with typical microfluidic conditions because of pressure build-up,4 making the technique unsuitable for in-line monitoring or process analytical technology (PAT). As an alternative to avoid the water absorption band, the D2O exchange of the solvent has been used. However, the thermodynamics and reaction kinetics of protein are known to be different in H2O and D2O, which can make the interpretation difficult.3

To overcome these limitations caused by low-intensity lamp-based IR light sources, more intense light sources have been used. These include synchrotron radiation and the more accessible external-cavity quantum cascade laser (EC-QCL).5 EC-QCL sources have been extensively used for IR microscopy6,7 and IR spectroscopy.8,9 The EC-QCL provides fast-scanning monochromatic light with high intensity ranging from mid- to far-IR wavelengths.10 This intense monochromatic light allows for the optical path length to increase up to 38 m for IR spectrum measurements of protein in a non-deuterated aqueous solution.9 The increased path length and the high light intensity significantly lowers the detection limit to <1 mg/mL for the amide I band. However, the reported spectral range is limited to the amide I band between 1600 cm-1 and 1700 cm-1. Microfluidic modulation spectroscopy is a different approach to improve the signal-to-noise ratio by measuring the signal difference as the sample and the solvent fluids are alternated in a microfluidic transmission cell.11 Recently, Lendl *et al.* reported IR spectra of proteins at concentrations as low as 0.1 mg/mL over the spectral range to 1500 cm-1 and 1700 cm-1 to cover both the amide I and II bands by using dual-beam balanced detection.12 To avoid the saturation at the amide II region, they selectively lowered the laser intensity using a set of short-pass and long-pass color filters. However, the transmission spectrum of the color filters is not easily adjustable for various sample conditions or a broader spectral range. Also, employing multiple color filters can result in a significant reduction in light intensity arriving at the detector.

Here, we introduce a new approach to use the maximum dynamic range of the detection system of QCL-IR spectroscopy over a broader spectral range with a significantly enhanced signal-to-noise ratio (SNR). Our approach is based on a conceptually simple but optically advanced technique that pre-compensates the solvent absorption by adaptively adjusting the incident light intensity during wavelength scanning. Using this *adaptive* solvent absorption compensation (SAC) technique, we can use the full dynamic range of the detection system at each wavelength. This adaptive spectral adjustment technique can be used for various sample conditions with various path lengths and even various solvents. The available frequency range can be extended as wide as the laser and the detector allow. This new technique can be easily combined with the other QCL-based IR spectroscopy methods by simply inserting a spectrum adjusting unit into the incident beam path. Our presented SAC-IR spectra of protein solutions between 1350 cm-1 and 1770 cm-1 show good linearity over three orders of magnitude in concentration with a lower limit of 0.1 mg/mL, which is >100 times greater SNR than non-SAC methods for the wide wavelength range.

Using this high-sensitivity optical method, we study the *in-situ* isothermal kinetics of a protein fibrillation reaction, which is known to accompany dramatic changes in secondary structure.13 Protein fibrillation has been studied by various analytical methods. Electron microscopy or atomic force microscopy can observe individual fibrils and provide the actual shape and size distributions, but it cannot monitor the fibrillation occurring in the solution phase. Differential scanning calorimetry measures the transition temperatures and their corresponding heat exchange during fibrillation, but it does not provide any direct information on structural changes.14 Far-ultraviolet circular dichroism (far-UV-CD) is a spectroscopic method that has been used often for the identification of the secondary structure.15 However, it requires buffer exchange in case of the presence of optically active excipients in the original formulation. Also, due to its strong electronic transitions of the protein in the far-UV region, researchers often need to dilute a solution below 1 mg/mL for stable measurement. Both buffer exchange and solution dilution can significantly alter the overall kinetics of peptide aggregation reactions, which are quite often affected by peptide concentration.15 Alternatively, the thioflavin T (ThT) fluorescence method has been widely used to monitor *in-situ* the progress of amyloid fibrillation by taking advantage of the property of fluorescence quantum yield increase of ThT upon binding to rigid amyloid fibrils.16 The ThT method, which does not require buffer exchange or dilution, has been performed in numerous kinetics studies and *in-vivo* imaging and diagnostics. However, the ThT measurements only report the amount of the final structure that is in the amyloid form, so any structural information on the initial or intermediate species is not available. Also, the technique has rarely been used to quantify the absolute amount of amyloid fibrils because of variability in labeling efficiency depending on peptide solution conditions.

On the other hand, IR absorption spectroscopy can quantify the amounts of multiple mixed secondary structures of the native and amyloid species during protein fibrillation. However, as mentioned earlier, because of the limited sensitivity of peptides in aqueous solutions, the conventional FT-IR approach has not been widely used to study fibrillation mechanisms. Recently, QCL-IR spectroscopy has shown that secondary conformations of polypeptide and protein are affected by temperature, pH, and concentration.17,18 However, to our knowledge, no reports have been published demonstrating *in-situ* measurements of isothermal kinetics of amyloid fibrillation in an aqueous solution at elevated temperature. It is a significant experimental challenge to perform stable *in-situ* absorption measurements through a heated transmission cell for longer than several hours because even slight thermal expansion of the sample cell can be significant compared to the path length for conventional FT-IR. Although the longer path length used in QCL-IR spectroscopy ameliorates thermal expansion effects, temperature-associated fluctuations in the enclosure and possible formation of microbubbles make it challenging to acquire absorption spectra for several hours as is required for a low-concentration solution.

In this paper, we demonstrate that the SAC-IR spectroscopy can monitor *in-situ* isothermal kinetics of fibril formation in insulin solutions prepared at two clinically relevant concentrations of 3.5 mg/mL and 17.5 mg/mL.19 Insulin, a peptide hormone that regulates blood sugar levels, has been used to treat diabetes for several decades. Insulin amyloid formation is still one of the pharmaceutical challenges, including adverse health effects, reduced efficacy, shortened storage time, and limited drug delivery options.20 Using the observed *in-situ* IR results, we discuss kinetic models of isothermal insulin fibrillation using time-resolved amounts of native form and fibrils for two different insulin concentrations. In the Discussion section, we present the SNR analysis for various noise sources to understand how the SAC-IR spectroscopy can enhance the SNR and the sensitivity.

