nature physics

LETTERS

PUBLISHED ONLINE: 28 AUGUST 2017 | DOI: 10.1038/NPHYS4241

# Molecules cooled below the Doppler limit

S. Truppe, H. J. Williams, M. Hambach, L. Caldwell, N. J. Fitch, E. A. Hinds, B. E. Sauer and M. R. Tarbutt*

**Magneto-optical trapping and sub-Doppler cooling have been essential to most experiments with quantum degenerate gases, optical lattices, atomic fountains and many other applications. A broad set of new applications await ultracold molecules[1], and the extension of laser cooling to molecules has begun[2–6]. A magneto-optical trap (MOT) has been demonstrated for a single molecular species, SrF[7–9], but the sub-Doppler temperatures required for many applications have not yet been reached. Here we demonstrate a MOT of a second species, CaF, and we show how to cool these molecules to 50 μK, well below the Doppler limit, using a three-dimensional optical molasses. These ultracold molecules could be loaded into optical tweezers to trap arbitrary arrays[10] for quantum simulation[11], launched into a molecular fountain[12,13] for testing fundamental physics[14–18], and used to study collisions and chemistry[19] between atoms and molecules at ultracold temperatures.**

We first focus on the MOT, which is likely to become a workhorse for cooling molecules just as it is for atoms. Previously, only SrF had been trapped this way. For SrF, two types of MOT have been developed, a d.c. MOT where the lifetime was short and the temperature high[7,8], and a radiofrequency (rf) MOT where longer lifetimes and lower temperatures were achieved[9,20]. In the rf MOT, optical pumping into dark states is avoided by rapidly reversing the magnetic field and the handedness of the MOT laser. It has been suggested that the detrimental effects of dark states can also be avoided in the d.c. MOT by driving the cooling transition with two oppositely polarized laser components, one red- and one blue-detuned[21]. We use this dual-frequency technique to make a d.c. MOT of CaF and find that it works just as well as the rf MOT. Thus, we demonstrate a MOT of a second molecular species, which is important for applications of ultracold molecules, and also verify the effectiveness of this new scheme.

Figure 1 illustrates the experiment, which is described in more detail in Methods. A pulse of CaF molecules produced at time $t = 0$ is emitted from a cryogenic buffer gas source, then decelerated by frequency-chirped counter-propagating laser light, and finally captured in the MOT between $t = 16$ and 40 ms. Figure 2a shows the molecules in the MOT, imaged on a charge-coupled device (CCD) camera by collecting their fluorescence. We estimate that there are $(1.3 \pm 0.3) \times 10^4$ molecules in this MOT (see Methods), with a peak density of $n = (1.6 \pm 0.4) \times 10^5$ cm$^{-3}$. These are similar to the best values achieved for SrF[20]. To determine the MOT lifetime, we fit the decay of its fluorescence to a single exponential. Figure 2b shows this lifetime as a function of the scattering rate. The lifetime is typically 100 ms and decreases with higher scattering rate, suggesting loss by optical pumping to a state not addressed by the lasers. We do not see the precipitous drop in lifetime observed at low scattering rate in the d.c. MOT of SrF[9]. To watch the molecules oscillate in the trap, we push them radially by pulsing on the slowing laser for 500 μs at $t = 50$ ms, and then image them for 1 ms after a fixed delay, $\tau$. Figure 2c shows $\rho(\tau)$, the mean radial displacement of the molecules as a function of this delay. Describing $\rho(\tau)$ by the damped harmonic oscillator equation, $\rho'' + \beta\rho' + \omega^2\rho = 0$, we determine an oscillation frequency of $\omega = 2\pi \times (94.4 \pm 0.2)$ Hz and a damping constant of $\beta = 390 \pm 4\ \mathrm{s}^{-1}$. Similar values were found in SrF MOTs (refs 8,9). The oscillation frequency is within 20% of the value predicted from rate-equation simulations[21], where it was established that the dual-frequency mechanism is mainly responsible for the trapping force. By contrast, the damping constant is 14 times smaller than predicted by these simulations. The reduced damping may be caused by polarization-gradient forces which oppose the Doppler cooling force in the molecule MOT and dominate at low velocities[22]. This is in contrast to normal atomic MOTs, where both forces damp the motion.

We turn now to the temperature, which we measure by turning off the MOT and then imaging the cloud after various free-expansion times $\tau$. Figure 2d shows the mean squared widths in the axial and radial directions versus $\tau^2$, together with fits to the model described in Methods, which give temperatures of $T_z = 11.0 \pm 0.3$ mK and $T_\rho = 12.3 \pm 0.5$ mK. A 10–20% variation between the two directions is typical of all our data. Hereafter, we use the geometric mean temperature, $T = T_\rho^{2/3} T_z^{1/3}$, giving $T = 11.4 \pm 0.3$ mK in this case. This temperature is similar to that observed in SrF MOTs (refs 7–9). We may compare this to the Doppler temperature (see Methods), which for our intensity and detuning is 830 μK, 14 times lower than measured. The elevated temperature is consistent with the reduced damping constant already noted above.

This MOT has a phase-space density of $\rho = h^3 n/(2\pi m k_\mathrm{B} T)^{3/2} = (1.5 \pm 0.4) \times 10^{-15}$. We now show how we increase this by lowering

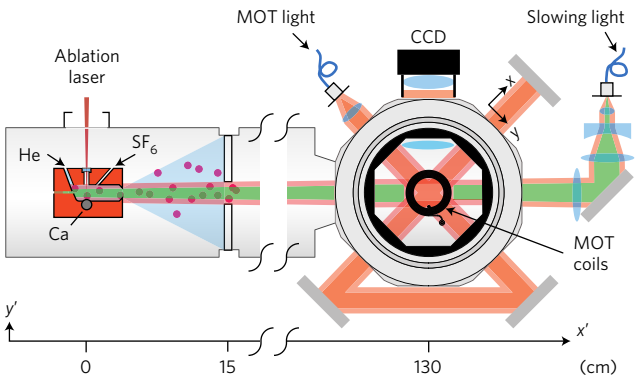


**Figure 1 | Schematic of the experiment.** Pulses of CaF molecules, with a mean forward speed of ~150 m s$^{-1}$, are emitted from a cryogenic buffer gas source. The molecules are slowed to ~10 m s$^{-1}$ using frequency-chirped counter-propagating laser light, captured in a MOT, then cooled to sub-Doppler temperatures in an optical molasses.

