Supporting Information


Valentina Sacchetti,‡ J. Ramos-Soriano‡, Beatriz M. Illescas,* ‡ M. Teresa González,* ‡ Dongzhe Li,§ Lucía Palomino-Ruiz,†‡, Irene R. Márquez,§ Edmund Leary*,¶ Gabino Rubio-Bollinger,§ Fabian Pauly,* ‡ Nicolás Agraït,‡,¶ Nazario Martín*,†,‡

† Departamento de Química Orgánica, Fac. CC. Químicas, Universidad Complutense de Madrid, Av. Complutense s/n, 28040 Madrid, Spain
‡ IMDEA-Nanoscience, Campus Cantoblanco, 28049 Madrid, Spain
§ Department of Physics, University of Konstanz, 78457 Konstanz, Germany
¶ Departamento de Química Orgánica, Universidad de Granada, C. U. Fuentanueva, Avda. Severo Ochoa s/n, 18071 Granada, Spain
† Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa 904-0495, Japan

Table of Contents .......................................................................................................................... S1
Materials and Methods.................................................................................................................. S2
Synthesis and characterization ..................................................................................................... S2
NMR titration ................................................................................................................................. S3
STMBJ measurements .................................................................................................................. S5
Theoretical details ......................................................................................................................... S12
References ...................................................................................................................................... S13
Materials and Methods

Reagents and solvents were purchased as reagent grade and used without further purification. Compounds 3 and 4 were prepared as previously reported. All solvents were dried by means of standard protocols with sodium and benzophenone as indicator. Silica gel 60 (230-400 mesh, 0.015-0.04 mm) for column chromatography was purchased from Merck. Thin Layer Chromatography (TLC) was performed on aluminium sheets coated with silica gel 60 F254 purchased from Merck and visualized by UV light. IR spectra (in units of cm^{-1}) were measured on an ATI Mattson Genesis Series FTIR instrument. 1H and 13C NMR spectra were recorded either on a Bruker Avance-300 or a Bruker Avance AMX-700 and reported as chemical shifts (δ) in ppm relative to tetramethylsilane (δ = 0) at room temperature, unless another temperature is indicated. Spin multiplicities are reported as a singlet (s), broad singlet (br s), doublet (d), triplet (t) and quartet (q) with proton-proton coupling constants (J) given in Hz, or multiplet (m). 1H-NMR spectra are reported in this order: chemical shift; multiplicity; coupling constant(s); number of protons. Matrix-assisted laser desorption ionization (MALDI) mass spectrometry (MS) was performed on a Bruker Ultraflex spectrometer using dithranol as matrix. ESI-mass spectra were recorded with an Esquire 6000 ESI-ion trap from Bruker Daltonics using MeOH/1% formic acid as the solvent system. Cl-mass spectra were recorded with an LTQ-Orbitrap XL (Thermo Fisher, Waltham, MA) using MeOH/CH₂Cl₂ as the solvent system.

Synthesis and Characterization

4-(methylthio)benzimidamide (2)

A suspension of NH₄Cl (1.05 g, 19.70 mmol), previously desiccated overnight and under scrupulous anhydrous conditions, in dry toluene (10 mL) was cooled at -78 °C and Ar/vacuum cycles were repeated 5 times. After warming to 0 °C, a solution of AlMe₃ (2M in hexane, 9.9 mL, 19.70 mmol) was added to the suspension. When the addition was completed, the reaction mixture was allowed to warm to room temperature and stirred for 2 h under Ar until gas evolution had ceased. A previously degassed solution of 4-methylthiobenzonitrile (200 mg, 1.31 mmol) in dry toluene (10 mL), was added to the aluminum amide complex mixture, and the resulting solution was heated to 90 °C for 78 h. After this period, the heating was stopped, the reaction was quenched with MeOH and the solvent removed under reduced pressure. The crude was purified by column chromatography on silica (CH₂Cl₂/MeOH, 9:1) to afford 4-(methylthio)benzamidinium chloride (230 mg, 87%) as a white powder 1H NMR (300 MHz, CDCl₃ + CD₃OD) δ 8.80 (br s, 1H, NH), 8.63 (br s, 1H, NH), 7.57 (d, J = 8.1, 2H, H-Ar), 7.22 (d, J = 8.3, 2H, H-Ar), 2.41 (s, 3H, CH₃). A suspension of the resulting amидinium salt in CH₂Cl₂ (10 mL) was washed with an aqueous solution of NaOH (10% w/v) and extracted with CH₂Cl₂ (3 × 30 mL). The organic phase was dried over anhydrous MgSO₄, filtered and concentrated to obtain the corresponding neutral form 2 (160 mg, 86%) as a white powder. FT-IR (KBr): 3476, 3433, 3321, 3045, 2918, 2276, 1648, 1590, 1421, 1187, 1080, 766; 1H NMR (300 MHz, CDCl₃) δ: 7.52 (d, J = 8.4, 2H, H-Ar), 7.25 (d, J = 8.2, 2H, H-Ar), 5.16 (br s, 3H, NH), 2.50 (s, 3H, CH₃); 13C NMR (75 MHz, DMSO-d₆) δ: 161.9 (CN), 140.3 (Cipso-
Ar), 132.5 (C<sub>ipso</sub>-Ar), 127.0 (C-Ar), 125.0 (C-Ar), 14.4 (CH<sub>3</sub>)

