**Supplementary Information for**

**QSAR model and microscopic mechanism analysis of dyes removal by coagulation of aluminum chloride under alkaline condition**

Meilan Zhang a,b, Shengnan Qin b, Yujia Tan c, Zhemin Shen c,[[1]](#footnote-1)\*

a. The State Key Laboratory of Pollution Control and Resource Reuse, School of Environmental Science and Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, China

b. Shanghai Laogang Waste Disposal Co. Ltd., Shanghai 201302, China.

c. School of Environmental Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

**\*Corresponding author** **contact.**

Zhemin Shen

Dongchuan Road, Minhang District, Shanghai, 200240, China.

Tel: +86-21-54745262, Fax: +86-21-54742863

Email: zmshen@sjtu.edu.cn

**Totally 22 pages including 3 Text and 6 Tables.**

**Table of Contents**

**Text S1.** Calculation details of parameters. (Page S3)

**Text S2.** Y test procedure. (Page S3)

**Text S3.** Applicability domain. (Page S3)

**Table S1.** The information of dyes used in the experiment. (Page S4-S5)

**Table S2.** The information of inorganic compounds used in the experiment. (Page S6)

**Table S3.** Definition of molecular parameters used in this study. (Page S7-S9)

**Table S4-1.** The molecular parameters of 38 kinds of dyes. (Page S10-S11)

**Table S4-2.** The molecular parameters of 38 kinds of dyes. (Page S12-S13)

**Table S4-3.** The molecular parameters of 38 kinds of dyes. (Page S14-S15)

**Table S4-4.** The molecular parameters of 38 kinds of dyes. (Page S16-S17)

**Table S5.** Correlation analysis between dye removal rate and 46 molecular parameters. (Page S18-S19)

**Table S6.** Experimental and predicted values of removal rate at pH=10. (Page S20-S21)

**References.** (Page S22)

**Text S1. Calculation details of parameters**

A total of 20 quantum chemical parameters were calculated by using Gaussian 09 and Material Studio 7.0 based on the density functional theory (DFT). In Gaussian 09, the calculation method and basis set were B3LYP and 6-311G (d, p), respectively. The parameters included the total energy of a molecule (EB3LYP), dipole moment (μ), energy of the highest occupied molecular orbital (EHOMO), energy of the lowest unoccupied molecular orbital (ELUMO), the gap energy (EGAP), the most positive partial charge on a hydrogen atom (qH+), negative partial charge on a carbon atom (qC-), and positive partial charge on a hydrogen atom linked with a carbon atom (q(CH)+). In Material Studio 7.0, the B3LYP method was selected to match the results of Gaussian 09. The maximum force, displacement and iterations were set as 0.004 Ha/Å, 500 and 0.3 Å, respectively. The SCF tolerance was set to 10-6 a.u. The bond order values (BOn, BOx) and Fukui indexes (f(+)x/n, f(−)x/n, f(0)x/n) were obtained from the out mol files. In addition, structural information to distinguish compound properties was also recorded, such as the relative molecular mass (MM) of organic compounds, the number of carbon atoms (NC), the number of hydrogen atoms (NH), the number of oxygen atoms (NO), the number of nitrogen atoms (NO), and the ratio of oxygen and carbon atoms (NO:C), the ratio of carbon and hydrogen atoms (NC:H), the ratio of nitrogen and carbon atoms (NN:C) in a molecule.

**Text S2. Y test procedure**

First, the dependent variables (Rexp) were rearranged in the training set twenty times randomly. Then the original independent variable matrix was kept unchanged to get 20 new models. Finally, the R2 and QINT2 values of the 20 new models were obtained according to the method mentioned above. If R2 and QINT2 of the new models are smaller than that of the original model, it means that the original model has good validity and robustness.

**Text S3. Applicability domain**

According to the OECD (Organization for Economic Co-operation and Development) principles, a QSAR model should have a defined domain of applicability (OECD, 2007). The domain of application for a QSAR model describes whether the model will predict an endpoint for a specific chemical with a given reliability. It helps the users of the model to judge whether the prediction for a new chemical is reliable or not, the predictions for only those compounds that fall into this domain could be considered reliable (Melagraki and Afantitis, 2014). In this study, the applicability domain (APD) was visualized by using the Williams plot (Gramatica, 2007). It helps estimate whether the compound is in the applicability domain according to the standardized residuals (σ) and leverage () values, and then judge the accuracy of the prediction (Gupta and Basant, 2017; Ortiz et al., 2017). The standardized residual (σ) is defined as (S-E1).

(S-E1)

where is the Rexp value of the ith compound and is the Rpre value; n is the number of compounds in the training set. Leverage can be calculated as (S-E2).

(S-E2)

where is a vector of descriptor for the ith compound and *X* is the descriptor matrix from the training set. The warning leverage () is defined as (S-E3).

(S-E3)

where m is the number of predictor variables, and n is the number of training compounds.

**Table S1.** The information of dyes used in the experiment.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Compounds | Abbreviation | Molecular formula | CAS No. | Molecular mass | Purity | Source |
| 1 | Acid Green 1 | AG1 | C30H15FeN3Na3O15S3 | 19381-50-1 | 878.45 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 2 | Rhodamine B | RB | C28H31ClN2O3 | 81-88-9 | 479.01 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 3a | Cresol Red | CR | C21H17NaO5S | 1733-12-6 | 382.43 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 4 | Acid Blue 93 | AB93 | C37H27N3Na2O9S3 | 28983-56-4 | 799.80 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 5 | Acid sky blue A | ASBA | C37H35N2NaO6S2 | 3486-30-4 | 690.80 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 6 | Bromophenol Blue | BB | C19H10Br4O5S | 115-39-9 | 669.96 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 7 | Methylene Blue trihydrate | MBT | C16H24ClN3O3S | 7220-79-3 | 373.90 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 8 | Cresol Purple | CP | C21H18O5S | 2303-01-7 | 382.43 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 9a | Crystal Violet | CV | C25H30ClN3 | 548-62-9 | 407.99 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 10 | Disperse Red 16 | DR16 | C24H26IN5O7 | 61968-52-3 | 519.94 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 11 | Methyl Red | MR | C15H15N3O2 | 493-52-7 | 269.30 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 12 | Bromocresol green | BG | C21H14Br4O5S | 76-60-8 | 698.01 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 13a | Mordant Black 17 | MB17 | C20H13N2O5SNa | 2538-85-4 | 416.39 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 14 | Acid Orange 10 | AO10 | C16H10N2Na2O7S2 | 1936-15-8 | 452.37 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 15 | Basic Brown 1 | BB1 | C18H20Cl2N8 | 10114-58-6 | 419.31 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 16 | Metanil Yellow | MY | C18H14N3NaO3S | 587-98-4 | 375.38 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 17 | Curcumin | Curcumin | C21H20O6 | 458-37-7 | 368.38 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 18a | Tris 4-aminophenyl methanol | T4-AM | C19H19N3O | 467-62-9 | 305.37 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 19 | Direct Red 28 | DR28 | C32H22N6Na2O6S2 | 573-58-0 | 696.66 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 20 | Azure B | AB | C15H16ClN3S | 531-55-5 | 305.83 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 21 | Fuchsin Basic | FB | C20H20ClN3 | 3248-93-9 | 301.38 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 22a | Isatin | Isatin | C8H5NO2 | 91-56-5 | 147.13 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 23 | Methyl Orange | MO | C14H14N3NaO3S | 547-58-0 | 327.33 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 24 | Indigo | Indigo | C16H10N2O2 | 482-89-3 | 262.26 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 25 | Acid Orange 7 | AO7 | C16H11N2NaO4S | 633-96-5 | 350.32 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 26 | Acid Orange 74 | AO74 | C16H12N5NaO7S | 10127-27-2 | 441.35 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 27a | Carmine | Carmine | C22H20O13 | 1390-65-4 | 492.39 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 28 | Ponceau S | PS | C22H12N4Na4O13S4 | 6226-79-5 | 760.57 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 29 | Mordant Yellow 1 | MY1 | C13H8N3NaO5 | 584-42-9 | 309.21 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 30 | Acid scarlet | AS | C18H14N2Na2O7S2 | 3761-53-3 | 480.42 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 31 | Reactive Yellow 3 | RY3 | C27H18ClN8Na3O10S3 | 6539-67-9 | 815.09 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 32a | Eriochrome Black T | EBT | C20H12N3NaO7S | 1787-61-7 | 461.38 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 33 | Fuchsin Acid | FA | C20H17N3Na2O9S3 | 3244-88-0 | 585.54 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 34 | Acid Black 1 | AB1 | C22H14N6Na2O9S2 | 1064-48-8 | 616.49 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 35 | Brilliant Red X 3B | BRX-3B | C19H10Cl2N6Na2O7S2 | 17804-49-8 | 615.33 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 36 | Basic Blue 26 | BB26 | C33H32N3Cl | 2580-56-5 | 506.09 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 37a | Acid chrome blue K | ACBK | C16H9N2Na3O12S3 | 3270-25-5 | 586.40 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 38 | Reactive Black 5 | RB5 | C22H16N2O11S3Na2 | 12225-25-1 | 626.55 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 39 | Disperse Red 60 | DR60 | C20H13NO4 | 12223-37-9 | 331.32 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 40 | Vat Blue 4 | VB4 | C28H14N2O4 | 81-77-6 | 442.42 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 41 | Vat Violet 1 | VV1 | C34H14Cl2O2 | 1324-55-6 | 525.38 | AR | Sinopharm Chemical Reagent Co., Ltd. |

