Data-driven modelling for electrolyte optimisation in dye-sensitised solar cells and photochromic solar cells.

Johan Liotier*[a], Antonio J. Riquelme [a], Valid Mwalukuku [a], Quentin Huaulmé [a], Yann Kervella [a], Renaud Demadrille* [a] and Cyril Aumaître*[a]

[a] IRIG-SyMMES, Université Grenoble Alpes/CEA/CNRS, Grenoble, 38000, France.
E-mail: renaud.demadrille@cea.fr.
E-mail: cyril.aumaitre@cea.fr
E-mail: johan.liotier@livmats.uni-freiburg.de

Table of content

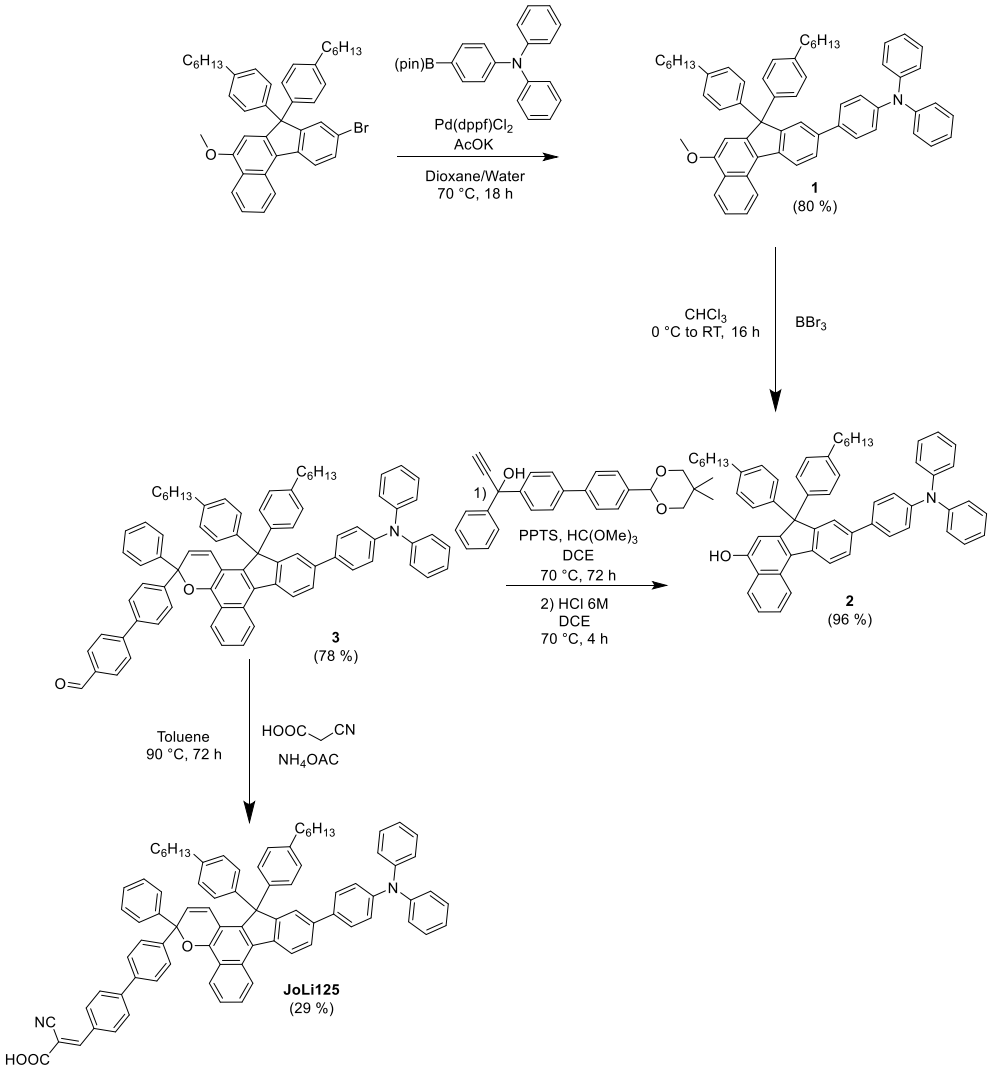
I. Machine learning program

All the different versions of the python program, the jupyter notebook and the raw datas are available in our Github repository

Github : https://github.com/JohanLiotier/MachineLearning_Optimisation.git

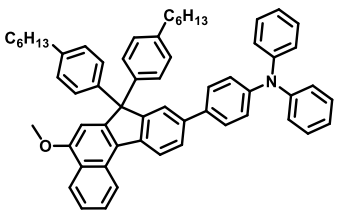
II. Dyes synthesis

9-bromo-7,7-bis(4-hexylphenyl)-5-methoxy-7H-benzo[c]fluorine and 1-(4'-(5,5-dimethyl-1,3-dioxan-2-yl)-[1,1'-biphenyl]-4-yl)-1-phenylprop-2-yn-1-ol have been synthesis according to literature[1,2]



Supplementary figure 1: Synthesis of **JoLi125**.

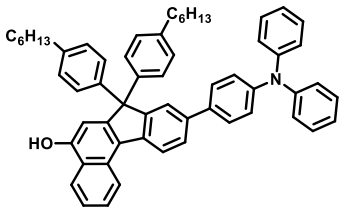
Compound **1**:



9-bromo-7,7-bis(4-hexylphenyl)-5-methoxy-7H-benzo[c]fluorene (190 mg, 0.29 mmol, 1.0 eq) and 4-(Diphenylamino)phenylboronic acid pinacol ester (131 mg, 0.35 mmol, 1.2 eq) were dissolved in dioxane (5 mL). The reaction mixture is degassed by argon bubbling during 20 minutes. $Pd(dppf)Cl_2$ (4.31 mg, 5.9 µmol, 2 mol%) is then added followed by a 1M solution of potassium acetate (0.88 mL,

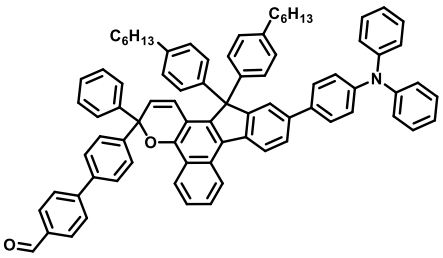
0.88 mmol, 3.0 eq). The reaction mixure is then heated to 70 °C and stirred at this temperature during 18 hours. The mixture is then cooled down to room temperature. Water (5 mL) and ethyl acetate (5 mL) are added and stirred at room temperature for 10 minutes. The phases are separated. The aqueous water is extracted three times with ethyl acetate and the organic phases are combined. The organic phase is washed with water and brine, dried over Na_2SO_4, filtered off and concentrated under vacuum. The resulting crude product is purified by column chromatography (neat Hexane to Hexane/DCM 80/20) to obtain compound **1** as a white foam (190 mg, 0.23 mmol, 80%). **^{1}H-NMR (CDCl3, 400 MHz)** δ (ppm): 8.77 (d, J = 8.4 Hz, 1H), 8.38 (dd, J = 8.4, 1.0 Hz, 1H), 8.32 (d, J = 8.1 Hz, 1H), 7.76 – 7.62 (m, 3H), 7.57 (dd, J = 8.3, 1.0 Hz, 1H), 7.54 – 7.47 (m, 2H), 7.33 – 7.25 (m, 5H), 7.22 (d, J = 8.3 Hz, 4H), 7.19 – 7.11 (m, 6H), 7.11 – 7.01 (m, 6H), 6.93 (s, 1H), 3.96 (s, 3H), 2.65 – 2.50 (m, 4H), 1.66 – 1.55 (m, 4H), 1.46 – 1.23 (m, 12H), 0.98 – 0.84 (m, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 155.73, 153.35, 151.46, 147.71, 146.95, 142.69, 141.32, 140.36, 137.78, 135.29, 130.16, 129.27, 128.30, 128.26, 127.68, 127.34, 126.86, 125.82, 125.74, 124.89, 124.35, 123.98, 123.89, 123.22, 122.87, 122.18, 102.44, 65.35, 55.73, 35.57, 31.73, 31.35, 29.18, 22.62, 14.11. **Elemental Analysis** (calcd, found for $C_{60}H_{59}NO$): C (88.96, 88.19), H (7.34, 7.22), N (1.73, 1.84).

Compound **2**:



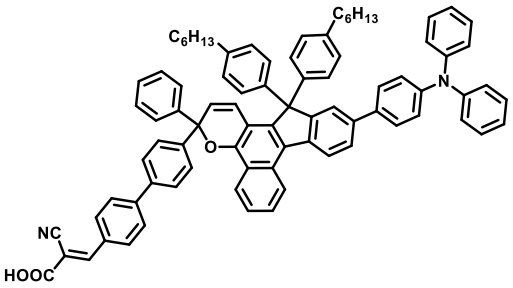
A solution of compound **1** (190 mg, 0.235 mmol, 1.0 eq) in anhydrous chloroform (5 mL) is added dropwise at 0°C BBr3 (1.0 M in DCM, 0.47 mL, 0.47 mmol, 2.0 eq). The reaction mixture is allowed to warm up at room temperature for 20 hours. The reaction mixture is poured on saturated aqueous $NaHCO_3$ solution (10 mL), diluted with DCM (10 mL) and the mixture is stirred at room temperature for 10 minutes. The aqueous layer is extracted with DCM. The combined organic layer is washed with water and brine, dried over Na_2SO_4, filtered off and concentrated under vacuum. The resulting crude product is purified by column chromatography (Hexane/DCM 80/20 to 60/40) to afford compound **2** as a grey foam (180 mg, 0.226 mmol, 96%). **^{1}H-NMR (Acetone-d6, 400 MHz)** δ (ppm): 9.20 (s, 1H), 8.81 (d, J = 8.6 Hz, 1H), 8.37 (dd, J = 10.5, 8.4 Hz, 3H), 7.78 – 7.65 (m, 4H), 7.62 – 7.44 (m, 4H), 7.27 (tt, J = 4.2, 2.0 Hz, 5H), 7.19 (d, J = 8.3 Hz, 5H), 7.11 – 7.02 (m, 15H), 7.01 (s, 2H), 2.58 – 2.47 (m, 4H), 1.53 (dd, J = 14.9, 7.2 Hz, 4H), 1.36 – 1.20 (m, 12H), 0.90 – 0.77 (m, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 153.59, 153.27, 151.92, 147.70, 147.07, 142.94, 141.24, 140.47, 137.43, 130.39, 129.36, 128.22, 128.18, 127.52, 127.45, 125.78, 125.74, 124.57, 124.25, 123.88, 123.85, 123.61, 123.42, 123.03, 122.17, 106.53, 65.06, 35.20, 31.51, 31.35, 22.35, 13.45. **Elemental Analysis** (calcd, found for $C_{59}H_{57}NO$): C (89.01, 88.43), H (7.22, 7.08), N (1.76, 1.16).

Compound **3**:

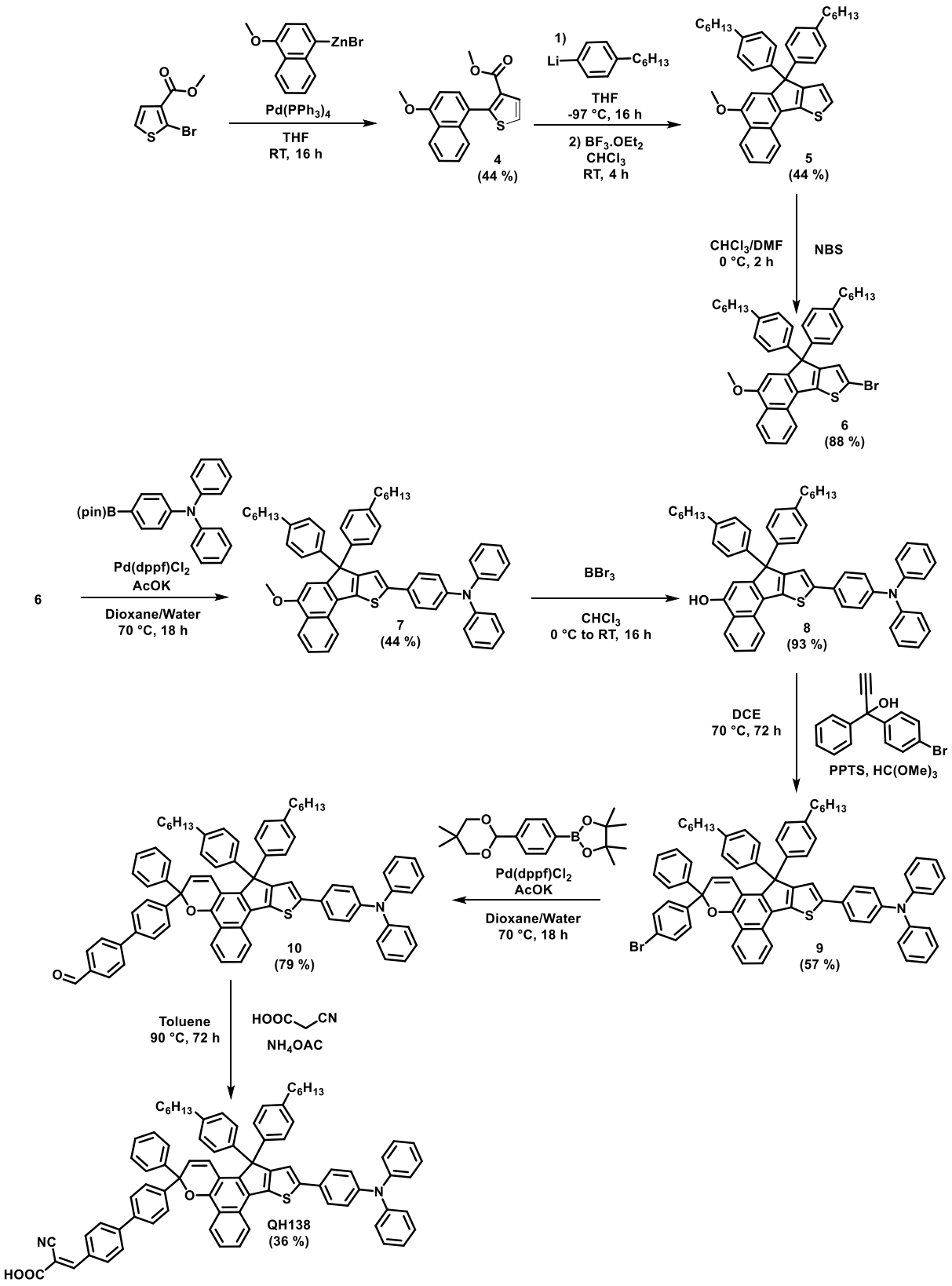


To a solution of compound **2** (150 mg, 0.188 mmol, 1.0 eq) in anhydrous DCE (10 mL) are successively added Compound **4.6** (90.1 mg, 0.226 mmol, 1.2 eq), PPTS (15.8 mg, 62.8 µmol, 0.33 eq) and trimethylorthoformate (0.06 mL, 0.565 mmol, 3.0 eq). The mixture is stirred at 70°C for 72 hours. A 6M hydrochloric acid solution in water (10 mL) and ethanol (5mL) are then added and the reaction mixture is stirred at 60 °C during 4 hours. Once cooled down to room temperature, the reaction mixture is poured on water and DCM is added. The layers are separated. The organic layer is washed with water, dried over Na2SO4, filtered off and concentrated under vacuum. The resulting product is purified by column chromatography (Hexane/DCM: 80/20 to 60/40) to afford compound **3** as a green solid (160 mg, 0.147 mmol, 78%). **^{1}H-NMR (CDCl3, 400 MHz)** δ (ppm): 10.05 (s, 1H), 8.75 (d, J = 8.2 Hz, 1H), 8.52 (d, J = 8.0 Hz, 1H), 8.25 (d, J = 8.2 Hz, 1H), 7.93 (d, J = 8.2 Hz, 1H), 7.72 – 6.95 (m, 36H), 6.90 – 6.76 (m, 1H), 5.88 (t, J = 6.7 Hz, 1H), 2.59 (dd, J = 12.1, 6.6 Hz, 4H), 1.60 (d, J = 6.8 Hz, 4H), 1.41 – 1.22 (m, 12H), 0.89 (dd, J = 14.6, 6.8 Hz, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 191.90, 155.87, 147.69, 147.31, 144.66, 144.08, 141.35, 141.33, 141.28, 141.24, 139.77, 139.70, 139.69, 139.58, 138.73, 135.20, 130.25, 130.14, 129.27, 129.02, 128.93, 128.14, 128.04, 127.81, 127.68, 127.60, 127.33, 127.02, 126.98, 125.57, 125.29, 124.43, 124.40, 124.35, 124.13, 123.93, 123.90, 123.34, 123.31, 122.89, 122.10, 121.24, 115.04, 35.59, 31.75, 31.38, 29.20, 22.63, 14.12. **Elemental Analysis** (calcd, found for $C_{81}H_{71}NO_2$): C (89.22, 81.17), H (6.56, 6.09), N (1.28, 0.52).