Centre for Cold Matter, Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2AZ, UK. *e-mail: m.tarbutt@imperial.ac.uk

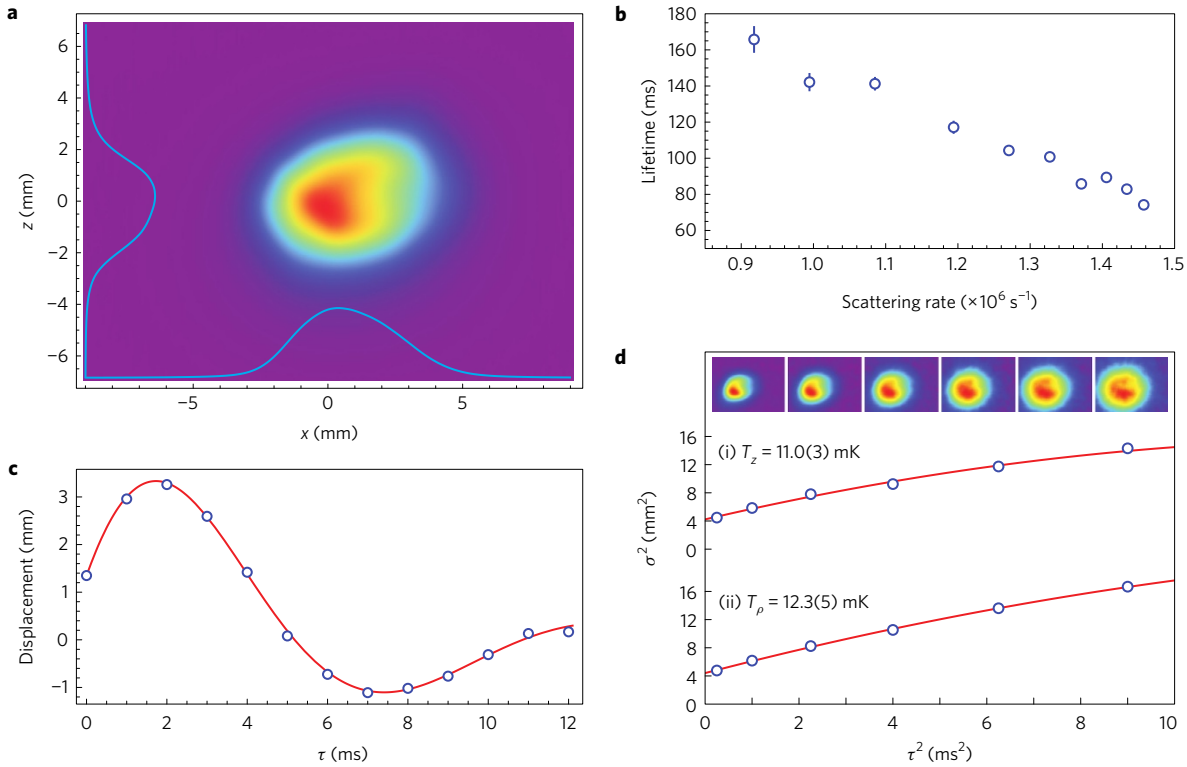


**Figure 2 | Characterization of the MOT. a**, Fluorescence image of the MOT, averaged over 100 shots with 50 ms exposure starting at $t = 40$ ms. Also shown are the 1D axial and radial density profiles obtained by integrating the image along each axis. Gaussian fits to these yield the axial and radial centres and r.m.s. widths, and associated $1\sigma$ standard deviations. **b**, MOT lifetime versus scattering rate. The lifetime, and its $1\sigma$ standard deviation, are found by fitting the decay of the MOT fluorescence to an exponential decay. The scattering rate is controlled via the intensity of the main MOT laser and determined as described in Methods. **c**, Radial displacement of the cloud versus time after pushing it. Each point is obtained by summing 50 images each of 1 ms exposure. Solid line: fit to the motion of a damped harmonic oscillator. **d**, Temperature measurement using the free-expansion method. The mean squared width of the cloud is plotted against the square of the free-expansion time, for (i) the axial and (ii) the radial directions. Each point is obtained by summing 50 images each of 1 ms exposure. The top row shows typical images. Solid lines: quadratic fit to the data (see Methods). Where error bars are not visible in **b**–**d**, they are smaller than the point size.

the temperature. In the rf MOT of SrF, lowering the laser intensity reduced the temperature from 10 mK to 400 μK without loss of molecules[9], and even to 250 μK (ref. 20), although at substantial cost to the number and density. This method was not useful for cooling the d.c. MOT of SrF because of its short lifetime at low intensity. By contrast, the lifetime of our d.c. CaF MOT increases at lower intensity, so this method is open to us. We decrease the power of the main MOT laser between $t = 50$ and 70 ms, hold it for 5 ms, then measure the MOT temperature. Figure 3 shows both the temperature and the size of the MOT versus final intensity, together with the expected Doppler temperature, $T_D$ (see equation (6) in Methods). At 9.2 mW cm$^{-2}$ we find a minimum of 960 μK, which is about $4T_D$. Extending the ramp to lower intensities increases the temperature again. Optimization of the detuning and the bias magnetic field lowers the minimum temperature to about 500 μK, but we do not pursue that further here. Figure 3 shows that the MOT size first decreases as the intensity decreases, but then grows once the intensity is below 50 mW cm$^{-2}$. At the minimum MOT size the density is $2.5 \times 10^5$ cm$^{-3}$. The maximum phase-space density in the MOT is obtained near 25 mW cm$^{-2}$, where it is $\rho = (3.4 \pm 0.9) \times 10^{-14}$, almost the same as the compressed rf MOT of SrF[20]. This ability to reach similar temperatures and phase-space densities in a d.c. MOT as in an rf MOT is an important advance.