**EXPERIMENTAL DETAILS**

Bovine serum albumin (BSA, >96 %, Sigma-Aldrich) and Dulbecco’s phosphate-buffered saline (DPBS, Thermo-Fisher Scientific) were used as received. A stock solution of BSA in DPBS was prepared at a 100 mg/mL concentration. Different BSA solution concentrations were prepared by diluting the BSA stock solution into DPBS to the desired concentration. The NISTmAb reference material (RM 8671), at a protein concentration of 10 mg/mL in 12.5 mmol L-1 L-histidine and 12.5 mmol L-1 L-histidine HCl (pH = 6), was used as received. Stock solutions of BSA and NISTmAb in the same histidine buffer were prepared and mixed to create (BSA:NISTmAb) mass ratios from 0 % to 100 % in 20 % increments. The total protein concentration of each mixture was 2 mg/mL.

Human recombinant insulin powder (Sigma-Aldrich) was used as received. The insulin was dissolved in hydrochloric acid solution (pH = 2) at 3.5 mg/mL and 17.5 mg/mL concentration, which are equivalent to clinically available insulin solutions labeled as U-100 and U-500, respectively.

An EC-QCL (MIRcat, DRS Daylight Solutions) was operated at a repetition rate of 20 kHz with a pulse width of 1 s, corresponding to the duty cycle of 2 %. The QCL head was thermo-electrically cooled with additional cooling provided by an external chiller. The laser wavelength was scanned from 1420 cm-1 to 1750 cm-1 in 2 cm-1 steps. For the spectrum adjuster unit, two acousto-optic modulators (AOM, Brimrose Corp.) were used in tandem to increase the intensity range of the incident light. The diffracted beam of the order of *n* = 1 from the first AOM was diffracted again by the second AOM but at the diffraction order of *n* = –1. The opposite diffraction orders of the two AOMs compensate for the wavelength-dependent Bragg diffraction of a single AOM. The fast response time (> 1s) of AOMs allows the spectrum adjuster to respond to wavelength changes even at the rapid scanning speed of an EC-QCL. The diffraction efficiency of the AOMs was controlled by a computer via a 16-bit DAC multifunction I/O device (PCIe-6374, National Instruments).

The transmission spectrum of the QCL output at full laser power and maximum diffraction efficiency of the tandem AOMs was measured with a sample cell filled with water. The detection system parameters were optimized to maximize signal at wavenumbers producing minimum transmitted light intensity. The transmission efficiency, *i.e.,* the diffraction efficiency of the tandem AOMs, was programmed so that the light intensity reaching the detector was uniform over all wavelengths. The transmission efficiency was controlled by two AOM controllers connected to the DAC card in the operating computer. The light was attenuated by a ratio from 1:1 to 1:3200, where water absorbance is minimal. A single AOM calibration was sufficient for multiple measurements as long as measurements were performed with the sample cell and the solvent.

The spectrum-adjusted beam was split into two beams by a 50:50 beam splitter: one beam (signal) was directed through a microfluidic flow cell containing a protein solution or a solution blank, and the other beam (reference) was directed through a microfluidic flow cell containing only the solvent blank. The dynamically adjusted intensity of the incident light entering each liquid cell varied from 1 mW to 1 W depending on wavelength. Both liquid cells in the signal and reference paths were non-demountable flow cells (GS20572, Specac), which have a 26 m lead spacer between two 3 mm thick CaF2 windows. The beams transmitted through the sample cell and the reference cell were independently focused by 10 cm focal-length lenses on two thermoelectrically cooled mercury cadmium telluride (MCT) detectors (MCT-7-TE4, Infrared Associates), which have an element with an area of 1 mm × 1 mm. For the dual-beam detection mode, two identically configured boxcar integrators (SR200, Stanford Research Systems) were used to detect the signals acquired from the two MCT detectors simultaneously. The data acquisition time was 500 ms at each wavenumber step, which corresponds to the total acquisition time of 3.3 min for a spectrum over 400 cm-1 with a 2 cm-1 increment. For the isothermal measurement of the insulin solutions, two identically configured lock-in amplifiers (SR830, Stanford Research Systems) were used in place of the boxcar integrators. A 16-bit multifunction I/O device (PCIe 6374, National Instruments) simultaneously transferred the signal from the boxcars or the lock-in amplifiers to the computer. The noise characteristics of the boxcar and the lock-in systems were indistinguishable in the current acquisition configuration.

The entire system was placed in an enclosure and constantly purged with dry air to reduce the humidity in the box. Water circulated from a temperature-controlled through a heating jacket surrounding the sample cell controlled the temperature. For measurements of BSA and NISTmAb solutions, the bath temperature was set to 20 ºC. For the isothermal measurements of insulin solutions, the bath temperature was set to 80 ºC, with a fluctuation of < 0.2 ºC. This resulted in sample cell temperatures of 22 ºC and 74 ºC, respectively, for the two protein solutions.

**RESULTS**

**Solvent absorption compensation (SAC) approach to IR spectroscopy**

Figure 1a shows the experimental scheme of the SAC-IR spectroscopy system implementing a conventional dual-beam absorption spectroscopy setup. An adaptive SAC unit, consisting of two tandem acousto-optic modulators, is inserted into the incident beam path to dynamically adjust the fraction of the light intensity passed while the wavelength of the EC-QCL is scanned. The rationale for the SAC and its impacts on an absorption spectrum are presented in Figures 1b–1g. The absorption spectra are acquired with and without applying the SAC for an identical aqueous solution of bovine serum albumin (BSA) at a concentration of 10 mg/mL. In conventional absorption spectroscopy, the spectrum of the light transmitted through a sample cell filled with water (Figure 1c) shows >3000 times intensity reduction at 1644 cm-1, where the water bending mode is located. In Figure 1d, the amide II band at 1550 cm-1 shows a recognizable absorption peak, but the amide I band near the 1644 cm-1 region is dominated by the noise, which shows the strong effect of light absorption by water. This type of dominating noise for a weak signal is considered to be unavoidable in a detection system that covers both strong and weak signals due to the limited dynamic range.

