The final publication is available at

https://doi.org/10.1126/science.abk3106

Additional Information

Continuous-Wave Frequency Upconversion with a Molecular Optomechanical Nanocavity

Wen Chen,[1] Philippe Roelli,[1†], Huatian Hu,[2] Sachin Verlekar,[1]
Sakthi Priya Amirtharaj,[1] Angela I. Barreda, [3] Tobias J. Kippenberg,[1]
Miroslavna Kovylina,[4] Ewold Verhagen,[5] Alejandro Martínez,[4]
Christophe Galland[1*]

[1]Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland

[2]Hubei Key Laboratory of Optical Information and Pattern Recognition
Wuhan Institute of Technology, Wuhan 430205, China

[3]Institute of Applied Physics, Abbe Center of Photonics
Friedrich Schiller University Jena, Albert-Einstein-Str. 15, 07745 Jena, Germany

[4]Nanophotonics Technology Center, Universitat Politècnica de València
Camino de Vera s/n, 46022 Valencia, Spain

[†]Present address: Nano-optics group, CIC nanoGUNE BRTA, San Sebastián, Spain

[5]Center for Nanophotonics, AMOLF, Science Park 104, 1098 XG Amsterdam, Netherlands

[*]To whom correspondence should be addressed; E-mail: chris.galland@epfl.ch

Frequency upconversion is a cornerstone of electromagnetic signal processing, analysis and detection. It is used to transfer energy and information from one frequency domain to another where transmission, modulation or detection is technically easier or more efficient. Optomechanical transduction is emerging as a flexible approach to coherent frequency upconversion; it has been successfully demonstrated for conversion from radio- and microwaves (kHZ to GHz) to optical fields. Nevertheless, optomechanical transduction of multi-THz and mid-infrared signals remains an open challenge. Here, we utilize molecular cavity optomechanics to demonstrate upconversion of sub-microwatt continuous-wave signals at $\sim$32 THz into the visible domain at ambient conditions. The device consists in a plasmonic nanocavity hosting a small number of molecules. The incoming field resonantly drives a collective molecular vibration, which imprints an optomechanical modulation on a visible pump laser and results in Stokes and anti-Stokes upconverted Raman sidebands with sub-natural linewidth, indicating a coherent process. The nanocavity offers 13 orders of magnitude enhancement of upconversion efficiency per molecule compared to free space, with a measured phonon-to-photon internal conversion efficiency larger than 10^{-4} per milliwatt of pump power. Our results establish a flexible paradigm for optomechanical frequency conversion using molecular oscillators coupled to plasmonic nanocavities, whose vibrational and electromagnetic properties can be tailored at will using chemical engineering and nanofabrication.

Introduction   Our ability to generate, detect and analyse electromagnetic signals spanning the full spectrum from radiofrequency up to visible light or even X-rays governs technological progress in areas ranging from information processing, telecommunication networks, material characterisation, broadband imaging and molecular sensing. In this broad context, frequency conversion plays an essential role by allowing detection and modulation of signals that are in a frequency band where suitable technologies are not widely accessible. A prominent example is that of electromagnetic waves with frequencies in the range from $\sim$1 to $\sim$100 THz (*1*), which we will refer to here as infrared (IR) waves in the broad sense. They find applications ranging from homeland security and molecular analysis of gases, chemicals and biological tissues (*2, 3*), to thermal imaging and non-destructive material inspection (*4*), to astronomical surveys (*5*). Yet, they represent a technological frontier for signal detection and processing, with a scarce offer of commercial devices in comparison with telecommunication or visible bands. Spectroscopy of IR signals is typically performed using tunable laser sources (*6*) or broadband sources coupled to a Fourier-transform IR (FTIR) spectrometer. These methods require bulky apparatuses or expensive coherent sources, and are typically limited to slow acquisition rates. Moreover, they employ IR detectors (*7*) that are more noisy than visible and near-infrared (VIS/NIR) detectors, motivating recent efforts to perform FTIR using NIR light with a nonlinear interferometer (*8*).

As a powerful and flexible approach to analyze long-wavelength electromagnetic signals, frequency upconversion is being intensely studied for imaging and spectroscopy applications (*9*). The main motivation for frequency upconversion in the context of spectroscopy is to leverage inexpensive, fast and low-noise detector technologies available in the VIS/NIR domain (*10*). This method has the additional advantage of allowing, in principle, coherent mapping of a quantum state from one frequency to another, with broad applications in classical and quantum information technologies. It offers a way to connect different information processing nodes over long distances via optical fibers (*11–15*). Frequency upconversion of IR signals can be accomplished by three-wave mixing in a bulk crystal with a large effective second order nonlinearity. Delicate phase matching, large pump powers and cm-long crystals are typically needed to reach high efficiencies (*16–20*), even though plasmonic gap modes have recently been leveraged to dramatically enhance nonlinear effects (*21–24*). A second-order nonlinearity is also present at interfaces, where inversion symmetry is broken. This fact has been broadly leveraged in ultrafast nonlinear spectroscopy to probe the properties and dynamics of molecular layers on various surfaces (*25, 26*) with applications extending to biologically relevant membranes (*27*). In this context, resonant excitation of vibrational modes is responsible for molecule-specific signatures

3

in the three-wave mixing signal, which offers deep insights into molecular structure, dynamics, and molecule-surface interactions (*25, 26*). However, such techniques require substantial peak powers only accessible with femto- or picosecond pulses and have therefore remained spectroscopic tools used by an expert community.

Optomechanical cavities have recently emerged as promising candidates to realise quantum coherent frequency conversion (*14, 15, 28*). In this implementation, the signal of interest resonantly drives a mechanical oscillator, itself parametrically coupled to a laser-driven optical cavity, which results in modulation sidebands at the sum- and difference-frequencies (called anti-Stokes and Stokes sidebands, respectively). This approach offers a number of advantages, such as the resonant enhancement of nonlinear response at the mechanical frequency and the parametric enhancement of conversion efficiency with intracavity pump power. It is also highly versatile and has been demonstrated with mechanical resonance frequencies ranging from kHz (*29, 30*) to GHz (*31–37*). In a different approach, modulations on THz waves have been read out optically through a MHz-frequency mechanical resonator (*38*). Molecular oscillators constitute a new frontier in cavity optomechanics (*39, 40*), as they enable to reach multi-THz resonant frequencies and room-temperature quantum coherent operation (*41*). Moreover, they can be coupled to plasmonic nanocavities with deep sub-wavelength mode volumes, thereby enabling optomechanical coupling rates in excess of 1 THz (*42*). Despite these promising features, devices based on molecular cavity optomechanics have yet to be demonstrated.

Here, we experimentally demonstrate the upconversion of a 32 THz (9.3 μm wavelength) continuous wave IR signal into the visible domain using a sub-wavelength molecular optomechanical cavity at ambient conditions. Our upconversion scheme operates with micro-Watt level continuous wave signal and pump beams, which is a disruptive departure from alternative methods, and allows high-resolution spectroscopy of the IR signal thanks to the coherent nature of optomechanical transduction. This novel regime of operation is achieved by coupling a monolayer of molecular oscillators to a doubly resonant plasmonic gap nanocavity, which supports deep-sub-wavelength mode volumes and simulated field enhancement factors in excess of 500 and 100 at IR and VIS frequencies, respectively, from which an overall enhancement of upconversion efficiency per molecule by more than 13 orders of magnitude compared to free space is predicted and experimentally confirmed. In our proof-of-principle experiment, internal optomechanical upconversion efficiency per unit of pump power are on the order of 10^{-4} per milliwatt with an estimated 10% outcoupling efficiency, while the IR photon-to-vibration conversion efficiency is on the order of 10^{-5}, corresponding to end-to-end conversion efficiencies above 10^{-12}
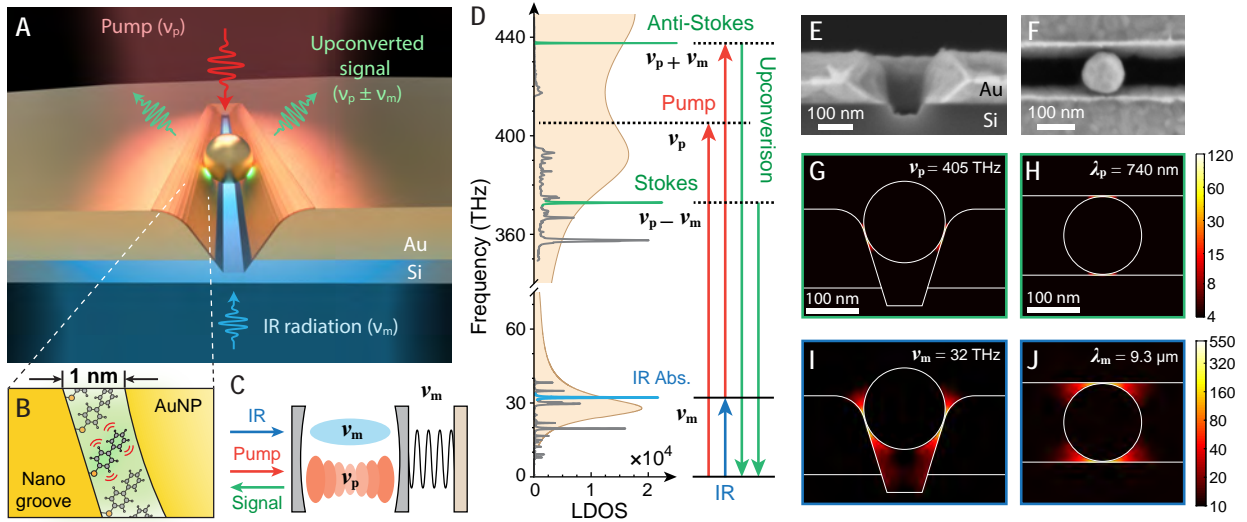
4

Figure 1: **Molecular optomechanical upconversion concept**. (**A**) Illustration of a nanoparticle-in-groove molecular optomechanical cavity which confines both IR (frequency ν_m) and visible (frequency ν_p) fields into the ~1-nm-thick molecular layer (**B**). (**C**) Molecular vibrations couple resonantly to the IR field because they change the molecular dipole moment, while the change in the molecule's polarizability leads to a parametric coupling to the visible field, realising an optomechanical cavity. (**D**) Resonant vibrational levels and electromagnetic fields involved in the upconversion together with the computed radiative local density of state (LDOS) inside the nanocavity, along with the measured Raman scattering and simulated IR absorption (cf. Sec. 1.5) of the molecules. (**E**) Cross sectional view of scanning electron microscope (SEM) images of a fabricated nano-groove. (**F**) Top view SEM image of a nanoparticle-in-nanogroove. (**G-J**) Simulated electromagnetic field enhancement factor for incident plane waves at $\nu_p = 405$ THz (740 nm, **G,H**) and $\nu_m = 32$ THz (9.3 μm, **I,J**), both polarized orthogonal to the groove.

for 10 μW pump power; we present several routes to further improve these figures.

Device description   The idea of a molecular optomechanical platform for upconversion was first proposed in (*39*) and its theoretical performance was analysed in (*10*), showing the feasibility of single-photon sensitivity at frequencies down to few THz even at ambient conditions. In contrast to plasmomechanical resonators that have been realised e.g. with dimer antennas built on a nanobeam oscillator (*43*) or with nanoparticles trapped in a plasmonic hotspot (*44*), the mechanical resonator in a molecular optomechanical cavity consists in a collective molecular vibration, which is parametrically coupled to the nanocavity through its Raman polarizability (*39*). This approach allows reaching mechanical frequencies in the 1–100 THz range and optomechanical coupling rates above 1 THz (*42*).

