

SUPPORTING INFORMATIONS

Sum Frequency Generation theory and P3HT molecular modeling

The interfacial systems studied in this work can be treated as a three-layer system composed of two centrosymmetric media (i.e., the polymer bulk and the surrounding medium, air or water) and an interfacial layer (i.e., the polymer/water or the polymer/air interface) with different refractive indexes (n_1 , n_2 , n' respectively) (Figure S1).

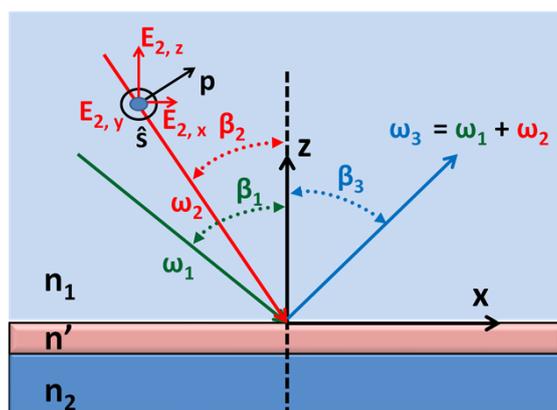


Figure S1. General sketch of an interfacial system treated as a three-layer structure composed of two media and an interfacial layer.

Upon irradiation of two optical fields E_1 and E_2 with frequencies ω_1 and ω_2 , respectively, a second-order nonlinear polarization $P_{SFG}^{(2)}(\omega_3 = \omega_1 + \omega_2)$ is generated [1-4]:

$$P_{SFG_i}^{(2)}(\omega_3 = \omega_1 + \omega_2) = \sum_j^{x,y,z} \sum_k^{x,y,z} \chi_{effijk}^{(2)}(\omega_3 = \omega_1 + \omega_2) : E_{j,1}(\omega_1) E_{k,2}(\omega_2) \quad (S1)$$

where $\chi_{eff}^{(2)}$ is the effective second-order non-linear susceptibility of the interface, a third rank tensor composed of 27 elements with specific symmetry properties.

$\chi_{eff}^{(2)}$ is expressed by:

$$\chi_{eff}^{(2)} = [\hat{e}(\omega_3) \cdot L(\omega_3)] \cdot \chi^{(2)} : [L(\omega_1) \cdot \hat{e}(\omega_1)] [L(\omega_2) \cdot \hat{e}(\omega_2)] \quad (S2)$$

where $\hat{e}(\omega_i)$ is the unit polarization vector of the electric field $E_i(\omega_i)$ and $L(\omega_i)$ is the Fresnel factor matrix at frequency ω_i , which takes into account how light is transmitted as a function of the incidence angle and wavelength.

Under the electric-dipole approximation, the non-linear polarization generated in the two centrosymmetric media vanishes due to inversion symmetry; the dominating source of radiation for SFG signal intensity in the reflected direction (I_{SFG}) is thus given only by the interfacial polarization sheet. I_{SFG} is expressed by [1-4]:

$$I_{SFG}(\omega_3 = \omega_1 + \omega_2) = \frac{8\pi^3 \omega_3^2 \sec^2 \beta_3}{c^3 n_1(\omega_3) n_1(\omega_1) n_1(\omega_2)} |\chi_{eff}^{(2)}|^2 I_1(\omega_1) I_2(\omega_2) \quad (S3)$$

where $n_i(\omega)$ is the refractive index of media i at frequency ω , β_3 is the reflection angle of the sum-frequency field, $I_1(\omega_1)$ and $I_2(\omega_2)$ are the intensities of the two input fields. One should notice the proportionality relation between I_{SFG} and $|\chi_{eff}^{(2)}|^2$.

Through a set of polarizers, the fields of incidence can be set in two components along the directions parallel and perpendicular to the xz incidence plane, identified by the versors \hat{p} and \hat{s} , respectively. In the case of an azimuthally isotropic interface, there are only 4 independent non-vanishing components of $\chi^{(2)}$; more specifically, in the adopted cartesian coordinates system, they are $\chi_{xxz}^{(2)} = \chi_{yyz}^{(2)}$, $\chi_{xxx}^{(2)} = \chi_{yzy}^{(2)}$, $\chi_{zzx}^{(2)} = \chi_{zyy}^{(2)}$ and $\chi_{zzz}^{(2)}$. They can be deduced by measuring the SFG signal intensity for four different input and output polarization combinations, namely: *SSP* (referring to S -polarized sum-frequency field E_3 , S -polarized E_1 , and P -polarized E_2 , respectively), *SPS*, *PSS*, and *PPP*. In more detail, the effective susceptibilities for these polarization combinations are expressed by the following equations:

$$\chi_{eff,SSP}^{(2)}(\omega_3) = L_{yy}(\omega_3) L_{yy}(\omega_1) L_{zz}(\omega_2) \sin \beta_2 \chi_{yyz}$$

$$\chi_{eff,SPS}^{(2)}(\omega_3) = L_{yy}(\omega_3) L_{zz}(\omega_1) L_{yy}(\omega_2) \sin \beta_1 \chi_{yzy}$$

$$\chi_{eff,PSS}^{(2)}(\omega_3) = L_{zz}(\omega_3) L_{yy}(\omega_1) L_{yy}(\omega_2) \sin \beta_3 \chi_{zyy}$$

$$\chi_{eff,PPP}^{(2)}(\omega_3)$$

$$= -L_{xx}(\omega_3) L_{xx}(\omega_1) L_{zz}(\omega_2) \cos \beta_3 \cos \beta_1 \sin \beta_2 \chi_{xxz} - L_{xx}(\omega_3) L_{zz}(\omega_1) L_{xx}(\omega_2) \cos \beta_3 \sin \beta_1 \cos \beta_2 \chi_{xxx} + L_{xx}(\omega_3) L_{xx}(\omega_1) L_{xx}(\omega_2) \sin \beta_3 \cos \beta_1 \sin \beta_2 \chi_{xxx} + L_{zz}(\omega_3) L_{zz}(\omega_1) L_{zz}(\omega_2) \sin \beta_3 \sin \beta_1 \sin \beta_2 \chi_{zzz}$$