ESI-MS: m/z calcd for C<sub>8</sub>H<sub>10</sub>N<sub>2</sub>S: 166.1, found: 167.1 [M+H]<sup>+</sup>; ESI-HRMS: m/z calcd for C<sub>8</sub>H<sub>11</sub>N<sub>2</sub>S [M+H]<sup>+</sup>: 167.0637, found: 167.0637.

4'-{methylthio}-[1,1'-biphenyl]-4-carboxylic acid (1b)

A mixture of 4-bromobenzoic acid (120 mg, 0.61 mmol), 4-(methylthio)phenylboronic acid (154 mg, 0.91 mmol), (iPr)<sub>2</sub>NH (171 µL, 1.22 mmol), Pd(OAc)<sub>2</sub> (35 mg, 0.15 mmol) in H<sub>2</sub>O/dioxane (2:1, 3 mL) was allowed to react at 100°C for 15 min. The reaction mixture was washed with brine (10 mL) and extracted with ethyl acetate (3 × 20 mL). The organic phase was dried over anhydrous MgSO<sub>4</sub>, filtered and concentrated. The crude was purified by column chromatography on silica (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1) to afford 1b (138 mg, 93%) as a white powder.

FT-IR (KBr): 3071, 2920, 2672, 2550, 1676, 1600, 1423, 1298, 818, 768; <sup>1</sup>H NMR (300 MHz, DMSO-d<sub>6</sub>) δ: 12.99 (br s, 1H, COOH), 8.00 (d, J = 8.4, 2H, H-Ar), 7.78 (d, J = 8.4, 2H, H-Ar), 7.69 (d, J = 8.5, 2H, H-Ar), 7.36 (d, J = 8.5, 2H, H-Ar), 2.52 (s, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (75 MHz, DMSO-d<sub>6</sub>) δ: 167.2 (COOH), 143.6 (C<sub>ipso</sub>-Ar), 138.8 (C<sub>ipso</sub>-Ar), 135.3 (C<sub>ipso</sub>-Ar), 130.0 (C-Ar), 129.5 (C<sub>ipso</sub>-Ar), 127.3 (C-Ar), 126.3 (C-Ar), 126.2 (C-Ar), 14.5 (CH<sub>3</sub>); CI-MS: m/z calcd for C<sub>14</sub>H<sub>12</sub>O<sub>2</sub>S: 244.1, found: 245.1 [M+H]<sup>+</sup>; CI-HRMS: m/z calcd for C<sub>14</sub>H<sub>13</sub>O<sub>2</sub>S [M+H]<sup>+</sup>: 245.0631, found: 245.0630.

NMR Titration

Binding constants were calculated using Sigma Plot software by fitting the proton signals to a 1:1 association model. A solution of the amidine derivative 2 was titrated with increasing amounts of the corresponding carboxylic acid derivative 1a or 1b. In both cases, K<sub>a</sub> values > 10<sup>4</sup> M<sup>-1</sup> were obtained.
Figure S1. NMR tritation of 2 with increasing amounts of 1a in DMSO-d$_6$ (0; 2.66·10$^{-4}$; 5.27·10$^{-4}$; 7.85·10$^{-4}$; 1.04·10$^{-3}$; 1.29·10$^{-3}$; 1.54·10$^{-3}$; 1.81·10$^{-3}$; 2.07·10$^{-3}$; 2.32·10$^{-3}$; 2.57·10$^{-3}$; 3.10·10$^{-3}$; 3.61·10$^{-3}$; 4.13·10$^{-3}$; 4.63·10$^{-3}$; 5.17·10$^{-3}$; 5.68·10$^{-3}$; 6.19·10$^{-3}$; 6.69·10$^{-3}$; 7.21·10$^{-3}$ and 7.73·10$^{-3}$ M). The equivalents added in each spectrum are indicated in the lower part, which is a magnified view of the region exhibiting prominent peaks in the upper plot.
Figure S2. NMR titration of 2 with increasing amounts of 1b in DMSO-\textit{d}$_6$. (0; 8.43·10$^{-5}$; 1.7·10$^{-4}$; 2.56·10$^{-4}$; 3.42·10$^{-4}$; 4.28·10$^{-4}$; 4.71·10$^{-4}$; 5.14·10$^{-4}$; 5.57·10$^{-4}$; 6.00·10$^{-4}$; 6.43·10$^{-4}$; 6.86·10$^{-4}$; 7.30·10$^{-4}$; 7.73·10$^{-4}$; 8.16·10$^{-4}$; 8.59·10$^{-4}$; 1.07·10$^{-3}$ and 1.29·10$^{-3}$ M). The equivalents added in each spectrum are indicated in the lower part, which shows a different view of the upper plot.