To lower $T$ and increase $\rho$ further, we transfer into a three-dimensional blue-detuned optical molasses where we find remarkably effective sub-Doppler cooling. Sub-Doppler cooling usually occurs in a light field with non-uniform polarization, where optical pumping between sub-levels sets up a friction force at low

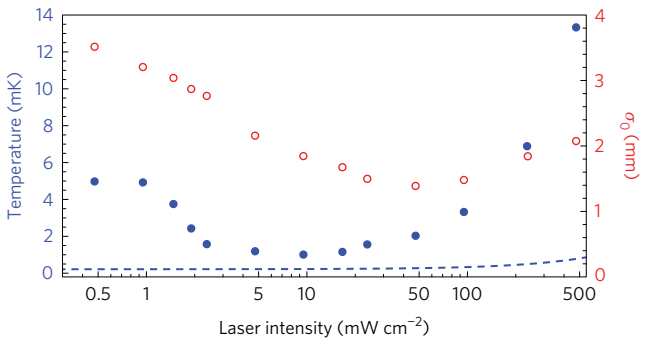


**Figure 3 | Cooling the MOT by ramping down the intensity.** Temperature (blue filled circles) and geometric mean r.m.s. width of the MOT (red open circles) versus total intensity of the MOT laser at the end of the ramp. Dashed line: Doppler temperature (Methods, equation (6)). Error bars are $1\sigma$ standard deviations of fits similar to those in Fig. 2d, and are smaller than the points.

velocity much stronger than the Doppler force[23,24]. Details of the polarization-gradient cooling mechanism depend on the ground- and excited-state angular momenta, $F$ and $F'$. In our case $F \geq F'$ (Supplementary Fig. 1), so there are dark ground states that cannot couple to the local laser polarization, and bright states that can. A bright-state molecule loses kinetic energy on moving into blue-detuned light, and pumps to a dark state when the light has sufficient intensity. The dark molecule moves through changing polarization

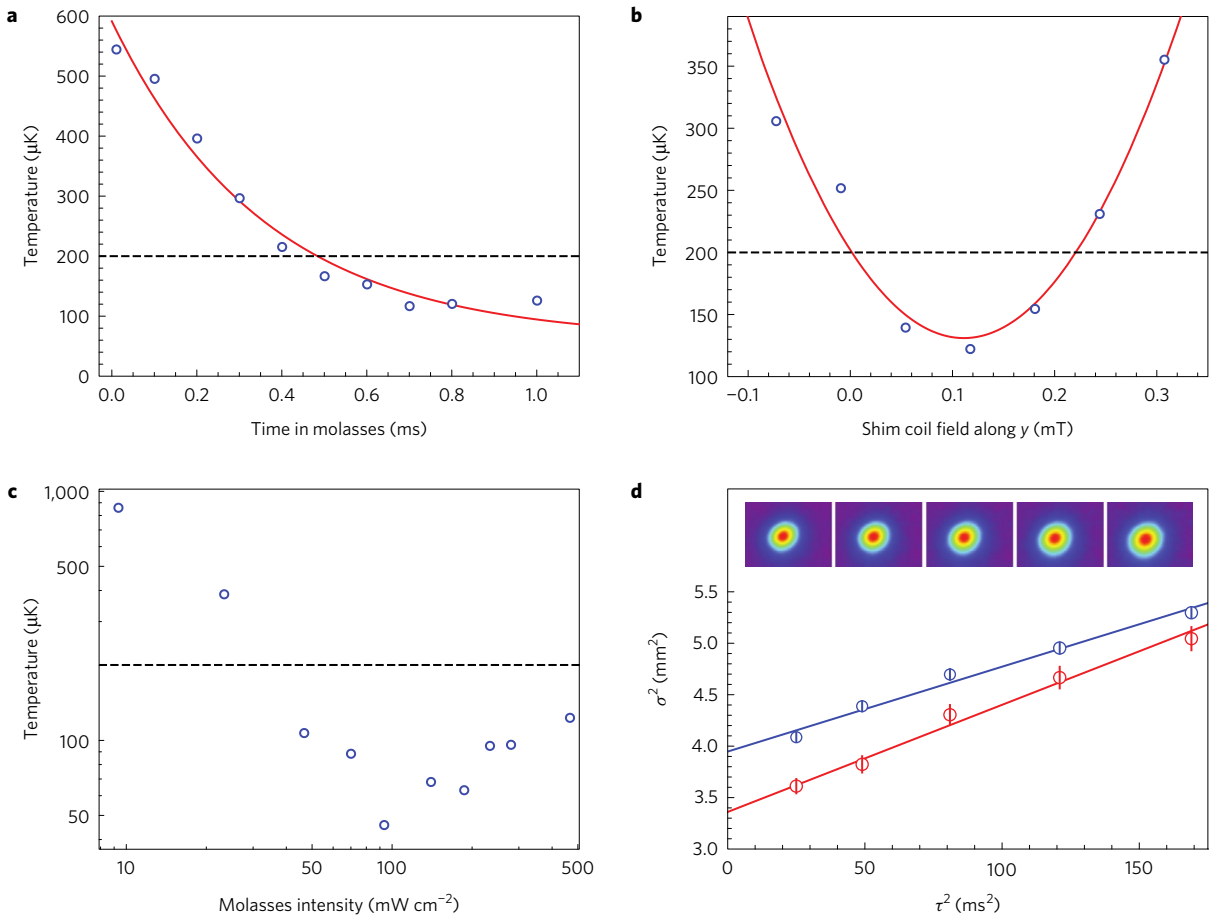


**Figure 4 | Sub-Doppler cooling. a**, Temperature versus time in molasses, when intensity is 460 mW cm$^{-2}$. Solid line: fit to an exponential decay towards a base temperature, giving a $1/e$ time constant of $361 \pm 2$ μs. Dashed line: minimum Doppler temperature. **b**, Temperature versus magnetic field produced at the cloud by one of three shim coils during the molasses phase. Intensity is 460 mW cm$^{-2}$ and molasses hold time is 5 ms. Solid line: quadratic fit giving curvature of $5{,}740 \pm 30$ μK mT$^{-2}$. **c**, Temperature versus molasses intensity, with molasses hold time of 5 ms. **d**, Temperature measurement after a period of 5 ms in a molasses. Intensity is 93 mW cm$^{-2}$. Inset: images at each time point, from which cloud sizes in the axial (red) and radial (blue) directions are obtained from Gaussian fits. Error bars are $1\sigma$ standard deviations. Solid lines: straight line fits. Temperature from six repeated measurements is $55 \pm 2$ μK.