a: Samples in the external test set.

**Table S2.** The information of inorganic compounds used in the experiment.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| No. | Compounds | Molecular formula | CAS No. | Molecular mass | Purity | Source |
| 1 | Magnesium Chloride | MgCl2 | 7786-30-3 | 95.21 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 2 | Sodium Hydroxide | NaOH | 1310-73-2 | 40.00 | AR | Sinopharm Chemical Reagent Co., Ltd. |
| 3 | Sulfuric Acid | H2SO4 | 7664-93-9 | 98.08 | AR | Sinopharm Chemical Reagent Co., Ltd. |

**Table S3.** Definition of molecular parameters used in this study.

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Descriptor | Note | Unit |
| 1 | ESPMAX | Maximum electrostatic potential of molecular surface | a.u. |
| 2 | ESPMEAN | Average electrostatic potential of molecular surface | a.u. |
| 3 | ESPMIN | Minimum electrostatic potential of molecular surface | a.u. |
| 4 | ESP0.25 | The value of the electrostatic potential at the quarter position of the molecular surface | a.u. |
| 5 | ESP0.75 | The value of molecular surface electrostatic potential at three fourths position | a.u. |
| 6 | E(B3LYP) | The total energy of a molecule (EB3LYP) | a.u. |
| 7 | μ | Dipole moment | Debye |
| 8 | qH+ | The most positive partial charge on a hydrogen atom | e |
| 9 | q(CH+)n | Minimum value of positive partial charge on a hydrogen atom linked with a carbon atom | e |
| 10 | q(CH+)x | Maximum value of positive partial charge on a hydrogen atom linked with a carbon atom | e |
| 11 | q(C-)n | Minimum value of negative partial charge on a carbon atom | e |
| 12 | q(C-)x | Maximum value of negative partial charge on a carbon atom | e |
| 13 | ⅀q(O) | Total charge of oxygen atom in molecule | e |
| 14 | ⅀q(N) | Total charge of nitrogen atom in molecule | e |
| 15 | ⅀q(O+N) | The sum of the total charges of the oxygen and nitrogen atoms in molecule | e |
| 16 | ⅀q(H) | Total charge of hydrogen atom in molecule | e |
| 17 | ⅀q(H)/NH | Average charge of a single hydrogen atom | - |
| 18 | ⅀q(-) | The sum of all the negative charges of a molecule | e |
| 19 | ⅀q(-)/NC | The average negative charge shared by a single carbon atom in a molecule | - |
| 20 | EHOMO | Energy of the highest occupied molecular orbital | eV |
| 21 | ELUMO | Energy of the lowest unoccupied molecular orbital | eV |
| 22 | ESUM | Sum of ELUMO and EHOMO, ESUM=ELUMO+EHOMO | eV |
| 23 | EGAP | Gap of ELUMO and EHOMO, EGAP=ELUMO-EHOMO | eV |
| 24 | EGAP2 | The square of EGAP | - |
| 25 | I | Ionization energy | eV |
| 26 | A | Electron affinity | eV |
| 27 | η | Global hardness | eV |
| 28 | S | Global softness | - |
| 29 | χ | Absolute electronegativity | eV |
| 30 | CP | Chemical potential | eV |
| 31 | ω | Global electrophilicity power | eV |
| 32 | BOn | Minimum value of bond order | - |
| 33 | BOx | Maximum value of bond order | - |
| 34 | f(+)n | Minimum value of nucleophilic Fukui index | e |
| 35 | f(+)x | Maximum value of nucleophilic Fukui index | e |
| 36 | f(-)n | Minimum value of electrophilic Fukui index | e |
| 37 | f(-)x | Maximum value of electrophilic Fukui index | e |
| 38 | f(0)n | Minimum value of free radical Fukui index | e |
| 39 | f(0)x | Maximum value of free radical Fukui index | e |
| 40 | SAA | Surface Area (Approx) | - |
| 41 | V | Molecular volume | - |
| 42 | HE | Hydration Energy | kcal/mol |
| 43 | Log P | Oil-water partition coefficient | - |
| 44 | R | Refractivity | - |
| 45 | P | Polarizability | - |
| 46 | MW | Molecular Weight | g/mol |

