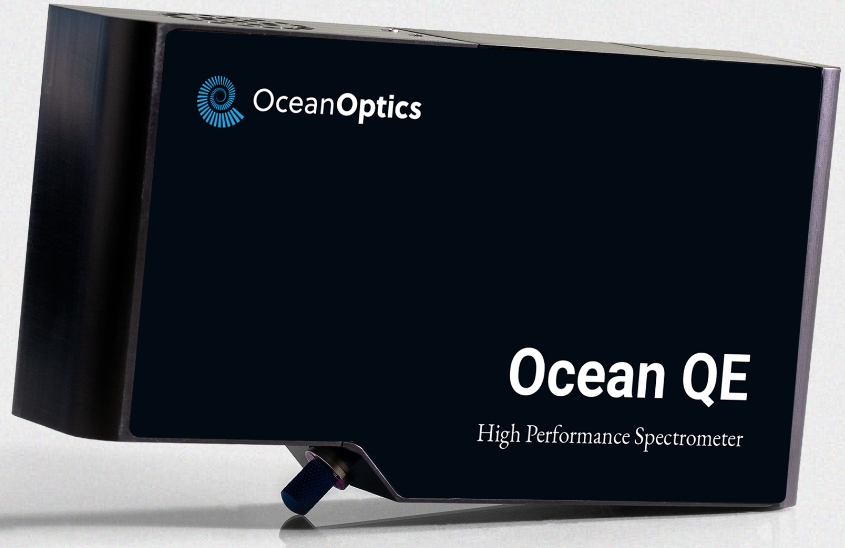
Getting Your Protein from Raman

Application Note



KEYWORDS

- Anaerobic bacterium
- Bovine collagen protein
- Microbial processes

TECHNIQUES

- Absorbance
- Raman

APPLICATIONS

- Protein monitoring
- Concentration validation
- Species identification

In this application note, we evaluate the use of 785 nm Raman analysis as a complement to traditional UV-visible spectroscopy in analyzing proteins. Both techniques offer valuable analytical insight into complex bioengineering processes for applications including medical diagnostics and pharmaceuticals development.

Burger King fast food restaurants may not be quite as popular as they once were, but their philosophy of "Have It Your Way" has been adopted by many other modern industries. Don't like hotels? Stay at an Airbnb that fits your specific preferences. Don't like watching scheduled programming? Stream whatever show you want, whenever you want. Think medicines are too generic? Have a custom treatment engineered and produced based on your specific genome.

This last example has become a hot topic in recent years, with public genome tests and wearable health devices providing unprecedented data around a single individual (1). Creating these custom treatments often requires using bioreactors with special monitoring capabilities beyond the typical generic systems producing generic compounds. For decades, those systems have used

US +1 727-733-2447
EUROPE +31 26-3190500
ASIA +86 21-6295-6600

info@oceanoptics.com • **www.oceanoptics.com**

UV-vis spectroscopy to monitor the concentration of proteins and other molecules in the mix, and as a result, this method has become relatively inexpensive, easy, and trusted.

Figure 1 shows an example of one such UV-vis system, using an SR-2XR spectrometer to monitor bovine collagen protein from 1-100 g/L and using 310 nm as a concentration correlation. Note the dramatic increase in UV response while the visible and NIR regions remain extremely steady.

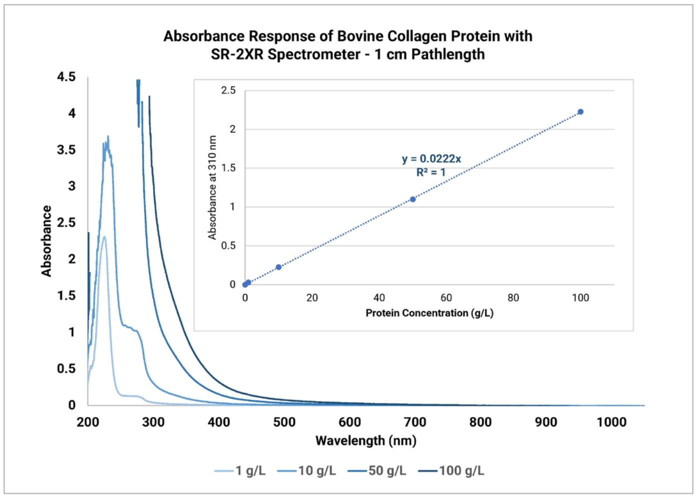


Figure 1. The absorbance of a protein measured with a UV-vis spectrometer. Note the strong UV response.