In contrast, the SAC method pre-compensates the light absorption by water by adjusting the transmittance adaptively at each frequency, so the light intensity transmitted through the solvent is nearly constant over the scanned frequency range. The equalized light intensity arriving at the detector makes it possible for a detection system to resolve the absorption signal with the maximum dynamic range. Figure 1e shows the pre-compensated spectrum of the incident light before the sample cell, while Figure 1f shows the spectrum of the transmitted light through the sample cell filled with water. The resulting absorption spectrum in Figure 1g shows that the SAC approach improves the SNR not only in the amide I region but also over the entire spectral range scanned, compared to the data collected using the conventional non-SAC method in Figure 1d. By using the fast responding (< ms) AOMs, this SAC method can be applied to fast scanning QCL-IR spectroscopy methods demonstrated in other reports.9,12 The AOMs unit can be replaced with other transmission controlling devices,21 depending on the response time and the dynamic range of intensity attenuation.

Diagram

Description automatically generated

**Figure 1**. (a) A dual-beam absorption measurement setup for the adaptive solvent absorption compensation (SAC)-IR spectroscopy. EC-QCL denotes external-cavity quantum cascade laser; AOM, acousto-optic modulator; BS, beam splitter; and TE-MCT, thermo-electric cooled mercury cadmium telluride detector. In a conventional approach, (b) the spectrum of incident light before the sample cell in the signal beam; (c) the spectra of light transmitted through the sample cell filled with PBS buffer and a BSA solution of 10 mg/mL in PBS; and (d) the absorbance spectrum of the BSA solution. In the SAC approach, (e) the spectrum of the adjusted incident light before the sample; (f) the spectra of light transmitted through the sample cell filled with the same solvent and solution as (c); and (g) the absorbance spectrum of the BSA solution from (f). The laser and the detection configuration are optimized to (f), and a neutral density filter (optical density = 3) is used for the measurement of (b), (c), and (e). The path length is 26 mm for both the sample cell and the reference cell. The acquisition time is 3.3 min for each spectrum over 400 cm-1 with a 2 cm-1 increment.

**Improved sensitivity for protein in aqueous solutions**

Figures 2a and 2b show SAC-IR absorption spectra of BSA solutions at ten different concentrations, spanning three orders of magnitude from 100 mg/mL to 0.1 mg/mL. The presented absorption spectra are directly calculated from the measured transmitted spectra of the PBS buffer and the BSA solutions *without any post-processing methods*, such as baseline-detrending, filtering, and smoothing, which are often implemented for both FT-IR measurements3 and QCL-IR measurements.12,17 Figure 5c shows the molar absorption coefficient (**) spectra, which is calculated by the Beer–Lambert law, ** = *A*/*cL*, where *A*, *c*, and *L* are absorbance, concentration, and path length, respectively. The overlapping ** of different solutions confirms the absorbance linearity in the wide concentration range. In Figure 2d, the absorbance values measured at two different frequencies show good linearity to concentration in both linear–linear and log–log plots. The observed linearity demonstrates the robustness of this SAC-IR spectroscopy for the wide concentration range (≈ three orders of magnitude) and the wide frequency range (> 330 cm-1) without any post-processing.

Chart, histogram

Description automatically generated

**Figure 2.** (a, b) Absorption spectra of BSA solutions at various concentrations acquired by the SAC-IR spectroscopy without any post data processing. (c) Molar absorption coefficient spectra at three representative BSA concentrations in (a, b). (d) Log–log plots of absorbance versus BSA concentration at the single frequencies corresponding to the amide I (1656 cm-1) and amide II peaks (1548 cm-1). The identical data are displayed as linear–linear plots in the inset. The straight lines are linear fits of the scatter data with the Beer–Lambert law. The gray arrows indicate the detection ranges of the protein amide I band of the conventional FT-IR method and the new SAC-IR method.

The error bars indicated at each concentration correspond to the standard deviation calculated at single frequencies from three or four consecutive measurements. At all concentrations, the standard deviations at both frequencies are ≈ 0.3 mOD, as displayed in Figure 1d. Thus, the SNR values of both bands turn out to be ≈ 2 for a solution of 0.1 mg/mL. We also calculate SNR using the area of the bands. Averaging absorbance over 30 cm-1, which corresponds to the full-width-half-maximum (FWHM) of the amide I and II bands, produces SNR of 6 and 5 for the amide I and II bands, respectively, from the 0.1 mg/mL BSA solution, which meets the criterion for the limit of detection (SNR ≥ 3). The lowest concentration demonstrated by the SAC approach is considered as 100 times lower than the practical detection limit of conventional FT-IR measurements. However, it is difficult to make a fair comparison of the detection limit between the two spectroscopy approaches because of inherent differences in the observed frequency range, the signal collection/process, and the nature of associated noise sources. Even among the QCL-based laser scanning methods, a fair comparison in the detection limit can be made only after the equivalent frequency range with equivalent post data processing. We compare results of SAC and non-SAC measurements using the identical system. While the amide I band observed by the non-SAC configuration shows SNR < 1 in Figure 1d, the amide I band observed by the SAC method shows SNR > 200 in Figure 1g. This significant enhancement in SNR can be explained by taking advantage of the maximum dynamic range at all wavelengths in the SAC approach, which will be discussed further later in the Discussion section.