**Table S4-1.** The molecular parameters of 38 kinds of dyes.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Dyes | ESPMAX | ESPMEAN | ESPMIN | ESP0.25 | ESP0.75 | E(B3LYP) | μ | qH+ | q(CH+)n | q(CH+)x | q(C-)n | q(C-)x |
| (a.u.) | (a.u.) | (a.u.) | (a.u.) | (a.u.) | (a.u.) | (Debye) | (e) | (e) | (e) | (e) | (e) |
| 1 | AG1 | 88.3370 | 0.0409 | -0.0855 | -0.0223 | 44.1890 | -5388.7400 | 16.5020 | 0.286 | 0.259 | 0.286 | -0.222 | 0.429 |
| 2 | RB | 145.2300 | 0.0198 | -0.1340 | -0.0571 | 72.6249 | -1881.4300 | 8.7880 | 0.482 | 0.190 | 0.280 | -0.581 | 0.438 |
| 3 | CR | 69.4820 | 0.0305 | -0.0882 | -0.0289 | 34.7562 | -1585.5559 | 4.9301 | 0.467 | 0.201 | 0.224 | -0.605 | 0.346 |
| 4 | AB93 | 209.1500 | 0.0360 | -0.1194 | -0.0417 | 104.5930 | -3787.2394 | 13.6871 | 0.499 | 0.205 | 0.243 | -0.327 | 0.244 |
| 5 | ASBA | 61.8780 | 0.0304 | -0.0998 | -0.0347 | 30.9542 | -2790.3819 | 22.4467 | 0.483 | 0.179 | 0.252 | -0.600 | 0.182 |
| 6 | BB | 101.7400 | 0.0434 | -0.0859 | -0.0212 | 50.8917 | -11801.1500 | 7.6560 | 0.479 | 0.212 | 0.245 | -0.257 | 0.290 |
| 7 | MBT | 48.4330 | 0.0394 | -0.1202 | -0.0404 | 24.2362 | -1643.4005 | 12.0854 | 0.290 | 0.191 | 0.290 | -0.366 | 0.256 |
| 8 | CP | 89.9200 | 0.0345 | -0.0814 | -0.0235 | 44.9772 | -1585.6311 | 4.7888 | 0.463 | 0.196 | 0.236 | -0.592 | 0.356 |
| 9 | CV | 37.1390 | 0.0247 | -0.0924 | -0.0339 | 18.5819 | -1595.4592 | 4.0784 | 0.219 | 0.174 | 0.219 | -0.350 | 0.193 |
| 10 | DR16 | 59.8830 | 0.0363 | -0.1156 | -0.0397 | 29.9596 | -2152.9718 | 10.9835 | 0.391 | 0.158 | 0.258 | -0.692 | 0.852 |
| 11 | MR | 49.3700 | 0.0307 | -0.1194 | -0.0443 | 24.7004 | -895.5266 | 6.1273 | 0.474 | 0.185 | 0.226 | -0.351 | 0.818 |
| 12 | BG | 94.8580 | 0.0487 | -0.0809 | -0.0161 | 47.4534 | -11879.7800 | 6.1390 | 0.477 | 0.212 | 0.249 | -0.598 | 0.294 |
| 13 | MB17 | 30.6940 | 0.0309 | -0.0859 | -0.0275 | 15.3624 | -1816.4024 | 6.8500 | 0.490 | 0.198 | 0.240 | -0.254 | 0.381 |
| 14 | AO10 | 62.3150 | 0.0382 | -0.0790 | -0.0204 | 31.1766 | -2373.1568 | 14.6745 | 0.491 | 0.211 | 0.266 | -0.289 | 0.406 |
| 15 | BB1 | 37.8410 | 0.0319 | -0.0907 | -0.0294 | 18.9364 | -1135.0034 | 4.9925 | 0.423 | 0.192 | 0.224 | -0.337 | 0.242 |
| 16 | MY | 59.9010 | 0.0332 | -0.0925 | -0.0297 | 29.9671 | -1645.0402 | 8.6591 | 0.383 | 0.199 | 0.238 | -0.264 | 0.201 |
| 17 | Curcumin | 76.3770 | 0.0329 | -0.1096 | -0.0384 | 38.2049 | -1263.8883 | 3.3578 | 0.480 | 0.169 | 0.219 | -0.353 | 0.484 |
| 18 | T4-AM | 108.0700 | 0.0184 | -0.0888 | -0.0352 | 54.0442 | -975.1907 | 3.1488 | 0.457 | 0.194 | 0.218 | -0.257 | 0.285 |
| 19 | DR28 | 63.1530 | 0.0346 | -0.0920 | -0.0287 | 31.5938 | -3134.0328 | 12.9468 | 0.396 | 0.192 | 0.262 | -0.262 | 0.262 |
| 20 | AB | 59.8440 | 0.0332 | -0.1371 | -0.0519 | 29.9386 | -1604.0848 | 11.8733 | 0.384 | 0.187 | 0.236 | -0.367 | 0.306 |
| 21 | BF | 75.2530 | 0.0350 | -0.1378 | -0.0514 | 37.6440 | -1398.9075 | 22.1782 | 0.420 | 0.198 | 0.265 | -0.596 | 0.271 |
| 22 | Isatin | 124.9000 | 0.0481 | -0.0971 | -0.0245 | 62.4740 | -513.1963 | 5.9116 | 0.409 | 0.206 | 0.219 | -0.254 | 0.622 |
| 23 | MO | 60.5770 | 0.0223 | -0.0964 | -0.0371 | 30.2997 | -1492.4559 | 8.0654 | 0.230 | 0.186 | 0.230 | -0.351 | 0.211 |
| 24 | Indigo | 71.5760 | 0.0541 | -0.0904 | -0.0182 | 35.8150 | -875.9119 | 0.0002 | 0.434 | 0.203 | 0.215 | -0.251 | 0.507 |
| 25 | AO7 | 91.9710 | 0.0322 | -0.0895 | -0.0287 | 46.0016 | -1587.4971 | 7.9904 | 0.489 | 0.200 | 0.230 | -0.251 | 0.387 |
| 26 | AO74 | 46.8060 | 0.0409 | -0.0829 | -0.0210 | 23.4234 | -1977.9964 | 7.5116 | 0.492 | 0.208 | 0.266 | -0.607 | 0.552 |
| 27 | Carmine | 46.7930 | 0.0383 | -0.0975 | -0.0296 | 23.4156 | -1828.8050 | 8.2625 | 0.510 | 0.148 | 0.238 | -0.596 | 0.814 |
| 28 | PS | 128.6800 | 0.0478 | -0.0761 | -0.0141 | 64.3639 | -4285.0131 | 12.4768 | 0.437 | 0.215 | 0.243 | -0.305 | 0.440 |
| 29 | MY1 | 41.0070 | 0.0490 | -0.1009 | -0.0260 | 20.5280 | -1203.0840 | 7.7001 | 0.457 | 0.198 | 0.250 | -0.297 | 0.791 |
| 30 | AS | 91.9530 | 0.0345 | -0.0845 | -0.0250 | 45.9938 | -2451.7962 | 9.1499 | 0.499 | 0.191 | 0.227 | -0.584 | 0.354 |
| 31 | RY3 | 89.8770 | 0.0432 | -0.1104 | -0.0336 | 44.9601 | -3032.0056 | 9.0962 | 0.491 | 0.210 | 0.401 | -0.279 | 0.683 |
| 32 | EBT | 73.5120 | 0.0324 | -0.0902 | -0.0289 | 36.7722 | -2022.1368 | 11.5672 | 0.476 | 0.188 | 0.244 | -0.252 | 0.447 |
| 33 | FA | 55.8340 | 0.0461 | -0.0966 | -0.0253 | 27.9401 | -3133.2784 | 2.5151 | 0.499 | 0.203 | 0.239 | -0.599 | 0.225 |
| 34 | AB1 | 65.6870 | 0.0358 | -0.0792 | -0.0217 | 32.8614 | -2973.6926 | 10.0412 | 0.428 | 0.210 | 0.246 | -0.289 | 0.491 |
| 35 | BRX-3B | 95.6890 | 0.0418 | -0.0761 | -0.0171 | 47.8654 | -3626.9843 | 12.4315 | 0.502 | 0.206 | 0.243 | -0.231 | 0.656 |
| 36 | BB26 | 105.8700 | 0.0271 | -0.0638 | -0.0184 | 52.9485 | -1902.8361 | 9.3990 | 0.419 | 0.175 | 0.252 | -0.344 | 0.199 |
| 37 | ACBK | 83.1190 | 0.0471 | -0.0760 | -0.0145 | 41.5830 | -3309.2926 | 5.2536 | 0.503 | 0.202 | 0.239 | -0.282 | 0.406 |
| 38 | RB5 | 101.4400 | 0.0349 | -0.0956 | -0.0304 | 50.7374 | -3286.1672 | 8.8369 | 0.441 | 0.228 | 0.277 | -0.591 | 0.477 |

