Peptide bonds detection via graphene nanogaps: a proof of principles study

Aldo Eugenio Rossini[a], Fabrizio Gala[a], Mauro Chinappi [b] and Giuseppe Zollo[a*]

Solid-state nanopores and nanogaps are emerging as promising tools for single molecule analysis. 2D materials, such as graphene, can potentially reach the spatial resolution needed for nucleic acid and protein sequencing. In the context of the density functional theory atomistic modeling and non equilibrium Green function calculation, we show that glicyne based polypeptide chains translocating across a nano-gap between two semi-infinite graphene nano-ribbons, leaves a specific transverse current signature for each peptide bond. Projected density of states and bond current analyses reveal a complex scenario with a role played by the adjacent α-carbons and side chains and by the the orbitals of the partially resonant double bond involving C, N and O atoms of the peptide bond. In this context, specific fingerprints of the atoms involved in the peptide bonds are found. The same scenario is evidenced also for peptides involving also alanine residues. The signal measured can be considered as a specific fingerprint of peptide bonds between small and neutral aminoacids with no polar/charge effects. On this basis, a newly conceived nano-device made of a graphene based array of nano-gap is proposed as a possible route to approach peptide sequencing with atomic resolution.

1 Introduction

Single-molecule techniques for biomolecule characterization are of critical importance for both fundamental biology and biomedical application. Nanoscale pores and gaps on biological and solid-state substrates are emerging as enabling technologies to achieve reliable and fast protocols for various biochemical analyses such as DNA sequencing[1,2], protein sequencing[3–5], protein and peptide folding[6–10] and detection of post-translational modification[11,12]. In this framework, graphene is currently considered for single molecule sensing of protein and nucleic acids for both resistive pulse[4,13–15] and transverse current[16–19] protocols and a number of different device configurations have been presented in literature. In a typical set-up, a transverse voltage is applied between two electrodes laying on the membrane plane where a single pore (or gap) is present[20]. When a portion of a molecule occupies the space between the gap, the transverse current between the electrode is altered. The intensity of the transverse current depends on the electronic structure of the molecule portion in the gap, allowing, in principle, to characterize the translocating molecules. Experimental studies have shown that approaches based on transverse current modulation are able to distinguish among the different DNA-bases[21,22] (proofing the sequencing protocol originally presented in[23]), and to detect amino acids (AAs) and post-translational modifications[11]. However still much work is needed to solve practical problems and improve the signal-to-noise ratio[24]. Concerning modeling, first principles atomistic simulations based on the density functional theory[25,26] (DFT) are nowadays state of the art techniques to obtain reliable predictions concerning the interactions between surfaces or nano-structures with AAs[27,28] which is the premise for possible future paths to get a full recognition of such important biological molecules as proteins.

In this paper we show, using state of the art atomistic modeling and theoretical calculations, that it is possible, in principle, to detect single peptide bonds (PBs) of a polypeptide chain translocating across a suitable array of Graphene Nano-Ribbon (GNR) nano-gaps by measuring the transverse electronic current flowing across the nano-gap. The presented results pave the way for developing newly conceived nano-devices for AAs characterization and protein sequencing by means of GNR based nano-devices. A sketch of the proposed nano-device is schematically shown in Fig. 1: the device is constituted by an array of GNRs deposited onto a suitable solid substrate where a sub-nanometer nearly rectangular pore is fabricated exploiting the present capability of nanotechnology manufacturing techniques[3]. In particular, in the recent experimental literature there are examples demonstrating that the exceptional degree of control at the atomistic scale required by the proposed device is not far. Indeed the present nanotechnology allow the realization of graphene nano-gaps down to the sub-nanometer size[11,24,29–31] and the synthesis, by bottom-up strategies, of narrow GNRs[32,33] of similar size as the one here supposed (but slightly wider GNRs could be employed as well). However the needed full atomistic control

[a] *Dipartimento di Scienze di Base e Applicate per l'Ingegneria (Sezione di Fisica), Università di Roma "La Sapienza", Via A. Scarpa 14–16, 00161 Rome, Italy.*
[b] *Dipartimento di Ingegneria Industriale, Università di Roma Tor Vergata, Via Del Politecnico 1,00133 Roma*
* email address: giuseppe.zollo@uniroma1.it.

at the sub-nanometer length scale required to build such a device is still challenging presently. Moreover, sub-nanometer gaps might be problematic for GNR electrodes that requires a further control of the bias[31]. For such a device, the rectangular shape of the pore imposes preferential AAs configurations during translocation with the PBs chain at the center and the AAs side chains along the long side of the rectangular pore. Hence this implies that α-carbons and the PBs occupy the central nano-gap while the AAs side chains the adjacent ones so that, in principle, various transverse tunneling current signals could be collected for large residues while PBs and small residues signals could be measured at the central nano-gap. With the present geometry and the usage of GNRs, then, the most limiting factors of nano-pores in detecting AAs, factors mainly related to the range of their size, are overcame. Indeed, differently from the DNA bases, the AAs side chain size d_{sc} spans a large range of values, from $d_{sc} \simeq 1.1$ Å for glycine (Gly) to $d_{sc} \simeq 8.3$ Å for arginine, which makes nanopores unsuitable for AAs detection. Another important aspect of the present device is that the signal dark current is zero (see below the transmission function) thus increasing the signal to noise ratio with respect to a possible AAs detection by graphene nanopores via transverse electronic current.

In the present study, as a fundamental step in the assessment of the proposed device, we have focused on the PB recognition in a Gly homo-peptide by a single GNR. The current signals coming from the PBs between Glycines are affected by a very small AA side chain (Gly side chain is just an hydrogen atom) and are unaffected by nearby charge sources (as in the case of charged or polar AAs). Hence, the signal here studied can be considered as the most favorable case to evidence specific PB fingerprints. The results on the Gly homo-peptide are tested on Ala homo-peptide and Ala-Gly peptides with alternating Ala and Gly residues.

2 Theoretical methods

The atomistic model we adopted is basically made of two semi-infinite graphene nano-ribbons with zig-zag edges (ZGNR) and the peptide that translocates across the nano-gap (see Fig.1(b)). As already mentioned, most of the present study is focussed on a homo-peptide with three Glycines. The chosen ZGNR is characterized by a gap in the transmission function around the Fermi energy with zero dark-current[34]. The idea is to exploit the ZGNR conductance properties to detect current signals from the peptide bonds and the amino acids side chains that might be processed until recognition of the amino acids or, at least, of the AA type such as, f.i. polar, neutral, charged and so on.