Our experiment is conceptually presented in Fig. 1A-D. To obtain extreme field confinement in nm-scale regions as well as large enhancement factors at all frequencies involved in the upconversion, we develop a new type of plasmonic nanocavity that we dub the nanoparticle-in-groove (Fig. 1E-F). It consists of a single Au nanoparticle (150 nm nominal diameter) placed inside a nano-groove etched in a Au film and covered by a monolayer of biphenyl-4-thiol (BPhT) acting as spacer and molecular oscillator (Fig. 1B-C). Our structure is engineered to support colocalized plasmonic resonances at IR ($\sim$30 THz) and visible frequencies, which can be excited under normal incidence illumination (more details are given in SM Sec. 1.4 and Fig. S9). The IR resonance frequency is governed by the length of the nano-groove (*45, 46*), which was chosen as 2 μm to match a vibrational mode of BPhT molecules at 32 THz with large IR and Raman cross sections (see SM Sec. 1.5). Two broad resonances in the visible domain are supported by the nanoparticle coupled to the walls of the Au nano-groove, analogous to the nanoparticle-on-mirror geometry (*47, 48*). Near fields are confined in the nanometer-wide gaps formed by the molecular layer (see Fig. 1G-J). This design provides enhancement factors for IR absorption and Raman scattering up to $\times 10^5$ and 10^8, respectively (estimated from the second and fourth power of the maximal field enhancements). Overall, our geometry is expected to boost the conversion efficiency by a factor on the order of 10^{13} for the molecules located in the region of highest field, compared to free space.

Results   When a pump laser tuned at 405 THz (740 nm) is focused on the sample (Fig. 2A), the parametric interaction with the molecular vibrations generates Raman sidebands at lower (Stokes) and higher (anti-Stokes) frequencies (Fig. 2B). The giant enhancement factor described above makes it possible to detect the Raman signal from few hundreds molecules (*49*), as estimated from the nanocavity physical size and molecular layer density (*50*) (see SM). Without IR beam incident on the device (black solid line in Fig. 2B), the Stokes signal is dominated by spontaneous emission of phonon-photon pairs while the anti-Stokes signal originates from the upconversion of thermal vibrations (*51*). At room temperature ($T = 25°C$), the thermal occupancy of the vibrational mode at $\nu_m = 32$ THz is $n_{\text{th}} = \left(\exp\left(\frac{h\nu_m}{k_B T} \right) - 1 \right)^{-1} \simeq 5.8 \times 10^{-3}$, where h and k_B are Planck's and Boltzman's constants.

When the IR beam from a quantum cascade laser (QCL) is focused through the Si substrate onto the backside of the device (Fig. 2A), we observe the amplification of a single resonant peak in the Stokes and anti-Stokes sidebands (green line in Fig. 2B), whose linewidth is much narrower than the natural linewidth of spontaneous Raman scattering (Fig. 2D,E). The latter
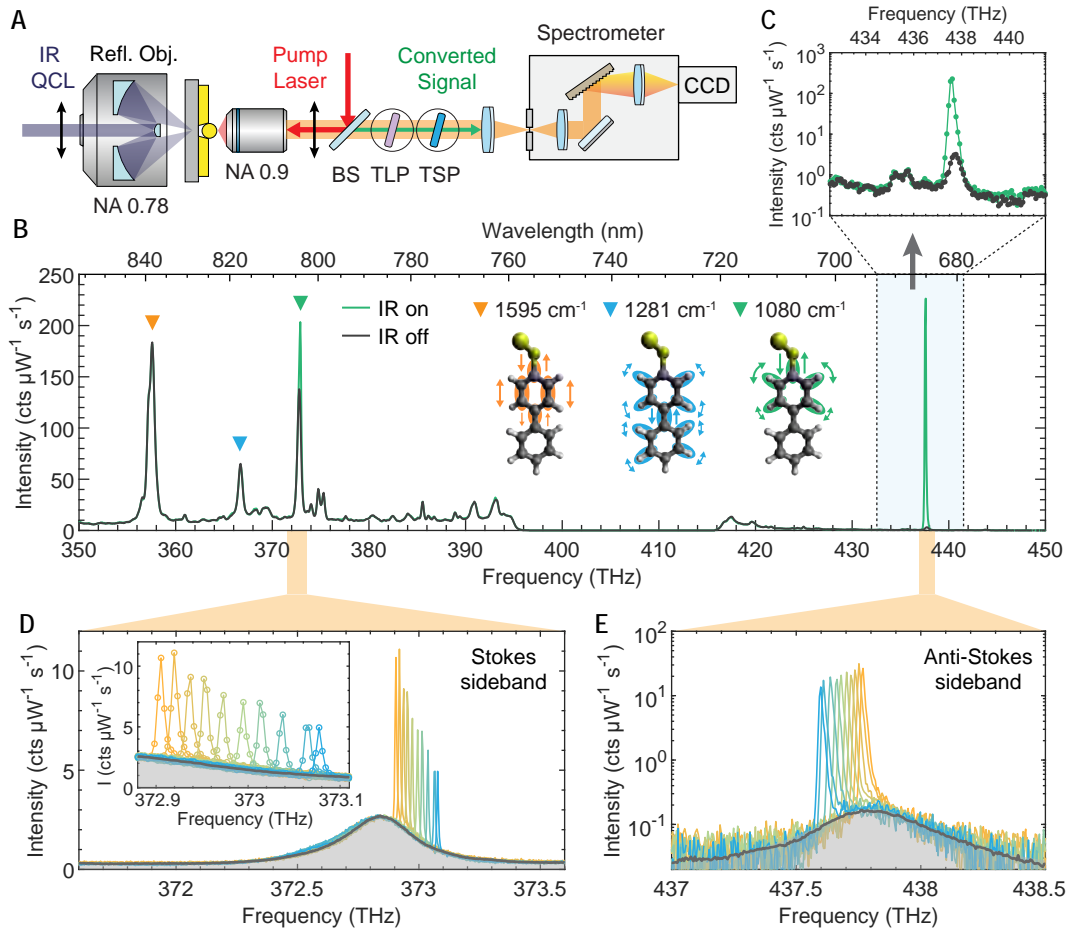
Figure 2: **Molecular optomechanical transduction from 32 THz to the visible domain.** (**A**) Schematic of the measurement configuration. A reflective objective with numerical aperture (NA) 0.78 focuses the IR beam through the silicon substrate, while a refractive objective (NA 0.9) focuses the visible pump beam and collects the Raman signal, which is directed to a spectrometer after blocking the pump light with tunable short- or long-pass filters (TSP, TLP). The polarization of both pump and IR radiation is perpendicular to the nanogroove (black arrows) for effective excitation of the plasmonic resonances. BS: beam splitter. (**B**) Low-resolution, full-range Raman spectra from a single nanocavity under 10 μW pump power (740 nm wavelength) without (black line) and with (green line) incoming IR radiation (530 μW) resonant with the vibrational mode at 1080 cm^{-1} or about 32 THz. The inset shows the displacements of typical Raman-active modes labelled with orange, blue and green triangles. Acquisition time, 10 s. (**C**) Closer view on the anti-Stokes sideband of panel **B** plotted on a logarithmic vertical scale. (**D,E**) High resolution (0.23 cm^{-1}, 7 GHz) Stokes (**D**) and anti-Stokes (**E**) spectra observed when tuning the signal beam (range limited by the QCL capabilities), all normalized to 175 μW incoming IR power (pump: 10 μW). Acquisition time for, 10 s. The grey line in **D** and **E** are the spectra without IR radiation, where the grey areas are the corresponding Lorentzian fits to the spontaneous Stokes and anti-Stokes scattering peaks. The inset of **D** shows the enlarged view of the resolution-limited lineshape of the upconverted signal (see also Figs. S6,S7).

is likely dominated by homogeneous broadening, as the peak is well fit by a Lorentzian (see Sec. 1.7 and . S7 in SM). Altogether, these observations are compatible with a coherent upconversion process. As we tune the frequency of the incoming field the upconverted signal shifts accordingly, and its measured linewidth is found to be limited by that of our spectrometer, or about 7 GHz (cf. Fig. S7). The relative conversion efficiency vs. detuning is plotted in Fig. S7 and confirms that upconversion is assisted by the vibrational mode.

We interpret these results as the manifestation of optomechanical transduction, where a collective molecular vibrational mode is resonantly driven by the nanocavity-enhanced 32 THz incoming field. We can describe the resulting vibrational state by a displaced thermal state with a coherent amplitude α and a corresponding mean phonon number $n_{\mathrm{coh}} = |\alpha|^2$. This oscillation, which is coherent with the IR drive, is mapped onto the Raman sidebands of the pump laser, where the IR signal can be analysed and detected using a standard optical spectrometer, camera or single photon counting module. A lower bound for n_{coh} is estimated from the ratio of the Raman peak areas with vs. without IR drive. It is a lower bound as not all molecules that contribute to the spontaneous Raman peak do also contribute to upconversion, due to the imperfect overlap of IR and VIS plasmonic modes. For most nanocavities, we find IR-driven phonon numbers $n_{\mathrm{coh}} > 0.1$ for 500 to 600 μW IR power, equivalent to $\frac{n_{\mathrm{coh}}}{n_{\mathrm{th}}} > 10$ (cf. Table S2). This figure is compatible with an enhancement of IR absorption cross-section per molecule by more than 5 orders of magnitude, as simulated in Fig. 1I,J and detailed in Sec. 1.6 in SM, corresponding to IR photon to vibration conversion efficiency above 10^{-5}.