**Table S4-2.** The Quantum parameters of 38 kinds of dyes.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Dyes | ⅀q(O) | ⅀q(N) | ⅀q(O+N) | ⅀q(H) | ⅀q(H)/NH | ⅀q(-) | ⅀q(-)/NC | EHOMO | ELUMO | ESUM | EGAP | EGAP2 |
| (e) | (e) | (e) | (e) | - | (e) | - | (eV) | (eV) | (eV) | (eV) | - |
| 1 | AG1 | -3.542 | 0.045 | -3.497 | 4.211 | 0.281 | -3.497 | -0.117 | -6.3700 | -0.3810 | -6.7510 | 5.9890 | 35.8681 |
| 2 | RB | -1.788 | -0.927 | -2.715 | 6.406 | 0.207 | -2.715 | -0.097 | -6.2340 | -0.3270 | -6.5610 | 5.9070 | 34.8926 |
| 3 | CR | -2.947 | 0.000 | -2.947 | 4.328 | 0.255 | -2.947 | -0.140 | -6.7647 | 0.2993 | -6.4653 | 7.0640 | 49.8998 |
| 4 | AB93 | -6.673 | -1.811 | -8.484 | 5.101 | 0.189 | -8.484 | -0.229 | -7.1538 | -0.8435 | -7.9973 | 6.3102 | 39.8190 |
| 5 | ASBA | -5.401 | -0.891 | -6.292 | 8.626 | 0.246 | -6.292 | -0.170 | -4.0789 | -0.3020 | -4.3810 | 3.7769 | 14.2649 |
| 6 | BB | -3.782 | 0.000 | -3.782 | 2.787 | 0.279 | -3.782 | -0.199 | -6.8600 | -1.8510 | -8.7110 | 5.0090 | 25.0901 |
| 7 | MBT | 0.000 | -1.162 | -1.162 | 3.791 | 0.158 | -2.026 | -0.127 | -3.7279 | -1.2844 | -5.0123 | 2.4435 | 5.9709 |
| 8 | CP | -3.839 | 0.000 | -3.839 | 4.332 | 0.241 | -3.839 | -0.183 | -6.8000 | 0.0408 | -6.7592 | 6.8408 | 46.7972 |
| 9 | CV | 0.000 | -1.341 | -1.341 | 5.839 | 0.195 | -1.515 | -0.061 | -6.2912 | 1.1456 | -5.1456 | 7.4368 | 55.3055 |
| 10 | DR16 | -3.683 | -0.992 | -4.675 | 5.657 | 0.218 | -4.675 | -0.195 | -6.5823 | -3.6871 | -10.2694 | 2.8953 | 8.3825 |
| 11 | MR | -1.277 | -0.854 | -2.131 | 3.307 | 0.220 | -2.131 | -0.142 | -6.5715 | -0.2694 | -6.8408 | 6.3021 | 39.7161 |
| 12 | BG | -3.873 | 0.000 | -3.873 | 4.525 | 0.323 | -3.873 | -0.184 | -6.6970 | -1.7970 | -8.4940 | 4.9000 | 24.0100 |
| 13 | MB17 | -4.349 | -0.476 | -4.825 | 3.261 | 0.251 | -4.825 | -0.241 | -6.4408 | -0.9660 | -7.4068 | 5.4749 | 29.9740 |
| 14 | AO10 | -6.590 | -0.518 | -7.108 | 2.517 | 0.252 | -7.108 | -0.444 | -6.8545 | -2.3646 | -9.2191 | 4.4898 | 20.1584 |
| 15 | BB1 | 0.000 | -4.035 | -4.035 | 5.096 | 0.255 | -4.035 | -0.224 | -5.1374 | -1.9075 | -7.0449 | 3.2299 | 10.4325 |
| 16 | MY | -3.047 | -1.014 | -4.061 | 3.150 | 0.225 | -4.061 | -0.226 | -6.7130 | -0.4027 | -7.1157 | 6.3102 | 39.8190 |
| 17 | Curcumin | -3.652 | 0.000 | -3.652 | 4.693 | 0.235 | -3.652 | -0.174 | -6.1225 | 0.4054 | -5.7170 | 6.5279 | 42.6137 |
| 18 | T4-AM | -0.750 | -2.337 | -3.087 | 5.075 | 0.267 | -3.087 | -0.162 | -2.3021 | -2.6068 | -4.9089 | -0.3048 | 0.0929 |
| 19 | DR28 | -5.501 | -2.409 | -7.910 | 5.207 | 0.237 | -7.910 | -0.247 | -6.4925 | -0.5633 | -7.0558 | 5.9293 | 35.1563 |
| 20 | AB | 0.000 | -1.337 | -1.337 | 3.507 | 0.219 | -2.097 | -0.140 | -5.2272 | -3.4667 | -8.6939 | 1.7606 | 3.0995 |
| 21 | BF | 0.000 | -2.181 | -2.181 | 5.317 | 0.266 | -3.007 | -0.150 | -3.4585 | -3.7769 | -7.2354 | -0.3184 | 0.1014 |
| 22 | Isatin | -1.010 | -0.623 | -1.633 | 1.255 | 0.251 | -1.633 | -0.204 | -7.3687 | 0.0952 | -7.2735 | 7.4640 | 55.7110 |
| 23 | MO | -2.178 | -0.873 | -3.051 | 2.873 | 0.205 | -3.051 | -0.218 | -6.6667 | -0.4354 | -7.1021 | 6.2313 | 38.8293 |
| 24 | Indigo | -1.170 | -1.120 | -2.290 | 2.538 | 0.254 | -2.290 | -0.143 | -7.0096 | 0.2068 | -6.8028 | 7.2164 | 52.0758 |
| 25 | AO7 | -3.697 | -0.493 | -4.190 | 2.629 | 0.239 | -4.190 | -0.262 | -6.6123 | -1.2762 | -7.8885 | 5.3361 | 28.4737 |
| 26 | AO74 | -5.073 | -0.505 | -5.578 | 3.221 | 0.268 | -5.578 | -0.349 | -6.9959 | -3.4667 | -10.4626 | 3.5293 | 12.4557 |
| 27 | Carmine | -8.822 | 0.000 | -8.822 | 6.552 | 0.328 | -8.822 | -0.401 | -6.6504 | -3.7306 | -10.3810 | 2.9197 | 8.5249 |
| 28 | PS | -10.538 | -0.678 | -11.216 | 2.944 | 0.245 | -11.216 | -0.510 | -2.8844 | -4.2313 | -7.1157 | -1.3469 | 1.8143 |
| 29 | MY1 | -2.988 | 0.092 | -2.896 | 2.038 | 0.255 | -2.896 | -0.223 | -6.1334 | -3.5538 | -9.6871 | 2.5796 | 6.6544 |
| 30 | AS | -6.752 | -0.396 | -7.148 | 3.248 | 0.232 | -7.148 | -0.397 | -6.4300 | 0.1742 | -6.2558 | 6.6041 | 43.6143 |
| 31 | RY3 | -5.900 | -4.487 | -10.387 | 5.331 | 0.296 | -10.387 | -0.385 | -3.2490 | -3.2463 | -6.4953 | 0.0027 | 0.0000 |
| 32 | EBT | -5.495 | -0.315 | -5.810 | 3.876 | 0.323 | -5.810 | -0.291 | -6.1987 | -0.5170 | -6.7157 | 5.6817 | 32.2812 |
| 33 | FA | -8.759 | -2.222 | -10.981 | 4.829 | 0.284 | -10.981 | -0.549 | -3.1891 | -3.3116 | -6.5007 | -0.1224 | 0.0150 |
| 34 | AB1 | -7.434 | -1.203 | -8.637 | 3.740 | 0.267 | -8.637 | -0.393 | -7.0449 | -3.8395 | -10.8844 | 3.2055 | 10.2749 |
| 35 | BRX-3B | -6.248 | -2.266 | -8.514 | 2.750 | 0.275 | -8.514 | -0.448 | -3.5701 | -3.1755 | -6.7456 | 0.3946 | 0.1557 |
| 36 | BB26 | 0.000 | -1.553 | -1.553 | 7.335 | 0.229 | -2.284 | -0.069 | -6.1442 | 0.1497 | -5.9946 | 6.2939 | 39.6132 |
| 37 | ACBK | -11.126 | -0.428 | -11.554 | 2.834 | 0.315 | -11.554 | -0.722 | -3.2109 | -2.8463 | -6.0572 | 0.3646 | 0.1330 |
| 38 | RB5 | -5.899 | -1.349 | -7.248 | 4.666 | 0.292 | -7.248 | -0.329 | -4.3565 | -5.4449 | -9.8014 | -1.0884 | 1.1847 |