Following the scheme of Fig. 1, in our calculations each lead has been modeled as a semi-infinite symmetric graphene nanoribbon with zig-zag edges (ZGNR), 5.0 Å wide and lying in the z direction. The nano-gap has a hydrogenated armchair pattern at the border and the distance between the two semi-infinite ZGNRs is 5.0 Å (distance between H atoms).The amino acids configurations crossing the gap have been identified using a NVT classical steered molecular dynamics (SMD) simulations[35] at $T = 300$ K in water ambient. Then, conformation along the translocation pathway have been selected, water molecules have been removed and the ionic degrees of freedom of the system have been relaxed at $T = 0\ K$, in the context of the density functional theory (DFT) in order to approach a local energy minimum.

The total energy DFT calculations scheme includes a generalized gradient approximation based on the Perdew-Burke-Ernzerhof formula[36](PBE) for the electron exchange and correlation potential $V_{xc}[n(\mathbf{r})]$ and norm-conserving pseudopotentials built with the Troullier-Martins scheme[37] in the framework of a plane-wave basis set expansion. DFT calculations have been performed using the QUANTUM ESPRESSO package[38], with a plane-wave energy cutoff of 70 Ry for the wave functions and an artificial smearing of 0.005 Ry, while structural optimization has been achieved in a cubic box of 14.7x14.7x36.16 Å^3 using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method[39] together with the Hellmann-Feynman forces acting on the ions; an empirical dispersion force field[40] has been included to handle the long range interactions. The chosen box size is large enough to prevent any spurious interaction arising from the periodic boundary conditions (PBC), thus calculations have been performed using the Γ point for the Brillouin zone sampling, and the ionic minimization has been carried on until the convergence threshold of 0.001 a.u. for the total force was achieved.

Quantum transport calculations have been performed using the non-equilibrium Green's function (NEGF) method as implemented in the TRANSIESTA code[41,42]. Again TM norm-conserving pseudopotentials have been employed, together with a DZP localized basis set. A 1x1x50 Monkhorst-Pack[43] k-grid was used for electrodes. The electronic temperature was set to 0 K. The transport properties of the ZGNR junction have been evaluated by means of the Hamiltonian decomposition:

$$H = \begin{pmatrix} H_L + \Sigma_L & V_L & 0 \\ V_L^\dagger & H_S & V_R \\ 0 & V_R^\dagger & H_R + \Sigma_R \end{pmatrix} \qquad (1)$$

where, $H_{L(R)}$ and H_S are, respectively, the decoupled Hamiltonians of left (right) electrode and the scattering region, $V_{L(R)}$ represents the interaction of left (right) electrode with the scattering region and $\Sigma_{L(R)}$ is the self-energy and describes the coupling of the left (right) lead with the scattering region.

For an external bias voltage V applied along the z direction,

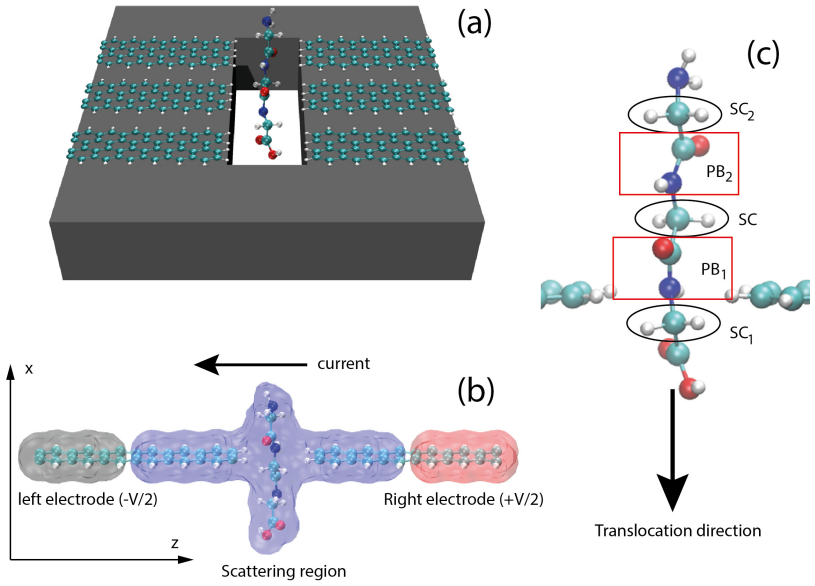

Fig. 1 The proposed device is constituted by an array of GNR nano-gaps deposited onto a substrate with a nanopore (a). Tunneling current signals could be collected between the electrodes depending on the size and chemistry of the AAs. The central nano-gap should be devoted to collect the PB related signals. The atomistic model adopted in the present work to calculate the current using the NEGF linear approach is reported in (b) while a zoom of the scattering region is in panel (c). Only the central GNR (without any additional substrate) is simulated. C, O, N atoms are respectively in cyan, red and blue. The gray, red and blue haloes represent respectively the left and the right electrodes and the central scattering region.

the current $I(V)$ is given by Landauer-Büttiker formula[44]

$$I(V) = \frac{2e}{h}\int_{-\infty}^{+\infty} d\varepsilon\, T(\varepsilon,V)\times[f(\varepsilon-\mu_L)-f(\varepsilon-\mu_R)] \qquad (2)$$

where $f(\varepsilon)$ is the Fermi-Dirac distribution function and $\mu_{L(R)}$ is the electrochemical potential of the left(right) electrode. $T(\varepsilon,V)$ is the transmission coefficient that is calculated from the Hamiltonian of Eq. (1), with the external bias voltage V included in the interaction terms, and is defined as:

$$T(\varepsilon) = \mathrm{Tr}\left[G(\varepsilon)\Gamma_L(\varepsilon)G^{\dagger}(\varepsilon)\Gamma_R(\varepsilon)\right] \qquad (3)$$

where $G(\varepsilon) = \lim_{\eta\to 0^+}(\varepsilon + i\eta - H)^{-1}$ is the Green's function of the system and $\Gamma_{L(R)}(\varepsilon) = i[\Sigma_{L(R)}(\varepsilon) - \Sigma^{\dagger}_{L(R)}(\varepsilon)]$ is the left(right) coupling function[45].

