Supporting Information
Richness of molecular junction configurations revealed by tracking a full pull-push cycle

Tamar Yelin, Sudipto Chakrabarti, Ayelet Vilan, and Oren Tal*

Chemical and Biological Physics Department, Weizmann Institute of Science, 76100 Rehovot, Israel
* corresponding author

Section S1- Materials and experimental methods

We perform our experiments in a mechanically controllable break-junction setup (MCBJ) placed in a liquid helium cryostat (4.2 K). We fabricate the sample by attaching a notched wire of Ag (99.997%, 0.1 mm, Alfa Aesar) to a flexible substrate (1-mm-thick phosphor-bronze covered by a 100 µm insulating Kapton film). We use a three-point bending mechanism to bend the substrate. Consequentially, the wire is broken at the notch, exposing (in cryogenic vacuum) two ultra-clean atomically sharp tips that are used as the electrodes of the junction. Vanadocene is purchased from Stream Chemicals (> 95% purity). Benzene, anthracene and tetracene are purchased from Sigma-Aldrich (> 99.9% purity). We introduce the molecules into the metallic junction using either a molecular source located in front of the junction, or from a heated pipe attached to an external vacuum tube that contains the molecules[1–4]. We use a piezoelectric element (PI P-882 PICMA) to tune the bending of the substrate and control the distance between the Ag electrodes with sub-Å resolution. The voltage to the piezoelectric element is driven by a data acquisition (DAQ) card (NI-PCI6221 or NI-PCI4461) connected to a piezo driver (Piezomechanik SVR 150/1). For the conductance measurements, we provide direct current (DC) bias voltage to the junction from the DAQ card, and use a divider (10:1) to improve signal-to-noise ratio. The resulting current is amplified by a current preamplifier (SR570 or Femto DLPCA-200) and recorded by the DAQ card at a sampling rate of 50–200 kHz. The obtained conductance is given by the measured current divided by the applied voltage.

Section S2 - Conditional conductance analysis

To detect the sub-ensemble of conductance traces that show a molecular signature from a given ensemble of traces, we perform conditional conductance analysis. We first define a conductance-range, $[G_{low} - G_{high}]$ that covers the specific conductance feature of interest, and check in which of the traces a significant amount of data points is measured in this range. For that, we extract for each trace t, the number of measured $G_{t,i}$ values within the given G range: $N_t = \sum_i G_{low} \leq G_{t,i} \leq G_{high}$. We then compare N_t to a threshold value N_0, that we define, that separates between traces that we consider as including or not including the conductance feature ($N_t \geq N_0$ and $N_t < N_0$, respectively). Fig. S1 demonstrates the conditional conductance analysis of the data set of Ag/vanadocene, used to identify the push traces that include any of the molecular features. Table S1 lists the number of traces where either the $1G_0$ or the molecular conductance feature are detected (columns b and d, respectively), out of the total amount of traces (column a) in each set for the junctions of Fig. 1 in the main text.

