Supplementary Information

High-Speed Coherent Raman Fingerprint Imaging of Biological Tissues

Charles H. Camp Jr.,[1] Young Jong Lee,[1] John M. Heddleston,[2] Christopher M. Hartshorn,[1]
Angela R. Hight Walker,[2] Jeremy N. Rich,[3] Justin D. Lathia,[4] and Marcus T. Cicerone[1,*]

[1]*Biosystems and Biomaterials Division, National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD 20899, USA*
[2]*Semiconductor and Dimensional Metrology Division, National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD 20899, USA*
[3]*Department of Stem Cell Biology and Regenerative Medicine, Cleveland Clinic, 9500 Euclid Ave, Cleveland, OH 44195, USA*
[4]*Department of Cellular and Molecular Medicine, Cleveland Clinic, 9500 Euclid Ave, Cleveland, OH 44195, USA*

I. 2-COLOUR AND 3-COLOUR EXCITATION MECHANISMS

The broadband coherent anti-Stokes Raman scattering (BCARS) system uses a hybrid 2-colour/3-colour approach to excite Raman transitions spanning from the Raman fingerprint region ($< 1{,}800$ cm^{-1}) to beyond 3,600 cm^{-1}. Raman energies above $\approx 2{,}100$ cm^{-1} are excited with a 2-colour mechanism in which the pump and probe are degenerate– the excitation mechanism most often used in coherent anti-Stokes Raman scattering (CARS), multiplex CARS (MCARS), and BCARS experiments. The fingerprint region, on the other hand, is excited with intrapulse 3-colour excitation. With this 3-colour excitation, the pump and Stokes sources are degenerate (intrapulse excitation) and the excitation profile is determined by the permutations of available frequencies (energies) of light[1]. Because the highest number of permutations is for closely spaced frequencies, the 3-colour excitation profile increases with decreasing wavenumber. It is for this reason that intrapulse 3-colour excitation is particularly well suited for stimulating Raman transitions within the vibrational fingerprint region. Additionally, the excitation profile of the 3-colour mechanism decreases with increasing wavenumber and with practical consideration of the bandwidth afforded by the 16 fs supercontinuum (SC) laser, we designed our system so that as the 3-colour excitation profile drops off within the Raman quiescent region ($\approx 1{,}800$ cm^{-1} to 2,600 cm^{-1}), the 2-colour excitation profile emerges, exciting Raman transitions ranging from $\approx$ 2,100 cm^{-1} to 3,600 cm^{-1}. With this system utilizing two distinct excitation schemes, it is instructive to analyze the differences between 2-colour and 3-colour excitation and to understand the effects on the resonant (Raman) and nonresonant signal generation. In this supplemental section, we will highlight 3 key points:

- With increasing pump and Stokes source bandwidths and fixed average power, the intensity of 3-colour excited components rises and 2-colour excited components falls
- The resonant-to-nonresonant signal ratio is primarily tied to probe source excitation parameters, not the pump and Stokes sources, under typical operating conditions
- The probe source spectral characteristics predominantly determine the spectral resolution of CARS instruments

With illumination of these key findings, one can appreciate the power afforded by the presented BCARS system and understand its strength for spectroscopic detection.

The earliest examination of the effects of source bandwidth in CARS microscopy demonstrated a marked increase of NRB generation with increasing source bandwidth but a far more limited increase in the resonant (Raman) component under the assumption of fixed source energies[2]. Under this framework, decreasing the source bandwidths stood to improve the chemical sensitivity and specificity for isolated Raman peaks. That work, although presenting a general mathematical model for the CARS generation process, focused on narrowband CARS with single-element detection; 2-colour excitation; and pump, Stokes, and probe sources with the same bandwidths. Since the time of this early contribution, other works have presented theoretical and experimental results concerning source characteristics, such

*Electronic address: cicerone@nist.gov

as chirp[3–6], polarization[6,7], and general pulse shaping paradigms for enhanced excitation efficiency[8,9]. Although the latter two studies implemented a 3-colour excitation method, these works did not address the uniqueness of the excitation mechanism. More recently, a comparison of 2-colour and 3-colour excitation was presented[1], highlighting distinct characteristics of the two. In this supplementary section, we will expand upon this early comparison in an effort to clarify the mechanistic distinctions. To begin, we will develop a classical model of the CARS mechanism and demonstrate analytical solutions available for certain simplified scenarios, comparing 2-colour and 3-colour excitation. It should be noted that this classical model does not account for effects such as source depletion and saturation of vibrational and electronic states, but under typical operating conditions, it should adequately describe signal generation. Additionally, we will present supporting simulations and experimental results.

For the CARS process, the output intensity, $I_{CARS}(\omega_{as})$, at anti-Stokes frequency ω_{as} is proportional to the squared-modulus of the 3^{rd}-order nonlinear polarization[10], $P^{(3)}(\omega_{as})$:

$$I_{CARS}(\omega_{as}) \propto |P^{(3)}(\omega_{as})|^2 \tag{1}$$

$$P^{(3)}(\omega_{as}) \propto \iiint \chi^{(3)}(\omega_{as};\omega_p,-\omega_s,\omega_{pr})\times E_p(\omega_p)E_s^*(\omega_s)E_{pr}(\omega_{pr})\delta(\omega_{as}-\omega_p+\omega_s-\omega_{pr})d\omega_p d\omega_s d\omega_{pr}, \tag{2}$$

where $\chi^{(3)}(\omega_{as};\omega_p,-\omega_s,\omega_{pr})$ is the nonlinear susceptibility; $E_p(\omega_p)$ is the pump electric field; $E_s(\omega_s)$ is the Stokes electric field; $E_{pr}(\omega_{pr})$ is the probe electric field; and ω_p, ω_s, and ω_{pr} are the pump, Stokes, and probe frequencies, respectively. We will describe the nonlinear susceptibility as an addition of chemically-nonspecific and chemically-specific terms:

$$\chi^{(3)}(\omega_{as};\omega_p,-\omega_s,\omega_{pr}) = \chi_{NR} + \chi_R(\omega_{as}) = \chi_{NR} + \sum_m \frac{A}{\Omega_m - (\omega_p-\omega_s) - i\Gamma_m}, \tag{3}$$

where $\chi_R(\omega_{as})$ is the nonlinear susceptibility for the resonant components and χ_{NR} is the nonlinear susceptibility for the nonresonant component that generates the nonresonant background (NRB). It should be noted that the nonresonant component, under near-infrared excitation, is typically assumed to be real and slowly varying in frequency (although not necessarily constant). Within the expansion of $\chi_R(\omega_{as})$, A_m, Ω_m, Γ_m describe the Lorentzian profile of the mth Raman peak at frequency Ω_m with half-width Γ_m. It should be noted that (to a first approximation) the imaginary component of the nonlinear susceptibility, $\Im\{\chi^{(3)}\}$, is proportional to spontaneous Raman response (spectra). For notational clarity, we will abbreviate the form of the nonlinear susceptibility to $\chi^{(3)}(\omega_p-\omega_s)$. Applying this to equation (2) and integrating over the pump frequencies simplifies the description of the nonlinear polarization:

$$P^{(3)}(\omega_{as}) \propto \int \chi^{(3)}(\omega_{as}-\omega_{pr}) \left[\int E_s^*(\omega_s)E_p(\omega_{as}+\omega_s-\omega_{pr})d\omega_s\right] E_{pr}(\omega_{pr})d\omega_{pr}, \tag{4}$$

which may be written in a more tractable form:

$$P^{(3)}(\omega_{as}) \propto \left\{\chi^{(3)}(\omega_{as})\left[E_s(\omega_{as}) \star E_p(\omega_{as})\right]\right\} * E_{pr}(\omega_{as}), \tag{5}$$

where '$\star$' and '$*$' are the cross-correlation and convolution operations, respectively. From equation (5), in view of the relation in equation (1), we can establish some intuitive insights about the CARS generation process. Firstly, the cross-correlation term describes the energy (frequency) profile available for material excitation. As the cross-correlation term spectrally broadens, so too will the range of Raman transitions that can be stimulated. Secondly, if the cross-correlation term is sufficiently broad (larger than a Raman lineshape), the probe source bandwidth will determine the spectral resolution of the system, i.e., the narrower the probe bandwidth, the narrower the recorded CARS lineshape–converging to a full-width at half-maximum (FWHM) of $2\Gamma_m$ for the mth peak. Finally, if the cross-correlation term becomes an autocorrelation, as in the case of intrapulse 3-colour excitation, the material excitation profile will be necessarily centered at $\omega_{as}=0$ and symmetric about this point; although, actual measurements correspond to $\omega_{as}>0$.

If we assume all sources have Gaussian spectral profiles and real envelopes, i.e., are transform limited and temporally centered, we may write the cross-correlation term:

$$E_s(\omega) \star E_p(\omega) = \int E_s^*(\omega')E_p(\omega+\omega')d\omega' \tag{6}$$

$$= \frac{E_{s0}E_{p0}\sigma_p\sigma_s\sqrt{2\pi}}{\sqrt{\sigma_p^2+\sigma_s^2}} e^{-\frac{(\omega-\omega_{p0}+\omega_{s0})^2}{2(\sigma_p^2+\sigma_s^2)}}, \tag{7}$$

where $E_s(\omega)$ and $E_p(\omega)$ are the Stokes and pump fields with amplitudes E_{s0} and E_{p0}, center frequencies ω_{s0} and ω_{p0}, and 1/e-intensity half-widths of σ_s and σ_p, respectively. For the special case of 3-colour excitation, in which the pump and Stoke sources are degenerate, the cross-correlation is an autocorrelation:

$$E_{p,s}(\omega) \star E_{p,s}(\omega) = |E_{p0,s0}|^2 \sigma_{p,s} \sqrt{\pi} e^{-\frac{\omega^2}{4\sigma_{p,s}^2}}, \tag{8}$$

where we have noted the degeneracy of the pump and Stokes fields by using $E_{p,s}(\omega)$, with amplitude $E_{s0,p0}$, 1/e-intensity half-width $\sigma_{p,s}$, and frequency offset $\omega_{p0,s0}$ (although all offset-frequency terms cancel out). For a Gaussian field, $A(\omega)$, of the form $A_0 \exp\{-\omega^2/2\sigma^2\}$, the average power, $\mathcal{P}_A$, is proportional to $|A_0|^2\sigma\sqrt{\pi}$; therefore, from equations (7) and (8):

$$E_s(\omega) \star E_p(\omega) \propto \sqrt{\mathcal{P}_p \mathcal{P}_s} \frac{\sqrt{2\sigma_s\sigma_p}}{\sqrt{(\sigma_p^2+\sigma_s^2)}} e^{-\frac{(\omega-\omega_{p0}+\omega_{s0})^2}{2(\sigma_p^2+\sigma_s^2)}} \tag{9}$$

$$E_{p,s}(\omega) \star E_{p,s}(\omega) \propto \mathcal{P}_{p0,s0} e^{-\frac{\omega^2}{4\sigma_{p,s}^2}}. \tag{10}$$

For equation (10), the material excitation from 3-colour stimulation, the maximum amplitude occurs at $\omega = 0$. Additionally, this maximum amplitude is constant regardless of source bandwidth with fixed average power. At spectral positions shifted from the origin, the material response will increase with source bandwidth. The case for 2-colour excitation differs in that the maximum material response will occur at $\omega = \omega_{p0} - \omega_{s0}$, and will decrease with increasing source bandwidth. For the scenario representative of most multispectral CARS experiments, in which the Stoke source is significantly broader than the pump source, $\sigma_s >> \sigma_p$, equation (9) simplifies:

$$E_s(\omega) \star E_p(\omega)|_{\sigma_s >> \sigma_p} \propto \sqrt{\mathcal{P}_p \mathcal{P}_s} \sqrt{\frac{2\sigma_p}{\sigma_s}} e^{-\frac{(\omega-\omega_{p0}+\omega_{s0})^2}{2(\sigma_p^2+\sigma_s^2)}}. \tag{11}$$

Under these conditions, with fixed average power, the 2-colour material excitation maximum amplitude will drop $\propto 1/\sqrt{\sigma_s}$.

To evaluate these findings and their effects on the total output signal, we will evaluate equation (5) for the case of a nonresonant material (with the simplification of a real, constant-valued χ_{NR}):

$$I_{CARS}(\omega) \propto \left| \{\chi_{NR} [E_s(\omega) \star E_p(\omega)]\} * E_{pr}(\omega) \right|^2. \tag{12}$$

Using the cross-correlation and autocorrelation terms in equations (16) and (17) and applying a Gaussian field probe, the 2-colour, $I_{2C}(\omega)$, and 3-colour, $I_{3C}(\omega)$, CARS signals may be written:

$$I_{2C}(\omega) \propto \left| \frac{2\pi\chi_{NR} E_{s0}^* E_{p0,pr0}^2 \sigma_{p,pr}^2 \sigma_s}{\sqrt{2\sigma_{p,pr}^2+\sigma_s^2}} e^{-\frac{(\omega-2\omega_{p0,pr0}+\omega_{s0})^2}{2(2\sigma_{p,pr}^2+\sigma_s^2)}} \right|^2 \tag{13}$$

$$= \frac{4\pi^2\chi_{NR}^2 |E_{s0}|^2 |E_{p0,pr0}|^4 \sigma_{p,pr}^4 \sigma_s^2}{2\sigma_{p,pr}^2+\sigma_s^2} e^{-\frac{(\omega-2\omega_{p0,pr0}+\omega_{s0})^2}{(2\sigma_{p,pr}^2+\sigma_s^2)}} \tag{14}$$