Chart

Description automatically generated

**Figure 3.** (a) IR spectra for mixture solutions of BSA and NISTmAb at a total concentration of 2 mg/mL. (b) Classical least squares (CLS) decomposition of a binary protein mixture (0.4 mg/mL BSA and 1.6 mg/mL NISTmAb) using two absorption spectra of the pure BSA and NISTmAb solutions. The bottom plot is the residual of the mixture spectra after the CLS decomposition. (c) CLS analysis results of the mixture solutions. The closed circles denote the predicted concentration of NISTmAb, and the open squares are the predicted concentration of BSA. The error bars indicate the standard deviation of three consecutive measurements for each analyte solution. The dotted straight lines are for the guidance of coincidence of the actual and predicted concentrations.

**Quantitation of secondary structures** **in protein mixtures**

In addition to the detection limit, quantitation of secondary structure is another important application of mid-IR spectroscopy for protein solutions. Secondary structure band fitting relies heavily on a flat baseline, which is difficult in FT-IR even after post data processing. Using a mixture of two proteins that have different compositions of -helical and -sheet, we demonstrated this technique to determine the concentration of each component quantitatively without any post data processing. BSA is known to be rich in -helical structure,22 while immunoglobulin G (IgG) is known to be -sheet rich.23 For IgG, we use NISTmAb (NIST RM 8671), an IgG reference material of IgG. Stock solutions of BSA and NISTmAb are prepared at 2 mg/mL and mixed by the volume ratios of 0.2, 0.4, 0.6, and 0.8. Thus, the total concentration is the same (2 mg/mL) for the pure stock solutions and the mixture solutions. Notably, the total protein concentration of 2 mg/mL is below the detection range of conventional FT-IR. The classical least squares (CLS) method is used to decompose the acquired absorption spectra of the mixture into a linear combination of two spectra of single-component solutions,

(1)

where *A*BSA(**) and *A*NISTmAb(**) are the absorption spectra of the pure BSA and NISTmAb solutions, respectively; ** is frequency; and *f*BSA and *f*NISTmAb are the predicted mass fractions of BSA and NISTmAb, respectively. Figure 3b shows an example of spectral decomposition for a mixture spectrum using the two basis functions of *A*BSA(**) and *A*NISTmAb(**). The residual spectrum at the bottom of Figure 3b shows that the decomposition works very well with an average deviation of 0.3 mOD over the frequency range. From the CLS analysis of the mixed protein spectra, the predicted concentrations of BSA and NISTmAb are plotted as a function of their actual concentrations. In Figure 3c, the actual concentrations and the predicted concentrations based on the CLS analysis exhibit decent linearity for both BSA (rich in -helix) and NISTmAb (rich in -sheet), suggesting that this advanced IR absorption measurement method can be used for quantitative analysis of mixed or pure protein solutions that contain different secondary structures even at a low concentration.

***In-situ* SAC-IR measurement of isothermal insulin fibrillation at elevated temperature**

We apply the SAC-IR spectroscopy for *in-situ* monitoring of both reactants and products of an insulin fibrillation reaction at two clinically relevant concentrations without seeds at a high temperature. The elevated temperature (74 ºC) makes a set of measurements of a fibrillation reaction completed within a day. The most clinically relevant concentration of insulin is 3.5 mg/mL (0.61 mmol/L, labeled as U-100); however, for some specific patient conditions, a higher concentration solution of 17.5 mg/mL (3.05 mmol/L, labeled as U-500) is used. Figures 4a and 4d show the *in-situ* time-resolved absorption spectra of the two insulin solutions at 17.5 mg/mL and 3.5 mg/mL, respectively, acquired at 74 ºC. In both sets of spectra, a peak centered at 1655 cm-1 decreases while a new peak at 1624 cm-1 increases with time. The 1655 cm-1 peak corresponds to -helix, which is known to be the dominant secondary structure of the native insulin.3,24 The rising peak at 1624 cm-1 can be assigned to the parallel or anti-parallel -sheets found in fibrillar aggregates.3,24

Before discussing the results, it must be noted that, unlike the earlier non-post-processed spectra, this series of time-resolved spectra in Figures 4a and 4b are baseline-detrended by a line of two absorbance values at 1480 cm-1 and 1750 cm-1. This first-order detrending compensates for slow absorbance drift due to long-term fluctuations caused by the water-heated sample cell in the sealed enclosure during the long acquisition time of 18 h. For comparison, with the sample cell unheated, the long-term absorbance fluctuation is smaller than 1 mOD for 4 h, not requiring any post-processing. First, we perform qualitative analyses of the series of time-resolved absorption spectra using principal component analysis (PCA)25 and two-dimensional correlation spectroscopy (2DCOS).26 These two methods can determine how many independent spectral components can be identified in the time-resolved spectra. A PCA result of the 17.5 mg/mL solution spectra shows that one principal component is as large as 90 % of the whole variance [see Supporting Information], suggesting that an observed spectrum can be described as a linear combination of two spectral components that are correlated in the opposite direction. A similar result is observed from a 2DCOS analysis of the identical data [see Supporting Information]. While the synchronous plot shows the two underlying peaks at 1624 cm-1 and 1655 cm-1 are counter-correlated, the asynchronous plot shows no other spectral species. These results of PCA and 2DCOS qualitatively indicate that the observed IR spectra can be explained by two dominant spectral species, which correspond to the -helix of the native insulin and the -sheet of the insulin fibril.