(S4)

where β_i is the incidence angle of the optical field E_i , and $L_{xx}(\omega_i)$, $L_{yy}(\omega_i)$ and $L_{zz}(\omega_i)$ are the diagonal elements of $L(\omega_i)$, given by

$$\begin{aligned} L_{xx}(\omega_i) &= \frac{2n_1(\omega_i)\cos\gamma_i}{n_1(\omega_i)\cos\gamma_i + n_2(\omega_i)\cos\beta_i} \\ L_{yy}(\omega_i) &= \frac{2n_1(\omega_i)\cos\beta_i}{n_1(\omega_i)\cos\beta_i + n_2(\omega_i)\cos\gamma_i} \\ L_{zz}(\omega_i) &= \frac{2n_2(\omega_i)\cos\beta_i}{n_1(\omega_i)\cos\gamma_i + n_2(\omega_i)\cos\beta_i} \left(\frac{n_1(\omega_i)}{n'(\omega_i)} \right)^2 \end{aligned} \quad (\text{S5})$$

In the above equations, γ_i is the refracted angle given by the Snell law, according to the system geometry:

$$n_1(\omega_i)\sin\beta_i = n_2(\omega_i)\sin\gamma_i \quad (\text{S6})$$

It is not straightforward to measure the refractive index of the interfacial layer, $n'(\omega_i)$, so this is usually supposed to be equal either to $n_1(\omega_i)$ or $n_2(\omega_i)$.

In the case where the interface is composed of molecules, $\chi^{(2)}$ is related to the molecular hyperpolarizability tensor $\alpha^{(2)}$ by the relationship:

$$\chi_{eff\,ijk}^{(2)} = N_S l_{ii}(\omega_3) l_{jj}(\omega_1) l_{kk}(\omega_2) \sum_{l,m,n}^{a,b,c} \langle (\hat{i} \cdot \hat{l})(\hat{j} \cdot \hat{m})(\hat{k} \cdot \hat{n}) \rangle \alpha_{lmn}^{(2)} \quad (\text{S7})$$

where N_S is the surface density of molecules. $(\hat{i}, \hat{j}, \hat{k})$ generically indicate the possible combination sets of the unit vectors $(\hat{x}, \hat{y}, \hat{z})$ along the lab coordinates (x,y,z) . $(\hat{l}, \hat{m}, \hat{n})$ generically indicate the possible combination sets of the unit vectors $(\hat{a}, \hat{b}, \hat{c})$ along the molecular coordinate systems (a,b,c) .

The diagonal elements of the tensor $l(\omega_i)$, namely $l_{ii}(\omega_i)$, $l_{jj}(\omega_i)$ and $l_{kk}(\omega_i)$, describe the microscopic local field correction, which will be neglected in our analysis ($l(\omega_i)$ will be taken as the identity matrix). The angular brackets denote an average over the molecular orientation distribution.

The formalism which describes, on a microscopic scale, the expression of the hyperpolarizability tensor $\alpha^{(2)}$ is based on the second-order time-dependent perturbation theory for the density matrix.

For the infrared radiation resonant with the frequencies of vibrational modes q , it is possible to derive following approximate expression for the resonant contribution [1-8]:

$$\alpha_{rlmn}^{(2)} = \frac{2}{\hbar} \sum_q \frac{\left[\frac{\partial \alpha_{glm}}{\partial Q_q} \right]_0 \left[\frac{\partial \mu_{g0g1n}}{\partial Q_q} \right]_0 |\langle \Theta_{g0} | Q_l | \Theta_{g1} \rangle|^2}{[\omega_{g1g0}^{(S)} - \omega_2 - i\Gamma_{g1g0}^{(S)}]} \quad (\text{S8})$$

where the subscript r indicates the resonant contribution, $\omega_{g1g0}^{(S)}$ is the frequency of the vibrational

mode of the molecule, $\frac{\partial \mu_{g0g1n}}{\partial Q_q}$ is the derivative of the infrared dipole operator μ with respect to the normal mode coordinate Q_q of the q^{th} vibrational mode, acting between an excited vibrational state $g1$ and the fundamental vibrational state $g0$ (g indicate the ground electronic state, while the

numbers 0 and 1 refer to the ground and first excited vibration state), $\frac{\partial \alpha_{glm}}{\partial Q_q}$ is the derivative of the specified element of the electronic polarizability α_g with respect to Q_q . The dispersion of the polarizability when the visible excitation is resonant with electronic transitions of the studied material is accounted by the following expression [1-8]:

$$\alpha_{glm} = \frac{1}{\hbar} \sum_e \frac{\mu_{gem} \mu_{egl}}{\omega_{eg} - \omega_2 - \omega_1} + \frac{\mu_{gel} \mu_{egm}}{\omega_{eg} + \omega_2 + \omega_1} \quad (\text{S9})$$

In this equation, the operator μ acts on transitions from the ground electronic state g to the excited electronic states e , and ω_{eg} is the resonant frequency of the electronic transition. The infrared absorption and Raman scattering have their respective intensities directly proportional to the square

modulus of $\frac{\partial \mu_{g0g1n}}{\partial Q_q}$ and $\frac{\partial \alpha_{glm}}{\partial Q_q}$. From Eq. S8, both must be non-zero in order to have SFG activity of the considered vibrational mode. It should be noted from Eqs. S8 and S9 that if the visible and SFG frequencies are away from electronic resonances, the hyperpolarizability tensor $\alpha_{lmn}^{(2)}$, and therefore the nonlinear susceptibility $\chi_{ijk}^{(2)}$ will be symmetric in the first two indices, which implies that for an isotropic surface $\chi_{zyz}^{(2)} = \chi_{zyy}^{(2)}$, so that PSS and SPS polarization combinations yield the same

tensor components. Although this is not strictly true in our experiments, we will make this simplifying assumption.