This tried-and-true UV-vis method fits the bill for many mainstream processes, but more complex synthesis requires more advanced technologies. Historically, Raman spectroscopy is rather expensive compared to absorbance/concentration approaches, but this is starting to change in recent years (2). Raman spectroscopy started off being quite costly, but its popularity has worked to drive down cost, and the lower cost has thus worked to further drive popularity.

Comparing UV-vis and Raman Spectroscopy for Protein Analysis

Raman is a different animal than standard broadband spectroscopy, with unique difficulties in the setup and acquisitions. For example, one has to consider the laser focal distance and laser power, as well as the software settings of integration time, averaging, and boxcar (more on these later). And even after all that, the user may end up with data trends that are not intuitive or necessarily even coherent. This all sounds like a lot of time, money, and trouble to generate what looks like mere noise.

But indeed, there is value in using Raman as a complement to the traditional UV-vis approach, both for concentration validation and species identification. Let's take a look at the same 1-100 g/L protein samples from Figure 1 but scanned with a QE *Pro* 785 nm Raman system.

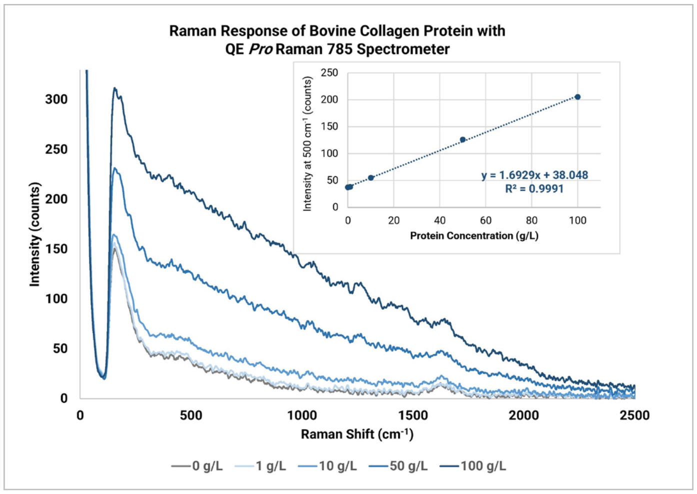


Figure 2. The Raman response of the protein is measured with a 785 nm Raman spectrometer. The results are rich in detail.

There's a good bit to unpack here (Figure 2). Perhaps the first item to note is the x-axis; this is no longer showing wave***length*** but rather wave***number***, as a function of 785 nm as the reference energy level. Raman works on the principle of measuring energy shifts away from some known excitation energy (i.e., your laser), so the units here convert for that reason. Next, observe how broad the response is. Our UV-vis example isolated itself to the left side of the graph, but here we see concentration increases everywhere. This is due to the large nature of protein molecules, which give much broader and general emissions versus small organic molecules with sharp peaks that correspond to specific functional groups.

This broad response gives us the ability to use a single wavenumber for concentration correlation, as shown in the inlaid plot using 500 cm^{-1} (Figure 2), or to use a broader area such as the integral over many wavenumbers. But either way this approach is inherently less sensitive to concentration than the UV-vis method, and by a lot. The equation for resolution can be summarized as:

$$Resolution = \frac{((2*STDEV(Input_{low}))*\Delta Output)}{ABS(\Delta Input)}$$

For the UV-vis approach our input is absorbance units (AU) and for Raman it is intensity (counts); in each case the output we are looking for is g/L protein. If we assume the UV-vis system offers a standard deviation of 0.0005 AU, and using the 0.0222

AU/(g/L) slope from the Figure 1 calibration plot, we see that particular system achieving about 0.09 g/L protein resolution, or ≈100 mg/L. Doing the same for Raman with the assumption of 10 counts standard deviation and the slope of 1.7 counts/(g/L), we resolve to 12 g/L, or about 2-orders-of-magnitude worse than the UV-vis method.