For quantitative analysis of the amount of native (-helix) and fibril (-sheet) forms of insulin, we measure the area of underlying peaks in the amide I band. As shown in Figures 4c and 4d, two Gaussian functions yield good fits even with center frequencies and bandwidths shared for all spectra. [Additional fitting results can be found in Supporting Information.] As discussed earlier, the Gaussian peak centered at 1655 cm-1 corresponds to the -helix conformation and the other peak at 1624 cm-1, to the -sheet conformation.3 The peak at 1655 cm-1 may include non--helix components. However, one can assume that *change* in the area of the 1655 cm-1 peak is proportional to *change* in the mass fraction of -helix conformation. Likewise, *change* in the area of the 1624 cm-1 peak is proportional to *change* in the mass fraction of fibrillated -sheet conformation. Figures 4c and 4d show the area plots as a function of time calculated from the absorption spectra of the 17.5 mg/mL and 3.5 mg/mL solutions, respectively. In both data sets, the -helix peak area decays with time monotonically while the -sheet peak area rises. One may notice that the -helix peak area in Figure 4e reaches a non-zero steady-state value after *t* = 8 h, which can be explained by the existence of non-aggregated native forms or -helices in fibrils. A study of single insulin fibrils by nano-FT-IR spectroscopy reported that insulin fibrils contain -helices in the non--stacked region.24 The apparent decay and rise times are faster for the 17.5 mg/mL insulin solution than for the 3.5 mg/mL. The faster fibrillation at a higher concentration is consistent with previous reports on the aggregation kinetics of other peptide solutions.27

Another concentration effect can be found in the ratio of the -sheet area to the -helix area at steady-state. The -sheet peak area may not be directly proportional to the mass ratio of -sheet. Similarly, the -helix peak area may not be proportional to the mass ratio of unreacted monomers to produced fibrils. However, the ratio of the -sheet area to the -helix area should be the same if the steady-state equilibrium is identical for both concentrations. The observed ratio of the -sheet area to the -helix area is smaller at the high concentration (Figure 4e) than at the low concentration (Figure 4f) even after the unfinished reaction in Figure 4f. This may be simply due to more precipitation in the higher concentration solution, but the -sheet area does not show any continuing decay after it reaches a plateau. Also, no visible precipitate is observed in the sample cell after each experiment. Other possible explanations are either that the fibril mass fraction is lower for the higher concentration solution or that the structure distribution of fibrils is different, neither of which can be confirmed by this IR measurement alone. Among very few reports on the concentration effects on the conformational equilibrium of peptide fibrillation, one indirectly related report shows that -sheet intensity of peptide solutions measured by far-UV CD exhibits a non-linear dependence on the initial monomer concentration.13

Chart, line chart, histogram

Description automatically generated

**Figure 4.** Time-resolved absorption spectra of insulin solutions at (a) 17.5 mg/mL and (b) 3.5 mg/mL concentrations, at 74 °C and pH = 2. Each spectrum is baseline detrended by absorbances at 1480 cm-1 and 1750 cm-1. (c) and (d) Examples of peak fitting of the amide I band with two Gaussian functions. The center frequency and the FWHM of the peak corresponding to the -helix are 1655 cm-1 and 25 cm-1, respectively, while those of the peak corresponding to the -sheet are 1624 cm-1 and 16 cm-1, respectively. All time-resolved spectra are fitted with the same center frequencies and bandwidths. (e) and (f) The areas of the two underlying peaks are plotted as a function of time. The peak areas are globally fitted for the high and low concentrations using two representative kinetics models. The solid lines, from the Finke–Watzky (F–W) model (Eqs. 5 and 6), are the best fit with *k*1 = 6.9 × 10-2 h-1 and *k*2 = 1.3 × 106 L mol-1h-1. The dashed lines, from the nucleation–elongation (N–E) model (Eqs. 7 and 8), are the best fit with (*k*n *k*+) = 1.6 × 105 L mol-1h-2 and *n*c = 1.

**Kinetic models for insulin fibrillation**

Protein fibrillation is generally considered a complex multi-step process involving multiple intermediate states in multiple pathways. Despite the complexity, most of the proposed reaction mechanisms are based on the two parallel paths: nucleation and growth.27,28

: nucleation (2)

: growth (3)

where *A*, *I*, and *B* are the reactant (the native form), the intermediate (the partially unfolded form), and the product (the fibril form), respectively; and *k*0, *k*1, and *k*2 are the rate constants of the unfolding, the nucleation, and the growth reactions, respectively. Unlike the ThT method based on a single metric for the mass of fibrils, the IR spectra presented in Figure 4 show the time profiles of the initial form (-helix), the final form (-sheet), and possibly the intermediate forms. As we discussed in the previous section, the overall spectral changes are dominated by two constrained spectral components of -helix and -sheet. That indicates either the mass fraction of intermediate states is negligible, or the spectrum of *I* is indistinguishable from that of *A* or *B*. If *k*0 << *k*1, the former case will be true. The latter case can also be explained by a partially unfolded intermediate state that cannot be distinguished by IR techniques. In either case, we can simplify the nucleation reaction as

: nucleation (4)

One simple analytical solution of the two-pathway coupled reactions is provided by the Finke–Watzky (F–W) model,28 where the growth reaction can occur from any size nucleus and at any site on the fibrils. The analytical solutions of the mass fractions of *A* and *B* are respectively

(5)

(6)

where [*A*]u and [*B*]u are added to the original form of the F–W solution for consideration of the contributions of unchanged -helices and -sheets, respectively. The solid lines in Figures 4c and 4d are the best fits using Eqs. 7 and 8.

Another kinetic equation can be found from the nucleation–elongation (N–E) model, which considers homogeneous nucleation followed by linear polymerization.27,29 In the absence of fragmentation or other secondary nucleation pathways, the mass fraction of filaments can be expressed in a simple form as

(7)

(8)

where with *k*n the nucleation rate constant, *k*+ the elongation rate constant, and *n*c the critical nucleus size. The dashed lines in Figures 4c and 4d are the best fits using Eqs. 8 and 9 from the N–E model. The best fit values are (*k*n *k*+) = 1.6 × 105 Lmol-1 h-2 and *n*c = 1, where the integer value of *n*c is determined by evaluating the fitting results of (*k*n *k*+) with each integer value for *n*c.

From only the results of two concentrations, it is difficult to know which kinetic model explains insulin fibrillation better or whether a more extended form of a kinetic model is required. However, this *in-situ* kinetic analysis based on two time-resolved spectral components provides extensive information on the reaction kinetics and the molecular conformations during the reaction.