Furthermore, the SFG spectroscopy usually shows other contributions (from the molecules and the substrate) with resonances outside the region of interest, leading to the non-resonant part $\alpha_{nr lmn}^{(2)}$, which in many cases is significant. Therefore $\alpha_{lmn}^{(2)}$ is considered to be the sum of these two contributions:

$$\alpha_{lmn}^{(2)} = \alpha_{r lmn}^{(2)} + \alpha_{nr lmn}^{(2)} \quad (\text{S10})$$

With IR-visible SFG, if the IR frequency ω_2 is near vibrational resonances, $\alpha^{(2)}$ and $\chi^{(2)}$ can be written as:

$$\alpha^{(2)} = \alpha_{nr}^{(2)} + \sum_q \frac{\alpha_q^{(2)}}{\omega_2 - \omega_q - i\Gamma_q} \quad (\text{S11})$$

$$\chi_{eff}^{(2)} = \chi_{nr eff}^{(2)} + \sum_q \frac{\chi_{q eff}^{(2)}}{\omega_2 - \omega_q - i\Gamma_q} \quad (\text{S12})$$

where $\alpha_q^{(2)}$ ($\chi_{q eff}^{(2)}$), ω_q and Γ_q denote the strength, the resonant frequency and the damping constant of the q^{th} vibrational mode, respectively, and the subscript *nr* refers to the non-resonant contribution, which is defined as [1-8]:

$$\chi_{nr eff}^{(2)}(\omega) = Ge^{iF} \quad (\text{S13})$$

G and F being its modulus and phase, respectively. Thus, the expression describing the SFG experimental curves for a given polarization combination of the input/output beams is:

$$|\chi_{eff}^{(2)}(\omega_2)|^2 = \left| \chi_{nr eff}^{(2)} + \sum_q \frac{\chi_{q eff}^{(2)}}{\omega_2 - \omega_q - i\Gamma_q} \right|^2 \quad (\text{S14})$$

Fitting the SFG spectrum with the above expression, it is possible to estimate $\chi_{q eff}^{(2)}$.

In order to reduce the number of free parameters we adopted two expedients. We made the reasonable assumption that for a fixed mode q , the parameters ω_q and Γ_q have the same values for all polarization combinations, i.e.:

$$\omega_{q,SSP} = \omega_{q,SPS} = \omega_{q,PPP} \quad (\text{S15})$$

$$\Gamma_{q,SSP} = \Gamma_{q,SPS} = \Gamma_{q,PPP} \quad (\text{S16})$$

while the parameters $\chi_{q\text{ eff}}^{(2)}$ were allowed to vary for each polarization combination. In this way, the number of parameters is strongly reduced and the fitting procedure is more robust. However, the number possible vibrational modes involved for P3HT in the spectral range investigated is high and even after this reduction in the number of adjustable parameters, the fitting procedure might not converge to a suitable solution. For this reason, we started the fitting procedure by considering only the principal mode and obtaining a first rough fit using this simplified model. Once a suitable initial guess is obtained, the model is iteratively refined by adding the remaining modes until a satisfactory match to the experimental data is achieved. If the experimental results showed little or negligible contribution of a particular mode q for one of the polarization combinations, the corresponding intensity parameter $\chi_{q\text{ eff}}^{(2)}$ is fixed at 0 in order to further reduce the dimension of the fitting parameters set. All the fitting procedures have been performed by implementing model and formulas in OriginLab and using the dedicated nonlinear fitting tool.

In many cases of SHG and SFG, $\alpha^{(2)}$ can be associated with a well-defined section or moiety of the surface molecules. If $\alpha_{lmn}^{(2)}$ is known, then the average orientation of the moiety can often be deduced from measurements of $\chi_{q\text{ eff}}^{(2)}$. In simpler cases, it can be associated to the whole molecule at the surface of interest.

If $\alpha_{q\text{ lmn}}^{(2)}$ for the particular studied material is known, then the average orientation of the selected moiety may be deduced by comparing theoretical values of $\chi_{q\text{ eff}}^{(2)}$ for all polarization combinations with the experimental ones obtained by fitting the SFG spectra. It is thus of fundamental importance in the analysis of SFG spectra to hypothesize a specific molecular model of the considered system. The orientation of the molecular axis system with respect to the axis of the laboratory system can be expressed as a function of the Euler angles θ , φ and ψ :

$$\hat{a} = (-\sin\psi\sin\varphi + \cos\psi\cos\theta\cos\varphi)\hat{x} + (\sin\psi\cos\varphi + \cos\psi\cos\theta\sin\varphi)\hat{y} + (\cos\psi\sin\theta)\hat{z}$$

$$\begin{aligned}\hat{b} &= (-\cos\psi\sin\varphi - \sin\psi\cos\theta\cos\varphi)\hat{x} + (\cos\psi\cos\varphi - \sin\psi\cos\theta\sin\varphi)\hat{y} + (-\sin\psi\sin\theta)\hat{z} \\ \hat{c} &= (-\sin\theta\cos\varphi)\hat{x} + (-\sin\theta\sin\varphi)\hat{y} + \cos\theta\hat{z}\end{aligned}\quad (\text{S17})$$

In the specific case of rr-P3HT, the c axis is selected along the polymer chain axis, while the a axis is in the plane of the thiophene ring (ac plane) and the b axis is perpendicular to it, as depicted in Figure S2.

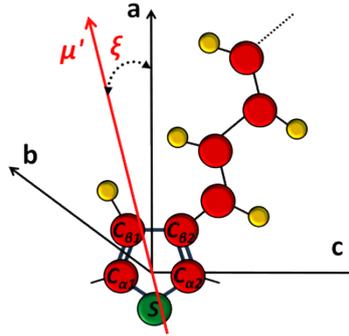


Figure S2. Molecular reference system (a, b, c) for a rr-P3HT monomer; the orientation of the dipole moment derivative μ' is also shown.

As a first approximation, it is possible to assume the electronic polarizabilities α_{aa} , α_{bb} and α_{cc} for the P3HT monomer to be the same of those for the conjugated core of Polythiophene (PT). However, due to the different symmetry group (C_s for P3HT and C_{2v} for PT), one should also take into account that P3HT has a nonvanishing α_{bc} polarizability component, although it should be small and will be neglected here. The molecular model of rr-P3HT takes into account the SFG signal $\chi_q^{(2)eff}$ of the symmetric stretching mode of $C_\alpha = C_\beta$ ($\sim 1450\text{ cm}^{-1}$, according to several reports in literature [9-10]). We will name this contribution as $\chi_{C=C}^{(2)eff}$. For PT, the infrared dipole moment derivative

$\frac{\partial\mu}{\partial Q_{C=C}} = \mu'$ lies in the plane of the thiophenic ring along the a axis (hence perpendicular to c axis of the chain). However, in the case of P3HT it is expected that μ' is slightly inclined with respect to the a axis, due to the presence of the alkyl chain R, which confers greater inertia to the carbon β_2 . Thus, μ' can be projected along the a and c axes, with components, respectively, given by $\mu'_a = \mu' \cos(\xi)$

and $\mu'_c = -\mu' \sin(\xi)$ (Figure S2, right panel). The model assumes $\xi = 5^\circ$ as a reasonable inclination.