$$\propto \frac{4\sqrt{\pi}\chi_{NR}^2 \mathcal{P}_s \mathcal{P}_{p,pr}^2 \sigma_{p,pr}^2 \sigma_{p,pr} \sigma_s}{2\sigma_{p,pr}^2+\sigma_s^2} e^{-\frac{(\omega-2\omega_{p0,pr0}+\omega_{s0})^2}{(2\sigma_{p,pr}^2+\sigma_s^2)}} \tag{15}$$

$$I_{3C}(\omega) \propto \frac{4\pi^2\chi_{NR}^2 |E_{p0,s0}|^4 |E_{pr0}|^2 \sigma_{p,s}^4 \sigma_{pr}^2}{2\sigma_{p,s}^2+\sigma_{pr}^2} e^{-\frac{(\omega-\omega_{pr0})^2}{(2\sigma_{p,s}^2+\sigma_{pr}^2)}} \tag{16}$$

$$\propto \frac{4\sqrt{\pi}\chi_{NR}^2 \mathcal{P}_{p,s}^2 \mathcal{P}_{pr} \sigma_{p,s}^2 \sigma_{pr}}{2\sigma_{p,s}^2+\sigma_{pr}^2} e^{-\frac{(\omega-\omega_{pr0})^2}{(2\sigma_{p,s}^2+\sigma_{pr}^2)}}. \tag{17}$$

Comparing these equations, the 2-colour maximum signal generation occurs at $\omega = \omega_{p0} + \omega_{pr0} - \omega_{s0}$ and under the condition of a significantly broader Stokes source than pump-probe, the maximum CARS signal will fall $\propto 1/\sigma_s$. For the 3-colour excitation case, the maximum signal occurs at frequency $\omega = \omega_{pr}$, and, also, under the case of a relatively narrow probe source and fixed average source powers, the maximum CARS signal will remain constant with increasing pump-Stokes bandwidth. Away from this maximum, with increasing pump-Stokes bandwidth, the CARS signal will rise $\propto \exp\{-(\omega-\omega_{pr0})^2/(2\sigma_{p,s}^2+\sigma_{pr}^2)\}$. To demonstrate these findings and to apply equation (5) to the more general

case of a material with resonant and nonresonant components, we simulated the CARS signal generation process. In particular, the pump, Stokes, and probe source average powers were held fixed, and the probe intensity FWHM was set to 12 cm^{-1}. For each simulation a single Raman peak (A = 1, 2Γ = 20) was simulated +100 cm^{-1} offset from the maximum excitation wavenumber, which for the 3-colour case is 0 cm^{-1} and for 2-colour stimulation is 3,100 cm^{-1}. The nonresonant susceptibility was set to a constant value of 1 ($\max \Im\{\chi^{(3)}\}/\chi_{NR} = 0.1$). Figure S1a shows the normalized intensity of the CARS spectrum (2-colour and 3-colour normalized to peak intensity independently) and a variable broadband source. As the theory suggests, with increasing bandwidth, the 3-colour regime continues to gain spectral breadth with the response at each spectral component continually rising. Conversely, the 2-colour component dramatically decreases in intensity. Figures S1b and S1c show the evolution of the resonant component with increasing bandwidth. In the 3-colour excitation case, the resonant stimulation rises with increasing bandwidth, as detailed in Figure S1d. In the 2-colour case, the peak intensity at 3,100 cm^{-1} rises as the excitation profile broadens to encompass the entire Raman lineshape. As the excitation profile surpasses this width, the intensity begins to fall, as plotted in Figure S1e. In both 2-colour and 3-colour excitation, the spectral resolution, a convolution of the probe source and the Raman nonlinear susceptibility, is the same: $\approx$ 33 cm^{-1}. In addition to the resonant signal, Figures S1d and S1e describe the strong NRB generation that is $> 100\times$ more intense than the resonant case. In both 2-colour and 3-colour excitation regimes, the NRB and resonant signal generation follow similar evolution as a function of pump-Stokes source bandwidth. In fact, as shown in Figure S1f, the nonresonant-to-resonant signal ratios are fixed (standard deviation $\pm$ 0.03 %) with sufficiently broad excitation sources ($>$ 160 cm^{-1}). It should be noted, however, that this ratio can be modulated by changing the probe bandwidth. Finally, if we look at the integrated total signal (resonant and nonresonant contributions), which is proportional to total signal power generated, as shown in Figure S1g, the 2-colour CARS signal is relatively constant with source bandwidths approximately $>$ 160 cm^{-1}. 3-colour excitation, on the other hand, generates a linearly increasing total signal with increasing source bandwidth. Although it may seem odd that there is an increase in signal with increased bandwidth and that energy is not conserved; in actuality there is just an increase in efficiency. This occurs only in the limit of no excitation source depletion and no vibrational and electronic state saturation.

To experimentally evaluate some of these simulated and theoretical points, including the counterintuitive behavior of the 3-colour signal increase with increased continuum bandwidth, we experimentally measured the BCARS spectrum from a glass microscope slide, which is predominantly NRB. To adjust the bandwidth, a slit was inserted into the prism compressor of the SC source and adjusted to reduce or increase available bandwidth (the slit was inserted on the red-side of the SC spectrum; thus, as bandwidth changes, so too does ω_{so}). A thin-film metal-on-glass attenuator was used to adjust the average power of the SC source to maintain constant average power (6.7 mW), regardless of bandwidth. Figure S2a shows the recorded BCARS spectra at various SC bandwidths (FWHM). As predicted, with increasing bandwidth, the Raman spectral region excited with 3-colour excitation increases universally with increasing source bandwidth. Figure S2b shows the evolution of signal intensity as a function of source bandwidth at 3 particular wavenumbers. The measured response qualitatively agrees with the simulated results in Figure S1d. Within the 2-colour region, the signal evolves with many similarities to the previously presented simulations. Figure S2c shows the signal generation at 2 individual frequencies. For 2,450 cm^{-1}, which is stimulated under all bandwidth settings, the signal decays with increasing bandwidth. The 2,800 cm^{-1} signal, on the other hand, is not stimulated under narrow source bandwidth conditions; thus, as the slit is opened, the signal rises, plateaus, then decays. Figure S2d shows the total integrated signal as a function of bandwidth. The 2-colour signal is relatively flat and is similar to the findings in Figure S1g, but does show some increase with increasing bandwidth. Possible explanations are SC source chirp and higher-order dispersion, which were not incorporated into the previous simulations, and source spectral profiles being other than Gaussian. The 3-colour signal, conversely, shows a dramatic increase with increasing bandwidth, which also agrees with theoretical and simulated findings. The exact shape of the rise deviates from predictions, and may, too, be partially accounted for by unsimulated SC dispersion and non-Gaussian source spectral envelope shapes. Overall, the experimental results qualitatively agree with the simulations and clearly demonstrate the theoretically predicted trends.

