Supporting Information:

Can Electron Transport through a Blue-Copper Azurin be Coherent? An ab-initio Study

Carlos Romero-Muñiz,[†,‡,#] María Ortega,[‡,#] J. G. Vilhena,[¶] Ismael Díez-Pérez,[§] Rubén Pérez,[‡,∥] Juan Carlos Cuevas,[*,‡,∥] and Linda A. Zotti[*,⊥,‡,∥]

†Department of Physical, Chemical and Natural Systems, Universidad Pablo de Olavide, Ctra. Utrera Km. 1, E-41013 Seville, Spain

‡Departamento de Física Teórica de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain

¶Department of Physics, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

§Department of Chemistry, Faculty of Natural & Mathematical Sciences, King's College London, Britannia House, 7 Trinity Street, London SE1 1DB, UK

∥Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, E-28049 Madrid, Spain

⊥Departamento de Física Aplicada I, Escuela Politecnica Superior, Universidad de Sevilla, Seville E40011, Spain

#C. R.-M. and M. O. contributed equally to this work.

E-mail: juancarlos.cuevas@uam.es; lzotti@us.es

Contents

S1   Details about the S-S bridge

In Figure S1, we analyze the bridge between the two cysteines in close proximity to the gold surface.
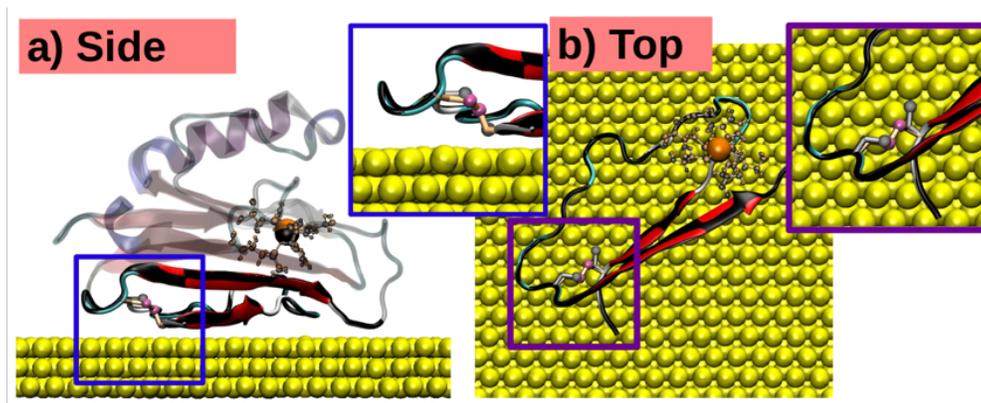

Figure S1: Comparison of the Azurin equilibrium configurations obtained upon adsorption when the S-S bridge is present (coloured in red-pink) and when it is not (coloured in black). From this figure, we conclude that the effect of the S-S bridge on the Azurin general structure upon adsorption is minimal. It is worth mentioning that the relative position of the cysteines on top of the gold substrate are in agreement with previous DFT calculations[S1] regardless the presence of the disulfide bond.

S2   Further details of transport calculations

As explained in the main manuscript, the calculation of the transmission function of the protein-based junctions is done in this work by making use of non-equilibrium Green's functions (NEGF) techniques. The approach used here differs slightly from previous ones that we have employed in the past. For this reason, we present in this section a detailed description of the transport method.

The starting point for the calculation of the transmission function is the Hamiltonian **H** obtained with DFT, as described in the main manuscript. Let us recall that in these calculations we consider a finite system comprising the protein (azurin), a gold substrate made of about 969 atoms (3 atomic layers) and a gold STM tip modeled with 252 atoms (7 atomic layers), see Fig. S2a. In order to use this input to simulate an infinite molecular junction and the necessary Green's functions for the determination of the transmission, we proceed as follows. First, and using the local character of our basis functions, we partition the basis states into L, C, and R ones, according to the division of the contact geometry,
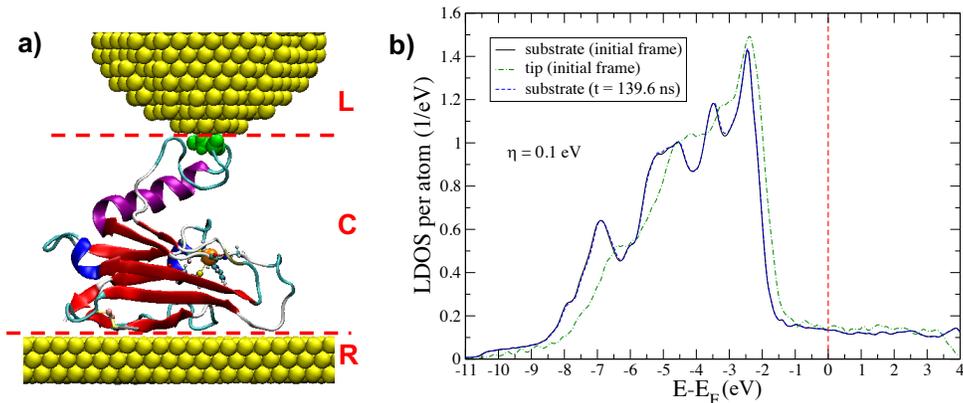

Figure S2: (a) Schematics of the partition of the system into L, C, and R parts for the transport calculations. (b) Lead density of states (DOS) as a function of energy (measured with respect to the Fermi energy E_F). The back solid line corresponding to the DOS of states in the gold substrate for the initial geometry in the blinking simulations, the green dashed-dotted line to the tip DOS of that geometry, and the dashed blue line to the gold substrate in another blinking geometry ($t = 139.6$ ns). The DOS in all cases has been normalized to the number of atoms in the corresponding lead and the broadening parameter used was $\eta = 0.1$ eV.