Compound **JoLi125**:

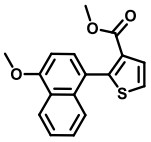


Compound **3** (80.0 mg, 73.4 µmol, 1.0 eq) is dissolved in Toluene (10 mL). Cyanoacetic acid (38.2 mg, 0.449 mmol, 6.1 eq) and ammonium acetate (90.5 mg, 1.17 mmol, 16.0 eq) is then added and the reaction mixture is stirred at 100 °C for 16 h. Water and EtOAc are added to the mixture, which is further stirred at room temperature for 10 minutes. The layers are separated and the aqueous layer is extracted once with EtOAc. The combined organic layer is dried over Na2SO4, filtered off and concentrated under vacuum. The resulting product is purified by column chromatography (silica gel, neat DCM to DCM/MeOH: 90/10) to afford compound **JoLi125** as a grey solid (25 mg, 21.6 µmol, 29%). **^{1}H-NMR (CDCl3, 400 MHz)** δ (ppm): 8.82 (d, J = 9.0 Hz, 1H), 8.58 – 8.50 (m, 1H), 8.36 (d, J = 8.0 Hz, 1H), 8.22 (br, 1H), 8.03 (br, 2H), 7.79 – 7.39 (m, 17H), 7.39 – 7.17 (m, 13H), 7.17 – 6.96 (m, 13H), 6.88 (d, J = 9.7 Hz, 1H), 5.96 (d, J = 9.7 Hz, 1H), 2.69 – 2.59 (m, 4H), 1.70 – 1.58 (m, 4H), 1.43 – 1.27 (m, 12H), 1.00 – 0.83 (m, 6H). **HRMS (MALDI-TOF):** calcd. for $C_{84}H_{72}N_2O_3$, 1156.553; found 1156.555.



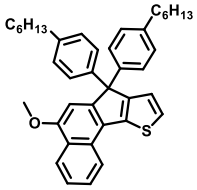
Supplementary Figure 2: Synthesis of **QH138**

Compound **4**:



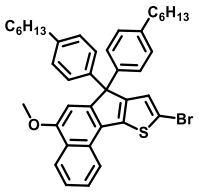
To a solution of 1-bromo-4-methoxynaphthalene (7.50 g, 31.6 mmol, 1.00 eq) in degassed and anhydrous THF (100 mL) is added dropwise at -78°C *n*-BuLi (2.5 M in hexanes, 12.6 mL, 33.2 mmol, 1.05 eq). The reaction mixture is stirred at -78°C for 1 hour before a solution of $ZnBr_2$ (7.84 g, 34.8 mmol, 1.10 eq) in anhydrous THF (20 mL) is added dropwise. The mixture is allowed to warm up to 0°C and is further stirred at this temperature for 1 hour. Methyl 2-bromothiophene-3-carboxylate (6.64 g, 30.1 mmol, 0.95 eq) and $Pd(PPh_3)_4$ (1.10 g, 0.949 mmol, 3.0 mol%) are successively added and the mixture is allowed to warm up at room temperature and is stirred at room temperature for 20 hours. The reaction mixture is poured on 2M aqueous HCl (100 mL) and diluted with diethyl ether (50 mL). **4** is obtained as white crystals (1.11 g, 2.99 mmol, 53%). **^{1}H RMN ($CDCl_3$, 400MHz) δ (ppm):** 8.33 (ddd, J = 10.2, 5.1, 4.1 Hz, 1H), 7.64 – 7.62 (m, 1H), 7.61 – 7.58 (m, 1H), 7.50 (dt, J = 3.0, 1.5 Hz, 1H), 7.47 (t, J = 1.8 Hz, 1H), 7.45 (dd, J = 3.0, 1.5 Hz, 1H), 7.44 – 7.42 (m, 1H), 7.36 (d, J = 5.4 Hz, 1H), 6.87 (d, J = 7.9 Hz, 1H), 4.07 (s, 3H), 3.51 (d, J = 1.8 Hz, 3H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 163.40, 156.11, 149.20, 133.38, 130.36, 129.17, 128.22, 126.85, 125.35, 125.22, 124.51, 123.34, 122.17, 102.91, 55.56, 51.40. **MS (ESI [M + H]$^+$)**: calcd. for $C_{17}H_{14}O_3S$+H, 299.1; found 289.9.

Compound **5**:



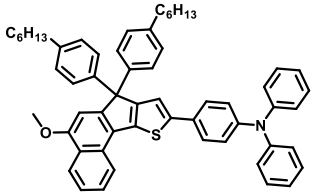
To a solution of 1-Bromo-4-hexylbenzene (7.77 g, 32.2 mmol, 2.30 eq) in anhydrous THF (120 mL) is added dropwise at -95°C n-BuLi (2.5 M in hexane, 12.9 mL, 32.2 mmol, 2.30 eq). The reaction is stirred at this temperature for 45 min before Compound **4** (4.18 g, 14.0 mmol, 1.00 eq) is added in one portion as a solid. The mixture is further stirred at this temperature for 30 min and then allowed to reach RT overnight (19 h). The reaction mixture is poured on HCl 2 M (30 mL), diluted in Et_2O (100 mL) and the mixture is stirred at RT 10 min. The organic layer is washed with water and brine, dried over Na_2SO_4, filtered off and concentrated under vacuum to obtain an intermediate alcohol. The intermediate alcohol is dissolved in anhydrous chloroform (200 mL) and $BF_3.OEt_2$ (4.5 mL, 5.17 g, 35.95 mmol, 2.00 eq) is added dropwise at room temperature. The reaction mixture is stirred 1h30. Water (150 mL) is added and the mixture is stirred at room temperature for 10 minutes. The layers are separated. The organic layer is washed with water and brine, dried over Na_2SO_4, filtered off and concentrated to dryness. The resulting oil is purified by column chromatography (neat PE to PE/DCM 95/5) to afford compound **5** as a grey foam (3.61 g, 6.30 mmol, 44 %). **^{1}H-NMR ($CDCl_3$, 298 K, 400 MHz)** δ (ppm): 8.33 – 8.29 (m, 1H), 8.18 – 8.14 (m, 1H), 7.65 (ddd, J = 8.2, 6.9, 1.3 Hz, 1H), 7.55 – 7.49 (m, 1H), 7.32 (d, J = 4.9 Hz, 1H), 7.19 – 7.15 (m, 4H), 7.10 (d, J = 4.9 Hz, 1H), 7.05 (d, J = 8.4 Hz, 4H), 6.93 (s, 1H), 3.96 (s, 3H), 2.60 – 2.52 (m, 4H), 1.64 – 1.53 (m, 4H), 1.41 – 1.25 (m, 10H), 0.95 – 0.85 (m, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 155.35, 154.71, 152.20, 141.90, 141.62, 140.37, 134.08, 133.89, 128.93, 128.77, 128.70, 128.61, 128.58, 128.18, 128.14, 127.51, 127.30, 126.66, 126.57, 125.37, 125.31, 124.74, 123.20, 122.84, 103.37, 63.99, 55.88, 35.83, 31.99, 31.97, 31.60, 29.39, 22.88, 14.38. **MS (ESI [M + H]$^+$)**: calcd. for $C_{40}H_{44}OS$+H, 573.3; found 573.3.

Compound **6**:



To a solution of **5** (3.61 g, 6.30 mmol, 1.00 eq) in chloroform (50 mL) is added dropwise at 0 °C a solution of NBS (1.16 g, 6.49 mmol, 1.03 eq) in DMF (15 mL). The reaction mixture is stirred at 0 °C for 2 hours. A saturated $NaHCO_3$ solution (50 mL) is then added and the mixture is stirred at RT for 10 minutes. Layers are separated and the organic layer is washed thrice with water, dried over Na_2SO_4, filtered off and concentrated to dryness. The crude product is purified by column chromatography (neat PE) to afford compound **6** as a light yellow oil (3.65 g, 5.60 mmol, 88 %). **^{1}H RMN ($CDCl_3$, 400MHz) δ (ppm):** 8.32 (d, J = 8.5 Hz, 1H), 7.99 (d, J = 8.3 Hz, 1H), 7.65 (ddd, J = 8.2, 6.9, 1.2 Hz, 1H), 7.54 (ddd, J = 8.2, 6.9, 1.1 Hz, 1H), 7.17 (d, J = 8.2 Hz, 4H), 7.12 (s, J = 4.4 Hz, 1H), 7.09 (d, J = 8.2 Hz, 4H), 6.92 (s, 1H), 3.97 (s, 3H), 2.68 – 2.54 (m, 5H), 1.61 (dt, J = 15.5, 7.6 Hz, 4H), 1.43 – 1.28 (m, 13H), 0.91 (t, J = 6.8 Hz, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 154.82, 153.81, 151.17, 141.77, 141.03, 140.44, 128.55, 128.53, 127.89, 127.35, 127.02, 126.12, 125.71, 125.38, 125.15, 124.37, 123.12, 112.07, 103.04, 64.69, 55.86, 35.66, 31.81, 31.42, 29.22, 22.71, 14.20. **MS (ESI [M + H]$^+$)**: calcd. for $C_{40}H_{43}OS$+H, 651.2; found 651.2.

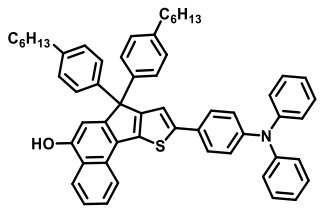
Compound **7**:



Compound **6** (1.36 g, 2.09 mmol, 1.0 eq) and 4-(Diphenylamino)phenylboronic acid pinacol ester (0.798 g, 2.15 mmol, 1.03 eq) were dissolved in dioxane (20 mL). The reaction mixture is degassed by argon bubbling during 20 minutes. $Pd(dppf)Cl_2$ (121 mg, 165 µmol, 2 mol%) is then added followed by a 1M solution of potassium acetate (6.26 mL, 6.26 mmol, 3.0 eq). The reaction mixure is then heated to 70 °C and stirred at this temperature during 18 hours. The mixture is then cooled down to room temperature. Water (10 mL) and ethyl acetate (10 mL) are added and stirred at room temperature for 10 minutes. The phases are separated. The aqueous water is extracted three times with ethyl acetate and the organic phases are combined. The organic phase is washed with water and brine, dried over Na_2SO_4, filtered off and concentrated under vacuum. The resulting crude product is purified by column chromatography (neat Hexane to Hexane/DCM 80/20) to obtain compound **7** as a white foam (753 mg, 0.922 mmol, 44 %). **^{1}H RMN ($CDCl_3$, 400MHz)** δ (ppm): 8.33 (d, J = 8.3 Hz, 1H), 8.20 (d, J = 8.2 Hz, 1H), 7.67 (ddd, J = 8.2, 6.9, 1.2 Hz, 1H), 7.60 – 7.52 (m, 3H), 7.33 – 7.27 (m, 5H), 7.26 – 7.22 (m, 4H), 7.18 – 7.14 (m, 4H), 7.13 – 7.04 (m, 8H), 6.96 (s, 1H), 3.98 (s, 3H), 2.64 – 2.54 (m, 4H), 1.68 – 1.56 (m, 4H), 1.44 – 1.24 (m, 13H), 0.91 (t, J = 6.8 Hz, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 156.06, 154.60, 151.63, 147.65, 146.98, 145.79, 141.60, 141.57, 138.74, 129.65, 129.40, 128.51, 128.07, 127.22, 127.15, 126.72, 126.14, 125.32, 125.20, 124.73, 124.45, 124.14, 123.07, 123.05, 118.01, 103.24, 64.25,

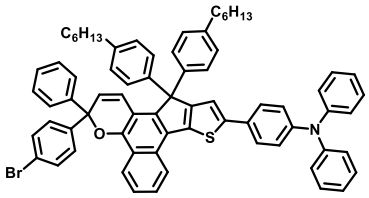
55.89, 35.70, 31.84, 31.47, 29.25, 22.73, 14.24. **MS (ESI [M + H]$^+$)**: calcd. for $C_{58}H_{57}NOS$+H, 816.4; found 816.5.

Compound **8**:



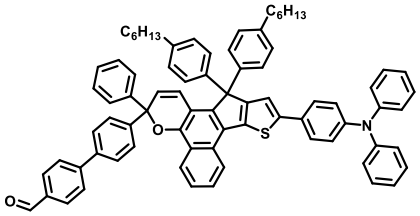
A solution of compound **7** (730 mg, 0.890 mmol, 1.0 eq) in anhydrous chloroform (25 mL) is added dropwise at 0°C BBr3 (1.0 M in DCM, 2.20 mL, 2.20 mmol, 2.0 eq). The reaction mixture is allowed to warm up at room temperature for 20 hours. The reaction mixture is poured on saturated aqueous $NaHCO_3$ solution (20 mL), diluted with DCM (20 mL) and the mixture is stirred at room temperature for 10 minutes. The aqueous layer is extracted with DCM. The combined organic layer is washed with water and brine, dried over Na_2SO_4, filtered off and concentrated under vacuum. The resulting crude product is purified by column chromatography (Hexane/DCM 80/20 to 60/40) to afford compound **8** as a yellow foam (673 mg, 0.831 mmol, 93%). **^{1}H RMN ($CDCl_3$, 400MHz)** δ (ppm): 8.97 (dd, *J* = 1.6, 0.9 Hz, 1H), 8.30 (d, *J* = 6.5 Hz, 1H), 8.15 (d, *J* = 6.4 Hz, 1H), 7.61 (s, 3H), 7.46 (d, *J* = 17.5 Hz, 2H), 7.35 – 7.16 (m, 8H), 7.16 – 6.98 (m, 12H), 6.91 (s, 1H), 2.58 (t, *J* = 7.7 Hz, 4H), 1.60 (dd, *J* = 14.4, 7.2 Hz, 4H), 1.48 – 1.23 (m, 10H), 0.92 (qd, *J* = 5.2, 4.1 Hz, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 155.54, 152.66, 152.12, 147.73, 146.93, 145.26, 141.97, 141.16, 138.94, 129.86, 129.19, 128.14, 128.02, 127.44, 126.74, 125.86, 125.30, 124.21, 123.36, 122.87, 118.16, 107.29, 63.97, 35.53, 31.80, 31.65, 29.15, 22.60, 13.65. **MS (ESI [M + H]$^+$)**: calcd. for $C_{57}H_{55}NOS$+H, 802.4; found 802.5.