Next, we characterise the power dependence of the upconverted signal, by varying first the power of the pump laser at 740 nm (Fig. 3A). We observe a linear dependence of the upconverted Raman signal, in adequation with the expected parametric upconversion process. Moreover, in the measurement conditions of Fig. 2, the spontaneous (no IR drive) Stokes and anti-Stokes signals also grow linearly with pump power. An optomechanical description of plasmon-enhanced Raman scattering (*10, 39, 52*) predicts three main regimes for the pump power dependence of the anti-Stokes signal: (i) at low power a linear regime dominated by thermal noise, then (ii) a quadratic regime where the vibrational population increases linearly with laser power due to quantum backaction (also called vibrational pumping) (*40, 42, 53*), and finally (iii) a phonon-stimulated regime is expected, dominated by dynamical backaction amplification of the vibration (*39*). The data in Fig. 3A are compatible with regime (i) above, i.e. we do not observe a superlinear dependence of the anti-Stokes signal on laser power in the range used in the upconversion experiment (blue shaded area). Fig. 3B shows how the upconverted signal scales with
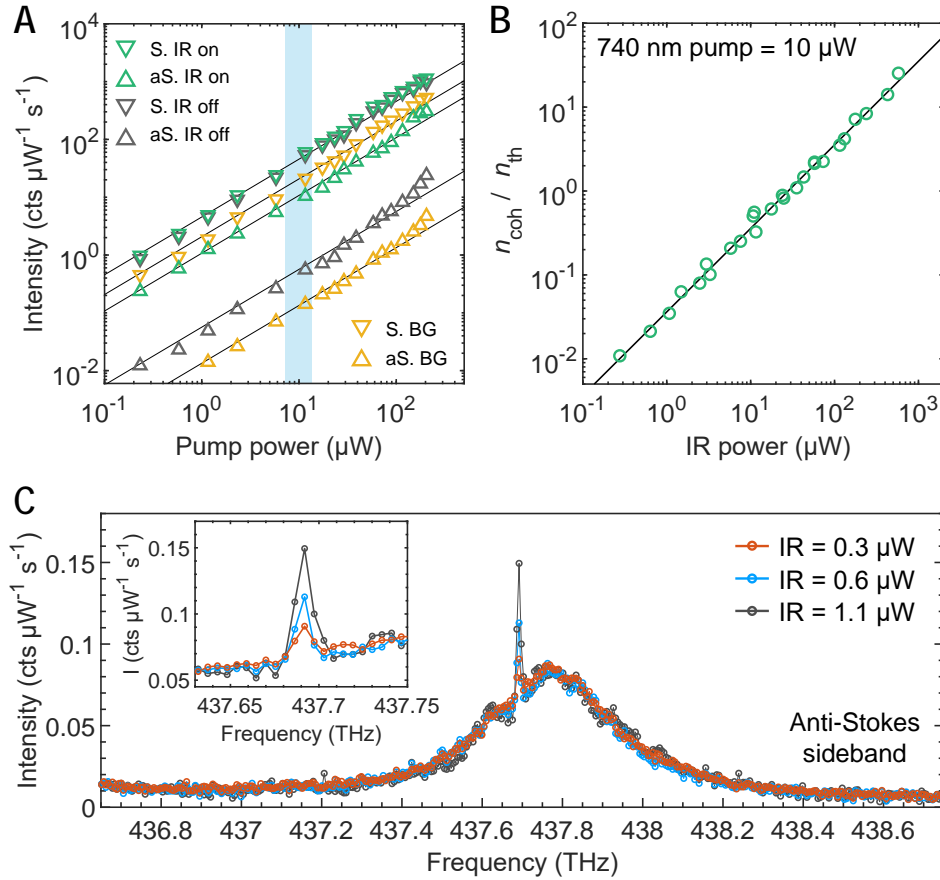
Figure 3: **Dependence of upconverted signal on mid-IR and pump powers.** (**A**) Measured Raman signal as a function of pump power (tuned at 740 nm). Triangles pointing up (resp. down) correspond to anti-Stokes (resp. Stokes) sideband. Green symbols correspond to an incident IR power of 600 μW, while grey symbols are taken without IR drive. The spectrally broad background underlying the Raman peaks is shown with orange symbols. Black lines are linear fits for reference. The blue shaded area denotes the power range used in typical upconversion measurements presented here. (**B**) Ratio of IR-driven (n_{coh}) to thermal (n_{th}) phonon occupancies as a function of the IR radiation power on the nanocavity (measured with 10 μW pump power, tuned at 740 nm). The black line shows a linear fit. (**C**) Detection of sub-μW IR signals using high-resolution spectroscopy on the anti-Stokes sideband.

incoming mid-IR power, with results compatible with a linear dependence, as expected for a resonant drive below saturation.

Finally, we quantify the external IR to visible conversion efficiency of our device by spectrally filtering the anti-Stokes sideband and sending it to a single photon counting module, with independently calibrated detection efficiency (results are summarised in Table. S2). Taking into account the calibrated detection efficiency, we infer that the upconverted anti-Stokes photon rate collected by the objective reaches up to 200 kHz per nanocavity for 600 μW incident IR power (corresponding to $n_{\mathrm{IR}} \simeq 2.8 \times 10^{16}$ photons per seconds) and 10 μW pump power. This measurement yields a conversion efficiency from incoming IR photon to outgoing visible photon collected by our objective on the order of 10^{-12} (see Fig S4), or 10^{-10} mW^{-1} as normalised per unit pump power. We note that the only 6% of the power emitted in the near field is collected by our objective, as detailed in Fig. S9E, so that the internal efficiency is at least 10 times larger. Despite the low efficiency, the coherent nature of the process allows us to reliably detect incoming IR powers down to few hundreds nanowatts, as shown in Fig. 3C – a figure which would further improve with the resolution of the spectrometer.

Discussion To conclude this Report, we discuss a few avenues for future research and development in molecular optomechanical frequency conversion. A possible route for efficiency improvement consists of increasing the number of molecules contributing to optomechanical interaction without sacrificing on field enhancement, which could be achieved by increasing the number of nanoparticles in a singe nano-groove, optimising the shape and facet size of the nanoparticles, increasing the density of grooves and possibly leveraging lattice resonances. In this spirit, new dual-resonant metamaterials (*54–56*) with nanocavities as unit cells could be developed to enhance IR absorption (*57*) and upconversion efficiency. We note that even with a single nanocavity, internal conversion efficiencies approaching 0.1 were predicted in Ref. (*10*) by operating at higher pump powers under pulsed excitation (*58*), allowing for single THz/mid-IR photon detection under pulsed pumping.

We also predict that leveraging strong light-matter coupling to boost the IR absorption and/or Raman scattering will become a particularly fruitful area of research for upconversion applications. In the vibrational strong coupling regime and under critical cavity input coupling, most absorbed radiation can be transfered to vibrational polaritons (*59*), so that even single-photon-level signals could be efficiently upconverted (*10*). Conversely, there has been recent indications that the formation of exciton-polaritons in nanocavities similar to the one used here

10

can further enhance the resonant Raman cross-section (*60*), offering a way to improve the internal upconversion efficiency. In addition to plasmonic resonance engineering, molecular engineering (*61*) also offers promising perspectives for improving IR and Raman cross-section, controlling the conversion bandwidth, and addressing a broader range of frequencies down to the THz domain. Optical phonons in 2D materials are other candidates for reaching lower frequencies, while two-photon driving of a vibration could be explored to extend operation to frequencies lower than that of the vibrational resonance (*62, 63*).

For a better fundamental understanding of degree of coherence and added noise in the transduction process, measurements of first- and second-order correlation functions as in e.g. (*36*) can be performed. Besides, broader tuning of the IR and pump tones and higher resolution spectroscopy should reveal new information about the molecule-nanocavity interaction and clarify possible non-resonant contributions to upconversion related to the electronic response of metal and molecules (*25, 26*). Looking ahead, the sub-wavelength size of our device will allow for the construction of chip-scale pixel arrays, spectrometers and hyperspectral imagers for THz and IR wavelengths, with the possibility to excite and detect the upconverted signals through dielectric waveguides (*64*). And since the transduction mechanism is bidirectional, it can be used for coherent downconversion as well, allowing optical generation of long-wave radiation assisted by molecular systems (*65, 66*).

References

1. M. Tonouchi, *Nature Photonics* **1**, 97 (2007).

2. S. D. Bruyne, M. M. Speeckaert, J. R. Delanghe, *Critical Reviews in Clinical Laboratory Sciences* **55**, 1 (2018).

3. Y. Bai, J. Yin, J.-X. Cheng, *Science Advances* **7**, eabg1559 (2021).

4. F. Ciampa, P. Mahmoodi, F. Pinto, M. Meo, *Sensors* **18**, 609 (2018).

5. T. L. Roellig, *et al.*, *Journal of Astronomical Telescopes, Instruments, and Systems* **6**, 041503 (2020).

6. F. K. Tittel, D. Richter, A. Fried, *Solid-State Mid-Infrared Laser Sources*, I. T. Sorokina, K. L. Vodopyanov, eds., Topics in Applied Physics (Springer, Berlin, Heidelberg, 2003), pp. 458–529.

7. A. Rogalski, *Infrared and Terahertz Detectors* (CRC Press, Boca Raton, 2019), third edn.

8. C. Lindner, *et al.*, *Optics Express* **29**, 4035 (2021).

9. A. Barh, P. J. Rodrigo, L. Meng, C. Pedersen, P. Tidemand-Lichtenberg, *Advances in Optics and Photonics* **11**, 952 (2019).

10. P. Roelli, D. Martin-Cano, T. J. Kippenberg, C. Galland, *Physical Review X* **10**, 031057 (2020).

11. A. G. Radnaev, *et al.*, *Nature Physics* **6**, 894 (2010).

12. S. Zaske, *et al.*, *Physical Review Letters* **109**, 147404 (2012).

13. A. P. Higginbotham, *et al.*, *Nature Physics* **14**, 1038 (2018).

14. N. Lauk, *et al.*, *Quantum Science and Technology* **5**, 020501 (2020).

15. N. J. Lambert, A. Rueda, F. Sedlmeir, H. G. L. Schwefel, *Advanced Quantum Technologies* **3**, 1900077 (2020).

16. K. Karstad, *et al.*, *Optics and Lasers in Engineering* **43**, 537 (2005).

17. P. Tidemand-Lichtenberg, J. S. Dam, H. V. Andersen, L. Høgstedt, C. Pedersen, *JOSA B* **33**, D28 (2016).

18. S. Junaid, *et al.*, *Optics Express* **26**, 2203 (2018).

19. Y.-P. Tseng, C. Pedersen, P. Tidemand-Lichtenberg, *Optical Materials Express* **8**, 1313 (2018).

20. C. Pedersen, A. A. S, P. Tidemand-Lichtenberg, *Smart Photonic and Optoelectronic Integrated Circuits XXII* (International Society for Optics and Photonics, 2020), vol. 11284, p. 112841A.

21. M. P. Nielsen, X. Shi, P. Dichtl, S. A. Maier, R. F. Oulton, *Science* **358**, 1179 (2017).

22. S. Ummethala, *et al.*, *Nature Photonics* **13**, 519 (2019).

23. Y. Salamin, *et al.*, *Nature Communications* **10**, 5550 (2019).

24. R. Gordon, *Nanotechnology* **30**, 212001 (2019).

25. Y. R. Shen, *Nature* **337**, 519 (1989).

26. C. S. Tian, Y. R. Shen, *Surface Science Reports* **69**, 105 (2014).

27. S. Roke, G. Gonella, *Annual Review of Physical Chemistry* **63**, 353 (2012).

28. Y. Chu, S. Gröblacher, *Applied Physics Letters* **117**, 150503 (2020).

29. T. Bagci, *et al.*, *Nature* **507**, 81 (2014).

30. I. Moaddel Haghighi, N. Malossi, R. Natali, G. Di Giuseppe, D. Vitali, *Physical Review Applied* **9**, 034031 (2018).