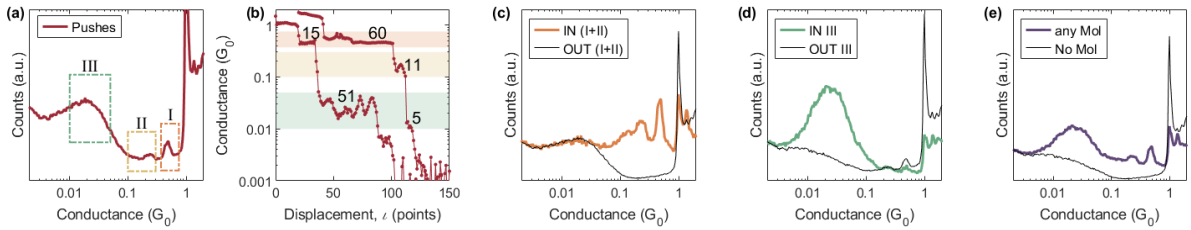

Figure S1: Demonstration of conditional conductance histograms on Ag/vanadocene (same data as Figs. 2-3 of main text). **(a)** The red curve is the push conductance histogram of Ag/vanadocene. The dashed rectangles represent the regions $[G_{low} - G_{high}]$ that we choose to represent each molecular peak: peak I - orange, peak II - yellow, peak III - green. **(b)** Example of N_t counting for two arbitrary traces: the shaded background covers the G-range marked by the dashed rectangles in (a). The number of counts for each trace at each conductance range (N_t) are specified. (e.g., $N_{t,I} = 15$ and 60 for the left and right traces, respectively). **(c-e)** Conditional conductance histograms. Each histogram is composed from the sub-set of traces that according to the conditional analysis either contain or do not contain a certain molecular feature (colored and black histogram, respectively). The relevant conductance features in the panel are: molecular conductance peaks I+II (c), molecular conductance peak III (d), any of the molecular conductance peaks (e).

Junction	(a) Number of traces	(b) Traces with $1G_0$ feature	(c) Traces with Jump to $1G_0$	(d) Traces with mol. feature	(e) Traces with Jump to mol. feature
bare Ag	9993	7714	7261		
Ag/anthracene	9977	8800	5924	8240	6500
Ag/tetracene	9886	4868	3135	5689	4385
Ag/benzene	14922	8287	7494		

Table S1: The table lists for each of the junctions in Fig. 1 in the main text **(a)** the number of traces in each data set, **(b,d)** the number of traces out of the set where a certain conductance feature is detected (conditional conductance analysis), and **(c,e)** the number of traces out of the set where a jump to a certain conductance feature is detected (conditional jump-to contact (J2C) analysis, section S3).

Section S3 - Jump to contact (J2C) detection

In principle, conductance "jumps" are abrupt changes in the conductance as a function of inter electrode displacement, manifested as peaks in the derivative of the conductance with respect to displacement. In practice, however, molecular conductance traces pose few inherent obstacles for a reliable jump detection, including a large dynamic range, and fluctuating data. We demonstrate these challenges, as well as our approach to confront them in Fig. S2, where the independent parameter (x-axis) is given in terms of the running index of the measured point, i, proportional to the displacement.

To address the large dynamic range, the dependent parameter is $\ln G$ rather than G, as shown in panel (a). The derivative: $d(\ln G)/d(i)$ is shown in (b) by the dark-green line. A "jump-segment" (magenta and green markers in Fig. S2a) is any region where the derivative value exceeds a certain tolerance threshold. Here, the tolerance is set to 0.1 in units of: $\Delta \ln G / \Delta i$ as marked by the horizontal dotted line in Fig. S2.b. The derivative peaks with values above this threshold (bright green dots in Fig. S2b) yield the location of the jump segments. Although

this approach works well for high-G values ($i < \sim 200$, left to the vertical dashed-line in Fig. S2), it clearly fails for low-G region (right to the vertical line), with numerous false-detection of noise fluctuations.

In principle, noisy fluctuations can be reduced by smoothing, but it can also eliminate fine features, and therefore undesired. To overcome such false detection, we use the direct derivative method (Fig. S2b) to define the onset and final values of the jump segments, yet combine it with an additional "filter" analysis, demonstrated in Fig. S2c, that distinguishes between *valid* and *false* jumps (magenta and green segments in Fig. S2a, respectively). We use the following filter-function:

$$U_i = \ln\left(std([G_{i-n} \quad \cdots \quad G_{i+n}])\right) \qquad \text{(S1)}$$