see Fig. S2a. Typically, the L part (or left lead) corresponds to the STM tip, the central part C to the protein (and eventually part of the electrodes), and the R part (or right lead) corresponds to the gold substrate. With this partition, the Hamiltonian (or single-particle Fock) matrix $\mathbf{H}$, as well as the overlap matrix $\mathbf{S}$, can be written in the block form

$$
\mathbf{H} = \begin{pmatrix} \mathbf{H}_{\mathrm{LL}} & \mathbf{H}_{\mathrm{LC}} & \mathbf{0} \\ \mathbf{H}_{\mathrm{CL}} & \mathbf{H}_{\mathrm{CC}} & \mathbf{H}_{\mathrm{CR}} \\ \mathbf{0} & \mathbf{H}_{\mathrm{RC}} & \mathbf{H}_{\mathrm{RR}} \end{pmatrix}.
\tag{1}
$$

Notice that here it is assumed that there is no direct coupling between the leads (L and R parts), something that we explicitly verified in every calculation. Using standard NEGF techniques, the energy-dependent transmission $\tau(E)$ can be expressed as[S2,S3]

$$
\tau(E) = \mathrm{Tr}\left[\mathbf{\Gamma}_{\mathrm{L}}\mathbf{G}_{\mathrm{CC}}^{\mathrm{r}}\mathbf{\Gamma}_{\mathrm{R}}\mathbf{G}_{\mathrm{CC}}^{\mathrm{a}}\right],
\tag{2}
$$

where the retarded Green's function is given by

$$
\mathbf{G}_{CC}^r(E) = \left[(E + i\eta_{\mathrm{C}})\mathbf{S}_{CC} - \mathbf{H}_{CC} - \mathbf{\Sigma}_L^r(E) - \mathbf{\Sigma}_R^r(E)\right]^{-1},
\tag{3}
$$

where η_{C} is an infinitesimally small positive real number and $\mathbf{G}_{\mathrm{CC}}^{\mathrm{a}} = \left[\mathbf{G}_{\mathrm{CC}}^{\mathrm{r}}\right]^{\dagger}$. The self-energies in the previous equation adopt the form

$$
\mathbf{\Sigma}_X^{\mathrm{r}}(E) = (\mathbf{H}_{\mathrm{C}X} - E\mathbf{S}_{\mathrm{C}X})\,\mathbf{g}_{XX}^{\mathrm{r}}(E)\,(\mathbf{H}_{X\mathrm{C}} - E\mathbf{S}_{X\mathrm{C}}).
\tag{4}
$$

On the other hand, the scattering rate matrices that enter the expression of the transmission are given by $\mathbf{\Gamma}_X(E) = -2\mathrm{Im}\left[\mathbf{\Sigma}_X^{\mathrm{r}}(E)\right]$. Finally, the functions $\mathbf{g}_{XX}^{\mathrm{r}}(E)$ are the electrode Green's functions with $X = \mathrm{L}, \mathrm{R}$. These electrode Green's functions describe the electronic structure of the (uncoupled) leads, which strictly speaking are infinite systems. In the past, to compute these Green's function we have modeled those functions as surface Green's functions

of ideal semi-infinite crystals, which can be computed with decimation techniques.[S3] In this work, and for computational convenience, we have followed a different approach and employed the so-called absorbing boundary conditions. In this approach the idea is simply to introduce a phenomenological broadening in the electrode Green's functions that are computed in practice with the Hamiltonians of the finite regions L and R defined in the partition above. This means that these functions are computed as follows

$$\mathbf{g}^{\mathrm{r}}_{XX}(E) = [(E + i\eta_X)\mathbf{S}_{XX} - \mathbf{H}_{XX}]^{-1},\tag{5}$$

where η_X is the phenomenological broadening parameter for $X = \mathrm{L}, \mathrm{R}$. Roughly speaking, this parameter should be on the order of the mean level spacing in our finite electrodes and, ideally, those electrode have to be big enough such that the choice of that parameter has not a strong influence in the electrodes density of states. In our calculations, we have found that a good compromise is $\eta_{\mathrm{L}} = \eta_{\mathrm{R}} = \eta = 0.1$ eV. In Fig. S2b we show a typical example for the density of states in both the substrate and the STM tip computed with the absorbing boundary conditions. Here, we can see that the description of the electronic structure of the leads is quite satisfactory and one can recognize all the expected features in the electronic structure of a gold system: (i) flat density of states around the Fermi energy due to the contribution of the s-orbitals, (ii) large density of states located a few eV below the Fermi energy due to the contribution of the d-orbitals, etc. Moreover, in that figure we also show that the density of states of the substrate computed with our method is basically identical in different junction realizations.