The hyperpolarizability elements of interest for this mode are therefore given by:

$$\begin{aligned}
\alpha_{C=C,aaa}^{(2)} &= \alpha'_{aa}\mu'_a \\
\alpha_{C=C,bbb}^{(2)} &= \alpha'_{bb}\mu'_b \\
\alpha_{C=C,cca}^{(2)} &= \alpha'_{cc}\mu'_a \\
\alpha_{C=C,aac}^{(2)} &= \alpha'_{aa}\mu'_c \\
\alpha_{C=C,bbc}^{(2)} &= \alpha'_{bb}\mu'_c \\
\alpha_{C=C,ccc}^{(2)} &= \alpha'_{cc}\mu'_c
\end{aligned} \tag{S18}$$

And the $\chi_{C=Cijk}^{(2)}$ components can be finally expressed as [11]:

$$\begin{aligned}
\chi_{C=C_{yyz}}^{(2)}(\theta, \psi) &= \frac{\mu' \cos \xi \sin \theta \cos(\psi)}{2} [(1 - \sin^2 \theta \cos^2(\psi)^2) \alpha'_{aa} + (\cos(\psi)^2 + \sin(\psi)^2 \cos^2 \theta) \alpha'_{bb} + \sin^2 \theta \alpha'_{cc}] - \frac{\mu' \sin \xi \cos \theta}{2} \\
& \quad [(1 - \sin^2 \theta \cos^2(\psi)^2) \alpha'_{aa} + (\cos(\psi)^2 + \sin(\psi)^2 \cos^2 \theta) \alpha'_{bb} + \sin^2 \theta \alpha'_{cc}]
\end{aligned}$$

$$\begin{aligned}
\chi_{C=C_{zyz}}^{(2)}(\theta, \psi) &= \frac{\mu' \cos \xi \sin \theta \cos(\psi)}{2} [(1 - \sin^2 \theta \cos^2(\psi)^2) \alpha'_{aa} - \sin^2 \theta \sin(\psi)^2 \alpha'_{bb} - \cos^2 \theta \alpha'_{cc}] - \frac{\mu' \sin \xi \cos \theta \sin^2 \theta}{2} [-\cos \theta]
\end{aligned}$$

$$\chi_{C=C_{zzz}}^{(2)}(\theta, \psi) = \mu' (\cos \xi \sin \theta \cos(\psi) - \sin \xi \cos \theta) [\sin^2 \theta \cos^2(\psi)^2 \alpha'_{aa} + \sin^2 \theta \sin(\psi)^2 \alpha'_{bb} + \cos^2 \theta \alpha'_{cc} + \cos^2 \theta \alpha'_{cc}] \tag{S19}$$

We now define the torsion angle δ between adjacent monomer rings and the chain twist versor $\hat{\tau}$,

perpendicular to the median plane of two consecutive monomers, which makes an angle $\gamma = \psi + \frac{\delta}{2}$

with the vertical plane defined by the z and c axes, as illustrated in Figure S3. The P3HT model

assumes that the chain does not have inversion symmetry perpendicular to the c axis, that is, the

thiophene rings are lying along two planes, mutually oriented by a torsion angle δ , where each plane

contains half of monomers. Therefore, $\hat{\tau}$ is the direction of preferential torsion for the thiophene

rings at the interface. A simpler model with effectively planar chains (a distribution of δ symmetric about 180°) would only have contribution from $\alpha_{C=C,ccc}^{(2)}$ in Eqs. S18, and cannot explain the experimental results [11].

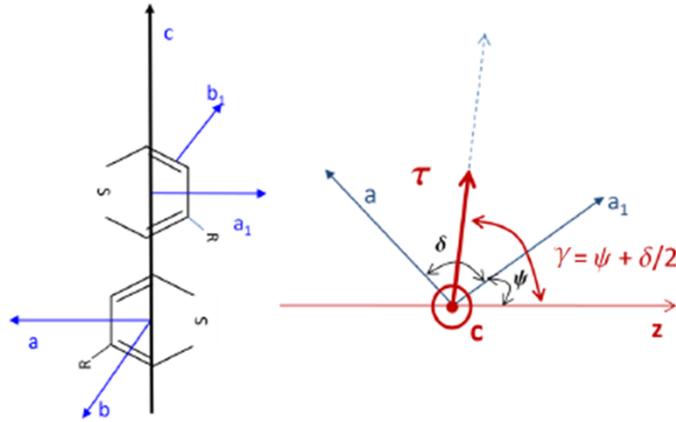


Figure S3. Definition of the angles δ , γ and ψ in the molecular coordinates system.

By considering the torsion between the thiophene rings of adjacent monomers and by using the above defined angles γ and δ , we obtain the independent $\chi_{C=Cijk}^{(2)}$ elements as a function of chain tilt θ and twist γ , with δ and ξ as fixed parameters [11]:

$$\chi_{C=C_{yyz}}^{(2)T}(\theta, \gamma, \delta, \xi) = \frac{1}{2} \left(\frac{\mu' \cos \xi \sin \theta \cos \left(\gamma - \frac{\delta}{2} \right)}{2} \right) \left[\left(1 - \sin^2 \theta \cos \left(\gamma - \frac{\delta}{2} \right) \right) \alpha_{aa}' + \left(\cos \left(\gamma - \frac{\delta}{2} \right) \right)^2 + \sin \left(\gamma - \frac{\delta}{2} \right)^2 \cos^2 \theta \right) \alpha_{bb}' + \sin \theta \cos \left(\gamma - \frac{\delta}{2} \right) \alpha_{cc}' \right]$$