and switches non-adiabatically back to a bright state, preferentially at low intensity where the energies of dark and bright states are similar. This cooling method[25], often called 'grey molasses', has been used to cool atoms. Magnetically induced laser cooling[24] is similar, but uses a magnetic field instead of polarization gradients so that Larmor precession transfers molecules between dark and bright sub-levels. This mechanism has been used to cool molecules in one dimension[2,6], but sub-Doppler temperatures were not reached.

The procedure we use to transfer from MOT to molasses is detailed in Methods. Figure 4a shows the molecules cooling in the molasses towards a base value of ~100 μK, which is half the Doppler limit. The $1/e$ time constant is $361 \pm 2$ μs, implying a damping constant of 1,400 s$^{-1}$. We find that the base temperature is sensitive to all three components of the background magnetic field. Figure 4b shows the variation of temperature versus one field component after optimizing the other two. We find a quadratic dependence on magnetic field, consistent with the model outlined in Methods. Figure 4c shows the temperature versus the laser intensity during the molasses phase. The temperature has a minimum of 46 μK near 100 mW cm$^{-2}$, increases rapidly at lower intensities and more gradually at higher intensities. The temperature dependencies shown in Fig. 4a–c are all similar to those observed in atomic grey molasses[25,26], so we conclude that the cooling process is indeed the grey molasses mechanism. Figure 4d shows the thermal expansion of a cloud after cooling for 5 ms in a 93 mW cm$^{-2}$ molasses. The average of six such temperature measurements gives $T = 55 \pm 2$ μK. To within our 5% uncertainty, no molecules are lost between the initial MOT and this ultracold cloud, excepting loss due to the MOT lifetime. The cloud now has $n = (0.8 \pm 0.2) \times 10^5$ cm$^{-3}$ and $\rho = (2.4 \pm 0.6) \times 10^{-12}$, 1,500 times higher than in the initial MOT and 40 times higher than previously achieved for laser-cooled molecules. It also exceeds that obtained by other molecule cooling methods that have reached sub-millikelvin temperatures[13,27].

Single molecules from this ultracold gas could be loaded into low-lying motional states of microscopic optical tweezer traps and formed into regular arrays[10] for quantum simulation[11]. They could be loaded into chip-based electric traps and coupled to transmission line resonators, forming the elements of a quantum processor[28]. By mixing the molecules with atoms, it will be possible to explore collisions, chemistry[19] and sympathetic cooling[29] in the ultracold regime. Our cooled molecules could be used to search for a time variation of the electron-to-proton mass ratio[16], while application of the methods to other amenable molecules will advance measurements of electric dipole moments[12,17] and nuclear anapole moments[18]. All these applications are feasible with the phase-space densities achieved in the present work. For other applications, it is desirable to increase the density towards $10^{12}$ cm$^{-3}$, as often achieved for ultracold atomic gases. Major increases in density are likely to come by compressing the MOT[20] and from more efficient slowing methods[30] along with transverse cooling[2]

prior to slowing. The resulting dense, ultracold sample would be an ideal starting point for sympathetic or evaporative cooling to quantum degeneracy.

## Methods

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of this paper.

