Supplementary Information (SI) for Physical Chemistry Chemical Physics.

Identifying Contributions Responsible for the Nonlinear Optical Property Differences in Functionalized Hexaphyrins with Explainable Machine-Learning Methods.

Eline Desmedt, Michiel Jacobs, Mercedes Alonso,* and Freija De Vleeschouwer*

Department of General Chemistry: Algemene Chemie (ALGC), Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium

E-mail: Mercedes.Alonso.Giner@vub.be; Freija.De.Vleeschouwer@vub.be

Contents

Depolarization ratio

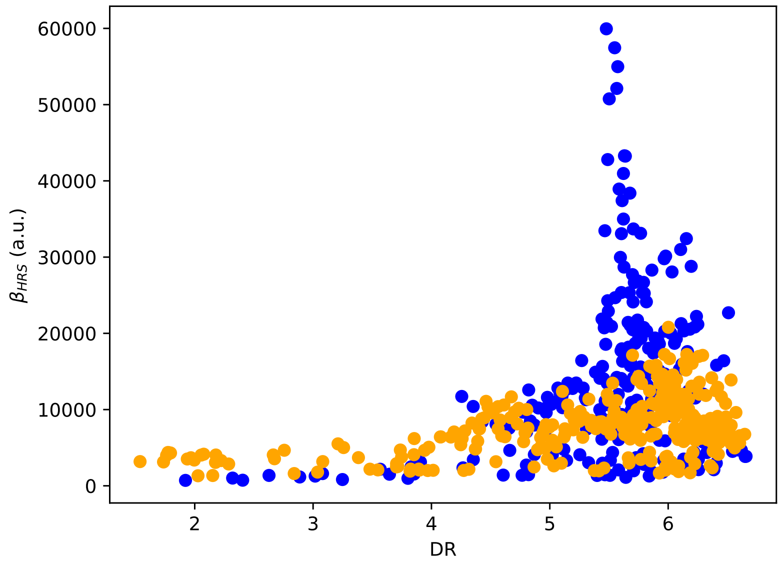

Figure S1: Scatter plot of the β_{HRS} response in a.u. versus the depolarization ratio DR. Data points are colored in blue and orange for [26]- and [30]-hexaphyrins, respectively.

Training and Test set

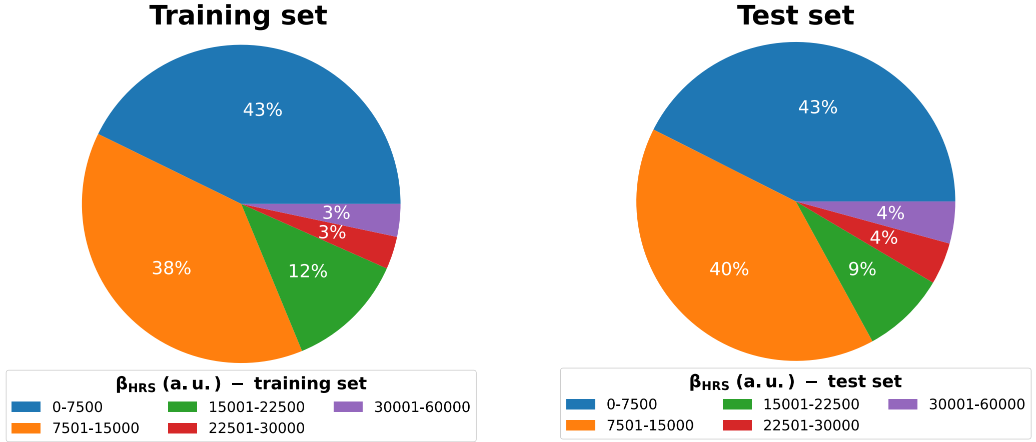

Figure S2: Pie diagram of the β_{HRS} response of the training (left) and test (right) set divided into groups ranging 7500 a.u.

Hyperparameters

Table S1: Tuned hyperparameters of all generated models.

Model	Kernel	α	degree	γ	coef0
Model 1 -	polynomial	0.21596172262555952	3	0.39811283148884463	0.6620232579248944
Model 9 features	polynomial	0.3534700903458985	4	0.10972086480126685	1.8307728246239008
Model 8 features	polynomial	0.21237948689683517	4	0.10168244123472461	1.8873825033896925
Model 7 features	polynomial	0.042812608257912776	4	0.11284573203779522	1.3081768404916778
Model 2	polynomial	0.23397666361424996	4	0.21964487154491796	1.726214658320214
Model 5 features	polynomial	0.2732328568080413	4	0.2840189622843563	1.8402448087866812