$$\chi_{C=C_{yzy}}^{(2)T}(\theta, \gamma, \delta, \xi) = \frac{1}{2} \left(\frac{\mu' \cos \xi \sin \theta \cos \left(\gamma - \frac{\delta}{2} \right)}{2} \right) \left[\left(1 - \sin^2 \theta \cos \left(\gamma - \frac{\delta}{2} \right) \right) \alpha_{aa}' - \sin^2 \theta \sin \left(\gamma - \frac{\delta}{2} \right)^2 \alpha_{bb}' - \cos^2 \theta \alpha_{cc}' \right] - \frac{\mu' \sin \xi \cos \theta}{2} \alpha_{cc}'$$

$$\chi_{C=C_{zzz}}^{(2)T}(\theta, \gamma, \delta, \xi) = \frac{1}{2} \left(\mu' \left(\cos \xi \sin \theta \cos \left(\gamma - \frac{\delta}{2} \right) - \sin \xi \cos \theta \right) \right) \left[\sin^2 \theta \cos \left(\gamma - \frac{\delta}{2} \right)^2 \alpha_{aa}' + \sin^2 \theta \sin \left(\gamma - \frac{\delta}{2} \right)^2 \alpha_{bb}' + \cos^2 \theta \alpha_{cc}' \right] + \mu' \sin \xi \cos \theta \alpha_{cc}'$$

S(20)

It is worth noting that, thanks to the specific molecular symmetry properties of P3HT, the molecular orientation respect to the laboratory coordinates can be specified by only two angular coordinates, θ and γ , as mentioned in the main paper . The adopted molecular model for P3HT admits δ near 170° [12] and takes into account the values for polarizability derivatives $\alpha_{aa}' = 363.09$; $\alpha_{bb}' = 108.07$; $\alpha_{cc}' = 948.92$, as calculated for oligomers with 7 thiophene monomers [13]. The obtained expressions still contain a unknown scalar parameter, μ' , which acts as a multiplicative factor. Figure S4 is a plot of the three independent $\chi_C = c_{ijk}^{(2)T} / \mu'$ elements as a function of chain orientation.

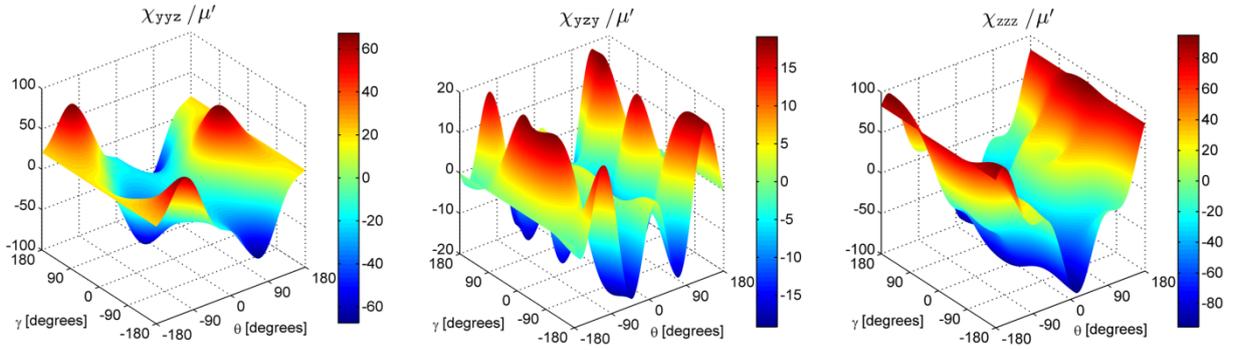


Figure S4. Three-dimensional plots of the nonlinear susceptibility components as a function of a rr-P3HT chain orientation at the interface.

The definition of the ratios q_1 and q_2 and q_3 , as listed in the equations below, allows eliminating the unknown parameter μ' .

$$q_1(\theta, \gamma) = \frac{|\chi_C = c_{eff,SPS}^{(2)T}|}{|\chi_C = c_{eff,SSP}^{(2)T}|}$$

$$q_2(\theta, \gamma) = \frac{|\chi_C = c_{eff,SSP}^{(2)T}|}{|\chi_C = c_{eff,PPP}^{(2)T}|}$$

$$q_3(\theta, \gamma) = \frac{|\chi_C = c_{eff,SPS}^{(2)T}|}{|\chi_C = c_{eff,PPP}^{(2)T}|}$$

(S21)

At this point, based on the adopted molecular model for rr-P3HT, it is possible to derive a set of angular function of these ratios for each considered interface, and finally, based on comparison with the experimental values (obtained by fitting the SFG spectra), to identify admissible region of values for (θ, γ) , i.e., the most probable molecular orientations.

Finally, the expression for $\chi_{eff}^{(2)}$ reported above (Eq. S4) have been corrected by taking into account the reflection losses due to all other interfaces that the beam may go through before reaching the interface of interest, in the following way:

$$\begin{aligned}
\chi_{eff,SSP}^{(2)}(\omega_3) &= t_{S3}t_{S1}t_{P2}L_{yy}(\omega_3)L_{yy}(\omega_1)L_{zz}(\omega_2)\sin\beta_2\chi_{yyz} \\
\chi_{eff,SPS}^{(2)}(\omega_3) &= t_{S3}t_{P1}t_{S2}L_{yy}(\omega_3)L_{zz}(\omega_1)L_{yy}(\omega_2)\sin\beta_1\chi_{yzy} \\
\chi_{eff,PPS}^{(2)}(\omega_3) &= t_{P3}t_{S1}t_{S2}L_{zz}(\omega_3)L_{yy}(\omega_1)L_{yy}(\omega_2)\sin\beta_3\chi_{zyy} \\
\chi_{eff,PPP}^{(2)}(\omega_3) &= t_{P3}t_{P1}t_{P2}[-L_{xx}(\omega_3)L_{xx}(\omega_1)L_{zz}(\omega_2)\cos\beta_3\cos\beta_1\sin\beta_2\chi_{xxz} - L_{xx}(\omega_3)L_{zz}(\omega_1)L_{xx}(\omega_2)\cos\beta_3\sin\beta_1\cos\beta_2\chi_{xxz}]
\end{aligned}
\tag{S22}$$

where the transmission Fresnel factors were calculated from the following equations for the S- and P- polarizations:

$$\begin{aligned}
t_{Si} &= \frac{2n_i\cos\beta_i}{n_i\cos\beta_i + n_r\cos\eta_i} \\
t_{Pi} &= \frac{2n_i\cos\eta_i}{n_i\cos\eta_i + n_r\cos\beta_i}
\end{aligned}
\tag{S23}$$

being β and η the incident and the refracted angle, respectively.

