**Predictive Quantitative Structure Toxicity Relationship Study on Avian Toxicity of Some Diverse Agrochemical Pesticides by Monte Carlo Method: QTR on Pesticides**

*Amit K. Halder, Natural Science Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India*

*Achintya Saha, Department of Chemical Technology, University of Calcutta, Kolkata, India*

*Tarun Jha, Natural Science Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India*

*Table S1. Modeling dataset used for QSTR model development along with their distributions in three training/test set combinations (Split 1-3)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Cpd** | **Name of pesticides** | **pLD50 (mmol)** | **Split1** | **Split2** | **Split3** |
| **C1** | (S)-Dimethenamid | -0.588 | Train | Train | Test |
| **C2** | 2-(Hydroxymethyl)amino)ethanol | -1.282 | Train | Train | Train |
| **C3** | 2,4-D Isopropyl Ester | -0.854 | Train | Test | Train |
| **C4** | 3,5 Dimethyl-1-(hydroxymethyl)pyrazole | -0.779 | Train | Train | Test |
| **C5** | 3-Iodo-2-propynyl butyl carbamate (IPBC) | -0.426 | Train | Train | Train |
| **C6** | 4,5-Dichloro-1,2-dithio-3-one | -0.121 | Train | Test | Train |
| **C7** | 4-Aminopyridine | 0.798 | Train | Train | Test |
| **C8** | 4-Chloro-3-cresol | -1.033 | Train | Train | Train |
| **C9** | Acetochlor | 0.741 | Train | Train | Train |
| **C10** | Alachlor | -0.755 | Train | Test | Test |
| **C11** | Aldicarb | 2.325 | Train | Train | Train |
| **C12** | Ametryn | -0.996 | Train | Test | Test |
| **C13** | Amitraz degradate | 0.616 | Train | Train | Train |
| **C14** | Antimycin A | 1.148 | Train | Test | Train |
| **C15** | Azinphos-methyl | 0.983 | Train | Train | Test |
| **C16** | BCDM Hydantoin, 1-bromo-3-chloro-5,5-dimethyl | -0.647 | Train | Train | Train |
| **C17** | Bis (3-aminopropyl)dodecylamine) Lonzabac 12 | 1.079 | Train | Train | Train |
| **C18** | Bromacil | -0.935 | Train | Train | Train |
| **C19** | Bromethalin | 1.719 | Train | Train | Train |
| **C20** | Carbofuran | 1.642 | Train | Train | Train |
| **C21** | Chlorethoxyfos | 1.079 | Train | Train | Test |
| **C22** | Chlorfenapyr AC 303,268 degradate | 1.218 | Train | Train | Train |
| **C23** | Chlorobenzilate (canceled in U.S.) | -0.271 | Train | Test | Train |
| **C24** | Chloroprop acid (3-CPA) | -0.817 | Train | Train | Train |
| **C25** | Chlorpyrifos oxon | 1.58 | Train | Train | Test |
| **C26** | Clodinafop-propargyl | -0.619 | Train | Train | Train |
| **C27** | Clomazone | -1.02 | Train | Train | Train |
| **C28** | Coumaphos | 2.187 | Train | Train | Test |
| **C29** | Cyproconazole | 0.289 | Train | Test | Train |
| **C30** | Dicamba | 0.01 | Train | Train | Test |
| **C31** | Dichlorprop P (2,4-DP-p) | -0.178 | Train | Train | Train |
| **C32** | Dichlorvos | 1.4 | Train | Train | Train |
| **C33** | Dienochlor | -0.172 | Train | Train | Train |
| **C34** | Dimethenamid | -0.84 | Train | Train | Train |
| **C35** | Dipropyl isocinchomeronate MGK 326 repellent | -0.73 | Train | Train | Train |
| **C36** | Disulfoton | 1.359 | Train | Train | Test |
| **C37** | Endosulfan sulfate metabolite(PC Code 079404) | 0.983 | Train | Train | Train |
| **C38** | Endothall | -0.429 | Train | Train | Train |
| **C39** | Esfenvalerate | 0.042 | Train | Train | Test |
| **C40** | Fenamiphos | 2.278 | Train | Test | Train |
| **C41** | Fenitrothion | 1.005 | Train | Train | Train |
| **C42** | Fenthion | 1.593 | Train | Test | Train |
| **C43** | Fluchloralin (uses canceled in U.S.) | -1.294 | Train | Train | Test |
| **C44** | Folpet | 0.732 | Train | Train | Train |
| **C45** | Fosthiazate | 1.427 | Train | Test | Train |
| **C46** | Hallcomid | -0.905 | Train | Train | Train |
| **C47** | Hydramethylnon | -0.568 | Train | Train | Train |
| **C48** | Indoxacarb IN-JT333-20 Degradate | -0.486 | Train | Train | Train |
| **C49** | MCPA acid | -0.274 | Train | Test | Train |
| **C50** | Mefenoxam | -0.546 | Train | Train | Test |
| **C51** | Methiocarb | 1.061 | Train | Train | Train |
| **C52** | Methyl parathion | 1.542 | Train | Test | Test |
| **C53** | N6-Benzyladenine | -0.851 | Train | Train | Train |
| **C54** | Omethoate (Dimethoate Degradate) | 1.346 | Train | Train | Train |
