

Analyst

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Pushing the detection limit of infrared spectroscopy for structural analysis of dilute protein samples

Maurizio Baldassarre¹ and Andreas Barth¹

Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

Keywords: Fourier-transform infrared spectroscopy; protein structure; amide I; infrared filters; cellulose; germanium; pyruvate kinase.

Abbreviations used: FT-IR, Fourier-transform infrared; amide I', amide I in ²H₂O medium; RMS, root mean square; SNR, signal-to-noise ratio.

Abstract

Fourier-transform infrared spectroscopy is a powerful and versatile tool to investigate the structure and dynamics of proteins in solution. The intrinsically low extinction coefficient of the amide I mode, the main structure-related oscillator, together with the high infrared absorptivity of aqueous media, requires that proteins are studied at high concentrations ($> 10 \text{ mg L}^{-1}$). This may represent a challenge in the study of aggregation-prone proteins and peptides, and questions the significance of structural data obtained for proteins physiologically existing at much lower concentrations. Here we describe the development of a simple experimental approach that increases the detection limit of protein structure analysis by infrared spectroscopy. Our approach relies on custom-made filters to isolate the amide I region ($1700\text{--}1600 \text{ cm}^{-1}$) from irrelevant spectral regions. The sensitivity of the instrument is then increased by background attenuation, an approach consisting in the use of a neutral density filter, such as a non-scattering metal grid, to attenuate the intensity of the background spectrum. When the filters and grid are combined, a 2.4-fold improvement in the noise level can be obtained. We have successfully tested this approach using a highly diluted solution of pyruvate kinase in deuterated medium (0.2% w/v), and found that it provides spectra of a quality comparable to those recorded with a 10-fold higher protein concentration.

¹**Corresponding authors:** Stockholm University, Department of Biochemistry and Biophysics, Svante Arrhenius väg 16C, SE-106 91 Stockholm, Sweden. E-mail addresses: maurizio.baldassarre@dbb.su.se, andreas.barth@dbb.su.se.

Contents

1	Introduction	2
2	Materials and Methods	3
2.1	Preparation of samples for infrared spectroscopy	3
2.2	Infrared spectroscopy	4
3	Results and Discussion	4
3.1	Potential filters for use in infrared spectroscopy of proteins	4
3.2	The cellulose filter	6
3.3	The cellulose-germanium filter	7
3.4	Background attenuation	8
3.5	Application to the analysis of highly diluted protein samples	9
4	Conclusions	11
5	Acknowledgments	12
	References	12

1 Introduction

Since the appearance of instruments based on the Michelson interferometer and digital systems for fast Fourier transformation of the data, infrared spectroscopy has represented an invaluable tool in the study of protein structure, stability and dynamics.¹⁻⁶ This technique offers several advantages over other types of optical spectroscopies, including (1) the ability to simultaneously probe the secondary structure and flexibility of proteins (for samples prepared in $^2\text{H}_2\text{O}$),⁷⁻¹² (2) the ability to follow binding and enzymatic processes through changes in the absorptions of amino acid side chains and ligands/substrates,¹³⁻¹⁷ and (3) the possibility to study proteins both in liquid and solid phases due to its low susceptibility to scattering phenomena.¹⁸⁻²⁰ More and more interest is growing regarding the latter aspect, for it allows to shed light on complex biological processes related to human health. These include, among others, amyloidoses and misfolding diseases, which often lead to formation of insoluble protein aggregates.^{20,21} Infrared spectroscopy, however, suffers from low sensitivity because vibrational transitions have intrinsically low extinction coefficients, ϵ . Typical values of ϵ (in $\text{M}^{-1} \text{cm}^{-1}$) are $\sim 200\text{--}300$ and $\sim 300\text{--}400$ for the C=O stretching vibration in carboxylic acids and amides, respectively, and even lower for vibrations arising from less polarised bonds.^{5,22}

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

These values are one to three orders of magnitude lower than those usually measured for electronic transitions. The main consequence is that the chromophores of interest must be analysed at concentrations typically between 1 and 100 mM.⁵

The amide I band (1700–1600 cm⁻¹) is the main band of interest for structural analysis of proteins. It arises mainly from the C=O stretching vibration of the peptide bond (~80%) with minor contributions from the out-of-phase CN stretching vibration, the CCN deformation and the NH in-plane bend.^{5,23} The latter is responsible for the sensitivity of the amide I band to *N*-deuteration of the backbone. The C=O group in peptide groups is sterically more hindered than it is in side chains, therefore the corresponding infrared band is narrower and displays a higher peak absorption. Typical values of ϵ in H₂O are 600–800 for α -helices (1650–1640 cm⁻¹), and 800–1000 for β -sheets (1630–1620 cm⁻¹).²⁴ In ²H₂O, the ϵ of α -helices is slightly decreased (450–485 M⁻¹ cm⁻¹), but it remains approximately the same for β -sheets.²⁴ Because of the abundance of peptide bonds in proteins, the latter are typically studied at concentrations of approximately 10–20 mg mL⁻¹ (100–200 mM peptide groups). Although most globular proteins are soluble at such a concentration, this may represent a challenge in the study of proteins and peptides that are prone to aggregate or that have undergone partial misfolding,^{25,26} as well as question the significance of structural data obtained for proteins physiologically existing at much lower concentrations.^{27–29}

Research on the soluble oligomers of the small (4.3–4.5 kDa) A β ₄₀ and A β ₄₂ peptides calls for new ways to obtain high-quality infrared spectra of 100–200 μ M peptide solutions (~4–8 mM peptide groups). The planned use of peptides isotopically labelled at specific positions will require, eventually, to push this limit even further. In this work, we report on the use of cost-effective, custom-made infrared filters that allow to improve the amide I' analysis of a model α/β protein, pyruvate kinase (PK). One of these filters, regenerated cellulose, is inexpensive but performs very well by removing unwanted frequencies (> 2800 and < 1500 cm⁻¹) from the spectrum of PK. We have coupled a cellulose filter to a commercial long wave pass, coated germanium filter, and obtained a significant increase in the signal-to-noise ratio. Further improvement has been achieved by inserting a non-scattering metal grid on the background position in our instrument. This allows to take advantage of the full sensitivity of our detector without any non-linearity issues. We have successfully applied these concepts to the study of a very diluted (0.2% and 0.02% w/v) PK solutions.