In order to provide an estimate of the admissible range for the molecular orientation, we compare the experimental results for $\chi_{C=c_{eff}}^{(2)}$ obtained from the fitting procedure with the simulated values for $\chi_{C=c_{eff}}^{(2)}$ for each chain orientation according to the molecular model (Eqs. S20 and S22). For each polarization (*SPS*, *SSP*, *PPP*) we consider just the contribution of the principal mode (C=C symmetric stretch at $\sim 1450 \text{ cm}^{-1}$), i.e. the expected value $\bar{\chi}_{C=c,eff}^{(2)}$ and the corresponding standard

deviation $\sigma_{C = Ceff}$. The unknown parameter μ' , is eliminated by considering the ratios of the signal components:

$$\begin{aligned}
 q_1 &= \frac{\bar{\chi}_{C = Ceff,SPS}}{\bar{\chi}_{C = Ceff,SSP}}, \\
 q_2 &= \frac{\bar{\chi}_{C = Ceff,SSP}}{\bar{\chi}_{C = Ceff,PPP}}, \\
 q_3 &= \frac{\bar{\chi}_{C = Ceff,SPS}}{\bar{\chi}_{C = Ceff,PPP}}
 \end{aligned} \tag{S24}$$

For each pair of effective susceptibilities with fitted values $\bar{\chi}_1$ and $\bar{\chi}_2$ and standard deviations σ_1 and σ_2 , for the ratio q it holds

$$\bar{q} = \frac{\bar{\chi}_1}{\bar{\chi}_2} \tag{S25}$$

$$\sigma_q = \frac{\bar{\chi}_1}{\bar{\chi}_2} \sqrt{\left(\frac{\sigma_1}{\bar{\chi}_1}\right)^2 + \left(\frac{\sigma_2}{\bar{\chi}_2}\right)^2} \tag{S26}$$

In order to identify the compatible molecular orientations, we define the parameter

$$\Delta_r = \max_{i=1,2,3} \frac{|q_i - q_i^{mod}|}{\sigma_{q_i}} \tag{S27}$$

which represents the maximum discrepancy between the experimental values q_i and the expected ones from the molecular model, q_i^{mod} , in units of standard deviation. Admissible regions are identified by values of Δ_r close to zero, denoted in red/orange in Figures 3 in the main manuscript and Figures S5 and S6 in following section.

The following table illustrates the refractive index used for the Fresnel factors calculation:

	CaF ₂	Air	Water	ITO	rr-P3HT
ω_1 (532 nm)	1.44	1	1.32	1.94	1.95
ω_2 (1450 cm ⁻¹ , ~6900 nm)	1.37	1	1.32	1.7	1.9
ω_3 (~490nm)	1.44	1	1.33	1.98	1.7

Analysis of thick samples

The previous description applies when the SFG signal is generated by a single interfacial region n' (typically a monolayer). In the case of samples fabricated by spin coating, like in the present study, possible contributions from the buried polymer/substrate interface may strongly affect the SFG signal from the polymer/water(air) interface, and several schemes have been used to separate these contributions [14-16]. It is thus important to discriminate among them, and in our study we will exploit the strong absorption of the visible and SFG beams in the polymer film, by taking into account two different cases, depending on the sample thickness t and penetration depth l_i of the light beams:

i) $t > l_i$

ii) $t \ll l_i$

With the beams incoming from the transparent substrate (CaF_2 – medium 1), in the first case, the contribution of the second interface (polymer/medium 2) can be neglected due to the strong attenuation of both the incident visible beam and the reflected SFG beam in the bulk material. Hence, the model of a very thin interfacial layer (monolayer) can be safely applied simply by assuming that the interfacial layer coincides with the first interface (medium 1/polymer), while the second medium is the bulk polymer (so that $n' = n_2$).

In the second case, one should take into account that the measured signal is the superposition of the contribution from the first interface (medium 1/P3HT) and from the second interface (P3HT/medium 2). In fact, if we assume that a linear superposition of the effects applies, which means neglecting phase shifts and absorption due to beam propagation in the very thin polymer film, for the case of the *SSP* polarization the following relation holds:

$$\chi_{eff,SSP}^{thin} = F_{SSP} L_{SSP} (\chi_{yyz}^{first} + \chi_{yyz}^{second}) = F_{SSP} L_{SSP} \chi_{yyz}^{first} + F_{SSP} L_{SSP} \chi_{yyz}^{second} = \chi_{eff,SSP}^{thick} + \chi_{eff,SSP}^{second} \quad (\text{S28})$$

Where F_{SSP} and L_{SSP} are the factors in equation S22 including the transmission Fresnel factors of previous interfaces (products of $t_{S,P}$'s), and angular terms and Fresnel factors L_{ii} , respectively. The

above formula clearly shows that the effective nonlinearity for *SSP* polarization measured in the case a thin sample is the sum of the effective nonlinearity that would be obtained with a sample having a thick P3HT layer and the contribution from the second interface (P3HT/water(air)).

Similarly for the *SPS* and *PPP* polarization, it holds

$$\chi_{eff,SPS}^{thin} = \chi_{eff,SPS}^{thick} + \chi_{eff,SPS}^{second} \quad (S29)$$

$$\chi_{eff,PPP}^{thin} = \chi_{eff,PPP}^{thick} + \chi_{eff,PPP}^{second} \quad (S30)$$

At this point, using the information obtained on the “thick” samples, χ^{thick} , and repeating the fitting procedure also for the “thin” samples, yielding χ^{thin} , we use the above equations to get the contribution from the second interface, χ^{second} . These χ^{second}_{ijk} are analyzed with the same molecular model to obtain the molecular orientation at the second interface, the polymer/water(air) interface. In our work, we were in regime (ii), close to the thin film limit, since we had to employ a configuration with light incident from the substrate, due to the strong attenuation of the IR beam in water. In this case, we had to take into account contribution from the substrate/polymer interface, and subtract it out of the thin film measurements, as in Eqs. S28-S30. We thus carried out a complete analysis also in the case of a thick sample (approx. 300 nm) (Figure S5), which allowed us to isolate specific SFG contributions of the CaF₂/P3HT interface and to consider them in the analysis of the thin samples. It should be noted that if the sample were thick enough, the two sets of spectra should have been identical, since the second polymer/water(air) interface would give no contribution. The small differences in the intensity ratios among different polarization combinations may be due to a finite thickness of the polymer film, to interference effects in the mid-IR (which is not so strongly absorbed in the polymer), or to partial penetration of water within the polymer film, slightly affecting the molecular orientation. However, the angular plots of Δ_r for the two cases [Figure S5(c) and (d)] are similar, confirming that our analysis is robust and gives meaningful results.