**DISCUSSION**

**Signal-to-noise ratio analyses of the non-SAC and the SAC approaches**

Here, we discuss why the equalized transmission spectrum from the SAC approach can enhance the SNR of the resulting absorption spectrum compared to the results from the conventional non-SAC approach. For quantitative discussion, we consider the transmission spectra of a protein solution with variable path lengths, which allows us to find not only the optimum path length but also the SNR for variable transmission intensity. For a spectroscopic measurement, the following three noise sources are typically discussed in terms of the scaling with transmitted light intensity, *I*.30,31

Laser noise (LN): (9)

Shot noise (SN): (10)

Detector noise (DN): (11)

The error propagation function of transmission intensity noise (*d*) to absorbance noise (*sA*) can be expressed as

(12)

For the sake of discussion, we assume that the incident light intensity is constant at *I*0 over the scanning frequency range. We calculate *A* and (\*\* denotes LN, SN, or DN) at the amide I frequency (1656 cm-1) as a function of path length, as shown in Figures 5b and 5d, for the non-SAC and the SAC approaches, respectively. We notice that, for all three noise-limiting cases tested, the curves for each noise source are similar between the non-SAC and the SAC measurements, which makes any of the three noise sources unable to explain the observed significant enhancement in SNR by the SAC approach.

**Chart, map

Description automatically generated**

**Figure 5**. Comparison of noise and SNR analyses between the non-SAC and the SAC approaches using various noise sources. (a, b) Transmission spectra of water (the dashed lines) and a 10 mg/mL BSA solution (the solid lines) calculated for the non-SAC and the SAC approaches. The path length is 26 mm. The incident light intensity (unmodified) is set to be 107, relative to the detector noise. *I*max is determined by the transmitted light intensity at 1780 cm-1. (c) Plots of the absorbance noises () at the amide I frequency at 1656 cm-1 as a function of path length for the non-SAC configuration. The superscript in denotes the noise sources: DRN denotes dynamic range noise; SN, shot noise; LN, laser noise; and DN, detector noise. The absorbance signal of BSA (the black dashed line) is plotted for reference. (d) Plots of the absorbance noises at the amide I frequency for the SAC configuration. (e, f) The SNR plots at the frequencies of 1656 cm-1 (amide I) and 1548 cm-1 (amide II) are compared between the non-SAC and the SAC results. For the noise calculations above, we set the noise parameters as DR = 5000 for *d*DRN; *a* = 2 × 10-4 for *d*LN; *b* = 0.01 for *d*SN; and *g* = 1 for *d*DN.

Dynamic range (DR) is the ratio of the largest signal to the smallest detectable signal. The lower limit of the smallest detectable signal is determined by the digital resolution of the analog-to-digital converter, but typically the smallest detectable signal corresponds to the noise-equivalent signal of the detection system.30 Then, the dynamic range noise can be expressed as,

Dynamic range noise (DRN): (13)

which yields with Eq. 12 for various path lengths. It is noted that and are determined by *I*max, which is significantly affected by the frequency scanning range. In the non-SAC configuration (Figure 5a), *I*max is determined by the weakest absorbance of the solvent in the frequency range (at 1780 cm-1). In contrast, in the SAC configuration (Figure 5c), *I*max corresponds to the transmission intensity at the frequency of the strongest solvent absorption (at 1644 cm-1). The non-SAC result in Figure 5b shows a rapid decrease in with path length while the SAC result in Figure 5d shows that is nearly insensitive to path length. The difference in between the non-SAC and the SAC results becomes greater than three orders of magnitude as path length reaches 26 mm, the path length used for the experiment. This large difference in is consistent with the orders of magnitude enhancement in SNR by the SAC approach. It also strongly suggests that is the dominant noise source in the tested absorption measurement condition.

Based on the four noise sources, we calculate the total absorbance noise () and the corresponding SNR as a function of path length at different frequencies.

(14)

In Figure 5e, the SNR plots are calculated at the amide I frequency (1656 cm-1) and compared between the SAC and the non-SAC approaches. The SNR from the non-SAC approach reaches the maximum at path length around 6 m and decreases rapidly, which is consistent with the path length commonly used for FT-IR measurements. In contrast, the SNR from the SAC approach keeps increasing until path length reaches ≈ 30 m. If the maximum SNR of the SAC at 30 m is compared to the maximum SNR of the non-SAC at 6 m, the SNR of the SAC is still > 10 times greater than that of the non-SAC. If compared at the SAC-optimum path length at 30 m, the SNR of the SAC approach becomes >1000 times greater than that of the non-SAC approach. Although the actual values of the optimum path length and the maximum SNR enhancement can vary depending on the actual noises of the experimental conditions, this significant enhancement of SNR is consistent with the demonstrated experiment results in Figure 1.

Similarly, we analyze the errors and the SNR at the amide II frequency (1548 cm-1) both in the SAC and non-SAC approaches. Detailed error plots of the four noise sources and the resulting SNR results are presented in the Supporting Information. In the non-SAC approach, the contribution of the dominant noise source, , is significantly reduced at the amide II frequency because the transmittance of ≈ 10 % at the amide II frequency is much larger compared to ≈ 0.003% at the amide I frequency. As a result, the SNR curve shows a much greater value at the amide II frequency than at the amide I frequency, as shown in Figures 5e and 5f. In contrast, in the SAC approach, all the noise characteristics are almost identical between the amide I and II frequencies because the transmitted light intensity is equalized at all wavelengths. The only difference is that the SNR of the amide II is smaller by half than that of the amide I, due to the smaller absorbance of the amide II band. A comparison between the SAC and the non-SAC results of the amide II band in Figure 5f shows that the SNR is higher in the SAC approach than in the non-SAC approach until the dynamic range noise becomes smaller than the detector noise at a long path length. This analysis result is also consistent with the observed incremental enhancement in SNR for the amide II band in Figure 1.