2 Materials and Methods

Pyruvate kinase (PK) from rabbit muscle and glycerol (99.5%) were from Sigma-Aldrich. Deuterium oxide (99.9% ²H) was from Cambridge Isotope Laboratories. Silicone-based vacuum grease was from Dow Corning. Regenerated cellulose membranes (25- μ m thick) were purchased from Membrane Filtration Products.

2.1 Preparation of samples for infrared spectroscopy

Lyophilised PK was resuspended directly in ²H₂O to final concentrations of 0.2, 2 and 20 mg mL⁻¹ (0.02, 0.2 and 2% w/v, respectively), mixed gently to complete dissolution

1
2
3
4
5 and incubated at 4 °C for 1–2 h to allow for H/²H exchange. Protein samples (3 μL)
6 were deposited on a CaF₂ window with a trough of 55 μm, covered with a flat CaF₂
7 window and placed in the spectrometer. A thin layer of vacuum grease applied between
8 the windows on the outer circumference ensured long-term protection from moisture
9 contamination and evaporation. The same sample (at one of the three concentrations
10 reported above) was used when recording spectra with different filter settings. Glycerol
11 and vacuum grease samples were prepared in a similar way.
12
13

14 2.2 Infrared spectroscopy

15
16 Infrared spectra were recorded with a Tensor 37 Fourier-transform spectrometer (Bruker
17 Optics, Germany), equipped with a liquid N₂-cooled HgCdTe detector and continuously
18 purged with CO₂-free, dry air. Samples were mounted on a two-position sample shuttle,
19 which allowed for interleaved acquisition of sample and background spectra. A waiting
20 time of 20 min was allowed after inserting the sample to ensure complete purging. The
21 infrared cell was kept at 20 °C throughout the measurement by means of an external
22 water bath circulator. Interferograms were recorded at a resolution of 2 cm⁻¹, apodised
23 using a 3-term Blackman-Harris apodisation function and Fourier-transformed with a
24 zero-filling factor of 2. Eight consecutive interferograms were averaged to obtain a single
25 sample or background spectrum. This procedure yields infrared absorption spectra that
26 are intrinsically noisy. However, the use of a sample shuttle ensured that the “noise” in
27 the spectra does not originate from changes in water vapour content between sample and
28 background spectra, but rather from detector noise.
29
30

31 Infrared spectra were recorded and analysed using the OPUS software from the in-
32 strument manufacturer. Second derivative spectra were calculated using a smoothing
33 length of 9 data points (approx. 9 cm⁻¹). For the sample at the lowest concentration tested
34 (0.2 mg mL⁻¹), additional second derivative spectra were calculated with a smoothing
35 length of 17 data points. The noise levels were calculated between 1900 and 1800 cm⁻¹ in
36 the second derivative spectra, using a built-in function of OPUS, and are reported as root
37 mean square (RMS). Protein spectra shown throughout the text are uncorrected spectra,
38 *i.e.* the absorption spectrum of the deuterated buffer was not subtracted. Arithmetic op-
39 erations between two or more spectra lead to increased noise levels, which would have
40 made quality assessment of our filters more difficult.
41
42
43
44
45
46
47

48 3 Results and Discussion

49 3.1 Potential filters for use in infrared spectroscopy of proteins

50
51 Most infrared spectrometers meant to operate in the mid-infrared spectral range () are
52 equipped with either Deuterated Tri-Glycine Sulfate (DTGS) or Mercurium Cadmium
53 Telluride (HgCdTe, or MCT) detectors. DTGS detectors are inexpensive and show high
54 linearity. They suffer, however, from low sensitivity and slow response rates since de-
55 tection of IR light is based on physical heating of a resistive element.³⁰ Photoconductive
56 detectors, such as MCT detectors, typically show response rates several orders of mag-
57 nitude higher than DTGS detectors, and are up to ten times as sensitive, making them
58
59
60