Compound **9**:



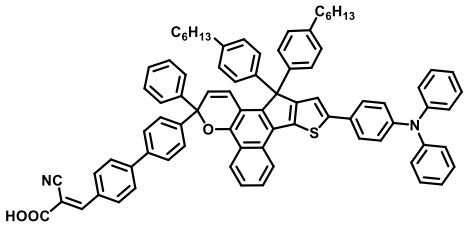
General procedure for chromenisation was followed with **8** (672 mg, 0.838 mmol, 1.00 eq), 1-(4-bromophenyl)-1-phenylprop2-yn-1-ol (481 mg, 1.67 mmol, 2.00 eq), PPTS (42 mg, 0.168 mmol, 20 mol%), $HC(OMe)_3$ (0.17 mL, 1.67 mmol, 2.00 eq) and DCE (20 mL) to afford compouns **9** as a green solid (517 mg, 0.482 mmol, 57 %). **^{1}H RMN ($CDCl_3$, 400MHz)** δ (ppm): 8.41 (d, J = 8.3 Hz, 1H), 8.16 (d, J = 8.2 Hz, 1H), 7.68 – 7.59 (m, 1H), 7.56 (ddd, J = 8.1, 5.8, 2.3 Hz, 1H), 7.51 (d, J = 8.6 Hz, 2H), 7.41 (dt, J = 3.8, 2.1 Hz, 2H), 7.39 – 7.34 (m, 2H), 7.34 – 7.20 (m, 17H), 7.14 (d, J = 7.6 Hz, 4H), 7.12 – 7.03 (m, 9H), 6.83 (d, J = 9.8 Hz, 1H), 5.93 – 5.84 (m, 1H), 2.63 (dd, J = 15.2, 6.3 Hz, 4H), 1.71 – 1.60 (m, 4H), 1.46 – 1.29 (m, 13H), 0.98 – 0.88 (m, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 158.75, 147.63, 147.55, 147.06, 146.79, 146.33, 144.00, 143.44, 141.64, 141.62, 138.49, 138.32, 137.96, 131.25, 129.46, 129.41, 129.17, 129.06, 128.78, 128.39, 128.36, 128.33, 128.27, 128.00, 127.77, 127.54, 127.09, 127.06, 126.99, 126.16, 125.80, 125.44, 125.05, 125.00, 124.46, 124.07, 123.10, 122.91, 121.85, 117.48, 115.86, 81.93, 63.92, 35.74, 31.87, 31.57, 31.52, 29.31, 29.29, 22.77, 22.76, 14.28, 14.26. **MS (ESI [M + H]$^+$)**: calcd. for $C_{72}H_{64}NOSBr$+H, 1070.4; found 1070.4.

Compound **10**:



A solution of compound **9** (0.369 mg, 0.344 mmol, 1.00 eq) in a mixture of 1,4-dioxane (20 mL) and 1M aqueous AcOK (0.861 mL, 0.861 mmol, 2.50 eq) is degassed 20 minutes by gentle bubbling with Argon. 4-(4,4,5,5-tetramethyl-1,3,2- dioxaborolan-2-yl)-benzaldehyde (120 mg, 0.517 mmol, 1.2 eq) and Pd(dppf)Cl2 (14 mg, 17 µmol, 2.0mol%) are successively added and the reaction mixture is stirred at 80°C for 16h. Water and EtOAc are added to the mixture, which is further stirred at room temperature for 10 minutes. The layers are separated and the aqueous layer is extracted once with EtOAc. The combined organic layer is dried over Na2SO4, filtered off and concentrated under vacuum. The resulting residue is purified by column chromatography (silica gel, Hexane/DCM 90/10 to 70/30) to afford compound **10** as a green solid (302 mg, 0.275 mmol, 79 %). **^{1}H RMN ($CDCl_3$, 400MHz)** δ (ppm): 10.06 (s, 1H), 8.45 (d, J = 8.2 Hz, 1H), 8.14 (d, J = 7.9 Hz, 1H), 7.98 – 7.89 (m, 2H), 7.73 – 7.67 (m, 2H), 7.64 – 7.59 (m, 1H), 7.55 (ddd, J = 8.2, 5.5, 1.2 Hz, 1H), 7.50 (s, 5H), 7.46 (ddd, J = 8.6, 3.5, 2.0 Hz, 3H), 7.33 – 7.18 (m, 14H), 7.15 – 7.02 (m, 12H), 6.83 (d, J = 9.8 Hz, 1H), 5.96 (d, J = 9.8 Hz, 1H), 2.66 – 2.56 (m, 4H), 1.69 – 1.58 (m, 4H), 1.40 – 1.26 (m, 10H), 0.95 – 0.84 (m, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 191.83, 158.72, 147.57, 147.05, 146.69, 144.83, 144.21, 141.63, 141.56, 138.85, 138.49, 138.34, 135.30, 130.30, 129.39, 129.15, 128.83, 128.79, 128.42, 128.37, 128.34, 128.30, 128.25, 127.86, 127.80, 127.74, 127.65, 127.12, 127.06, 126.13, 125.77, 125.76, 125.42, 124.97, 124.45, 124.44, 124.35, 124.04, 123.09, 123.04, 122.96, 117.48, 82.14, 35.72, 31.86, 31.52, 31.51, 29.29, 29.28, 22.74, 14.24. **MS (ESI [M + H]$^+$)**: calcd. for $C_{79}H_{69}NO_2S$+H, 1096.5; found 1096.5.

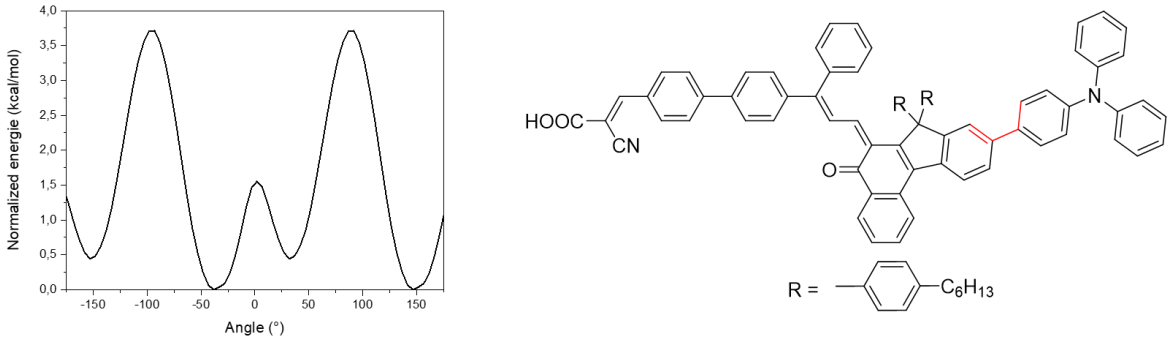
Compound **QH138**:



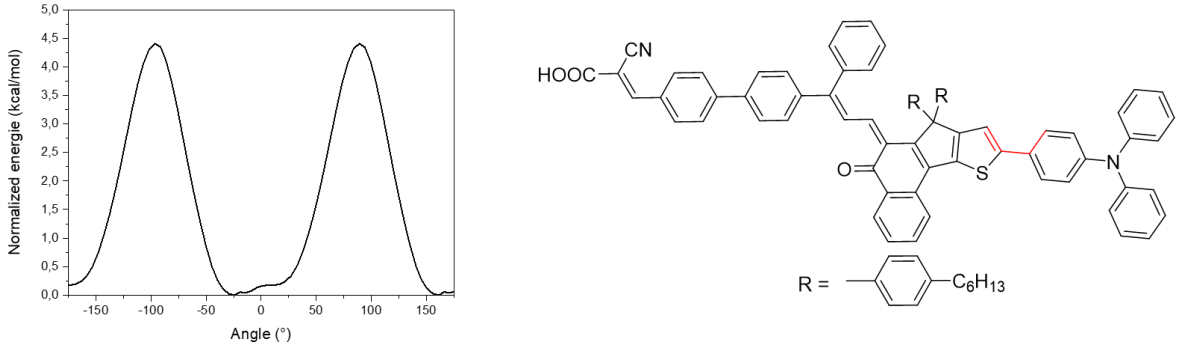
Compound **10** (302 mg, 275 µmol, 1.0 eq) is dissolved in Toluene (10 mL). Cyanoacetic acid (234 mg, 2.75 mmol, 10 eq) and ammonium acetate (40.6 mg, 0.527 mmol, 2.0 eq) is then added and the reaction mixture is stirred at 100 °C for 72 h. Water and EtOAc are added to the mixture, which is further stirred at room temperature for 10 minutes. The layers are separated and the aqueous layer is extracted once with EtOAc. The combined organic layer is dried over Na2SO4, filtered off and concentrated under vacuum. The resulting product is purified by column chromatography (silica gel, neat DCM to DCM/MeOH: 90/10) to afford compound **QH138** as a green solid (122 mg, 105 µmol, 38 %). **^{1}H RMN ($CDCl_3$, 400MHz)** δ (ppm): 8.48 (d, J = 8.2 Hz, 1H), 8.27 (s, 1H), 8.13 (t, J = 8.4 Hz, 3H), 7.79 (d, J = 8.2 Hz, 2H), 7.70 – 7.60 (m, 3H), 7.56 (t, J = 8.2 Hz, 5H), 7.51 – 7.45 (m, 2H), 7.42 (s, 1H), 7.31 – 7.19 (m, 17H), 7.19 – 7.13 (m, 8H), 7.13 – 7.07 (m, 9H), 7.07 – 6.99 (m, 4H), 6.87 (d, J = 9.9 Hz, 1H), 6.06 (d, J = 9.8 Hz, 1H), 2.68 – 2.56 (m, 4H), 1.71 – 1.58 (m, 4H), 1.48 – 1.27 (m, 12H), 0.91 (dt, J = 9.5, 7.1 Hz, 6H). **^{13}C-NMR (CDCl3, 100 MHz)** δ (ppm): 159.03, 147.64, 147.46, 147.04, 146.85, 146.19, 144.70, 144.23, 141.32, 141.30, 138.62, 138.47, 138.45, 137.47, 131.17, 129.51, 129.11, 128.72, 127.99, 127.96, 127.88, 127.63, 127.60, 127.35, 127.22, 126.92, 126.85, 126.58, 125.79, 125.52, 125.09,

125.02, 124.55, 124.15, 123.88, 122.82, 122.74, 122.52, 117.73, 115.92, 82.11, 63.85, 35.48, 31.74, 31.64, 31.63, 29.18, 24.77, 24.57, 24.37, 24.17, 23.97, 22.54, 20.56, 13.49. **MS (ESI [M + H]$^+$)**: calcd. for $C_{82}H_{70}N_2O_3S$+H, 1163.5; found 1163.5.

III. Torsion angle analysis

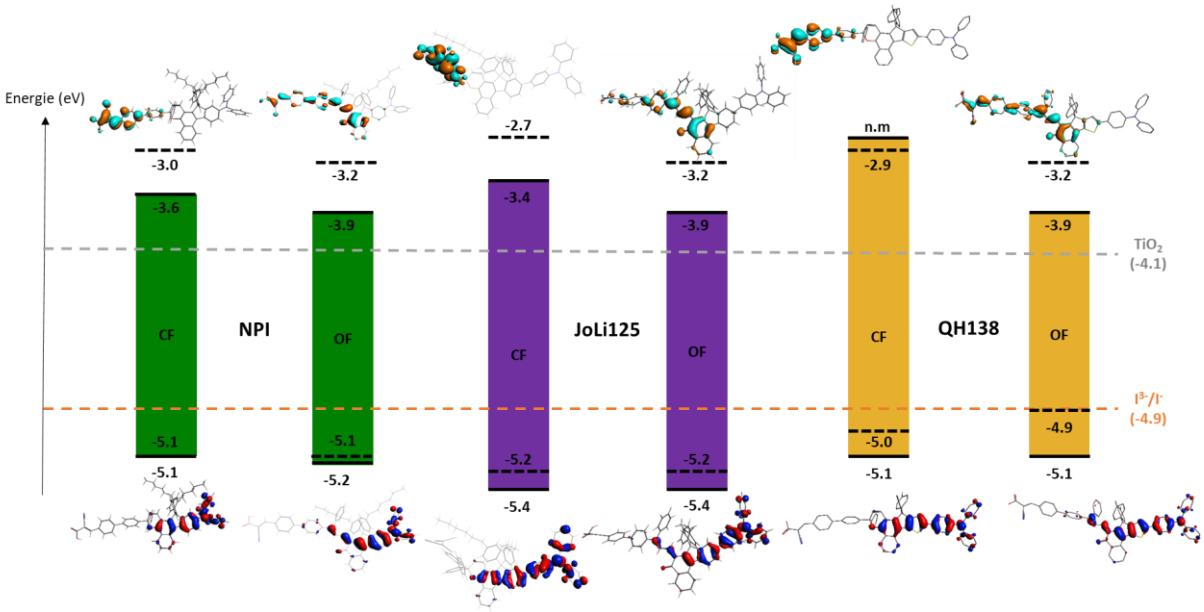


Supplementary figure 3: Binding energy as a function of the torsion angle of the bond between the indenonaphthalene unit and the triphenylamine unit of **JoLi125** calculated with the RevPBE functional (base: tzvp).



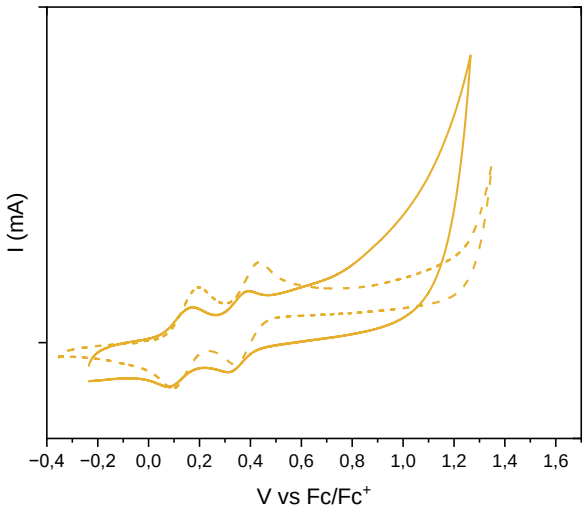
Supplementary figure 4 : Binding energy as a function of the torsion angle of the bond between the indenonaphthalene unit and the triphenylamine unit of **QH138** calculated with the RevPBE functional (base: tzvp).