Received 25 January 2017; accepted 21 July 2017; published online 28 August 2017

## References


1. Carr, L. D., DeMille, D., Krems, R. V. & Ye, J. Cold and ultracold molecules: science, technology and applications. *New J. Phys.* **11,** 055049 (2009).
2. Shuman, E. S., Barry, J. F. & DeMille, D. Laser cooling of a diatomic molecule. *Nature* **467,** 820–823 (2010).
3. Hummon, M. T. *et al*. 2D magneto-optical trapping of diatomic molecules. *Phys. Rev. Lett.* **110,** 143001 (2013).
4. Zhelyazkova, V. *et al*. Laser cooling and slowing of CaF molecules. *Phys. Rev. A* **89,** 053416 (2014).
5. Hemmerling, B. *et al*. Laser slowing of CaF molecules to near the capture velocity of a molecular MOT. *J. Phys. B* **49,** 174001 (2016).
6. Kozyryev, I. *et al*. Sisyphus laser cooling of a polyatomic molecule. *Phys. Rev. Lett.* **118,** 173201 (2017).
7. Barry, J. F., McCarron, D. J., Norrgard, E. B., Steinecker, M. H. & DeMille, D. Magneto-optical trapping of a diatomic molecule. *Nature* **512,** 286–289 (2014).
8. McCarron, D. J., Norrgard, E. B., Steinecker, M. H. & DeMille, D. Improved magneto-optical trapping of a diatomic molecule. *New J. Phys.* **17,** 035014 (2015).
9. Norrgard, E. B., McCarron, D. J., Steinecker, M. H., Tarbutt, M. R. & DeMille, D. Submillikelvin dipolar molecules in a radio-frequency magneto-optical trap. *Phys. Rev. Lett.* **116,** 063004 (2016).
10. Barredo, D., de Léséleuc, S., Lienhard, V., Lahaye, T. & Browaeys, A. An atom-by-atom assembler of defect-free arbitrary 2d atomic arrays. *Science* **354,** 1021–1023 (2016).
11. Micheli, A., Brennen, G. K. & Zoller, P. A toolbox for lattice-spin models with polar molecules. *Nat. Phys.* **2,** 341–347 (2006).
12. Tarbutt, M. R., Sauer, B. E., Hudson, J. J. & Hinds, E. A. Design for a fountain of YbF molecules to measure the electron's electric dipole moment. *New J. Phys.* **15,** 053034 (2013).
13. Cheng, C. *et al*. Molecular fountain. *Phys. Rev. Lett.* **117,** 253201 (2016).
14. Hudson, J. J. *et al*. Improved measurement of the shape of the electron. *Nature* **473,** 493–496 (2011).
15. Baron, J. *et al*. Order of magnitude smaller limit on the electric dipole moment of the electron. *Science* **343,** 269–272 (2014).
16. Kajita, M. Variance measurement of $m_p/m_e$ using cold molecules. *Can. J. Phys.* **87,** 743–748 (2009).
17. Hunter, L. R., Peck, S. K., Greenspon, A. S., Alam, S. S. & DeMille, D. Prospects for laser cooling TlF. *Phys. Rev. A* **85,** 012511 (2012).
18. Cahn, S. B. *et al*. Zeeman-tuned rotational-level crossing spectroscopy in a diatomic free radical. *Phys. Rev. Lett.* **112,** 163002 (2014).
19. Krems, R. V. Cold controlled chemistry. *Phys. Chem. Chem. Phys.* **10,** 4079–4092 (2008).
20. Steinecker, M. H., McCarron, D. J., Zhu, Y. & DeMille, D. Improved radio-frequency magneto-optical trap of SrF molecules. *Chem. Phys. Chem.* **17,** 3664–3669 (2016).
21. Tarbutt, M. R. & Steimle, T. C. Modeling magneto-optical trapping of CaF molecules. *Phys. Rev. A* **92,** 053401 (2015).
22. Devlin, J. A. & Tarbutt, M. R. Three-dimensional Doppler, polarization-gradient, and magneto-optical forces for atoms and molecules with dark states. *New J. Phys.* **18,** 123017 (2016).
23. Dalibard, J. & Cohen-Tannoudji, C. Laser cooling below the Doppler limit by polarization gradients: simple theoretical models. *J. Opt. Soc. Am. B* **6,** 2023–2045 (1989).
24. Ungar, P. J., Weiss, D. S., Riis, E. & Chu, S. Optical molasses and multilevel atoms: theory. *J. Opt. Soc. Am. B* **6,** 2058–2071 (1989).
25. Boiron, D., Triché, C., Meacher, D. R., Verkerk, P. & Grynberg, G. Three-dimensional cooling of cesium atoms in four-beam gray optical molasses. *Phys. Rev. A* **52,** R3425–R3428 (1995).
26. Fernandes, D. R. *et al*. Sub-Doppler laser cooling of fermionic $^{40}$K atoms in three-dimensional gray optical molasses. *Europhys. Lett.* **100,** 63001 (2012).
27. Prehn, A., Ibrügger, M., Glöckner, R., Rempe, G. & Zeppenfeld, M. Optoelectrical cooling of polar molecules to submillikelvin temperatures. *Phys. Rev. Lett.* **116,** 063005 (2016).
28. André, A. *et al*. A coherent all-electrical interface between polar molecules and mesoscopic superconducting resonators. *Nat. Phys.* **2,** 636–642 (2006).
29. Lim, J., Frye, M. D., Hutson, J. M. & Tarbutt, M. R. Modeling sympathetic cooling of molecules by ultracold atoms. *Phys. Rev. A* **92,** 053419 (2015).
30. Fitch, N. J. & Tarbutt, M. R. Principles and design of a Zeeman-Sisyphus decelerator for molecular beams. *Chem. Phys. Chem.* **17,** 3609–3623 (2016).


## Acknowledgements

We thank J. Devlin for his assistance and insight. We are grateful to J. Dyne, G. Marinaro and V. Gerulis for technical assistance. The research has received funding from EPSRC under grants EP/I012044, EP/M027716, and EP/P01058X/1, and from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement 320789.

## Author contributions

All authors contributed to all aspects of this work.

## Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to M.R.T.

## Competing financial interests

The authors declare no competing financial interests.

## Methods

**Laser cooling scheme.** Supplementary Fig. 1 shows the energy levels in CaF relevant to the experiment, and the branching ratios between them. The two excited states, $A^2\Pi_{1/2}$ and $B^2\Sigma^+$, have decay rates of $\Gamma = 2\pi \times 8.3$ MHz (ref. 31) and $2\pi \times 6.3$ MHz (ref. 32), respectively. The main slowing laser ($\mathcal{L}^s_{00}$) drives the $B^2\Sigma^+(v'=0) \leftarrow X^2\Sigma^+(v''=0)$ transition at 531.0 nm. Population that leaks to $v''=1$ during the slowing is returned to the cooling cycle by a repump slowing laser ($\mathcal{L}^s_{10}$) that drives the $A^2\Pi_{1/2}(v'=0) \leftarrow X^2\Sigma^+(v''=1)$ transition at 628.6 nm. The hyperfine interval of about 20 MHz in the lowest level of the $B^2\Sigma^+$ excited state is inconvenient for making a MOT, and our modelling[21] suggests that a higher capture velocity is obtained by using the $A^2\Pi_{1/2}$ state where the hyperfine structure is unresolved. The MOT uses four lasers, denoted $\mathcal{L}_{ij}$, to drive the $A^2\Pi_{1/2}(v'=j) \leftarrow X^2\Sigma^+(v''=i)$ transitions. These are $\mathcal{L}_{00}$ at 606.3 nm (called the main MOT laser in the main paper), $\mathcal{L}_{10}$ at 628.6 nm, $\mathcal{L}_{21}$ at 628.1 nm and $\mathcal{L}_{32}$ at 627.7 nm. All lasers drive the P(1) component so that rotational branching is forbidden by the electric dipole selection rules[33]. Each of the levels of the X state shown in Supplementary Fig. 1 is split into four components due to the spin–rotation and hyperfine interactions. The splittings for the $v''=0$ state are shown in the figure, while those for the other states are similar. Radiofrequency (rf) sidebands are added to each laser (see discussion of Supplementary Fig. 2 below) to ensure that all these components are addressed. The frequency component of $\mathcal{L}_{00}$ that addresses the upper $F=1$ state has the opposite circular polarization to the other three. Thus, when the overall detuning of $\mathcal{L}_{00}$ is negative, as in Supplementary Fig. 1, the $F=2$ component is driven simultaneously by two frequencies with opposite circular polarization, one red- and the other blue-detuned. This configuration produces the dual-frequency MOT described in ref. 21. The simulations presented there suggest that most of the confining force in the MOT is due to this dual-frequency effect.