In this supplementary section, we described several similarities and differences between 2-colour and 3-colour excitation. Most importantly we showed that with increasing bandwidth (under the condition of a fixed probe source bandwidth, and fixed average source powers), 3-colour excitation response rises, and 2-colour excitation signal will eventually decay. Additionally, the theory and simulated results indicated that for excitation sources with stimulation profiles fully encompassing a particular Raman band, the probe source determines the spectral resolution, with the use of an infinitely narrow probe source generating a resonant signal with the same bandwidth as the Raman lineshape. In view of these two characteristics, if the probe source bandwidth, however, were not fixed but varied with the pump and Stokes source bandwidths, as described in early theoretical CARS treatments[2], the resonant and nonresonant signals would evolve differently. Figure S3a shows the total 2-colour and 3-colour generated signal as a function of source bandwidths when the pump, Stokes, and probe sources have the same bandwidth under fixed average source powers. Under these conditions with increasing bandwidths, the resonant signal increases but at a decreasing rate, whereas the

nonresonant signal increases at an increasing rate. Additionally, if one compares the nonresonant-to-resonant signal ration, as shown in Figure S3b, the ratio is continually increasing with source bandwidths. Although the generation of NRB is not necessarily an impediment to recovering the resonant signal in spectroscopic CARS techniques, the increasing probe bandwidth will increasingly deteriorate the spectral resolution, smearing out the Raman Lorentzian lineshapes.

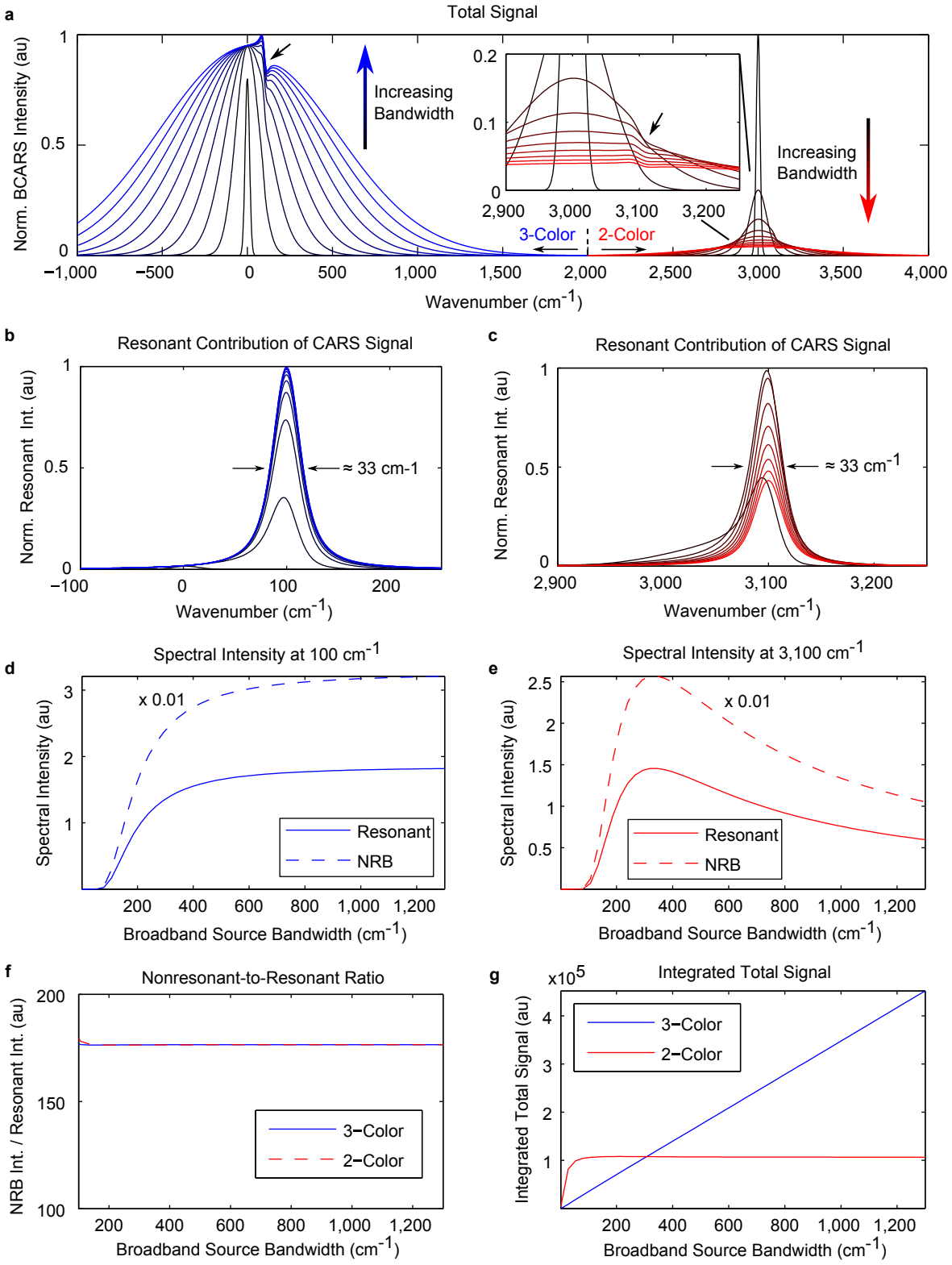

Figure S1: **Simulated effect of SC source bandwidth on NRB generation with a narrowband probe source. a,** Normalized BCARS spectrum from a nonresonant material with 2-colour and 3-colour excitation mechanisms with varying broadband source bandwidth and fixed average power. **b,** Resonant component signal from 3-colour excitation. **c,** Resonant component signal from 2-colour excitation. **d,** BCARS signal frequency response for 3-colour excitation. **e,** BCARS signal frequency response for 2-colour excitation. **f,** NRB-to-resonant signal component ratio. **g,** Integrated total signal.