Feature importance analysis

As we suspected in the main manuscript that Model 1 is overfit, we applied a strategy based on feature importances and MAE changes to carefully select and omit features from our next model. At closer inspection of Table 4 in the main manuscript, which contains the feature importances of Model 1, we observe that $|\Delta\mu|$ has the lowest feature importance of all ten features. Given its low importance, $|\Delta\mu|$ is removed from the feature set and a new model is trained with only 9 features. Even though the MAE values of the full training and test sets increase slightly by around 30 a.u., the MAE of the cross-validation set significantly drops, as displayed in Figure S3 by the bar plot on the one hand (left-hand axis) and the line plot showing the difference in MAE between the current model and its predecessor on the other hand (right-hand axis). Thus, it is decided to omit $|\Delta\mu|$ from the input features. For the next model, we remove BOA as this descriptor shows the lowest feature importance (Table S3). Based on the drop in MAE for the test set, the minimal MAE increase for the training set and the cross-validation set performing not much worse than Model 1, BOA removal is confirmed. We continue the same procedure by evaluating changes in model statistics for training, cross-validation and test sets. After consecutively removing 4 features with respect to Model 1, we finally see a big jump in MAE for the model with 5 features having eliminated Δ_{L}. This increase is quite significant for all sets (training, cross-validation and test). We thus opted to continue with the model with 6 features, hereafter named as Model 2 (cf. main manuscript) and containing the following input features: Δ_{HL}, q_{CT}, Δ_{L}, μ_{01}, ΔE_{S1}, and D_{CT}.

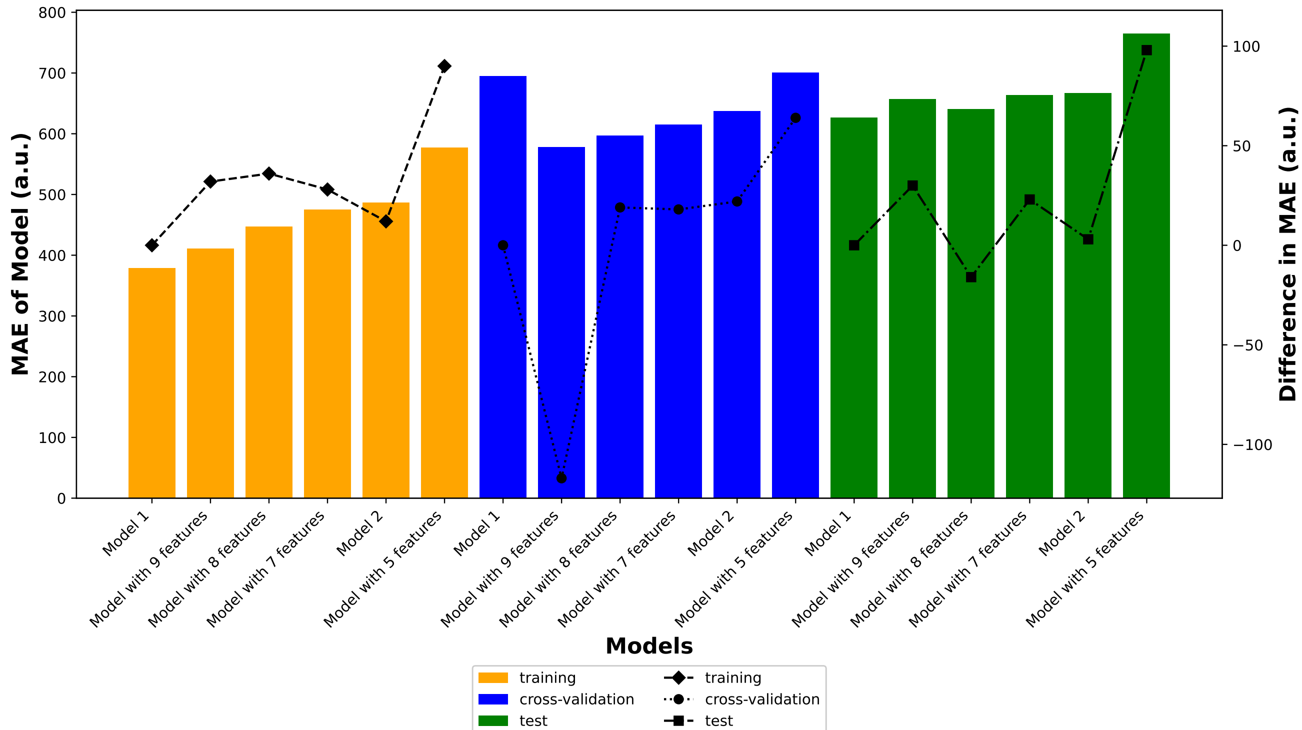

Figure S3: Model statistics after feature importance analysis for training set, cross-validation set and test set. As referred to the left-hand axis, the MAE of every consecutive model is displayed as a bar plot for all three data sets. As referred to the right-hand axis, the MAE change between the current and previous models (differing in 1 input feature) is plotted as a line plot. Note the difference in scale and level 0 for both axes.

Table S2: MAE of training, cross-validation and test set of the additional models.

Models	MAE (Train)	MAE (cross-validation)	MAE (Test) in a.u.
Model 1	379	695	627
Model 9 features	411	578	657
Model 8 features	447	597	641
Model 7 features	475	615	664
Model 2	487	637	667
Model 5 features	577	701	765

Table S3: Feature permutation importance of additional models with the differences in statistics between the actual dataset together with the permuted datasets and their standard deviations in a.u.

