Models Report – 2 parameters

## Neat

### Only Boronic Acids

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.1.2. + dip\_y | 0.7502635 | 0.6550705 | 0.0979900 |
| output ~ X.2.9. + B5 | 0.7523212 | 0.4994522 | 0.1156619 |

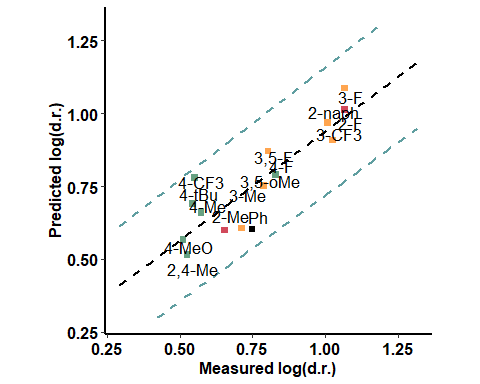
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0284190 | 26.751640 | 0.0000000 |
| X.1.2. | 0.1240787 | 0.0294255 | 4.216711 | 0.0011959 |
| dip\_y | 0.1287997 | 0.0294255 | 4.377149 | 0.0009008 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6138099 | 0.1070997 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6366031 | 0.1012295 |

***Top Ranked Full Model***



### With Catalyst - Close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ NBO.N.19 + diff.B1.O16 | 0.6684152 | 0.6025137 | 0.0923752 |
| output ~ Dist.8..14. + Dist.16..17. | 0.6849151 | 0.5438933 | 0.1078794 |

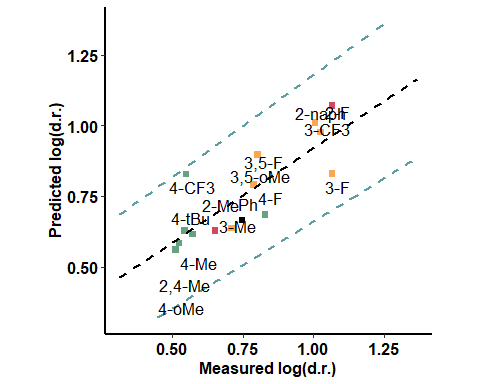
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0327466 | 23.216357 | 0.0000000 |
| NBO.N.19 | 0.1606125 | 0.0343340 | 4.677943 | 0.0005342 |
| diff.B1.O16 | 0.0770560 | 0.0343340 | 2.244306 | 0.0444510 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4914689 | 0.1235114 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.547022 | 0.110013 |

***Top Ranked Full Model***



### With Catalyst - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ dip\_y + diff.H14.O18 | 0.7555236 | 0.5734527 | 0.1034078 |

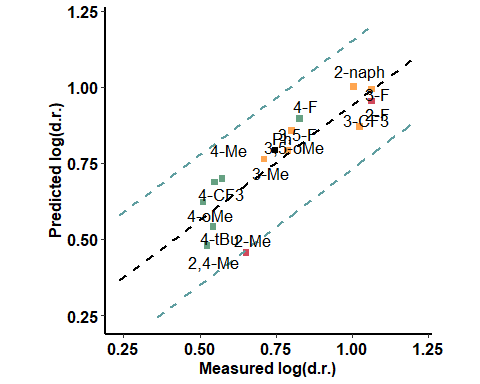
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0281182 | 27.037899 | 0.0000000 |
| dip\_y | 0.1068162 | 0.0305356 | 3.498088 | 0.0043973 |
| diff.H14.O18 | -0.1773930 | 0.0305356 | -5.809382 | 0.0000835 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5445424 | 0.1149639 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5596977 | 0.1086652 |

***Top Ranked Full Model***



### BA + Pentanone

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.2.9. + NBO.O.9 | 0.7416260 | 0.6443590 | 0.0948581 |
| output ~ X.2.9. + diff.O9.H16 | 0.6997166 | 0.6195804 | 0.0937588 |

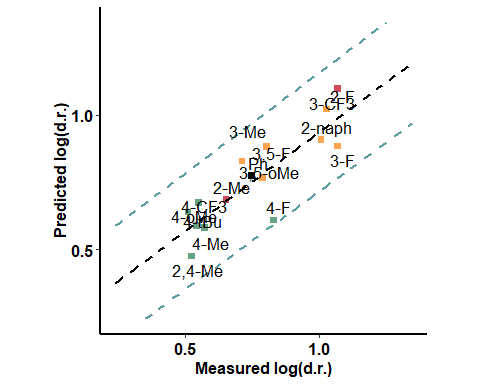
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0289063 | 26.300683 | 0.0000000 |
| X.2.9. | 0.2386137 | 0.0441253 | 5.407635 | 0.0001581 |
| NBO.O.9 | -0.1071345 | 0.0441253 | -2.427959 | 0.0318511 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6028835 | 0.1072041 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6241444 | 0.1005212 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - oxygen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.18..20. + dip\_y | 0.5730471 | 0.4039036 | 0.1174817 |
| output ~ Dist.2..9. + B1 | 0.5697977 | 0.3272534 | 0.1497725 |

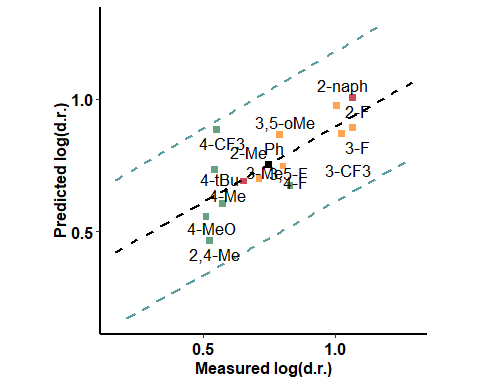
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0371585 | 20.459802 | 0.0000000 |
| Dist.18..20. | 0.1510292 | 0.0407446 | 3.706729 | 0.0029987 |
| dip\_y | -0.1090000 | 0.0407446 | -2.675199 | 0.0202156 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3743475 | 0.1293962 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3844913 | 0.1231027 |

***Top Ranked Full Model***



### BA + Aldehyde - pi interaction - hydrogen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.16..17. + diff.O17.C18 | 0.7049055 | 0.5944731 | 0.1055027 |
| output ~ Dist.18..19. + diff.C18.H19 | 0.7084203 | 0.5591932 | 0.1079523 |

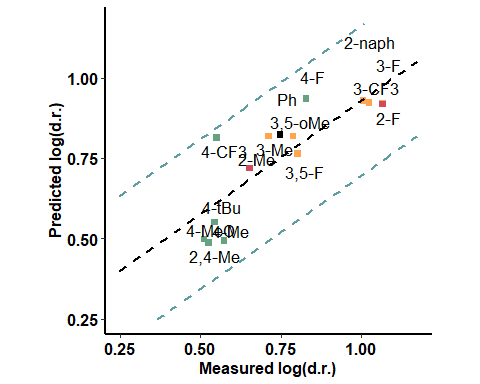
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0308922 | 24.609956 | 0.0000000 |
| Dist.16..17. | 0.1308435 | 0.0330176 | 3.962841 | 0.0018837 |
| diff.O17.C18 | 0.0825206 | 0.0330176 | 2.499290 | 0.0279518 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5279509 | 0.1191171 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.560222 | 0.111728 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - aldehyde side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ para + Dist.16..17. | 0.6512250 | 0.4952162 | 0.1185252 |
| output ~ para + NBO.O.1 | 0.6511833 | 0.4907146 | 0.1242047 |

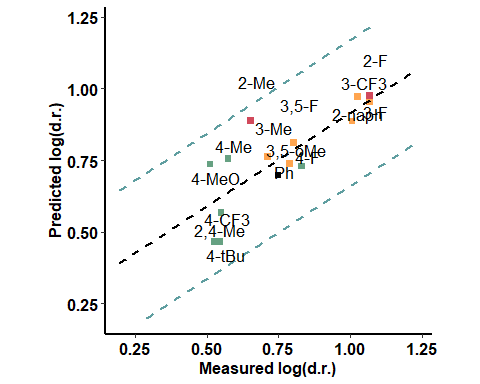
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0335847 | 22.636993 | 0.0000000 |
| para | 0.1172961 | 0.0354264 | 3.310975 | 0.0062132 |
| Dist.16..17. | -0.0950124 | 0.0354264 | -2.681964 | 0.0199643 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4601792 | 0.1292474 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4773838 | 0.1229769 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - opposite to aldehyde