| **C55** | Oryzalin | -0.165 | Train | Test | Train |
| **C56** | Oxythioquinox | 0.078 | Train | Train | Test |
| **C57** | Paradichlorobenzene | -1.039 | Train | Train | Train |
| **C58** | Paranitrophenol | -0.618 | Train | Test | Test |
| **C59** | Pirimiphos-methyl | 0.883 | Train | Train | Train |
| **C60** | Prallethrin | -0.591 | Train | Train | Train |
| **C61** | Propachlor | 0.381 | Train | Train | Train |
| **C62** | Propiconazole | -0.917 | Train | Train | Test |
| **C63** | Propoxur Baygon 2 Bait formulation | -0.682 | Train | Test | Train |
| **C64** | Spiroxamine | -0.27 | Train | Train | Train |
| **C65** | Thiodicarb | -0.756 | Train | Train | Train |
| **C66** | Tolfenpyrad | 0.645 | Train | Train | Train |
| **C67** | Tralomethrin | -0.678 | Train | Train | Train |
| **C68** | Trichlorfon | 1.06 | Train | Train | Train |
| **C69** | Trichloro-s-triazinetrione | -0.858 | Train | Train | Test |
| **C70** | Triclopyr butoxyethyl ester (Garlon 4) | -0.377 | Train | Test | Train |
| **C71** | Trimethacarb | -0.09 | Train | Train | Train |
| **C72** | Uniconazole | -0.7 | Train | Train | Train |
| **C73** | Fluazinam | -0.583 | Train | Test | Test |
| **C74** | Diazinon oxon (metabolite of diazinon) | 1.761 | Train | Train | Train |
| **C75** | Dimethoxane | -0.959 | Train | Train | Train |
| **C76** | Dinoseb acid (Cancelled in U.S.) | 0.779 | Train | Test | Train |
| **C77** | Isofenphos | 1.599 | Train | Train | Train |
| **C78** | Monocrotophos (Cancelled in U.S.) | 2.374 | Train | Train | Train |
| **C79** | 1,2-Benzisothiazolin-3-one, 2-butyl | -0.796 | Train | Test | Train |
| **C80** | Ethephon | -0.74 | Train | Train | Test |
| **C81** | Naphthalene | -1.322 | Train | Train | Train |
| **C82** | Toxaphene (Cancelled in U.S.) | 0.683 | Train | Test | Train |
| **C83** | Phostebupirim | 1.195 | Train | Train | Test |
| **C84** | Propanil | 0.035 | Train | Train | Train |
| **C85** | Triclosan | -0.455 | Train | Train | Test |
| **C86** | Etridiazole (Terrazole) | -0.355 | Train | Test | Train |
| **C87** | Grotan | -0.841 | Train | Train | Train |
| **C88** | Bifenazate | -0.536 | Train | Test | Test |
| **C89** | Thiazopyr | -0.684 | Train | Train | Train |
| **C90** | Sulprofos | 0.836 | Train | Train | Train |
| **C91** | Iprodione | -0.45 | Train | Test | Train |
| **C92** | Imidacloprid | 0.225 | Train | Train | Train |
| **C93** | Ipconazole | -0.46 | Train | Train | Test |
| **C94** | Furfural | 0.053 | Train | Train | Train |
| **C95** | Octhilinone (Acticide OIT) | -0.255 | Train | Test | Train |
| **C96** | 1,2-Benzisothiazolin-3-one | -0.611 | Train | Train | Test |
| **C97** | Amicarbazone | -0.911 | Test | Train | Train |
| **C98** | Atrazine desethylatrazine metabolite | -0.552 | Test | Train | Train |
| **C99** | Bendiocarb | 1.07 | Test | Test | Train |
| **C100** | Chlorophacinone | -0.121 | Test | Train | Train |
| **C101** | Chlorpyrifos | 1.04 | Test | Train | Train |
| **C102** | Cyromazine | -1.031 | Test | Test | Train |
| **C103** | DBNPA | -0.165 | Test | Train | Test |
| **C104** | Dicloran (DCNA) | -0.638 | Test | Train | Train |
| **C105** | Diethyl toluamide (DEET) | -0.857 | Test | Train | Train |
| **C106** | Dimethoate | 1.292 | Test | Train | Train |
| **C107** | Endosulfan | 0.986 | Test | Test | Train |
| **C108** | Isazophos methyl CGA 12223 | 1.411 | Test | Train | Train |
| **C109** | Mecoprop-P | -0.405 | Test | Train | Train |
| **C110** | Phorate | 1.571 | Test | Test | Test |
| **C111** | TCMTB | -0.443 | Test | Train | Train |
| **C112** | Tetrahydro-2(nitromethylene)-2H-1, 3-thiazine | -1.155 | Test | Train | Train |
| **C113** | Diazinon | 1.58 | Test | Test | Train |
| **C114** | Oxamyl | 0.748 | Test | Train | Train |
| **C115** | Pentachlorophenol | -0.372 | Test | Train | Train |
| **C116** | Methomyl | 0.826 | Test | Train | Test |
| **C117** | Hexazinone | -0.95 | Test | Train | Train |
| **C118** | Inert Surfadone(LP100) | -1.085 | Test | Train | Train |
| **C119** | Terbufos | 1.004 | Test | Test | Train |
| **C120** | Tetraconazole | 0.453 | Test | Train | Train |
| **C121** | 1,3 Dichloropropene (Telone soil fumigant) | -0.137 | Test | Train | Train |