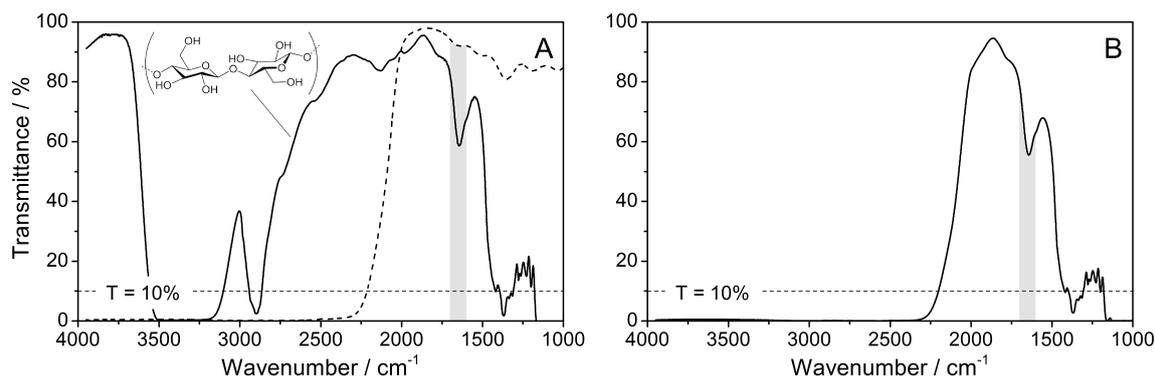


Figure (1) Infrared transmission spectra of cellulose and germanium filters. (A) Transmission spectrum of a 25- μm thick regenerated cellulose sheet (continuous line) and a Ge filter (dashed line); (B) Transmission spectrum of the combined cellulose-germanium filter. In panel A the chemical structure of cellulose is shown. The area highlighted in grey denotes the region where the amide I band is observed.

well suited for most biochemical or biophysical applications.³¹ The increased sensitivity of MCT detectors means, however, that they become non-linear and saturate more easily. This is particularly true in Fourier-transform spectrometers, where light intensities of all wavelengths generated by the source reach the detector simultaneously. Because the interest in most measurements is often restricted to specific regions of the infrared spectrum, the need to limit the light throughput to avoid detector non-linearity and/or saturation means that the sensitivity in these regions is sub-optimal. This drawback can be avoided by the use of optical filters, *i.e.* filters that “shape” the emission spectrum of the source so that light in irrelevant regions of the spectrum is blocked. This allows to increase the light intensity, and consequently the sensitivity, in the regions of interest.

In order to be used as filters to improve structural analysis of proteins by infrared spectroscopy, optical filters must exhibit strong, saturating absorption above or below the amide I band (1700–1600 cm^{-1}). The ability of a filter to block light in irrelevant spectral ranges on both sides of a band of interest, such as in bandpass filters, is a desirable property. In an attempt to develop custom-made, cost-effective filters for the improved analysis of protein structures, we have screened for compounds containing functional groups that are expected to strongly absorb infrared radiation in regions both above and below the amide I region, while poorly absorbing within the latter. These include compounds containing Si—O (1150–1000 cm^{-1}), Si—C (~ 1260 cm^{-1}) and C—H (3000–2900 cm^{-1}) bonds, such as polydimethylsiloxane, and compounds containing C—O (1300–1000 cm^{-1}) and O—H (3500–3200 cm^{-1}) bonds, such as alcohols. The transmittance spectra of two candidate compounds, silicone-based vacuum grease and glycerol, are shown in Supplementary Figure S1. Comparison of the spectra of vacuum grease and glycerol reveals that although both compounds can be used as filters to improve the analysis of the amide I band, the latter is potentially a better filter because it generates a narrower clear range around the amide I region and, at the same time, absorbs more light in other regions of the mid-IR spectrum. The main drawback associated with the use of glycerol as an infrared filter arises from the fact that it is fluid, and thus needs to be sandwiched in an IR-transparent support, such as two CaF_2 windows. A compound that is similarly rich

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

in hydroxyl groups, but has the advantage of being solid and stable, is cellulose. Regenerated cellulose, such as that used in dialysis membranes, is thin (25 μm) and resistant, and its infrared transmission spectrum, shown in Figure 1A, is similar to that of glycerol (Supplementary Figure S1B). This makes cellulose a more promising compound and prompts further tests on its potential use as a filter for the analysis of protein structures.

3.2 The cellulose filter

In order to test whether the use of cellulose as a filter improves the detection of the amide I band components, we recorded infrared absorption spectra of 20 mg mL⁻¹ (2% w/v) pyruvate kinase (PK) in deuterated medium without any filter and with a filter obtained from a 25- μm thick dialysis membrane made of regenerated cellulose. The spectra were recorded at different apertures to explore the entire linear range of our HgCdTe detector so as to find the conditions that yield the lowest noise levels, and thus the highest signal-to-noise ratio. This is necessary because all filters reduce the intensities of both the sample and background spectra. The second derivative spectra of PK recorded without any filter and with a cellulose filter are shown in Figure 2, panels A and B, respectively.

In the absence of any filter, an aperture of 1.0 mm yields the best spectrum, as judged by visual inspection of the spectra region between 1900 and 1800 cm⁻¹ in panel A (Figure 2). For this spectrum, the noise calculated between 1900 and 1800 cm⁻¹ in four independent measurements is $(6.55 \pm 0.69) \times 10^{-6}$ (Table 1). A smaller aperture (0.5 mm) increases the noise level (grey trace), while a larger one (1.5 mm) distorts the spectrum and introduces larger noise levels than with an aperture of 0.5 mm (data not shown). The second derivative spectrum of PK shows three distinct signals in the amide I' region. The band at 1652 cm⁻¹ can be assigned to α -helices, while the band at 1630 cm⁻¹, together with the weaker bands at 1686 and 1682 cm⁻¹ can be assigned to anti-parallel β -sheets.^{5,32-35} The shoulder at 1644 cm⁻¹ might arise from highly hydrated α -helices, such as those most exposed to the solvent, as well as from residues in unstructured regions, such as loops. A detailed structural analysis of PK is beyond the purpose of this work. However, the type and relative abundance of the main types of secondary structures are consistent with the crystal structure of the protein (PDB ID: 1PKN),³⁶ as well as with previous infrared spectroscopy studies of PK.³⁷ The signals below 1610 cm⁻¹ originate from amino acid side chain vibrations. These include the bands at 1606 and 1515 cm⁻¹, arising from vibrations of the guanidyl and phenyl moieties of arginines and tyrosines, respectively, and those at 1584 and 1575 cm⁻¹, arising from the carboxylate groups of deprotonated aspartates and glutamates, respectively.²² The broader signal at ~ 1550 cm⁻¹ arises from the residual amide II band, *i.e.* from the in-plane N—H bending vibration of peptide groups that are not accessible to the deuterated solvent and, thus, have not undergone H/²H exchange.³⁸