The filter-function as shown in Fig. S2c is based on the standard deviation (std) of the conductance G over a $(2n+1)$-long interval ($n = 3$, normally). The choice of std rather than mean or extreme values is suitable to distinguish abrupt jumps from mild-tilted regions, as in the tunneling region. Comparing panels (b) and (c) of Fig. S2, it is visible that the filter function U in panel (c) is much more rounded than the direct derivative in panel (b) and therefore inappropriate for defining the extent of the jump segment, yet, it correctly flattens meaningless fluctuations at low-G region and yields clear peaks where there are valid jumps. Since the U-function has a decaying base-line, the filtering criterion is not by absolute value of U but by peak-prominence with respect to a base line, ΔU, defined as the difference between peak's maximum and the higher of the two adjacent local minima (magenta vertical lines in Fig. S2c). Here, we set a criterion of $\Delta U_{tol} = \ln 3 \approx 1$; and only U-peaks with *prominence* $> \Delta U_{tol}$ are taken as valid jump. In practice, this criterion means that the standard deviation across the jump region is at least 3 times larger than the standard-deviation of its adjacent regions. Finally, we add a third criterion that the net ratio between high to low G values of the jump is at least 1.5 (i.e., at least 50% change). This last criterion is invalid for high-G regions (e.g., a minimal jump of at least 50% would discard a jump from 3 to $4\ G_0$), however it is appropriate in the current context focusing on molecular region.

The example of Fig. S2 detects a total of 5 valid jumps, where each jump is characterized by two conductance values: Onset_G and Final_G. The number of jumps per trace is not limited and varies from trace to trace. The code collects Onset_G values from all the traces into a single array and creates a **J2C histogram** composed of these values. Similarly, a different independent histogram of Final_G is created. (Fig. 1 e-h in the main text).

Note that the values in the collective Onset_G array are paired with those of the Final_G array (each pair is the onset and final conductance values of a specific jump). This pairing is used to create **conditional J2C histograms**. We can select jump-pairs with Final_G values within a chosen G-window ($G_{min} \leq \mathrm{Final\_G} \leq G_{max}$) and then draw the histogram for the corresponding subset of Onset_G, or vice-versa. The conditional J2C analysis is used in the main text in Fig. 1 e-h (colored regions) to discern specifically the jumps that resulted with a molecular conductance value, or with the atomic $1G_0$ value. Table S1 specifies the number of traces where J2C with Final_G at a specific conductance range is detected (columns c and e), out of the total amount of traces (column a), for each of the junctions.

Combining conditional J2C analysis with a preliminary conditional conductance analysis (section S2) allows us to first find the traces that show the molecular conductance feature, and then only for these relevant traces, to find the traces presenting a jump with Final_G at the conductance range of interest. Such a combined analysis is used in Fig. 3a in the main text, to show the sub- Final_G histogram of jumps with their Onset_G value in the range corresponding to conductance peak III.