In the method described above, the partition of the system into L, C, and R is somewhat arbitrary and therefore, it is important to show that the computed transmission function does not depend on the details of this division. In all the calculations presented in this work, we have considered that the central part consists of the protein and the system is partitioned as shown in Fig. S3a, which we call "cut 0". To check the validity of this choice,
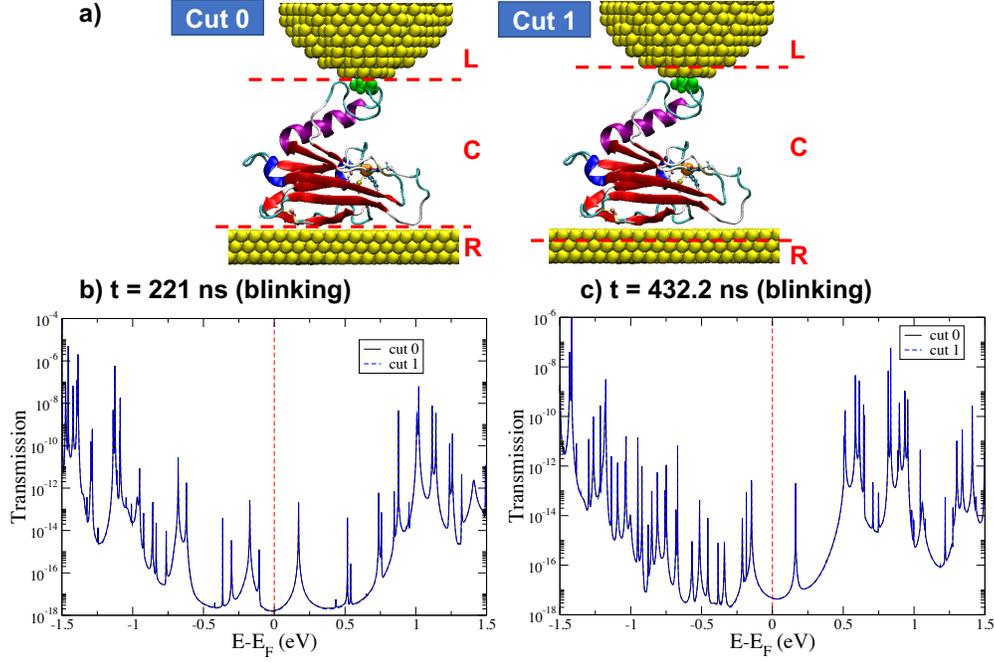
Figure S3: (a) Two different partitions of the protein junctions into L, C, and R parts. For "cut 0" (left), the central part only includes the protein, while in the "cut 1" (right) also the first atomic layers of the substrate and tip are included in the central part. (b) Transmission as a function of energy (measured with respect to the Fermi energy E_F) for a geometry obtained in the blinking simulations ($t = 221$ ns). The two curves correspond to the two types of partition described in panel a. The red vertical line indicates the position of the Fermi energy. (c) The same as in panel b, but for another geometry from the blinking simulations ($t = 432.2$ ns).

we have compared the results obtained in this way with those obtained by assuming that the central part also includes part of the electrodes. In Fig. S3b-c we show an example of such a comparison of the transmission functions obtained for two junctions from the blinking simulations using "cut 0" and a division in which the central part includes 1 atomic layer of both the substrate and the tip, which we call "cut 1", see right structure in Fig. S3a. As one can see, there is hardly any difference between the results of those two partitions, which constitutes a very stringent test of the validity of our whole approach.

S3   Lateral indentation MD Simulations

In order to obtain an atomistic understanding of the formation of a gold-azurin-gold junction when one of the two gold electrodes (the tip) approaches the protein sideways, we study this process with molecular dynamics (MD) simulations. The simulation protocol consisted of a two stage process: a) Free adsorption of the azurin to an Au(111) surface; b) Lateral indentation of the as adsorbed azurin with an Au(111) tip. In both cases: *i)* the energy of the structures was minimized to avoid steric clashes using a combination of steepest descent and conjugate gradient methods; *ii)* the pressure was stabilized using a 2 ns long NPT simulation (note that all simulations are performed in liquid medium); and finally *iii)* the production runs were performed in the NVT ensemble. Regarding the free adsorption stage, the protein is initially positioned at 1 nm from the Au(111) surface and is then allowed to freely adsorb to the surface in a 150 ns long NVT MD simulation (see Ref. S4 for more details) As for the lateral indentation stage, we start from the equilibrium configuration obtained from the previous stage (free adsorption). The gold tip (with the corresponding solvating water) is introduced at ~ 5.5 nm away from the azurin copper ion in the x direction, i.e. $d_{t-Cu} \sim 5.5$ nm (see Fig. S4a). Thus, the tip-protein interaction at the beginning of the simulation is avoided. Then, we start to reduce the tip-protein distance by translating the tip at a velocity of 0.05 m/s in the x direction, i.e. parallel to one of the gold surface axis. When $d_{t-Cu} \sim 2.5$ nm, i.e. the protein and the tip are only separated by one hydration layer, we reduce the tip velocity to 0.025 m/s, see Fig. S4a. Thus, the breaking of the azurin/tip first hydration layer (water with the smallest mobility) is better captured as we use further equilibration time for its description. Once the protein is firmly anchored to the tip ($d_{t-Cu} \sim$ 2 nm, see more details in the paragraph below), we reduce again the tip velocity up to 0.0125 m/s, see Fig. S4a, for a proper characterization of the protein stability on this junction. The simulation is continued up to $d_{t-Cu} \sim 1.4$ nm, see Fig. S4a. It is worth noting that the difference between the z coordinates of the tip bottommost layer and the surface topmost layer is initially set to $d_{t-s} \sim 1.9$ nm, and is held fixed during the whole

MD simulation.