So why pay over $10k for a sanity check that's worse by 2-orders-of-magnitude? Complex mixtures are where the Raman system shines, because the ability to pull out traces of small molecules can make-or-break an expensive custom synthesis process.

Exploring Raman Analysis Benefits

Let's take a look at a real-world example where *Clostridium autoethanogenum* is used in the carbon-negative production of isopropanol (3). This anaerobic bacterium produces ethanol via a well-known pathway but can be further utilized in the production of isopropanol and acetone. Bacteria may generate these small organic molecules as waste, which then may become reactants with other components in the mix, perhaps yielding unexpected products.

Before we get to the Raman part, let's start again with the UV-vis method and see how the introduction of organic solvents affects those trends. We've been calling this the UV-vis method so far because most protein activity is within those lower wavelengths, but the extended-range Ocean SR model spectrometer used here has response up to 1100 nm, well into the NIR. The added value of this becomes clear very quickly (Figure 3).

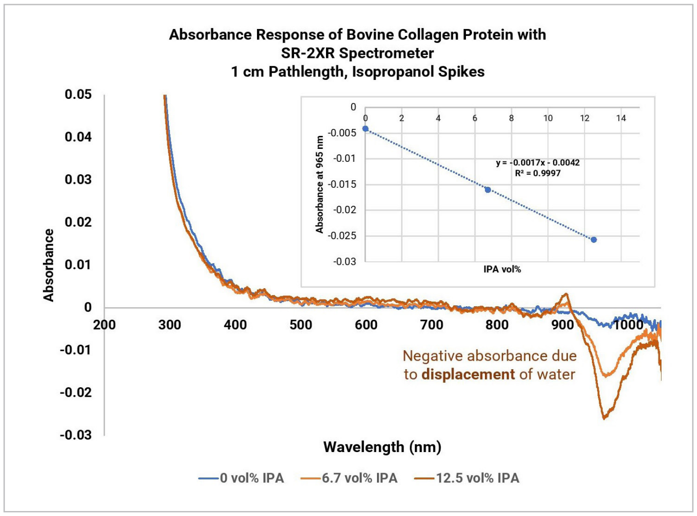


Figure 3. Here is the UV-vis absorbance response of a protein solution with isopropanol added.

Note the negative absorbance in the NIR region; this is not directly due to the presence of isopropanol but rather the absence of water. Within a constant pathlength there is only so much real estate for light interaction, and as alcohol concentration goes up some water molecules are displaced, thus creating a negative dip in the water-absorbing regions. But also note the calibration plot is still very linear with a 0.9997 R^2 value; we can still use these numbers for IPA concentration correlation ***if*** we know where they are coming from. This is where Raman analysis comes in.

The Raman measurement tells you *what* is causing the water displacement in the NIR so you can confidently turn the UV-vis-NIR absorbance numbers into something meaningful. Take a look at a protein Raman scan with and without isopropanol spiked-in at 6.7 vol% (Figure 4).