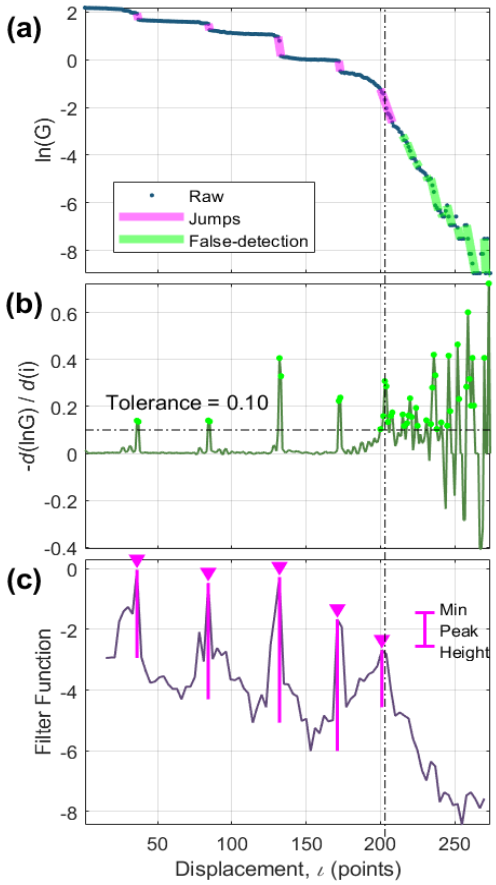

Figure S2: Jumps-detection procedure, illustrated on a single push conductance trace of Ag/anthracene. (a) Raw data of the conductance trace, shown as $\ln G$ *vs*. index of the measured point. **(b)** A simple derivative of the conductance trace. (c) A standard-deviation filtering function (Eq. S1) analysis of the conductance trace. In (a), the magenta-marked segments are the 'jump-segments', as detected by the simple derivative method shown in panel (b). Green-marked segments are false-detection by the derivative method before additional filtering. In (b), green dots mark values larger than the chosen tolerance value (0.1), marked by horizontal dash-dot line; In (c), the dark-purple line is the filter function, U (Eq. S1) computed every 3 data-points. Magenta triangles mark peaks in the filter function higher than a selected tolerance value ($\ln 3$), a height shown by a vertical bar at the top-right corner. The vertical magenta lines underneath the triangles mark peak-height vs. baseline.

Section S4 - Length calibration

The displacement between the two metal-tips in a break-junction is modified by varying the voltage to the piezo-element (V_p), but the displacement is not measured directly. The physical displacement (ΔL) relates to the piezo voltage (V_p) by a calibration factor, κ:

$$\kappa = \frac{\Delta L}{\Delta V_p}\left(\frac{\text{Å}}{mV}\right) \qquad \text{(S2)}$$

The calibration factor, κ is extracted from the slope of the exponential fitting to the tunneling tail, as it is measured in the push conductance traces, according to the following relation:

$$\kappa\left(\frac{\text{Å}}{mV}\right) = -\frac{\Delta \ln G}{\Delta V_p(mV)}\Big/\beta \qquad \text{(S3)}$$

$$\beta = \frac{4\pi}{h}\sqrt{2qm_e\phi} \cong 1.0246\sqrt{\phi(eV)}\left(\text{Å}^{-1}\right)$$

where h is the Planck constant, q and m_e are the electron charge and mass and ϕ is the metal's work-function (WF). Taking $\phi = 4.46eV$ for Ag(111) work-function[5], yields $\beta_{Ag} = 2.16\text{Å}^{-1}$. An extreme uncertainty of $\pm 1eV$ in the WF, leads to an uncertainty of $\pm 0.26\text{Å}^{-1}$ in

β_{Ag} ($\beta_{Ag} = 1.91 - 2.39\text{Å}^{-1}$ or 12%), yielding a (reciprocal) uncertainty of up to 14% in κ. The exact work-function of Ag slightly varies between different references, due to changes in surface orientation, roughness, or contamination. Molecular adsorption (*c.f.* clean Ag) could in principle change the effective work-function[6], yet without strong electron-acceptors/donors, the change is expected to be mild. We note that the length calibrations in the paper are used only to present the traces, and do not have any implications to the analysis in the paper.

For each trace, we detect the slop of the tunneling region by a fit to linearity in the log scale (Fig. S3a). Each slope is translated into a calibration constant κ (Eq. S3). As shown in Fig. S3b, we then construct a histogram of all κ values of the traces that show a reliable tunneling identification (at least 15 data points long and linear regression coefficient, $R^2 > 0.8$). We find a fit to a Gaussian distribution, and the mean of the distribution is taken as the final κ value of the data set. The use of a Gaussian-fit instead of simple mean, reduces sensitivity to outlier values.