| **C122** | Bromoxynil heptanoate | 0.035 | Test | Train | Test |
| **C123** | Cyhexatin | 0.109 | Test | Train | Train |
| **C124** | MCPP acid (Mecoprop) | -0.518 | Test | Test | Test |
| **C125** | 2-Bromo-4'-hydroxyacetophenone(BHAP) | -0.488 | Test | Train | Train |
| **C126** | Metconazole cis | -0.419 | Test | Train | Train |
| **C127** | 4,4-Dimethyloxazolidine | -0.843 | Test | Train | Train |
| **C128** | Bensulide | -0.542 |  |  |  |
| **C129** | Temephos | 1.231 |  |  |  |
| **C130** | Ethion | 0.478 |  |  |  |
| **C131** | Sulfluramid | 0.046 |  |  |  |

*Table S2. SMILES notations of the modeling set compounds*

|  |  |  |
| --- | --- | --- |
| **Cpd** | **SMILES** | **pLD50 (mmol)** |
| **C1** | C1=C(C(=C(S1)C)N(C(COC)C)C(CCl)=O)C | -0.588 |
| **C2** | OCNCCO | -1.282 |
| **C3** | C1=C(C(=CC(=C1)Cl)Cl)OCC(OC(C)C)=O | -0.854 |
| **C4** | C1=C([N](N=C1C)CO)C | -0.779 |
| **C5** | C(NC(OCC#CI)=O)CCC | -0.426 |
| **C6** | O=C1C(=C(SS1)Cl)Cl | -0.121 |
| **C7** | C1=C(C=CN=C1)N | 0.798 |
| **C8** | C1=C(C(=CC=C1O)Cl)C | -1.033 |
| **C9** | C1=CC=C(C(=C1CC)N(C(CCl)=O)COCC)C | 0.741 |
| **C10** | C1=CC=C(C(=C1CC)N(C(CCl)=O)COC)CC | -0.755 |
| **C11** | CC(SC)(/C=N/OC(=O)NC)C | 2.325 |
| **C12** | C(NC1=NC(=NC(=N1)SC)NC(C)C)C | -0.996 |
| **C13** | C1=CC(=CC(=C1N=CN(C=NC2=C(C=C(C)C=C2)C)C)C)C | 0.616 |
| **C14** | [C@H]1(C(O[C@H]([C@@H]([C@H](C(O[C@@H]1C)=O)CCCCCC)OC(=O)CC(C)C)C)=O)NC(C2=CC=CC(=C2O)NC=O)=O | 1.148 |
| **C15** | C1=CC=CC2=C1C(=O)N(N=N2)CS[P](=S)(OC)OC | 0.983 |
| **C16** | BrN1C(C(N(C1=O)Cl)=O)(C)C | -0.647 |
| **C17** | NCCCN(CCCCCCCCCCCC)CCCN | 1.079 |
| **C18** | C(C(N1C(=O)C(=C(NC1=O)C)Br)C)C | -0.935 |
| **C19** | C1=C(C(=C(C(F)(F)F)C=C1[N+]([O-])=O)N(C2=C(C=C(C=C2Br)Br)Br)C)[N+]([O-])=O | 1.719 |
| **C20** | C1=CC=C2C(=C1OC(NC)=O)OC(C2)(C)C | 1.642 |
| **C21** | C(O[P](=S)(OCC)OC(C(Cl)(Cl)Cl)Cl)C | 1.079 |
| **C22** | C2=C(C1=C(C(=C([N]1COCC)C(F)(F)F)Br)C#N)C=CC(=C2)Cl | 1.218 |
| **C23** | ClOC(C(O)(C1=CC=CC=C1)C2=CC=CC=C2)=O | -0.271 |
| **C24** | C1=C(OC(C(O)=O)C)C=CC=C1Cl | -0.817 |
| **C25** | C1=C(C(=NC(=C1Cl)O[P](OCC)(OCC)=O)Cl)Cl | 1.58 |
| **C26** | [C@H](OC2=CC=C(OC1=NC=C(Cl)C=C1F)C=C2)(C(OCC#C)=O)C | -0.619 |
| **C27** | C1=CC=CC(=C1CN2C(C(C)(C)CO2)=O)Cl | -1.02 |
| **C28** | C1=C(O[P](OCC)(OCC)=S)C=CC2=C1OC(=O)C(=C2C)Cl | 2.187 |
| **C29** | C1=CC(=CC=C1Cl)C(O)(C[N]2C=NC=N2)C(C3CC3)C | 0.289 |
| **C30** | C1=CC(=C(C(=C1Cl)C(O)=O)OC)Cl | 0.01 |
| **C31** | [C@H](OC1=C(Cl)C=C(Cl)C=C1)(C(=O)O)C | -0.178 |
| **C32** | CO[P](OC=C(Cl)Cl)(OC)=O | 1.4 |
| **C33** | C1(=C(C(=C(C1(Cl)C2(Cl)C(=C(C(=C2Cl)Cl)Cl)Cl)Cl)Cl)Cl)Cl | -0.172 |
| **C34** | C1=C(C(=C(S1)C)N(C(COC)C)C(CCl)=O)C | -0.84 |
| **C35** | C1=CC(=NC=C1C(OCCC)=O)C(OCCC)=O | -0.73 |
| **C36** | C(S[P](OCC)(OCC)=S)CSCC | 1.359 |
| **C37** | O=[S]2(OCC1C3(Cl)C(Cl)(Cl)C(Cl)(C1CO2)C(=C3Cl)Cl)=O | 0.983 |
| **C38** | O=C([O-])C1C(C2OC1CC2)C([O-])=O | -0.429 |
| **C39** | [C@H](C(C)C)(C(=O)O[C@@H](C#N)C1=CC(=CC=C1)OC2=CC=CC=C2)C3=CC=C(C=C3)Cl | 0.042 |
| **C40** | C1=C(O[P](NC(C)C)(OCC)=O)C=CC(=C1C)SC | 2.278 |
| **C41** | C1=C(O[P](OC)(OC)=S)C=CC(=C1C)[N+]([O-])=O | 1.005 |
| **C42** | C1=C(O[P](OC)(OC)=S)C=CC(=C1C)SC | 1.593 |
| **C43** | C1=C(C(F)(F)F)C=C(C(=C1[N+](=O)[O-])N(CCC)CCCl)[N+](=O)[O-] | -1.294 |
| **C44** | C1=CC=CC2=C1C(N(SC(Cl)(Cl)Cl)C2=O)=O | 0.732 |
| **C45** | C(O[P](N1C(SCC1)=O)(SC(CC)C)=O)C | 1.427 |
| **C46** | C(C(N(C)C)=O)CCCCCCCC | -0.905 |
| **C47** | C3=C(\C=C\C(\C=C\C1=CC=C(C=C1)C(F)(F)F)=N/NC2=NCC(CN2)(C)C)C=CC(=C3)C(F)(F)F | -0.568 |
| **C48** | [C@]13(OCN(N=C1C2=CC=C(Cl)C=C2C3)C(=O)N(C4=CC=C(OC(F)(F)F)C=C4)C(OC)=O)C(OC)=O | -0.486 |
| **C49** | C1=C(C(=CC=C1Cl)OCC(O)=O)C | -0.274 |
| **C50** | C1=CC=C(C(=C1C)N(C(COC)=O)C(C(=O)OC)C)C | -0.546 |
| **C51** | C1=C(OC(NC)=O)C=C(C)C(=C1C)SC | 1.061 |
| **C52** | C1=CC(=CC=C1O[P](OC)(OC)=S)[N+]([O-])=O | 1.542 |
| **C53** | C(C1=CC=CC=C1)NC2=C3C(=NC=N2)N=C[NH]3 | -0.851 |
| **C54** | C(S[P](OC)(OC)=O)C(NC)=O | 1.346 |
| **C55** | C1=C([S](N)(=O)=O)C=C(C(=C1[N+](=O)[O-])N(CCC)CCC)[N+](=O)[O-] | -0.165 |
| **C56** | C2=C(C=CC3=NC1=C(SC(=O)S1)N=C23)C | 0.078 |
| **C57** | C1=C(C=CC(=C1)Cl)Cl | -1.039 |
| **C58** | [N+](=O)([O-])C1=CC=C(C=C1)O | -0.618 |