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.2..9. + B5 | 0.5912006 | 0.4547096 | 0.1286433 |
| output ~ diff.B2.O9 + B5 | 0.5512341 | 0.4285776 | 0.1308614 |

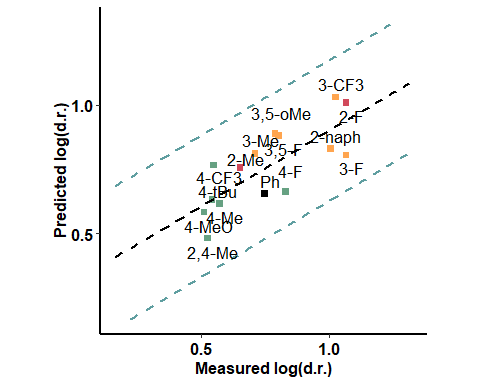
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0363600 | 20.909145 | 0.0000000 |
| Dist.2..9. | -0.1350464 | 0.0377053 | -3.581633 | 0.0037709 |
| B5 | 0.0882433 | 0.0377053 | 2.340343 | 0.0373624 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4214863 | 0.1398078 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4360454 | 0.1335854 |

***Top Ranked Full Model***



### Enamine - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ para + Dist.1..2. | 0.7076306 | 0.5792479 | 0.0948479 |
| output ~ para + diff.O1.B2 | 0.6849136 | 0.5356607 | 0.1031372 |

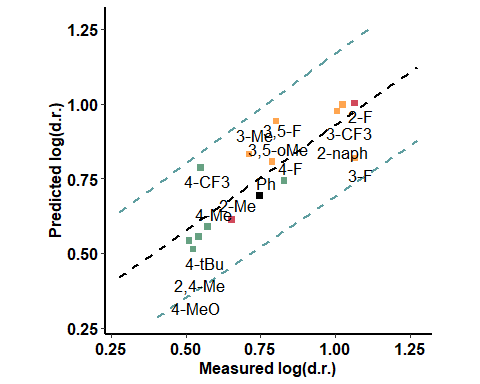
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0307492 | 24.724384 | 0.0000000 |
| para | 0.1146593 | 0.0326403 | 3.512814 | 0.0042797 |
| Dist.1..2. | -0.1046718 | 0.0326403 | -3.206825 | 0.0075365 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5222579 | 0.1065229 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5381343 | 0.1004736 |

***Top Ranked Full Model***



### Enamine - close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + NBO.H.10 | 0.6745607 | 0.4649195 | 0.1246208 |

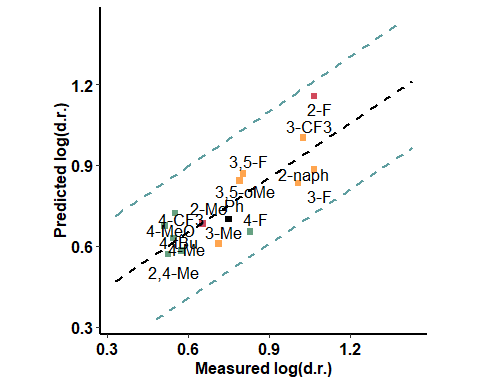
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7602560 | 0.0324417 | 23.434538 | 0.0000000 |
| Total | -0.1252952 | 0.0371527 | -3.372443 | 0.0055451 |
| NBO.H.10 | 0.1769896 | 0.0371527 | 4.763847 | 0.0004611 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4496495 | 0.1333502 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4523634 | 0.1295369 |

***Top Ranked Full Model***



## ACN

### Only Boronic Acids

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.2.3. + Dist.2..9. | 0.8549778 | 0.7987696 | 0.0465792 |

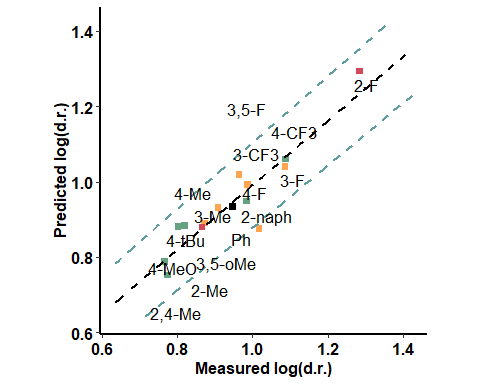
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0148379 | 63.655912 | 0.0000000 |
| X.2.3. | -0.0445812 | 0.0163465 | -2.727256 | 0.0183595 |
| Dist.2..9. | -0.1374663 | 0.0163465 | -8.409507 | 0.0000022 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.7596139 | 0.0545887 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.7792024 | 0.0507687 |

***Top Ranked Full Model***



### With Catalyst - Close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.8.14. + Dist.14..18. | 0.8234273 | 0.7638875 | 0.0531336 |
| output ~ Dist.1..16. + NBO.B.1 | 0.7991561 | 0.7319453 | 0.0601752 |

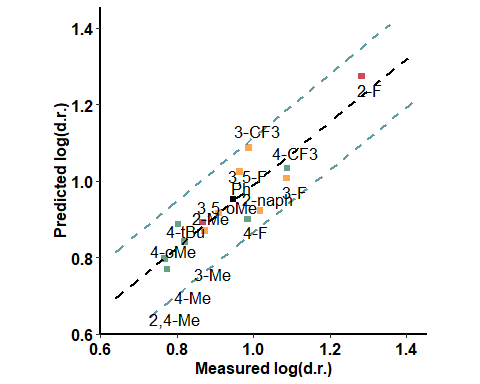
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0163725 | 57.689166 | 0.00e+00 |
| X.8.14. | -0.1047128 | 0.0176703 | -5.925917 | 6.97e-05 |
| Dist.14..18. | 0.1070219 | 0.0176703 | 6.056592 | 5.70e-05 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6678096 | 0.0687734 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.7106756 | 0.0614302 |

***Top Ranked Full Model***



### With Catalyst - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.19..20. + diff.C17.O18 | 0.8684463 | 0.7960916 | 0.0509759 |

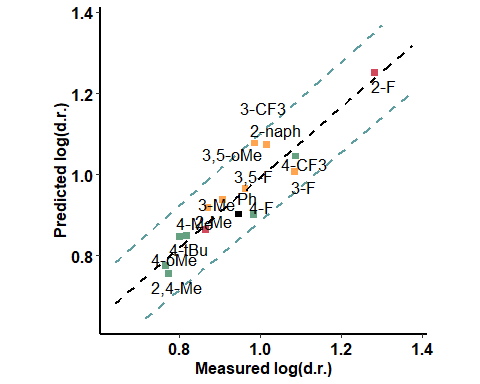
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0141321 | 66.835076 | 0.0000000 |
| Dist.19..20. | 0.1003721 | 0.0202799 | 4.949344 | 0.0003368 |
| diff.C17.O18 | -0.1777287 | 0.0202799 | -8.763794 | 0.0000015 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.7655674 | 0.056806 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.7815421 | 0.0538645 |

***Top Ranked Full Model***



### BA + Pentanone

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + diff.O1.H10 | 0.7983838 | 0.7466983 | 0.0557884 |
| output ~ Dist.1..2. + diff.O9.H16 | 0.8079846 | 0.7417208 | 0.0592688 |

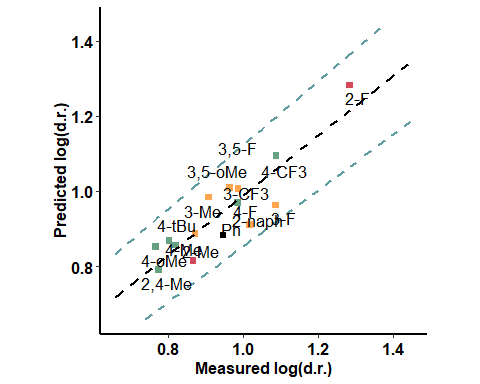
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0174951 | 53.987515 | 0.00e+00 |
| Total | 0.0859449 | 0.0189700 | 4.530573 | 6.89e-04 |
| diff.O1.H10 | -0.1196818 | 0.0189700 | -6.309005 | 3.89e-05 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6509652 | 0.0890807 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6937687 | 0.0732998 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - oxygen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.2..9. + dip\_y | 0.8969871 | 0.8489363 | 0.0436013 |
| output ~ Dist.2..9. + NBO.C.20 | 0.8901088 | 0.8385889 | 0.0372199 |