IV. Energy levels estimated with cyclic voltammetry and with DFT calculation

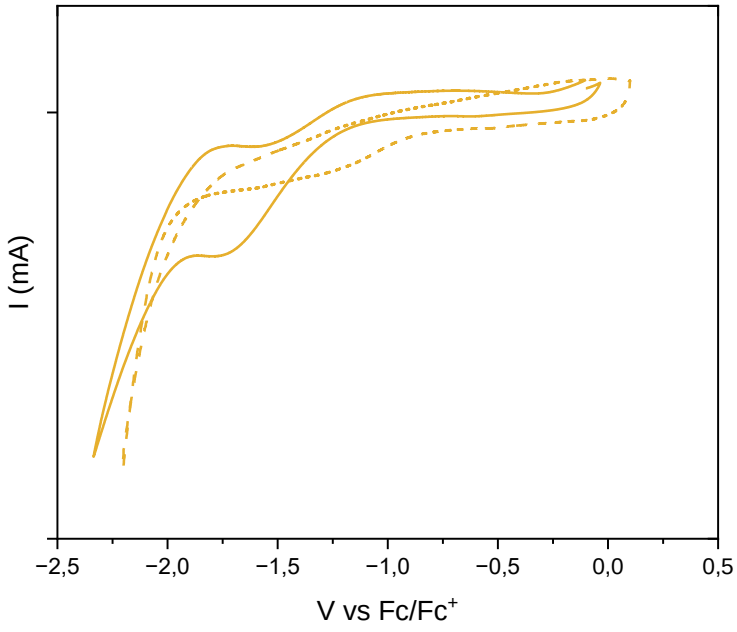


Supplementary figure 5: Energy levels estimated by cyclic voltammetry for the dyes NPI, JoLi125 and QH138 in closed form (CF) and open form (OF). (DCM, tBAPF6, at 25°C for CF and 0°C for OF, 100 $mV.s^{-1}$)

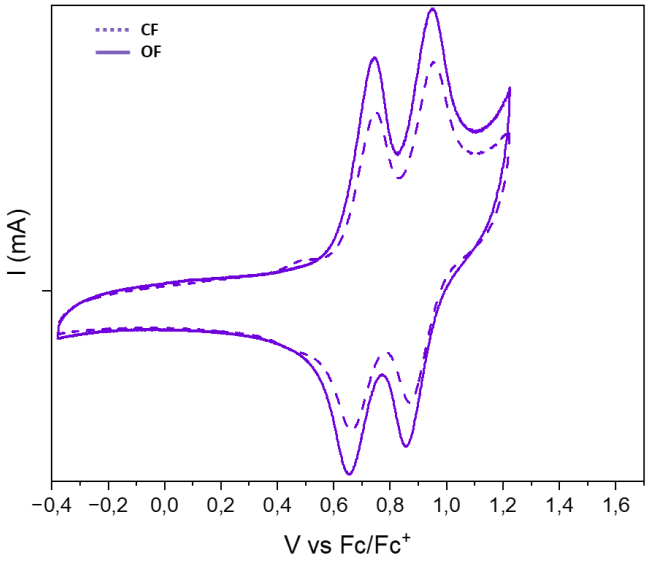
a. Cyclic voltamperometry of the dyes



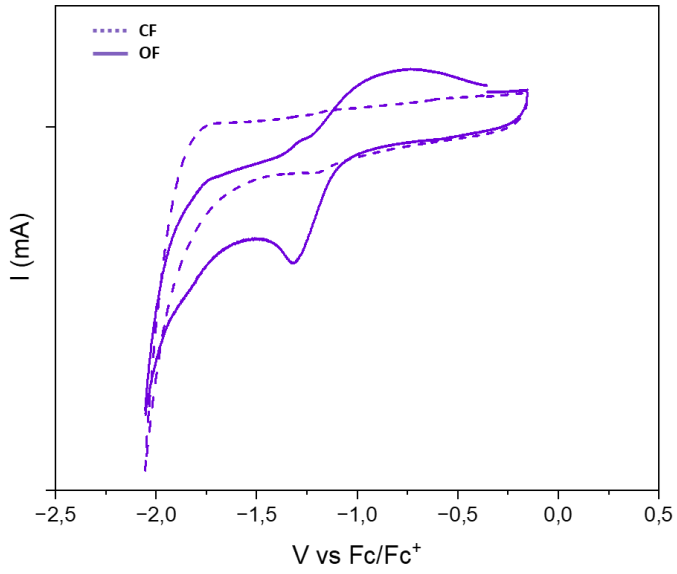
Supplementary figure 6: Cyclic voltammetry trace of compound **QH138** (-0.4 V to 1.6 V, DCM, $tBAPF_6$, 25 °C for the CF and 0 °C for the OF, 100 $mV.s^{-1}$) in the dark (dashed line) and after irradiation (plain line).



Supplementary figure 7: Cyclic voltammetry trace of compound **QH138** (-2.4 V to 0.0 V, DCM, $tBAPF_6$, 25 °C for the CF and 0 °C for the OF, 100 $mV.s^{-1}$) in the dark (dashed line) and after irradiation (plain line).



Supplementary figure 8: Cyclic voltammetry trace of compound **JoLi125** (-0.4 V to 1.6 V, DCM, tBAPF6, 25 °C for the CF and 0 °C for the OF, 100 $mV.s^{-1}$) in the dark (dashed line) and after irradiation (plain line).

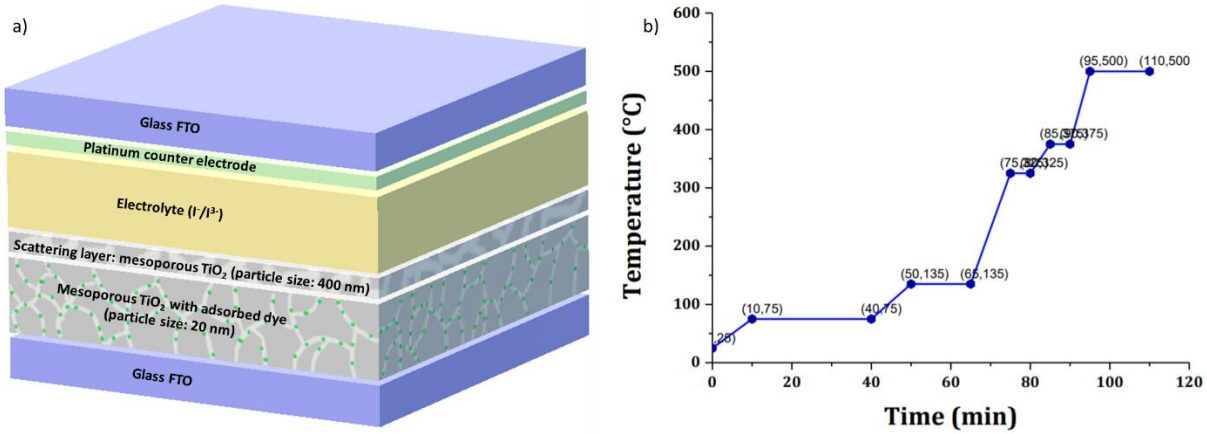


Supplementary figure 9: Cyclic voltammetry trace of compound **JoLi125** (-2.4 V to 0.0 V, DCM, $tBAPF_6$, 25 °C for the CF and 0 °C for the OF, 100 $mV.s^{-1}$) in the dark (dashed line) and after irradiation (plain line).

V. Cells fabrication and characterisation

a. Device fabrication

The device architecture is presented in Supplementary Figure 10a. The solar cells were prepared using the following procedure. TiO_2 thin films with a specific thickness and a total area of 0.36 cm^2 were screen printed in Solaronix (Switzerland) using a TiO_2 nanoparticle paste (Ti-Nanoxide HT/SP). Throughout the manuscript, opaque device refers to a device that includes an additional TiO_2 layer of about 3–4 µm in thickness above the mesoporous TiO_2 (Solaronix; Ti-Nanoxide R/SP). The active area of the solar cells (0.36 cm^2) were estimated from the printing masks and re-measured with a calliper. Beforehand, the electrodes were cleaned with absolute ethanol and dried under an argon flux. These photoanodes were then treated by immersion into a freshly prepared 4.1 mmol l^{-1} TiO_2 aqueous suspension at 70 °C for 20 min. The electrodes were then cooled to room temperature and rinsed with distilled water then absolute ethanol, followed by drying under an argon flux. The electrodes were then sintered under air at 500 °C for 20 min, following the heating procedure reported in Supplementary Figure 10b. The photoanodes were then cooled down to 80 °C and sensitized through immersion in the dyeing solution for 16 h at room temperature in the dark ([Dye] = 0.2 M; [chenodeoxycholic acid (CDCA)] as indicated; $CHCl_3$/tBuOH = 1/1 (vol/vol)). The drilled counter electrodes were coated with a thin layer of platisol (Solaronix) and charred under air at 500 °C using the same heating procedure as presented in Supplementary Figure 10b. The sensitized photoanode was rinsed with dichloromethane and absolute ethanol and dried with an argon flux. Both electrodes were then sealed together using a Surlyn thermoglueing polymer (60 µm thick) using a heating press at 105 °C for 16 s. The cell was then filled with an appropriate acetonitrile-based electrolyte via the pre-drilled hole using a vacuum pump. The electrolyte injection hole on the counter electrode was then sealed with the aid of Surlyn underneath the thin glass cover using heat. A contact along the cell edges was created.

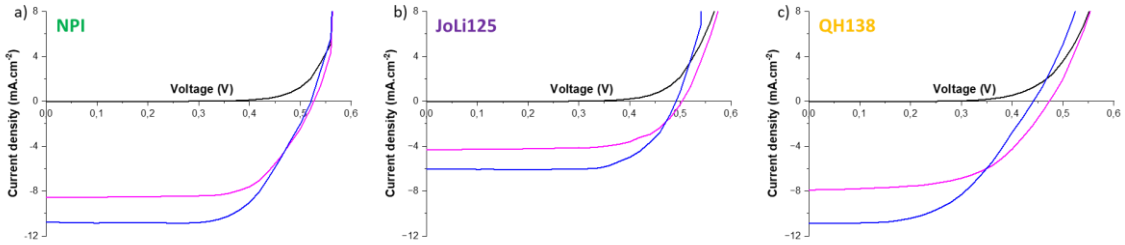


Supplementary Figure 10: a) Device architecture; b) heating procedure for electrode calcination.

Before measurements, the AM 1.5 G simulator (Newport class AAA) was calibrated using a reference silicon photodiode equipped with an infrared-cut-off filter (KG3; Schott). This reference photodiode consisted of a readout device and a 2 cm × 2 cm calibrated solar cell made from monocrystalline silicon with a KG3 window. The cell was also equipped with a thermocouple assembled in accordance with IEC 60904-2. The certification is accredited by the National Institute of Standards and Technology to the ISO-17025 standard and is traceable to the National Renewable Energy Laboratory. The current–voltage characteristics of the cells were measured under dark and under the AM 1.5 G (1,000 W m^{-2}) irradiation condition, which was achieved by applying an external potential bias to the cell while measuring the generated photocurrent with a Keithley model 2400 digital source meter. Measurement

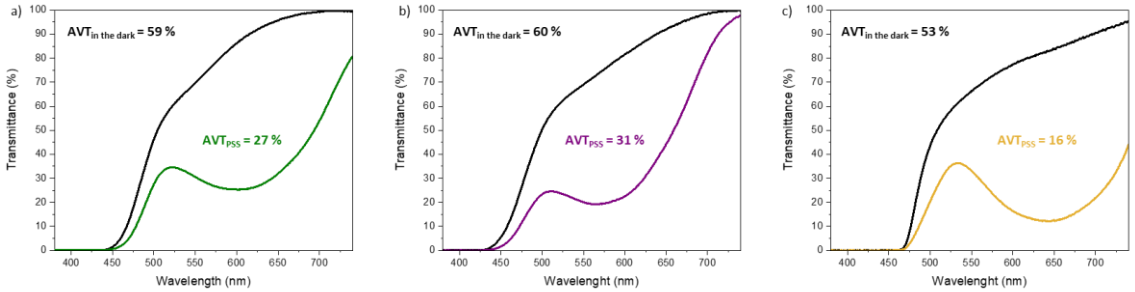
for the cells was from +0.7 to −0.2 V, divided into 45 points, with a speed of 20 mV s^{-1}. The devices were masked before the measurements to attain an illuminated active area of 0.36 cm^2.

b. J-V curves of NPI, JoLi125, QH138



Supplementary figure 11: J=f(V) curves in the dark (black), after a few seconds of irradiation (pink) and at PSS (blue) for cells made with compound **NPI**, **JoLi125**, **QH138** (Dye/CDCA ratio: 1/10).

c. Transmittance curves of the cells made with NPI, JoLi125, QH138



Supplementary figure 12: Transmittance in the dark (black) and at the PSS (coloured) for cells made with compound **NPI**, **JoLi125**, **QH138** (Dye/CDCA ratio: 1/10).

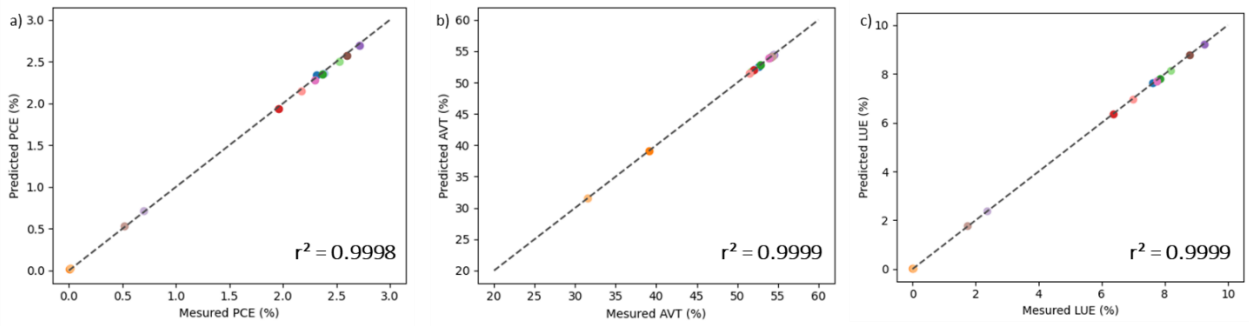
VI. Optimisation of the iodine based electrolyte

a. First run of experiments characterisation for the optimisation of the iodine electrolyte

Experiment Nb	[I_2] (mmol·L^{-1})	[LiI] (mol·L^{-1})	[LiTFSI] (mol·L^{-1})	PCE (%)	AVT (%)
1-1	90	0.5	0.15	2.31	52.64
1-2	120	0.5	0.15	2.38	52.13
1-3	90	0.1	0.15	0.0022	39.12
1-4	120	0.1	0.15	0.0107	31.5
1-5	90	0.3	0	2.37	52.89
1-6	120	0.3	0	2.52	51.75

1-7	90	0.3	0.3	1.96	52
1-8	120	0.3	0.3	2.17	51.47
1-9	60	0.5	0	2.72	54.48
1-10	60	0.1	0	0.69	54.44
1-11	60	0.5	0.3	2.60	54.1
1-12	60	0.1	0.3	0.52	54.26
1-13	60	0.3	0.15	2.29	53.88

Table S1 : Table listing the conditions used to make the electrolytes for the 13 experiments in the first optimisation round as well as the PCE, AVT before irradiation and LUE_p obtained for each electrolyte.

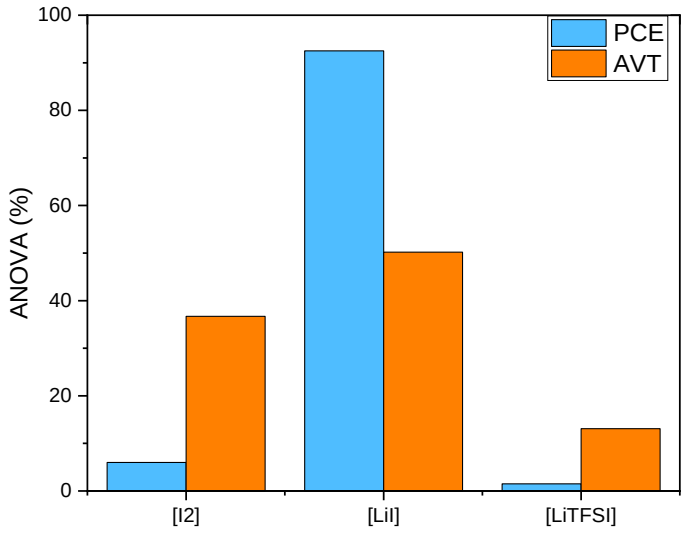


Supplementary figure 13: Verification curves for the regression obtained for the PCE (a), AVT (b) and LUE_p (c) analysis on the first set of experiments. It does correspond that our learning curves for the fitting of gamma, epsilon and c hyperparameters are sufficient for a good correlation in our model

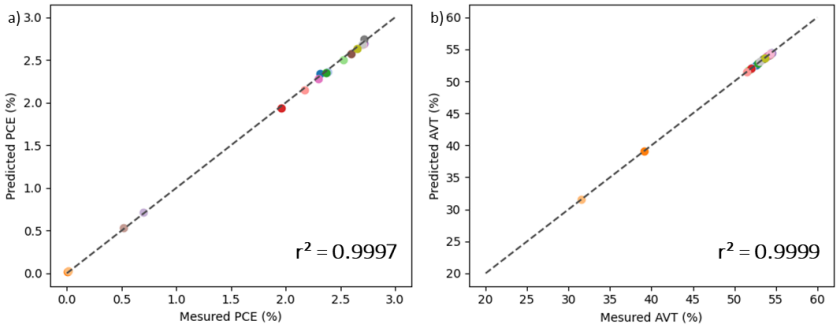
b. Second run of experiments characterisation for the optimisation of the iodine electrolyte

Experiment Nb	$[I_2]$ (mmol·L^{-1})	[LiI] (mol·L^{-1})	[LiTFSI] (mol·L^{-1})	PCE (%)	AVT (%)	LUE_p
2-1	60	0.4	0	2.70	54.32	9.9
2-2	60	0.45	0.05	2.72	53.51	9.1
2-3	60	0.4	0.1	2.70	53.05	9.0
2-4	60	0.5	0.1	2.66	53.72	8.9

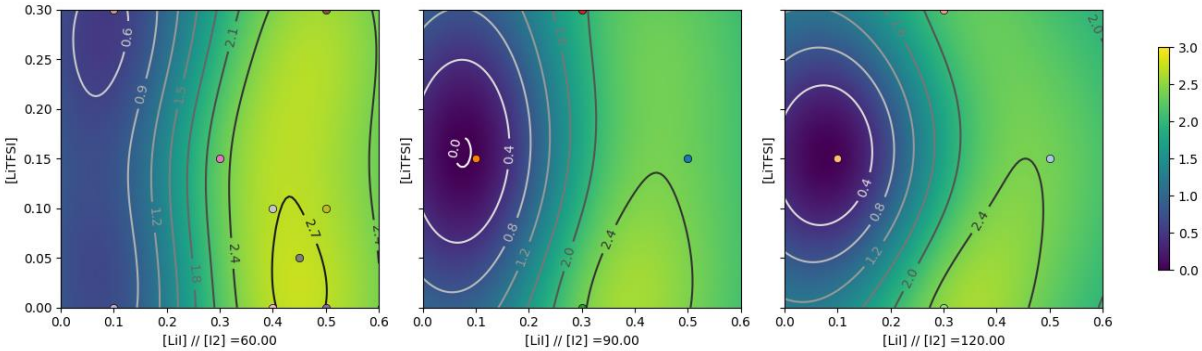
Table S2: Table listing the conditions used to make the electrolytes for the 4 experiments in the second optimisation round, as well as the PCE and AVT before irradiation obtained for each electrolyte.