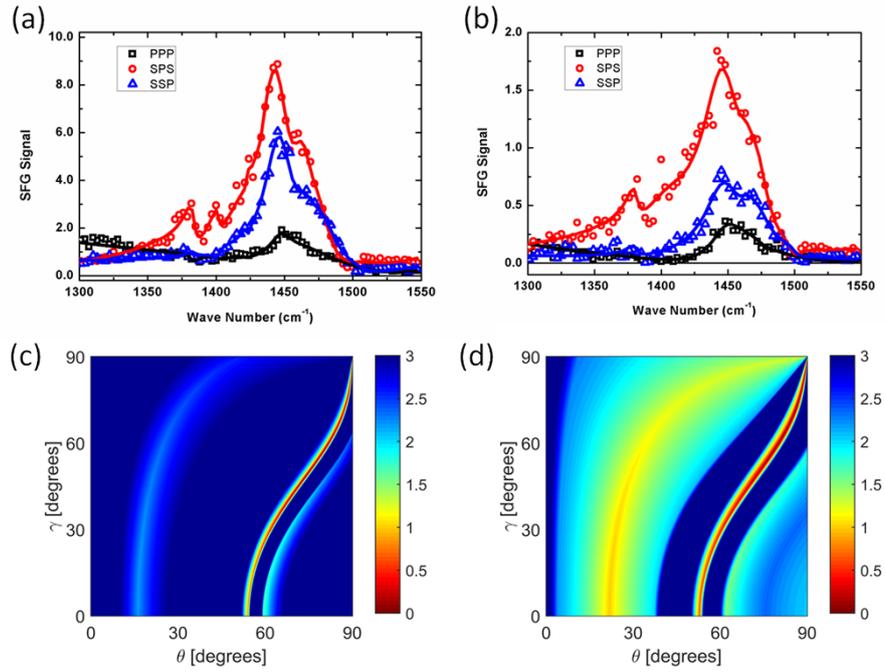


Figure S5. SFG spectra and corresponding angular plots of Δ_r for a thick polymer layer (approximately 300 nm), exposed to air (panels (a) and (c)) and water (panels (b) and (d)), with light beams incident from the CaF_2 substrate. From these data the specific contribution of the substrate ($\text{CaF}_2/\text{P3HT}$ interface) was calculated, and taken into account in the analysis of thin films.

Finally, for the sake of completeness, we report also the results obtained in the study of the air/polymer interface for a thick sample, with the beams incoming from the air side, fully confirming the results reported in the main text (Figure S6).

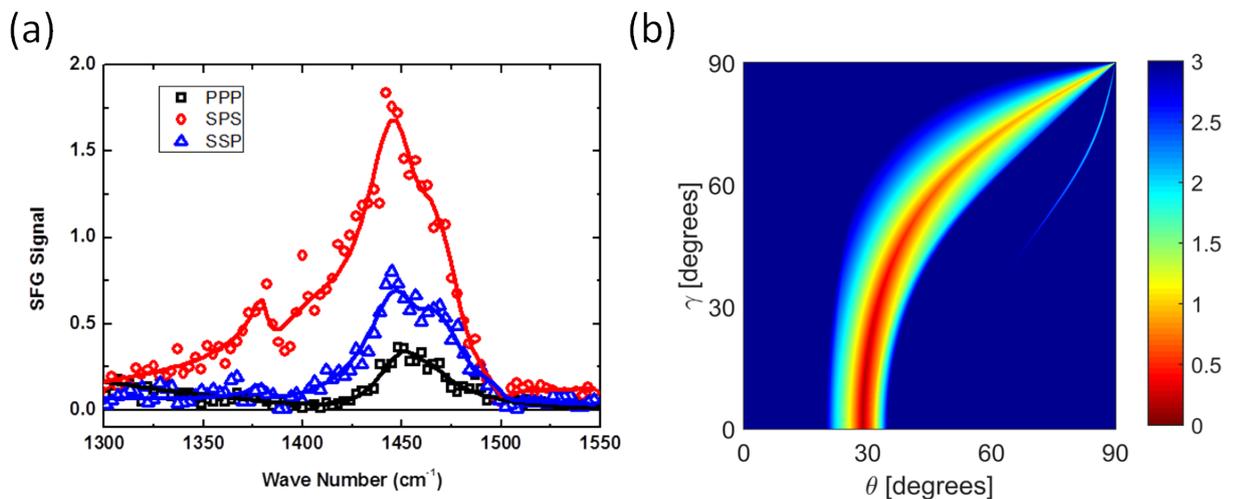


Figure S6. SFG spectra and angular plot of Δ_r obtained for a thick rr-P3HT sample, for light beams incident from the air/polymer interface.

Surface energy calculation from molecular dynamics simulations

First of all we built the reference bulks for P3HT and water. The P3HT bulk polymer was formed by 60 chains with the same chemical termination of the polymer at the surface (methyl group). The bulk was equilibrated at room temperature and pressure (NPT ensemble) for 10 ns. After the equilibration, the volumetric mass density (i.e. the polymer mass per unit volume) is found to be $\rho=1.06 \text{ g/cm}^3$, in good agreement with the experimental values of 1.01-1.10 g/cm^3 [18, 19]. The average cell size is $90.68 \times 124.86 \times 44.46 \text{ \AA}^3$. The bulk P3HT was then relaxed by a 10000-steps conjugate gradient minimization as implemented in NAMD in order to obtain a reference energy of the bulk polymer, $E_{bulk}=N_p\mu_p$ where N_p is the number of molecules in the considered polymer crystal and μ_p is the energy density, i.e. the energy per molecule.