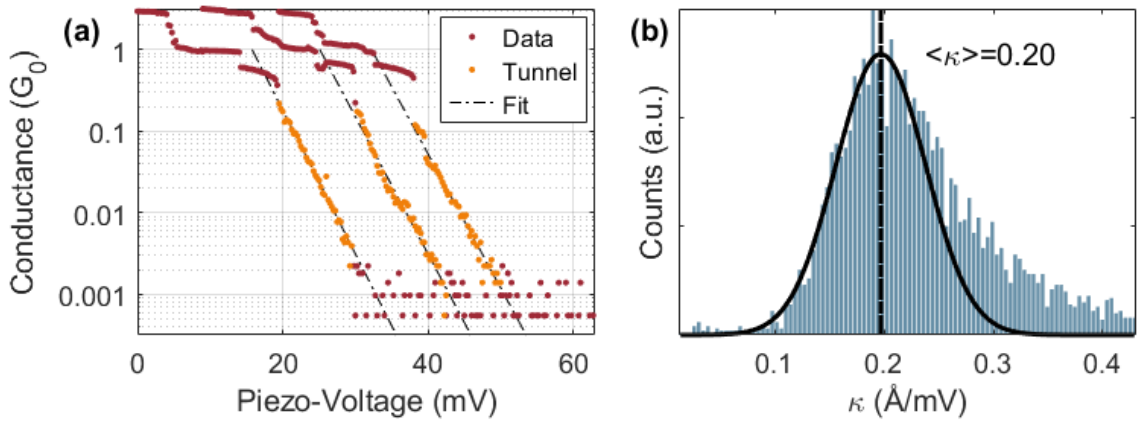

Fig. S3: Length calibration by conductance dependence on piezo voltage, demonstrated for Ag/anthracene junction. (a) Push conductance traces shown as conductance in a log scale, as a function of piezo voltage. Red and orange dots are the measured data points. Orange dots are those detected as the tunneling region. Dashed lines are linear fits to the tunneling region. (**b)** A histogram of the calibration constants κ relating the inter-electrode displacement to the applied piezo voltage, as found from the traces in the data set. The black curve is a fit to Gaussian (normal) distribution and the vertical dashed line marks the mean of the Gaussian.

Section S5 – Critical distance before jump

In order to get an estimation for the difference in the distance at which the J2C occurs with and without the molecules, we use the basic tunneling-length attenuation relation $G = G_t e^{-\beta L}$, and assume that the parameters G_t and β are relatively unchanged with and without molecule presence. This assumption means that also in the presence of molecules, the tunneling is mainly characterized as tunneling through vacuum between the Ag electrodes. (for estimation of possible changes in β up to 12% due to the presence of molecule see section S4). Under these assumptions, the ratio between Onset_G values with and without molecules is related to the difference in the critical length, L_{onset}, of the J2C:

$$\frac{G_{onset}^{Ag}}{G_{onset}^{Ag/Mol}} = \frac{0.3}{0.1} = \frac{exp\left(-\beta L_{onset}^{Ag}\right)}{exp\left(-\beta L_{onset}^{Ag/Mol}\right)}$$

Rearranging we get: $exp\left[-\beta\left(L_{onset}^{Ag} - L_{onset}^{Ag/Mol}\right)\right] = \frac{0.3}{0.1} = 3$

$$L_{onset}^{Ag/Mol} - L_{onset}^{Ag} = \frac{1}{\beta}\ln 3 = \frac{1.1}{(2.16 \pm 0.26)\text{Å}^{-1}} = 0.5\text{Å} \pm 14\%$$

References

1. T. Yelin, R. Vardimon, N. Kuritz, R. Korytár, A. Bagrets, F. Evers, L. Kronik and O. Tal, *Nano Lett.*, 2013, **13**, 1956–1961.
2. D. Rakhmilevitch, S. Sarkar, O. Bitton, L. Kronik and O. Tal, *Nano Lett.*, 2016, **16**, 1741–1745.
3. T. Yelin, R. Korytár, N. Sukenik, R. Vardimon, B. Kumar, C. Nuckolls, F. Evers and O. Tal, *Nat. Mater.*, 2016, **15**, 444–449.
4. A. N. Pal, D. Li, S. Sarkar, S. Chakrabarti, A. Vilan, L. Kronik, A. Smogunov and O. Tal, *Nat. Commun.*, 2019, **10**, 1–8.
5. M. Chelvayohan and C. Mee, *J. Phys. C Solid State Phys.*, 1982, **15**, 2305.
6. D. M. Alloway, A. L. Graham, X. Yang, A. Mudalige, R. Colorado Jr, V. H. Wysocki, J. E. Pemberton, T. R. Lee, R. J. Wysocki and N. R. Armstrong, *J. Phys. Chem. C*, 2009, **113**, 20328–20334.