| **C59** | C1=C(N=C(N=C1O[P](=S)(OC)OC)N(CC)CC)C | 0.883 |
| **C60** | [C@H]1(C([C@H]1C=C(C)C)(C)C)C(OC2C(=C(C(=O)C2)CC#C)C)=O | -0.591 |
| **C61** | C1=C(N(C(CCl)=O)C(C)C)C=CC=C1 | 0.381 |
| **C62** | C1=C(Cl)C(=CC=C1Cl)C2(OCC(CCC)O2)C[N]3C=NC=N3 | -0.917 |
| **C63** | C1=CC=CC(=C1OC(NC)=O)OC(C)C | -0.682 |
| **C64** | C(N(CCC)CC)C1OC2(OC1)CCC(CC2)C(C)(C)C | -0.27 |
| **C65** | CN(SN(C(O\N=C(SC)/C)=O)C)C(O\N=C(SC)/C)=O | -0.756 |
| **C66** | C2=C(CNC(=O)C1=C(Cl)C(=N[N]1C)CC)C=CC(=C2)OC3=CC=C(C=C3)C | 0.645 |
| **C67** | [C@@H]1(C([C@H]1C(C(Br)(Br)Br)Br)(C)C)C(=O)O[C@H](C#N)C2=CC(=CC=C2)OC3=CC=CC=C3 | -0.678 |
| **C68** | CO[P](C(C(Cl)(Cl)Cl)O)(OC)=O | 1.06 |
| **C69** | O=C1N(C(N(Cl)C(N1Cl)=O)=O)Cl | -0.858 |
| **C70** | C1=C(Cl)C(=NC(=C1Cl)OCC(OCCOCCCC)=O)Cl | -0.377 |
| **C71** | C1=C(C=C(C)C(=C1C)C)OC(NC)=O | -0.09 |
| **C72** | [C@@H](\C(=C/C1=CC=C(C=C1)Cl)[N]2N=CN=C2)(C(C)(C)C)O | -0.7 |
| **C73** | C1=C(C(=C(C(=C1C(F)(F)F)Cl)[N+](=O)[O-])NC2=C(C=C(C(F)(F)F)C=N2)Cl)[N+](=O)[O-] | -0.583 |
| **C74** | C1=C(N=C(N=C1O[P](OCC)(=O)OCC)C(C)C)C | 1.761 |
| **C75** | CC1CC(OC(C)O1)OC(C)=O | -0.959 |
| **C76** | C1=C([N+](=O)[O-])C=C(C(=C1C(CC)C)O)[N+](=O)[O-] | 0.779 |
| **C77** | C1=C(O[P](=S)(OCC)NC(C)C)C(=CC=C1)C(OC(C)C)=O | 1.599 |
| **C78** | C\C(O[P](OC)(OC)=O)=C/C(=O)NC | 2.374 |
| **C79** | C(CCC)N1C(C2=C(S1)C=CC=C2)=O | -0.796 |
| **C80** | C([P](=O)(O)O)CCl | -0.74 |
| **C81** | C1=C2C(=CC=C1)C=CC=C2 | -1.322 |
| **C82** | C(C1(C2(C(C(C(C1=C)(C2(Cl)Cl)Cl)Cl)Cl)Cl)CCl)Cl | 0.683 |
| **C83** | C1=C(O[P](=S)(OC(C)C)OCC)C=NC(=N1)C(C)(C)C | 1.195 |
| **C84** | C1=C(C=CC(=C1Cl)Cl)NC(CC)=O | 0.035 |
| **C85** | C1=CC(=CC(=C1OC2=CC=C(C=C2Cl)Cl)O)Cl | -0.455 |
| **C86** | C(OC1=NC(=NS1)C(Cl)(Cl)Cl)C | -0.355 |
| **C87** | C(C1NC(NC(N1)CCO)CCO)CO | -0.841 |
| **C88** | C2=C(C1=CC=CC=C1)C=CC(=C2NNC(OC(C)C)=O)OC | -0.536 |
| **C89** | CC(C)CC1=C(C(=NC(=C1C(=O)OC)C(F)F)C(F)(F)F)C2=NCCS2 | -0.684 |
| **C90** | C1=C(O[P](SCCC)(OCC)=S)C=CC(=C1)SC | 0.836 |
| **C91** | C2=C(N1C(=O)N(C(NC(C)C)=O)CC1=O)C=C(C=C2Cl)Cl | -0.45 |
| **C92** | C1=CC(=NC=C1CN2C(=NCC2)N[N+](=O)[O-])Cl | 0.225 |
| **C93** | C3=C(CC1C(C(CC1)C(C)C)(O)C[N]2N=CN=C2)C=CC(=C3)Cl | -0.46 |
| **C94** | C(C1=CC=CO1)=O | 0.053 |
| **C95** | C(N1C(C=CS1)=O)CCCCCCC | -0.255 |
| **C96** | C1=CC=CC2=C1C(NS2)=O | -0.611 |
| **C97** | CC(NC(N1N=C(N(C1=O)N)C(C)C)=O)(C)C | -0.911 |
| **C98** | C(NC1=NC(=NC(=N1)Cl)NC(C)C)C | -0.552 |
| **C99** | C1=CC=C(C2=C1OC(O2)(C)C)OC(NC)=O | 1.07 |
| **C100** | C4=C(C(C(=O)C1C(=O)C2=C(C1=O)C=CC=C2)C3=CC=CC=C3)C=CC(=C4)Cl | -0.121 |
| **C101** | C1=C(Cl)C(=NC(=C1Cl)O[P](OCC)(OCC)=S)Cl | 1.04 |
| **C102** | C1(=NC(=NC(=N1)N)N)NC2CC2 | -1.031 |
| **C103** | O=C(C(C#N)(Br)Br)N | -0.165 |
| **C104** | C1=C(C=C(Cl)C(=C1Cl)N)[N+](=O)[O-] | -0.638 |
| **C105** | C1=C(C(N(CC)CC)=O)C=CC=C1C | -0.857 |
| **C106** | C(S[P](OC)(OC)=S)C(NC)=O | 1.292 |
| **C107** | O=[S]2OCC1C3(Cl)C(Cl)(Cl)C(Cl)(C1CO2)C(=C3Cl)Cl | 0.986 |
| **C108** | CC([N]1C(=NC(=N1)O[P](=S)(OC)OC)Cl)C | 1.411 |
| **C109** | [C@H](OC1=CC=C(Cl)C=C1C)(C(=O)O)C | -0.405 |
| **C110** | C(S[P](OCC)(OCC)=S)SCC | 1.571 |
| **C111** | C1=CC=CC2=C1N=C(SCSC#N)S2 | -0.443 |
| **C112** | O=[N+](\C=C1/NCCCS1)[O-] | -1.155 |
| **C113** | S=[P](OCC)(OCC)OC1=NC(=NC(=C1)C)C(C)C | 1.58 |
| **C114** | CN(C(=O)/C(SC)=N/OC(=O)NC)C | 0.748 |
| **C115** | C1(=C(Cl)C(=C(O)C(=C1Cl)Cl)Cl)Cl | -0.372 |
| **C116** | C\C(SC)=N/OC(=O)NC | 0.826 |
| **C117** | CN1C(=O)N(C(=O)N=C1N(C)C)C2CCCCC2 | -0.95 |
| **C118** | C(N1C(CCC1)=O)CCCCCCC | -1.085 |
| **C119** | C(S[P](OCC)(OCC)=S)SC(C)(C)C | 1.004 |
| **C120** | C2=C(C(C[N]1C=NC=N1)COC(C(F)F)(F)F)C(=CC(=C2)Cl)Cl | 0.453 |
| **C121** | ClC=CCCl | -0.137 |
| **C122** | C1=C(C(=C(C=C1C#N)Br)OC(CCCCCC)=O)Br | 0.035 |
| **C123** | [Sn](C1CCCCC1)(C2CCCCC2)(C3CCCCC3)|O | 0.109 |
| **C124** | C1=CC(=CC(=C1OC(C(O)=O)C)C)Cl | -0.518 |
| **C125** | BrCC(=O)C1=CC=C(C=C1)O | -0.488 |
| **C126** | C1=NC=N[N]1CC2(O)C(CCC2CC3=CC=C(Cl)C=C3)(C)C | -0.419 |
| **C127** | CC1(COCN1)C | -0.843 |
| **C128** | C1=C([S](NCCS[P](OC(C)C)(OC(C)C)=S)(=O)=O)C=CC=C1 | -0.542 |
| **C129** | C1=C(O[P](OC)(OC)=S)C=CC(=C1)SC2=CC=C(O[P](OC)(OC)=S)C=C2 | 1.231 |
| **C130** | C(S[P](OCC)(OCC)=S)S[P](OCC)(OCC)=S | 0.478 |