**Table S4-3.** The Quantum parameters of 38 kinds of dyes.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Dyes | I | A | η | S | χ | CP | ω | BOn | BOx |
| (eV) | (eV) | (eV) | - | (eV) | (eV) | (eV) | - | - |
| 1 | AG1 | 6.3700 | 0.3810 | 2.9945 | 0.3339 | 3.3755 | -3.3755 | 1.9025 | 0.6380 | 1.6190 |
| 2 | RB | 6.2340 | 0.3270 | 2.9535 | 0.3386 | 3.2805 | -3.2805 | 1.8219 | 0.9620 | 1.5300 |
| 3 | CR | 6.7647 | -0.2993 | 3.5320 | 0.2831 | 3.2327 | -3.2327 | 1.4794 | 0.8053 | 1.9669 |
| 4 | AB93 | 7.1538 | 0.8435 | 3.1551 | 0.3169 | 3.9987 | -3.9987 | 2.5339 | 0.8733 | 1.9361 |
| 5 | ASBA | 4.0789 | 0.3020 | 1.8884 | 0.5295 | 2.1905 | -2.1905 | 1.2704 | 0.8928 | 1.8400 |
| 6 | BB | 6.8600 | 1.8510 | 2.5045 | 0.3993 | 4.3555 | -4.3555 | 3.7873 | 0.9500 | 1.4150 |
| 7 | MBT | 3.7279 | 1.2844 | 1.2218 | 0.8185 | 2.5061 | -2.5061 | 2.5703 | 0.9081 | 1.5532 |
| 8 | CP | 6.8000 | -0.0408 | 3.4204 | 0.2924 | 3.3796 | -3.3796 | 1.6696 | 0.7973 | 1.9465 |
| 9 | CV | 6.2912 | -1.1456 | 3.7184 | 0.2689 | 2.5728 | -2.5728 | 0.8901 | 0.9152 | 1.4864 |
| 10 | DR16 | 6.5823 | 3.6871 | 1.4476 | 0.6908 | 5.1347 | -5.1347 | 9.1064 | 0.2245 | 1.8366 |
| 11 | MR | 6.5715 | 0.2694 | 3.1510 | 0.3174 | 3.4204 | -3.4204 | 1.8564 | 0.1311 | 1.8636 |
| 12 | BG | 6.6970 | 1.7970 | 2.4500 | 0.4082 | 4.2470 | -4.2470 | 3.6810 | 0.9540 | 1.4190 |
| 13 | MB17 | 6.4408 | 0.9660 | 2.7374 | 0.3653 | 3.7034 | -3.7034 | 2.5051 | 0.1448 | 1.8283 |
| 14 | AO10 | 6.8545 | 2.3646 | 2.2449 | 0.4455 | 4.6095 | -4.6095 | 4.7325 | 0.1031 | 1.8270 |
| 15 | BB1 | 5.1374 | 1.9075 | 1.6150 | 0.6192 | 3.5225 | -3.5225 | 3.8415 | 0.8921 | 1.6153 |
| 16 | MY | 6.7130 | 0.4027 | 3.1551 | 0.3169 | 3.5578 | -3.5578 | 2.0060 | 0.9382 | 1.8329 |
| 17 | Curcumin | 6.1225 | -0.4054 | 3.2640 | 0.3064 | 2.8585 | -2.8585 | 1.2517 | 0.8652 | 1.8016 |
| 18 | T4-AM | 2.3021 | 2.6068 | -0.1524 | -6.5625 | 2.4544 | -2.4544 | -19.7669 | 0.9381 | 1.4235 |
| 19 | DR28 | 6.4925 | 0.5633 | 2.9646 | 0.3373 | 3.5279 | -3.5279 | 2.0991 | 0.9240 | 1.8159 |
| 20 | AB | 5.2272 | 3.4667 | 0.8803 | 1.1360 | 4.3470 | -4.3470 | 10.7330 | 0.9080 | 1.5582 |
| 21 | BF | 3.4585 | 3.7769 | -0.1592 | -6.2820 | 3.6177 | -3.6177 | -41.1088 | 0.9467 | 1.4845 |
| 22 | Isatin | 7.3687 | -0.0952 | 3.7320 | 0.2680 | 3.6368 | -3.6368 | 1.7720 | 0.8848 | 1.9188 |
| 23 | MO | 6.6667 | 0.4354 | 3.1157 | 0.3210 | 3.5510 | -3.5510 | 2.0236 | 0.9328 | 1.8232 |
| 24 | Indigo | 7.0096 | -0.2068 | 3.6082 | 0.2771 | 3.4014 | -3.4014 | 1.6032 | 0.9396 | 1.7672 |
| 25 | AO7 | 6.6123 | 1.2762 | 2.6680 | 0.3748 | 3.9442 | -3.9442 | 2.9154 | 0.1506 | 1.8294 |
| 26 | AO74 | 6.9959 | 3.4667 | 1.7646 | 0.5667 | 5.2313 | -5.2313 | 7.7542 | 0.1305 | 1.8315 |
| 27 | Carmine | 6.6504 | 3.7306 | 1.4599 | 0.6850 | 5.1905 | -5.1905 | 9.2273 | 0.1254 | 1.8978 |
| 28 | PS | 2.8844 | 4.2313 | -0.6735 | -1.4848 | 3.5578 | -3.5578 | -9.3977 | 0.1417 | 1.7902 |
| 29 | MY1 | 6.1334 | 3.5538 | 1.2898 | 0.7753 | 4.8436 | -4.8436 | 9.0944 | 0.1034 | 1.6288 |
| 30 | AS | 6.4300 | -0.1742 | 3.3021 | 0.3028 | 3.1279 | -3.1279 | 1.4815 | 0.8274 | 1.8532 |
| 31 | RY3 | 3.2490 | 3.2463 | 0.0014 | 734.9969 | 3.2476 | -3.2476 | 3876.0498 | 0.8418 | 1.9493 |
| 32 | EBT | 6.1987 | 0.5170 | 2.8408 | 0.3520 | 3.3578 | -3.3578 | 1.9845 | 0.2292 | 1.8300 |
| 33 | FA | 3.1891 | 3.3116 | -0.0612 | -16.3333 | 3.2504 | -3.2504 | -86.2788 | 0.1005 | 1.9097 |
| 34 | AB1 | 7.0449 | 3.8395 | 1.6027 | 0.6239 | 5.4422 | -5.4422 | 9.2397 | 0.1051 | 1.8392 |
| 35 | BRX-3B | 3.5701 | 3.1755 | 0.1973 | 5.0689 | 3.3728 | -3.3728 | 28.8317 | 0.2293 | 1.8445 |
| 36 | BB26 | 6.1442 | -0.1497 | 3.1470 | 0.3178 | 2.9973 | -2.9973 | 1.4274 | 0.9107 | 1.5019 |
| 37 | ACBK | 3.2109 | 2.8463 | 0.1823 | 5.4851 | 3.0286 | -3.0286 | 25.1553 | 0.8272 | 1.8365 |
| 38 | RB5 | 4.3565 | 5.4449 | -0.5442 | -1.8375 | 4.9007 | -4.9007 | -22.0654 | 0.7973 | 1.8947 |