**Set-up and procedures.** Figure 1 illustrates the experiment. A short pulse of CaF molecules is produced at $t=0$ by laser ablation of a Ca target in the presence of $SF_6$. These molecules are entrained in a continuous 0.5 sccm flow of helium gas cooled to 4 K, producing a pulsed beam with a typical mean forward velocity of 150 m s$^{-1}$. The beam exits the source through a 3.5-mm-diameter aperture at $x'=0$, passes into the slowing chamber through an 8-mm-diameter aperture at $x'=15$ cm, and then through a 20-mm-diameter, 200-mm-long differential pumping tube whose entrance is at $x'=90$ cm, reaching the MOT at $x'=130$ cm. The pressure in the slowing chamber is $6\times10^{-8}$ mbar, and in the MOT chamber is $2\times10^{-9}$ mbar. The experiment runs at a repetition rate of 2 Hz.

The beam is slowed using the methods described in ref. 34. The slowing light is combined into a single beam, containing 100 mW of $\mathcal{L}^s_{00}$ and 100 mW of $\mathcal{L}^s_{10}$. This beam has a $1/e^2$ radius of 9 mm at the MOT, converging to 1.5 mm at the source. The changes in frequency and intensity of $\mathcal{L}^s_{00}$ are illustrated in the timing diagram in Supplementary Fig. 3. The initial frequency of $\mathcal{L}^s_{00}$ is set to a detuning of $-270$ MHz so that molecules moving at 145 m s$^{-1}$ are Doppler-shifted into resonance. The light is switched on at $t=2.5$ ms and frequency chirped at a rate of 23 MHz ms$^{-1}$ between 3.4 ms and 15 ms. The frequency of $\mathcal{L}^s_{10}$ is not chirped, which differs from the procedure used previously[34]. Instead, it is frequency broadened as described below, and its centre frequency detuned by $-200$ MHz. Both $\mathcal{L}^s_{00}$ and $\mathcal{L}^s_{10}$ are turned off at $t=15$ ms. A 0.5 mT magnetic field, directed along $y'$, is applied throughout the slowing region and is constantly on.

The MOT light is combined into a single beam containing 80 mW of $\mathcal{L}_{00}$, 100 mW of $\mathcal{L}_{10}$, 10 mW of $\mathcal{L}_{21}$ and 0.5 mW of $\mathcal{L}_{32}$. This beam is expanded to a $1/e^2$ radius of 8.1 mm, and then passed through the centre of the MOT chamber six times, first along $y$, then $x$, then $z$, then $-z$, then $-x$, then $-y$. In this paper, intensity refers always to the full intensity at the MOT due to $\mathcal{L}_{00}$, which is six times the intensity of the $\mathcal{L}_{00}$ beam entering along $y$. The light is circularly polarized each time it enters the chamber, and returned to linear polarization each time it exits, following ref. 7. For any given frequency component of the light, the handedness is the same for each pass in the horizontal plane, but opposite in the vertical direction. All MOT lasers have zero detuning, apart from $\mathcal{L}_{00}$, which has variable detuning $\Delta$. The MOT field gradient, which is 2.9 mT cm$^{-1}$ in the axial direction, is produced by a pair of anti-Helmholtz coils inside the vacuum chamber. The MOT vanishes when this field gradient is reversed or turned off. Three bias coils, with axes along $x'$, $y'$ and $z$, are used to tune the magnetic field in the MOT region to trap the most molecules. The MOT fluorescence at 606 nm is collected by a lens inside the vacuum chamber and imaged onto a CCD camera with a magnification of 0.5. An interference filter blocks background light at other wavelengths.

Supplementary Fig. 2 shows the frequency spectrum of each laser. For $\mathcal{L}_{00}$, a 73.5 MHz electro-optic modulator (EOM) generates the sidebands that drive the $F=2$, $F=0$ and lower $F=1$ states, while a 48 MHz acousto-optic modulator (AOM) generates the light of opposite polarization to address the upper $F=1$ state[4]. The rf sidebands for $\mathcal{L}_{10}$, $\mathcal{L}_{21}$, $\mathcal{L}_{32}$ and $\mathcal{L}^s_{00}$ are generated using 24 MHz EOMs. We spectrally broaden $\mathcal{L}^s_{10}$ to approximately 500 MHz using three consecutive EOMs, one driven at 72 MHz, one at 24 MHz and one at 8 MHz. For each laser, we find the frequency that maximizes the laser-induced fluorescence (LIF) from the molecular beam when that laser is used as an orthogonal probe. These frequencies define zero detuning for each laser, with the exception of $\mathcal{L}_{00}$. For $\mathcal{L}_{00}$, we find that there is a critical frequency where an observable MOT is only formed in half of all shots. These large fluctuations in molecule number seem to come from a strong sensitivity to the approximately 1 MHz frequency fluctuations of $\mathcal{L}_{00}$ at this frequency. We define this critical frequency to be zero detuning, $\Delta=0$. At $\Delta=-0.25\Gamma$ the MOT is stable, and at $\Delta=0.25\Gamma$ there is never a MOT (recall that $\Gamma=2\pi\times8.3$ MHz). We find that this method is more sensitive and reproducible than maximizing the LIF. When $\mathcal{L}_{00}$ is used as an orthogonal probe, the LIF is maximized at $\Delta=0.25\Gamma\pm0.5\Gamma$. We observe MOTs when $0>\Delta>-1.8\Gamma$, and we load the most molecules when $\Delta=-0.75\Gamma$, the value used for all data in this paper.