Model 9 features	ΔMAE (Train)	ΔMAE (Test)
Δ_{HL}	1843 ± 77	1731 ± 134
D_{CT}	4151 ± 144	3901 ± 236
$\Delta_{\mathrm{L+1\_H}}$	1240 ± 48	911 ± 87
Δ_{L}	1805 ± 80	1721 ± 149
μ_{01}	1560 ± 64	1375 ± 100
q_{CT}	2009 ± 85	1642 ± 177
ΔE_{S1}	1348 ± 52	1207 ± 92
BOA	765 ± 51	570 ± 81
μ	863 ± 42	653 ± 68
Model 8 features	**ΔMAE (Train)**	**ΔMAE (Test)**
Δ_{HL}	1996 ± 139	2074 ± 87
D_{CT}	3958 ± 139	3770 ± 223
$\Delta_{\mathrm{L+1\_H}}$	1363 ± 52	1117 ± 90
Δ_{L}	1926 ± 88	1886 ± 151
μ_{01}	1873 ± 75	1683 ± 127
q_{CT}	1940 ± 77	1649 ± 154
ΔE_{S1}	1312 ± 54	1187 ± 106
μ	940 ± 46	780 ± 74
Model 7 features	**ΔMAE (Train)**	**ΔMAE (Test)**
Δ_{HL}	1825 ± 84	1647 ± 141
D_{CT}	4338 ± 183	4121 ± 320
$\Delta_{\mathrm{L+1\_H}}$	1284 ± 44	1172 ± 77
Δ_{L}	1753 ± 85	1750 ± 157
μ_{01}	2804 ± 130	2719 ± 227
q_{CT}	2134 ± 110	1661 ± 202
ΔE_{S1}	1951 ± 83	1846 ± 169
Model 2	**ΔMAE (Train)**	**ΔMAE (Test)**
Δ_{HL}	4735 ± 186	4347 ± 295
D_{CT}	4730 ± 214	4498 ± 360
Δ_{L}	1247 ± 72	1043 ± 106
μ_{01}	2982 ± 154	2898 ± 267
q_{CT}	2152 ± 133	1776 ± 243
ΔE_{S1}	2028 ± 90	1960 ± 195