From this simulation results, we can conclude that the tip-protein contact in the lateral indentation scenario is promoted via two different azurin regions: the α-helix arm and the azurin hydrophobic patch, see Figs. S4b-d. Concerning the first of these two regions, i.e. the α-helix arm, we observe that it is firmly anchored to the tip at $d_{t-Cu} \sim 2.3$ nm via the L68 and D69 amino acids (see Fig. S4b). This anchorage point is reinforced at $d_{t-Cu} \sim$
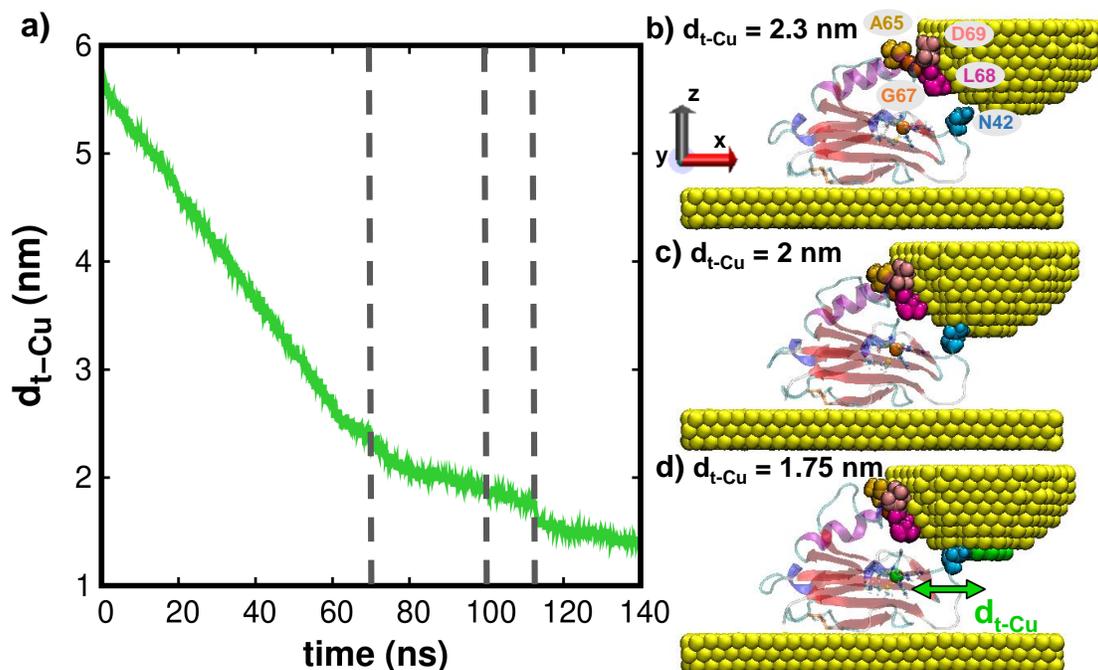

Figure S4: Description of the lateral indentation MD simulation. (a) Evolution of the distance between the tip and the azurin copper ion (d_{t-Cu}) during the simulation time. The d_{t-Cu} value is defined as the difference between the x coordinate of the bottommost tip layer geometric center and the copper ion (see schematic representation in (d)). (b-d) MD simulations snapshots at three representative azurin-tip configurations during the lateral indentation. The selected frames are highlighted via a grey dashed line in panel (a).The surface/tip Au atoms are represented as yellow spheres, and the protein is represented using its secondary structure (β-sheets in red, α-helix in purple, turns in cyan and random coils in gray). The five main anchoring amino acids in the protein-tip contact are represented accordingly to their role in the junction stability during the simulation, i.e. first anchoring residues (L68, D69) coloured in pink shades, α-helix residues (G67, A65) coloured in orange shades and the N42 residue, which belongs to the azurin hydrophobic patch[S4,S5] (near to the copper coordination sphere), coloured in cyan. Note that the water molecules were not represented for clarity purposes.

2 nm when other two amino acids of this region (A65,G67) contacts the tip, see Fig. S4c. Interestingly, this last protein-tip contact is accompanied by the lose of $\sim 3\%$ of the α-helix secondary structure content, see Fig. S5. Regarding the second anchorage point, i.e. the azurin hydrophobic patch, we observe that it contacts the tip at $d_{t-Cu} \sim 2$ nm, i.e before the $\alpha-$helix arm. The main anchoring amino-acid of this region is the N42, see Fig. S4c-d. It is worth noting that the N42 residue is very close to the copper coordination sphere. In fact, it is directly connected to this region via a short and flexible random coil, see Fig. S4d. This connection implies the inferring of a direct tension to the azurin cooper ion as a consequence of the N42-tip contact, which is translated into a sharp change of d_{t-Cu} when the N42 amino-acid slides over the tip at $d_{t-Cu} \sim 1.75$ nm, see Figs. S4a and S4d.