| **C131** | CCNS(=O)(=O)C(C(C(C(C(C(C(C(F)(F)F)(F)F)(F)F)(F)F)(F)F)(F)F)(F)F)(F)F | 0.046 |

*Table S3. Validation dataset used for QSTR model external validation*

|  |  |  |  |
| --- | --- | --- | --- |
| Cpd | Name | SMILES | pLD50 (mmol) |
| **V1** | Arsanilic acid | C1=C([As](O)(O)=O)C=CC(=C1)N | -0.203 |
| **V2** | Benfluralin | C1=C(C(F)(F)F)C=C(C(=C1[N+](=O)[O-])N(CCCC)CC)[N+](=O)[O-] | -0.695 |
| **V3** | Cymoxanil | C(NC(NC(\C(=N\OC)C#N)=O)=O)C | -1.055 |
| **V4** | Cyprodinil | C1=C(N=C(N=C1C2CC2)NC3=CC=CC=C3)C | -0.948 |
| **V5** | CYROMAZINE | C1(=NC(=NC(=N1)N)N)NC2CC2 | -1.022 |
| **V6** | Deltamethrin | [C@H]1(C([C@@H]1C(O[C@@H](C2=CC(=CC=C2)OC3=CC=CC=C3)C#N)=O)(C)C)C=C(Br)Br | -0.649 |
| **V7** | Dikegulac sodium | [C@]23(O[C@@H]1[C@@H](OC(OC1)(C)C)[C@@H]2OC(O3)(C)C)C([O-])=O | -0.882 |
| **V8** | Ethephon | C([P](=O)(O)O)CCl | -0.870 |
| **V9** | Ethylenethiourea | C1NC(NC1)=S | -1.343 |
| **V10** | Fenhexamid | C1=CC(=C(Cl)C(=C1NC(=O)C2(CCCCC2)C)Cl)O | -0.821 |
| **V11** | Flonicamid | C1=CN=CC(=C1C(F)(F)F)C(=O)NCC#N | -0.941 |
| **V12** | Flumetsulam | C1=CC=C(F)C(=C1F)N[S](=O)(=O)C2=N[N]3C(=N2)N=C(C=C3)C | -0.840 |
| **V13** | Fluthianet methyl | C1=C(C(=CC(=C1SCC(=O)OC)Cl)F)N=C3N2N(CCCC2)C(S3)=O | -0.746 |
| **V14** | FLUTOLANIL | C1=C(C(=CC=C1)C(F)(F)F)C(NC2=CC(=CC=C2)OC(C)C)=O | -0.791 |
| **V15** | Flutriafol | C3=C(C(C1=CC=C(F)C=C1)(C[N]2C=NC=N2)O)C(=CC=C3)F | -0.430 |
| **V16** | Halofenozide | C2=C(C(=O)NN(C(=O)C1=CC=CC=C1)C(C)(C)C)C=CC(=C2)Cl | -0.833 |
| **V17** | Maneb | C(NC([S-])=S)CNC(=S)[S-] | -0.644 |
| **V18** | Methazole | C2=C(N1C(=O)N(C)C(O1)=O)C=CC(=C2Cl)Cl | -0.922 |
| **V19** | Metribuzine | CC(C)(C)C1=NN=C(N(C1=O)N)SC | -0.282 |
| **V20** | Norflurazone | C2=C(N1C(C(=C(NC)C=N1)Cl)=O)C=CC=C2C(F)(F)F | -0.387 |
| **V21** | ORYZALIN | C1=C([S](N)(=O)=O)C=C(C(=C1[N+](=O)[O-])N(CCC)CCC)[N+](=O)[O-] | -0.165 |
| **V22** | Oxythioquinox | C2=C(C=CC3=NC1=C(SC(=O)S1)N=C23)C | 0.078 |
| **V23** | Phostebupirim | C1=C(O[P](=S)(OC(C)C)OCC)C=NC(=N1)C(C)(C)C | 1.050 |
| **V24** | Prometon | COC1=NC(=NC(=N1)NC(C)C)NC(C)C | -1.002 |
| **V25** | Siduron | C1=CC=CC=C1NC(NC2C(CCCC2)C)=O | -0.986 |

*Table S4. Attributes with steady positive and negative contributions in the best model*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | ID | SAk | Probe 1 | Probe 2 | Probe 2 | NSs | NSv | Defect[SAk] |
| 1 | 246 | EC2-H...7... | 1.571 | 1.301 | 0.913 | 81 | 28 | 0.0005 |
| 2 | 576 | VS2-H...6... | 0.663 | 0.736 | 1.473 | 81 | 28 | 0.0005 |
| 3 | 40 | ++++O---B2== | 3.545 | 3.801 | 3.526 | 79 | 29 | 0.001 |
| 4 | 186 | BOND10000000 | 2.640 | 3.126 | 2.503 | 79 | 28 | 0.0008 |
| 5 | 149 | C...=....... | 0.947 | 0.359 | 0.349 | 75 | 25 | 0.0003 |
| 6 | 126 | C...(...C... | 3.133 | 3.091 | 3.322 | 65 | 22 | 0.0004 |
| 7 | 443 | O...=....... | 1.173 | 1.767 | 1.597 | 64 | 21 | 0.0001 |
| 8 | 467 | NNC-C...312. | 2.204 | 2.392 | 2.486 | 64 | 24 | 0.0012 |
| 9 | 35 | ++++N---B2== | 0.016 | 0.241 | 0.028 | 63 | 23 | 0.001 |
| 10 | 562 | VS2-C...8... | 0.627 | 0.828 | 0.550 | 63 | 20 | 0.0001 |
| 11 | 207 | EC2-C...20.. | 2.599 | 2.142 | 2.492 | 62 | 23 | 0.0011 |
| 12 | 476 | NNC-H...110. | 0.624 | 0.933 | 0.905 | 62 | 28 | 0.0029 |
| 13 | 442 | O...=...(... | 2.702 | 2.349 | 2.720 | 61 | 20 | 0.0001 |
| 14 | 494 | NNC-O...202. | 0.402 | 0.104 | 0.400 | 61 | 23 | 0.0013 |
| 15 | 492 | NNC-O...101. | 1.734 | 1.552 | 1.276 | 57 | 18 | 0.0002 |
| 16 | 100 | =...C...(... | 3.303 | 3.325 | 3.350 | 55 | 20 | 0.001 |
| 17 | 448 | O...C....... | 0.098 | 0.153 | 0.099 | 54 | 20 | 0.0011 |
| 18 | 154 | C...=...C... | 0.734 | 0.928 | 1.308 | 53 | 17 | 0.0001 |
| 19 | 245 | EC2-H...6... | 0.899 | 0.926 | 0.899 | 53 | 15 | 0.001 |
| 20 | 541 | VS2-C...12.. | 1.572 | 1.454 | 2.016 | 53 | 13 | 0.002 |
| 21 | 575 | VS2-H...5... | 1.183 | 1.241 | 1.291 | 53 | 15 | 0.001 |
| 22 | 102 | =...C...1... | 3.054 | 3.546 | 3.724 | 52 | 20 | 0.0014 |
| 23 | 539 | VS2-C...10.. | 0.688 | 0.434 | 0.275 | 52 | 18 | 0.0006 |
| 24 | 158 | C...C...(... | 2.028 | 2.273 | 1.948 | 50 | 17 | 0.0004 |
| 25 | 150 | C...=...1... | 3.672 | 3.726 | 3.678 | 48 | 22 | 0.003 |
| 26 | 560 | VS2-C...6... | 0.901 | 0.533 | 0.528 | 48 | 14 | 0.0008 |