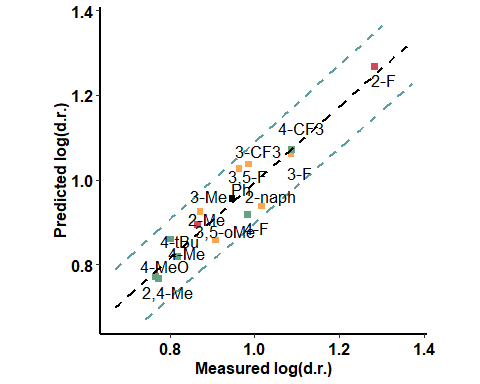
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0125055 | 75.528377 | 0.0000000 |
| Dist.2..9. | -0.1177076 | 0.0131926 | -8.922225 | 0.0000012 |
| dip\_y | -0.0418457 | 0.0131926 | -3.171897 | 0.0080413 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.8321263 | 0.0482246 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.8389818 | 0.0461439 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - hydrogen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.18.19. + NBO.O.1 | 0.9018949 | 0.8684135 | 0.0427732 |
| output ~ NBO.O.1 + NBO.O.17 | 0.9062933 | 0.8552073 | 0.0428522 |

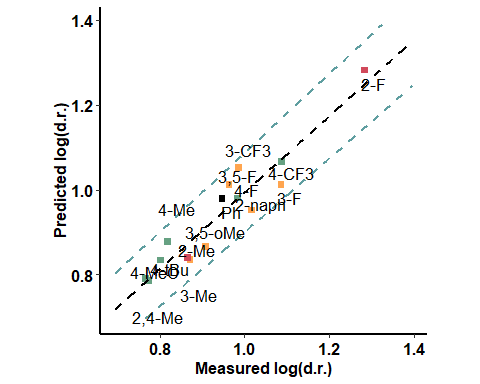
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0122039 | 77.394515 | 0.00e+00 |
| X.18.19. | -0.0729285 | 0.0126645 | -5.758496 | 9.04e-05 |
| NBO.O.1 | 0.1161605 | 0.0126645 | 9.172123 | 9.00e-07 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.8356473 | 0.0496494 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.8527622 | 0.0467147 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - aldehyde side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.9..16. + NBO.O.9 | 0.7822189 | 0.6928175 | 0.0695343 |
| output ~ Dist.2..9. + Dist.9..16. | 0.7637982 | 0.6243849 | 0.0751802 |

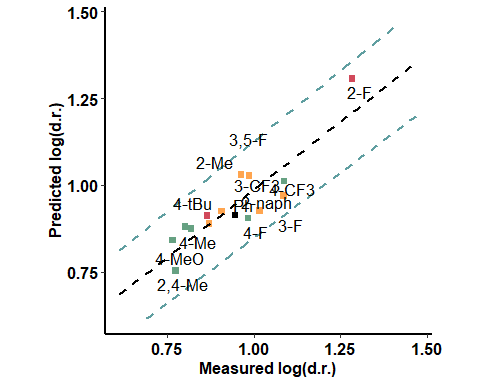
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0181830 | 51.945259 | 0.0000000 |
| Dist.9..16. | 0.2120352 | 0.0322972 | 6.565129 | 0.0000267 |
| NBO.O.9 | 0.1719997 | 0.0322972 | 5.325530 | 0.0001807 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6637378 | 0.07748 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6799558 | 0.0725864 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - opposite to aldehyde

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.9..16. + NBO.O.9 | 0.7599594 | 0.6901046 | 0.0680400 |
| output ~ Dist.2..9. + Dist.9..16. | 0.7636245 | 0.5847323 | 0.0788087 |

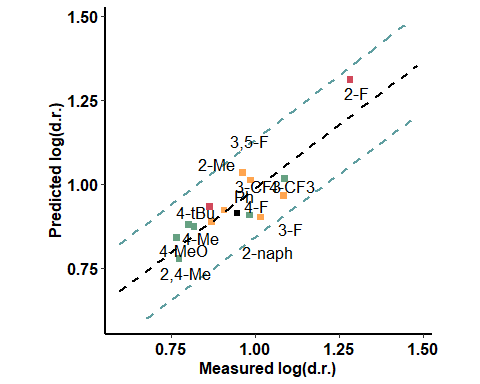
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0190896 | 49.478182 | 0.0000000 |
| Dist.9..16. | 0.0828234 | 0.0236594 | 3.500661 | 0.0043765 |
| NBO.O.9 | 0.1457957 | 0.0236594 | 6.162284 | 0.0000486 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5764615 | 0.0902935 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6204771 | 0.0810264 |

***Top Ranked Full Model***



### Enamine - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.1..2. + B1 | 0.8053431 | 0.7372605 | 0.0570261 |
| output ~ X.1.2. + Dist.1..2. | 0.7866995 | 0.7148957 | 0.0608419 |

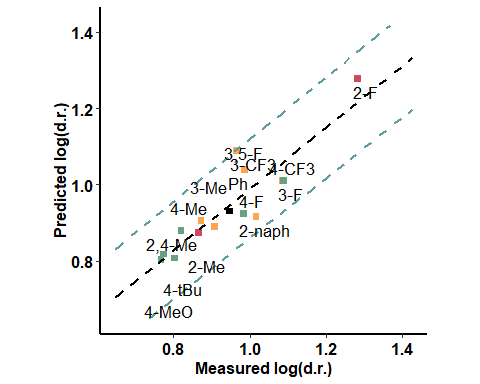
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0171905 | 54.944117 | 0.0000000 |
| Dist.1..2. | -0.1202990 | 0.0179879 | -6.687766 | 0.0000224 |
| B1 | -0.0218504 | 0.0179879 | -1.214726 | 0.2478327 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6115763 | 0.076565 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6610066 | 0.0699747 |

***Top Ranked Full Model***



### Enamine - close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.2..9. + Dist.9..16. | 0.7768762 | 0.6797322 | 0.067488 |
| output ~ Dist.2..9. + Dist.16..36. | 0.7704016 | 0.6700345 | 0.068579 |

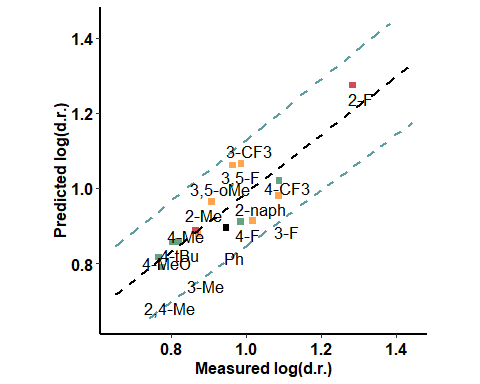
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9445187 | 0.0184046 | 51.319587 | 0.0000000 |
| Dist.2..9. | -0.1528354 | 0.0288622 | -5.295346 | 0.0001898 |
| Dist.9..16. | -0.0441940 | 0.0288622 | -1.531205 | 0.1516461 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6242363 | 0.0764223 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6471299 | 0.0723483 |

***Top Ranked Full Model***



## Methanol

### Only Boronic Acids

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.2..3. + L | 0.4873112 | 0.2284430 | 0.1113156 |
| output ~ X.9.16. + Dist.1..10. | 0.4564576 | 0.1695005 | 0.1092643 |

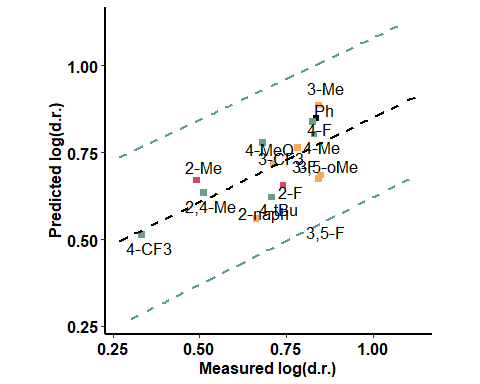
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0309442 | 22.932499 | 0.0000000 |
| Dist.2..3. | -0.1209816 | 0.0399098 | -3.031378 | 0.0104398 |
| L | -0.1198623 | 0.0399098 | -3.003331 | 0.0109984 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.226258 | 0.1182279 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2285191 | 0.1142997 |