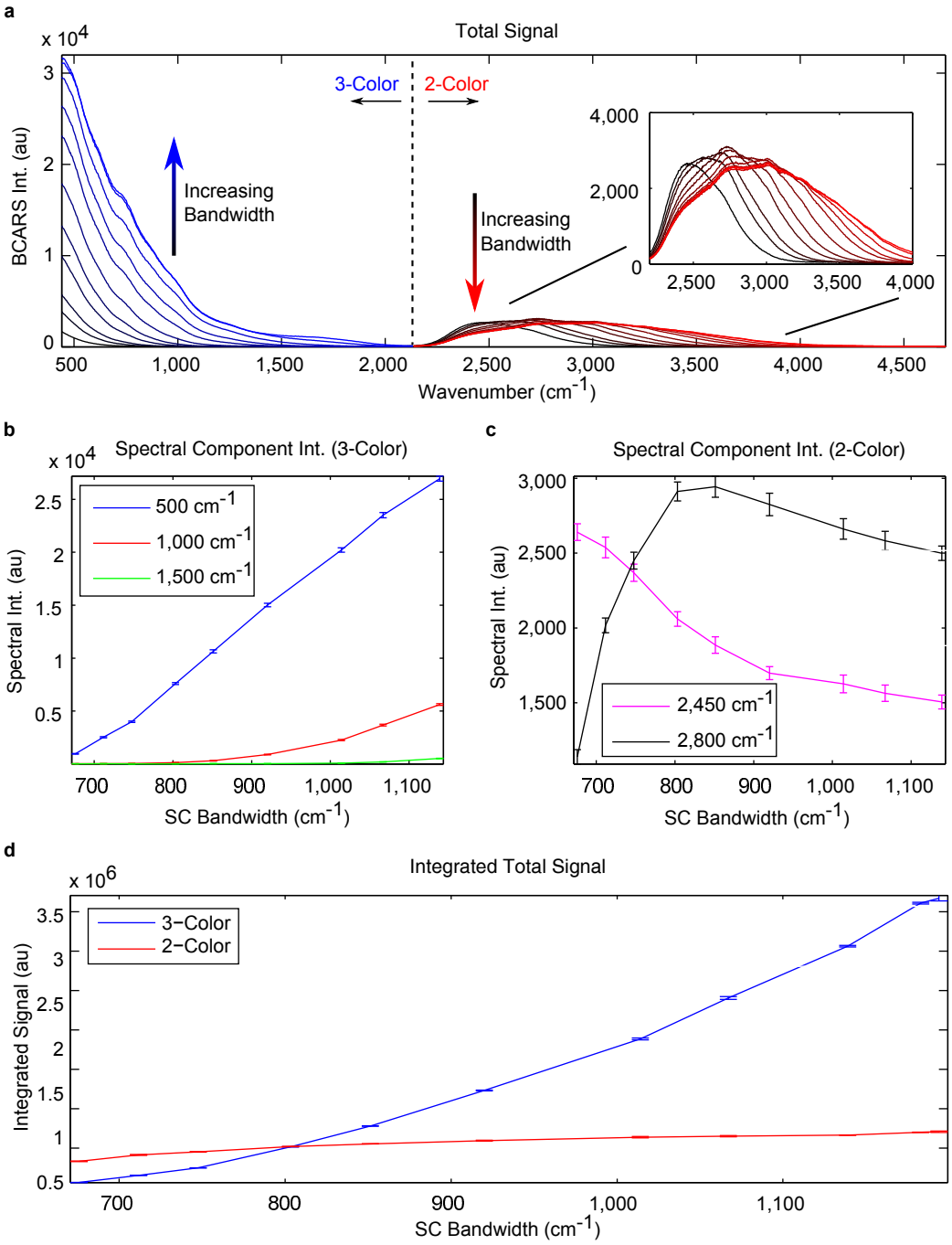

Figure S2: **Experimentally measured effect of SC source bandwidth on NRB generation. a,** BCARS spectrum of glass slide (primarily NRB) with different SC bandwidths. **b,** BCARS intensity as a function of SC bandwidth for 3-colour excitation. **c,** BCARS intensity as a function of SC bandwidth for 2-colour excitation. **d,** Integrated spectral intensity as a function of SC bandwidth. **b-d,** Error bars show ±1 standard deviation.

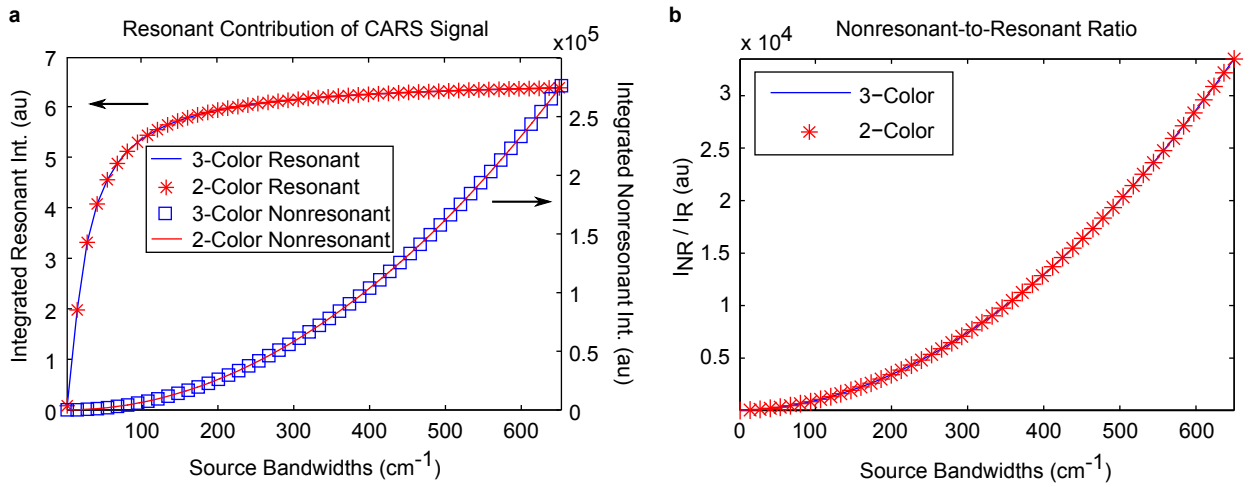

Figure S3: **Simulated CARS signal generation with laser sources of equal bandwidths. a,** Total integrated CARS signal as a function of pump/Stokes/probe source bandwidths. **b,** Evolution of nonresonant-to-resonant signal component ratio with increasing source bandwidths.

II. THE NONRESONANT BACKGROUND (NRB) AS HETERODYNE AMPLIFIER

The NRB has often been viewed as an impediment to acquiring resonant CARS signals. In this section, we will theoretically describe and demonstrate with simulations that the NRB can benefit the resonant components of the CARS signal by heterodyne amplification of weak resonant signals above the noise floor of a detector.

For demonstrative purposes, we will consider a hypothetical BCARS system probing a single complex Lorentzian peak:

$$I_{CARS}(\omega) = |\chi_R(\omega)|^2 + |\chi_{NR}(\omega)|^2 + 2\chi_{NR}(\omega)\Re\{\chi_R(\omega)\} \tag{18}$$

$$\chi_R(\omega) = \frac{A}{\omega - \Omega - i\Gamma} \tag{19}$$

$$\chi_{NR}(\omega) \approx \chi_{NR} \in \mathbb{R}. \tag{20}$$

The CARS signal intensity, I_{CARS}, is proportional to the squared modulus of the total third-order nonlinear susceptibility convolved with the probe source field, but for simplicity we have assumed a delta function-like probe source and a real, constant value for the cross-correlation of the pump and Stokes sources so that we can encapsulate the pump and Stokes source intensities into the nonlinear susceptibility amplitudes. With these assumptions we define the signal-to-noise ratio (SNR) as:

$$\begin{aligned} SNR(\omega) &= \frac{|\chi_R(\omega)|^2 + 2\chi_{NR}(\omega)\Re\{\chi_R(\omega)\}}{\sqrt{|\chi_R(\omega)|^2 + |\chi_{NR}(\omega)|^2 + 2\chi_{NR}(\omega)\Re\{\chi_R(\omega)\} + N_R^2}} \\ &= \frac{[A^2 + 2\chi_{NR}A(\omega - \Omega)]/[(\omega - \Omega)^2 + \Gamma^2]}{\sqrt{[A^2 + 2\chi_{NR}A(\omega - \Omega)]/[(\omega - \Omega)^2 + \Gamma^2] + \chi_{NR}^2 + N_R^2}}, \end{aligned} \tag{21}$$

where N_R is the read-out noise of the detector (which includes all non-time, non-signal dependent noises, such as read noise), and the dark noise is assumed negligible as it is typically orders of magnitude smaller than readout noise and shot noise with modern cooled-CCD cameras over short acquisition times.