Using a cellulose filter and the same aperture value (grey spectrum, panel B), the noise increases to $(7.65 \pm 1.38) \times 10^{-6}$. Increasing the aperture to 1.5 mm, however, decreases the noise level to $(4.71 \pm 0.50) \times 10^{-6}$. In order to quantitatively compare the different conditions, we define the Improvement Factor (IF) as follows:

$$\text{IF} = \frac{\text{noise}^{\text{ref}}}{\text{noise}^i} \quad (1)$$

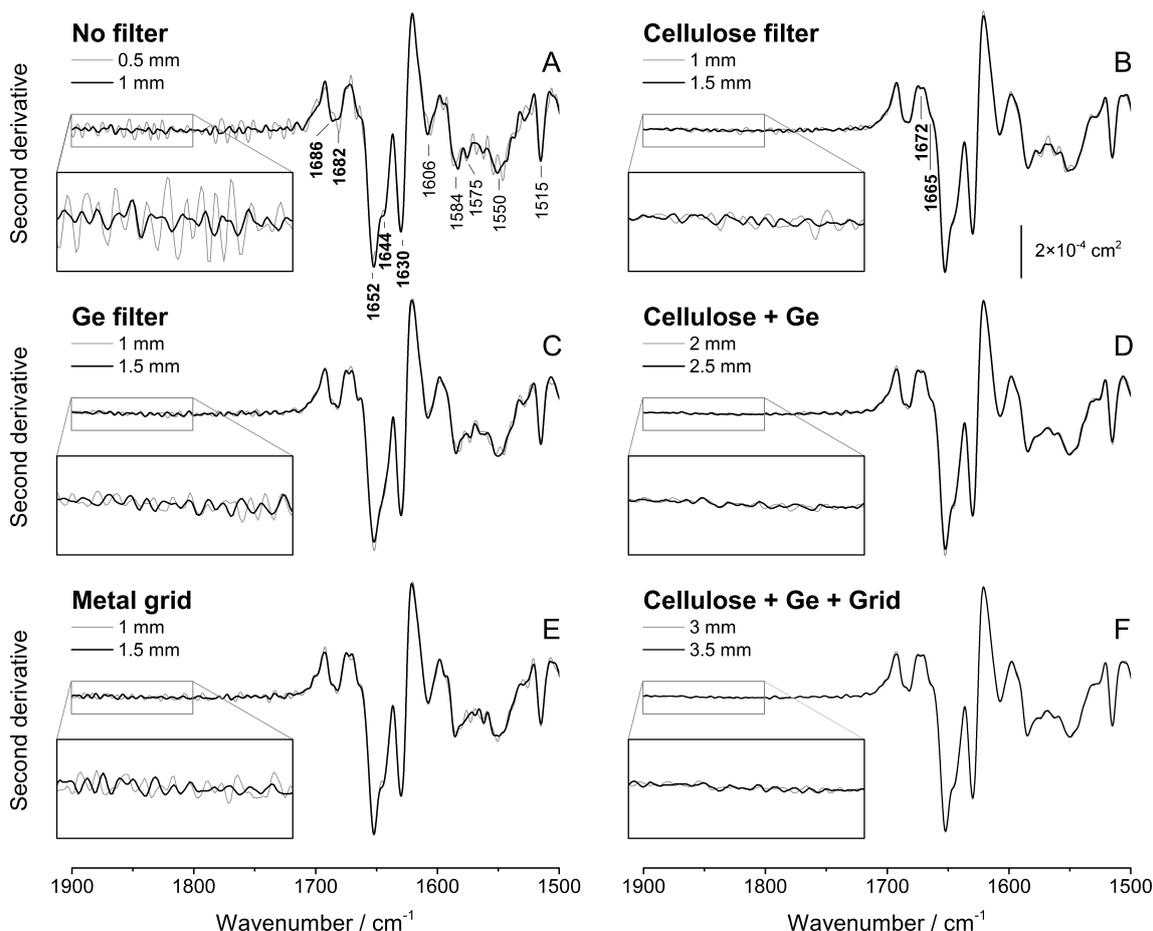


Figure (2) Second derivative spectra of 2% w/v PK in deuterated medium recorded using several filter combinations. The box in each panel provides an enlarged view of the 1900–1800- cm^{-1} region used for determination of the noise levels. All boxes have the same scale. The origin of the labelled signals is described in the text. Bold labels denote signals within the amide I' band.

where noise^{ref} is the noise of the reference spectrum, which we consider to be the second derivative of the absorption spectrum recorded without any filter and an aperture value of 1.0 mm, and noise^i is the noise level (expressed as RMS) of a spectrum with a particular filter and a particular aperture value. The IF of the spectrum recorded with the cellulose filter and an aperture of 1.5 mm is 1.39 (Table 1). In this spectrum, two more signals, probably arising from turn-like structures, can be detected at 1672 and 1665 cm^{-1} .