***Top Ranked Full Model***



### With Catalyst - Close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.14..18. + dip\_y | 0.6566038 | 0.383088 | 0.0995996 |

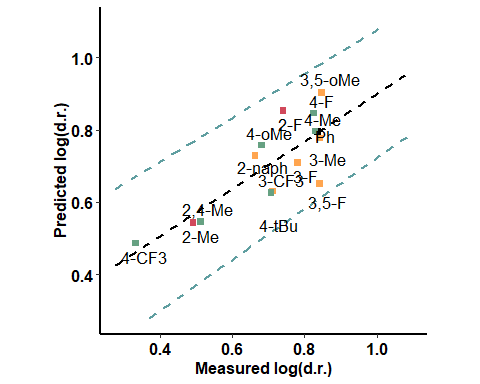
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0253250 | 28.020802 | 0.0000000 |
| Dist.14..18. | 0.1154762 | 0.0278342 | 4.148713 | 0.0013497 |
| dip\_y | 0.1015918 | 0.0278342 | 3.649887 | 0.0033272 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3368284 | 0.1074321 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3563787 | 0.1021568 |

***Top Ranked Full Model***



### With Catalyst - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.8..14. + L | 0.6877579 | 0.5648867 | 0.068771 |
| output ~ X.1.8. + dip\_y | 0.6851683 | 0.4564739 | 0.087374 |

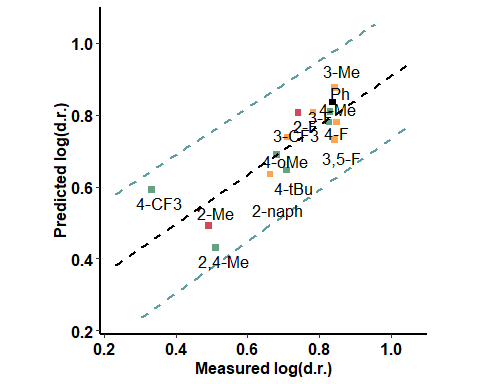
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0241489 | 29.385465 | 0.0000000 |
| Dist.8..14. | -0.1296404 | 0.0271915 | -4.767682 | 0.0004581 |
| L | -0.0991157 | 0.0271915 | -3.645101 | 0.0033565 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5017991 | 0.0889416 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5300659 | 0.0817515 |

***Top Ranked Full Model***



### BA + Pentanone

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ cross + Dist.9..16. | 0.5703362 | 0.3409087 | 0.1008972 |
| output ~ Dist.2..3. + L | 0.5713759 | 0.3257863 | 0.1033847 |

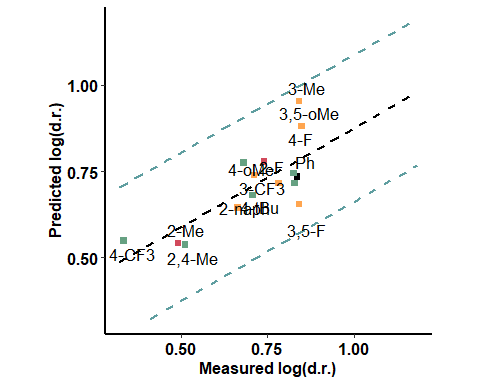
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0283280 | 25.050356 | 0.0000000 |
| cross | 0.0944118 | 0.0304279 | 3.102801 | 0.0091423 |
| Dist.9..16. | -0.0988260 | 0.0304279 | -3.247869 | 0.0069840 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3166677 | 0.1113414 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3214076 | 0.1058011 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - oxygen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.9..16. + L | 0.6620972 | 0.4737708 | 0.0906331 |

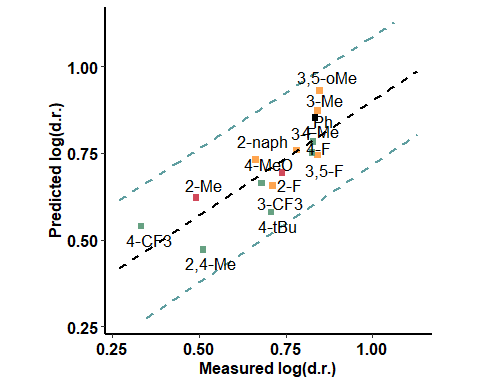
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0251216 | 28.247655 | 0.0000000 |
| Dist.9..16. | -0.1262061 | 0.0282996 | -4.459648 | 0.0007796 |
| L | -0.0993057 | 0.0282996 | -3.509089 | 0.0043091 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4402504 | 0.0984614 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4495994 | 0.0948711 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - hydrogen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.1..10. + L | 0.7248318 | 0.568937 | 0.0707027 |
| output ~ X.18.20. + L | 0.6504974 | 0.442189 | 0.0874867 |

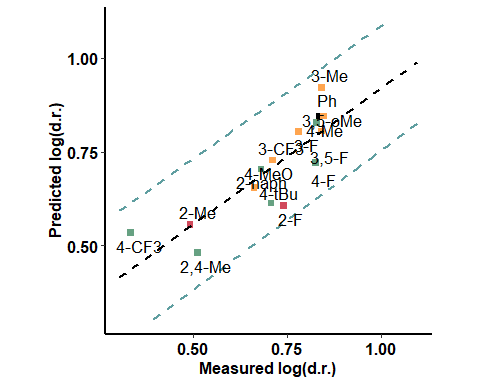
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0226700 | 31.302511 | 0.0000000 |
| Dist.1..10. | -0.1402467 | 0.0268696 | -5.219523 | 0.0002149 |
| L | -0.1173549 | 0.0268696 | -4.367568 | 0.0009161 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5269259 | 0.0809355 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5409732 | 0.0764907 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - aldehyde side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.18.20. + NBO.H.19 | 0.6990670 | 0.4779786 | 0.0789922 |
| output ~ X.18.20. + Dist.18..19. | 0.6858546 | 0.4635106 | 0.0781010 |

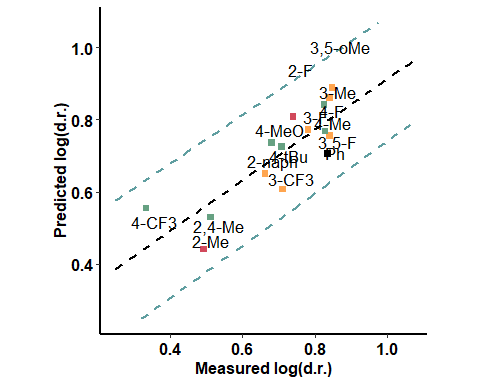
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0237076 | 29.932524 | 0.0000000 |
| X.18.20. | -0.2182619 | 0.0441527 | -4.943340 | 0.0003402 |
| NBO.H.19 | -0.2269562 | 0.0441527 | -5.140253 | 0.0002449 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4545821 | 0.0851356 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4630493 | 0.0819673 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - opposite to aldehyde

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.1.10. + dip\_x | 0.6561442 | 0.5032629 | 0.0819531 |
| output ~ X.18.20. + diff.H16.O17 | 0.6885971 | 0.4593459 | 0.0960803 |

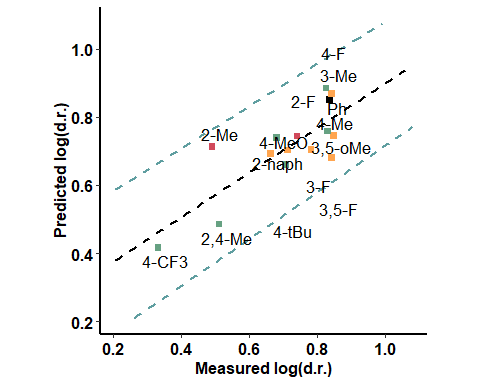
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0253420 | 28.002068 | 0.0000000 |
| X.1.10. | 0.1044065 | 0.0265931 | 3.926069 | 0.0020129 |
| dip\_x | -0.0889245 | 0.0265931 | -3.343888 | 0.0058459 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4488384 | 0.2491514 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4820735 | 0.110239 |