31. J. Bochmann, A. Vainsencher, D. D. Awschalom, A. N. Cleland, *Nature Physics* **9**, 712 (2013).

32. R. W. Andrews, *et al.*, *Nature Physics* **10**, 321 (2014).

33. A. Vainsencher, K. J. Satzinger, G. A. Peairs, A. N. Cleland, *Applied Physics Letters* **109**, 033107 (2016).

34. K. C. Balram, M. I. Davanço, J. D. Song, K. Srinivasan, *Nature Photonics* **10**, 346 (2016).

35. R. Van Laer, R. N. Patel, T. P. McKenna, J. D. Witmer, A. H. Safavi-Naeini, *APL Photonics* **3**, 086102 (2018).

36. M. Forsch, *et al.*, *Nature Physics* **16**, 69 (2020).

37. M. Mirhosseini, A. Sipahigil, M. Kalaee, O. Painter, *Nature* **588**, 599 (2020).

38. C. Belacel, *et al.*, *Nature Communications* **8**, 1578 (2017).

39. P. Roelli, C. Galland, N. Piro, T. J. Kippenberg, *Nature Nanotechnology* **11**, 164 (2016).

40. M. K. Schmidt, R. Esteban, F. Benz, J. J. Baumberg, J. Aizpurua, *Faraday Discussions* (2017).

41. S. T. Velez, V. Sudhir, N. Sangouard, C. Galland, *Science Advances* **6**, eabb0260 (2020).

42. F. Benz, *et al.*, *Science* **354**, 726 (2016).

43. R. Thijssen, T. J. Kippenberg, A. Polman, E. Verhagen, *Nano Letters* **15**, 3971 (2015).

44. P. Mestres, J. Berthelot, S. S. Aćimović, R. Quidant, *Light: Science & Applications* **5**, e16092 (2016).

45. F. J. Garcia-Vidal, L. Martin-Moreno, T. W. Ebbesen, L. Kuipers, *Reviews of Modern Physics* **82**, 729 (2010).

46. C. Huck, *et al.*, *ACS Photonics* **2**, 1489 (2015).

47. C. Ciracì, *et al.*, *Science* **337**, 1072 (2012).

48. J. J. Baumberg, J. Aizpurua, M. H. Mikkelsen, D. R. Smith, *Nature Materials* **18**, 668 (2019).

49. J. Langer, *et al.*, *ACS Nano* **14**, 28 (2020).

50. A. Ahmed, *et al.*, *ACS Photonics* **8**, 1863 (2021).

51. S. T. Velez, *et al.*, *Physical Review X* **9**, 041007 (2019).

52. M. K. Schmidt, R. Esteban, A. González-Tudela, G. Giedke, J. Aizpurua, *ACS Nano* **10**, 6291 (2016).

53. K. Kneipp, *et al.*, *Phys. Rev. Lett.* **76**, 2444 (1996).

54. H. Wang, J. Kundu, N. J. Halas, *Angewandte Chemie International Edition* **46**, 9040 (2007).

55. F. Le, *et al.*, *ACS Nano* **2**, 707 (2008).

56. N. S. Mueller, *et al.*, *ACS Nano* **15**, 5523 (2021).

57. Y. Yao, *et al.*, *Nano Letters* **14**, 6526 (2014).

58. A. Lombardi, *et al.*, *Physical Review X* **8**, 011016 (2018).

59. A. Shalabney, *et al.*, *Nature Communications* **6**, 5981 (2015).

60. X. Liu, *et al.*, *Nature Materials* pp. 1–6 (2021).

61. Y. Shao, *et al.*, *Physical Chemistry Chemical Physics* **8**, 3172 (2006).

62. S. Maehrlein, A. Paarmann, M. Wolf, T. Kampfrath, *Physical Review Letters* **119**, 127402 (2017).

63. A. J. Traverso, *et al.*, *Physical Review Letters* **120**, 063602 (2018).

64. J. Losada, *et al.*, *IEEE Journal of Selected Topics in Quantum Electronics* **25**, 1 (2019).

65. M. P. Hasselbeck, L. A. Schlie, D. Stalnaker, *Applied Physics Letters* **85**, 173 (2004).

66. J. del Pino, F. J. Garcia-Vidal, J. Feist, *Physical Review Letters* **117**, 277401 (2016).

67. P. B. Johnson, R. W. Christy, *Physical Review B* **6**, 4370 (1972).

68. R. L. Olmon, *et al.*, *Physical Review B* **86**, 235147 (2012).

69. M. A. Green, *Solar Energy Materials and Solar Cells* **92**, 1305 (2008).

70. J. Yang, J.-P. Hugonin, P. Lalanne, *ACS Photonics* **3**, 395 (2016).

71. E. C. L. Ru, P. G. Etchegoin, *Principles of Surface-Enhanced Raman Spectroscopy*, E. C. L. Ru, P. G. Etchegoin, eds. (Elsevier, Amsterdam, 2009), pp. 465–490.

72. Y. Zhang, J. Aizpurua, R. Esteban, *ACS Photonics* **7**, 1676 (2020).

73. P. Marian, T. A. Marian, *Physical Review A* **47**, 4474 (1993).

74. C. Artur, E. C. Le Ru, P. G. Etchegoin, *The Journal of Physical Chemistry Letters* **2**, 3002 (2011).

75. S. Yampolsky, *et al.*, *Nature Photonics* **8**, 650 (2014).

Acknowledgments

Supplementary materials

Supplementary materials

1 Materials and Methods

1.1 Nanocavity fabrication

The device where fabricated using a double polished silicon wafer (380-nm-thick) as the substrate. A 2-nm-thick Cr adhesion layer and a 150-nm-thick gold layer were thermally evaporated on the Si at a rate of 0.5 nm/s. An array of 150 nm wide and 2 μm long nano-grooves with unit cell 4×5 μm^2 was patterned by electron-beam lithography using poly methyl methacrylate (PMMA) as photoresist. After photoresist development, the the nano-grooves were etched by angled collimated ion-beam, resulting in the cross-section shown in Fig. 1E. The sample was then immersed in piranha solution (H$_2$SO$_4$: H$_2$O$_2$ = 3:1) for 5 min to remove PMMA, which also annealed the Au to smooth the surface. The wafer with nano-grooves was subsequently diced into chips. The single chips were cleaned with ozone for 30 min under ultraviolet light illumination and then put into ethanol solution overnight to remove Au oxide. The sample was then incubated into a 1 mMol solution of Biphenylthiol (BPhT, Sigma-Aldrich) in ethanol for 24 h at ambient conditions to form a self-assembled monolayer (*50*), followed by 5 rounds of ethanol rinsing and drying by nitrogen gas flow. Finally, a drop of aqueous solution containing 150 nm diameter gold nanoparticle (BBI solution) was deposited on the nano-groove region. After 10 min incubation, the liquid was removed by nitrogen gas flow and the sample was then rinsed by de-ionized water and dried by the gas flow. The nanoparticles were randomly distributed on the Au film and inside the grooves. We found that nanoparticles at any relative position inside a groove enables frequency upconversion (see the inset of Fig. S1A, and also the statistics in Table S2).

1.2 Optical setup and data acquisition

We built a custom transmission microscope as schematised in Fig. 2A and detailed in Fig. S2. The sample was at ambient temperature in air, mounted on closed-loop XYZ piezo-positioners (AttoCube). As shown in the inset of Fig. S1A, single nanoparticles inside nano-grooves can be observed directly under bright field white light illumination.

For the excitation of Raman scattering, the pump laser was focused by a refractive objective (Olympus, max. numerical aperture NA 0.9) from the side of the gold film. We used two types

of lasers as the pump light: (i) A continuous-wave Ti:Sapph laser (Model 3900 SpectraPhysics) tuned around 785 nm, where we can use a notch filter at 785 nm (Thorlabs) in detection to detect both Raman sidebands simultaneously; (ii) a C-WAVE OPO from Hubner Photonics tuned around 740 nm, where the spectra of Stokes and anti-Stokes sidebands were taken sequentially by introducing tunable long-pass and short-pass filters (Semrock), respectively. While the Ti:Sapph laser features a measured linewidth of 20 GHz or more with spectral jitter leading to a Gaussian lineshape, the C-WAVE has a specified linewidth of 1 MHz. Using the C-WAVE line we find that the resolution of our spectrometer is at best 7 GHz, defined as the measured FWHM (Fig. S7I), so that it is the limiting factor in the measurements of Fig. 2D-E.

Using a broadband half-wave plate, the polarization of the pump light was tuned perpendicular to the groove's long-axis to effectively excite the plasmonic modes in the visible domain (Fig. 2A). Without additional specification, the power of both lasers for all the measurement were kept at the level of ~ 10 μW on the sample to avoid slow nanocavity modifications and other nonlinearities that we observe above ~ 10 μW (see Fig. 3A). Reflected and scattered fields were collected through the same refractive objective, and then spectrally filtered to isolate Stokes and anti-Stokes Raman sidebands, using a notch or two different edge filters (see above), before being focused on the entrance slit of a spectrometer, or alternatively coupled to a telecom fiber and routed to super-conducting nanowire single photon counting modules held in a closed-cycle cryostat below 2 K. We use two spectrometers with 3 different configurations: a Kymera 193i (grating 600 line/mm) and Shamrock 750 (grating 150 line/mm) for broad spectral range and lower resolution (as in Fig. 2B-C and Fig. S3) and the Shamrock 750 with 1800 line/mm grating for higher resolution (used in Fig. 2D,E and Fig. S5B,C).

A single mode quantum cascade laser (QCL, Alpes Laser) provides the coherent IR signal for the upconversion measurements. Its specified tuning range is from ~ 1071 to ~ 1077 cm^{-1} (see detailed performance measured by upconversion in Fig. S8), close to the center of the Raman peak used in upconversion at ~ 1080 cm^{-1} (we recall that 1 cm$^{-1} \simeq 30$ GHz). The QCL is collimated by a ZnSe lens and then focused through the silicon substrate (approximately 30% transmission) on a nanocavity with a reflective objective (PIKE technologies, max. NA 0.78). To excite the IR cavity mode, the orientation of the nano-grooves is set perpendicular to the QCL polarisation. The alignment of the IR beam can be adjusted by observing the IR light reflected from the Au film on Si, where a part of the IR light was directed to a thermal camera (Seek Thermal) using an IR beam splitter (Thorlabs), allowing to check the beam quality (cf. Fig. S2). The IR beam size on the sample is ~ 10 μm.

Control measurements were also performed on (i) the nearby nanoparticle-on-mirror system outside a groove (inset of Fig. S1) and (ii) the bare nano-groove without nanoparticle; but also on a nanoparticle-in-nanogroove with either (iii) IR radiation or (iv) pump light polarised parallel to the long-axis of the groove. In all these cases, no upconverted signal could be observed, confirming the essential role of the overlapped plasmonic visible and IR resonances for efficient optomechanical transduction.

IR power dependence The QCL power can be tuned from ~3.6 mW to ~60 μW (the lower limit of the bolometer) before the reflective objective by cooperatively varying its working temperature and drive current (see Fig. S8). Given that the transmission through the objective and the Si substrate in our system are 55% and 30%, respectively, the actual IR power on the nanocavity can be adjusted in this way between ~600 μW and ~10 μW. Combing this tunability with a series of IR neutral density filters (Thorlabs), the lowest power in our measurement was ~0.3 μW (Fig. 3C). The IR power dependent spectra were taken using high resolution spectrometer on the anti-Stokes band excited by 740 nm pump light. We used acquisition times ranging from 10 to 600 s, with the upper limit set by the mechanical drift of the stage holding the sample.

IR frequency scanning For the measurements in high-resolution mode shown in Fig. 2D,E and S5B,C, the IR radiation was tuned from ~1071 to ~1077 cm^{-1} (32.1 to 32.3 THz) while the IR power is in the range from ~80 to ~200 μW on the nanocavity. The acquisition time for the Raman signal with IR radiation is 10s. The acquisition time for Anti-stokes and Stokes signal without IR radiation are 10 min and 5 min, respectively. The fluctuation of the raw spectra were offset based on keeping the spontaneous peak the same height. The differences in IR powers as a function of frequency was corrected by the following treatment: (i) the narrow peak from the net upconverted signal and the broad peak from the spontaneous Raman were separated mathematically; (ii) the narrow peak intensities for different QCL frequencies were normalised to ~175 μW by linearly scaling the spectra (given that the IR power dependence is linear, see Fig. 3B); (iii) the normalised narrow peaks were added back to their spontaneous background. From this measurement, we can also plot the relative upconversion efficiency vs. IR frequency, with results following the lineshape of spontaneous Raman scattering (Fig. S6)

Single photon counting measurement. To detect the photon flux of the signal, the anti-Stokes Raman line of interest is further filtered by the same combination of tunable short and

long-pass filters. The filtered beam is coupled into a telecom fiber connected to superconducting nanowire single-photon detectors (SNSPD, Single Quantum). Count rate is measured by a time-to-digital converter (ID900, ID Quantique). In typical experimental conditions ($\sim 10~\mu$W pump power on sample), after accounting for the measured detection efficiency, the thermal anti-Stokes photon rate collected by the objective is about 1 to 4 kHz (collection efficiency corrected, the same below), while the photon rate reaches more than 100 kHz for 600 μW IR power on the sample for the most efficient devices (see Table S2).