Although DFT is a ground state theory, not a steady state one, the above scheme, known as DFT-NEGF, is nowadays the most popular approach for steady-state transport in nanostructure and has been successfully applied in many cases. However, other schemes, that are formally correct, have been developed in the last years to encounter for discrepancies with the experimental results[46–48]. In any case, DFT-NEGF always provide a correct qualitative insight on the transport behavior, and often give current values close to the ones obtained with formally correct methods[48,49]. DFT and quantum transport calculations have been carried out in dry ambient, since the conductance of ZGNR nanopore does not vary significantly when the dynamical environment of water is removed from the system[50] (see Supplementary Sec. S1) due to the hydrophobic character of graphene based structures. Moreover, because of the narrow size of the ZGNR nano-gap, water molecules cannot be allocated between the leads together with the PBs but are expected to cross the nano-device near the lateral ZGNR nano-gaps (see Fig. 1) where the AAs side chain move across the device.

3 Results

3.1 Gly homo-peptide

The current signals have been measured on a piece of peptide with two bonds between three Glycines for an applied voltage $\Delta V = 1V$ between the leads. The present bias has been employed both in the theoretical and in the experimental literature concerning tunneling currents or molecular junctions[51–53] but, of course, it should be tested for breakdown in real systems, especially for GNR electrodes[31]. In any case we have not observed any density or wave-functions anomalous behavior up to a bias of 2 V. Because Gly side chain is a single hydrogen atom, the AAs side chain effects on the PB current signal should be, in principle, minimized; this circumstance, together

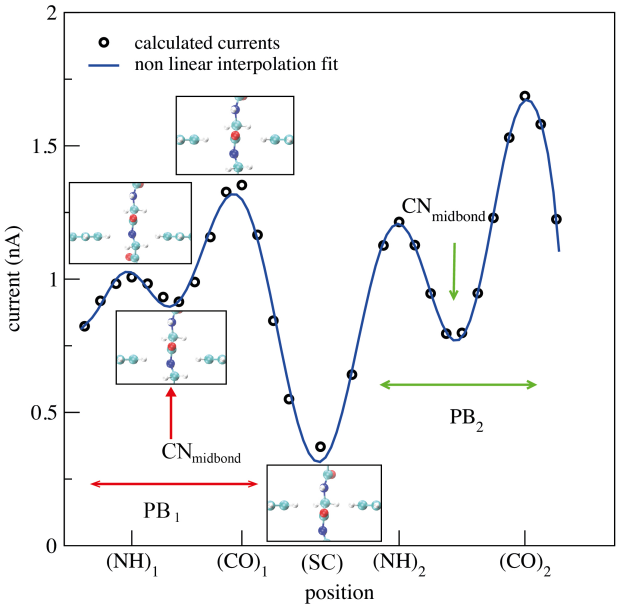

Fig. 2 Electron current measured in the ZGNR based nano-gap device while a Gly peptide with three AAs translocates across the nano-gap.

with the neutral and the non-polar nature of Gly, makes the present one a reference case of the PBs signal. In Fig.2, the current signal across the GNR nano-gap is reported as a function of the position of the peptide inside the nano-gap. Each peptide bond appears with two, well defined, current peaks. The calculated signals have been obtained in quasi-static conditions, i.e. with an idealized quasi-stationary translocation dynamics where the atoms are allowed to relax at $T = 0\ K$ at each translocation stage, see methods. The electronic current order of magnitude is comparable with typical tunneling current signals measured in nano-gap devices[11,18,24,54,55] but with the advantage of an atomic resolution.

The two peaks correspond approximately to the two positions of, respectively, the N and the O atoms of the peptide bond in the nano-gap and they are separated by a shallow current drop associated with the presence of the C-N bond in the nano-gap. In the following these configurations will be indicated as $(CO)_i$, $(NH)_i$ and $(CN)_i$, where the subscript indicates the peptide bond, PB_1 or PB_2. The signal in between two peptide bonds, i.e. with the Gly side chain in the nano-gap, is much smaller.

For comparison, we have calculated the current for the optimized configurations of one water molecule located in the gap. The resulting current is one order of magnitude smaller than maximum current signal obtained for the peptide and well below the absolute minimum at the *SC* configuration (Supplementary Sec. S1) so that the PB signals are separable from noise due to the water molecules that might eventually cross

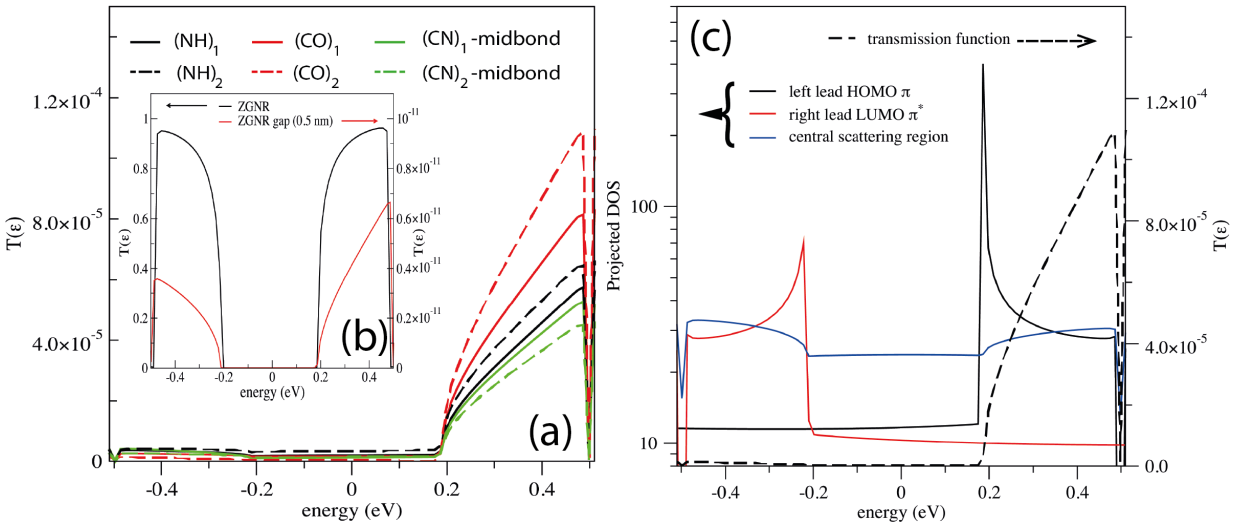

Fig. 3 Transmission coefficients at different positions during the translocation of the peptide with three Gly (a); the continuous and dashed curves refer to the first and the second PB respectively. The inset (b) shows the transmission functions of the ZGNR and of the void nano-gap. In (c) are reported the transmission function calculated with the nitrogen atom of the PB in the gap and the PDOS of the left and right leads and of the central scattering region.

the gap.