***Top Ranked Full Model***



### Enamine - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.1..10. + L | 0.5981772 | 0.4639326 | 0.0785754 |

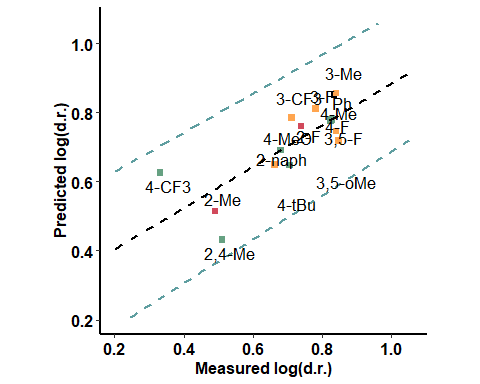
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0273949 | 25.903652 | 0.0000000 |
| Dist.1..10. | -0.1178298 | 0.0303279 | -3.885193 | 0.0021674 |
| L | -0.0889777 | 0.0303279 | -2.933855 | 0.0125140 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4033939 | 0.1017179 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4317359 | 0.0887341 |

***Top Ranked Full Model***



### Enamine - close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.2.9. + L | 0.5227355 | 0.3755615 | 0.0842177 |
| output ~ Dist.2..3. + L | 0.4794867 | 0.2163765 | 0.1138705 |

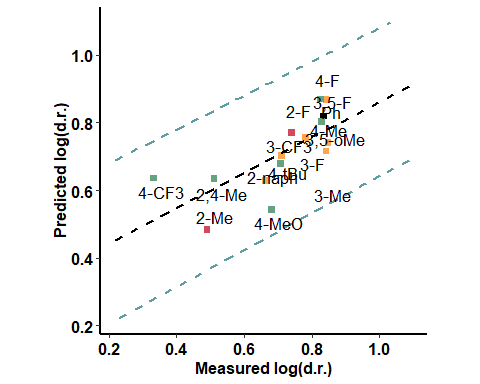
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7096271 | 0.0298560 | 23.768334 | 0.0000000 |
| X.2.9. | 0.1099819 | 0.0334978 | 3.283257 | 0.0065405 |
| L | -0.0899466 | 0.0334978 | -2.685150 | 0.0198471 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.339072 | 0.1005616 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3494414 | 0.0923993 |

***Top Ranked Full Model***



## Chloroform

### Only Boronic Acids

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.1.2. + diff.O1.H10 | 0.7063467 | 0.5249278 | 0.0997979 |
| output ~ X.2.9. + Dist.2..9. | 0.6799762 | 0.5177720 | 0.1004984 |

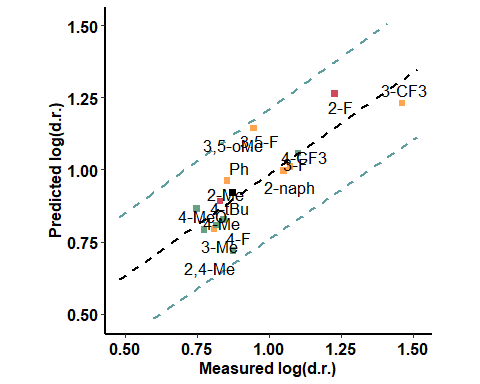
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0296959 | 32.047999 | 0.0000000 |
| X.1.2. | 0.1387619 | 0.0308544 | 4.497313 | 0.0007301 |
| diff.O1.H10 | 0.0783109 | 0.0308544 | 2.538079 | 0.0260300 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4856652 | 0.1135153 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5046355 | 0.1064469 |

***Top Ranked Full Model***



### With Catalyst - Close

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.19.20. + Dist.8..14. | 0.8365717 | 0.7100281 | 0.0737639 |
| output ~ Dist.8..14. + dip\_z | 0.8301853 | 0.6769120 | 0.0928861 |

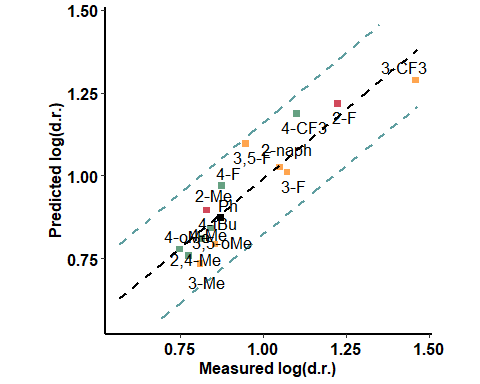
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0221535 | 42.959056 | 0.0000000 |
| X.19.20. | 0.0668236 | 0.0229337 | 2.913769 | 0.0129898 |
| Dist.8..14. | 0.1658221 | 0.0229337 | 7.230493 | 0.0000104 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6881715 | 0.082144 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6948214 | 0.0788665 |

***Top Ranked Full Model***



### With Catalyst - far

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + NBO.O.16 | 0.9105786 | 0.8313782 | 0.0600844 |
| output ~ Total + L | 0.8567410 | 0.7745529 | 0.0776621 |

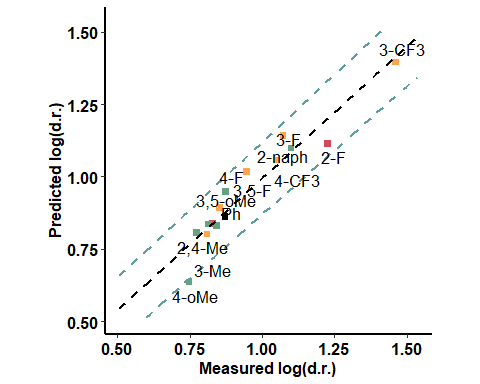
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0163870 | 58.076115 | 0.0000000 |
| Total | 0.1614673 | 0.0177496 | 9.096969 | 0.0000010 |
| NBO.O.16 | -0.0589669 | 0.0177496 | -3.322162 | 0.0060858 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.7533024 | 0.0723211 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.7807192 | 0.0673714 |

***Top Ranked Full Model***



### BA + Pentanone

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ NBO.H.16 + NBO.C.18 | 0.7881859 | 0.5923135 | 0.0992653 |

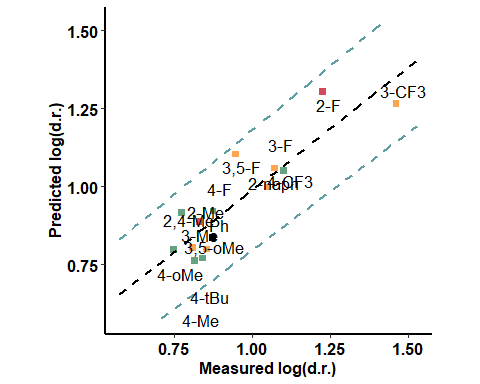
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0252206 | 37.734700 | 0.0000000 |
| NBO.H.16 | 0.6243529 | 0.1445792 | 4.318414 | 0.0009989 |
| NBO.C.18 | -0.4809641 | 0.1445792 | -3.326647 | 0.0060354 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.548231 | 0.1075869 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5743685 | 0.1026113 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - oxygen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.18.19. + Dist.2..9. | 0.8932486 | 0.4999122 | 2.042431 |

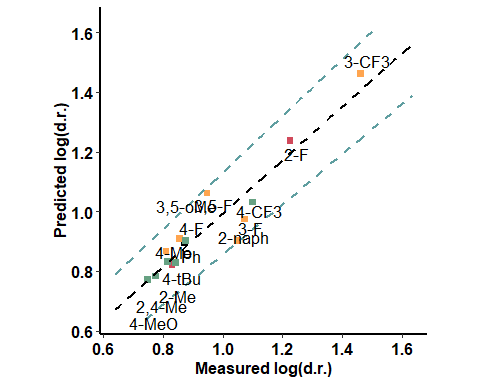
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0179046 | 53.153471 | 0.0000000 |
| X.18.19. | 0.1040733 | 0.0193622 | 5.375069 | 0.0001667 |
| Dist.2..9. | -0.1266026 | 0.0193622 | -6.538638 | 0.0000278 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4993602 | 1.950007 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4966702 | 1.949219 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - hydrogen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.18.19. + Dist.1..2. | 0.6947638 | 0.5417311 | 0.0872277 |
| output ~ Dist.1..2. + Dist.17..18. | 0.6868608 | 0.5405861 | 0.0784905 |