SHAP analysis: Test set

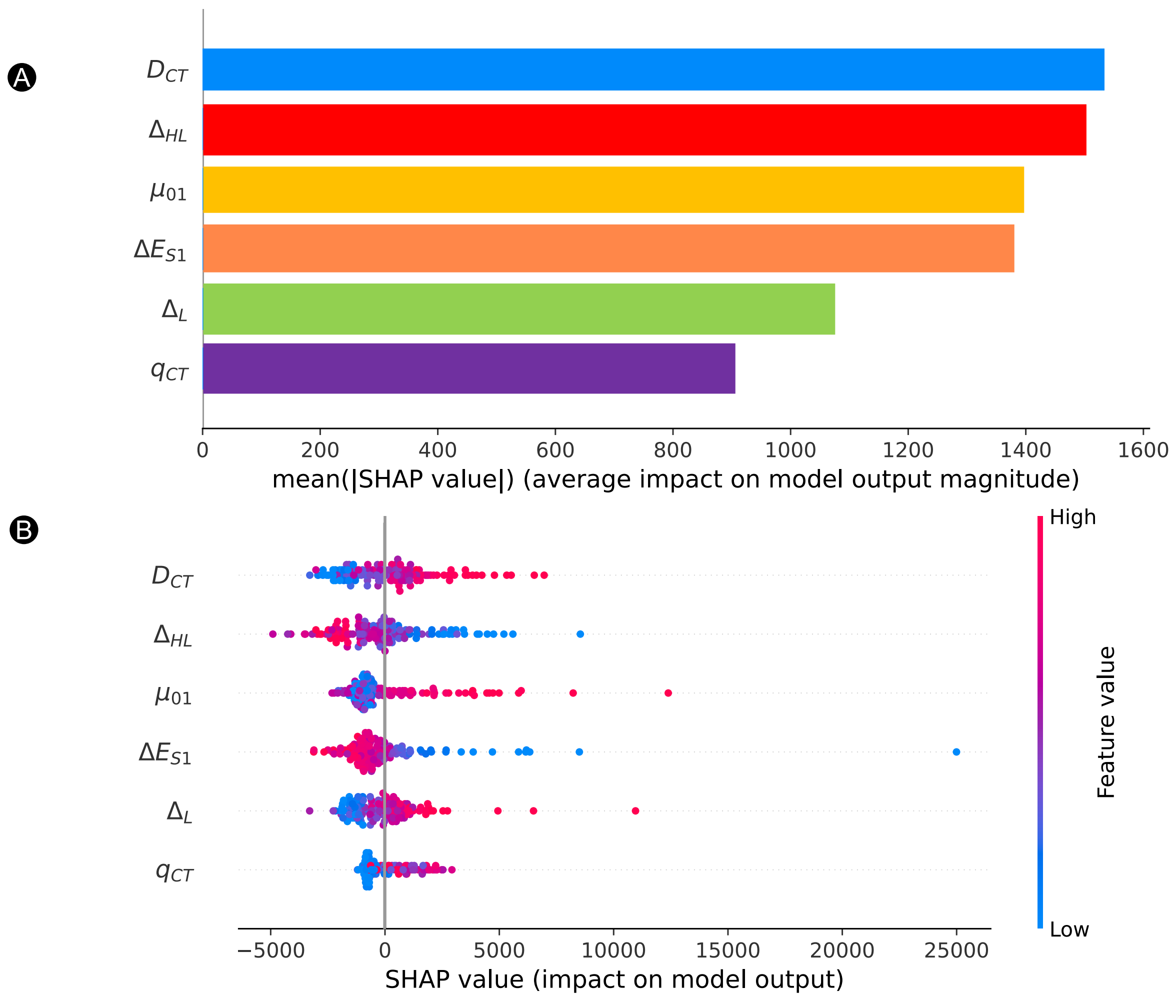

Figure S4: SHAP analysis of Model 2. (A) Bar plot containing the mean absolute SHAP value for each feature over all test set samples. (B) Beeswarm plot with SHAP values of all test set datapoints while highlighting the feature value.

Dependency plots

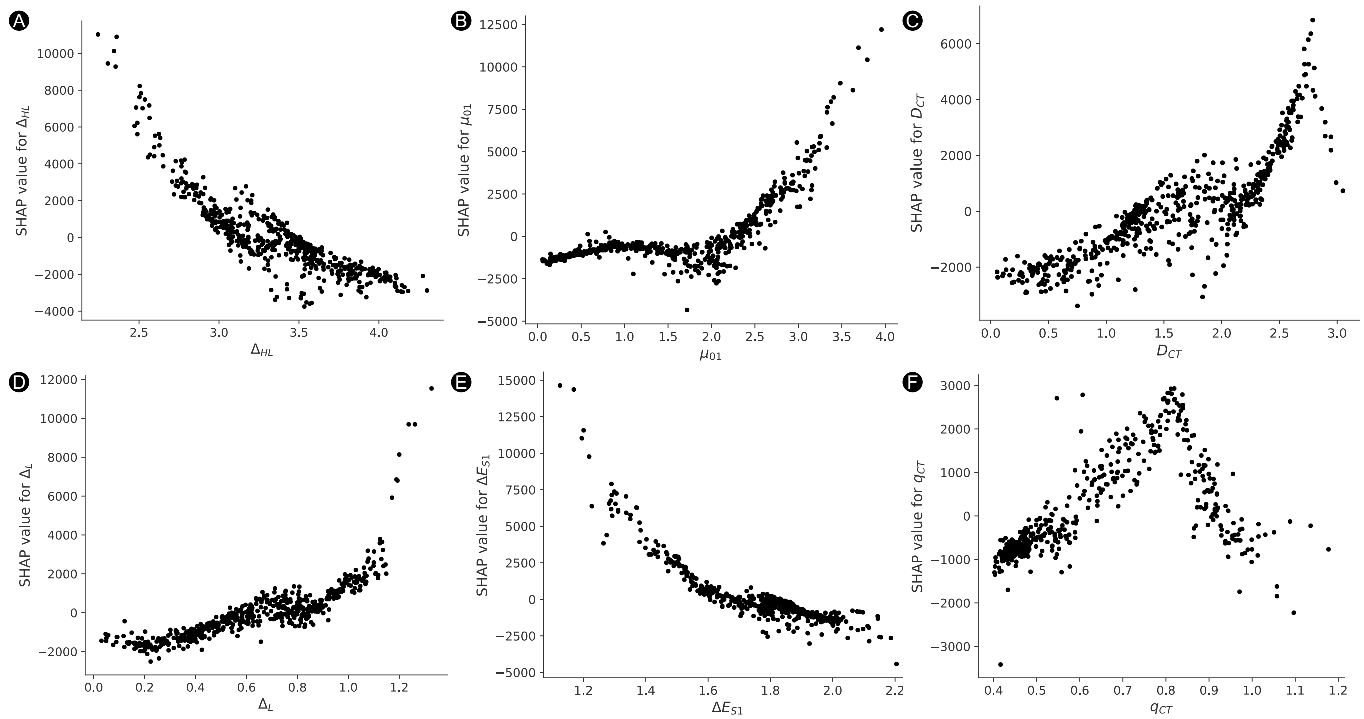

Figure S5: Dependency plots highlighting the feature SHAP *versus* the feature value for (A) Δ_{HL} (B) μ_{01}, (C) $\mathrm{D_{CT}}$, (D) Δ_{L}, (E) ΔE_{S1} and (F) $\mathrm{q_{CT}}$.