The same protocol was applied to the bulk water: we have equilibrated $N_w=3921$ water molecules at ambient conditions and then relaxed the system by a 10000-steps conjugate gradient minimization, to obtain the reference energy for the water $E_{water}=N_w\mu_w$, with μ_w being the energy per water molecule.

As for the free surfaces, the P3HT-vacuum systems were only relaxed by a 10000-steps conjugate gradient minimization in order to leave the chain orientation unaltered (θ and γ angles). The surface energy of the polymer *in vacuo* is then calculated as

$$\sigma_p = \frac{E_p(\gamma, \theta) - N_p\mu_p}{2A(\gamma, \theta)} \quad (\text{S31})$$

where $E_p(\gamma, \theta)$ is the configurational energy of the tilted polymer slab, $N_p\mu_p$ the energy of a bulk polymer containing the same number of molecules N_p and $A(\gamma, \theta)$ is the surface area.

P3HT-water systems were preliminarily equilibrated in the pseudo-NVT ensemble at room temperature in order to allow the water molecules to rearrange around the polymer chains. At this stage, the atomic positions of the P3HT molecules were kept fixed so preserving the initial γ and θ angles. After a very short simulation time (less than 100 ps) the configuration energy converged to a

constant value with oscillations smaller than $0.05 J/m^2$. The resulting configurations were then fully relaxed by a 10000-steps conjugate gradient minimization. The interface energy in water $\sigma_{p/w}$ is defined as

$$\sigma_{p/w} = \frac{E_{p/w}(\gamma, \theta) - N_p \mu_p - N_w \mu_w}{2A(\gamma, \theta)}, \quad (\text{S32})$$

where $E_{p/w}(\gamma, \theta)$ is the configuration energy of polymer/water system, N_w is the number of the water molecules in the system and $A(\gamma, \theta)$ is the area of the interface between the polymer slab and the water.

Finally, the following figure illustrates the schematic representation of the tilting angle θ and its projections over the (x, z) plane, used in the modeling of the surface energy, as detailed in the main text, in order to take into account the oscillating energy cost $\sigma_1(\theta)$ due to the intermolecular staggering.

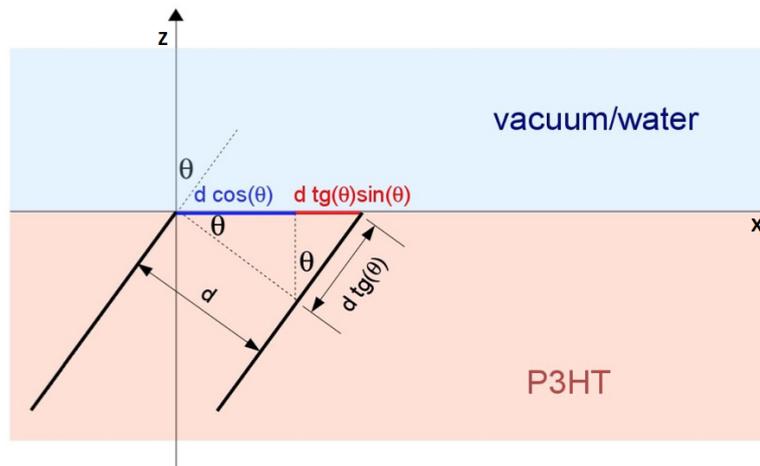


Figure S7. Schematic representation of the polymer orientation in the (x, z) plane used in the modeling of the formation energy. The tilting angle θ and interchain distance d are reported.

References

1. Y. R. Shen, *J. Phys. Chem. C*, 2012, 116, 15505.
2. A. G. Lambert, P. B. Davies, D. J. Neivandt, *Appl. Spectr. Rev.*, 2005, 40, 103.
3. F. Vidal, A. Tadjeddine, *Rep. Progr. Phys.*, 2005, 68, 1095.
4. H. F. Wang, L. Velarde, W. Gan, L. Fu, *Ann. Rev. Phys. Chem.*, 2015, 66, 189.
5. Y. R. Shen, V. Ostroverkhov, *Chem. Rev.*, 2006, 106, 1140.
6. X. Lu, S. A. Spanninga, C. B. Kristalyn, Z. Chen, *Langmuir*, 2010, 26, 14231.
7. C. Chen, J. Wang, M. A. Even, Z. Chen, *Macromolecules*, 2002, 35, 8093.
8. X. Zuang, P. B. Miranda, D. Kim, Y. R. Shen, *Phys. Rev. B*, 1999, 59, 12632.
9. P. Dhar, P. P. Khlyabich, B. Burkhardt, S. T. Roberts, S. Malyk, B. C. Thompson, A. V. Benderskii, *J. Phys. Chem. C*, 2013, 117, 15213.
10. T. C. Anglin, D. B. O'Brien, A. M. Massari, *J. Phys. Chem. C*, 2010, 114, 17629.
11. F. C. B. Maia, P. B. Miranda, *J. Phys. Chem. C*, 2015, 119, 7386.
12. S. Widge, Y. Matsuoka, M. Kurnikova, *J. Molec. Graph. Modell.*, 2008, 27, 34.
13. B. Champagne, D. H. Mosley, J. M. Andre, *J. Chem. Phys.*, 1994, 100, 2034.
14. X. Lu, L. Bolin, Z. Peizhi, X. Gi, L. Dawei, *Soft Matter* 2014, 10, 5390.
15. Y. Tong, Y. Zhao, N. Li, M. Osawa, P. B. Davies, S. Ye, *J. Chem. Phys.* 2010, 133, 034704.
16. G. Li, A. Dhinojwala, M. S. Yeganeh, *J. Phys. Chem. C* 2011, 115, 7554.
17. X. Lu, M. L. Clarke, D. Li, X. Wang, G. Xue, Z. Chen, *J. Phys. Chem. C* 2011, 115, 13759.
18. M. Brinkmann, P. Rannou, *Adv. Funct. Mat.*, 2007, 17, 101.
19. T. J. Prosa, M. J. Winokur, J. Moulton, P. Smith, A. J. Heeger, *Macromolecules*, 1992, 25, 4364.