Finally, we compare the change of the azurin secondary structure content occurring during this simulation with the ones observed in the other two scenarios, i.e. indentation and blinking, see Fig. S5. More concretely, we focus on the evolution of the α-helix content, as it is the most influenced by the tip interaction in the three scenarios (the α-helix arm is in the three cases in contact with the tip). As is shown in Fig. S5, the α-helix content
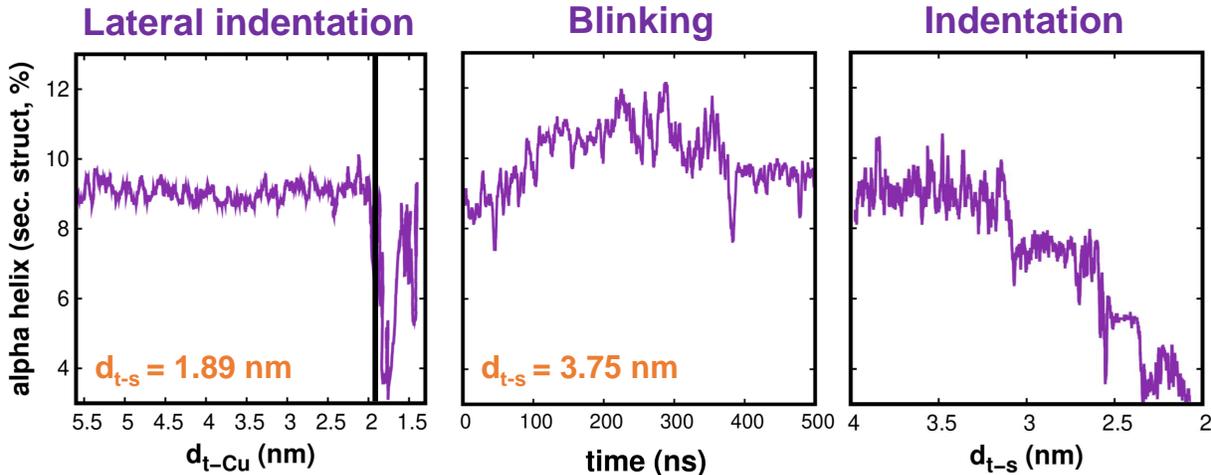

Figure S5: Evolution of the α-helix content during the MD simulation of the azurin junction formation for the three possible scenarios here analyzed, i.e lateral indentation (left panel), blinking (central panel) and top indentation (right panel). For the lateral indentation and blinking cases, where the distance between the tip and the surface (d_{t-s}) is fixed during the whole MD simulation, the corresponding d_{t-s} value is here specified.

does not decrease when the junction is formed due to a blinking process, but it significantly reduces in both the top and lateral indentation scenarios as the protein-tip distance decreases. However, the effect of the indentation over the secondary structure is much larger when the tip approaches the protein from the top (α-helix content decreases $\sim 6\%$) than when it approaches sideways (α-helix content changes a $\sim 3\%$). Considering that $d_{t-s} \sim 1.9$ nm during the whole lateral indentation simulation, which is even smaller that the minimum tip-surface distance achieved in the top indentation ($d_{t-s} \sim 2$ nm, see Fig. S5), this result manifests the lateral indentation as the best scenario for maintaining a close tip-surface distance without major altering the protein structure.

S4   Role of the Cu ion in the transport

It is generally believed that the transport through azurin proceeds through the Cu ion. We have indeed found in the analysis of the electronic structure of azurin-based junctions that the HOMO of the protein is usually the closest energy level to the Fermi energy of the junction.[S6] However, those studies also showed that the electronic states of the Cu ion are not very affected by the coupling to the gold electrodes.[S6,S7] This fact suggests a small hybridization between the Cu states and the metallic states of the leads, and let us recall that this hybridization is a necessary prerequisite for the Cu to participate in the charge transport. Therefore, it is not entirely clear a priori whether the Cu ion plays an important role in the coherent transport discussed in this work.
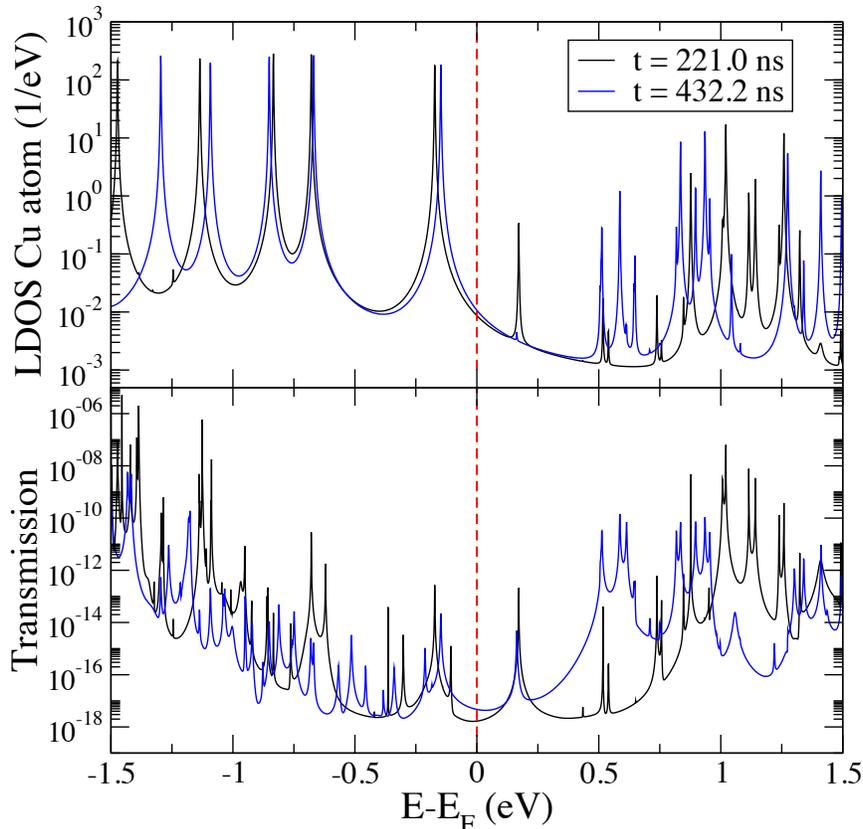

Figure S6: LDOS on the Cu ion (upper panel) and transmission function (lower panel) as a function of energy. The two curves correspond to two different snapshots in the blinking simulations.