**Table S4-4.** The Quantum parameters of 38 kinds of dyes.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Dyes | f(+)n | f(+)x | f(-)n | f(-)x | f(0)n | f(0)x | SAA | V | HE | Log P | R | P | MW |
| (e) | (e) | (e) | (e) | (e) | (e) | - | - | kcal/mol | - | - | - | g/mol |
| 1 | AG1 | -0.050 | 0.056 | -0.030 | 0.057 | -0.060 | 0.057 | 1324.87 | 1569.64 | -63.39 | -4.69 | 196.69 | 59.58 | 878.45 |
| 2 | RB | -0.005 | 0.063 | -0.007 | 0.038 | -0.006 | 0.062 | 784.67 | 1273.62 | 53.04 | 1.33 | 144.65 | 51.42 | 479.01 |
| 3 | CR | -0.013 | 0.115 | -0.010 | 0.069 | 0.002 | 0.064 | 672.74 | 1001.77 | -30.21 | 0.01 | 115.12 | 38.44 | 382.43 |
| 4 | AB93 | 0.000 | 0.084 | 0.002 | 0.063 | 0.001 | 0.064 | 1441.05 | 1787.65 | -53.41 | -1.48 | 221.88 | 69.53 | 799.80 |
| 5 | ASBA | -0.002 | 0.089 | -0.003 | 0.079 | -0.002 | 0.053 | 966.13 | 1671.90 | -19.20 | 2.68 | 207.14 | 68.20 | 690.80 |
| 6 | BB | -0.008 | 0.095 | -0.010 | 0.039 | 0.002 | 0.053 | 1020.18 | 1139.10 | -34.36 | -0.09 | 136.69 | 45.28 | 669.96 |
| 7 | MBT | 0.003 | 0.110 | 0.011 | 0.084 | 0.009 | 0.078 | 631.71 | 857.43 | -6.16 | 0.43 | 95.56 | 33.52 | 373.90 |
| 8 | CP | -0.009 | 0.101 | -0.011 | 0.069 | 0.003 | 0.057 | 714.00 | 957.55 | -31.58 | 0.01 | 115.12 | 38.44 | 382.43 |
| 9 | CV | 0.003 | 0.089 | 0.007 | 0.051 | 0.007 | 0.048 | 695.13 | 1179.43 | -4.42 | 1.73 | 133.27 | 46.68 | 407.99 |
| 10 | DR16 | 0.000 | 0.129 | 0.001 | 0.099 | 0.001 | 0.074 | 1105.93 | 1381.00 | 159.77 | -5.77 | 142.05 | 50.34 | 519.94 |
| 11 | MR | 0.009 | 0.113 | 0.002 | 0.086 | 0.006 | 0.071 | 600.90 | 789.20 | 43.09 | 0.40 | 89.61 | 29.73 | 269.30 |
| 12 | BG | -0.012 | 0.071 | -0.009 | 0.035 | 0.002 | 0.056 | 1007.33 | 1177.80 | -32.33 | 0.22 | 145.25 | 48.95 | 698.01 |
| 13 | MB17 | 0.005 | 0.105 | 0.005 | 0.052 | 0.005 | 0.065 | 765.00 | 945.12 | -27.64 | -1.29 | 119.85 | 37.57 | 416.39 |
| 14 | AO10 | 0.002 | 0.096 | 0.004 | 0.074 | 0.003 | 0.073 | 739.12 | 912.86 | -27.97 | -1.69 | 110.29 | 32.52 | 452.37 |
| 15 | BB1 | 0.003 | 0.078 | 0.007 | 0.048 | 0.006 | 0.053 | 836.58 | 977.42 | -39.32 | -4.53 | 123.23 | 39.28 | 419.31 |
| 16 | MY | 0.004 | 0.127 | 0.002 | 0.100 | 0.003 | 0.085 | 744.99 | 937.24 | -25.38 | -0.07 | 113.45 | 34.94 | 375.38 |
| 17 | Curcumin | 0.001 | 0.075 | 0.004 | 0.058 | 0.004 | 0.056 | 764.68 | 968.16 | 99.99 | -0.14 | 111.16 | 38.95 | 368.38 |
| 18 | T4-AM | 0.001 | 0.049 | 0.004 | 0.080 | 0.003 | 0.051 | 647.75 | 902.88 | -24.82 | 2.52 | 95.76 | 36.28 | 305.37 |
| 19 | DR28 | 0.002 | 0.077 | 0.004 | 0.053 | 0.003 | 0.052 | 1248.24 | 1539.32 | -50.07 | -3.05 | 201.54 | 62.15 | 696.66 |
| 20 | AB | 0.002 | 0.113 | 0.010 | 0.112 | 0.009 | 0.080 | 633.20 | 800.87 | -10.11 | 0.07 | 90.27 | 31.68 | 305.83 |
| 21 | BF | 0.003 | 0.096 | 0.006 | 0.102 | 0.005 | 0.071 | 685.98 | 885.78 | -24.86 | -0.07 | 106.62 | 37.06 | 301.38 |
| 22 | Isatin | 0.026 | 0.164 | 0.020 | 0.119 | 0.035 | 0.124 | 419.19 | 426.72 | 107.81 | -0.27 | 42.65 | 14.85 | 147.13 |
| 23 | MO | 0.006 | 0.125 | 0.003 | 0.099 | 0.005 | 0.088 | 635.72 | 839.03 | -18.53 | -0.34 | 93.86 | 28.95 | 327.33 |
| 24 | Indigo | 0.012 | 0.096 | 0.008 | 0.071 | 0.015 | 0.080 | 625.46 | 705.81 | 102.41 | -0.87 | 84.98 | 28.57 | 262.26 |
| 25 | AO7 | 0.005 | 0.113 | 0.004 | 0.077 | 0.004 | 0.077 | 688.05 | 833.62 | -25.32 | -0.34 | 100.05 | 30.75 | 350.32 |
| 26 | AO74 | 0.001 | 0.143 | 0.003 | 0.091 | 0.004 | 0.082 | 806.85 | 1012.14 | -40.08 | -4.67 | 112.99 | 36.61 | 441.35 |
| 27 | Carmine | -0.002 | 0.085 | -0.003 | 0.062 | 0.002 | 0.063 | 983.64 | 1082.02 | 140.77 | -4.57 | 116.79 | 43.52 | 492.39 |
| 28 | PS | 0.001 | 0.087 | 0.002 | 0.136 | 0.002 | 0.073 | 1210.02 | 1402.88 | -55.90 | -4.03 | 169.83 | 47.79 | 760.57 |
| 29 | MY1 | 0.002 | 0.153 | 0.005 | 0.088 | 0.009 | 0.082 | 717.30 | 754.82 | 25.21 | -3.84 | 83.65 | 27.64 | 309.21 |
| 30 | AS | 0.003 | 0.114 | 0.004 | 0.068 | 0.004 | 0.090 | 824.10 | 1031.00 | -31.27 | -1.39 | 118.86 | 36.19 | 480.42 |
| 31 | RY3 | 0.000 | 0.083 | 0.000 | 0.154 | 0.000 | 0.096 | 1108.74 | 1283.99 | 16.67 | -1.74 | 153.56 | 49.16 | 815.09 |
| 32 | EBT | 0.003 | 0.148 | 0.004 | 0.062 | 0.005 | 0.080 | 880.26 | 1022.53 | -40.43 | -5.46 | 125.90 | 39.86 | 461.38 |
| 33 | FA | 0.002 | 0.103 | 0.002 | 0.090 | 0.003 | 0.085 | 912.07 | 1172.85 | -42.48 | -4.02 | 137.44 | 42.39 | 585.54 |
| 34 | AB1 | 0.003 | 0.078 | 0.004 | 0.062 | 0.003 | 0.047 | 1006.65 | 1267.41 | -51.47 | -7.21 | 158.93 | 47.89 | 616.49 |
| 35 | BRX-3B | 0.007 | 0.084 | 0.001 | 0.160 | 0.005 | 0.088 | 1063.57 | 1211.17 | -39.57 | 0.29 | 144.09 | 45.26 | 615.33 |
| 36 | BB26 | 0.001 | 0.084 | 0.006 | 0.045 | 0.005 | 0.045 | 848.21 | 1366.01 | -13.37 | 2.07 | 171.05 | 58.85 | 506.09 |
| 37 | ACBK | 0.004 | 0.102 | 0.003 | 0.067 | 0.003 | 0.073 | 980.38 | 1080.06 | -50.90 | -5.10 | 123.75 | 35.57 | 586.40 |
| 38 | RB5 | 0.000 | 0.087 | 0.000 | 0.082 | 0.000 | 0.063 | 1066.66 | 1260.35 | 80.86 | -3.34 | 144.39 | 44.52 | 626.55 |