Force plots

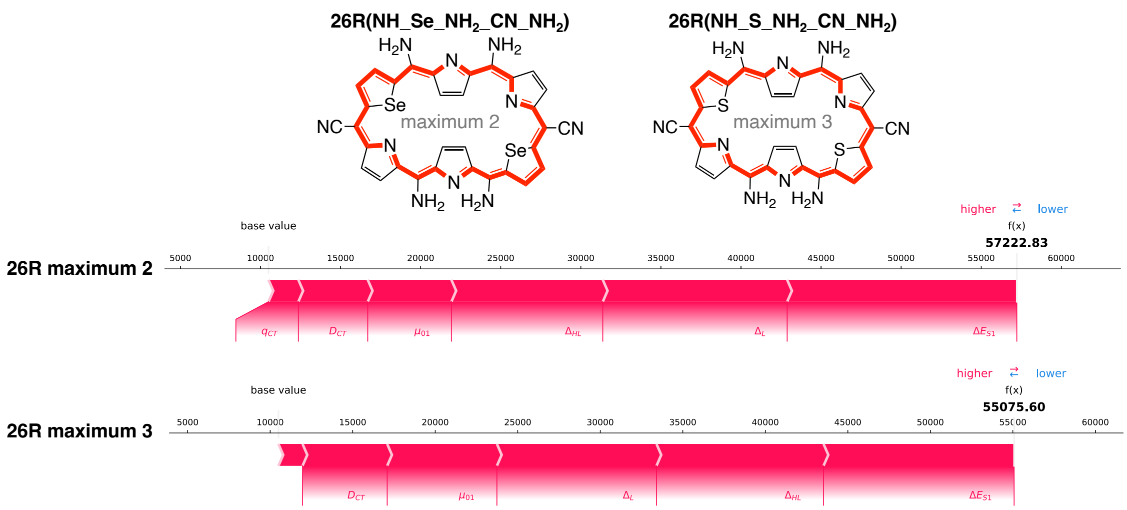

Figure S6: Force plot of the **26R** system with the second and third maximal β_{HRS} response denoted as **26R(NH_Se_NH$_2$_CN_NH$_2$)** and **26R(NH_S_NH$_2$_CN_NH$_2$)**, respectively. Features highlighted in red positively contribute with respect to the base value, while those highlighted in blue lower the NLO response prediction.

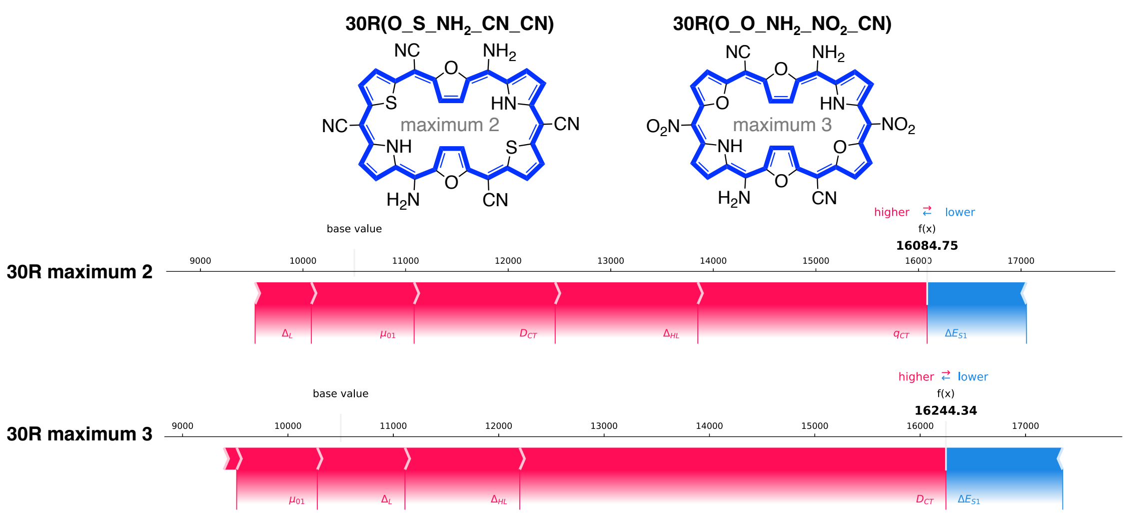

Figure S7: Force plot of the **30R** system with the second and third maximal β_{HRS} response denoted as **30R(O_S_NH$_2$_CN_CN)** and **30R(O_O_NH$_2$_NO$_2$_CN)**, respectively. Features highlighted in red positively contribute with respect to the base value, while those highlighted in blue lower the NLO response prediction.