**STMBJ measurements**

*Experimental details*

In order to form molecular junctions with the studied compounds, we used a made-in-house STM optimized to work in ambient conditions. We used commercial gold over glass substrates (Arrandee), and a freshly cut gold wire as tip. The experiments were performed in air, after immersing the substrates in 1 mM solutions of the compounds in dioxane. A constant bias voltage \( V \) of 230 mV was established between the tip and the substrate in all the experiments, and a 12 MΩ resistor was placed in series for protection. The current \( I \) through the STM substrate-tip circuit was measured with a made-in-house linear current-to-voltage converter with two amplification stages. In this work, if not noted otherwise, we used as gains of 5 × 10$^8$ and
These values allowed us to explore a range in conductance $G = I/V$ of 9 orders of magnitude between $10G_0$ and $10^{-8}G_0$ ($G_0 = 2e^2/h$).

The STM gold substrates were covered by the compound to be studied using the dip-casting technique. A just flamed-annealed substrate was immersed for several minutes in a 1 mM solution of the compound. As mentioned before, in this work we used dioxane as solvent for all the compounds.

Examples of individual $G$ vs $z$ traces, recorded during the separation part of the cycle, are shown in Figure S3. We recorded several thousands of them. In the traces, regions of constant conductance, i.e. plateaus, at values below $1G_0$ indicate the formation of stable molecular junctions between the STM tip and the substrate. They do not occur in all the $G$ vs $z$ traces. So in order to discriminate the molecular-junction signature, we separated the traces with plateaus from those without using a computer program. Our criterion to identify a plateau was that at any conductance below $0.5G_0$, a displacement $\Delta z$ larger than 8 Å was needed to produce a change in conductance of $\Delta \log(G/G_0) = 0.1$. As reference, the typical displacement needed for gold-air-gold tunneling to produce this conductance change is 0.2 Å. Using the traces that complied with this criterion, we built conductance histograms by accumulating the number of points $N_{\text{points}}$ measured in fixed $\Delta \log(G/G_0)$ intervals from thousands of traces. The y-axis in these histograms is normalized, as reported previously, as

$$N_{\text{norm}} = \frac{N_{\text{points}}}{N_{\text{curves}} \cdot v_p \cdot \Delta \log(G/G_0)}$$

where $N_{\text{curves}}$ is the total number of $G$ vs $z$ traces included in the histogram, and $v_p$ is the number of points recorded per unit of length in $z$. With this normalization, $N_{\text{norm}}$ is now just the inverse of the typical $\log(G/G_0)$ vs $z$ curve, that tells us the distance $z$ that we would need to move to produce a variation of 1 order of magnitude in conductance $G$, at each particular point Therefore, it has units of nanometers per order of magnitude that we abbreviate to nm/order-mag.

In addition, we built 2D conductance-distance histograms by accumulating the number of points measured in fixed $\Delta \log(G/G_0)$ and $z$ intervals from the traces. In this case, we used as zero for $z$ the position at which the conductance of $0.5G_0$ is reached (just after each gold contact is broken). In 1D histograms accumulation takes place only in fixed $\Delta \log(G/G_0)$ intervals, and no $z$ alignment needs to be performed. At those conductance values where plateaus are frequent in the $G$ vs $z$ traces, there will be a larger number of points recorded, and a peak will form in the 1D and 2D conductance histograms.

![Figure S3. Individual $G$ vs $z$ traces recorded while separating the STM tip from the substrate. Plateaus around $1G_0$ indicate the formation and breakage of a gold one-atom tip-substrate contact.](image-url)
while lower conductance plateaus signal the formation and rupture of molecular junctions between the tip and substrate.