| 27 | 2 | (...(....... | 1.221 | 1.327 | 1.399 | 46 | 15 | 0.0001 |
| 28 | 111 | =...O...(... | 1.225 | 1.649 | 1.301 | 46 | 18 | 0.0016 |
| 29 | 400 | N...C....... | 1.202 | 1.036 | 0.995 | 46 | 10 | 0.0028 |
| 30 | 148 | C...=...(... | 3.827 | 3.473 | 3.774 | 44 | 16 | 0.001 |
| 31 | 210 | EC2-C...23.. | 0.375 | 0.713 | 0.614 | 44 | 11 | 0.0019 |
| 32 | 173 | C...O...(... | 1.596 | 1.899 | 1.502 | 41 | 15 | 0.001 |
| 33 | 356 | Cl.......... | 0.063 | 0.171 | 0.174 | 41 | 15 | 0.001 |
| 34 | 62 | 1...C...(... | 2.028 | 1.954 | 2.197 | 38 | 9 | 0.0022 |
| 35 | 347 | Cl..(....... | 0.008 | 0.249 | 0.364 | 38 | 14 | 0.0011 |
| 36 | 558 | VS2-C...4... | 2.201 | 2.395 | 2.325 | 38 | 12 | 0.0002 |
| 37 | 468 | NNC-C...321. | 1.436 | 1.211 | 1.648 | 37 | 16 | 0.0025 |
| 38 | 57 | 1...(....... | 1.786 | 1.898 | 1.729 | 36 | 9 | 0.0019 |
| 39 | 625 | VS2-O...6... | 0.940 | 0.551 | 0.372 | 35 | 10 | 0.0009 |
| 40 | 137 | C...2....... | 0.803 | 1.398 | 1.284 | 34 | 10 | 0.0007 |
| 41 | 480 | NNC-Cl..101. | 1.098 | 0.833 | 0.736 | 34 | 15 | 0.0026 |
| 42 | 237 | EC2-H...10.. | 0.725 | 0.822 | 0.964 | 33 | 6 | 0.0039 |
| 43 | 244 | EC2-H...5... | 0.441 | 0.410 | 0.321 | 33 | 15 | 0.0029 |
| 44 | 574 | VS2-H...4... | 0.485 | 0.788 | 0.904 | 33 | 15 | 0.0029 |
| 45 | 579 | VS2-H...9... | 0.838 | 0.384 | 0.559 | 33 | 6 | 0.0039 |
| 46 | 118 | C...(...(... | 0.395 | 0.400 | 0.153 | 32 | 8 | 0.0019 |
| 47 | 248 | EC2-H...9... | 1.252 | 1.312 | 0.666 | 32 | 6 | 0.0037 |
| 48 | 543 | VS2-C...14.. | 1.187 | 0.795 | 1.224 | 32 | 4 | 0.0057 |
| 49 | 561 | VS2-C...7... | 0.873 | 0.339 | 0.288 | 32 | 12 | 0.0012 |
| 50 | 578 | VS2-H...8... | 0.638 | 0.686 | 1.578 | 32 | 6 | 0.0037 |
| 51 | 211 | EC2-C...24.. | 4.760 | 4.848 | 4.900 | 31 | 8 | 0.0017 |
| 52 | 447 | O...C...(... | 0.996 | 0.828 | 0.873 | 30 | 14 | 0.0032 |
| 53 | 45 | ++++S---B2== | 1.652 | 1.254 | 1.347 | 29 | 12 | 0.0021 |
| 54 | 212 | EC2-C...25.. | 3.703 | 3.604 | 3.299 | 29 | 10 | 0.0005 |
| 55 | 514 | S........... | 0.525 | 0.708 | 0.704 | 29 | 12 | 0.0021 |
| 56 | 8 | (...Cl..(... | 0.184 | 0.062 | 0.377 | 28 | 6 | 0.0029 |
| 57 | 43 | ++++O---S=== | 1.287 | 1.325 | 1.379 | 28 | 11 | 0.0016 |
| 58 | 261 | EC2-Cl..7... | 1.223 | 0.930 | 1.253 | 28 | 13 | 0.0031 |
| 59 | 591 | VS2-Cl..6... | 1.445 | 1.192 | 1.248 | 28 | 13 | 0.0031 |
| 60 | 214 | EC2-C...27.. | 2.703 | 2.403 | 2.398 | 27 | 3 | 0.0061 |
| 61 | 399 | N...C...(... | 4.404 | 4.402 | 4.504 | 27 | 4 | 0.0049 |
| 62 | 70 | 2...(....... | 0.662 | 1.224 | 1.225 | 26 | 6 | 0.0024 |
| 63 | 167 | C...N...(... | 4.549 | 4.396 | 4.597 | 26 | 3 | 0.006 |
| 64 | 300 | EC2-O...6... | 0.855 | 0.624 | 0.496 | 24 | 8 | 0.0003 |
| 65 | 484 | NNC-N...202. | 0.703 | 0.496 | 0.722 | 24 | 8 | 0.0003 |
| 66 | 544 | VS2-C...15.. | 3.721 | 3.933 | 3.942 | 24 | 2 | 0.0071 |
| 67 | 39 | ++++N---S=== | 0.484 | 0.449 | 0.449 | 23 | 8 | 0.0006 |
| 68 | 302 | EC2-O...8... | 1.734 | 0.773 | 0.811 | 23 | 9 | 0.0016 |
| 69 | 135 | C...1...C... | 3.413 | 3.745 | 3.549 | 22 | 8 | 0.001 |
| 70 | 488 | NNC-N...312. | 1.639 | 1.872 | 1.637 | 22 | 8 | 0.001 |
| 71 | 693 | [...O....... | 1.500 | 1.427 | 1.298 | 22 | 9 | 0.002 |
| 72 | 131 | C...1...(... | 1.083 | 1.129 | 1.017 | 21 | 5 | 0.0022 |
| 73 | 507 | S...(....... | 0.520 | 0.450 | 0.446 | 21 | 11 | 0.0043 |
| 74 | 420 | O...(...(... | 1.600 | 1.348 | 1.302 | 20 | 11 | 0.0047 |
| 75 | 73 | 2...C...(... | 2.597 | 2.051 | 2.251 | 19 | 2 | 0.0064 |
| 76 | 97 | =...2....... | 2.047 | 2.228 | 3.014 | 19 | 8 | 0.0022 |
| 77 | 136 | C...2...(... | 3.796 | 3.271 | 3.498 | 19 | 3 | 0.0046 |
| 78 | 139 | C...2...=... | 2.242 | 3.025 | 3.029 | 19 | 8 | 0.0022 |
| 79 | 239 | EC2-H...12.. | 0.999 | 1.002 | 1.000 | 19 | 2 | 0.0064 |
| 80 | 569 | VS2-H...11.. | 0.999 | 1.090 | 1.304 | 19 | 2 | 0.0064 |
| 81 | 103 | =...C...2... | 0.403 | 0.749 | 0.705 | 18 | 6 | 0.0003 |
| 82 | 42 | ++++O---P=== | 1.602 | 1.596 | 1.650 | 17 | 9 | 0.0044 |
| 83 | 44 | ++++P---B2== | 1.600 | 1.598 | 1.595 | 17 | 9 | 0.0044 |
| 84 | 151 | C...=...2... | 2.723 | 3.047 | 2.937 | 17 | 8 | 0.0032 |
| 85 | 384 | N...1....... | 0.289 | 0.754 | 0.354 | 17 | 3 | 0.004 |
| 86 | 460 | P........... | 1.796 | 1.728 | 1.724 | 17 | 9 | 0.0044 |