The relative amplitude of the resonant and nonresonant components strongly influences the role that the NRB has on the SNR. Under the condition that the resonant component of the signal is significantly larger than the nonresonant component, the maximum SNR is:

$$SNR(\omega = \Omega) = \frac{A^2/\Gamma^2}{\sqrt{[A^2/\Gamma^2 + \chi_{NR}^2 + N_R^2}} \tag{22}$$

$$\lim_{A^2/\Gamma^2 \to \infty} SNR(\omega = \Gamma) = \frac{A}{\Gamma}; \tag{23}$$

thus, the nonresonant component only contributes noise to the signal (i.e., the SNR falls with increasing nonresonant contribution). Practically, this condition may be met in spectral regions of high oscillator density such as within the CH-stretch region of the Raman spectrum, where the resonant component is extremely strong. In other spectral regions, the nonresonant component has a different effect on SNR. For $|\chi_R|^2 << \chi_{NR}^2$ and at the maximum of the dispersed Raman lineshape ($\omega = \Gamma + \Omega$), the SNR is given as:

$$SNR(\omega = \Gamma + \Omega) = \frac{2\chi_{NR}A\Gamma/2\Gamma^2}{\sqrt{2\chi_{NR}A\Gamma/2\Gamma^2 + \chi_{NR}^2 + N_R^2}} = \frac{\chi_{NR}A/\Gamma}{\sqrt{\chi_{NR}A/\Gamma + \chi_{NR}^2 + N_R^2}} \tag{24}$$

$$\lim_{\chi_{NR}^2 \to \infty} SNR(\omega = \Gamma + \Omega) = \frac{A}{\Gamma}. \tag{25}$$

This relationship demonstrates that a large NRB may amplify the signal above the readout noise level when necessary. Additionally, in both the large resonant signal limit and the small resonant signal limit, the SNR asymptotically approaches A / Γ. Figure S4 shows the calculated SNR for a system with A / Γ = 2 for (a) a low-noise detector (similar to many cooled scientific CCD cameras) and (b) a higher-noise detector. In both cases, the SNR with no nonresonant contribution is below one, but with an increasing nonresonant-to-resonant ratio, the maximum SNR approaches A / Γ; ergo, the NRB does not affect signal quality. Under these conditions, capturing the CARS signal requires a relatively large NRB. This is further demonstrated in Figure S5 for a system with A / Γ = 5 with a noisy detector. Figure S5a shows the calculated SNR plot beginning with an SNR below 1 but increasing towards 5. Figures S5b-d show the CARS spectrum for a single simulated peak with noise contributions from the shot noise (both resonant and nonresonant contributions) and readout noise. In S5b, the case of no NRB, the SNR is ≈ 0.25; thus, the signal is completely masked. In the other extreme with the NRB being 100 times larger than the maximum resonant contribution, the SNR approaches 5 and the signal is easily discernible.

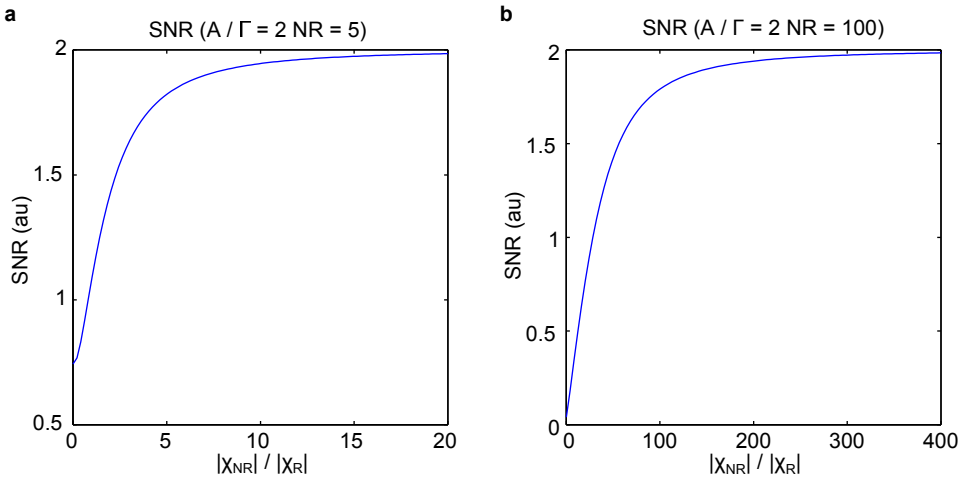

Figure S4: **SNR improvement with increasing NRB. a,** SNR evolution as a function of the nonresonant-to-resonant contribution ratio with a low-noise detector ($N_R = 5$) and ,**b**, a high-noise detector ($N_R = 100$).