3.1.1 Transmission coefficients

To explain the observed currents, we have analyzed the transmission coefficient functions $T(\varepsilon)$ corresponding to (CO), (NH) and (CN) configurations. The data are reported in Fig. 3(a) where the continuous (dashed) curves refer to the PB_1 (PB_2).

The transmission coefficient curves show that the contribution to the current comes from the states in the energy range $[0.2\ \text{eV} \leq \varepsilon \leq 0.5\ \text{eV}]$; moreover, according to the current values of Fig. 2, for each PB the largest $T(\varepsilon)$ values are obtained for $(CO)_\text{i}$ configurations (the CO group is the ZGNR gap), the second largest occur for $(NH)_\text{i}$ configurations and the smallest ones are obtained with the *CN* mid-bond in the middle point of the ZGNR gap. For comparison, we reported in Fig. 3(b) the transmission functions of the empty nano-gap showing that no transmission occurs in this case. Moreover, the projected density of states (PDOS), Fig. 3(c), evidences that the rising of $T(\varepsilon)$ occurs in correspondence with the rising of the PDOS belonging to the scattering central region and to the left hand side ZGNR HOMO π-like states, provided that the PDOS of the right hand side ZGNR LUMO π^*-like states are not zero. The channel decomposition of the transmission coefficient[56,57] calculated at $V = 1\ V$ indicates the existence of just one channel contributing to the whole transmission coefficient, independently on the configuration considered with either the *CO* bond or the *NH* bond at the center of the GNR gap.

3.1.2 Projected density of states

To evidence the main contributions to the transmission functions at the current peaks, the atomic PDOS decomposition in the localized basis set of the central scattering region have been grouped following the scheme of Fig.1(c); the main features are reported in Fig. 4 for the two PBs. For the $(NH)_\text{i}$ configurations, the largest density comes from the the Gly α-carbon and side chain that are the closest ones at the C-terminal side (SC_1 and SC for PB_1 and PB_2). We will indicate these groups as "lower" since they are just below $(NH)_\text{i}$ in the snapshot in figure 1. Minor contributions to the PDOS of the scattering region come from the PB_i itself (with similar NH_i and CO_i PDOS values) and the closest side chain group at the N-terminal side (indicated as "upper" side chain in the following), namely *SC* and SC_2 for PB_1 and PB_2 (see Fig. 1(c)). At the second peak occurring with CO_i in the nano-gap, the major contribution comes from the "upper" side chain (i.e. the closest one). Concerning the PB-related PDOS, the major contribution comes from the NH_i group, while the one from the CO_i group is always lower in the entire energy range. A more detailed analysis of the DOS projected on the PBs atomic orbitals evidences the leading role played by the p_z orbitals of N and C (see Supplementary Informations Sec. S2).

3.1.3 Bond currents

A deeper insight on the local transport phenomena occurring at the atomistic level can be achieved looking at the local currents flowing across the atomic sites[58–60]. Given a suitable

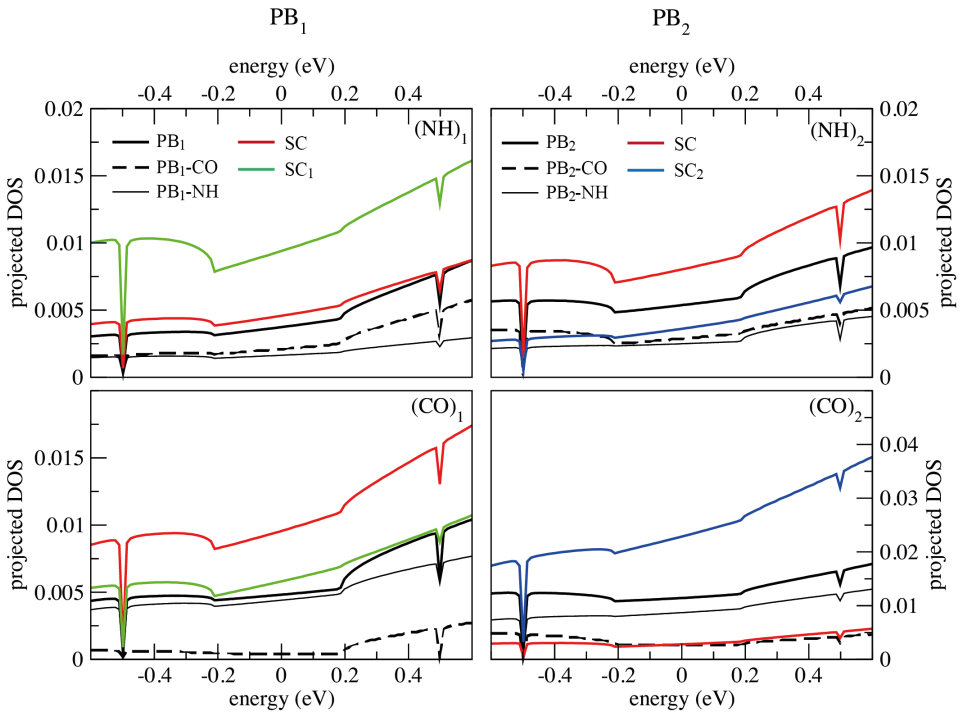

Fig. 4 Atomic PDOS decomposition corresponding to the four current maxima measured in the three Gly peptide. The cases with the *NH* and *CO* groups in the middle of the nano-gap for each of the two PBs are aligned in columns (with the same legend). The atomic PDOS have been grouped together for atoms belonging to the the same unit. The grouping units are evidenced in Fig. 1(c).

complete localized basis set $\{|\phi_{n,\gamma}\rangle\}$ in a subspace of the full device region where $|\phi_{n,\gamma}\rangle$ is an atomic orbital of type γ, centered at the atomic site n, the local current flowing between atoms n and n' is[59]:

$$I_{n,n'} = e\int d\varepsilon\,[f_L(\varepsilon) - f_R(\varepsilon)]\,Tr\left[G(\varepsilon)\Gamma_L(\varepsilon)G^{\dagger}(\varepsilon)J_{n,n'}\right] \quad (4)$$

with

$$f_L(\varepsilon) = f(\varepsilon - \mu_L);\;\; f_R(\varepsilon) = f(\varepsilon - \mu_R)$$

and

$$J_{n,n'} = \frac{1}{i\hbar}\sum_{n'\neq n}\left(P_n H P_{n'} - P_{n'} H P_n\right) \quad (5)$$

and the projection operator on the n site $P_n = \sum_\gamma |\phi_{n,\gamma}\rangle\langle\phi_{n,\gamma}|$.