1.3 Calibration of external collection and detection efficiency

In order to estimate the experimental conversion efficiency, we must take into account all collection and detection losses in the Raman signal path. For this purpose, we replace the interference filters (TLP, TSP and N785 in Fig. S2) in the detection path by neutral density filters with known optical densities, measure the laser power reflected from the sample, from which we compute the collected photon flux, and compare this photon flux to the count rate on the single photon counting module. This procedure yields a detection efficiency (including fiber coupling) of 10%. Then the reflected laser with known count rate was then directed to the spectrometer to get its spectrum and the counts from the CCD, enabling the conversion from the Raman peak intensity to photon flux. For a low-resolution upconverted anti-Stokes peak, 1 Hz from the CCD at the peak position corresponds to ~40-50 Hz count rate.

Moreover, from the simulations described above, we estimate than only 10% of the power upconverted inside the nanocavity is collected by our objective (the remainder is either radiated in other directions or dissipated in the metal, see Fig. S9).

1.4 Electromagnetic simulations

Electromagnetic simulations were performed with commercial FEM package COMSOL Multiphysics 5.2. The 3D structure contained a 155 nm gold nanoparticle situated on a 2 μm nano-groove whose trapezoid cross-section had a 180 nm top-base and a 65 nm bottom-base. The height of the nano-groove cross-section is 185 nm, including a 35 nm dip in the silicon substrate (determined from SEM images). The nano-groove was carved into a 150 nm thick gold layer which was situated on an infinitely thick silicon substrate. The dielectric function of gold follows the data from Johnson and Christy (optical domain, (67)) and Olmon et al. (IR domain, (68)). The refractive index of the silicon is set as 3.4 in the IR range, and follows Green

et al. in the visible range (*69*).

To calculate the radiative emission and infer the local density of states (LDOS, Fig.S9A, C), an electric dipole μ_0 was placed into the nanogap formed by the molecular layer. The radiated power P_{rad} collected by an objective (visible, NA = 0.9; infared, NA = 0.78) was normalized by the dipole radiation in vacuum P_0 as P_{rad}/P_0. The radiative emission from the dipole in the nanogap P_{rad} was calculated as the surface integral of the Poynting vector in the far field, (*70*) while the dipole radiation in vacuum reads $P_0 = \frac{|\mu_0|^2}{4\pi\varepsilon_0\varepsilon}\frac{n^3\omega^4}{3c^3}$. The ratio of radiated power with the objective's NA to the total vacuum emission is identified as the radiative LDOS enhancement multiplied by collection efficiency η_{coll}, $\rho_{\mathrm{rad}}/\rho_0 = P_{\mathrm{rad}}/P_0 \cdot \eta_{\mathrm{coll}}$. Radiative emission towards the silicon substrate was calculated similarly.

The nonradiative decay is attributed to Ohmic loss (absorption), which was calculated by the total decay subtracted by the radiative decay. Here, the total decay rate was calculated by a closed surface integral only containing the electric dipole in the gap. The antenna radiative efficiency can be obtained as $\eta_{\mathrm{rad}} = P_{\mathrm{rad}}/P_{\mathrm{tot}}$, where P_{tot} is the above-mentioned total dipole radiation in the nanogap.

1.5 Molecular calculations

Raman and infrared cross-sections are calculated with density functional theory (DFT) using GAUSSIAN09 (B3LYP/6-311++G(d,p) base). The choice of molecule adopted for the manuscript's estimations is based on the comparison of the DFT calculated vibrational frequencies of BPhT and BPhT bonded to gold (Au-BPhT) with the experimental results (Fig.S10A). The orientation of the Au-BPhT molecule with respect to the local fields and the linewidth of the vibrational modes are similarly adapted to best match experimental data. The vibrational modes' IR absorption cross-section (σ_{IR}, Fig.S10B) and differential Raman cross-section ($d\sigma_R/d\Omega$, Fig.S10C) are calculated from the derivatives of the polarizability tensor and electric dipole with respect to the displacements of the molecule's atoms, following the procedures described in Refs. (*10, 71*).

1.6 Theoretical predictions of conversion efficiency from calculated IR and Raman cross-sections

In this section we predict the upconverted Raman power P_{up} (or photon flux ϕ_{up}) for a given incoming infrared power P_{IR} and pump power P_0. We first perform the calculation for a single

molecule, and then assume that P_{up} scales linearly with the number of molecules $N = A_{\mathrm{cav}}\rho$, where A_{cav} is the effective nanocavity area and ρ the molecular surface density (*50*). Since the molecules occupy a region much smaller than the pump wavelength, the degree of spatial coherence is not expected to modify this scaling. Moreover, we operate in a pump power regime where collective amplification is negligible (*72*).

The IR-driven phonon occupancy n_{coh} created in excess of the thermal occupancy n_{th} is the ratio of phonon creation rate to decay rate: $n_{\mathrm{coh}} = \frac{\gamma_+}{\gamma_-}$, where each rate is expressed as:

$$\gamma_- = 1/\tau_{\mathrm{m}} \quad \text{and} \quad \gamma_+ = \frac{1}{h\nu_{\mathrm{IR}}}\frac{P_{\mathrm{IR}}}{A_{\mathrm{IR}}}\sigma_{\mathrm{IR}}$$

Here, we have introduced the following quantities:

- τ_{m} is the vibrational $1/e$ energy decay time (lifetime) (Unit: s)

- $h\nu_{\mathrm{IR}}$ is the energy of incoming IR photons and corresponding vibrational quanta (Unit: J)

- A_{IR} is the effective IR spot area (Unit: m^2). We can approximately link it to the numerical aperture for a diffraction limited spot by $A_{\mathrm{IR}} \simeq \pi D^2/4$ where D is the spot FWHM, which at the wavelength λ_{IR} is approximately: $D \simeq \frac{\lambda_{\mathrm{IR}}}{2NA}$

- σ_{IR} is the IR absorption cross section (Unit: m^2)

We can write $n_{\mathrm{coh}} = |\alpha|^2$ where α is the complex amplitude of a coherent state. The state of the collective molecular oscillator is expected to be a displaced thermal state (*73*), in agreement with the spectra from Fig. 2D, E in the main text.

Knowing n_{coh}, we express the corresponding upconverted Raman power:

$$P_{\mathrm{up}} = \frac{P_0}{A_0}\sigma_{\mathrm{R}}n_{\mathrm{coh}}$$

Here, we have introduced the following quantities:

- A_0 is the effective pump laser spot area (Unit: m^2). As above, we can write $A_0 \simeq \pi D^2/4$ with $D \simeq \frac{\lambda_0}{2NA}$

- σ_{R} is the total Raman cross section (Unit: m^2)

From this quantity, we can find the expression of the conversion efficiency η per unit of pump power (Unit: W^{-1}), considering N molecules equally contributing to the upconversion, and integrating the upconverted power in the 4π solid angle:

$$\eta = N\frac{P_{\mathrm{up}}}{P_{\mathrm{IR}}\, P_0} = \frac{N\sigma_{\mathrm{IR}}\sigma_{\mathrm{R}}\,\tau_{\mathrm{m}}}{A_0 A_{\mathrm{IR}}\, h\nu_{\mathrm{IR}}}$$

For an isotropic Raman scattered field, the conversion efficiency when taking into account the collection through the objective's solide angle will be reduced by a factor $(1-\sqrt{1-(NA)^2})/2$, where NA is the numerical aperture (we consider the medium is air).

We can also compute the total upconverted photon rate ϕ_{up} per unit power in the pump and IR inputs (Unit: $s^{-1}W^{-2}$):

$$\phi_{\mathrm{up}} = N\frac{P_{\mathrm{up}}}{h\nu_{\mathrm{up}}\, P_{\mathrm{IR}}\, P_0} = \frac{N\sigma_{\mathrm{IR}}\sigma_{\mathrm{R}}\,\tau_{\mathrm{m}}}{A_0 A_{\mathrm{IR}}\, h\nu_{\mathrm{up}}\, h\nu_{\mathrm{IR}}}$$

These expressions are valid for molecules in free space, or bound to a bare gold film without cavity. In the latter case, the IR and Raman cross-section need to be appropriately modified to account for the effect of the thiol bound to gold, which provides a so-called chemical enhancement through partial charge transfer and changes in molecular orbitals. We take chemical enhancement into account in our molecular simulations (Sec. 1.5) by linking a gold atom to the sulfur atom of BPhT.

For a molecule coupled to a nanocavity, we expect the conversion efficiency to be enhanced by the factor $\tilde{F} = \beta_{\mathrm{IR}}^2\beta_0^2 F_{\mathrm{rad}}$, where the β's are field enhancement factors $|E_{loc}|/|E_{inc}|$ at the incoming IR and pump wavelengths, and F_{rad} is the radiative enhancement for a dipole placed inside the nanocavity, computed as explained in Section 1.4 and taking into account the finite collection efficiency by the objective.

Results of numerical estimates   Table S1 summarises the parameters extracted from molecular DFT and electromagnetic FEM simulations as explained in Sec. 1.5 and 1.4, and corresponding to the experimental setup presented in the main text and in Sec. 1.2. From these parameters, the expected nanocavity-enhanced upconversion efficiency is computed to be $\eta\cdot\tilde{F} \simeq 3\times 10^{-11}$, in reasonable agreement with the experimental efficiency on the order of 10^{-12} measured for typical nanocavities.

Remark   For a fair estimate of the enhancement factor, we should compare our upconversion efficiency to that of a hypothetical molecular monolayer on a flat gold substrate, for which all

23

Parameter	σ_{IR}	σ_{R}	ρ	A_{cav}	N	ν_{IR}	τ_{m}
Unit	cm^2	cm^2	cm^{-2}	cm^2	molecules	THz	ps
Value	7×10^{-18}	5×10^{-29}	5×10^{14}	1.6×10^{-12}	785	32.4	0.53
Parameter	β_{IR}	β_0	F_{rad}	ν_{up}			
Unit	factor	factor	factor	THz			
Value	100	400	1.2×10^4	414			

Supplementary Table S1: Parameter values extracted from simulations and used in the theoretical efficiency estimate. The vibrational lifetime τ_{m} is estimated from the measured spontaneous Raman linewidth of 300 GHz (Lorentzian fit).

molecules overlapping with the pump beam would contribute to upconversion. In this case, $\tilde{F}$ defined above must be multiplied by $N/N_0 = A_n/A_L$, where N_0 is the number of molecules covering the pump laser spot area A_L, while A_n is the effective area of the nanocavity containing the N active molecules. With our experimental parameters, we find $\tilde{F} \simeq 1.9 \times 10^{13}$ and $A_n/A_L \simeq 10^{-3}$, so that our effective upconversion enhancement factor is in excess of 10^{10}.

1.7 Possible intepretations of the IR-driven vs. natural linewidths

In this section, we discuss the interpretation of the resolution-limited linewidth of the upconverted signal as seen in Fig. 2, and show that it most likely reflects a coherently displaced thermal state of the collective molecular oscillator. Central to this discussion is the origin of the spontaneous Raman linewidth $\Delta\nu$ (on the order of 230 to 350 GHz as measured without IR drive), which is related to the vibrational decay time τ_{m} by $\tau_{\mathrm{m}} = 1/2\pi\Delta\nu$ and to the vibrational coherence time $\tau_{\mathrm{m,coh}}$ by $\tau_{\mathrm{m,coh}} = 1/\pi\Delta\nu = 2\tau_{\mathrm{m}}$. Two extreme physical situations can *a priori* be encountered (with any intermediate situation being possible):

(i) The spectral linewidth $\Delta\nu$ of the spontaneous Raman scattered field is dominated by *homogeneous broadening*, translating into $\tau_{\mathrm{m,coh}} \simeq 1$ to 1.4 ps for the coherence time of the collective vibrational mode.