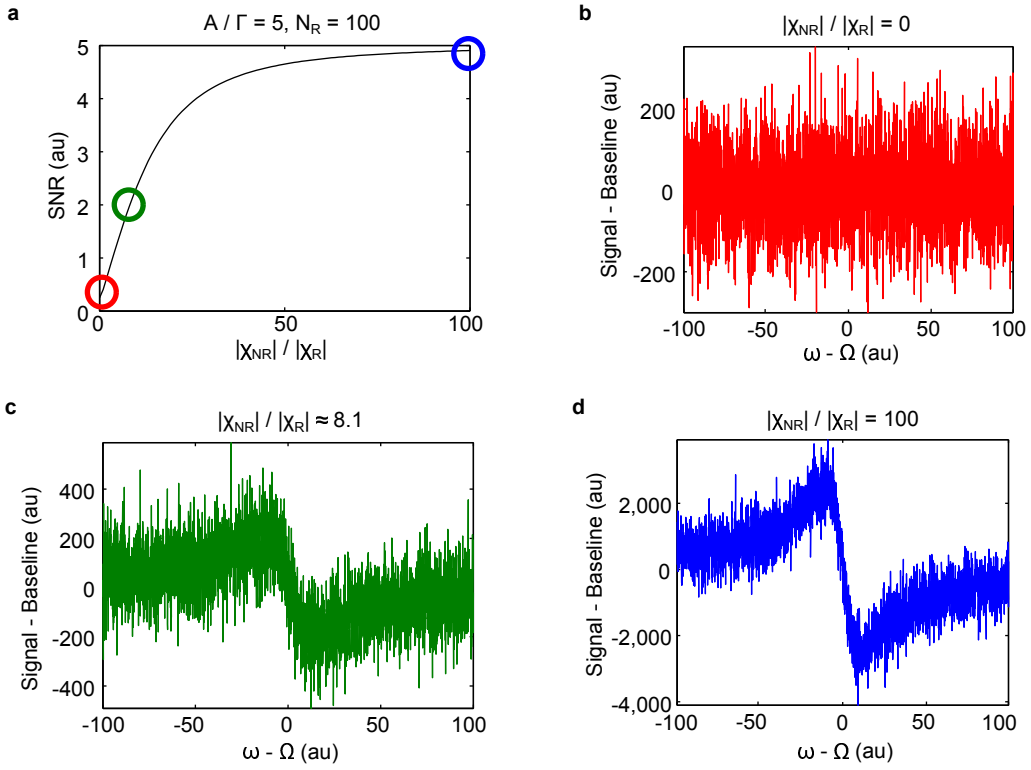

Figure S5: **NRB heterodyne amplification of signal spectrum. a,** SNR evolution as a function of the nonresonant-to-resonant contribution ratio with a high-noise detector ($N_R = 100$). The red, green, and blue circles denote the operational conditions for the simulations presented in sub-plots **b**, **c**, and **d**, respectively. **b,** Simulated spectra with $|\chi_{NR}|/|\chi_R| = 0$, $SNR \approx 0.25$. **c,** Simulated spectra with $|\chi_{NR}|/|\chi_R| \approx 8.1$, $SNR \approx 2$. **d,** Simulated spectra with $|\chi_{NR}|/|\chi_R| = 100$, $SNR \approx 4.9$.

III. SPONTANEOUS AND COHERENT RAMAN SPECTROSCOPY OF GLYCEROL

To compare the speed and sensitivity of the presented BCARS system to traditional spontaneous Raman spectroscopy, we recorded spectra of 99% glycerol (Sigma) using a Renshaw inVia confocal Raman microscope and the BCARS system. Although this demonstration is qualitatively instructive, a quantitative, thorough comparison between coherent and spontaneous Raman systems requires detailed calibration and spectral measurement of both instruments, and requires an intricate knowledge of each molecule under investigation– with particular attention to electronic characteristics and polarization response[12]. For spontaneous Raman acquisition, a large drop of glycerol was placed on a silicon wafer to reduce autofluorescence, and the sample was illuminated with 36 mW (on-sample) with 785 nm light from a photodiode. The Stokes photons were epi-detected (reflection mode) and recorded by the built-in CCD-equipped spectrometer (CCD: ∼40 - 60% quantum efficiency [QE] within fingerprint region; ∼5 - 20 % QE within CH-/OH-stretch region). Spectra were recorded with 200 ms (the minimum available on the system) or 3.5 s integration times. For BCARS spectroscopy, the glycerol was mounted using typical histological preparation, with the sample sealed between a glass coverslip and a glass slide. As described in the main text, the spectra were captured with 3.5 ms integration times by the CCD-spectrometer (QE ∼45% across the spectrum), and the Raman spectrum was retrieved using the time-domain Kramers-Kronig (TDKK) transform (a modified Hilbert transform). Figure S6 shows the fingerprint spectra collected with spontaneous Raman and coherent Raman techniques. Visual comparison indicates that most of the BCARS spectrum (<1,2000 cm^{-1}) shows similar noise performance to that of spontaneous Raman with 3.5 s integration times (a speed factor of 1000x). As the BCARS system response diminishes at higher energies, the signal quality degrades, but not to the levels seen with spontaneous Raman at 200 ms integration times. It should be noted that the spontaneous Raman system was incapable of recording fingerprint spectra from a glycerol sample mounted in glass (i.e., the sample used for BCARS analysis) due to the autofluorescence from the glass. Although for neat samples this may not be problematic, for histological investigations, this need be a special consideration when mounting specimens.

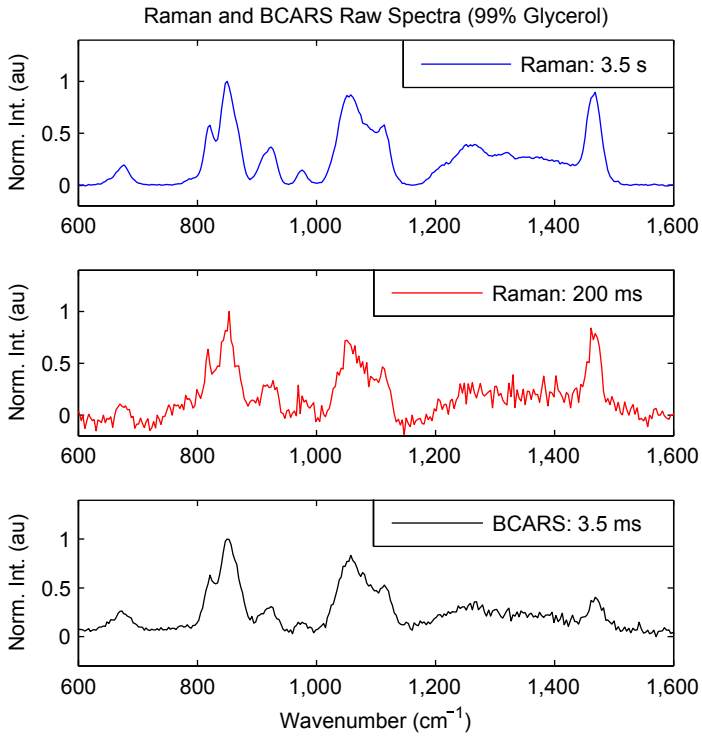

Figure S6: **Spontaneous and Coherent Raman Fingerprint Spectra of Glycerol.** Spontaneous Raman spectrum of 99% glycerol with 3.5 s (blue, top) and 200 ms (middle, red) integration times. Retrieved Raman spectrum using BCARS with 3.5 ms integration time (bottom, black).

Performing the same experiment within the longer CH-/OH-stretch Raman region shows an even more dramatic separation between the techniques (see Figure S7): the BCARS sample contains clearly less noise than the spontaneous Raman spectra– even at 3.5 s dwell times. A contributing factor to this is the low QE of the silicon CCD detector

at longer wavelengths of the inVia system. Additionally, as the raw BCARS signal intensity is proportional to the molecular concentration, N, as $N^2 + N$, there is a tremendous response from strong, dense molecular oscillators, such as those found within the CH-/OH-stretch region.