The main atomic currents at the left lead (electrons injection), calculated in the energy range $[-0.5\,eV \leq \varepsilon \leq 0.5\,eV]$ with the INELASTICA package[61], are reported in Table 1; the resulting scenario is sketched in Fig. 5 where the most relevant local atomic currents are represented by arrows for $(NH)_i$ (a) and $(CO)_i$ (b) configurations.

$(NH)_i$ configurations. Consistently with the corresponding PDOS, the bond current analysis confirms that the main contributions for configurations $(NH)_i$ come from the "lower" side chain groups, i.e. the closest ones at the C-terminal side,

Table 1 Main local atomic current contributions to the whole current at the left lead (electron injection from the left lead). As mentioned in the text, the subscript $i = 1,2$ indicate the first or the second PB. In parentheses the groups and the corresponding local currents calculated for PB_2.

configuration	atomic group	electron injection into the left lead bond current (nA)
$(NH)_i$	SC_1 (SC)	0.67 (0.857)
	$(NH)_i$	0.045 (0.152)
$(CO)_i$	SC (SC_2)	0.63 (0.773)
	$(NH)_i$	0.44 (0.529)
SC	$(NH)_1$	0.114
	SC	-0.044
	$(NH)_2$	0.146

namely SC_1 and SC for PB_1 and PB_2 respectively. Indeed, since these groups are off the ZGNR plane but close to it, the overlap of the corresponding atomic orbitals with the left lead π orbital is favored. Much minor contributions come from the upper residues and from the PBs themselves, the last ones, however, behaving in a rather complex way. Indeed a detailed analysis reveals that the *CO* group acts as reflectors while the *NH* groups may participate to the current transfer at a very small extent through the N_{p_z} orbital (Supplementary

Informations Sec S3).

$(CO)_i$ configurations. The same analysis of the second current peaks evidences that electrons are injected from the left lead mainly through both the "upper" α-carbon and side chain, i.e. the closest to the *CO* group at the N-terminal side, and, remarkably, through the *NH* group of the PB at almost the same extent as the "upper" residue. Then, what makes the difference between the two peaks in a given PB signal is the contribution from the *NH* group that is rather small in the case of the first peak (*NH* group in the gap) while is almost of the same magnitude as the side chain contribution in the case of the second peak. This is mainly due to the off-plane position of the *NH* group in this second case that favors the overlap and the hybridization between the N_{p_z} orbital and the π orbital of the left lead. A complete and detailed analysis of the atomic currents at both the leads reveals, moreover, the occurrence of charge transfer phenomena inside the peptide, the reflection of electrons at the *CO* sites and minor, but still meaningful, contributions to the current from the second nearest residues (see Supplementary Informations Sec. S3).

As a matter of fact, the atomic current analysis reveals that the main reason why the second peak of the PB fingerprint is larger than the first one resides in the larger contribution from the *NH* group of the PB itself. This behavior is related to the nature of the PB resonating double bond with partial occupation of the N_{p_z} orbital that, as a consequence, can participate to the electron transfer, provided it is off the ZGNR plane.

Finally, as a last comment to this part of the discussion, we emphasize that the difference between the second and the first current peaks in each PB is a clear fingerprint of the PB itself because it is due to the electrons flowing only across the *NH* group of the PB.

SC configuration. Let's now consider the configuration with the central Gly side-chain in the gap (*SC* configuration-see Fig. 1) where the current signal drops. The corresponding PDOS and schematic drawing of the current fluxes are reported in Fig. 6. Here we see that PB_1 has the largest PDOS followed by PB_2 while the side chain contribution is small in the whole energy range. The main atomic current fluxes at the left lead are reported in Table 1. Electrons are injected from the left lead mainly through the closest $(NH)_2$ group, that, however, does not show the largest PDOS (see Fig.6(a)), and $(NH)_1$ of PB_1 at nearly the same extent. Contrarily to the configurations of two current peaks, the central side chain *SC* behaves as a reflecting group. A more detailed analysis shows that a meaningful charge redistribution between the two PBs occurs after injection and also that *CO* groups may participate to the current flow at a minor, but still meaningful, extent (see Supplemetary Informations). Therefore, because the Gly residue does not participate to the current transfer, the measured signal is a fingerprint of the two PB involved only, with a

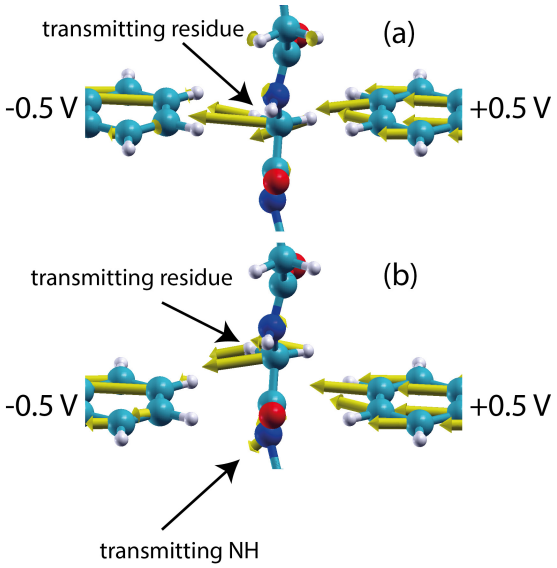

Fig. 5 Schematic drawing of the atomic current flows with the *NH* (a) and the *CO* (b) groups in the middle of the nano-gap. C, O, N atoms are respectively in cyan, red and blu and atomic currents are yellow arrows. Transmission occurs via the "upper" and the "lower" side chains and, in the case of the *CO* group in the gap (b) also via the nitrogen atom of the PB.