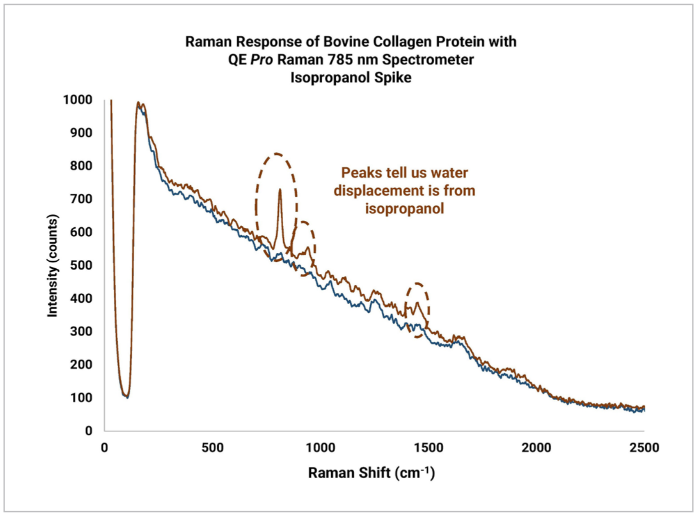


Figure 4. The Raman response of a protein solution with isopropanol additions clues us in to the reason for water displacement in the NIR region of the spectrum.

Those peak locations are fingerprints for isopropyl alcohol, and since we have no other fingerprints present we can assume our NIR response is due to this and not ethanol, acetone or similar. We now see the value in having complementary techniques, with one acting as leverage for the other.

Optimizing Raman Analysis Parameters

Earlier we mentioned the software settings of integration time, averaging, and boxcar. These can be tricky to get right for Raman since it's a game of timing and peak resolution. In general, boxcar averaging should be set very low for Raman measurements, perhaps 1-3 but no more than 5, so that peaks are not artificially muted. That said, zero boxcar also can be dangerous by creating "false peaks" from single-pixel behaviors. The main trade-off comes down to integration time and averaging, each of which changes the total scan time.

In Figure 5, the comparative plots show several variations of integration time and boxcar settings as we measured our 1-100 g/L protein samples spiked with 6.7 vol% isopropanol.