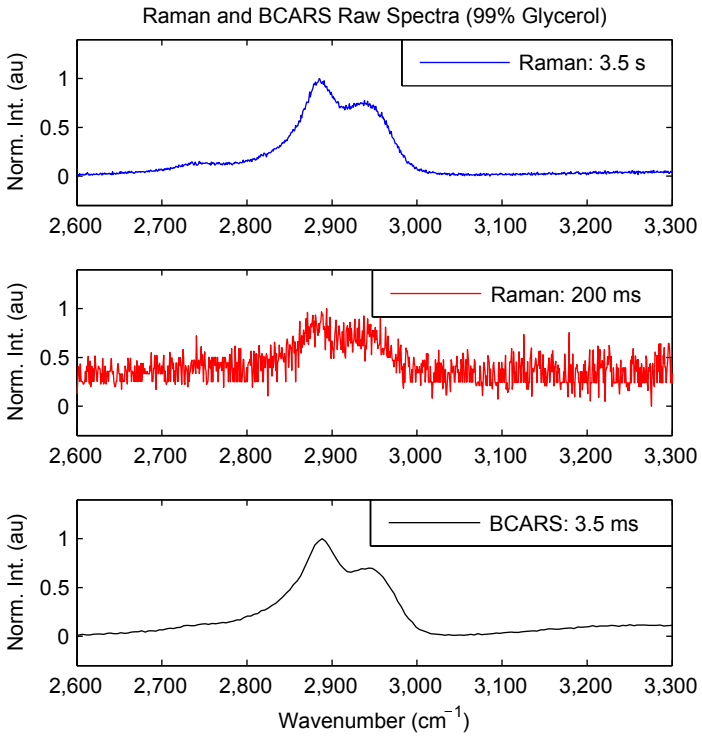

Figure S7: **Spontaneous and Coherent Raman CH-/OH-Stretch Spectra of Glycerol.** Spontaneous Raman spectrum of 99% glycerol with 3.5 s (blue, top) and 200 ms (middle, red) integration times. Retrieved Raman spectrum using BCARS with 3.5 ms integration time (bottom, black).

From 100 collected spectra under each experimental condition, we calculated the signal-to-noise ratio (SNR) at each Raman energy level as shown in Figure S8. Across the fingerprint region, the BCARS system provides an SNR greater than that collected with spontaneous Raman at 200 ms (a speed factor of 57×). At the lowest wavenumbers, the BCARS response is similar to that of spontaneous Raman at 3.5 s integration times (a speed factor of 1000×). Of course, slight differences in recorded peak intensities (due to polarization and system response affects, as previously mentioned) will modulate the exact SNR calculated, but these results indicate that the developed BCARS system is capable of high-speed acquisition with high SNR.

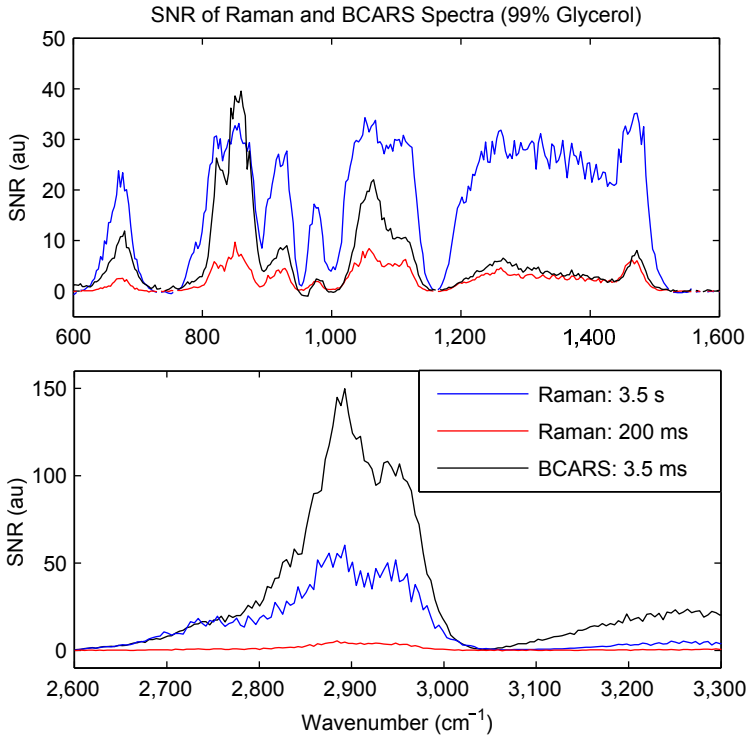

Figure S8: **SNR of Spontaneous and Coherent Raman Spectra of Glycerol.** SNR calculated from 100 spontaneous Raman spectra of 99% glycerol with 3.5 s (blue, top) and 200 ms (middle, red) integration times. SNR from 100 retrieved Raman spectra using BCARS with 3.5 ms integration time (bottom, black).

[1] Lee, Y. J., Liu, Y. & Cicerone, M. T. Characterization of three-color CARS in a two-pulse broadband CARS spectrum. *Opt. Lett.* **32**, 3370–3372 (2007).
[2] Cheng, J. X., Volkmer, A., Book, L. D. & Xie, X. S. An epi-detected coherent anti-Stokes Raman scattering (E-CARS) microscope with high spectral resolution and high sensitivity. *J. Phys. Chem. B* **105**, 1277–1280 (2001).
[3] Knutsen, K. P., Messer, B. M., Onorato, R. M. & Saykally, R. J. Chirped coherent anti-stokes Raman scattering for high spectral resolution spectroscopy and chemically selective imaging. *J. Phys. Chem. B* **110**, 5854–5864 (2006).
[4] Andresen, E. R., Paulsen, H. N., Birkedal, V., Thørgersen, J. & Keiding, S. R. Broadband multiplex coherent anti-Stokes Raman scattering microscopy employing photonic-crystal fibers. *J. Opt. Soc. Am. B* **22**, 1934–1938 (2005).
[5] Pestov, D. *et al.* Pulse shaping for mode-selective ultrafast coherent Raman spectroscopy of highly scattering solids. *J. Opt. Soc. Am. B* **25**, 768–772 (2008).
[6] Cheng, J. X., Volkmer, A., Book, L. D. & Xie, X. S. Multiplex coherent anti-Stokes Raman scattering microspectroscopy and study of lipid vesicles. *J. Phys. Chem. B* **106**, 8493–8498 (2002).
[7] Cheng, J. X., Book, L. D. & Xie, X. S. Polarization coherent anti-Stokes Raman scattering microscopy. *Opt. Lett.* **26**, 1341–1343 (2001).
[8] Oron, D., Dudovich, N. & Silberberg, Y. Femtosecond phase-and-polarization control for background-free coherent anti-Stokes Raman spectroscopy. *Phys. Rev. Lett.* **90**, 1–4 (2003).
[9] Dudovich, N., Oron, D. & Silberberg, Y. Single-pulse coherently controlled nonlinear Raman spectroscopy and microscopy. *Nature* **418**, 512–514 (2002).
[10] Gomez, J. S. Coherent Raman Spectroscopy. In *Modern Techniques in Raman Spectroscopy* (ed. Laserna, J. J.) 305–342 (John Wiley & Sons, Chichester, England, 1996).
[11] Hellwarth, R. Third-order optical susceptibilities of liquids and solids. *Prog. Quant. Electron.* **5**, 1–68 (1979).
[12] Mendelovici, E., Frost, R. L. & Kloprogge, T. Cryogenic Raman spectroscopy of glycerol. *J. Raman Spectrosc.* **1126**, 1121–1126 (2000).

Additional information

Any mention of commercial products or services is for experimental clarity and does not signify an endorsement or recommendation by the National Institute of Standards and Technology.