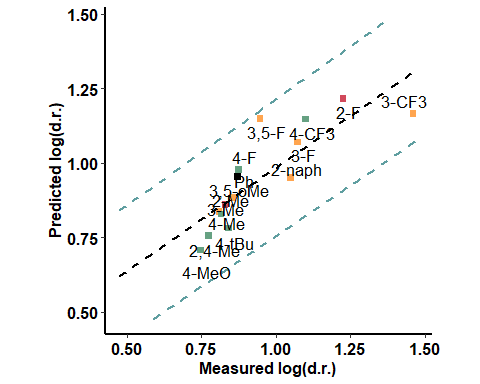
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0302759 | 31.434051 | 0.0000000 |
| X.18.19. | -0.0609723 | 0.0327582 | -1.861283 | 0.0873696 |
| Dist.1..2. | -0.1352892 | 0.0327582 | -4.129932 | 0.0013957 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4853537 | 0.1071252 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5054695 | 0.0984294 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - aldehyde side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.16..17. + NBO.H.10 | 0.8075570 | 0.6547173 | 0.0894582 |
| output ~ Dist.10..17. + NBO.H.10 | 0.7608458 | 0.6280473 | 0.0919192 |

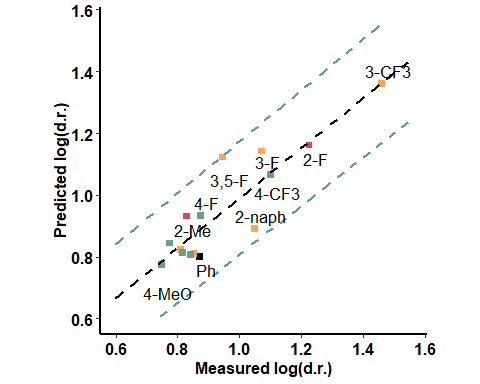
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0240397 | 39.588343 | 0.0000000 |
| Dist.16..17. | -0.2823593 | 0.0397995 | -7.094542 | 0.0000126 |
| NBO.H.10 | 0.2165431 | 0.0397995 | 5.440847 | 0.0001499 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6010178 | 0.1003702 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6290117 | 0.0946882 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - opposite to aldehyde

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.2..3. + diff.C18.H19 | 0.7284930 | 0.5572243 | 0.0966339 |
| output ~ NBO.C.20 + diff.C18.C20 | 0.7187271 | 0.5269840 | 0.1001709 |

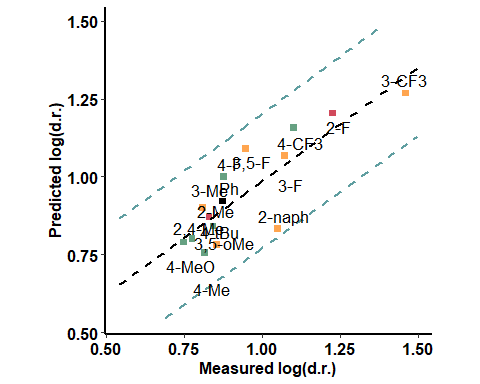
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0285542 | 33.329428 | 0.0000000 |
| Dist.2..3. | 0.1311625 | 0.0302822 | 4.331339 | 0.0009764 |
| diff.C18.H19 | -0.1368899 | 0.0302822 | -4.520473 | 0.0007012 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5049053 | 0.1091854 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.528301 | 0.1023787 |

***Top Ranked Full Model***



### Enamine - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + B5 | 0.8007607 | 0.7127119 | 0.0857386 |
| output ~ para + Total | 0.7954438 | 0.6829290 | 0.0834981 |

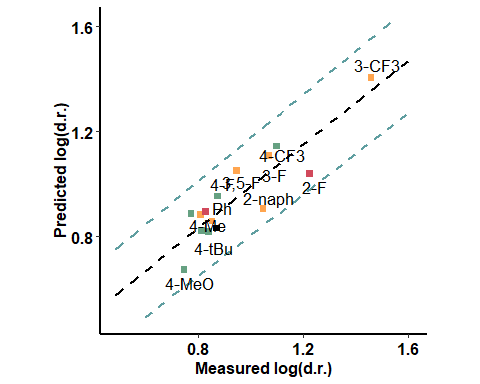
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0244606 | 38.907275 | 0.0000000 |
| Total | 0.1692428 | 0.0253730 | 6.670194 | 0.0000229 |
| B5 | 0.0599766 | 0.0253730 | 2.363797 | 0.0358032 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6765567 | 0.0938786 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6941914 | 0.0893493 |

***Top Ranked Full Model***



### Enamine - close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + NBO.H.10 | 0.759132 | 0.5977506 | 0.0830164 |

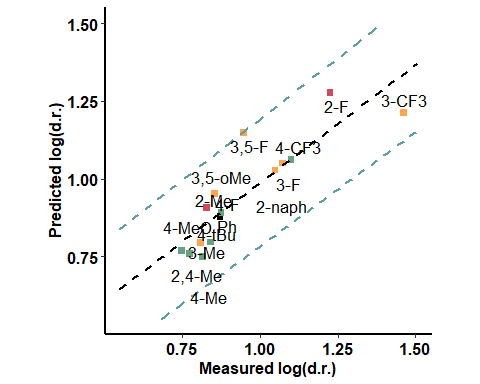
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.9516937 | 0.0268948 | 35.385782 | 0.0000000 |
| Total | -0.0676258 | 0.0308003 | -2.195621 | 0.0485164 |
| NBO.H.10 | 0.1888533 | 0.0308003 | 6.131542 | 0.0000509 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5304826 | 0.100116 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.562416 | 0.0914321 |

***Top Ranked Full Model***



## Hexane

### Only Boronic Acids

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.2.9. + B5 | 0.7807392 | 0.6338624 | 0.0840461 |

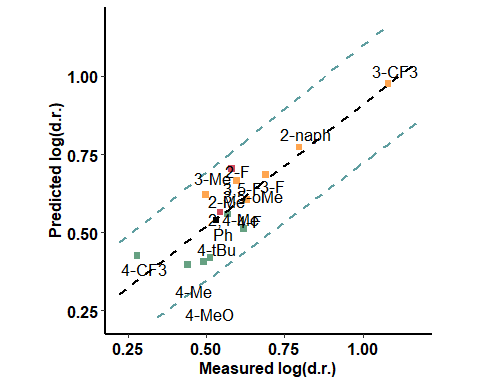
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0233012 | 25.329975 | 0.0000000 |
| X.2.9. | 0.1257364 | 0.0251956 | 4.990413 | 0.0003143 |
| B5 | 0.1381919 | 0.0251956 | 5.484766 | 0.0001396 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5817496 | 0.0922794 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6050105 | 0.0888198 |

***Top Ranked Full Model***



### With Catalyst - Close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + diff.B1.O16 | 0.6696 | 0.332226 | 0.1168222 |

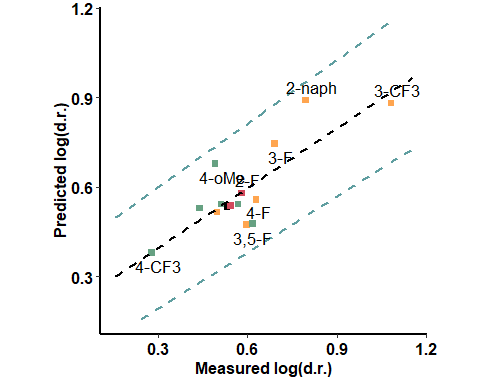
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0286034 | 20.634563 | 0.0000000 |
| Total | -0.1342640 | 0.0309348 | -4.340225 | 0.0009612 |
| diff.B1.O16 | 0.1082315 | 0.0309348 | 3.498699 | 0.0043924 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3319761 | 0.1312353 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3351846 | 0.1248125 |

***Top Ranked Full Model***



### With Catalyst - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Dist.19..20. + B5 | 0.6112505 | 0.3108150 | 0.1124397 |
| output ~ Dist.17..18. + B5 | 0.5824655 | 0.2988876 | 0.1132490 |

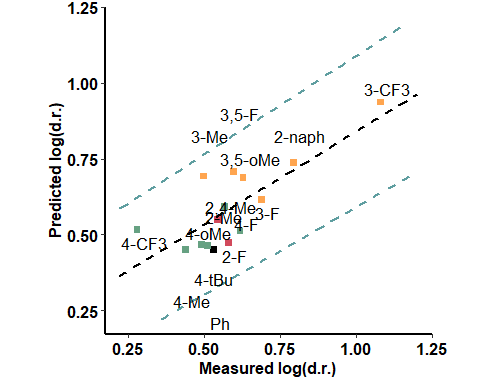
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0310265 | 19.023059 | 0.0000000 |
| Dist.19..20. | -0.0812907 | 0.0337947 | -2.405424 | 0.0331887 |
| B5 | 0.1414656 | 0.0337947 | 4.186025 | 0.0012629 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2852269 | 0.1274527 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.290438 | 0.1165023 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - oxygen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.18.19. + X.18.20. | 0.7054320 | 0.5783839 | 3.052036 |
| output ~ X.1.10. + X.18.19. | 0.7179386 | 0.5745548 | 1.596017 |