We determined the length of the plateaus in the $G$ vs $z$ traces by calculating the total displacement $\Delta z$ needed to change the junction conductance from $0.5G_0$ to a conductance value below the peak observed in the corresponding conductance histogram, and adding 0.4 nm to account for the typical gold retraction after the gold contact breaks.\(^3\) The length distributions as those of Figure 3 in the main text of Figures S4, S6 and S7 give the percentage ($p$) of curves in the measurement run with a plateau with length in each length interval. As criterion to determine the typical maximum length in our distributions, we used the largest length at which the distribution decays to 20% of the maximum probability, and the error is determined by allowing a change of $\pm 5\%$ to this probability value. Figure S7 shows the plateau length distributions for high $G$ and low $G$ plateaus of 1a-2 in comparison with those of compounds 2 and 3.

![Figure S4. Conductance histograms for compounds 1a, 2 and 4. The inset shows the plateau length distributions for the three compounds, with cross marks indicating the typical maximum plateau lengths. For compound 4, apart from the peak between $-3<\log(G/G_0)<-2$, there is a clear second peak between $-6<\log(G/G_0)<-4$. These two peaks are analyzed separately in Figures S5 and S6.](image)

![Figure S5. 2D conductance vs distance histograms for compound 4 built with (a) all traces exhibiting plateaus, (b) with plateaus only in $-3<\log(G/G_0)<-1$, or (c) with plateaus in $-6.5<\log(G/G_0)<-3.5$. Traces having plateaus in both intervals are included in this last group. The conductance range of the plateaus is indicated in the figures in brackets on a logarithmic scale.](image)
Figure S6. Conductance histograms for compounds 2 and 4 built from traces (a) with plateaus only in $-3.5 \leq \log(G/G_0) < -1$ or (b) with plateaus in $-6.5 \leq \log(G/G_0) < -3.5$. The conductance range of the plateaus is indicated in brackets on a logarithmic scale. The inset in (a) shows the plateau length distribution for both groups of traces for compound 2, demonstrating that the plateaus in the low-conductance peak (blue line) are about double of the length of those in the high-conductance peak (green line).

Figure S7. Plateau-length distributions of the high-$G$ and low-$G$ plateaus, observed for complex 1a·2 in comparison with those observed for compounds 2 and 3.

Current versus Voltage Curves
We recorded opening $G$ vs $z$ traces, stopping the movement of the STM tip each 0.016 nm to measure $I/V$ curves whenever the conductance was within the interval $10^{-7} G_0 - 7 \times 10^{-6} G_0$ (see for example Figure S8). The bias voltage was then ramped between $-1.2$ and $1.2$ V, and both the ramping-up and ramping-down curves were collected. The random orientation of the molecules between the electrodes means that the amidinium and carboxylate moieties will randomly connect to the STM tip or substrate so that we can expect $I/V$ curves with higher current at either positive or negative voltages. In order to visualize the asymmetry of a typical $I/V$
curve, we flipped individual \(IV\) curves before averaging or building histograms with our data to ensure that the higher current branch is always at the same side, which we chose to be at positive voltages. In particular, we did so in 43% of the cases, which is reasonably close to the expected 50%. Figure S9 displays the histograms and averages for the logarithm of the current and conductance for 1a-2 in comparison with the corresponding results recorded without molecules in a similar conductance range. Although the curvature of the conductance vs. voltage curves is similar for molecular junctions and tunneling, a higher asymmetry can clearly be observed for 1a·2.

We define the rectification ratio (RR) in the \(IV\) curves as the ratio between the absolute value of the current for voltages of different signs. In order to calculate the RR of the recorded \(IV\) curves, we first performed a second-order polynomial fit to the curve and used the resulting current values for \(\pm 1\) V. Examples of these fits to individual \(IV\) curves are shown below in Figure S10. In addition Figure S11 shows the histogram of the obtained RR values in our measurements. Note that by definition our RR is always larger or equal to 1. The RR is above 1.5 for 60% of the recorded \(IV\) curves, and above 2 for 30%, with an average of 1.9. This value agrees with the RR of the average \(IV\) curve of Figure S9(a).

**Figure S8.** Example of an individual \(G\) vs \(z\) trace for 1a·2 with \(IV\) curves recorded along the molecular conductance plateau. The open blue circles indicate the positions at which the STM tip was stopped and \(IV\) curves were recorded between -1.2 V and 1.2 V. The inset shows the average \(IV\) curve for the whole plateau with error bars indicating the variation along the plateau.