3.3 The cellulose-germanium filter

The previous results have shown that the use of cellulose as a filter reduces the noise levels at or near the amide I region. Inspection of the transmission spectrum of cellulose reveals, however, that large frequency intervals are allowed to pass through the filter. These regions do not contain information of interest for the structural analysis of proteins, and include the regions between 4000 and 3500 cm^{-1} and between 2800 and 2000 cm^{-1} (Figure 1A). In order to filter out these regions, we have coupled the cellulose filter to a long wave pass, coated germanium filter (LOT GmbH, Darmstadt, Germany) with a

Type of filter	Aperture (mm)	Noise (Avg \pm SD)/ 10^{-6}	IF
No filter	0.5	19.4 ± 1.60	0.34
	1.0	6.55 ± 0.69	1.00
Cellulose	1.0	7.65 ± 1.38	0.86
	1.5	4.71 ± 0.50	1.39
Germanium	1.0	5.37 ± 0.28	1.22
	1.5	3.32 ± 0.21	1.97
Cellulose + germanium	2.0	3.50 ± 0.10	1.87
	2.5	3.05 ± 0.04	2.15
Metal grid	1.0	8.11 ± 0.95	0.81
	1.5	5.17 ± 0.12	1.27
Cellulose + germanium + grid	3.0	2.85 ± 0.10	2.30
	3.5	2.76 ± 0.06	2.38

Table (1) Noise levels with different filter and aperture sets. The noise levels are reported as RMS and represent the average (Avg) and standard deviation (SD) of four independent measurements.

cut-off at $\sim 2200 \text{ cm}^{-1}$ (thereby referred to as “germanium” or “Ge” filter for simplicity). The transmission spectra of the germanium filter and the combined cellulose-germanium filter are shown in Figure 1, panels A (dashed line) and B, respectively. Comparison of the spectrum of the combined cellulose-germanium filter (panel B) with that of cellulose (panel A) reveals that the former generates a much narrower transmittance window around the amide I region (grey-shaded area). This is expected to reduce the noise and improve the SNR in the amide I region by allowing the use of larger apertures. Indeed, with the combined cellulose-germanium filter the lowest noise level [$(3.05 \pm 0.04) \times 10^{-6}$] is obtained with an aperture of 2.5 mm. This corresponds to an IF value of 2.15. The combined cellulose-germanium filter performs significantly better than the individual filters of which it is composed (Table 1 and panels B–D in Figure 2).

3.4 Background attenuation

The most straightforward way to improve the SNR in infrared spectra is to maximise the amount of light that reaches the detector. This can be achieved, for instance, by increasing the aperture on spectrometers equipped with variable aperture wheels. The properties of a HgCdTe detector dictates the largest aperture that can be used before linearity is lost. Additionally, because the background (or reference) spectrum is the one that transmits more light, the largest aperture threshold is defined by this spectrum, rather than by the sample spectrum. One way to overcome this limitation would be to attenuate the intensity of the reference spectrum with a neutral density filter so that its intensity is comparable to that of the sample spectrum, with the intensities of both still lying in the linear range of the detector. This introduces a constant, negative offset (ΔAbs) in the absorption spectrum of the sample, as shown in Supplementary Figure S2. In this way,

both spectra can be recorded under conditions where detector response is highest and still linear. The easiest way to perform this would be to record the background spectrum using an IR cell of equal pathlength filled with the buffer. Although this is in principle feasible and has the advantage of directly removing the contribution of the strong H₂O and ²H₂O bands, it might also introduce artifacts in the resulting spectrum if the two cells are not thermostatted at exactly the same temperature. Additionally, should setbacks arise during recording of the background spectrum, *e.g.* air bubbles in the cell, this would irreversibly ruin the entire measurement. It is more practical, though slower, to record sample and buffer spectra separately using the same cell and experimental conditions, and then manually adjust the subtraction factor.

To overcome this drawback, a metal grid (mesh size ~0.5 mm) transmitting only ~60% of the light was placed in the empty position of the sample shuttle, and the spectra of PK were recorded at different apertures. Figure 2E shows the second derivative spectra of PK recorded with apertures of 1.0 (grey trace) and 1.5 (black trace) mm without using any filter other than the grid. The spectrum recorded with an aperture of 1.0 mm has a higher noise level ($(5.17 \pm 0.12) \times 10^{-6}$) than the same spectrum recorded in the absence of the grid (Figure 2A). However, while 1.0 mm is the largest aperture that can be used without any type of filter, the presence of the metal grid on the reference position allows to increase this value to 1.5 mm, which results in a 1.27-fold improvement (*i.e.* reduction) in the noise level. In addition to this, the comparison of the spectra reveals that the metal grid does not introduce any artifact in the resulting spectrum. The absorption spectrum of the metal grid recorded with a 1-mm aperture consists, indeed, of a flat, featureless line in the spectral range under examination (see Supplementary Figure S2).

Figure 2F shows the second derivative spectra of PK recorded using a combination of all the filters described previously: the cellulose filter, the coated germanium filter and the metal grid on the reference position of the sample shuttle. Under these conditions, an aperture of 3.5 mm yields the lowest noise level: $(2.76 \pm 0.06) \times 10^{-6}$ (Table 1). This represents a 2.4-fold improvement in the noise level with respect to the spectrum recorded in the absence of any filter.