| 87 | 461 | P...[...(... | 1.603 | 1.601 | 1.897 | 17 | 9 | 0.0044 |
| 88 | 652 | [...(...O... | 4.055 | 4.004 | 4.001 | 17 | 7 | 0.002 |
| 89 | 696 | [...P....... | 1.604 | 1.998 | 1.898 | 17 | 9 | 0.0044 |
| 90 | 697 | [...P...[... | 1.599 | 1.851 | 1.596 | 17 | 9 | 0.0044 |
| 91 | 164 | C...C...C... | 1.751 | 1.786 | 1.700 | 16 | 3 | 0.0037 |
| 92 | 243 | EC2-H...4... | 2.733 | 2.729 | 3.005 | 16 | 7 | 0.0026 |
| 93 | 291 | EC2-O...15.. | 1.997 | 2.096 | 1.730 | 16 | 9 | 0.0049 |
| 94 | 481 | NNC-Cl..110. | 1.004 | 0.801 | 0.900 | 16 | 2 | 0.0057 |
| 95 | 497 | NNC-P...404. | 1.595 | 1.603 | 1.601 | 16 | 8 | 0.0038 |
| 96 | 500 | NNC-S...202. | 0.829 | 1.427 | 1.298 | 16 | 5 | 0.0003 |
| 97 | 509 | S...(...C... | 0.697 | 0.396 | 0.402 | 16 | 6 | 0.0012 |
| 98 | 573 | VS2-H...3... | 2.604 | 2.986 | 2.499 | 16 | 7 | 0.0026 |
| 99 | 615 | VS2-O...11.. | 2.178 | 1.937 | 2.198 | 16 | 9 | 0.0049 |
| 100 | 266 | EC2-N...14.. | 2.103 | 2.199 | 2.204 | 15 | 7 | 0.0032 |
| 101 | 290 | EC2-O...14.. | 3.691 | 3.600 | 3.248 | 15 | 2 | 0.0054 |
| 102 | 355 | Cl..(...Cl.. | 4.435 | 4.678 | 4.374 | 15 | 3 | 0.0033 |
| 103 | 594 | VS2-N...10.. | 0.877 | 0.629 | 0.749 | 15 | 7 | 0.0032 |
| 104 | 614 | VS2-O...10.. | 3.480 | 2.874 | 3.774 | 15 | 2 | 0.0054 |
| 105 | 687 | [...N....... | 0.684 | 0.622 | 0.827 | 15 | 7 | 0.0032 |
| 106 | 267 | EC2-N...15.. | 4.002 | 3.829 | 4.299 | 14 | 3 | 0.0029 |
| 107 | 397 | N...=...C... | 0.776 | 0.523 | 0.130 | 14 | 7 | 0.0038 |
| 108 | 523 | S...C....... | 3.523 | 2.522 | 3.400 | 14 | 6 | 0.0024 |
| 109 | 545 | VS2-C...16.. | 2.751 | 3.046 | 2.798 | 14 | 5 | 0.0008 |
| 110 | 47 | ++++S---P=== | 1.002 | 1.003 | 1.001 | 13 | 7 | 0.0045 |
| 111 | 463 | P...[...O... | 1.900 | 1.904 | 1.899 | 13 | 5 | 0.0014 |
| 112 | 493 | NNC-O...110. | 3.004 | 3.386 | 3.296 | 13 | 4 | 0.0004 |
| 113 | 547 | VS2-C...18.. | 2.178 | 2.025 | 2.002 | 13 | 2 | 0.0047 |
| 114 | 627 | VS2-O...8... | 2.497 | 2.122 | 2.452 | 13 | 5 | 0.0014 |
| 115 | 38 | ++++N---P=== | 1.903 | 1.897 | 1.903 | 12 | 4 | 0.0003 |
| 116 | 217 | EC2-C...30.. | 1.613 | 1.504 | 1.678 | 12 | 3 | 0.0019 |
| 117 | 595 | VS2-N...11.. | 1.598 | 1.295 | 1.295 | 12 | 2 | 0.0043 |
| 118 | 649 | [...(...C... | 0.600 | 0.595 | 0.751 | 12 | 4 | 0.0003 |
| 119 | 81 | 3........... | 0.530 | 0.273 | 0.084 | 11 | 4 | 0.001 |
| 120 | 178 | C...S...(... | 5.146 | 4.547 | 4.547 | 11 | 4 | 0.001 |
| 121 | 288 | EC2-O...12.. | 2.504 | 2.221 | 2.602 | 11 | 4 | 0.001 |
| 122 | 385 | N...1...C... | 0.841 | 0.187 | 0.950 | 11 | 1 | 0.0069 |
| 123 | 557 | VS2-C...3... | 0.754 | 0.501 | 0.525 | 11 | 5 | 0.0029 |
| 124 | 611 | VS2-N...7... | 1.503 | 2.065 | 1.776 | 11 | 8 | 0.0076 |
| 125 | 661 | [...1....... | 0.654 | 0.372 | 0.922 | 11 | 1 | 0.0069 |
| 126 | 677 | [...C....... | 0.162 | 0.189 | 0.376 | 11 | 3 | 0.0013 |
| 127 | 691 | [...O...(... | 1.299 | 1.602 | 1.650 | 11 | 5 | 0.0029 |
| 128 | 11 | (...O...(... | 4.852 | 4.504 | 5.347 | 10 | 5 | 0.0038 |
| 129 | 203 | EC2-C...16.. | 2.138 | 2.928 | 3.317 | 10 | 5 | 0.0038 |
| 130 | 219 | EC2-C...32.. | 1.327 | 1.017 | 1.275 | 10 | 0 | 1 |
| 131 | 260 | EC2-Cl..6... | 3.401 | 3.125 | 3.205 | 10 | 3 | 0.0006 |
| 132 | 268 | EC2-N...16.. | 1.763 | 1.471 | 2.446 | 10 | 6 | 0.0056 |
| 133 | 292 | EC2-O...16.. | 4.604 | 4.648 | 4.599 | 10 | 3 | 0.0006 |
| 134 | 305 | EC2-P...27.. | 2.629 | 2.277 | 2.524 | 10 | 4 | 0.0018 |
| 135 | 403 | N...C...=... | 2.801 | 2.804 | 2.202 | 10 | 3 | 0.0006 |
| 136 | 449 | O...C...1... | 3.854 | 4.105 | 3.946 | 10 | 0 | 1 |
| 137 | 489 | NNC-N...321. | 2.302 | 1.999 | 1.642 | 10 | 1 | 0.0065 |
| 138 | 499 | NNC-S...101. | 0.979 | 0.755 | 0.752 | 10 | 7 | 0.0072 |
| 139 | 520 | S...=....... | 0.774 | 1.078 | 0.397 | 10 | 7 | 0.0072 |
| 140 | 590 | VS2-Cl..5... | 3.346 | 3.051 | 2.945 | 10 | 3 | 0.0006 |
| 141 | 597 | VS2-N...13.. | 1.401 | 1.354 | 1.499 | 10 | 1 | 0.0065 |
| 142 | 610 | VS2-N...6... | 3.704 | 3.851 | 4.005 | 10 | 2 | 0.0033 |
| 143 | 612 | VS2-N...8... | 3.446 | 3.525 | 3.396 | 10 | 2 | 0.0033 |
| 144 | 616 | VS2-O...12.. | 4.725 | 4.600 | 4.702 | 10 | 3 | 0.0006 |
| 145 | 623 | VS2-O...4... | 1.971 | 2.124 | 2.017 | 10 | 6 | 0.0056 |
| 146 | 629 | VS2-P...11.. | 1.945 | 2.340 | 2.627 | 10 | 4 | 0.0018 |