In order to shed light on this issue we have studied the local density of states (LDOS) on the Cu ion and correlated that with the transmission functions in the different situations that we have investigated here. In Fig. S6 we show an example of this correlation for two representative geometries obtained during the blinking process. The relation between the energy dependences of both the LDOS on Cu and the junction transmission suggests that this ion plays indeed an important role on the transport. In particular, we typically find that, like in these cases, the highest occupied state in the Cu ion is the HOMO (and the closest level to the Fermi energy) in the molecular junction. In any case, the small hybridization between the Cu states and the metallic leads is apparent in the small width of the LDOS peaks, which is ultimately the reason for the very low conductances that we find in the
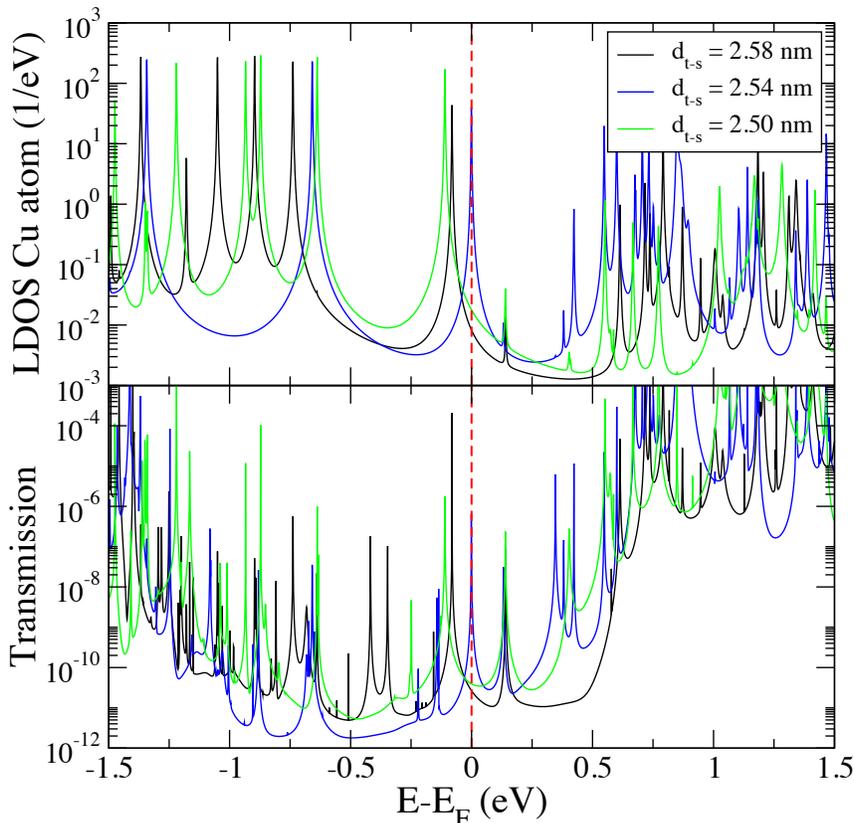

Figure S7: LDOS on the Cu ion (upper panel) and transmission function (lower panel) as a function of energy. The three curves correspond to different geometries obtained during the simulations of the indentation process of the STM tip.
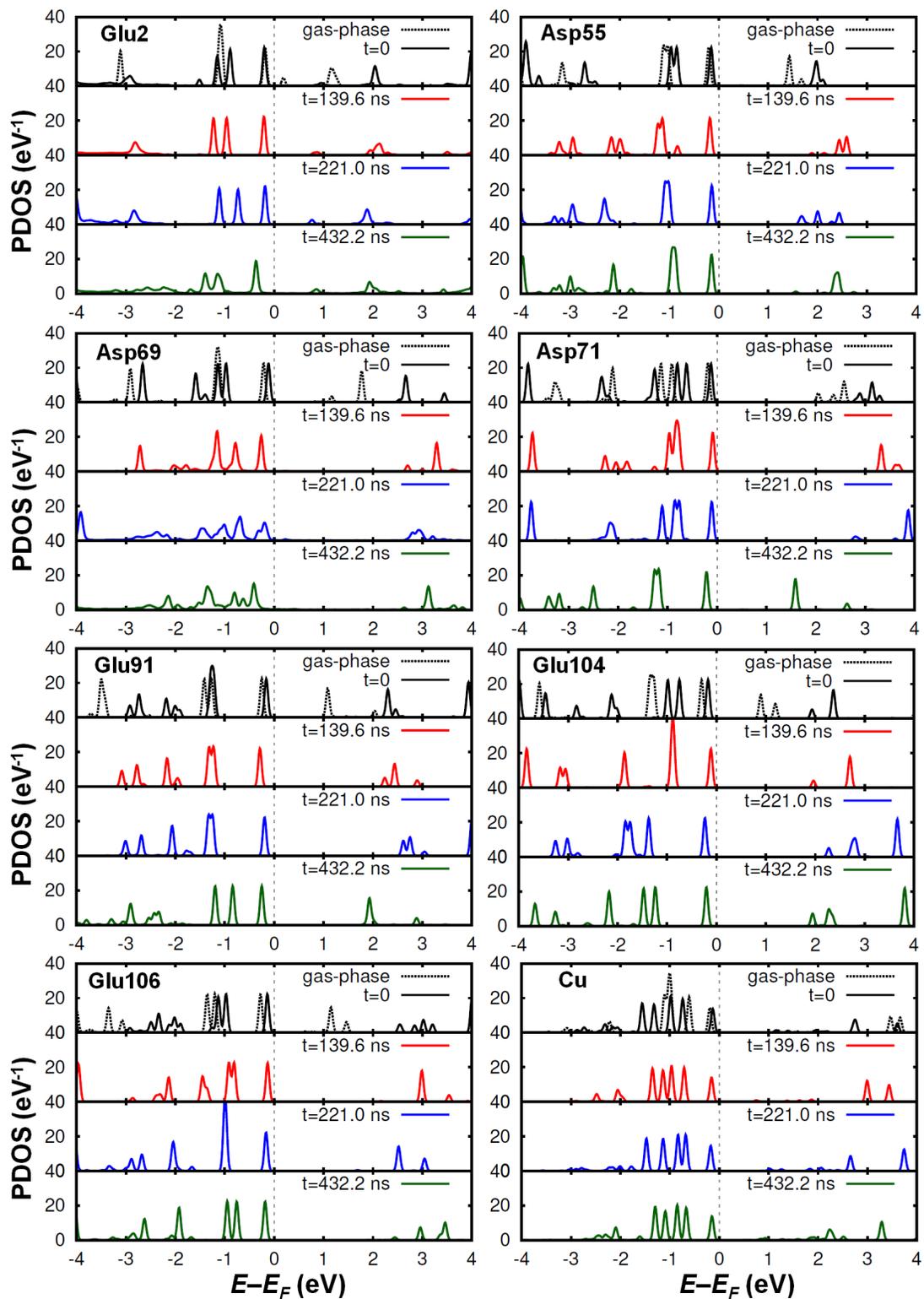
Figure S8: PDOS on selected residues for the four time frames of the blinking events analyzed in Figure 1 of the main text.