3.5 Application to the analysis of highly diluted protein samples

The results presented so far show that protein spectra of increasing quality can be obtained by using (1) a cellulose filter, (2) a combined cellulose-germanium filter and (3) a combined cellulose-germanium filter with the use of a metal grid when recording the background spectrum. To test whether this approach can be useful to increase the detection limit so as to study highly diluted protein samples, we recorded absorption spectra of PK at 2 mg mL^{-1} (0.2% w/v, ~20 mM peptide groups) and 0.2 mg mL^{-1} (0.02% w/v, ~2 mM peptide groups). These concentration values are one order and two orders of magnitude lower than the one commonly reported in the literature.^{10–12,39}

The second derivative spectrum of 0.2% w/v PK recorded without any filter and an aperture of 1.0 mm is shown in Figure 3A. Although it is possible to observe features in the amide I' region that are reminiscent of the bands described previously, the spectrum shows large noise levels with intensities comparable to those of the amide I' component bands. Because of the noisy appearance of the spectrum, the bands at 1686 and 1682

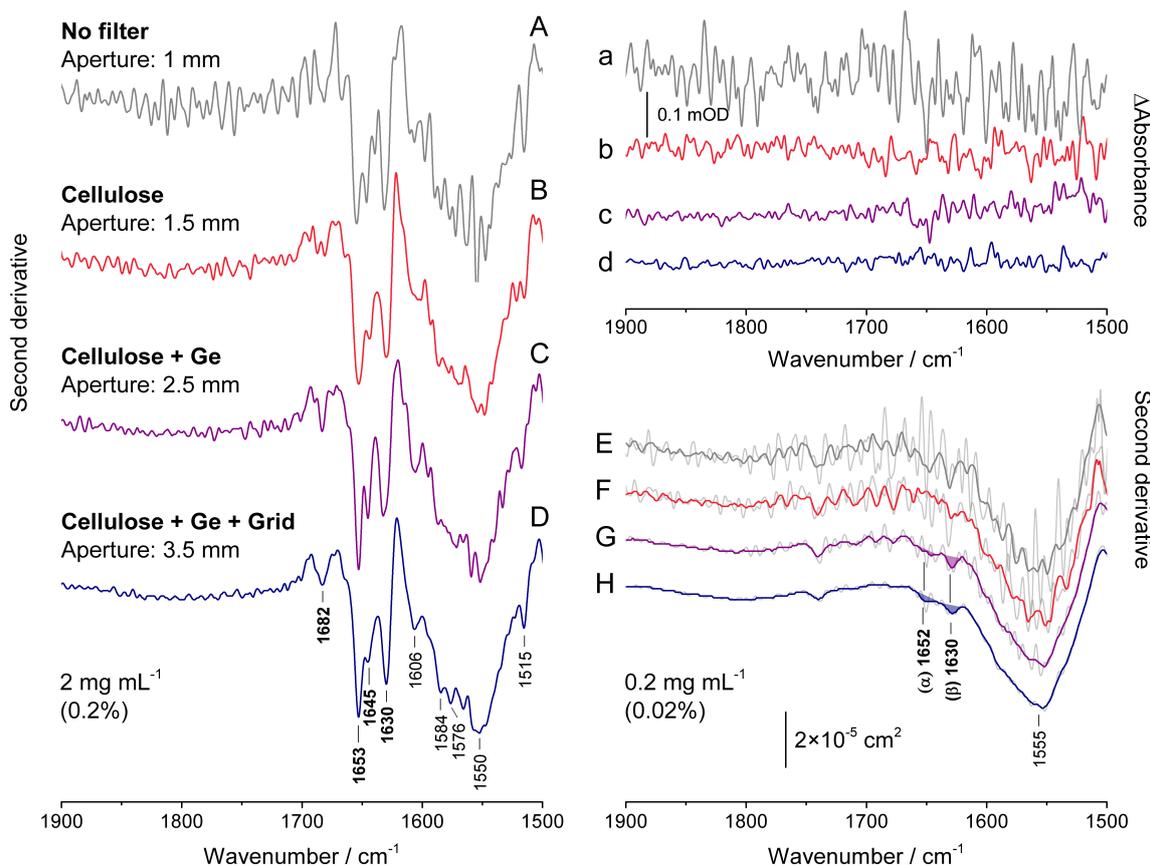


Figure (3) Second derivative spectra of 0.2% (A–D) and 0.02% w/v (E–H) PK in deuterated medium recorded using several filter combinations. Each spectrum in panels E–H is the average of four second derivatives spectra calculated using 9 (thin grey lines) and 17 (coloured lines) data points. The origin of the labelled peaks is described in the text. Bold labels and shading denote signals within the amide I' band. Panels A–H share the same scale. Panels a–d show the noise levels obtained by subtracting two consecutive absorption spectra recorded under each of the filter conditions reported in A–D. The difference absorption (Δ Ab) spectra show that the noise level is higher at lower frequencies because the cellulose filter has lower transmittance near 1500 cm^{-1} than near 1900 cm^{-1} .

cm^{-1} , for instance, cannot be observed. Using a cellulose filter and an aperture of 1.5 mm (panel B), the spectrum appears less noisy, and the individual bands contributing to the amide I' absorption can be better discriminated. However, uncertainties in the amino acid side absorptions remain. For instance, the arginine and tyrosine bands at 1606 and 1515 cm^{-1} cannot be assigned unequivocally. Coupling the cellulose filter to a commercial long wave pass, coated germanium filter further improves the quality of the spectrum. With an aperture of 2.5 mm (panel C), all the features of the amide I' band, with the exception of the 1686 and 1682 cm^{-1} couplet, can be discriminated. Correct assignment of both the amide I' and side chain bands, however, can only be obtained by introducing a metal grid on the reference position, thereby allowing to increase the aperture to 3.5 mm (panel D). The spectrum of PK recorded with this filter arrangement is in very good agreement with the one recorded using the same settings and a 10-fold higher protein concentration (Figure 2F). The higher absorption between 1600 and 1500

1
2
3
4
5 cm^{-1} , resulting in a more pronounced negative second derivative signal, arises from the
6 higher contribution of the “combination + libration” band of $^2\text{H}_2\text{O}$ ($\sim 1555 \text{ cm}^{-1}$) in the
7 low protein-concentration sample.