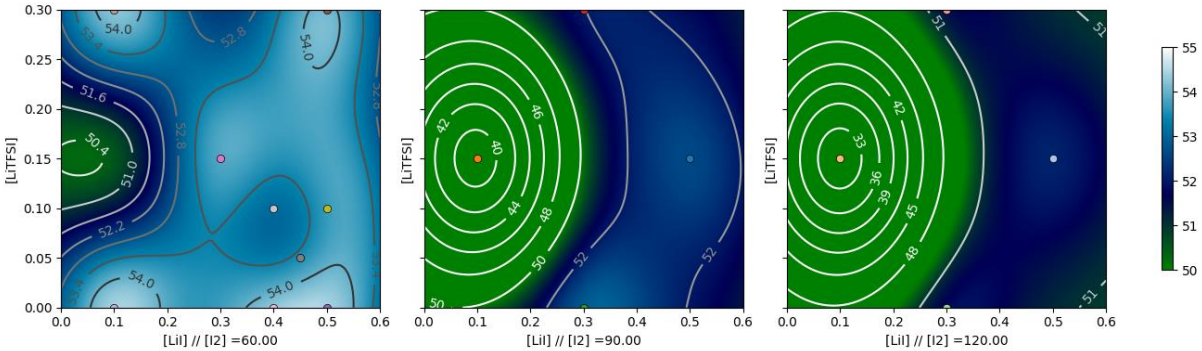
Supplementary Figure 14: Analysis of variance and factor evaluation of the second round of optimization of the Iodine-based electrolyte for **QH138**-dyed solar cells



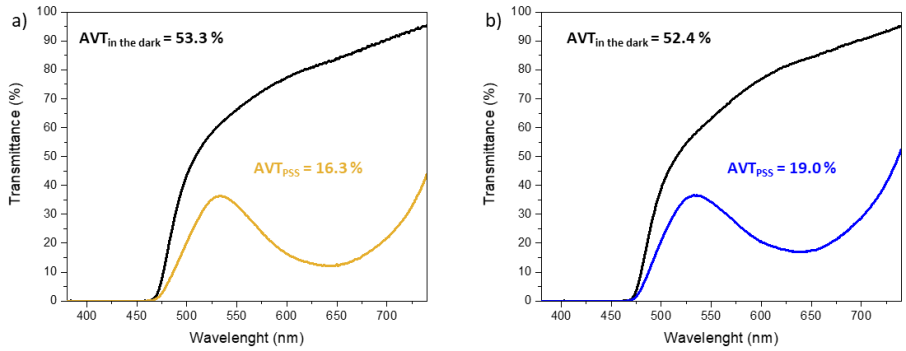
Supplementary Figure 15 : Verification curves for the regression obtained for the PCE (a) and AVT (b) analysis on the second set of experiments. . It does correspond that our learning curves for the fitting of gamma, epsilon and c hyperparameters are sufficient for a good correlation in our model



Supplementary Figure 16 : PCE Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the second batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the LiTFSI concentration. Each panel represents the effect of the two previous parameters at a defined iodine concentration (a: 60 mmol.L-1, b: 90 mmol.L-1, c: 120 mmol.L-1).



Supplementary figure 17 : AVT Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the second batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the LiTFSI concentration. Each panel represents the effect of the two previous parameters at a defined iodine concentration (a: 60 mmol.L-1, b: 90 mmol.L-1, c: 120 mmol.L-1).



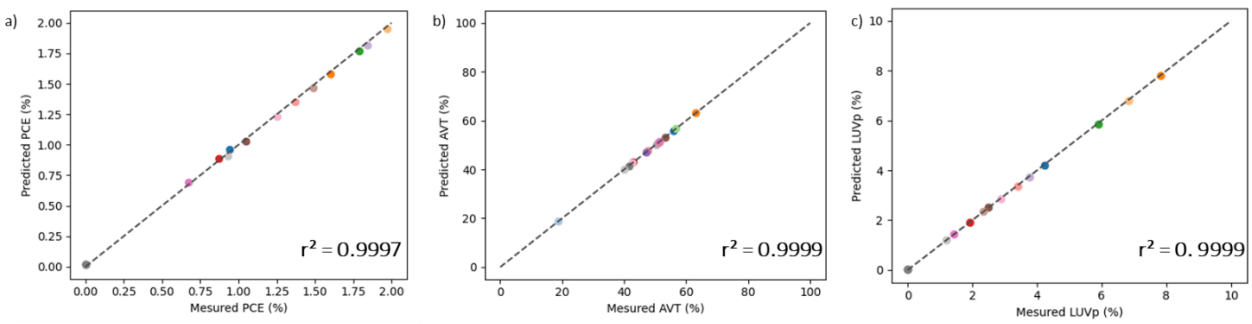
Supplementary figure 18: Transmittance spectra of QH138 cells with HM electrolyte (a) and EI_{opt} (b) in the dark (black) and at the PSS (coloured). (Dye/CDCA ratio: 1/10)

VII. Optimisation of the TEMPO based electrolyte

a. First run of experiments characterisation for the optimisation of the TEMPO based electrolyte

Experiment Nb	[TEMPO] (mol·L^{-1})	[TEMPO$^+$] (mol·L^{-1})	[LiI] (mol·L^{-1})	[LiTFSI] (mol·L^{-1})	PCE (%)	AVT_{closed} (%)	AVT_{open} (%)	C*	LUE_p
1-1	0.15	0.075	0.3	0	0.16	56	12.42	4.51	4.2
1-2	0.15	0.1	0	0.2	0.00	18	14.82	1.21	0
1-3	0.15	0.05	0.5	0.1	0.21	63	12.89	4.89	7.8
1-4	0.15	0.025	1	0.3	0.24	51	14.81	3.44	6.8
1-5	0.1	0.075	0.5	0.3	1.51	51	15.5	3.29	5.9
1-6	0.1	0.025	0	0	0.015	57	24	2.38	0
1-7	0.1	0.05	0.3	0.2	1.25	43	19.6	2.19	1.9
1-8	0.1	0.1	1	0.1	1.55	48	19.24	2.49	3.4
1-9	0.2	0.05	0	0.3	0.015	47	25.62	1.83	0
1-10	0.2	0.075	1	0.2	1.52	50	24.53	2.04	3.8
1-11	0.2	0.1	0.5	0	2.01	53	22.44	2.36	2.5
1-12	0.2	0.025	0.3	0.1	1.90	51	32.43	1.57	2.3
1-13	0.25	0.05	1	0	1.72	51	24.23	2.1	1.4
1-14	0.25	0.025	0.5	0.2	1.55	43	18.53	2.32	2.9
1-15	0.25	0.075	0	0.1	0.01	41	37.52	1.09	0
1-16	0.25	0.1	0.3	0.3	0.63	40	31.1	1.29	1.2

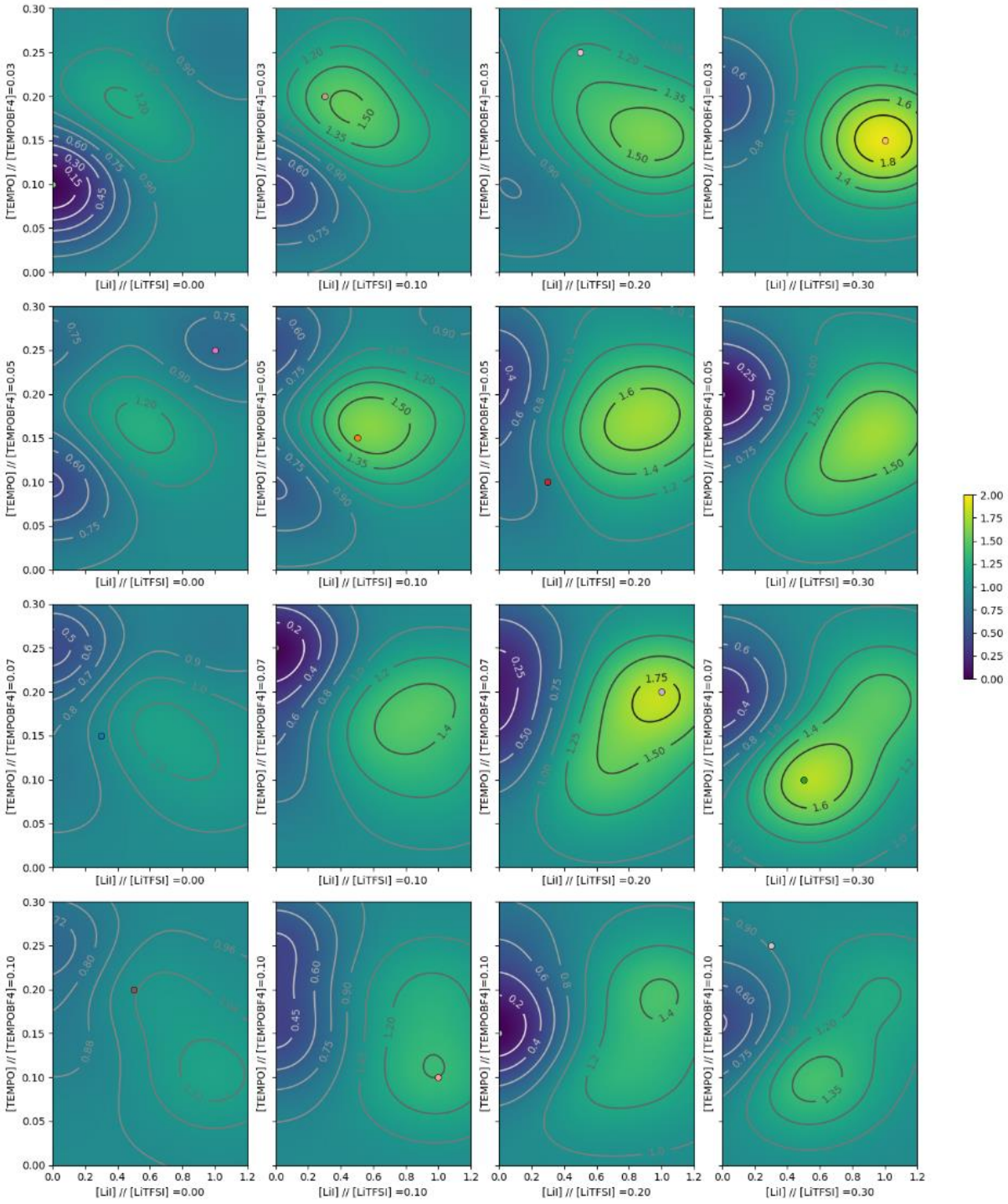
Table S 3 : Table listing the conditions used to make the electrolytes for the 16 experiments in the first optimisation round as well as the PCE at the 1st irradiation, AVT before and after irradiation and LUE_p obtained for each electrolyte.



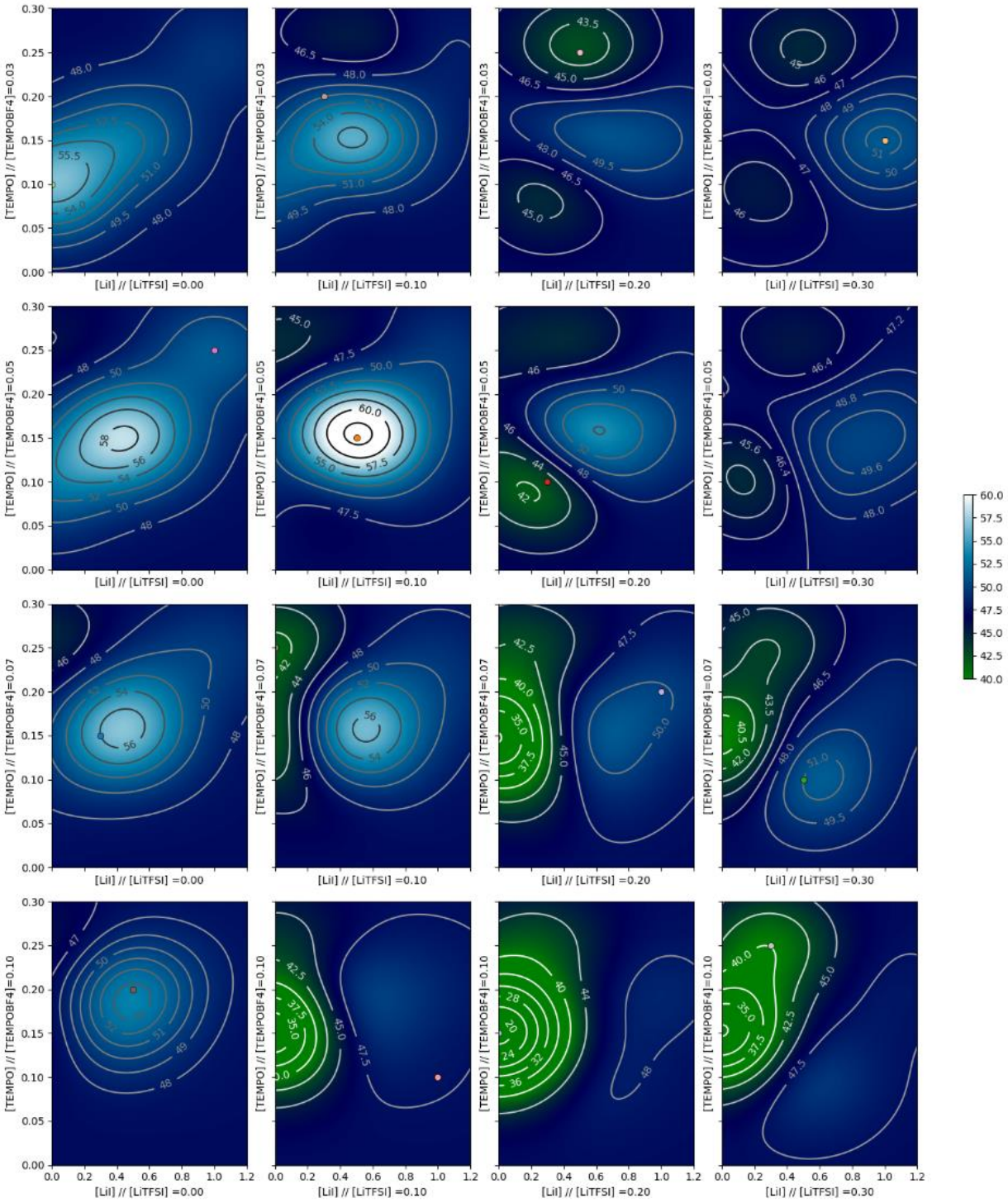
Supplementary figure 19: Verification curves for the regression obtained for the PCE (a), AVT (b) and LUE_p (c) analysis on the first set of experiments. . It does correspond that our learning curves for the fitting of gamma, epsilon and c hyperparameters are sufficient for a good correlation in our model