blinking process.

For completeness, we also show several results for the Cu LDOS and its relation to the transmission function in the the simulations of the tip indentation, see Fig. S7. In particular, we have chosen three of the geometries with the highest transmission values at the Fermi energy. As in the blinking case, the highest occupied Cu state seems to determine the transmission at the Fermi energy. In the case of the geometry with a tip-sample distance of 2.54 nm, the extraordinary transmission at the Fermi energy, as compared to the other geometries, is due to the fact that highest occupied Cu states lies practically at the Fermi energy.

It is worth noting that other residues than the Cu ion contribute to states aligned very close to the Fermi level. In Fig. S8 we show the projected density of states for these moieties throughout the blinking process described in Fig. 1 of the main text.

S5 Indentation

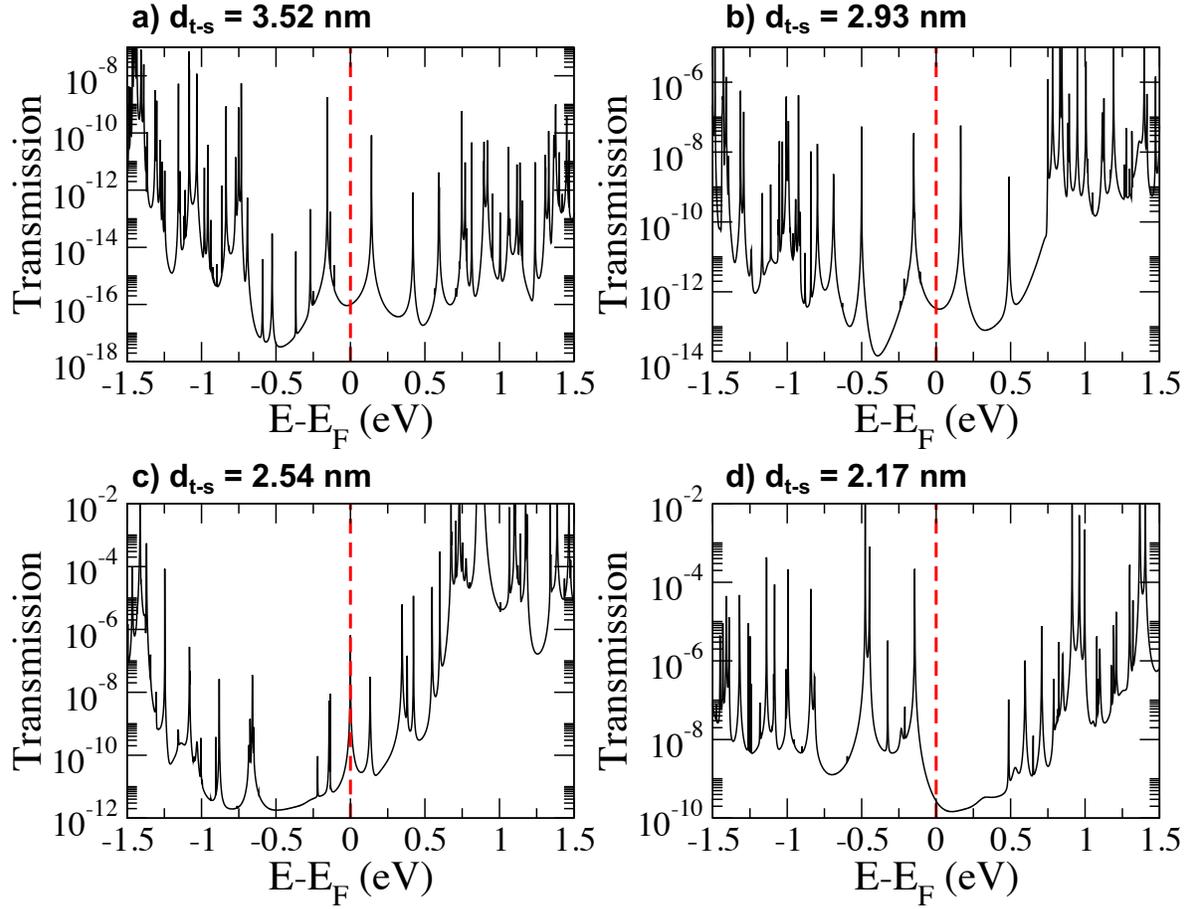In Fig. S9 we show some of the transmission curves obtained for the simulation of the indentation.