**Table S5.** Correlation analysis between dye removal rate and 46 molecular parameters.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. | Parameters | Unit | Pearson Correlation coefficient | Statistical significance (P) |
| 1 | ESPMAX | (a.u) | -0.106 | 0.528 |
| 2 | ESPMEAN | (a.u) | 0.272 | 0.098 |
| 3 | ESPMIN | (a.u) | 0.284 | 0.084 |
| 4 | ESP0.25 | (a.u) | 0.347\* | 0.033 |
| 5 | ESP0.75 | (a.u) | -0.106 | 0.528 |
| 6 | E(B3LYP) | (a.u) | 0.212 | 0.200 |
| 7 | μ | (Debye) | -0.136 | 0.417 |
| 8 | qH+ | (e) | 0.240 | 0.146 |
| 9 | q(CH+)n | (e) | -0.010 | 0.950 |
| 10 | q(CH+)x | (e) | 0.030 | 0.858 |
| 11 | q(C-)n | (e) | 0.153 | 0.358 |
| 12 | q(C-)x | (e) | 0.220 | 0.185 |
| 13 | ⅀q(O) | (e) | -0.364\* | 0.025 |
| 14 | ⅀q(N) | (e) | -0.216 | 0.193 |
| 15 | ⅀q(O+N) | (e) | **-0.435\*\*** | **0.006** |
| 16 | ⅀q(H) | (e) | -0.264 | 0.110 |
| 17 | ⅀q(H)/NH | - | **0.430\*\*** | **0.007** |
| 18 | ⅀q(-) | (e) | **-0.452\*\*** | **0.004** |
| 19 | ⅀q(-)/NC | - | **-0.607\*\*** | **0.000** |
| 20 | EHOMO | (eV) | 0.309 | 0.059 |
| 21 | ELUMO | (eV) | **-0.494\*\*** | **0.002** |
| 22 | ESUM | (eV) | -0.231 | 0.163 |
| 23 | EGAP | (eV) | **-0.473\*\*** | **0.003** |
| 24 | EGAP2 | - | -0.404\* | 0.012 |
| 25 | I | (eV) | -0.309 | 0.059 |
| 26 | A | (eV) | **0.494\*\*** | **0.002** |
| 27 | η | (eV) | **-0.473\*\*** | **0.003** |
| 28 | S | - | 0.161 | 0.333 |
| 29 | χ | (eV) | 0.231 | 0.163 |
| 30 | CP | (eV) | -0.231 | 0.163 |
| 31 | ω | (eV) | 0.161 | 0.335 |
| 32 | BOn | - | -0.298 | 0.069 |
| 33 | BOx | - | 0.240 | 0.146 |
| 34 | f(+)n | (e) | **0.481\*\*** | **0.002** |
| 35 | f(+)x | (e) | 0.221 | 0.182 |
| 36 | f(-)n | (e) | **0.483\*\*** | **0.002** |
| 37 | f(-)x | (e) | 0.397\* | 0.014 |
| 38 | f(0)n | (e) | 0.357\* | 0.028 |
| 39 | f(0)x | (e) | 0.376\* | 0.020 |
| 40 | SAA | - | -0.045 | 0.789 |
| 41 | V | - | -0.240 | 0.147 |
| 42 | HE | kcal/mol | 0.039 | 0.815 |
| 43 | Log P | - | -0.364\* | 0.025 |
| 44 | R | - | -0.227 | 0.170 |
| 45 | P | - | -0.311 | 0.058 |
| 46 | MW | g/mol | -0.089 | 0.595 |