We need to note that the error analysis of FT-IR is quite different from this scanning-wavelength approach. The FT method is based on time-dependent signals consisting of the entire wavelengths at the same time. The dynamic range of a detector cannot be directly transferrable to the dynamic range of a specific frequency intensity. This all-frequency signal is more likely dominated by a strong absorption band of a solvent, resulting in an effectively narrow dynamic range to resolve relatively weak analyte signals. Also, we note that our SAC approach equalizes the light intensity constant over the entire frequency range so that the signal remains within the *linearity dynamic range* of the detector. The limited linearity of absorbance to concentration has been a problem in quantitative analysis of the protein concentration in an aqueous solution. This new SAC method allows for quantitative IR spectra to be acquired over a wide wavelength range for a wide linear concentration range of protein in water.

**CONCLUSION**

We have demonstrated an adaptive SAC technique that can enhance the signal-to-noise ratio of QCL-IR absorption spectroscopy for the characterization of proteins in aqueous solutions by mitigating the high background absorption by water near the amide I band. Our advanced QCL-IR spectroscopy has good linearity of absorbance for the amide I and II bands of protein solutions from 100 mg/mL to 0.1 mg/mL, which corresponds to a higher concentration sensitivity than non-SAC configurations. We also have demonstrated that multivariate analysis of the broad spectra presents good predictability of the mass fractions of BSA and NISTmAb in mixture solutions. We apply this non-invasive, high-sensitivity method to monitor *in-situ* the secondary structure change during insulin fibrillation at elevated temperature. Results of the time-resolved IR spectral measurements at two representative clinical concentrations have been used to describe and analyze the isothermal kinetics using simple kinetic models for protein fibrillation. These SAC-IR spectroscopy results show a simple optical technique that enhances the sensitivity by two orders of magnitude without changing the detection system. The same SAC approach can be applied to other types of absorption spectroscopy that investigate samples in strongly absorbing solvents to improve the sensitivity and the detection limit of the methods. The longer path length allows for broader applications, including microfluidic sample introduction, which facilitates other applications in process analytical technology for biopharmaceutical production control.

ACKNOWLEDGMENTS

We thank the nSoft Consortium and the NIST Biomanufacturing Initiative for financial support. We thank Travis Gallagher for sharing the crystallographic data of NISTmAb based on the individual *F*c and *F*ab XRD structures with a grafted hinge. We thank Wyatt N. Vreeland and Curtis W. Meuse for critical reviews and comments on the manuscript.

ASSOCIATED CONTENT

**Supporting Information.** The results of PCA, 2DCOS, and peak fitting of the *in situ* IR spectra of insulin (Figures S1−S4). The SNR analysis results of the amide II band (Figure S5). This materialis available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

Young Jong Lee – orcid.org/0000-0001-7754-3001; Email: youngjong.lee@nist.gov

**Authors**

**Bonghwan Chon** – orcid.org/0000-0002-3882-1045

**Shuyu Xu** –orcid.org/0000-0001-6313-3773

Author Contributions

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

**Notes**

The authors declare that they have no conflict of interest.

**DISCLAIMER**

Certain commercial equipment, instruments, or materials are identified in this paper to foster understanding. Such identification does not imply endorsement by NIST, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

REFERENCES

(1) Baker, M. J.; Trevisan, J.; Bassan, P.; Bhargava, R.; Butler, H. J.; Dorling, K. M.; Fielden, P. R.; Fogarty, S. W.; Fullwood, N. J.; Heys, K. A.; Hughes, C.; Lasch, P.; Martin-Hirsch, P. L.; Obinaju, B.; Sockalingum, G. D.; Sulé-Suso, J.; Strong, R. J.; Walsh, M. J.; Wood, B. R.; Gardner, P.; Martin, F. L. Using Fourier Transform IR Spectroscopy to Analyze Biological Materials. *Nat. Protoc.* **2014**, *9* (8), 1771–1791.

(2) Jackson, M.; Mantsch, H. H. The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30* (2), 95–120.

(3) Yang, H.; Yang, S.; Kong, J.; Dong, A.; Yu, S. Obtaining Information about Protein Secondary Structures in Aqueous Solution Using Fourier Transform IR Spectroscopy. *Nat. Protoc.* **2015**, *10* (3), 382–396.

(4) Perro, A.; Lebourdon, G.; Henry, S.; Lecomte, S.; Servant, L.; Marre, S. Combining Microfluidics and FT-IR Spectroscopy: Towards Spatially Resolved Information on Chemical Processes. *React. Chem. Eng.* **2016**, *1* (6), 577–594.

(5) Childs, D. T. D.; Hogg, R. A.; Revin, D. G.; Rehman, I. U.; Cockburn, J. W.; Matcher, S. J. Sensitivity Advantage of QCL Tunable-Laser Mid-Infrared Spectroscopy over FTIR Spectroscopy. *Appl. Spectrosc. Rev.* **2015**, *50* (10), 822–839.

(6) Yeh, K.; Kenkel, S.; Liu, J. N.; Bhargava, R. Fast Infrared Chemical Imaging with a Quantum Cascade Laser. *Anal. Chem.* **2015**, *87* (1), 485–493.

(7) Zhang, D.; Li, C.; Zhang, C.; Slipchenko, M. N.; Eakins, G.; Cheng, J. Depth-Resolved Mid-Infrared Photothermal Imaging of Living Cells and Organisms with Submicrometer Spatial Resolution. *Sci. Adv.* **2016**, *2* (September), e1600521.

(8) Faist, J.; Capasso, F.; Sivco, D. L.; Sirtori, C.; Hutchinson, A. L.; Cho, A. Y. Quantum Cascade Laser. *Science* **1994**, *264* (5158), 553–556.

(9) Alcaráz, M. R.; Schwaighofer, A.; Kristament, C.; Ramer, G.; Brandstetter, M.; Goicoechea, H.; Lendl, B. External-Cavity Quantum Cascade Laser Spectroscopy for Mid-IR Transmission Measurements of Proteins in Aqueous Solution. *Anal. Chem.* **2015**, *87* (13), 6980–6987.