(ii) The spontaneous linewidth is dominated by *inhomogeneous broadening*, and each individual molecule has a much longer intrinsic coherence time – possibly longer than 40 ps, which would correspond to $\Delta\nu < 7.5$ GHz as observed for the upconversion peak.

In case (i), the resolution-limited, sub-10 GHz linewidth of the upconversion (i.e. IR driven)

Raman peak would find its explanation in the coherent nature of optomechanical transduction, which results in a coherently displaced thermal state of the oscillator. In case (ii), even if upconversion were fully incoherent, the resonant driving of a sub-ensemble of molecules could yield data similar to Fig. 2.

Discussion on coherence lifetime   As mentioned above, case (ii) would entail single molecule vibrational coherence times in excess of 40 ps. For molecules bound to a metal and at ambient temperature, this coherence time appears improbably long. For example, the linewidth of a specific Raman peak of individual Nile Blue molecules adsorbed on Ag surfaces was measured to be 3 to 3.5 cm^{-1} at room-temperature (*74*), corresponding to a coherence time on the order of 3 ps.

The only report we know of in the literature that could support much longer coherence times is Ref. (*75*), where the authors studied *trans*-1,2-bis-(4-pyridyl) ethylene (BPE) molecules embedded inside gold nanoparticle dimers encapsulated in porous silica shells. In Fig. 1d of Ref. (*75*) the continuous-wave Raman spectrum appears Gaussian, in agreement with its being dominated by inhomogeneous broadening. When performing time-resolved, surface-enhanced coherent anti-Stokes Raman spectroscopy (tr-SECARS), the authors occasionally found that the coherent beat signal between two nearby vibrational modes persisted for at least 10 ps, which they interpreted as instances of single-molecule signal free of inhomogenous broadening.

We note, however, that the > 10 ps persistence of the beat note in tr-SECARS does not imply that the continuous-wave Raman linewidth of a single molecule would be observed as smaller than 30 GHz. Indeed, the beat note in tr-SECARS is only sensitive to the relative phase between the two nearby vibrational mode, so that all common-noise is rejected. In particular, this measurement is not sensitive to common jitter of both Raman peaks. Finally, the molecules used in Ref. (*75*) are not bound to gold by thiol groups, as in our experiment. Thiol bounds are strong and modify the Raman and IR polarisability of the molecules, thereby most likely affecting also there vibrational lifetimes.

Discussion on spontaneous Raman lineshape   The most convincing and direct evidence we have in support of Case (i) (linewidth dominated by homogeneous broadening) is a careful analysis of the peak lineshape, as summarised in Fig. S7. More specifically, Figs. S7D and H show two typical spontaneous Stokes peaks and their best fits (in the sense of least squared error) with Lorentzian (red lines) and Gaussian (blue lines) functions. Visual inspection close

to the center and the edges of the peak favor the Lorentzian fit as more faithful. Note that the baseline is skewed by the presence of nearby Raman peaks, in particular at higher frequencies (i.e. smaller Raman shifts for the Stokes sideband). To be more quantitative, we conducted two statistical tests: Akaike's information criterion (AIC) and Bayesian information criterion (BIC) to elect the best of the two fits. Both tests clearly favor the Lorentzian fit, Fig. S7.

We conclude that the observed spontaneous linewidth is more likely dominated by homogeneous broadening. Consequently, the much narrower linewidth of the upconversion peak supports the coherent nature of the upconversion process. Nevertheless, we argue that it would be interesting to conduct future experiments to better estimated the extent of inhomogeneous broadening and to quantify the degree of coherence of the upconverted signal with respect to the incoming IR signal.
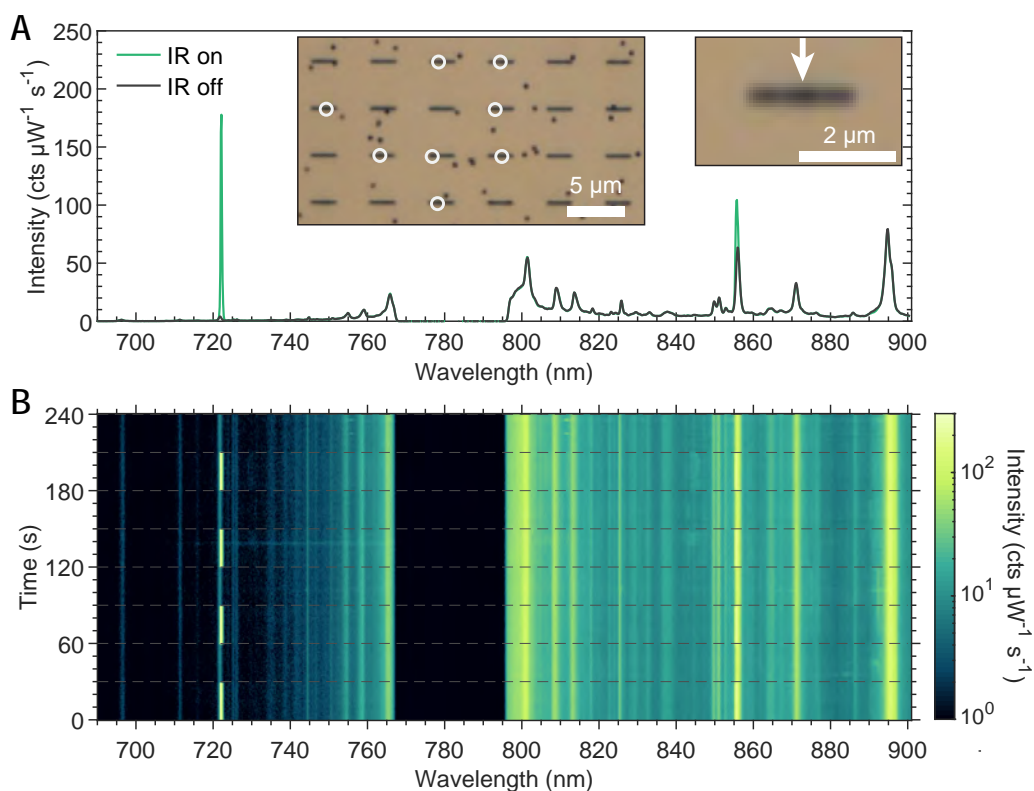
Supplementary Figure S1: **Optical image of the nanoparticle-in-nanogroove system and time series of the Raman spectra with and without IR radiation.** (**A,B**) the Averaged Raman spectra of a nanocavity with and without IR radiation (**A**), which were obtained from the time series of the Raman spectra (**B**) to show the stability and reproducibility of the upconversion measurement. The inset in **A** shows a typical region of the sample, where the single 150-nm-diameter nanoparticle inside the nanogroove (labelled by white circles) can be identified due to the strong scattering cross section of the nanoparticle. Note that no upconverted signal can be observed from the nearby nanoparticle on the Au film.

| Parameter	P_{IR} (μW)	P_0 (μW)	λ_0 (nm)	n_{coh} (aS)	n_{coh} (S)	cts_{aS} (kHz)	η (×10^{-12})
NP1	600	10	740	0.11	0.37	33.87	1.21
NP2	600	10	740	0.05	0.44	8.69	0.31
NP3	600	10	740	0.07	0.15	16.51	0.59
NP4	600	10	740	0.05	0.27	17.42	0.62
NP5	600	10	740	0.10	0.30	36.24	1.30
NP6	600	10	740	0.13	0.29	75.67	2.71
NP7	600	10	740	0.09	0.38	15.90	0.57
NP8	600	10	740	0.07	0.17	30.19	1.08
NP9	545	10	740	0.15	0.19	106.55	4.19
NP10	578	11.2	785	0.16	0.42	22.33	0.83
NP11	578	11.2	785	0.12	0.26	20.56	0.76
NP12	578	11.2	785	0.11	0.10	48.64	1.81
NP13	578	11.2	785	0.06	0.18	16.12	0.60
NP14	578	11.2	785	0.07	0.14	6.40	0.24
NP15	578	11.2	785	0.06	0.29	5.72	0.21
NP16	578	11.2	785	0.14	0.39	27.15	1.01
NP17	578	11.2	785	0.11	0.25	27.63	1.03
NP18	578	11.2	785	0.09	0.26	4.62	0.17
NP19	578	11.2	785	0.11	0.41	21.53	0.80
NP20	578	11.2	785	0.22	0.30	11.18	0.41
NP21	578	11.2	785	0.16	0.44	15.81	0.59
NP22	578	11.2	785	0.11	0.24	17.97	0.67
NP23	578	11.2	785	0.11	0.24	28.04	1.04
NP24	578	11.2	785	0.11	0.26	16.36	0.61
NP25	578	11.2	785	0.26	0.25	40.87	1.52
NP26	545	11.2	785	0.25	0.10	18.20	0.72
NP27	545	11.2	785	0.18	0.28	15.13	0.60
NP28	610	8.6	785	0.15	0.38	106.84	3.76
NP29	545	11.2	785	0.14	0.25	12.55	0.49
NP30	545	11.2	785	0.10	0.32	15.98	0.63
NP31	545	11.2	785	0.08	0.20	20.64	0.81