The complete procedure for cooling to the lowest temperatures is illustrated in Supplementary Fig. 3. The intensity of $\mathcal{L}_{00}$ is ramped down by a factor 100 to 4.6 mW cm$^{-2}$ between $t=50$ and 70 ms to lower the MOT temperature (see Fig. 3), while keeping the detuning at $-0.75\Gamma$. At $t=72$ ms, the shim coil currents are switched from those that load the most molecules in the MOT to those that give the lowest temperature in the molasses, thereby optimizing both molecule number and temperature. The MOT coils are turned off at $t=75$ ms. At $t=76$ ms we jump $\mathcal{L}_{00}$ to a detuning of $+2.5\Gamma$, and to a (variable) higher intensity, to form the molasses (see Fig. 4c). After allowing the molasses to act for a variable time (see Fig. 4a) $\mathcal{L}_{00}$ is turned off so that the cloud can expand for a variable time (see Fig. 4d) before it is imaged for 1 ms at full intensity with $\Delta=-0.75\Gamma$.

**Scattering rate and saturation intensity.** Despite the complexity of the multi-level molecule, it is useful to use a simple rate model[12] to predict some of the properties of the MOT, as done previously[9]. In this model, $n_g$ ground states are coupled to $n_e$ excited states, and the steady-state scattering rate is found to be

$$R_{sc} = \Gamma \frac{n_e}{(n_g+n_e) + 2\sum_{j=1}^{n_g}(1+4\Delta_j^2/\Gamma^2)I_{s,j}/I_j} \tag{1}$$

Here, $I_j$ is the intensity of the light driving transition $j$, $\Delta_j$ is its detuning, and $I_{s,j}=\pi hc\Gamma/(3\lambda_j^3)$ is the two-level saturation intensity for a transition of wavelength $\lambda_j$. In applying this model we need to include the $n_g=24$ Zeeman sub-levels of the $v=0$ and $v=1$ ground states, all of which are coupled to the same $n_e=4$ levels of the excited state. The $v=2$ and $v=3$ ground states can be neglected since they are repumped through other excited states with sufficient intensity that their populations are always small. Because $\mathcal{L}_{00}$ is detuned whereas $\mathcal{L}_{10}$ is not, and because the $\mathcal{L}_{10}$ intensity is always higher than the $\mathcal{L}_{00}$ intensity, the transitions driven by $\mathcal{L}_{10}$ make only a small contribution to the sum in equation (1) and we neglect them. This is a reasonable approximation at full $\mathcal{L}_{00}$ power, and a very good approximation once the power of $\mathcal{L}_{00}$ is ramped down. The 12 transitions driven by $\mathcal{L}_{00}$ have common values for $\Delta$ and $I_s$, and the total intensity, $I_{00}$, is divided roughly equally between them so that we can write $I_j = I_{00}/(n_g/2)$. With these simplifications, we can rewrite equation (1) in the form

$$R_{sc} = \frac{\Gamma_{eff}}{2}\frac{s_{eff}}{1+s_{eff}+4\Delta^2/\Gamma^2} \tag{2}$$

where

$$\Gamma_{eff} = \frac{2n_e}{n_g+n_e}\Gamma = \frac{2}{7}\Gamma \tag{3}$$

and

$$s_{eff} = \frac{2(n_g+n_e)}{n_g^2}\frac{I_{00}}{I_s} \tag{4}$$

Writing $s_{eff}=I_{00}/I_{s,eff}$, we find an effective saturation intensity of

$$I_{s,eff} = \frac{n_g^2}{2(n_g+n_e)}I_s = \frac{72}{7}I_s = 50\ \text{mW cm}^{-2} \tag{5}$$

We have measured the scattering rate at various $\mathcal{L}_{00}$ intensities, using the method described in the next paragraph. These measurements show that the scattering rate does indeed follow the form of equation (2), but suggest that $\Gamma_{eff}$ is roughly a factor of 2 smaller than equation (3) while $I_{s,eff}$ is roughly a factor of 2 smaller than equation (5). However, the determination of $I_{s,eff}$ is sensitive to the value of the detuning, which is imperfectly defined for the multi-level molecule. Fortunately, none of our conclusions depend strongly on knowing the values of either $\Gamma_{eff}$ or $I_{s,eff}$.

**Number of molecules.** To estimate the number of molecules in the MOT, we need to know the photon scattering rate per molecule. We measure this by switching off $\mathcal{L}_{21}$ and recording the decay of the fluorescence as molecules are optically pumped into $v''=2$. The decay is exponential with a time constant of $570\pm10$ μs at full $\mathcal{L}_{00}$ intensity. Combining this with our measured branching ratio of $(8.4\pm0.5)\times10^{-4}$

to $v''=2$ (ref. 35) gives a scattering rate of $2.1\times10^6\ \mathrm{s}^{-1}$. This is about 2.5 times below the value predicted by equation (2). The detection efficiency is $1.6\pm0.2\%$ and is determined by numerical ray tracing together with the measured transmission of the optics and the specified quantum efficiency of the camera. For the MOT shown in Fig. 2a, the detected photon count rate at the camera is $4.3\times10^8\ \mathrm{s}^{-1}$. From these values, we estimate that there are $(1.3\pm0.3)\times10^4$ molecules in this MOT. From one day to the next, using nominally identical parameters, and after optimization of the source, the molecule number varies by about 25%. We assign this uncertainty to all molecule number estimates in this paper.

**Temperature.** In the standard theory of Doppler cooling, the equilibrium temperature is reached when the Doppler cooling rate equals the heating rate due to the randomness of photon scattering. Because both rates are proportional to the scattering rate, the multi-level system is expected to have the same Doppler-limited temperature as a simple two-level system. This is

$$T_{\mathrm{D}}=-\frac{\hbar\Gamma^2}{8k_{\mathrm{B}}\Delta}(1+s_{\mathrm{eff}}+4\Delta^2/\Gamma^2) \tag{6}$$

We measure the temperature using the standard ballistic expansion method. We turn off the magnetic field and $\mathcal{L}_{00}$, then turn $\mathcal{L}_{00}$ back on after a delay time $\tau$ to image the cloud using a 1 ms exposure. For a thermal velocity distribution and an initial Gaussian density distribution of r.m.s. width $\sigma_0$, the density distribution after a free-expansion time $\tau$ is a Gaussian with a mean squared width given by $\sigma^2=\sigma_0^2+k_{\mathrm{B}}T\tau^2/m$, where $m$ is the mass of the molecule. Thus, a plot of $\sigma^2$ against $\tau^2$ should be a straight line whose gradient gives the temperature.