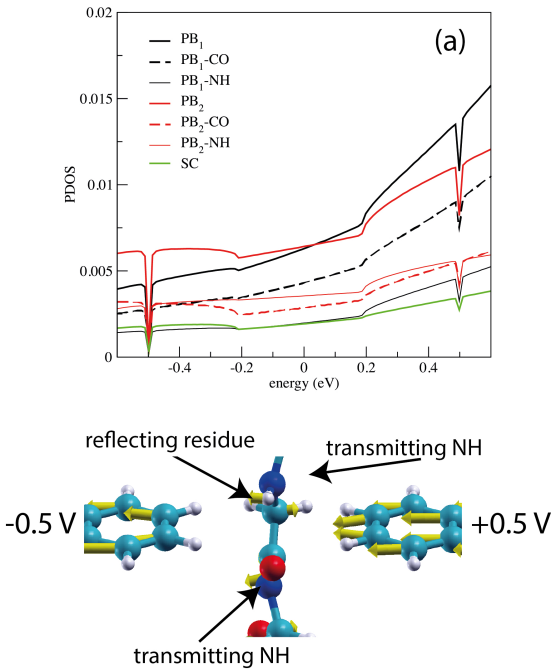

Fig. 6 PDOS (a) and schematic drawing of the atomic current flows (b) with the Gly residue in the middle of the nano-gap. C, O, N atoms are respectively in cyan, red and blu nd atomic currents are yellow arrows. Transmission occurs via the nitrogen atom of the PB while the atomic current are reflected from the central Gly residue.

prominent role of the "upper" $(NH)_2$ group. Interestingly, we observe that the total current measured for this configuration is nearly the same as the bond current contributions through the *NH* group at the second PB peaks ($(CO)_i$ configurations).

We can thus conclude the discussion on the Gly homo-peptide emphasizing that for the $(CO)_i$ configurations the electron injection through the PB *NH* group is maximized, reflecting a large overlap between the partially occupied N_{p_z} orbital and π, it is minimized for $(NH)_i$ configurations with the *NH* group in the gap, while, in the *SC* configuration, it separates in two branches across the upper and the lower *NH* groups, with nearly the same total amount as the one measured in the $(CO)_i$ case, thus reflecting a partial overlap between the N_{p_z} and π orbitals.

3.2 The Ala-Gly and Ala-Ala peptides.

In order to test the findings of the Gly homo-peptide case, we have examined two more cases including Ala residues. The first one is a tetra-peptide with alternating Gly and Ala AAs. The tunneling current signal obtained in this case is reported in Fig. 7

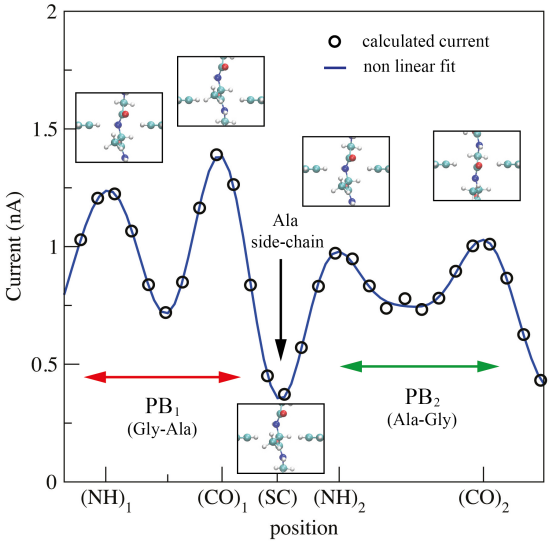

Fig. 7 Transverse tunneling current signal from the Gly-Ala peptide. Gly-Ala and the Ala-Gly PBs are considered with the Ala side chain in between.

We see that the main features found in the Gly homo-peptide are present also in this case. In particular, the PB current signals are, again, double peaked and the Gly-Ala-Gly transverse current evidences, as expected, a periodic behavior with a doubled period with respect to the Gly-Gly case: indeed the difference between the second and the first current maxima, that was nearly the same for PB_1 and PB_2 in the Gly case, now varies between the two adjacent PBs. The signal measured is of the same order of magnitude as in the Gly-Gly case that makes it, again, fully recognizable and separable from the noise due to water molecules. An important feature is that the minimum current values measured, $I \approx 0.4\ nA$, are nearly equal to the ones measured in the Gly case and correspond to the Gly and the Ala side chains in the middle of the gap. Hence, also in the Gly-Ala case the minimum current signal occurs with the suppression of the transport through the side chain and with the contributions of the upper and lower PBs, with a main role played by the (NH) groups. Looking at the first PB, we see that the current measured for $(CO)_1$ is larger than the one for $(NH)_1$ but their difference is smaller than the minimum current value measured (see the Gly case). This circumstance reflects the difference between the upper (Ala) and the lower (Gly) side chains. It should be outlined, also, that the narrow size of the GNRs here employed is probably unfit to capture the entire contribution from the Ala side chain if the peptide backbone is right in the GNR gap center. This enforces our idea of separating the signals from the side chain and the central peptide backbone. The same comments apply also for the second PB (Ala-Gly) where the two peaks are nearly identical. A further mapping of the PB signal in the Ala case with wider GNRs would clarify this secondary aspect.

A further confirmation of the main features observed in the Gly-Gly and the Gly-Ala cases comes from the Ala-Ala peptide. In this case, however, a reliable mapping of the entire signal could be attained only with wider GNRs electrodes. Indeed the peptide transversal overall dimension is, in this case, larger than 6 $\AA$, the adjacent side chains always lying on opposite directions with respect to the backbone axis. Therefore, if one side chain is within the gap, the other remains out of it making the atomistic model here adopted, with 5 $\AA$ wide GNR electrodes, suitable for a reliable mapping of just half of the signal, namely the ones related to the (NH) group and the side chain *SC* closest to it. The calculated transmission functions of these two configurations are reported in Fig. 8 and the corresponding current values are $I_{(NH)} \approx 1.44\ nA$ and $I_{(SC)} \approx 0.42\ nA$. We observe that, although $T(\varepsilon)$ is not negligible in the entire energy range, it increases from $\varepsilon \geq 0.2\ eV$ similarly to the Gly case. The current signal corresponding the *SC* configuration is nearly equal to the minimum current values measured in the previous two cases: then, also for the Ala peptide, the side chain contribution to the current is suppressed and the transport occurs mainly via the (NH) groups of the adjacent PBs. The peak signal is still of the order the *nA*, in this case a bit larger than the one measured in the previous cases, but, as already outlined, a full a complete mapping in this case requires the usage of wider GNRs.

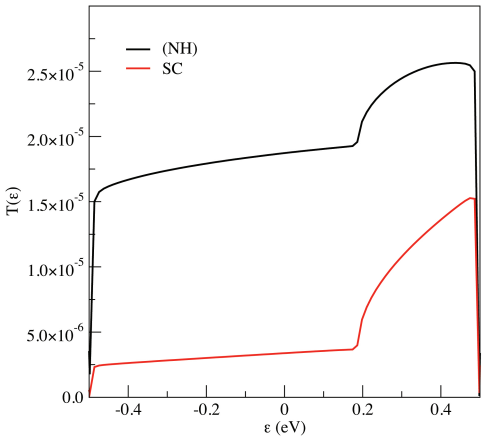

Fig. 8 Transmission functions calculated for the (NH) and the SC configurations in a Ala-Ala homo peptide. Full characterization of the PBs and side chains signals requires wider GNR electrodes.