**\*\*: P<0.01;**

**\*: P<0.05.**

**Table S6.** Experimental and predicted values of removal rate at pH=10.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Dyes | Rexp | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
| Rpre | Diff. | Rpre | Diff. | Rpre | Diff. | Rpre | Diff. |
| 1 | AG1 | 0.0325 | -0.1473 | 0.1798 | -0.0669 | 0.0994 | -0.0576 | 0.0901 | 0.0293 | 0.0032 |
| 2 | RB | 0.0919 | 0.1097 | -0.0178 | 0.1401 | -0.0482 | 0.0543 | 0.0376 | -0.0165 | 0.1084 |
| 3 | CR | 0.1146 | 0.2547 | -0.1401 | 0.3487 | -0.2341 | 0.3863 | -0.2717 | 0.3378 | -0.2232 |
| 4 | AB93 | 0.1310 | 0.3394 | -0.2084 | 0.2570 | -0.1260 | 0.2062 | -0.0752 | 0.1369 | -0.0059 |
| 5 | ASBA | 0.1823 | 0.4123 | -0.2300 | 0.4332 | -0.2509 | 0.4344 | -0.2521 | 0.3871 | -0.2048 |
| 6 | BB | 0.1943 | 0.3869 | -0.1926 | 0.1851 | 0.0092 | 0.1815 | 0.0128 | 0.1629 | 0.0314 |
| 7 | MBT | 0.2334 | 0.3752 | -0.1418 | 0.2965 | -0.0631 | 0.2864 | -0.0530 | 0.3580 | -0.1246 |
| 8 | CP | 0.2365 | 0.1994 | 0.0371 | 0.2845 | -0.0480 | 0.3405 | -0.1040 | 0.2889 | -0.0524 |
| 9 | CV | 0.2431 | 0.3847 | -0.1416 | 0.3573 | -0.1142 | 0.4320 | -0.1889 | 0.5567 | -0.3136 |
| 10 | DR16 | 0.2714 | 0.4095 | -0.1381 | 0.4078 | -0.1364 | 0.3695 | -0.0981 | 0.3899 | -0.1185 |
| 11 | MR | 0.3119 | 0.4142 | -0.1023 | 0.4507 | -0.1388 | 0.4125 | -0.1006 | 0.3640 | -0.0521 |
| 12 | BG | 0.3341 | 0.5820 | -0.2479 | 0.4138 | -0.0797 | 0.4147 | -0.0806 | 0.4195 | -0.0854 |
| 13 | MB17 | 0.4251 | 0.6698 | -0.2447 | 0.6917 | -0.2666 | 0.7359 | -0.3108 | 0.6962 | -0.2711 |
| 14 | AO10 | 0.4792 | 0.7700 | -0.2908 | 0.7770 | -0.2978 | 0.8068 | -0.3276 | 0.7871 | -0.3079 |
| 15 | BB1 | 0.4859 | 0.7248 | -0.2389 | 0.7635 | -0.2776 | 0.7976 | -0.3117 | 0.8129 | -0.3270 |
| 16 | MY | 0.5199 | 0.4824 | 0.0375 | 0.4988 | 0.0211 | 0.5303 | -0.0104 | 0.5566 | -0.0367 |
| 17 | Curcumin | 0.5521 | 0.5397 | 0.0124 | 0.5690 | -0.0169 | 0.5537 | -0.0016 | 0.5113 | 0.0408 |
| 18 | T4-AM | 0.5748 | 0.6635 | -0.0887 | 0.7310 | -0.1562 | 0.7747 | -0.1999 | 0.7587 | -0.1839 |
| 19 | DR28 | 0.5881 | 0.5917 | -0.0036 | 0.5640 | 0.0241 | 0.5873 | 0.0008 | 0.6173 | -0.0292 |
| 20 | AB | 0.6587 | 0.6049 | 0.0538 | 0.5861 | 0.0726 | 0.4975 | 0.1612 | 0.5351 | 0.1236 |
| 21 | FB | 0.6856 | 0.7005 | -0.0149 | 0.7458 | -0.0602 | 0.6338 | 0.0518 | 0.6647 | 0.0209 |
| 22 | Isatin | 0.6878 | 1.0180 | -0.3302 | 1.0204 | -0.3326 | 1.0537 | -0.3659 | 1.0918 | -0.4040 |
| 23 | MO | 0.6887 | 0.4228 | 0.2659 | 0.4223 | 0.2664 | 0.4504 | 0.2383 | 0.5864 | 0.1023 |
| 24 | Indigo | 0.6959 | 0.6969 | -0.0010 | 0.7397 | -0.0438 | 0.7904 | -0.0945 | 0.7880 | -0.0921 |
| 25 | AO7 | 0.7556 | 0.6098 | 0.1458 | 0.6320 | 0.1236 | 0.6646 | 0.0910 | 0.6212 | 0.1344 |
| 26 | AO74 | 0.7604 | 0.7555 | 0.0049 | 0.7949 | -0.0345 | 0.8225 | -0.0621 | 0.7997 | -0.0393 |
| 27 | Carmine | 0.7946 | 0.8771 | -0.0825 | 0.9977 | -0.2031 | 0.9434 | -0.1488 | 0.9420 | -0.1474 |
| 28 | PS | 0.7955 | 0.7343 | 0.0612 | 0.6836 | 0.1119 | 0.7064 | 0.0891 | 0.7331 | 0.0624 |
| 29 | MY1 | 0.7989 | 0.6744 | 0.1245 | 0.7191 | 0.0798 | 0.7182 | 0.0807 | 0.7081 | 0.0908 |
| 30 | AS | 0.8244 | 0.6625 | 0.1619 | 0.6500 | 0.1744 | 0.6762 | 0.1482 | 0.6364 | 0.1880 |
| 31 | RY3 | 0.8405 | 0.8148 | 0.0257 | 0.8580 | -0.0176 | 0.7745 | 0.0659 | 0.7796 | 0.0608 |
| 32 | EBT | 0.8412 | 0.9651 | -0.1238 | 1.0486 | -0.2074 | 1.0446 | -0.2034 | 1.0609 | -0.2197 |
| 33 | FA | 0.8627 | 0.9137 | -0.0510 | 0.9329 | -0.0703 | 0.8727 | -0.0101 | 0.8813 | -0.0186 |
| 34 | AB1 | 0.8762 | 0.8013 | 0.0749 | 0.8041 | 0.0721 | 0.8350 | 0.0412 | 0.8681 | 0.0082 |
| 35 | BRX-3B | 0.8796 | 0.7922 | 0.0875 | 0.7932 | 0.0864 | 0.8170 | 0.0626 | 0.7997 | 0.0799 |
| 36 | BB26 | 0.8901 | 0.5045 | 0.3857 | 0.5023 | 0.3878 | 0.6551 | 0.2350 | 0.6333 | 0.2569 |
| 37 | ACBK | 0.9494 | 1.1660 | -0.2166 | 1.2013 | -0.2519 | 1.1694 | -0.2200 | 1.2053 | -0.2559 |
| 38 | RB5 | 0.9887 | 0.7634 | 0.2253 | 0.7954 | 0.1933 | 0.7701 | 0.2186 | 0.8008 | 0.1880 |

**References**

Gramatica, P., 2007. Principles of QSAR models validation: internal and external. QSAR & Combinatorial Science. 26**,** 694-701.

Gupta, S., Basant, N., 2017. Modeling the pH and temperature dependence of aqueousphase hydroxyl radical reaction rate constants of organic micropollutants using QSPR approach. Environmental Science and Pollution Research. 24**,** 24936-24946.

Melagraki, G., Afantitis, A., 2014. Enalos InSilicoNano platform: an online decision support tool for the design and virtual screening of nanoparticles. RSC Advances. 4**,** 50713-50725.

OECD, 2007. Guidance Document On The Validation of (Quantitative) Structure-activity Relationships [(Q)SAR] Models.

Ortiz, E. V., et al., 2017. The conformation-independent QSPR approach for predicting the oxidation rate constant of water micropollutants. Environmental Science and Pollution Research. 24**,** 27366-27375.

1. \*Corresponding author. Tel.: +86 21 54741065

   E-mail address: zmshen@sjtu.edu.cn. [↑](#footnote-ref-1)