Experiment Nb	[TEMPO] ($mol \cdot L^{-1}$)	[$TEMPO^+$] ($mol \cdot L^{-1}$)	[LiI] ($mol \cdot L^{-1}$)	[LiTFSI] ($mol \cdot L^{-1}$)	PCE (%)	AVT_{closed} (%)	AVT_{open} (%)	C*	LUE_p
2-1	0.125	0.025	0.9	0.3	0.905	49	21	2.33	2.1
2-2	0.175	0.025	0.9	0.3	0.885	48	30	1.6	1.4
2-3	0.125	0.025	1.1	0.3	1.71	38	10	3.8	6.9
2-4	0.175	0.025	1.1	0.3	1.31	39	9	4.33	5.7
2-5	0.175	0.05	0.8	0.2	0.825	39	18	2.17	1.8
2-6	0.175	0.075	0.8	0.2	1.72	39	8	4.88	8.0
2-7	0.175	0.05	0.6	0.1	0.96	36	7	5.14	4.9

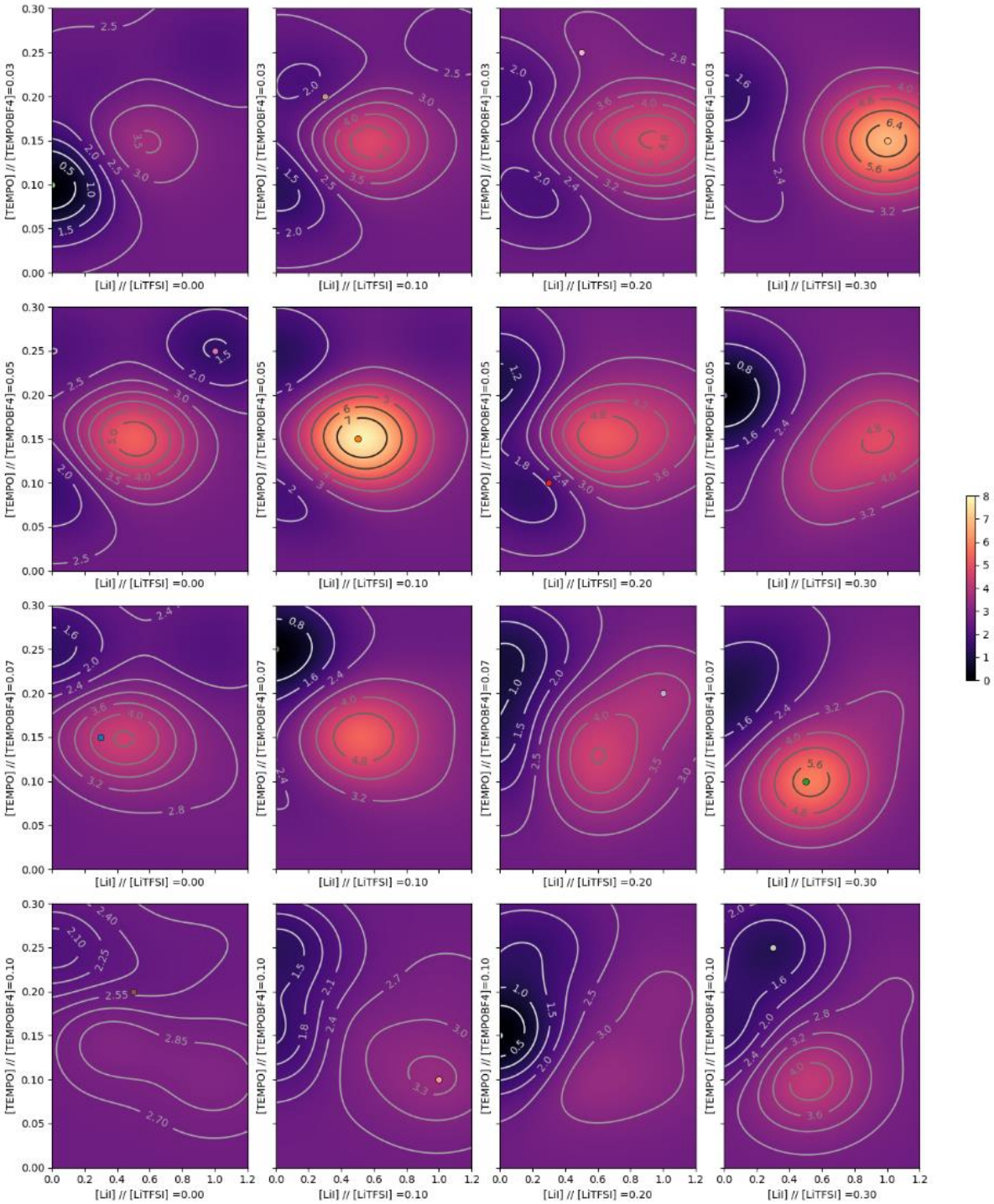
Table S 4 : Table listing the conditions used to make the electrolytes for the 7 experiments in the first optimisation round as well as the PCE at the 2nd irradiation, AVT before and after irradiation and LUE_p obtained for each electrolyte.



Supplementary figure 20: PCE Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the first batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the TEMPO concentration. Each panel represents the effect of the two previous parameters at a defined LiTFSI concentration (from left to right: [LiTFSI] = 0.0; 0.1; 0.2; 0.3 mol.L^{-1}) and TEMPO^{+} concentration (from top to bottom: [TEMPO+] = 0.025; 0.05; 0.075; 0.1 mol.L^{-1}).

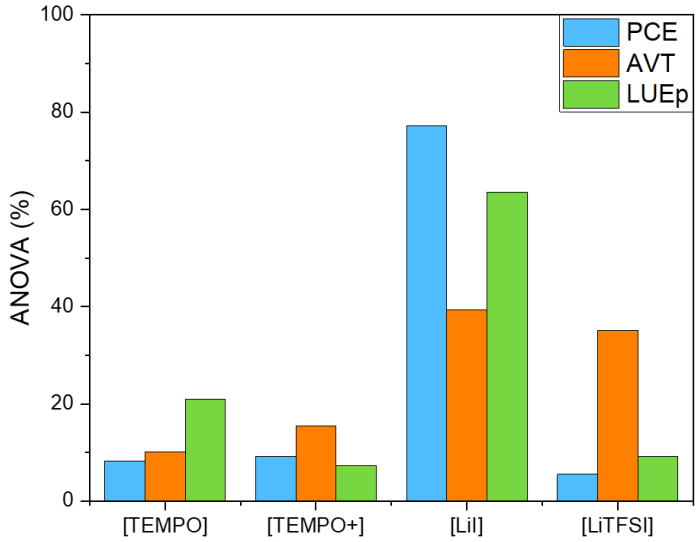


Supplementary figure 21: AVT Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the first batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the TEMPO concentration. Each panel represents the effect of the two previous parameters at a defined LiTFSI concentration (from left to right: [LiTFSI] = 0.0; 0.1; 0.2; 0.3 mol.L^{-1}) and TEMPO^{+} concentration (from top to bottom: [TEMPO+] = 0.025; 0.05; 0.075; 0.1 mol·L^{-1}).

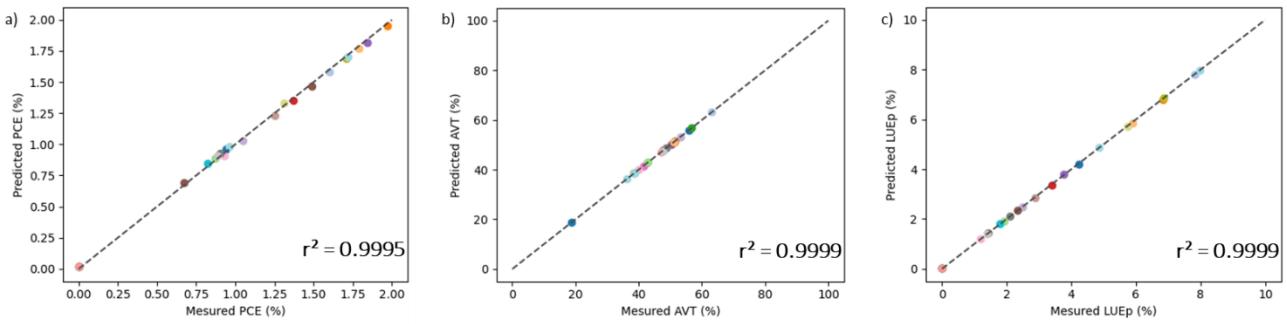


Supplementary figure 22: AVT Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the first batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the TEMPO concentration. Each panel represents the effect of the two previous parameters at a defined LiTFSI concentration (from left to right: [LiTFSI] = 0.0; 0.1; 0.2; 0.3 mol.L^{-1}) and TEMPO^{+} concentration (from top to bottom: [TEMPO+] = 0.025; 0.05; 0.075; 0.1 mol·L^{-1}).

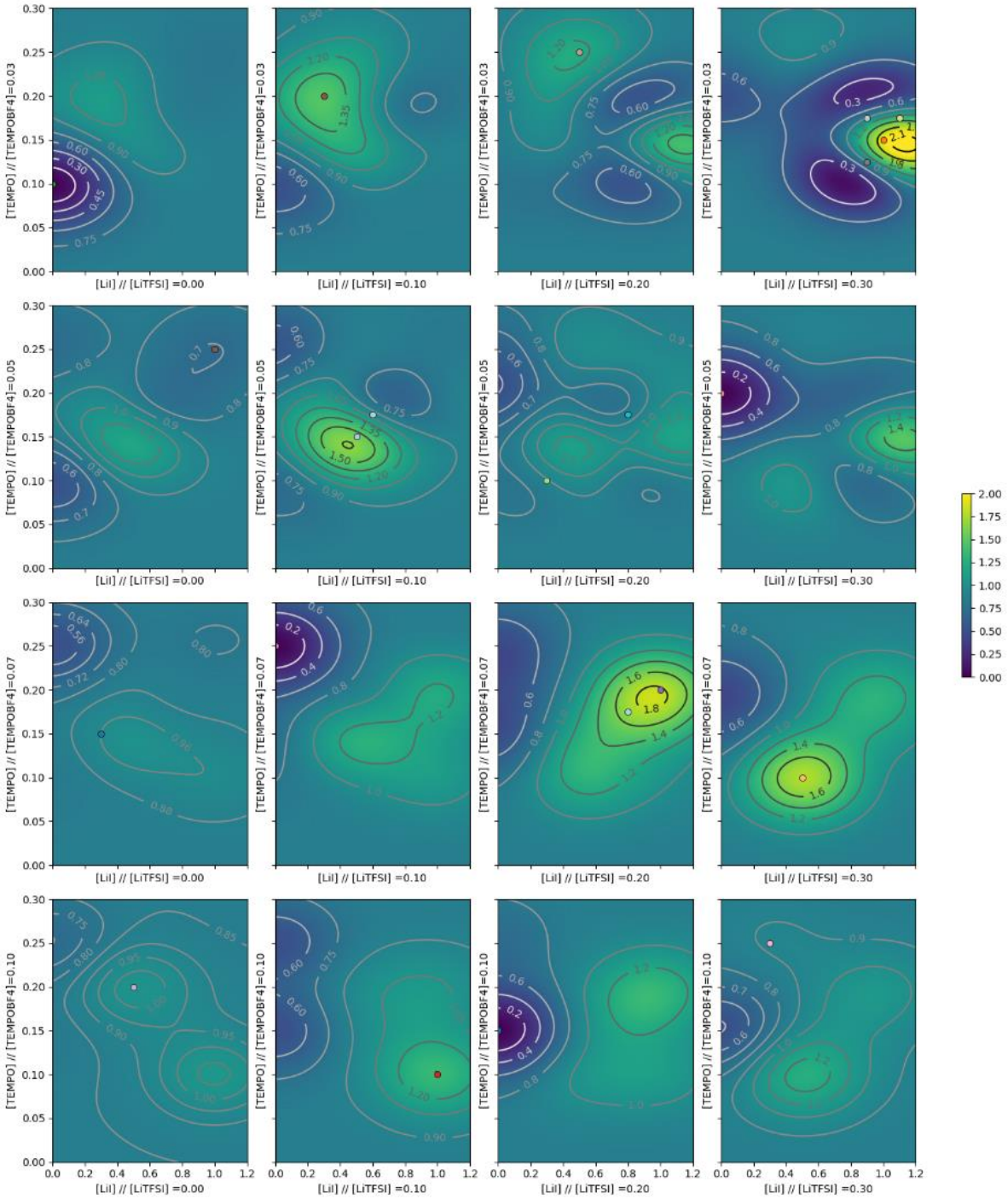
b. Second run of experiments characterisation for the optimisation of the TEMPO based electrolyte



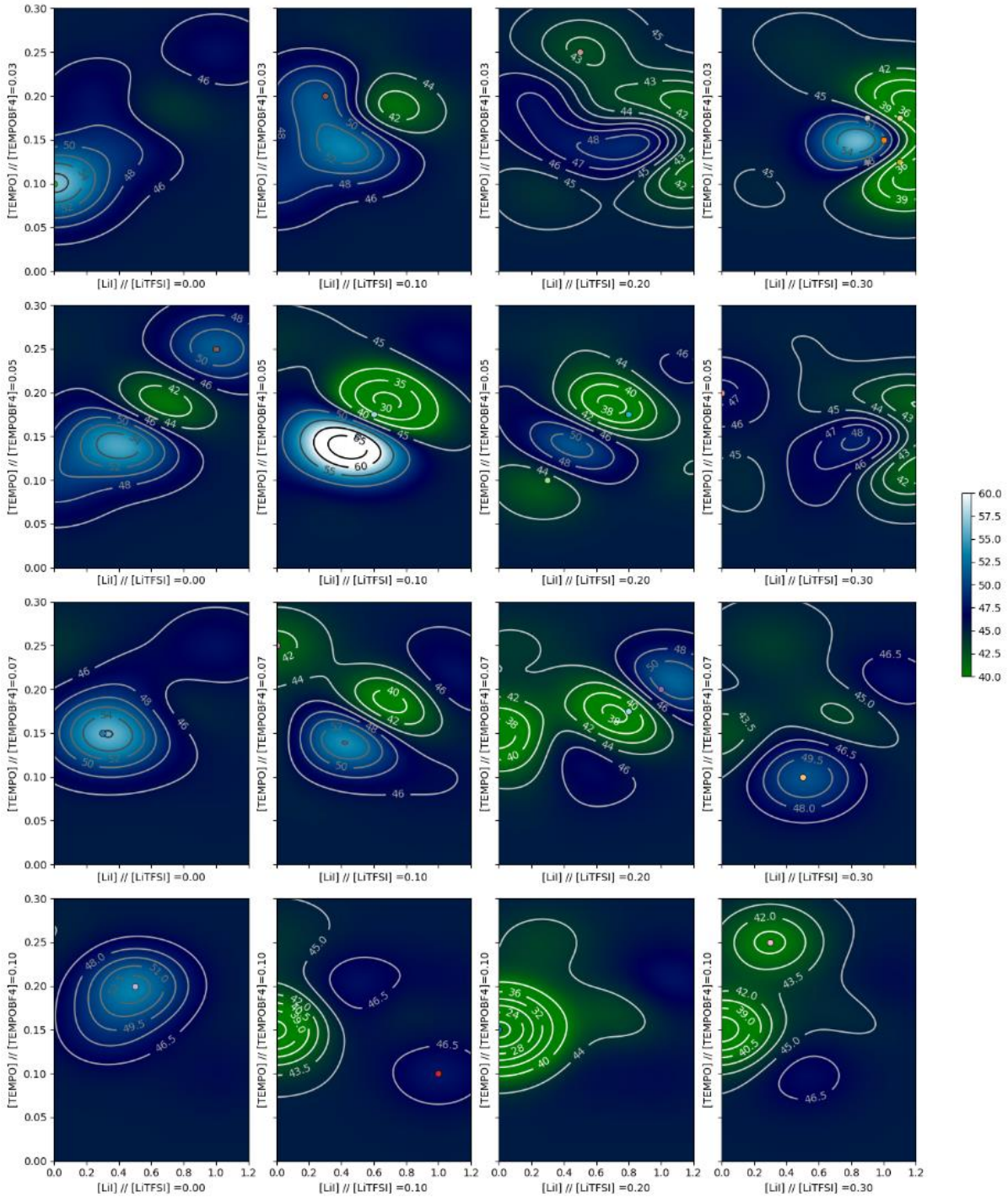
Supplementary figure 23: Analysis of variance and factor evaluation of the second round of optimization of the TEMPO-based electrolyte for **QH138**-dyed solar cells