(10) Tittel, F. K.; Richter, D.; Fried, A. Mid-Infrared Laser Applications in Spectroscopy. *Solid-State Mid-Infrared Laser Sources* **2007**, *516* (7), 458–529.

(11) Kendrick, B. S.; Gabrielson, J. P.; Solsberg, C. W.; Ma, E.; Wang, L. Determining Spectroscopic Quantitation Limits for Misfolded Structures. *J. Pharm. Sci.* **2019**, *109* (1), 933–936.

(12) Akhgar, C. K.; Ramer, G.; Żbik, M.; Trajnerowicz, A.; Pawluczyk, J.; Schwaighofer, A.; Lendl, B. The Next Generation of IR Spectroscopy: EC-QCL-Based Mid-IR Transmission Spectroscopy of Proteins with Balanced Detection. *Anal. Chem.* **2020**, *92* (14), 9901–9907.

(13) Hua, Q. X.; Weiss, M. A. Mechanism of Insulin Fibrillation: The Structure of Insulin under Amyloidogenic Conditions Resembles a Protein-Folding Intermediate. *J. Biol. Chem.* **2004**, *279* (20), 21449–21460.

(14) Sanchez-Ruiz, J. M. Theoretical Analysis of Lumry-Eyring Models in Differential Scanning Calorimetry. *Biophys. J.* **1992**, *61* (4), 921–935.

(15) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2007**, *1* (6), 2876–2890.

(16) Xue, C.; Lin, T. Y.; Chang, D.; Guo, Z. Thioflavin T as an Amyloid Dye: Fibril Quantification, Optimal Concentration and Effect on Aggregation. *R. Soc. Open Sci.* **2017**, *4* (1), 160696.

(17) Schwaighofer, A.; Alcaráz, M. R.; Araman, C.; Goicoechea, H.; Lendl, B. External Cavity-Quantum Cascade Laser Infrared Spectroscopy for Secondary Structure Analysis of Proteins at Low Concentrations. *Sci. Rep.* **2016**, *6* (April), 33556.

(18) Schwaighofer, A.; Alcaraz, M. R.; Lux, L.; Lendl, B. PH Titration of β-Lactoglobulin Monitored by Laser-Based Mid-IR Transmission Spectroscopy Coupled to Chemometric Analysis. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2020**, *226*, 117636.

(19) Segal, A. R.; Sayed, N. El. Are You Ready for More Insulin Concentrations? *J. Diabetes Sci. Technol.* **2015**, *9* (2), 331–338.

(20) Groenning, M.; Frokjaer, S.; Vestergaard, B. Formation Mechanism of Insulin Fibrils and Structural Aspects of the Insulin Fibrillation Process. *Curr. Protein Pept. Sci.* **2009**, *10* (5), 509–528.

(21) Schultz, B. J.; Mohrmann, H.; Lorenz-Fonfria, V. A.; Heberle, J. Protein Dynamics Observed by Tunable Mid-IR Quantum Cascade Lasers across the Time Range from 10 Ns to 1 S. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2018**, *188*, 666–674.

(22) Carter, D. C.; Ho, J. X. Structure of Serum Albumin. *Adv. Protein Chem.* **1994**, *45* (C), 153–176.

(23) Gokarn, Y.; Agarwal, S.; Arthur, K.; Bepperling, A.; Day, E. S.; Filoti, D.; Greene, D. G.; Hayes, D.; Kroe-Barrett, R.; Laue, T.; Lin, J.; McGarry, B.; Razinkov, V.; Singh, S.; Taing, R.; Venkataramani, S.; Weiss, W.; Yang, D.; Zarraga, I. E. Biophysical Techniques for Characterizing the Higher Order Structure and Interactions of Monoclonal Antibodies. *ACS Symp. Ser.* **2015**, *1201*, 285–327.

(24) Amenabar, I.; Poly, S.; Nuansing, W.; Hubrich, E. H.; Govyadinov, A. A.; Huth, F.; Krutokhvostov, R.; Zhang, L.; Knez, M.; Heberle, J.; Bittner, A. M.; Hillenbrand, R. Structural Analysis and Mapping of Individual Protein Complexes by Infrared Nanospectroscopy. *Nat. Commun.* **2013**, *4*, 2890.

(25) Bro, R.; Smilde, A. K. Principal Component Analysis. *Anal. Methods* **2014**, *6* (9), 2812–2831.

(26) Noda, I. Vibrational Two-Dimensional Correlation Spectroscopy (2DCOS) Study of Proteins. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2017**, *187*, 119–129.

(27) Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An Analytical Solution to the Kinetics of Breakable Filament Assembly. *Science* **2009**, *326* (5959), 1533–1537.

(28) Morris, A. M.; Watzky, M. A.; Agar, J. N.; Finke, R. G. Fitting Neurological Protein Aggregation Kinetic Data via a 2-Step, Minimal/"Ockham’s Razor" Model: The Finke-Watzky Mechanism of Nucleation Followed by Autocatalytic Surface Growth. *Biochemistry* **2008**, *47* (8), 2413–2427.

(29) Cohen, S. I. A.; Vendruscolo, M.; Welland, M. E.; Dobson, C. M.; Terentjev, E. M.; Knowles, T. P. J. Nucleated Polymerization with Secondary Pathways. I. Time Evolution of the Principal Moments. *J. Chem. Phys.* **2011**, *135* (6), 065105.

(30) Newbury, N. R.; Coddington, I.; Swann, W. Sensitivity of Coherent Dual-Comb Spectroscopy. *Opt. Express* **2010**, *18* (8), 7929–7945.

(31) Huber, M.; Trubetskov, M.; Hussain, S. A.; Schweinberger, W.; Hofer, C.; Pupeza, I. Optimum Sample Thickness for Trace Analyte Detection with Field-Resolved Infrared Spectroscopy. *Anal. Chem.* **2020**, *92* (11), 7508–7514.

TOC

Diagram

Description automatically generated