4 Conclusions

GNR based nano-device, made of an array of nano-gaps, have been considered and studied in the context of DFT for single PB detection in a Gly based poly-peptide. The peptide translocation across the nano-gap array allows, in principle, the detection of various signals with atomic resolution. Using NEGF method and the Landauer-Büttiker formula we show that PBs from Gly poly-peptides leave a clear tunneling current fingerprint made of two, well defined current peaks of the order of magnitude of *nA*. We have shown that this signal, a sort of a reference PB signal (because the AAs side chain are just hydrogen atoms without any charge or dipolar source), arises from one channeled transmission functions due to the superposition of the left lead pseudo-π HOMO, the right lead pseudo-π^* LUMO and the α-carbon, side chains and PB related orbitals, the last ones being the p_z orbitals directly involved in the resonant double bond of the PB. The extent of these different contributions depends on the relative positions of PBs and residues with respect to the GNR plane. The two peaks of the current signal, occurring with the *NH* and *CO* groups of the PB in the middle of the GNR nano-gap, are mainly due to the overlap of the closest off-plane residues and the π and π^* GNR states. Interestingly, the difference between the higher second peak $(CO)_{\mathrm{i}}$ and the lower first peak $(NH)_{\mathrm{i}}$ is due the atomic current flowing across the *NH* group that contributes to the total current through the off-plane overlap between N_{p_z} and the HOMO and LUMO GNR states. The PB contribution to the current is the only one that survives when the Gly residue is in the middle of the nano-gap giving the same current level as the N_{p_z} contribution at the second maximum. A prominent role is played by the N_{p_z} orbital but a complex scenario, involving also charge transfer across the atoms of the PB, emerges when looking at all the orbitals involved. In particular, wave-function reflection occurs at the *CO* atomic group whose extent depends on the specific configuration. The described general behavior is confirmed when the second smallest residue is considered in a peptide with alternating Gly and Ala residues that still shows a double peak current signal for each PB measured. A strong indication is also obtained that the same scenario occurs for the larger Ala-Ala homo-peptide but in this case a different atomistic setup is required. The reported results pave the way for further development of nano-devices made of graphene nano-gaps for AAs detection and protein sequencing.

However the generalization of such a procedure to generic poly-peptides, needs, of course, a huge long term effort because the scenario may change dramatically depending on the charge/polar state of the side chains, on the side chain size and chemistry and on the variety of the possible PBs between the different AAs.

5 Conflict of interest

There are no conflicts to declare

6 Acknowledgements

Computational resources have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff[62]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see http://www.cresco.enea.it/english for information. Part of this work has been supported by CINECA under the ISCRAC project IsC45_InteGrAa.

References

1 M. Jain, I. T. Fiddes, K. H. Miga, H. E. Olsen, B. Paten and M. Akeson, *Nature methods*, 2015.

2 G. F. Schneider and C. Dekker, *Nature biotechnology*, 2012, **30**, 326–328.

3 E. Kennedy, Z. Dong, C. Tennant and G. Timp, *Nat Nano*, 2016, **11**, 968–976.

4 J. Wilson, L. Sloman, Z. He and A. Aksimentiev, *Advanced Functional Materials*, 2016, **26**, 4830–4838.

5 A. Asandei, A. E. Rossini, M. Chinappi, Y. Park and T. Luchian, *Langmuir*, 2017, **33**, 14451–14459.

6 A. Oukhaled, L. Bacri, M. Pastoriza-Gallego, J.-M. Betton and J. Pelta, *ACS chemical biology*, 2012, **7**, 1935–1949.

7 D. Di Marino, E. L. Bonome, A. Tramontano and M. Chinappi, *The journal of physical chemistry letters*, 2015, **6**, 2963–2968.

8 L. Mereuta, M. Roy, A. Asandei, J. K. Lee, Y. Park, I. Andricioaei and T. Luchian, *Scientific reports*, 2014, **4**, 3885.

9 O. Tavassoly, S. Nokhrin, O. Y. Dmitriev and J. S. Lee, *Febs Journal*, 2014, **281**, 2738–2753.

10 D. Rodriguez-Larrea and H. Bayley, *Nature nanotechnology*, 2013, **8**, 288–295.

11 T. Ohshiro, M. Tsutsui, K. Yokota, M. Furuhashi, M. Taniguchi and T. Kawai, *Nature nanotechnology*, 2014, **9**, 835–840.

12 C. B. Rosen, D. Rodriguez-Larrea and H. Bayley, *Nature biotechnology*, 2014.

13 G. Goyal, Y. B. Lee, A. Darvish, C. W. Ahn and M. J. Kim, *Nanotechnology*, 2016, **27**, 495301.

14 E. L. Bonome, R. Lepore, D. Raimondo, F. Cecconi, A. Tramontano and M. Chinappi, *The Journal of Physical Chemistry B*, 2015.

15 A. Sarathy, H. Qiu and J.-P. Leburton, *The Journal of Physical Chemistry B*, 2016.

16 F. Traversi, C. Raillon, S. Benameur, K. Liu, S. Khlybov, M. Tosun, D. Krasnozhon, A. Kis and A. Radenovic, *Nature nanotechnology*, 2013, **8**, 939–945.

17 E. Paulechka, T. A. Wassenaar, K. Kroenlein, A. Kazakov and A. Smolyanitsky, *Nanoscale*, 2016, **8**, 1861–1867.

18 R. G. Amorim, A. R. Rocha and R. H. Scheicher, *The Journal-2016 of Physical Chemistry C*, 2016, **120**, 19384–19388.

19 J. Prasongkit, G. T. Feliciano, A. R. Rocha, Y. He, T. Osotchan, R. Ahuja and R. H. Scheicher, *Scientific Reports*, 2015, **5**, 17560 EP –.

20 M. Di Ventra and M. Taniguchi, *Nature nanotechnology*, 2016, **11**, 117–126.

21 S. Huang, J. He, S. Chang, P. Zhang, F. Liang, S. Li, M. Tuchband, A. Fuhrmann, R. Ros and S. Lindsay, *Nature nanotechnology*, 2010, **5**, 868–873.

22 M. Tsutsui, M. Taniguchi, K. Yokota and T. Kawai, *Nature nanotechnology*, 2010, **5**, 286–290.

23 J. Lagerqvist, M. Zwolak and M. Di Ventra, *Nano letters*, 2006, **6**, 779–782.

24 S. J. Heerema and C. Dekker, *Nature Nanotechnology*, 2016, **11**, 127 EP –.

25 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.