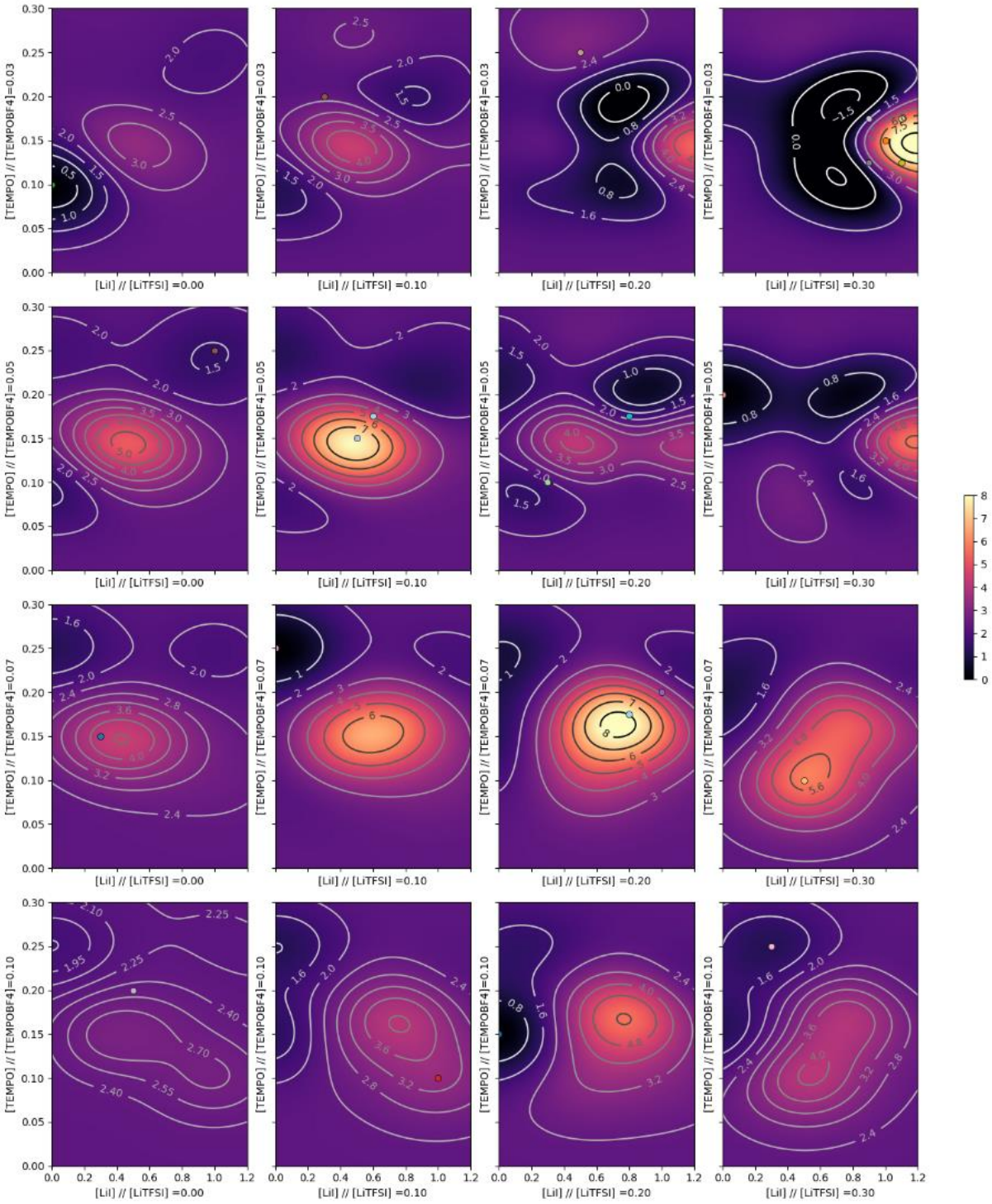
Supplementary figure 24 : Verification curves for the regression obtained for the PCE (a), AVT (b) and LUE_p (c) analysis on the second set of experiments. . It does correspond that our learning curves for the fitting of gamma, epsilon and c hyperparameters are sufficient for a good correlation in our model



Supplementary figure 25: PCE Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the second batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the TEMPO concentration. Each panel represents the effect of the two previous parameters at a defined LiTFSI concentration (from left to right: [LiTFSI] = 0.0; 0.1; 0.2; 0.3 mol.L^{-1}) and TEMPO^{+} concentration (from top to bottom: [TEMPO+] = 0.025; 0.05; 0.075; 0.1 mol·L^{-1}).



Supplementary figure 26: AVT Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the second batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the TEMPO concentration. Each panel represents the effect of the two previous parameters at a defined LiTFSI concentration (from left to right: [LiTFSI] = 0.0; 0.1; 0.2; 0.3 mol·L^{-1}) and TEMPO^{+} concentration (from top to bottom: [TEMPO+] = 0.025; 0.05; 0.075; 0.1 mol·L^{-1}).

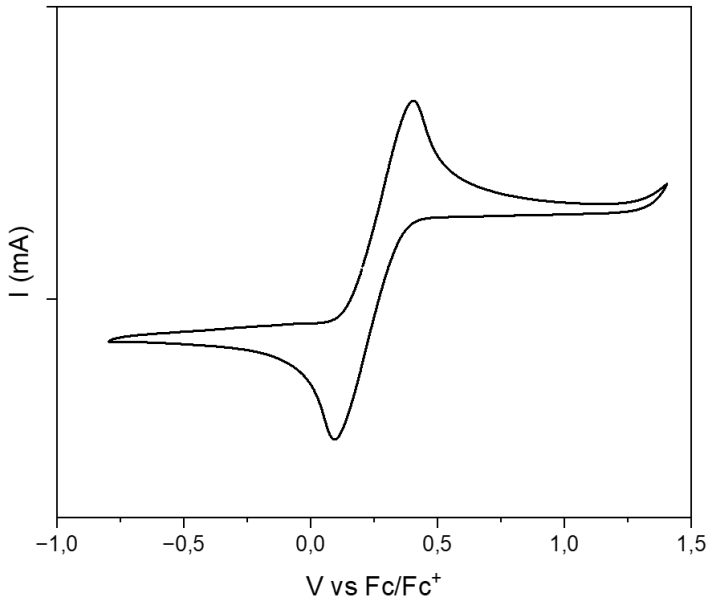


Supplementary figure 27: LUE_p Surfaces obtained by regression (SVR method; RBF kernel) on the parameters used in the second batch of experiments. The abscissa represents the LiI concentration and the ordinate represents the TEMPO concentration. Each panel represents the effect of the two previous parameters at a defined LiTFSI concentration (from left to right: [LiTFSI] = 0.0; 0.1; 0.2; 0.3 $mol.L^{-1}$) and $TEMPO^+$ concentration (from top to bottom: [TEMPO+] = 0.025; 0.05; 0.075; 0.1 $mol \cdot L^{-1}$).

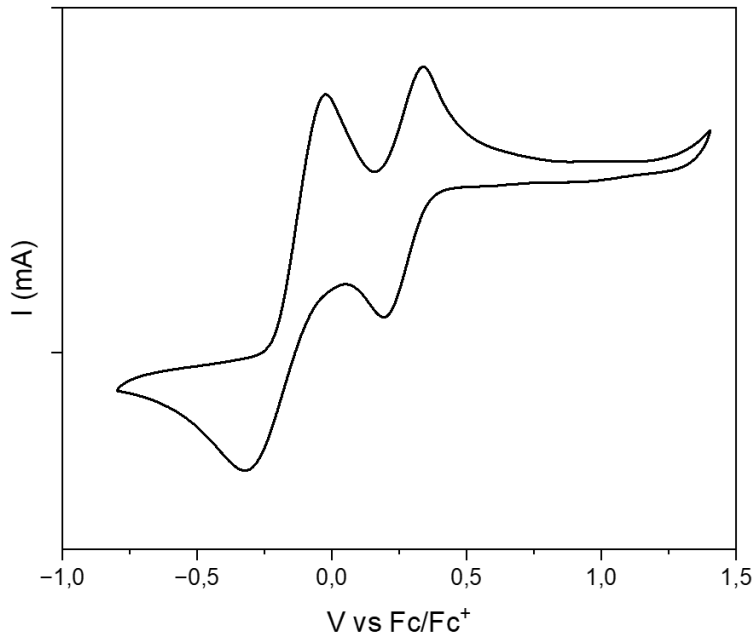
Electrolyte	J_{SC} (mA·cm^{-2})	V_{OC} (mV)	FF (%)	PCE (%)	AVT_{closed} (%)	AVT_{open} (%)	C^*	LUE_p
El_{opt}	4.36 (4.075 ± 0.40)	580 (582 ± 2)	85.0 (85.5 ± 0.5)	2.16 (2.03 ± 0.18)	55	13	4.22	**9.12**

Table S 5 : Table listing the optimal conditions obtained from final round of the TEMPO electrolyte optimization

c. Analysis of the newly optimized TEMPO electrolyte



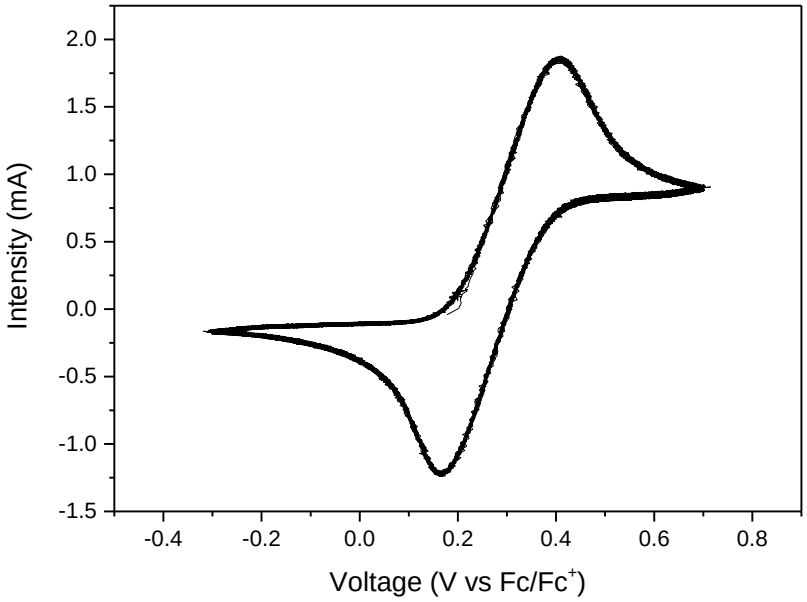
Supplementary figure 28: Cyclic voltammetry trace of Iodine (-1 V to 1.5 V, ACN, $tBAPF_6$, 25 °C, 100 mV·s^{-1})



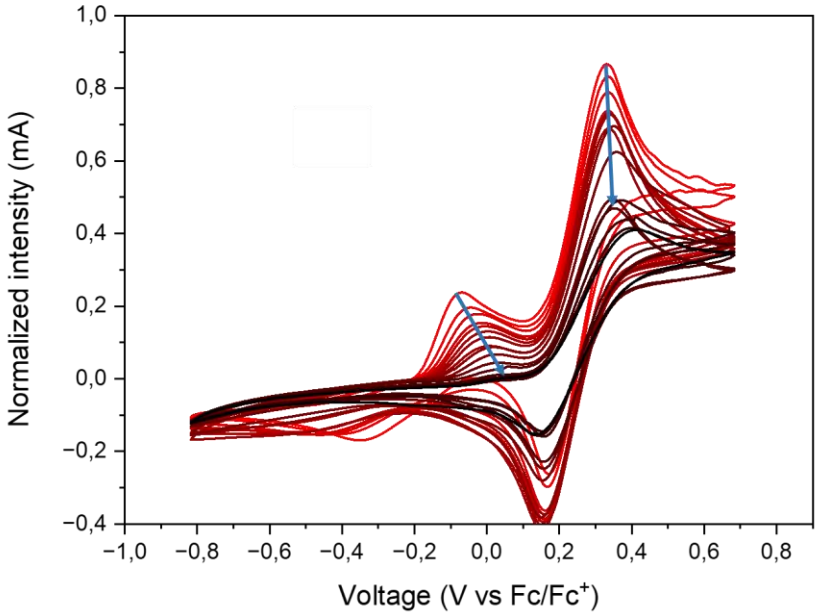
Supplementary figure 29: Cyclic voltammetry trace of TEMPO (-1 V to 1.5 V, ACN, $tBAPF_6$, 25 °C, 100 $mV \cdot s^{-1}$)

d. Stability of devices using TEMPO electrolyte

As described in the main text, the photochromic dye-sensitized solar cells using TEMPO-based electrolytes PCE declines under irradiation after a rapid rise, until stabilizing at PSS. To cast light into the degradation under light, cyclic voltammetry measurements were conducted in both the TEMPO/TEMPO+ couple and a representative TEMPO full electrolyte used in this optimisation, finding that although the redox couple itself remains stable after a large number of iterations, the electrolyte suffers from degradation, probably due to interaction between the TEMPO species and LiI (see Supplementary figures 30 and 31).