Supplementary Table S2: **Summary of experimental results on several nanocavities.** cts_{aS} is the net upconverted signal on the anti-Stokes band, converted from CCD counts of the spectrum based on the estimation in Section 1.3. the experimental end-to-end upconversion efficiency $\eta = cts_{aS}h\nu_{IR}/P_{IR}$, where h the Planck's constant, ν_{IR} the IR radiation frequency.
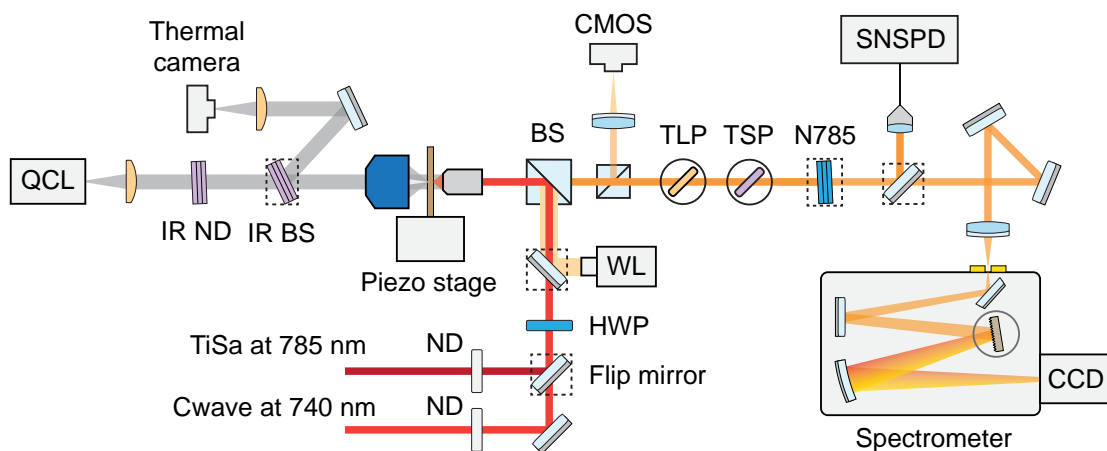
Supplementary Figure S2: **Schematic of the optical setup.** QCL: quantum cascade laser; ND: neutral density filter; BS: beam splitter; CMOS: complementary metal–oxide–semiconductor camera (mvBlueFOX); TLP: tunable long-pass filter; TSP: tunable short-pass filter; N785: notch filter at 785 nm; SNSPD: superconducting nanowire single-photon detector; WL: white light; HWP: half-wave plate; CCD: charge-coupled device.
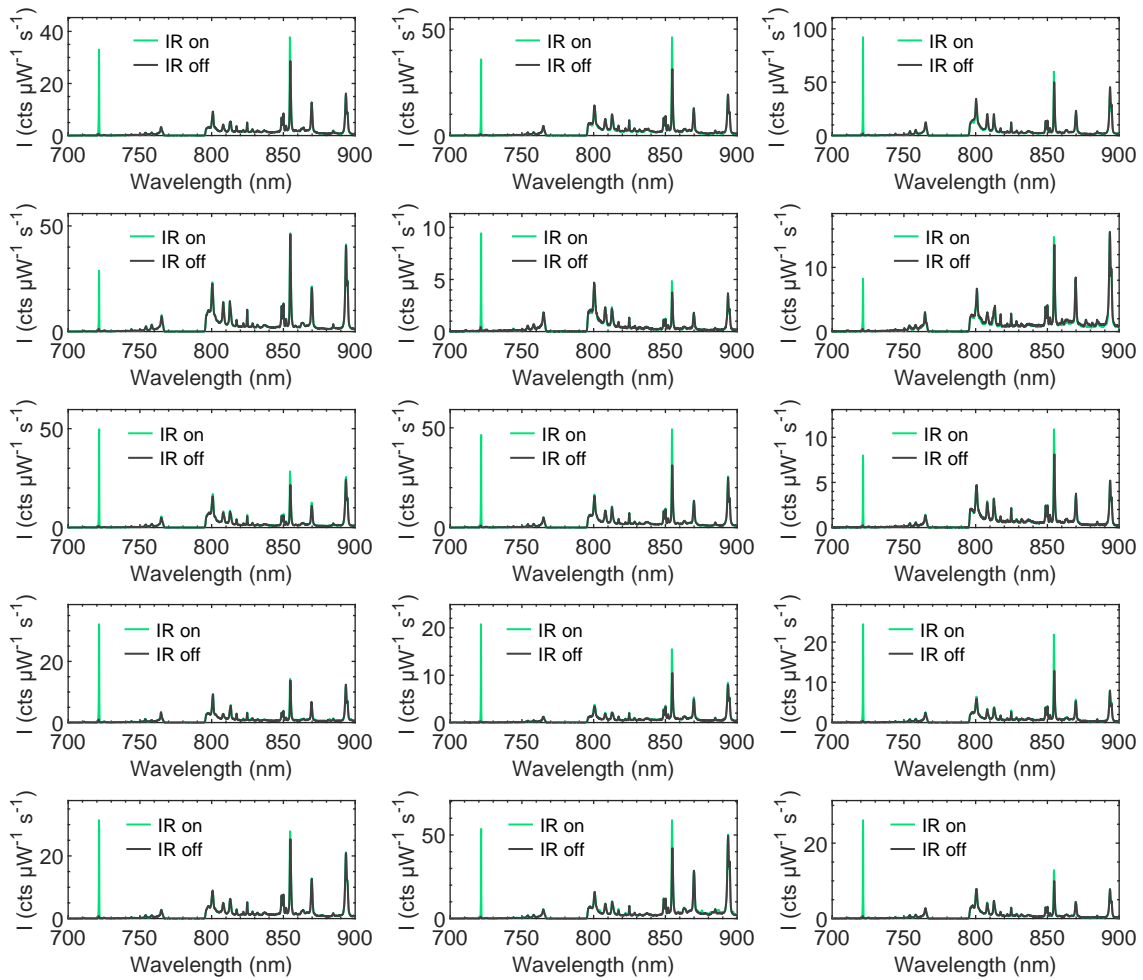
Supplementary Figure S3: **More examples of upconversion spectra from different nanocavities.** The pump light is a TiSa laser at 785 nm with power of 10 μW. IR radiation power, 580 μW.
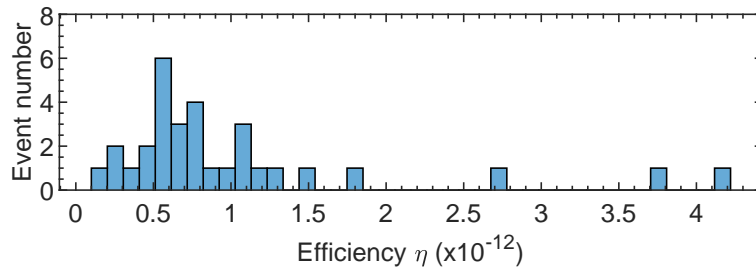

Supplementary Figure S4: **Histogram of the conversion efficiency from Table S1.**
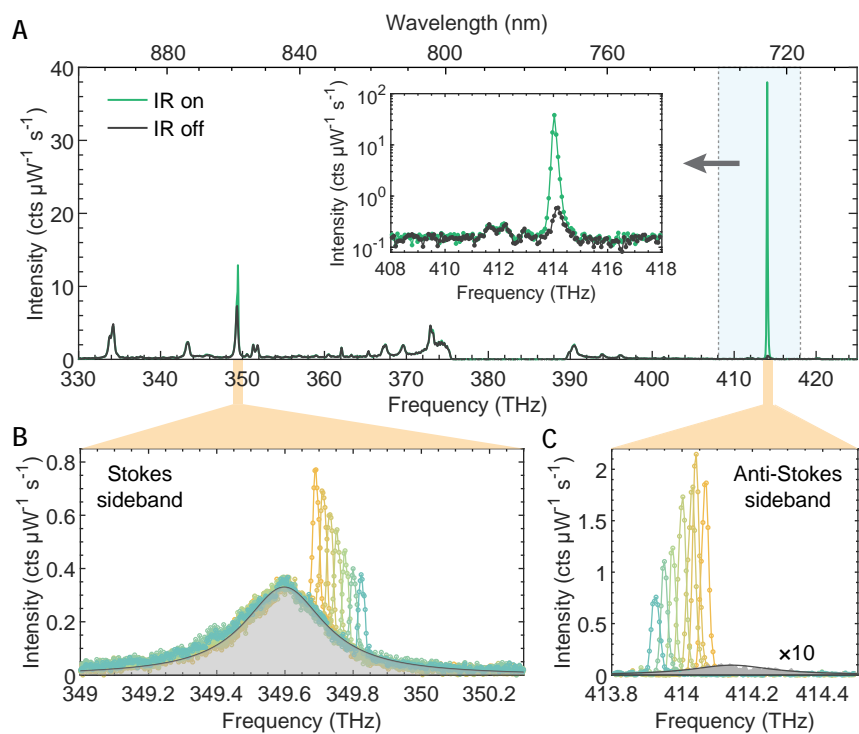
Supplementary Figure S5: **IR frequency scanning upconversion measurement using TiSa 785 nm pump.** (**A**) Low-resoluation, full-range Raman spectra from a single nanocavity without (black line) and with (green line) incoming IR radiation (550 μW on the nanocavity), for 11 μW pump power at 785 nm on the sample. The inset shows a log scale view of the anti-Stokes sideband. (**B,C**) High-resolution Stokes (**B**) and anti-Stokes (**C**) spectra observed when tuning the signal beam with normalized IR radiation power of 135 μW. 785 nm pump power: 11 μW. The grey line in **C** and **D** are the spectra without IR radiation, where the grey areas are the corresponding Lorentzian fits.
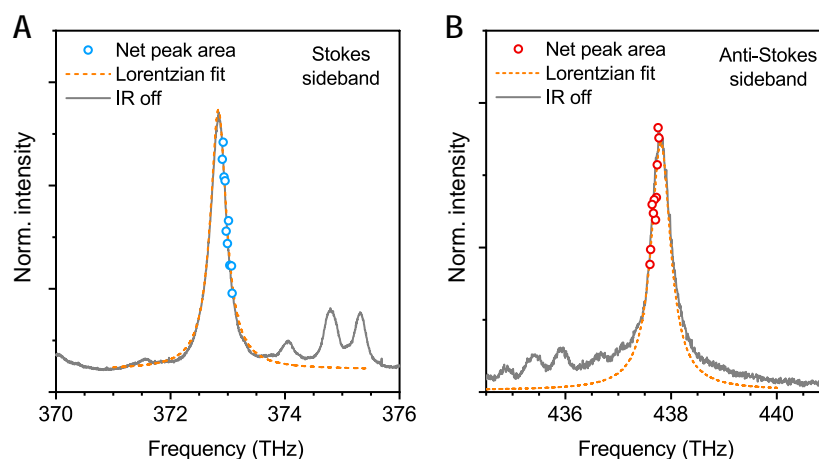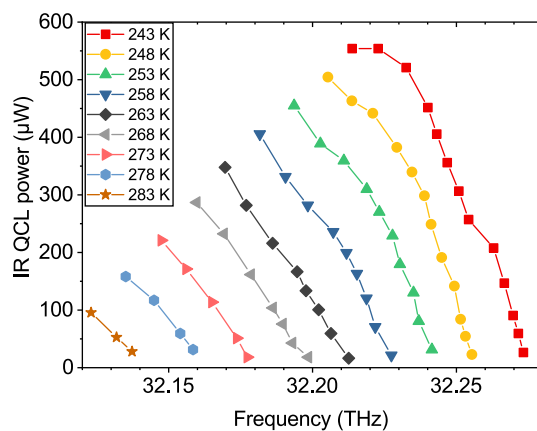
Supplementary Figure S6: **IR frequency dependent upconverted signal intensity versus the spontaneous Raman spectrum.** (**A,B**) Net peak area of the upconverted signal as the function of the peak frequency (obtained from the narrow peaks in Fig. 2 by Lorentian fit) and its comparison with the spontaneous Raman spectra on the Stokes (**A**) and anti-Stokes (**B**) sidebands. The data is the same as shown in Fig. 2C and D. The orange lines are the Lorentzian fits of the frequency dependent net peak area, where the peak position and linewidth for fitting are taken from the value of the corresponding spontaneous Raman peak. The intensities of the spontaneous Raman spectra were normalised to match the height of the orange curve for a direct comparison. The fitting profiles match well with the spontaneous Raman peak, suggesting that the upconversion intensity is determined by the cavity detuning of the phonon modes.
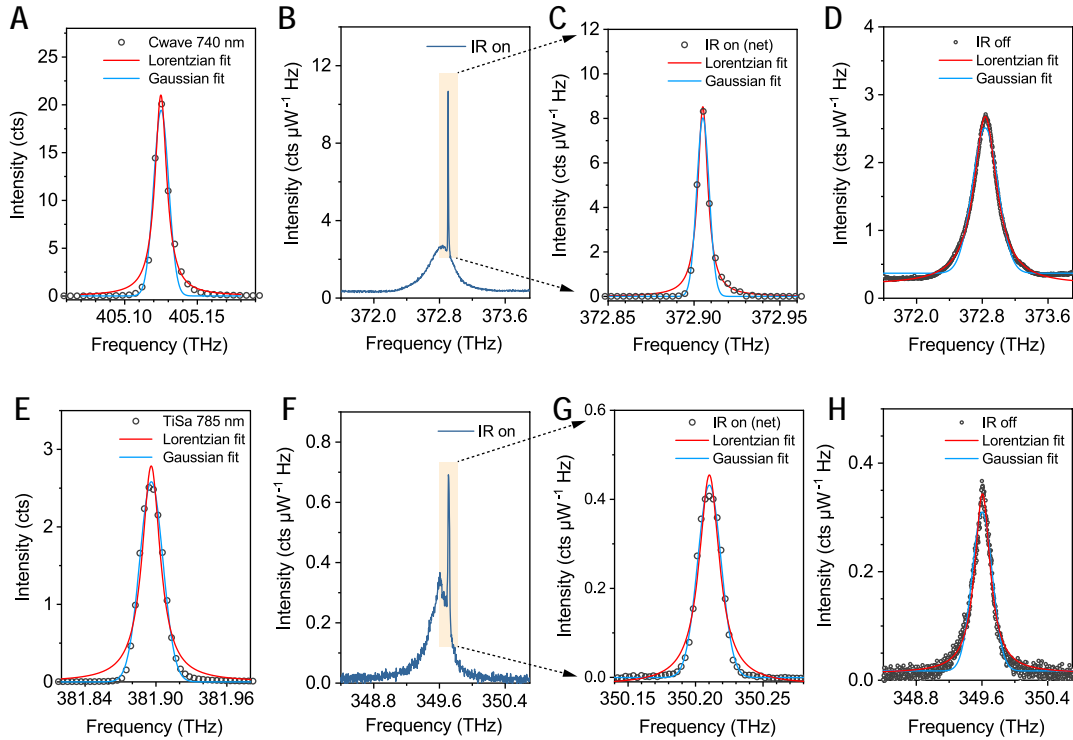
|			Lorentzian Fit			Gaussian Fit		
Pump light	Peak	FWHM	R-Square	AIC	BIC	R-Square	AIC	BIC
C-wave 740 nm	Laser	9.9 GHz	0.98212	8665.083	8693.206	0.9832	8537.477	8565.6
	IR on (net)	7 GHz	0.98421	-4915.705	-4887.582	0.98361	-4839.202	-4811.079
	IR off	337 GHz	0.99833	-1359.767	-1338.3	0.98666	-235.6223	-214.1552
Ti:Sapphire 785 nm	Laser	23.2 GHz	0.96987	-44087.53	-44059.41	0.9981	-49750.39	-49722.26
	IR on (net)	20.5 GHz	0.96783	-505.195	-495.2467	0.99547	-637.7492	-627.8009
	IR off	230 GHz	0.92101	-16983.76	-16955.64	0.91739	-16892	-16863.87