Figure S9: Transmission curves for the geometries obtained during the indentation process.

S6 Modeling the tip in the lateral indentation

In our DFT calculations, we have used a tip containing the same number of atoms for all cases. It was obtained, as explained in the main text, from the much larger tip employed in the MD simulation upon removal of the outermost layers. However, in the case of the lateral indentation, such a modification might in principle remove interaction between certain gold and protein atoms in the uppermost part of the junction. This could perhaps be a problem, especially for the geometries with the shortest protein-gold distance. To verify what the impact of this is on the transmission, we have compared the result given by using this kind of tip with that obtained by selecting the tip atoms in a different way, namely by applying a diagonal cut as shown in Fig. S10. The corresponding transmission curves do not show significant differences, therefore justifying our choice.
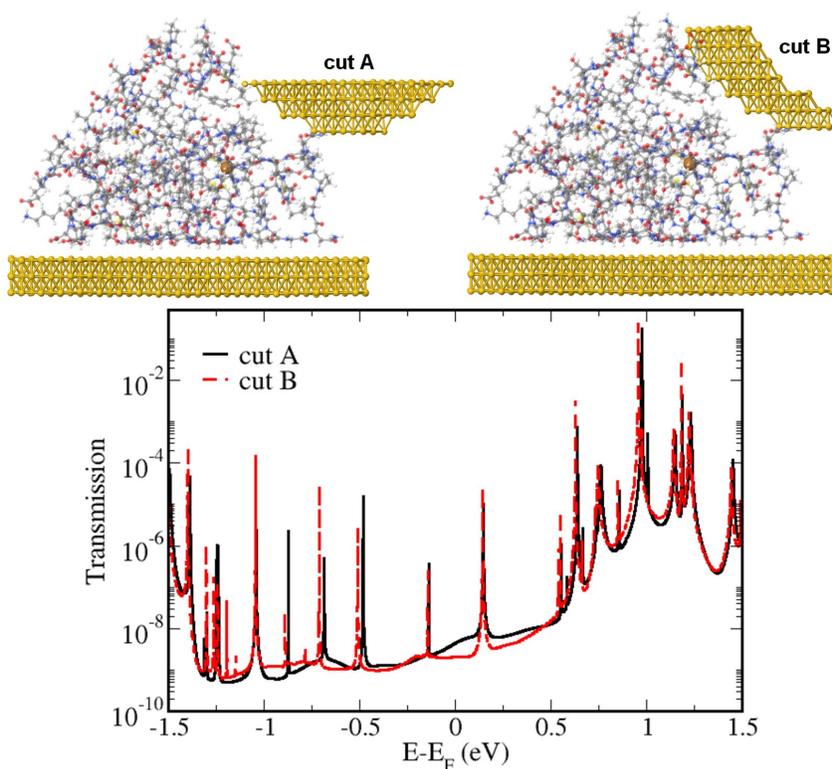

Figure S10: Comparison between the two tips tested for the lateral indentation.

References

(S1) Di Felice, R.; Selloni, A.; Molinari, E. DFT Study of Cysteine Adsorption on Au(111). *J. Phys. Chem. B* **2003**, *107*, 1151–1156.

(S2) Cuevas, J. C.; Scheer, E. *Molecular Electronics: An Introduction to Theory and Experiment*, 2nd ed.; World Scientific, 2017.

(S3) Pauly, F.; Viljas, J. K.; Huniar, U.; Häfner, M.; Wohlthat, S.; Bürkle, M.; Cuevas, J. C.; Schön, G. Cluster-Based Density-Functional Approach to Quantum Transport through Molecular and Atomic Contacts. *New J. Phys.* **2008**, *10*, 125019.

(S4) Ortega, M.; Vilhena, J. G.; Zotti, L. A.; Díez-Pérez, I.; Cuevas, J. C.; Pérez, R. Tuning Structure and Dynamics of Blue Copper Azurin Junctions via Single Amino-Acid Mutations. *Biomolecules* **2019**, *9*, 611.

(S5) Ruiz, M. P.; Aragonès, A. C.; Camarero, N.; Vilhena, J. G.; Ortega, M.; Zotti, L. A.; Pérez, R.; Cuevas, J. C.; Gorostiza, P.; Díez-Pérez, I. Bioengineering a Single-Protein Junction. *J. Am. Chem. Soc.* **2017**, *139*, 15337–15346.

(S6) Romero-Muñiz, C.; Ortega, M.; Vilhena, J. G.; Diéz-Pérez, I.; Cuevas, J. C.; Pérez, R.; Zotti, L. A. Mechanical Deformation and Electronic Structure of a Blue Copper Azurin in a Solid-State Junction. *Biomolecules* **2019**, *9*, 506.

(S7) Romero-Muñiz, C.; Ortega, M.; Vilhena, J. G.; Díez-Pérez, I.; Cuevas, J. C.; Pérez, R.; Zotti, L. A. Ab Initio Electronic Structure Calculations of Entire Blue Copper Azurins. *Phys. Chem. Chem. Phys.* **2018**, *20*, 30392–30402.