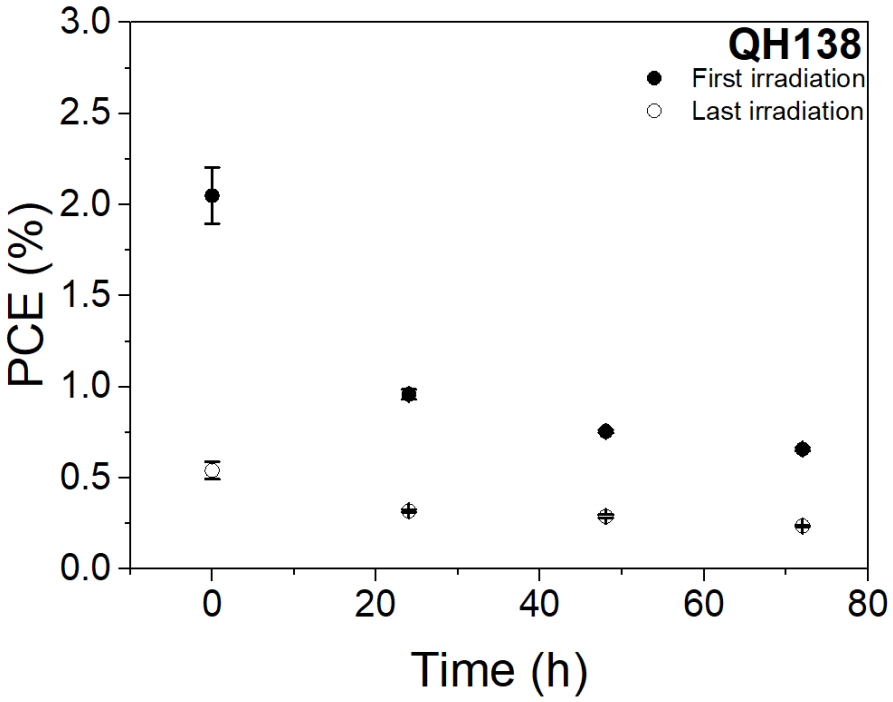
Supplementary figure 30: TEMPO/TEMPO$^+$ couple stability test by cyclic voltammetry over 300 cycles. 15 cycles are shown here, selected every 20 cycles (ACN, tBAPF$_6$, 25°C, 100mV·s^{-1}).



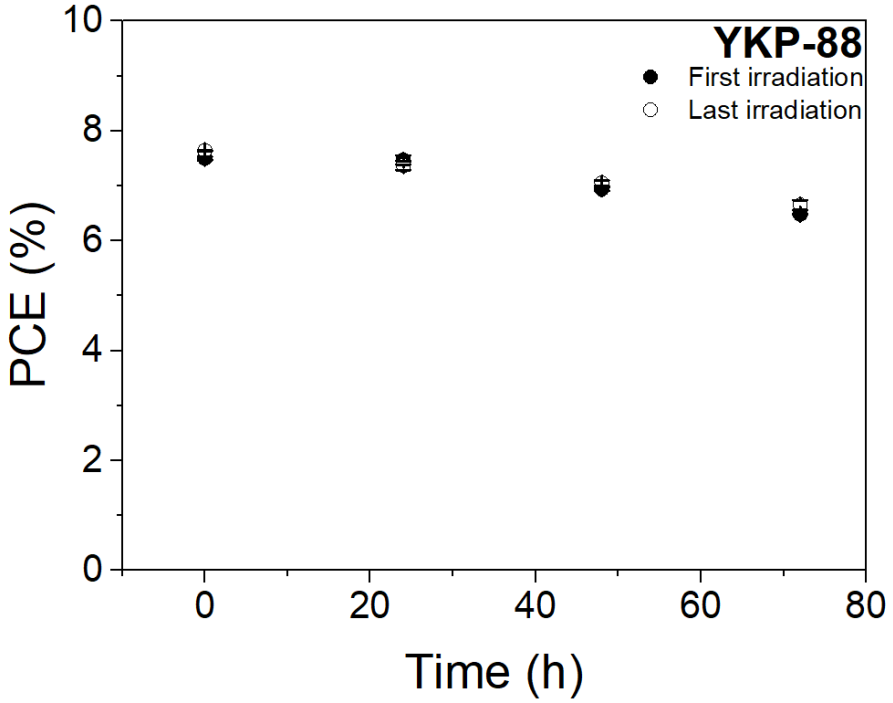
Supplementary figure 31: TEMPO electrolyte stability test by cyclic voltammetry over 300 cycles. 15 cycles are shown here, selected every 20 cycles (ACN, tBAPF$_6$, 25°C, 100mV·s^{-1}).

Furthermore, the observed PCE loss is partially recovered after keeping the devices under dark conditions during the time required for the photochromic dye to return to their uncoloured state. Nonetheless, even after this recovery the measured PCE shows a decreasing trend with time (see Supplementary figure 34). However, using a non-photochromic dye (YKP-88), produces devices that retain up to a 90% of the original efficiency during the same period of time (see Supplementary figure

34), evidencing that the decrease in efficiency observed in photochromic devices responds to interactions between the photochromic unit and the TEMPO species, which is currently under investigation but beyond the scope of this work.



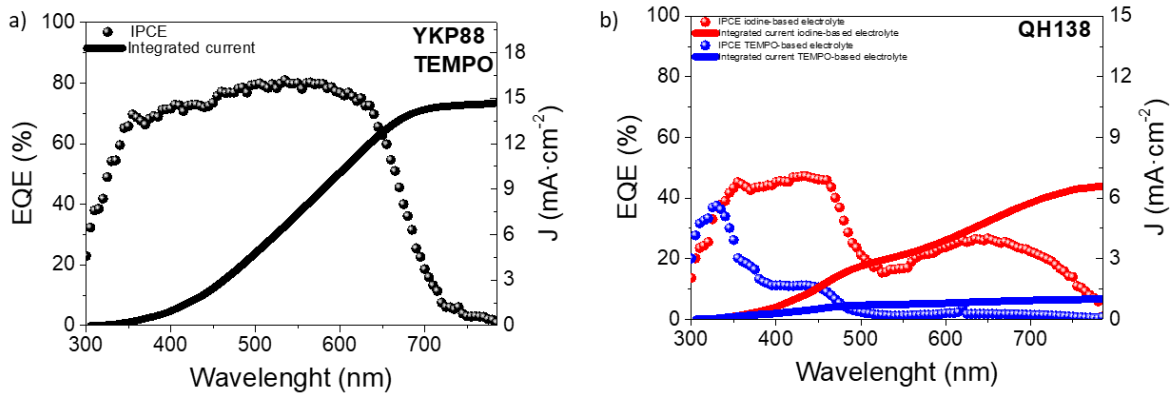
Supplementary figure 32: Efficiency before and after light soaking measured every 24 hours for 3 days for photochromic dye-sensitized solar cells.



Supplementary figure 33: Efficiency before and after light soaking measured every 24 hours for 3 days for non-photochromic dye-sensitized solar cells.

e. Optoelectronic characterization of optimized devices

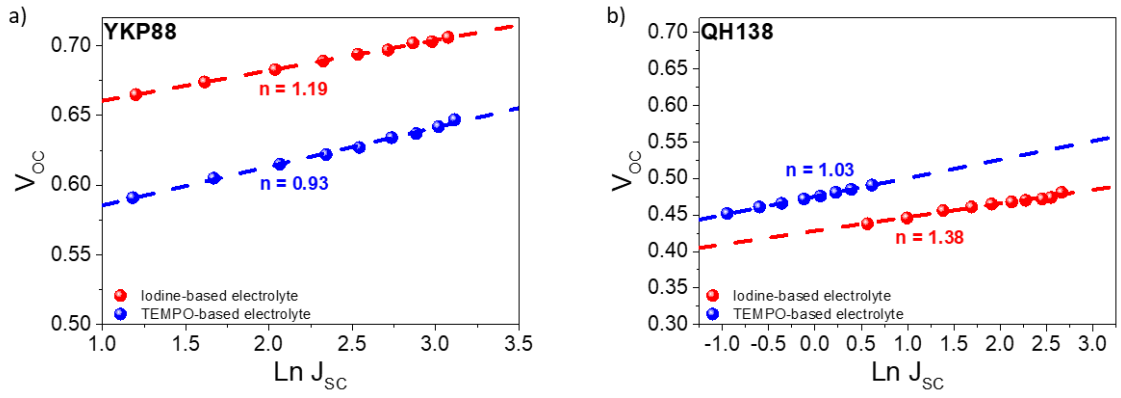
IPCE was measured using a Xenon lamp with a Newport monochromator, coupled to an Arkeo measurement system provided by Cicci Research S.L. The measurements were conducted using a white illumination bias and a chopper at 11 Hz, with a 5 nm resolution.



Supplementary figure 34: Incident photon-to-electron conversion efficiency and integrated current of transparent devices using non-photochromic YKP88 and photochromic QH138 dyes with the optimized iodine and TEMPO based electrolytes.

Light intensity dependent photovoltaic parameters were measured using a solar simulator LED based provided by Ossila, from 1000 $W{\cdot}m^{-2}$ to 100 $W{\cdot}m^{-2}$. The resulting open-circuit voltages were plotted versus the natural logarithm of the short-circuit current to extract the ideality factor using the following expression[3].

$$V_{OC} = \frac{nk_BT}{q}\ln(J_{SC})$$



Supplementary figure 35: Open-circuit voltage versus the natural logarithm of short-circuit current measured from 100 $W{\cdot}m^{-2}$ to 1000 $W{\cdot}m^{2}$ and the extracted ideality factor.

f. Calculation of average costs of optimised electrolytes

As the electrolyte were composed of commercial elements (Sigma Aldrich), an estimation of the average cost of the iodine and TEMPO based electrolytes could be done. In the case of the TEMPO/TEMPO$^+$ electrolyte, TEMPO$^+$ was synthetised, and the average cost can be estimated as 6.94 €/g[4].

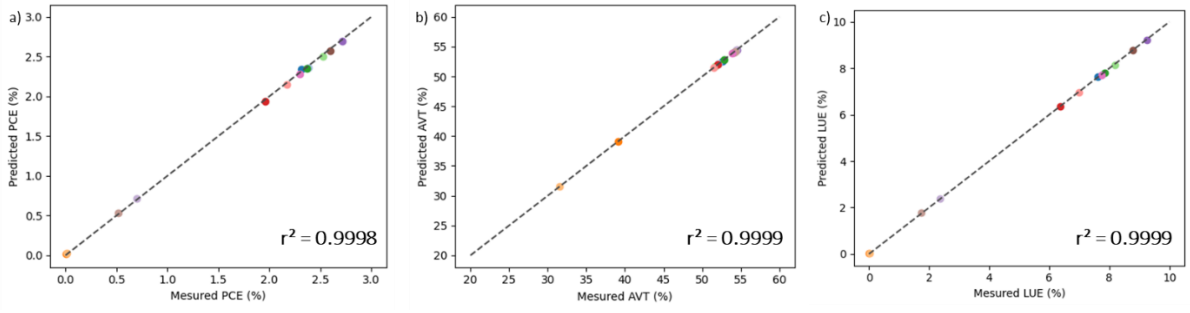
Species	Commercial price	Mg per ml of solution	€ per ml of solution
I_2	1€/g	15.23	0.015 €
LiI	9.23€/g	60.23	0.56 €
LiTFSI	8.24€/g	5.74	0.047 €
Total			**0.622 €**

Species	Commercial price	Mg per ml of solution	€ per ml of solution
TEMPO	5.52€/g	23.44	0.13 €
TEMPO$^+$	6.94 €/g	7.29	0.05 €
LiI	9.23€/g	153.93	1.42 €
LiTFSI	8.24€/g	86.13	0.71 €
Total			**2.31 €**

g. Additional information on data driven assisted modeling of the Electrolytes

The Python program described in our work is an adaptation and update of the work published by the Buriak group in ACS Nano in 2018 (Ref. 59). In our work, we maintained the same approach and experimental plan, incorporating our specific features for optimisation, as described in the supporting information of the original article.

Initially, the data is imported into a pandas DataFrame with labeled data. The entire dataset is utilized as the training set for a Support Vector Machine (SVM) regression machine learning algorithm, employing a radial basis function kernel from the Scikit-learn package. We explored various scoring methods, including train/test split, cross-validation, and Leave-One-Out Cross-Validation, but deliberately chose to retain the whole dataset for training. The model's score, measured against experimental values, provides insight into the trend of the predicted conditions and generate a calibration curve as follows:



While we acknowledge the importance of the model's score for evaluating the effectiveness of machine learning, we used it primarily as a guide for optimizing our features. To this end, we implemented a function to optimize the hyper parameters 'gamma', 'epsilon', and 'C'. Additionally, we created a function to detect the number of features to optimize (2, 3, or 4).

A value map is then generated by creating a matrix of feature values fed into the model. Each pixel in the map corresponds to a prediction, with colours indicating the value of the target features (PCE, AVT, LUE). These values are visualized using the matplotlib module in Python, and the fitted model is used to make PCE predictions for these values. The resulting map from the code would be comparable to Figure 3 in the main text.

1) Choice of the design of experiment

The number of parameters to be optimized n and the number of values m to be assigned to them influence the number of experiments to be carried out. The total number of experiments needed to carry out a complete factorial design, i.e. testing each combination of parameters, requires mn experiments. This method is very time-consuming and the number of experiments increases very quickly with the precision required. To reduce the number of these experiments, two types of DoE can be used: the fractional factorial and the Latin square.[5-6]
The first design requires m(n-k) experiments, with k a factor of the user's choice. Usually, k = 1, which divides the total number of experiments by the number of values assigned to the different factors. This fractional design is useful for studying a small number of factors but with many values to assign to them. This latest design greatly reduces the number of experiments to be carried out, which becomes only m×n experiments. This design consists of running orthogonal experiments where two different parameters are never found in two different experiments with identical values. This design allows rapid screening of a large number of parameters. However, the extrapolation and use of these data can be more complicated because the data obtained is more fragmentary.
Therefore, depending on the case, it will be necessary to choose between the different possible experimental designs in order to obtain solid results for analysis.

2) Working principle of SVM

A Support Vector Machine (SVM) model can be used for two purposes: classification (SVC) and regression (SVR). Both models are available in the python package Scikit Learn used in this study.[7] The first is used to spatially discriminate points on a map according to different factors. This method is also known as Support Vector Classification (SVC). The second is the one we are interested in here, and is used to derive a function for linking points by regression. In this case, the method is called Support Vector Regression (SVR). Its principle is based on the determination of a function allowing points to be linked within a certain limit.[8] n initial regression Y = f(x) is carried out and then the limits are established as a function of a parameter ε. These limits are established by f(x) + ε and f(x) - ε. All subsequent regressions will only take into account points within this limit and the new function Y = f2(x) must satisfy: -ε < Y - f(x) < +ε. This operation is repeated until the regression passes through a maximum number of points.
This type of model can use several kernels depending on the regression required. The kernel can be linear, polynomial, exponential, etc. In our case, the kernel chosen is a Radial Basis Function (RBF) kernel (Eq.1) . This type of function is a sum of non-linear functions such as Gaussians. This model has been widely used for Machine Learning and regression on non-linear data.[9]

Eq. 1 $$K(X\_1,X\_2)=exp(-\gamma\|X\_1-X\_2 \|^2)$$

where, 'γ' is the hyperparameter and $||X_1 - X_2||$ is the Euclidean Distance between two points X_1 and X_2.

References

1 Q. Huaulmé, V. M. Mwalukuku, D. Joly, J. Liotier, Y. Kervella, P. Maldivi, S. Narbey, F. Oswald, A. J. Riquelme, J. A. Anta and R. Demadrille, *Nat Energy*, 2020, **5**, 468–477.

2 J. Liotier, V. M. Mwalukuku, S. Fauvel, A. J. Riquelme, J. A. Anta, P. Maldivi and R. Demadrille, *Solar RRL*, 2022, **6**, 2100929.

3 J. P. Gonzalez-Vazquez, G. Oskam and J. A. Anta, *The Journal of Physical Chemistry C*, DOI:10.1021/jp306517g.

4 H. Richter and O. García Mancheño, *European Journal of Organic Chemistry*, 2010, **2010**, 4460–4467.

5 M. E. Piper, T. R. Schlam, D. Fraser, M. Oguss and J. W. Cook, in Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: Advanced Topics, eds. L. M. Collins and K. C. Kugler, Springer International Publishing, Cham, 2018, pp. 23–45.

6 K. Sirikasemsuk and K. Thachongthumla, in Proceedings of the 2019 7th International Conference on Information and Education Technology, Association for Computing Machinery, New York, NY, USA, 2019, pp. 196–199.

7 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, Journal of Machine Learning Research, 2011, 12, 2825–2830.

8 A. Sethi, Support Vector Regression Tutorial for Machine Learning, https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/, (accessed December 13, 2024).

9 M. J. D. Powell, Mathematical Programming, 1977, 12, 241–254.