Supplementary Figure S7: **Peak shape analysis by the comparison of Lorentzian and Gaussian fit.** (**A-H**) Typical spectra of the laser (**A, E**), Stokes peak under IR radiation (**B, F**), net upconverted signal (spontaneous peak substracted, **C, G**) and the Stokes peak without IR radiation (**D, H**), along with their Lorentzian and Gaussian fit. Spectra of **B, D** and **F, H** are taken from Fig. 2C and D and Fig. S5B and C, respectively. (**I**) The summary of the fit results from **A-H**, where R-square, Akaike's information criterion (AIC) and Bayesian information criterion (BIC) tests are used to identify either Lorenzian or Gaussian function fits best with the peak. For R-square the higher value the better, and for AIC and BIC the lower value the better (BIC: difference greater than 10 gives decisive conclusion), which are labelled with yellow shadow in the table.

Supplementary Figure S8: **QCL performance as probed by the upconversion spectra.**
Power and frequency of the IR radiation from QCL are dependent on its working temperature and current, which can be probed at the optical domain using high resolution upconversion spectrum (Fig. 2D).The figure shows frequency versus power (on the nanocavity) of the QCL under different temperatures while scanning the current, which agrees well with the FTIR results provided from the QCL supplier. The result is also used as reference for the IR frequency tuning (the power should be set to a small range, Fig. 2D and E) and IR power dependence (the frequency set to a small range, Fig. 3B) measurements.
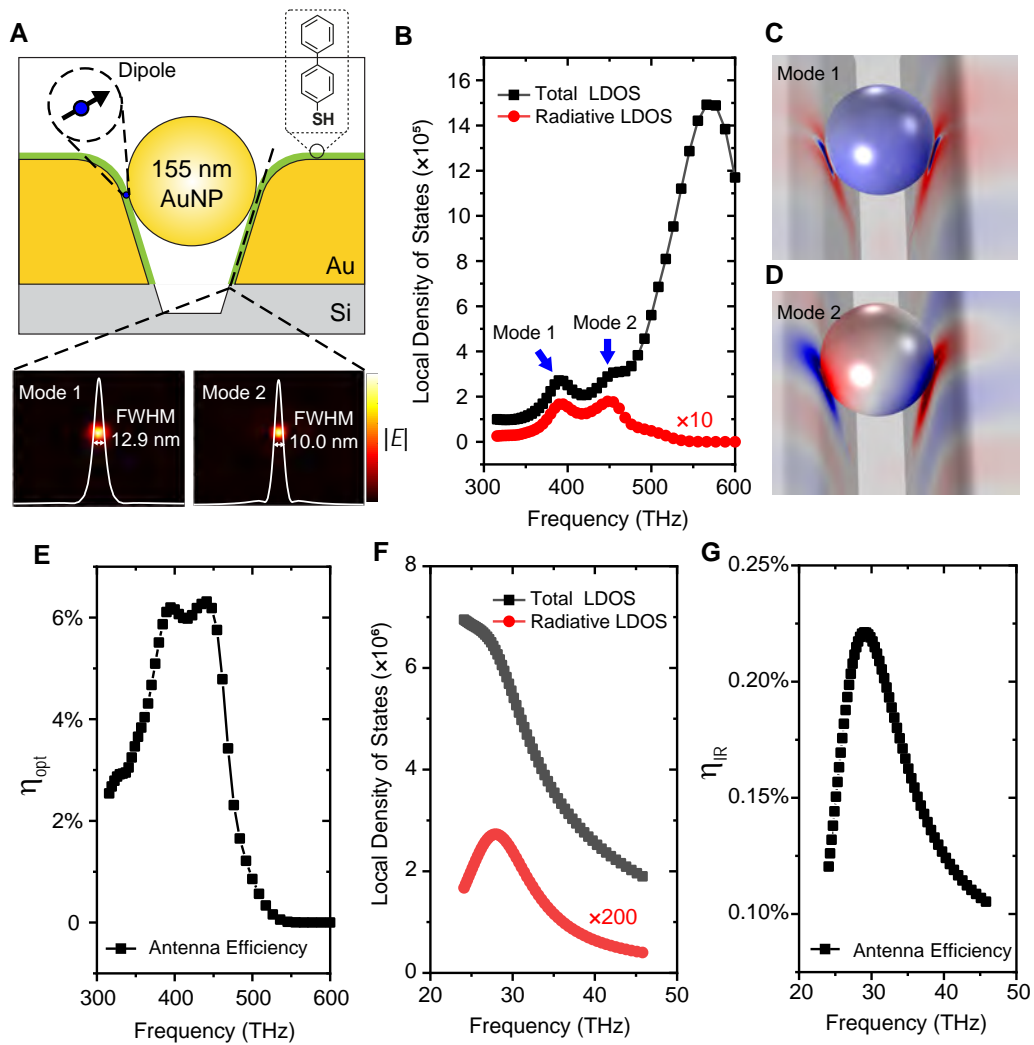
Supplementary Figure S9: **Simulated electromagnetic response of the nanocavity in the visible and infrared range**. The local density of states (LDOS) enhancement (**B, F**) and antenna efficiency (**E, G**) of the modes in visible (**B, E**) and infrared range (**F, G**), respectively, according to the simulation schematics (**A**) where an electric dipole was placed in the nanogap. The electric field distribution of two modes on a slice parallel to the nano-groove's sidewall were plotted (lower panel of **A**), whose mode confinement were demonstrated by their full width at half maximum (FWHM). **C, D** are the charge distribution of two modes near 392.3 THz and 449.5 THz shown in **B**.

35
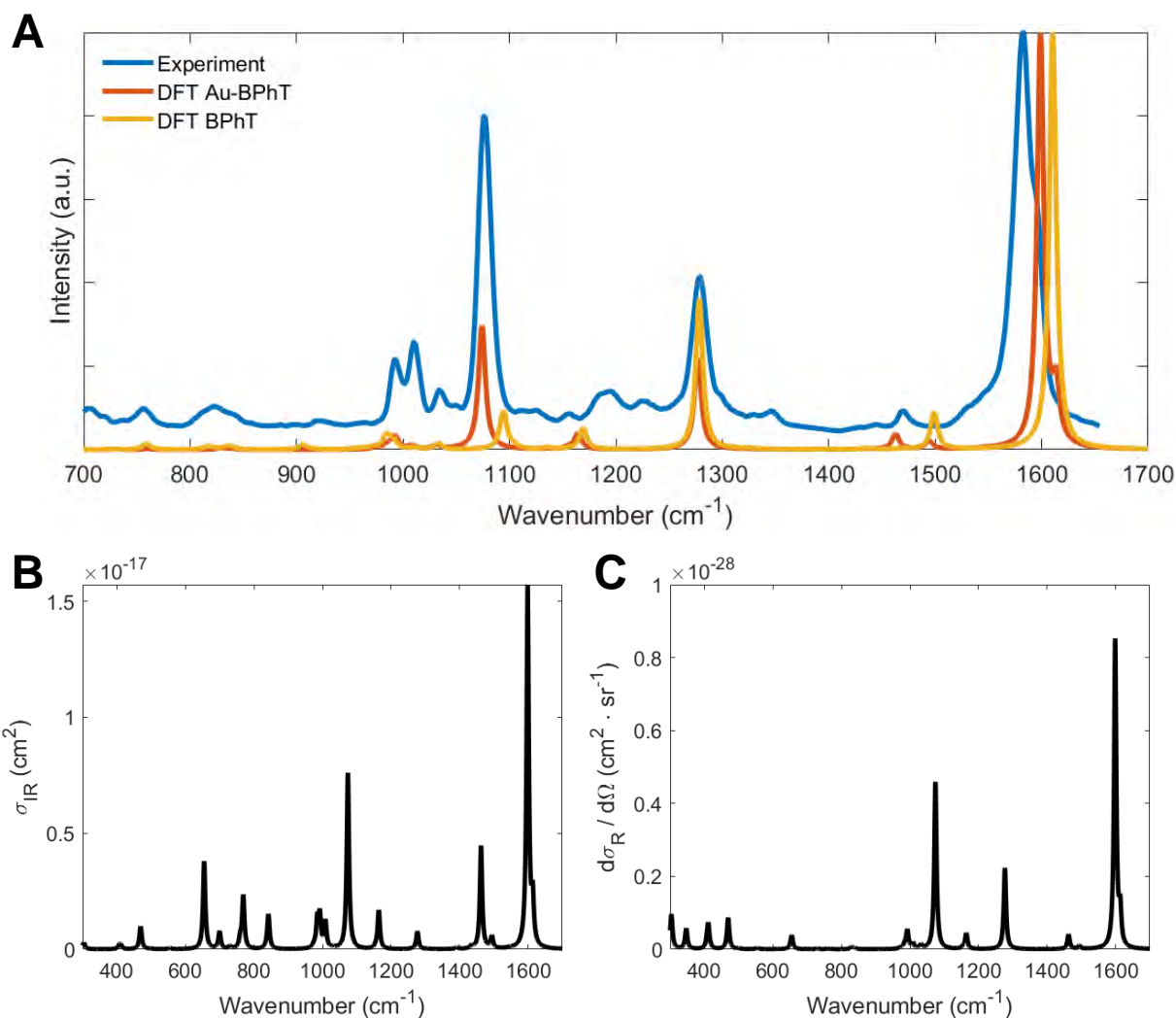
Supplementary Figure S10: **Simulated response of the molecule in the visible and infrared range**. (**A**) Comparison of experimental Raman spectra with DFT simulations of Au-BPhT and BPhT molecules. (**B**) Absorption cross-section and (**C**) differential Raman cross-section spectra of a single Au-BPhT molecule