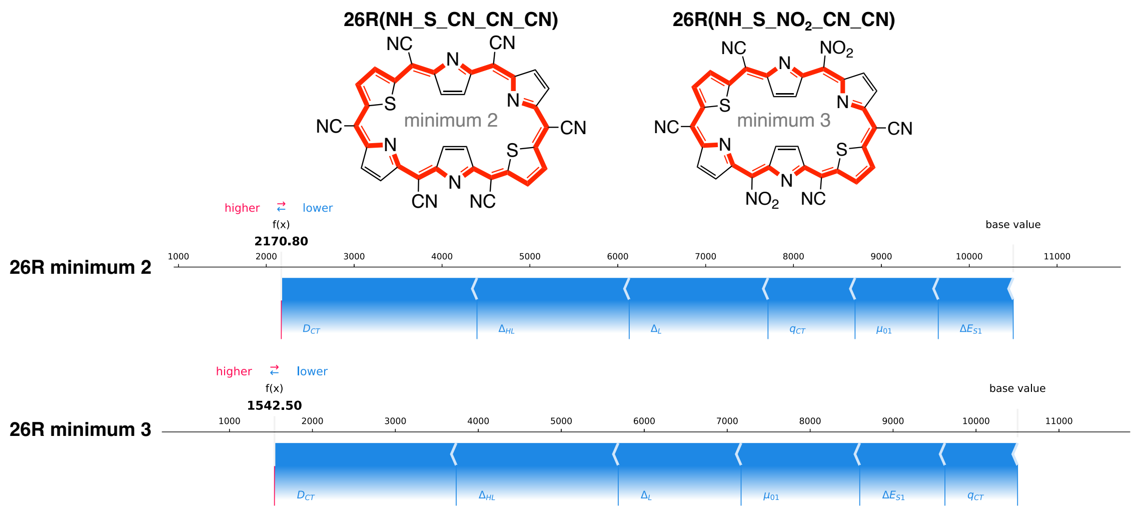

Figure S8: Force plot of the **26R** system with the second and third minimal β_{HRS} response denoted as **26R(NH_S_CN_CN_CN)** and **26R(NH_S_NO$_2$_CN_CN)**, respectively. Features highlighted in red positively contribute with respect to the base value, while those highlighted in blue lower the NLO response prediction.

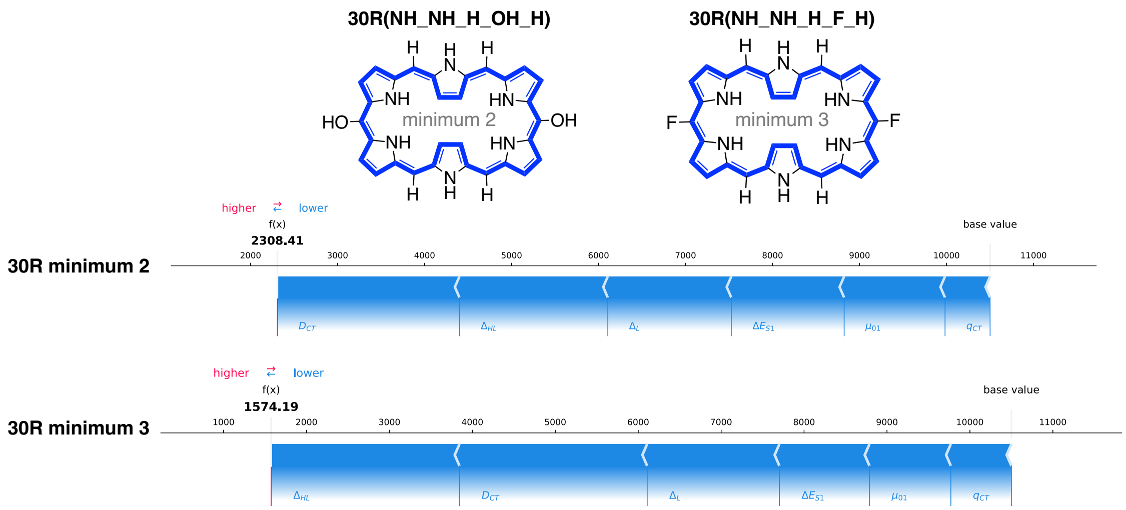

Figure S9: Force plot of the **30R** system with the second and third minimal β_{HRS} response denoted as **30R(NH_NH_H_OH_H)** and **30R(NH_NH_H_F_H)**, respectively. Features highlighted in red positively contribute with respect to the base value, while those highlighted in blue lower the NLO response prediction.