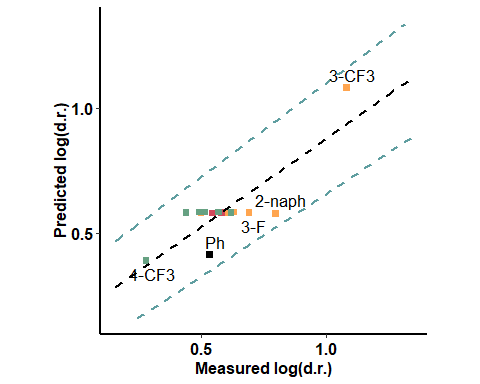
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0270079 | 21.853575 | 0.0000000 |
| X.18.19. | 0.1285421 | 0.0281339 | 4.568938 | 0.0006447 |
| X.18.20. | 0.0639465 | 0.0281339 | 2.272932 | 0.0422139 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5627528 | 3.186934 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.5658159 | 2.950431 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - hydrogen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ diff.B2.O9 + B5 | 0.5484099 | 0.2333914 | 0.1209806 |
| output ~ dip\_x + B5 | 0.4869764 | 0.2012582 | 0.1242570 |

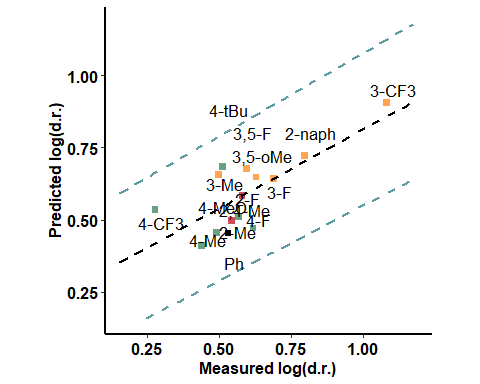
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0334402 | 17.649933 | 0.0000000 |
| diff.B2.O9 | -0.1433809 | 0.0431471 | -3.323070 | 0.0060755 |
| B5 | 0.1506338 | 0.0431471 | 3.491169 | 0.0044537 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2232037 | 0.131803 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2235731 | 0.1267003 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - aldehyde side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.18.19. + B5 | 0.5642604 | 0.2868490 | 0.1194131 |
| output ~ X.2.9. + B5 | 0.5633176 | 0.2811037 | 0.1177249 |

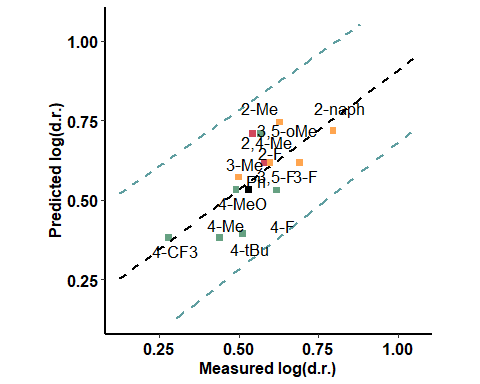
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0328481 | 17.968084 | 0.0000000 |
| X.18.19. | 0.0687508 | 0.0358645 | 1.916960 | 0.0793580 |
| B5 | 0.0952443 | 0.0358645 | 2.655673 | 0.0209583 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2744458 | 0.1700764 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2787109 | 0.1400443 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - opposite to aldehyde

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ diff.C18.H19 + B5 | 0.6759283 | 0.4534682 | 0.0958661 |

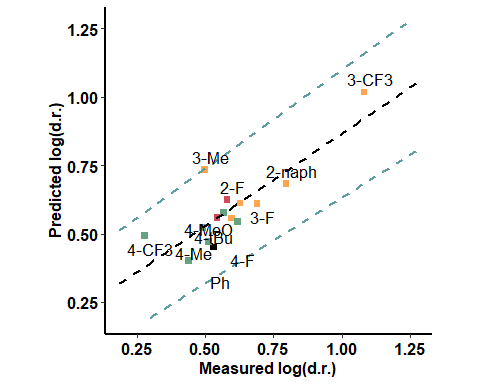
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0283281 | 20.835060 | 0.0000000 |
| diff.C18.H19 | -0.0925083 | 0.0309449 | -2.989449 | 0.0112858 |
| B5 | 0.1471882 | 0.0309449 | 4.756455 | 0.0004669 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3809066 | 0.1061728 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4135812 | 0.0999656 |

***Top Ranked Full Model***



### Enamine - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ dip\_z + B5 | 0.6124885 | 0.3204784 | 0.1230525 |
| output ~ X.1.2. + B5 | 0.5737977 | 0.2919757 | 0.1075037 |

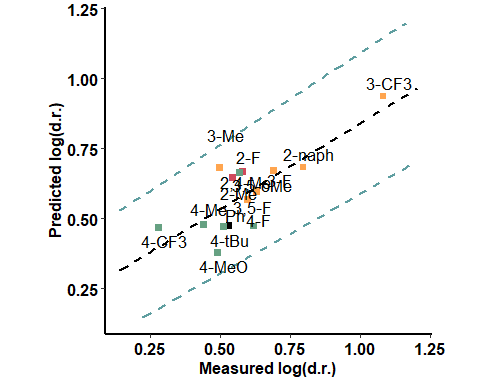
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0309770 | 19.053423 | 0.0000000 |
| dip\_z | -0.0968785 | 0.0340775 | -2.842888 | 0.0148171 |
| B5 | 0.1385923 | 0.0340775 | 4.066974 | 0.0015621 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2832877 | 0.1351236 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2966598 | 0.1266062 |

***Top Ranked Full Model***



### Enamine - close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.1.2. + B5 | 0.5322447 | 0.2874310 | 0.1168694 |
| output ~ para + X.2.3. | 0.4783236 | 0.2051382 | 0.1226364 |

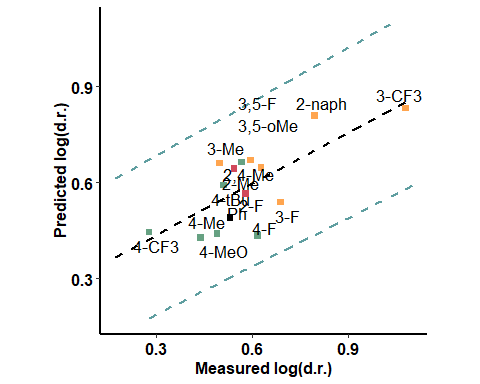
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5902181 | 0.0340335 | 17.342271 | 0.0000000 |
| X.1.2. | 0.1041529 | 0.0453229 | 2.298020 | 0.0403419 |
| B5 | 0.1674694 | 0.0453229 | 3.695031 | 0.0030635 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2306424 | 0.1382326 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2471289 | 0.1261854 |

***Top Ranked Full Model***



## Hexane - against ee

### Only Boronic Acids

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.2.9. + B5 | 0.6619781 | 0.3685013 | 0.1744393 |

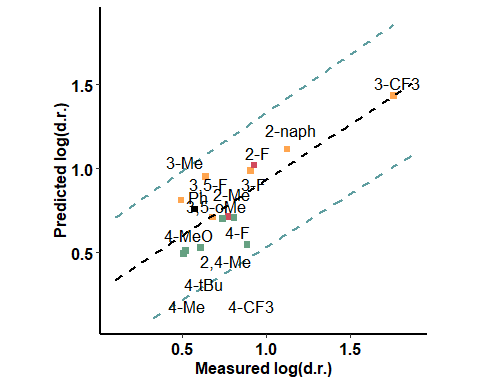
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0520187 | 15.37785 | 0.0000000 |
| X.2.9. | 0.2338040 | 0.0562479 | 4.15667 | 0.0013307 |
| B5 | 0.2019317 | 0.0562479 | 3.59003 | 0.0037132 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3230509 | 0.2046499 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3365788 | 0.1887413 |