8
9 In a further set of measurements, infrared absorption spectra of a 10-fold more di-
10 luted PK solution (0.2 mg mL^{-1} , or 0.02%) were recorded with the same filter settings
11 described previously. The second derivatives of these spectra, calculated with a smooth-
12 ing length of 9 data points, are shown as thin lines in Figure 3, panels E–H. Even with
13 the best performing filter settings (panel H, thin grey line), however, it is nearly impos-
14 sible to distinguish the amide I' component bands from the diffuse noise, indicating that
15 such a low protein concentration lies beyond the detection limit of our instrumentation.
16 Using a wider smoothing window (17 data points) when calculating the second deriva-
17 tives, however, leads to the appearance of two weak bands at 1652 and 1630 cm^{-1} in the
18 spectra recorded with the combined cellulose-germanium filters (panels G and H, thick
19 lines). Their positions correspond to α -helices and β -sheets, respectively. Weaker signals
20 observed previously, such as the high-frequency component of anti-parallel β -sheets, as
21 well as signals originating from amino acid side chains, cannot be observed. The presence
22 of the two main signals in the amide I' region suggests, however, that the main secondary
23 structure elements can be detected at protein concentrations as low as 0.02%.
24
25
26
27
28

29 4 Conclusions

30
31 In this work, we report on a methodological approach that allows for improved struc-
32 tural analysis of dilute protein samples by Fourier-transform infrared spectroscopy. This
33 approach is based on the use of custom-made filters that isolate the amide I band, *i.e.* the
34 main structure-related oscillator in proteins, from unwanted spectral frequencies. One
35 of the filters tested, regenerated cellulose, performs particularly well by filtering out in-
36 frared radiation both above and below the amide I band ($1700\text{--}1600 \text{ cm}^{-1}$), due to the
37 presence of the strong O—H (broad, $\sim 3500\text{--}3200 \text{ cm}^{-1}$) and C—H (sharp, $\sim 3000\text{--}2900$
38 cm^{-1}) stretching vibrations, and to the weaker C—H ($\sim 1470 \text{ cm}^{-1}$) bending and C—
39 O ($< 1200 \text{ cm}^{-1}$) stretching vibrations. Coupling this filter to a commercial long wave
40 pass, coated germanium filter further increases the signal-to-noise ratio. By using a non-
41 scattering metal grid to reduce the intensity of the background spectrum, an approach
42 that we name “background attenuation”, the quality of the infrared spectrum of pro-
43 teins can be dramatically increased since this allows to increase the amount of light that
44 reaches the detector by using larger apertures. This leads to a 2.4-fold decrease in the
45 noise level, corresponding to a 2.4-fold increase in the signal-to-noise ratio. To the best
46 of our knowledge, this represents the first report in literature describing the latter ap-
47 proach. If developed further, it may represent a substantial improvement of commercial
48 infrared spectrometers available nowadays. A layout of the experimental setup is shown
49 in Supplementary Figure S3.
50
51
52
53
54
55

56 The combined approaches described in this work may provide significant advantages
57 in the study of protein structures by allowing to study proteins at concentrations at least
58 one order of magnitude lower than those usually reported in the literature. Despite being
59 an advantage *per se*, this will prove to be particularly useful in the study of aggregation-
60 prone proteins and peptides, as well as those proteins that are physiologically present

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

at low concentrations. Additionally, this will enhance the analysis of proteins studied in normal water (H₂O), where the requirement for short (~5 μm) pathlengths leads to lower signal-to-noise ratios than in measurements performed in heavy water (²H₂O). Our approach will introduce significant advantages in the study of proteins and peptides isotopically labelled at specific positions by allowing to discriminate the weak signals arising from the few labelled groups from the majority of the unlabelled groups.

5 Acknowledgments

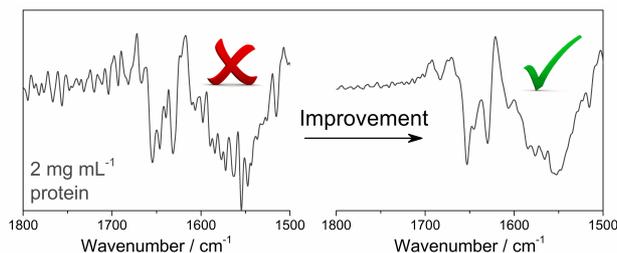
This work was supported by a grant from Alzheimerfonden (to A.B.) and a post-doctoral fellowship from Wenner-Gren Stiftelsen (to M.B.). Knut och Alice Wallenbergs Stiftelsen and Lars Hiertas Minne Stiftelsen are gratefully acknowledged for providing the spectrometer and the sample shuttle, respectively.