External test sets

Oudar and Chemla's model

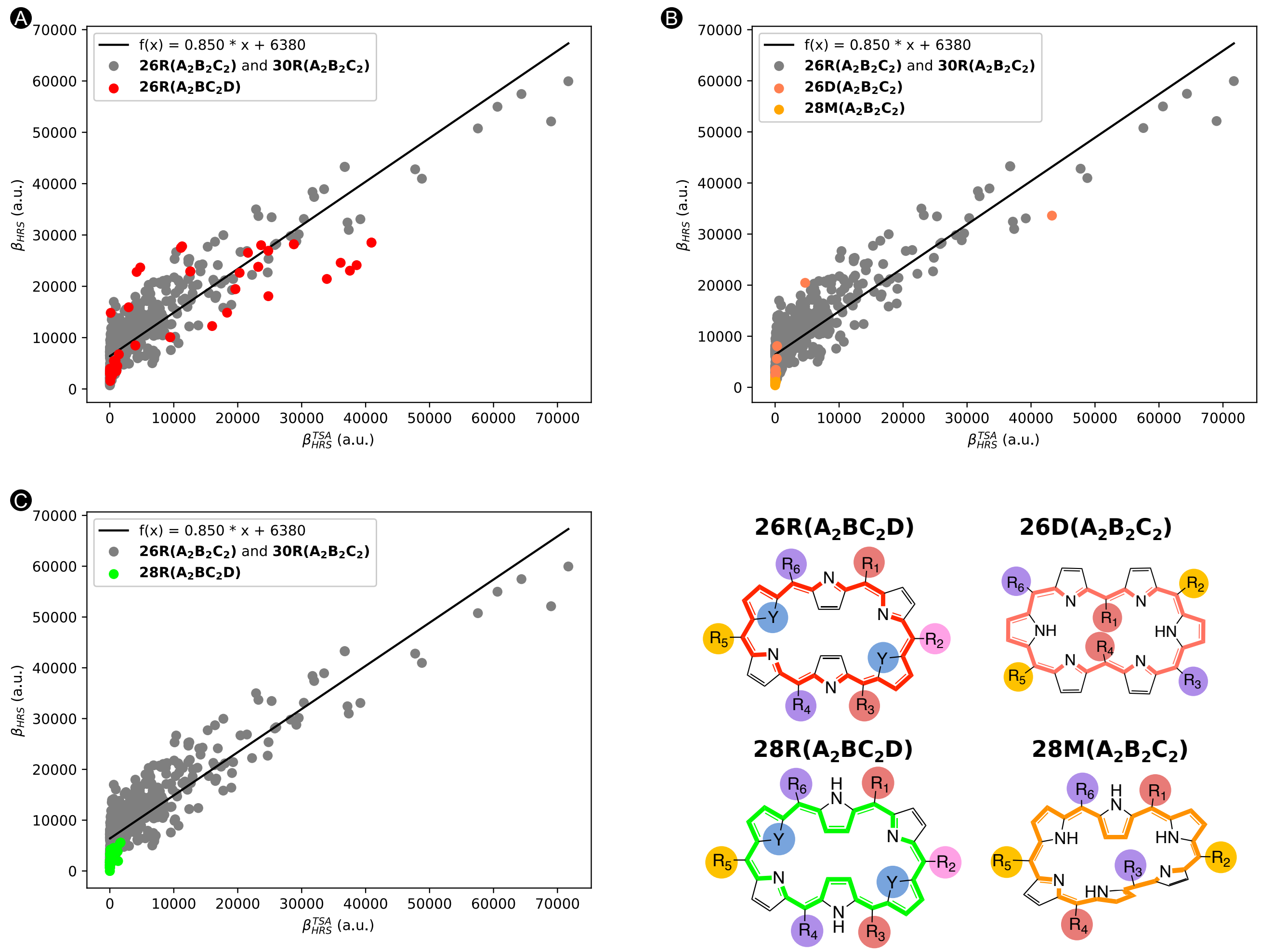

Figure S10: Scatter plot of the β_{HRS} response (in a.u.) versus the β_{HRS}^{TSA} (in a.u.). The linear regression line is portrayed by solid black line and its mathematical expression is given in the box. Data points from the initial dataset are highlighted in grey, but each new test set is coloured in red, salmon, orange and green for **26R(A_2BC_2D)**, **26D($A_2B_2C_2$)**, **28M($A_2B_2C_2$)** and **28R(A_2BC_2D)**, respectively.