There are several potential sources of systematic error in this measurement which we address here. First we consider whether the finite exposure time of 1 ms introduces any systematic error. While the image is being taken using the MOT light, the magnetic quadrupole field is off. Thus, there is no trapping force, but there is a velocity-dependent force which, according to ref. 22, may either accelerate or decelerate the molecules depending on whether their velocity is above or below some critical value. To quantify the effect, we have made temperature measurements using various exposure times. For a 12 mK cloud we estimate that the 1 ms exposure time results in an overestimate of the temperature by about $0.3\pm0.5$ mK. For a 50 μK cloud, the overestimate is about $0\pm0.3$ μK. These corrections are insignificant.

With a MOT that is centred on the light beams, the intensity of the imaging light is higher in the middle of the cloud than it is in the wings, making the cloud look artificially small. This skews the temperature towards lower values because the effect is stronger for clouds that have expanded. The error is mitigated by using laser beams that are considerably larger than the cloud and that strongly saturate the rate of fluorescence. Using a three-dimensional model of the MOT beams and equations (2) and (5) for the dependence of the scattering rate on intensity, we have simulated the imaging to determine the functional form of $\sigma^2(\tau^2)$ expected in our experiment. The model suggests that a simple $\tau^4$ correction—$\sigma^2=\sigma_0^2+k_{\mathrm{B}}T\tau^2/M+a_2\tau^4$—will fit well to all our ballistic expansion data, will recover the correct temperature, and will give a significantly non-zero $a_2$ for our $T\sim10$ mK data, but a negligible one for all our data where $T\leq1$ mK. We have investigated this in detail using $\sigma^2$ versus $\tau^2$ data with a higher density of data points (12 points, instead of our usual 6). For a hot MOT ($T\sim12$ mK), we see a nonlinear expansion and find that a fit to the above 'quadratic model' gives a temperature that is typically about 10% higher than a linear fit. For an ultracold molasses ($T\sim100$ μK) linear fits to the earliest six points and the latest six points of the ballistic expansion give the same temperature to within the statistical uncertainty of 4%. For the data in Figs 2 and 3 we use the quadratic model, while for the data in Fig. 4 we use the linear model. We determine the minimum temperature of $55\pm2$ μK from the weighted mean of six measurements at a molasses intensity of 93 mW cm$^{-2}$. None of these six datasets show a statistically significant nonlinearity in the plots of $\sigma^2$ versus $\tau^2$, so we use the linear model. Using the quadratic model instead of the linear one changes the temperature to $66\pm7$ μK.

The magnification of the imaging system may not be perfectly uniform across the field of view. This can alter the apparent size of the cloud as it drops under gravity and expands. We have measured the magnification across the whole field of view that is relevant to our data and find that the uniformity is better than 3%. At this level, the effect on the temperatures is negligible.

Finally, a non-uniform magnetic field can result in an expansion that does not accurately reflect the temperature. A magnetic field gradient accelerates the molecules, and since the magnetic moment depends on the hyperfine component and Zeeman sub-level, this acceleration is different for different molecules. We calculate that this effect contributes a velocity spread less than 0.5 cm s$^{-1}$ after 10 ms of free expansion. This is negligible, even for a 50 μK cloud. The second derivative of the magnetic field causes a differential acceleration across the cloud, but this effect is even smaller.

The temperature reached in the molasses depends on the background magnetic field, as shown in Fig. 4a. The sub-Doppler cooling mechanism relies on optical pumping into a dark state near the intensity maxima of the light field, and motion-induced non-adiabatic transitions back to a bright state near the intensity minima[25,36]. This process becomes ineffective if the Larmor precession time between bright and dark states is comparable to the time taken by a molecule to move from a minimum to a maximum. Therefore, in the presence of a magnetic field $B$, the cooling will not reduce the mean velocity much below $v\approx\mu\lambda B/(4h)$, where $\mu$ is the magnetic moment and $\lambda$ is the wavelength. Converting from mean speed to temperature we find the relation $T\approx(\pi m/8k_{\mathrm{B}})(\mu\lambda B/4h)^2$. Taking $\mu\approx\mu_{\mathrm{B}}$ we obtain $T=aB^2$ with $a\approx12{,}500\,\mu\mathrm{K\,mT}^{-2}$. Our data fits well to this quadratic dependence on $B$ and gives $a=5{,}740\pm30\,\mu\mathrm{K\,mT}^{-2}$, consistent with our model, given its approximate nature.

**Data availability.** Data underlying this article can be accessed from Zenodo (https://doi.org/10.5281/zenodo.264440) and may be used under the Creative Commons CCZero licence.

## References


31. Wall, T. E. *et al*. Lifetime of the $A(v'=0)$ state and Franck–Condon factor of the $A-X(0-0)$ transition of CaF measured by the saturation of laser-induced fluorescence. *Phys. Rev. A* **78,** 062509 (2008).
32. Dagdigian, P. J., Cruse, H. W. & Zare, R. N. Radiative lifetimes of the alkaline earth monohalides. *J. Chem. Phys.* **60,** 2330–2339 (1974).
33. Stuhl, B. K., Sawyer, B. C., Wang, D. & Ye, J. Magneto-optical trap for polar molecules. *Phys. Rev. Lett.* **101,** 243002 (2008).
34. Truppe, S. *et al*. An intense, cold, velocity-controlled molecular beam by frequency-chirped laser slowing. *New J. Phys.* **19,** 022001 (2017).
35. Williams, H. J. *et al*. Characteristics of a magneto-optical trap of molecules. Preprint at http://arxiv.org/abs/1706.07848 (2017).
36. Weidemüller, M., Esslinger, T., Ol'shanii, M. A., Hemmerich, A. & Hänsch, T. W. A novel scheme for efficient cooling below the photon recoil limit. *Europhys. Lett.* **27,** 109–114 (1994).