**Figure S9.** Histograms of the logarithm of the current (a) and conductance (b) versus voltage for 1a·2. (c) Corresponding histogram of the logarithm of the conductance versus voltage in the absence of molecules (just tunneling) in a similar conductance range as (b). In all figures the
averaged curve is superposed in black. The results for 1a·2 show a slight asymmetry when compared with just tunneling ones.

Figure S10. Four examples of individual $I/V$ curves for 1a·2. The recorded curves are displayed in blue and the second-order polynomial fit in green.

Figure S11. Histogram of the RR at 1 V for all individual $I/V$ curves recorded for 1a·2. The average RR value is 1.9.
Figure S12. (a,b) 2D histograms for 1b·2 focusing on the range -6<\log(G/G_0)<-3.5: Measurements recorded with (a) high gain (5 × 10^8 and 2.5 × 10^{10} V/A) and (b) low gain (4.3 × 10^7 and 2.1 × 10^9 V/A). (c) 2D histogram of the traces recorded for 2 displaying plateaus in the same conductance region (the corresponding 1D histogram is shown in Figure S6b). (d) 2D histogram for compound 1b.

Figure S13. Conductance histogram for 1b·2 built from traces displaying plateaus between -6<\log(G/G_0)<-3.5 (green) in comparison with histograms for 2 with plateaus in the same range (blue) and 1b (red).
Computational details

We performed the electronic structure calculations using the quantum-chemistry package TURBOMOLE 6.5 in the framework of density functional theory (DFT). We used the generalized-gradient approximation with the Perdew-Burke-Ernzerhof (PBE) parameterization of the exchange-correlation functional, the basis sets of split-valence-plus polarization quality def-SV(P) and the corresponding Coulomb fitting basis sets.

We first optimized geometries of all free molecules in the gas phase and analyzed their electronic structure. Then, the Au-molecule-Au junctions were built by placing the relaxed free molecules between two Au$_{20}$ clusters. After that, the molecule and the 4 Au tip atoms on each side of the junction closest to the molecule were relaxed, while the other Au atoms were fixed. Total energies were converged to a precision of better than $10^{-8}$ a.u., while the geometries were optimized, until the change in the maximum norm of the Cartesian gradient was below $10^{-5}$ a.u. The transmission function is obtained through the NEGF formalism as

$$\tau(E) = \text{Tr}[\Gamma_L(E)G^a(E)\Gamma_R(E)G^r(E)]$$

where $G^{r,a}$ are the retarded and advanced Green’s function and $\Gamma_{L,R}$ are the linewidth broadening matrices of the scattering region, stemming from left and right electrodes, evaluated as imaginary parts of the corresponding self-energies.

It is well known that DFT with semi-local exchange-correlation functionals is not able to describe accurately the energy-level alignment at the hybrid metal-molecule interfaces due to self-interaction errors and missing nonlocal image-charge effects. To avoid these problems, we employed a non-selfconsistent self-energy correction scheme, the so-called DFT+$\Sigma$ method. Previous studies showed that the DFT+$\Sigma$ approach predicts conductance values in good agreement with single-molecule experiments. Following Ref. 11, we carried out the DFT+$\Sigma$ procedure, which involves the following 3 steps: (i) Correct the gas phase HOMO and LUMO energies of the molecules by $\Delta_{\text{SCF}}$ calculations to get the IP and EA, (ii) estimate the image-charge correction $\Delta_{\text{occ}}$ for all the occupied states using the HO-MO charge distribution and likewise the LUMO charge distribution to obtain the correction $\Delta_{\text{vir}}$ for all the virtual states, (iii) apply the energy shift to all the occupied states via $\Sigma_{\text{occ}} = -\text{IP} - \epsilon_{\text{H}} + \Delta_{\text{occ}}$ and the corresponding one for the unoccupied states via $\Sigma_{\text{vir}} = -\text{EA} - \epsilon_{\text{L}} + \Delta_{\text{vir}}$. We note that the application of the DFT+$\Sigma$ procedure is facilitated here, since the molecules 1a, 2, 1b, 1a·2 and 1b·2 in isolation are identical to those in the junctions, i.e., no hydrogen or other atoms are removed at the anchoring groups.
Figure S14. Calculated density of states (DOS) of the gas-phase molecules 1a, 2, 1b, 1a·2 and 1b·2 at the DFT level. The energy is measured with respect to the position of the HOMO orbital. The broadening in the DOS plot was set to $\eta = 82$ meV. For each molecule we present in addition the spatial distributions of HOMO and LUMO orbital wavefunctions. Isosurfaces of positive and negative isovalues are shown in red and blue, respectively.
Figure S15. Energy-dependent transmission function (a) and local density of states (LDOS) on four different parts of the molecule (b) for 1a•2.

References