References

- [1] Goormaghtigh, E., Cabiaux, V., and Ruyschaert, J. M. *Subcell Biochem* **23**, 405–450 (1994).
- [2] Cooper, E. A. and Knutson, K. *Pharm Biotechnol* **7**, 101–143 (1995).
- [3] Arrondo, J. L. and Goñi, F. M. *Prog Biophys Mol Biol* **72**(4), 367–405 (1999).
- [4] Barth, A. and Zscherp, C. *Q Rev Biophys* **35**(4), 369–430 Nov (2002).
- [5] Barth, A. *Biochim Biophys Acta* **1767**(9), 1073–1101 Sep (2007).
- [6] Doglia, S. M., Ami, D., Natalello, A., Gatti-Lafranconi, P., and Lotti, M. *Biotechnol J* **3**(2), 193–201 Feb (2008).
- [7] Surewicz, W. K., Mantsch, H. H., and Chapman, D. *Biochemistry* **32**(2), 389–394 Jan (1993).
- [8] Costantino, H. R., Griebenow, K., Mishra, P., Langer, R., and Klibanov, A. M. *Biochim Biophys Acta* **1253**(1), 69–74 Nov (1995).
- [9] Carpenter, J. F., Prestrelski, S. J., and Dong, A. *Eur J Pharm Biopharm* **45**(3), 231–238 May (1998).
- [10] Pelton, J. T. and McLean, L. R. *Anal Biochem* **277**(2), 167–176 Jan (2000).
- [11] Manning, M. C. *Expert Rev Proteomics* **2**(5), 731–743 Oct (2005).
- [12] Kong, J. and Yu, S. *Acta Biochim Biophys Sin (Shanghai)* **39**(8), 549–559 Aug (2007).
- [13] Wharton, C. W. *Nat Prod Rep* **17**(5), 447–453 Oct (2000).
- [14] Thoenges, D. and Barth, A. *J Biomol Screen* **7**(4), 353–357 Aug (2002).
- [15] Kumar, S. and Barth, A. *Sensors (Basel)* **10**(4), 2626–2637 (2010).
- [16] Kumar, S. and Barth, A. *Biophys J* **98**(9), 1931–1940 May (2010).
- [17] Kumar, S. and Barth, A. *J Phys Chem B* **115**(20), 6784–6789 May (2011).
- [18] Chen, X., Knight, D. P., Shao, Z., and Vollrath, F. *Biochemistry* **41**(50), 14944–14950 Dec (2002).
- [19] Dzwolak, W., Smirnovas, V., Jansen, R., and Winter, R. *Protein Sci* **13**(7), 1927–1932 Jul (2004).

- 1
2
3
4
5 [20] Sarroukh, R., Goormaghtigh, E., Ruyschaert, J.-M., and Raussens, V. *Biochim Biophys Acta* **1828**(10), 2328–2338 Oct (2013).
6
7 [21] Tamm, L. K. and Tatulian, S. A. *Biochemistry* **32**(30), 7720–7726 Aug (1993).
8
9 [22] Barth, A. *Prog Biophys Mol Biol* **74**(3-5), 141–173 (2000).
10 [23] Fabian, H. and Mäntele, W. *Handbook of Vibrational Spectroscopy*, chapter Infrared Spectroscopy of Proteins Infrared Spectroscopy of Proteins Infrared Spectroscopy of Proteins. Wiley (2006).
11
12 [24] Venyaminov, S. Y. and Kalnin, N. N. *Biopolymers* **30**(13-14), 1259–1271 (1990).
13
14 [25] Fink, A. L. *Folding and Design* **3**(1), R9–23 (1998).
15
16 [26] Shivu, B., Seshadri, S., Li, J., Oberg, K. A., Uversky, V. N., and Fink, A. L. *Biochemistry* **52**(31), 5176–5183 Aug (2013).
17
18 [27] Jackson, M. and Mantsch, H. H. *Crit Rev Biochem Mol Biol* **30**(2), 95–120 (1995).
19 [28] Hauser, K. *Encyclopedia of Biophysics*, chapter Infrared-spectroscopy of protein folding, misfolding and aggregation, 1089–1095. Springer Berlin-Heidelberg (2013).
20
21 [29] Ami, D., Natalello, A., Lotti, M., and Doglia, S. M. *Microb Cell Fact* **12**, 17 (2013).
22 [30] Smith, B. C., editor. *Fundamentals of Fourier Transform Infrared Spectroscopy, Second Edition*, chapter How an FTIR works, 19–54. CRC Press (2011).
23 [31] Rogalski, A. *Infrared Physics & Technology* **43** (3–5), 187–210 (2002).
24 [32] Bandekar, J. *Biochim Biophys Acta* **1120**(2), 123–143 Apr (1992).
25 [33] Zuber, G., Prestrelski, S. J., and Benedek, K. *Anal Biochem* **207**(1), 150–156 Nov (1992).
26 [34] Arrondo, J. L., Muga, A., Castresana, J., and Goñi, F. M. *Prog Biophys Mol Biol* **59**(1), 23–56 (1993).
27 [35] Kumar, S. and Barth, A. *J Phys Chem B* **115**(39), 11501–11505 Oct (2011).
28 [36] Larsen, T. M., Laughlin, L. T., Holden, H. M., Rayment, I., and Reed, G. H. *Biochemistry* **33**(20), 6301–6309 May (1994).
29 [37] Yu, S., Lee, L. L.-Y., and Lee, J. C. *Biophys Chem* **103**(1), 1–11 Jan (2003).
30 [38] Baldassarre, M., Scirè, A., and Tanfani, F. *Biomedical Spectroscopy and Imaging* **1**, 247–259 (2012).
31 [39] Haris, P. I. and Chapman, D. *Methods Mol Biol* **22**, 183–202 (1994).
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table of content



Optical filters dramatically reduce the noise of protein infrared spectra and allow for a more accurate structural analysis of dilute protein samples