***Top Ranked Full Model***



### With Catalyst - Close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + NBO.O.18 | 0.7249018 | 0.3737202 | 0.1771361 |
| output ~ Total + diff.H14.O18 | 0.6997620 | 0.3502341 | 0.1963199 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0469279 | 17.046071 | 0.0000000 |
| Total | -0.2096349 | 0.0504998 | -4.151205 | 0.0013437 |
| NBO.O.18 | 0.2415795 | 0.0504998 | 4.783773 | 0.0004457 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3665489 | 0.1908736 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3675552 | 0.1859561 |

***Top Ranked Full Model***

### With Catalyst - far

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + diff.O16.C17 | 0.7672782 | 0.3967453 | 0.1729457 |
| output ~ Total + NBO.O.16 | 0.7653131 | 0.3063869 | 0.1942524 |

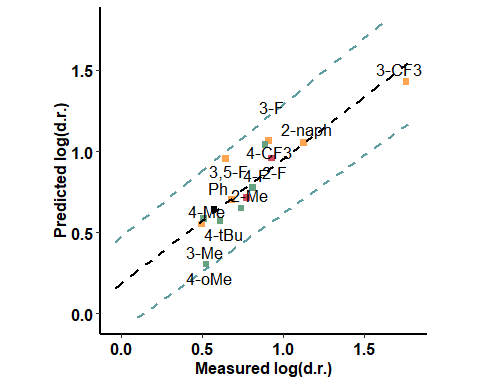
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0431624 | 18.533167 | 0.0000000 |
| Total | 0.2203704 | 0.0463757 | 4.751856 | 0.0004706 |
| diff.O16.C17 | 0.1250300 | 0.0463757 | 2.696026 | 0.0194518 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3853301 | 0.181412 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3900278 | 0.1766446 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - oxygen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ X.18.19. + diff.H10.O17 | 0.8017001 | 0.6902510 | 1.671395 |
| output ~ X.18.19. + NBO.H.19 | 0.7782123 | 0.6867141 | 2.019213 |

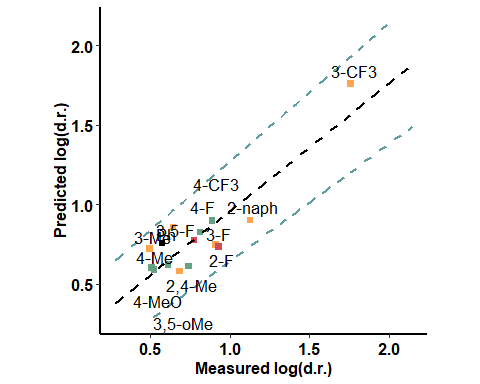
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0398427 | 20.077373 | 0.0000000 |
| X.18.19. | 0.2606542 | 0.0412709 | 6.315686 | 0.0000386 |
| diff.H10.O17 | 0.1112211 | 0.0412709 | 2.694903 | 0.0194923 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.6778579 | 5.92548 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.674292 | 4.498462 |

***Top Ranked Full Model***



### BA + Aldehyde (pi interaction) - hydrogen side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + diff.H10.O17 | 0.5156322 | 0.2522363 | 0.1940352 |
| output ~ Dist.17..18. + B5 | 0.4771292 | 0.1989185 | 0.2259979 |

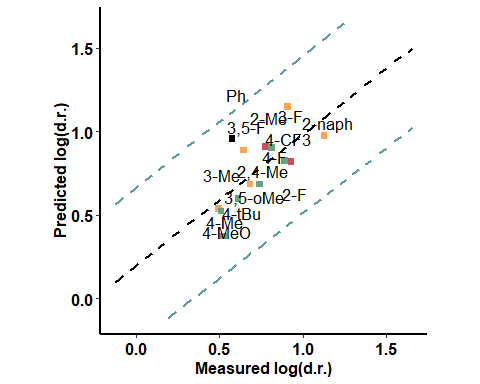
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0622694 | 12.846372 | 0.0000000 |
| Total | -0.2264213 | 0.1100461 | -2.057513 | 0.0620429 |
| diff.H10.O17 | 0.3718911 | 0.1100461 | 3.379413 | 0.0054741 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2279574 | 0.2171824 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2428908 | 0.2035005 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - aldehyde side

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + B5 | 0.6490579 | 0.3421465 | 0.1853301 |
| output ~ Dist.1..10. + Dist.16..17. | 0.6098284 | 0.2705932 | 0.1833607 |

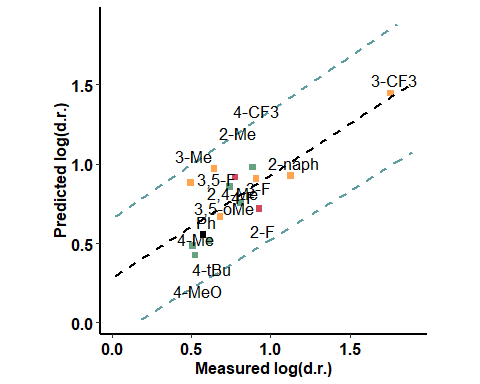
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0530036 | 15.092126 | 0.0000000 |
| Total | 0.1951871 | 0.0550626 | 3.544821 | 0.0040349 |
| B5 | 0.1868034 | 0.0550626 | 3.392564 | 0.0053426 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.317911 | 0.2047416 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3352989 | 0.1935397 |

***Top Ranked Full Model***



### BA + Aldehyde - H bond - opposite to aldehyde

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ diff.C18.H19 + B5 | 0.6637662 | 0.3151564 | 0.1711994 |

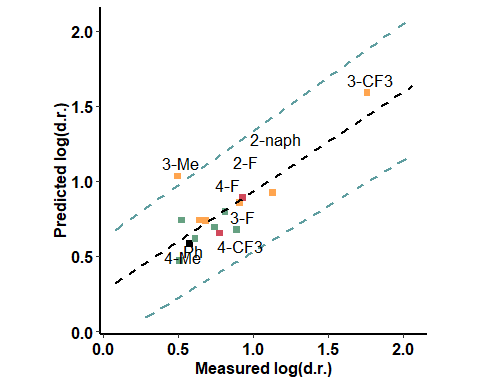
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0518810 | 15.418689 | 0.0000000 |
| diff.C18.H19 | -0.2080839 | 0.0566735 | -3.671628 | 0.0031974 |
| B5 | 0.2380773 | 0.0566735 | 4.200859 | 0.0012300 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.3125577 | 0.1907002 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2890364 | 0.1823723 |

***Top Ranked Full Model***



### Enamine - far from the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ Total + B5 | 0.7294911 | 0.4561942 | 0.1714146 |

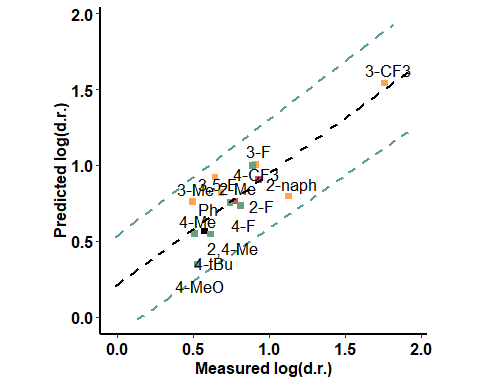
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0465348 | 17.190058 | 0.0000000 |
| Total | 0.2356612 | 0.0482707 | 4.882076 | 0.0003772 |
| B5 | 0.1560044 | 0.0482707 | 3.231865 | 0.0071944 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4344219 | 0.1819997 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.4488754 | 0.1761558 |

***Top Ranked Full Model***



### Enamine - close to the active site

|  |  |  |  |
| --- | --- | --- | --- |
| formula | R.sq | Q.sq | MAE |
| output ~ NBO.C.16 + diff.N24.C25 | 0.5837550 | 0.2885550 | 0.1949026 |
| output ~ NBO.C.16 + NBO.C.25 | 0.5748357 | 0.1919583 | 0.2045477 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7999364 | 0.0577247 | 13.857779 | 0.0000000 |
| NBO.C.16 | 0.3436096 | 0.0889212 | 3.864204 | 0.0022514 |
| diff.N24.C25 | -0.3367733 | 0.0889212 | -3.787325 | 0.0025888 |

***3 & 5 fold CV***

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2489813 | 0.2307942 |

|  |  |
| --- | --- |
| Q2 | MAE |
| 0.2694418 | 0.2065041 |

***Top Ranked Full Model***