26 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.

27 L. Agosta, G. Zollo, C. Arcangeli, F. Buonocore, F. Gala and M. Celino, *Phys. Chem. Chem. Phys.*, 2015, **17**, 1556–1561.

28 F. Buonocore, C. Arcangeli, F. Gala, G. Zollo and M. Celino, *The Journal of Physical Chemistry B*, 2015, **119**, 11791–11797.

29 A. V. Zaretski, B. C. Marin, H. Moetazedi, T. J. Dill, L. Jibril, C. Kong, A. R. Tao and D. J. Lipomi, *Nano Letters*, **15**, 635–640.

30 S. Lumetti, L. Martini and A. Candini, *Semiconductor Science and Technology*, 2017, **32**, 024002.

31 S. G. Sarwat, P. Gehring, G. Rodriguez Hernandez, J. H. Warner, G. A. D. Briggs, J. A. Mol and H. Bhaskaran, *Nano Letters*, 2017, **17**, 3688–3693.

32 J. Cai, P. Ruffieux, R. Jaafar, M. Bieri, T. Braun, S. Blankenburg, M. Muoth, A. P. Seitsonen, M. Saleh, X. Feng, K. Müllen and R. Fasel, *Nature*, 2010, **466**, 470 EP –.

33 M. Koch, F. Ample, C. Joachim and L. Grill, *Nature Nanotechnology*, 2012, **7**, 713 EP –.

34 Z. Li, H. Qian, J. Wu, B.-L. Gu and W. Duan, *Phys. Rev. Lett.*, 2008, **100**, 206802.

35 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale' and K. Schulten, *Journal of Computational Chemistry*, 2005, **26(16)**, 1781–1802.

36 J. Perdew, K. Burke and M.Ernzerhof, *Phys. Rev. Lett.*, 1996, **77(4)**, 3865–3868.

37 N. Troullier and J. Martins, *Phys. Rev. B*, 1991, **43**, 1993.

38 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys.: Condens. Matter*, 2009, **21(19)**, 395502–1–19.

39 R. Fletcher, *The Computer Journal*, 1970, **13(6)**, 317–22.

40 S. Grimme, *J. Comput. Chem.*, 2006, **27(15)**, 1787–99.

41 J. Soler, E. Artacho, J. Gale, A. García, J. Junquera, P. Ordejón and D. Sánchez-Portal, *Journal of Physics: Condensed Matter*, 2002, **14(11)**, 2745–2779.

42 K. Stokbro, J. Taylor, M. Brandbyge and P. Ordejón, *Annals of the New York Academy of Sciences*, 2003, **1006(1)**, 212–226.

43 H. Monkhorst and J. Pack, *Phys. Rev. B*, 1973, **13(5)**, 5188–92.

44 K. Stokbro, J. Taylor, M. Brandbyge, J.-L. Mozos and P. Ordejón, *Computational Material Science*, 2002, **27**, 151–160.

45 S. Datta, *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, 1995.

46 S. Kurth, G. Stefanucci, C.-O. Almbladh, A. Rubio and E. K. U. Gross, *Phys. Rev. B*, 2005, **72**, 035308.

47 G. Stefanucci and S. Kurth, *Nano Letters*, 2015, **15**, 8020–8025.

48 S. Liu, A. Nurbawono and C. Zhang, *Scientific Reports*, 2015, **5**, 15386 EP –.

49 C. Yam, X. Zheng, G. Chen, Y. Wang, T. Frauenheim and T. A. Niehaus, *Phys. Rev. B*, 2011, **83**, 245448.

50 G. Feliciano, C. Sanz-Navarro, M. Coutinho.Neto, P. Ordejón, R. H. Scheicher and A. R. Rocha, *Physical Review Applied*, 2015, **3**, 034003.

51 Y. Zhao, B. Ashcroft, P. Zhang, H. Liu, S. Sen, W. Song, J. Im, B. Gyarfas, S. Manna, S. Biswas *et al.*, *Nature nanotechnology*, 2014, **9**, 466–473.

52 H. L. McFarland, T. Ahmed, J.-X. Zhu, A. V. Balatsky and J. T. Haraldsen, *The Journal of Physical Chemistry Letters*, 2015, **6**, 2616–2621.

53 M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin and J. M. Tour, *Science*, 1997, **278**, 252–254.

54 H. W. C. Postma, *Nano Letters*, 2010, **10**, 420–425.

55 Y. Zhao, B. Ashcroft, P. Zhang, H. Liu, S. Sen, W. Song, J. Im, B. Gyarfas, S. Manna, S. Biswas, C. Borges and S. Lindsay, *Nature Nanotechnology*, 2014, **9**, 466 EP –.

56 M. Brandbyge, N. Kobayashi and M. Tsukada, *Physical Review B*, 1999, **60**, 17064.

57 B. Ludoph, M. Devoret, D. Esteve, C. Urbina and J. Van Ruitenbeek, *Physical review letters*, 1999, **82**, 1530.

58 G. C. Solomon, C. Herrmann, T. Hansen, V. Mujica and M. A. Ratner, *Nat Chem*, 2010, **2**, 223–228.

59 T. N. Todorov, *Journal of Physics: Condensed Matter*, 2002, **14**, 3049.

60 N. Papior, N. Lorente, T. Frederiksen, A. Garca and M. Brandbyge, *Computer Physics Communications*, 2017, **212**, 8 – 24.

61 T. Frederiksen, M. Paulsson, M. Brandbyge and A.-P. Jauho, *Phys. Rev. B*, 2007, **75**, 205413.

62 G. Ponti, F. Palombi, D. Abate, F. Ambrosino, G. Aprea, T. Bastianelli, F. Beone, R. Bertini, G. Bracco, M. Caporicci, B. Calosso, M. Chinnici, A. Colavincenzo, A. Cucurullo, P. dAngelo, M. D. Rosa, P. D. Michele, A. Funel, G. Furini, D. Giammattei, S. Giusepponi, R. Guadagni, G. Guarnieri, A. Italiano, S. Magagnino, A. Mariano, G. Mencuccini, C. Mercuri, S. Migliori, P. Ornelli, S. Pecoraro, A. Perozziello, S. Pierattini, S. Podda, F. Poggi, A. Quintiliani, A. Rocchi, C. Scio, F. Simoni and A. Vita, *IEEE HPCS*, 2014, **6903807**, 1030–1033.