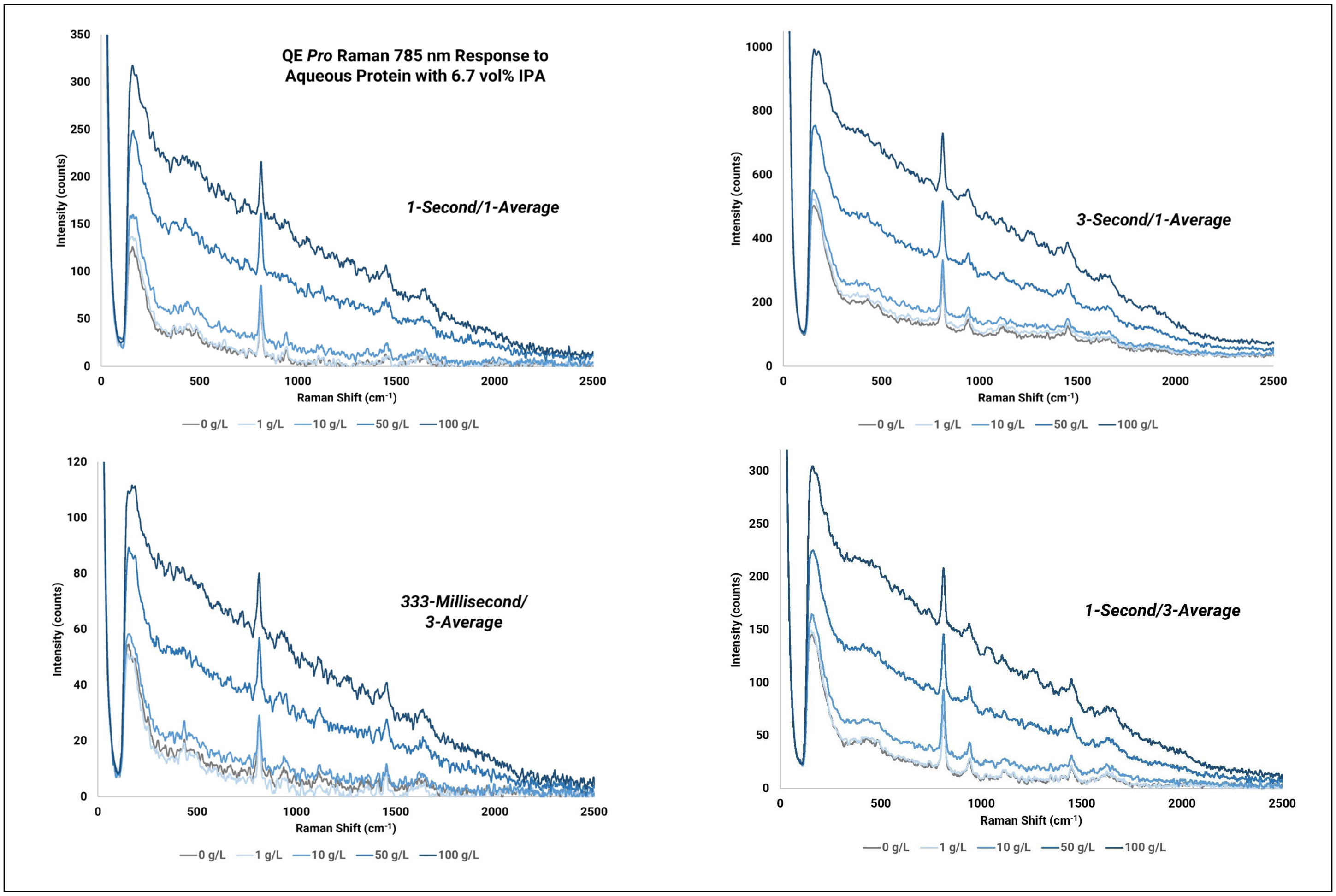


Figure 5. A comparative analysis of Raman emissions captured at variable integration times and boxcar averaging reveals trade-offs inherent to the system measurement settings.

The main observation is that the plots with longer scan times look smoother and less noisy, which will almost always be the case. Patience is a virtue. But you don't have all the time in the world, so for a given scan time do you give more weight to averaging or to integration time? (For an example application note where these considerations are played out, search *Raman Spectroscopy to Monitor UV Curing in Semiconductor Production* at oceanoptics.com.)

Averaging will typically do more for you than integration time, since integration will merely scale everything while averaging will add meaningful coherence. It may be difficult to see in Figure 5 but look at the plots with the peak below 1500 cm^{-1}. You can see how much sharper these are for the 333 msec/3-average scans versus the more jagged 1 sec/1-average scans. With each giving a total scan time of 1 second, the lower integration with higher averaging gives a more ideal output.

This sets the foundation for more complex mixtures where several organics are present, with Raman giving information on "what's there" and absorbance giving information on "how much." Running these methods in parallel gives the user notably more visibility over their process than with one method alone.

Summarizing UV-vis and Raman Protein Measurements

These are the primary takeaways from this study:

- Traditional UV-vis-NIR spectroscopy absorbance methods are better for concentration calculations than Raman methods. Use Raman to provide a general sanity check or validation of those numbers via an independent approach.
- The true value of Raman is in small molecule identification (qualification), which then feeds back into the absorbance method for quantification. This analysis becomes critical in modern and custom biosynthesis processes where microbes may generate multiple organics.
- For cleaner Raman peaks at a constant scan time, lean on averaging more than on integration time.

References

1. **Vallely-Gilroy, Caoimhe.** Personalized Medicine: *The Industry's Future. International Society for Pharmaceutical Engineering.* [Online] ISPE, December 2021. https://ispe.org/pharmaceutical-engineering/november-december-2021/personalized-medicine-industrys-future.

2. *New advances in using Raman spectroscopy for the characterization of catalysts and catalytic reactions.* **Hess, Christian.** 5, s.l. : Chemical Society Reviews, 2021, Vol. 50. doi.org/10.1039/D0CS01059F.

3. *Clostridium autoethanogenum isopropanol production via native plasmid pCA replicon.* **Nogle, Robert, et al.** s.l. : Frontiers in Bioengineering and Biotechnology, 2022, Vol. 10. doi.org/10.3389/fbioe.2022.932363.