Truth of predictions plots

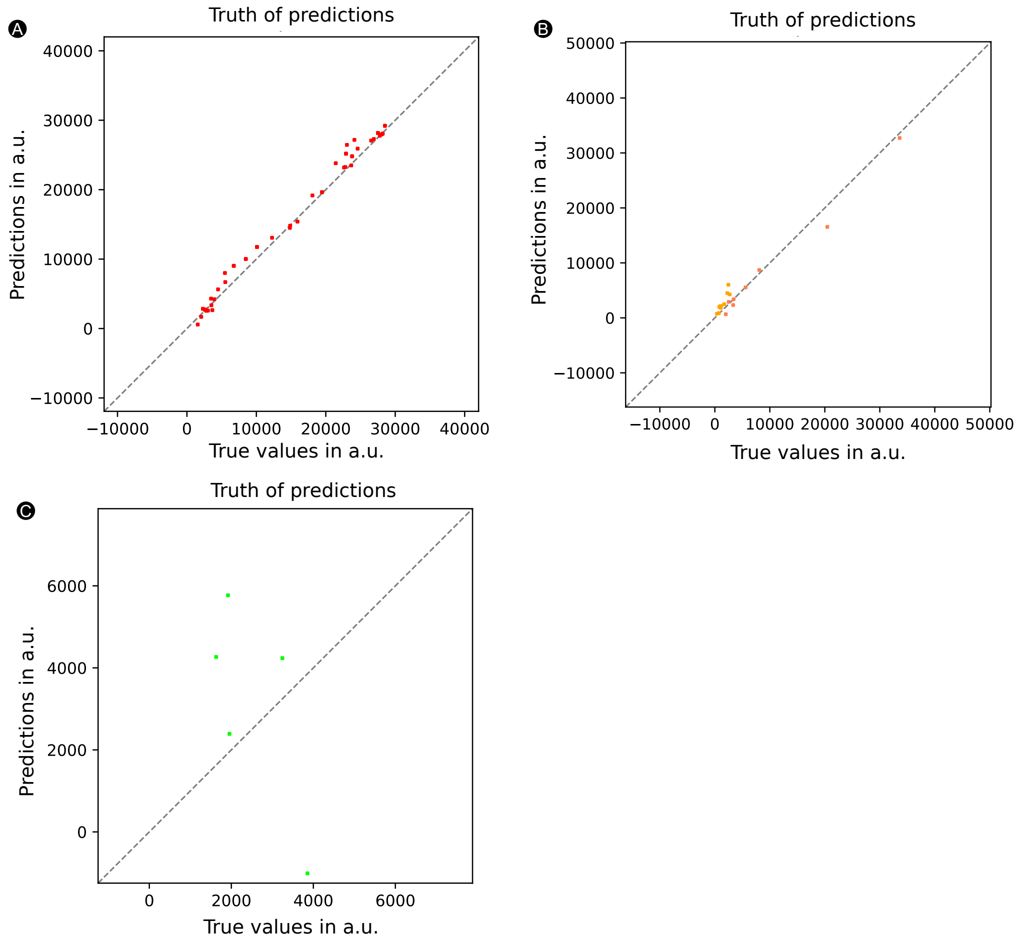

Figure S11: Truth of predictions plot: Scatter plot of True values ($i.e$, β_{HRS} based on quantum chemical calculations) and Predictions (β_{HRS} predicted by Model 2) on the new test sets (A) **26R(A_2BC_2D)**, (B) **26D($A_2B_2C_2$)** + **28M($A_2B_2C_2$)**, and (C) **28R(A_2BC_2D)**.

Dependency plots

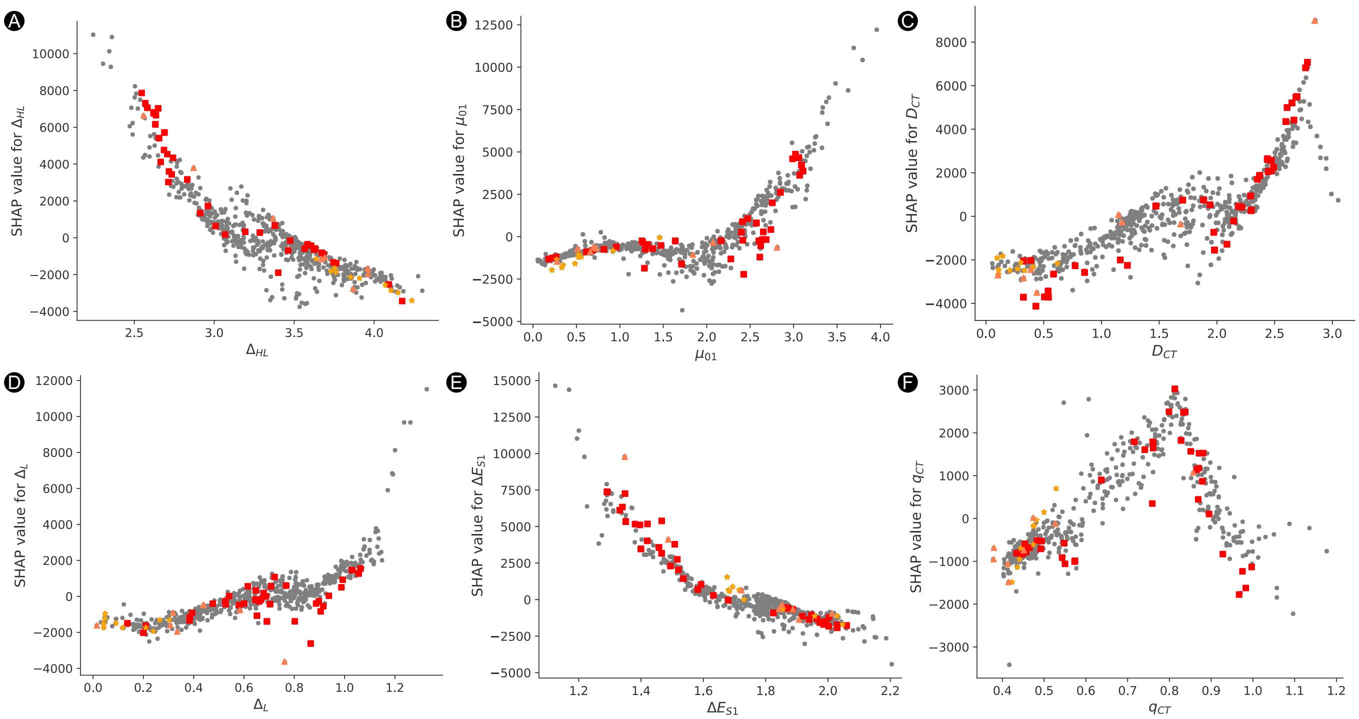

Figure S12: Dependency plots highlighting the feature SHAP *versus* the feature value for (A) Δ_{HL} (B) μ_{01}, (C) D_{CT}, (D) Δ_L, (E) Δ_{L+1_H} and (F) q_{CT}. Data points from the initial dataset are highlighted in grey, but each new test set is coloured in red, salmon, and orange for **26R(A_2BC_2D)**, **26D($A_2B_2C_2$)**, and **28M($A_2B_2C_2$)**, respectively.