| 147 | 648 | [...(...=... | 0.901 | 0.785 | 0.990 | 10 | 8 | 0.0085 |
| 148 | 663 | [...1...C... | 0.976 | 0.139 | 0.440 | 10 | 1 | 0.0065 |
| 149 | 112 | =...S...(... | 1.402 | 1.053 | 0.948 | 9 | 7 | 0.0083 |
| 150 | 122 | C...(...2... | 2.854 | 3.790 | 3.799 | 9 | 4 | 0.0027 |
| 151 | 303 | EC2-O...9... | 0.702 | 1.002 | 1.296 | 9 | 3 | 0.0003 |
| 152 | 319 | EC2-S...7... | 0.400 | 0.098 | 0.097 | 9 | 7 | 0.0083 |
| 153 | 430 | O...(...N... | 3.396 | 3.153 | 3.404 | 9 | 2 | 0.0027 |
| 154 | 437 | O...1....... | 2.199 | 2.200 | 2.401 | 9 | 4 | 0.0027 |
| 155 | 519 | S...=...(... | 1.073 | 1.296 | 1.054 | 9 | 7 | 0.0083 |
| 156 | 641 | VS2-S...6... | 0.401 | 0.101 | 0.396 | 9 | 7 | 0.0083 |
| 157 | 104 | =...C...3... | 1.122 | 0.548 | 0.774 | 8 | 3 | 0.0012 |
| 158 | 140 | C...2...C... | 0.787 | 1.886 | 1.520 | 8 | 1 | 0.0057 |
| 159 | 202 | EC2-C...15.. | 2.800 | 3.102 | 3.299 | 8 | 4 | 0.0038 |
| 160 | 269 | EC2-N...17.. | 4.625 | 4.265 | 4.980 | 8 | 1 | 0.0057 |
| 161 | 272 | EC2-N...20.. | 1.296 | 1.295 | 1.297 | 8 | 1 | 0.0057 |
| 162 | 273 | EC2-N...21.. | 1.199 | 1.297 | 1.052 | 8 | 2 | 0.0019 |
| 163 | 54 | /........... | 1.727 | 2.279 | 1.735 | 7 | 1 | 0.0051 |
| 164 | 65 | 1...Cl..(... | 2.229 | 2.398 | 2.855 | 7 | 4 | 0.0051 |
| 165 | 93 | =...(...1... | 4.896 | 4.603 | 4.728 | 7 | 0 | 1 |
| 166 | 113 | @........... | 0.922 | 0.498 | 0.771 | 7 | 1 | 0.0051 |
| 167 | 155 | C...@....... | 0.601 | 0.467 | 0.778 | 7 | 1 | 0.0051 |
| 168 | 270 | EC2-N...18.. | 1.300 | 1.298 | 1.004 | 7 | 1 | 0.0051 |
| 169 | 274 | EC2-N...22.. | 1.405 | 1.522 | 0.899 | 7 | 1 | 0.0051 |
| 170 | 277 | EC2-N...25.. | 1.151 | 1.001 | 1.529 | 7 | 0 | 1 |
| 171 | 286 | EC2-O...10.. | 1.979 | 2.311 | 1.749 | 7 | 2 | 0.0009 |
| 172 | 310 | EC2-S...13.. | 2.454 | 2.128 | 2.403 | 7 | 2 | 0.0009 |
| 173 | 414 | N...[...1... | 3.889 | 4.046 | 4.028 | 7 | 1 | 0.0051 |
| 174 | 431 | O...(...O... | 3.505 | 2.879 | 2.999 | 7 | 5 | 0.0074 |
| 175 | 438 | O...1...C... | 2.197 | 2.199 | 2.204 | 7 | 4 | 0.0051 |
| 176 | 598 | VS2-N...14.. | 3.153 | 3.198 | 3.692 | 7 | 1 | 0.0051 |
| 177 | 600 | VS2-N...16.. | 1.350 | 1.396 | 1.351 | 7 | 0 | 1 |
| 178 | 644 | VS2-S...9... | 2.097 | 2.152 | 2.200 | 7 | 2 | 0.0009 |
| 179 | 679 | [...C...@... | 0.546 | 0.971 | 0.845 | 7 | 1 | 0.0051 |
| 180 | 690 | [...N...[... | 3.097 | 3.098 | 3.474 | 7 | 3 | 0.0024 |
| 181 | 1 | #........... | 0.950 | 1.166 | 0.803 | 6 | 3 | 0.0038 |
| 182 | 6 | (...F...(... | 1.585 | 1.398 | 1.688 | 6 | 2 | 0.0003 |
| 183 | 12 | ++++B2--B3== | 0.340 | 1.100 | 0.720 | 6 | 3 | 0.0038 |
| 184 | 25 | ++++Br--B2== | 2.835 | 2.941 | 3.198 | 6 | 2 | 0.0003 |
| 185 | 28 | ++++Br--O=== | 2.676 | 2.901 | 2.448 | 6 | 2 | 0.0003 |
| 186 | 51 | -...[...(... | 6.204 | 5.460 | 5.604 | 6 | 4 | 0.0067 |
| 187 | 69 | 2...(...(... | 1.603 | 1.596 | 1.054 | 6 | 0 | 1 |
| 188 | 116 | C...#....... | 1.720 | 1.412 | 1.234 | 6 | 3 | 0.0038 |
| 189 | 185 | BOND00000000 | 2.848 | 2.580 | 3.254 | 6 | 0 | 1 |
| 190 | 190 | F...(...(... | 0.925 | 0.509 | 0.752 | 6 | 2 | 0.0003 |
| 191 | 194 | F...(...F... | 1.422 | 2.245 | 2.690 | 6 | 2 | 0.0003 |
| 192 | 264 | EC2-N...12.. | 3.748 | 3.821 | 3.796 | 6 | 1 | 0.0043 |
| 193 | 323 | H...@....... | 1.123 | 1.028 | 0.851 | 6 | 1 | 0.0043 |
| 194 | 324 | H...@...C... | 0.527 | 0.471 | 0.902 | 6 | 1 | 0.0043 |
| 195 | 335 | Br.......... | 1.378 | 1.350 | 1.276 | 6 | 2 | 0.0003 |
| 196 | 346 | Cl..(...(... | 6.050 | 5.890 | 6.148 | 6 | 1 | 0.0043 |
| 197 | 383 | N...1...(... | 2.305 | 2.504 | 2.401 | 6 | 2 | 0.0003 |
| 198 | 404 | N...C...C... | 1.545 | 1.602 | 1.846 | 6 | 0 | 1 |
| 199 | 426 | O...(...=... | 5.876 | 5.501 | 5.901 | 6 | 3 | 0.0038 |
| 200 | 440 | O...2....... | 1.196 | 1.489 | 0.246 | 6 | 1 | 0.0043 |
| 201 | 473 | NNC-F...101. | 1.202 | 0.880 | 1.724 | 6 | 2 | 0.0003 |
| 202 | 522 | S...C...(... | 0.403 | 0.299 | 0.096 | 6 | 5 | 0.009 |
| 203 | 682 | [...H...@... | 0.795 | 0.585 | 0.524 | 6 | 1 | 0.0043 |
| 204 | 699 | [...S....... | 1.848 | 1.901 | 1.597 | 6 | 3 | 0.0038 |
| 205 | 27 | ++++Br--N=== | 1.437 | 1.584 | 1.602 | 5 | 2 | 0.0018 |
| 206 | 41 | ++++O---B3== | 2.120 | 1.690 | 2.115 | 5 | 3 | 0.0056 |
| 207 | 106 | =...N...(... | 0.723 | 0.905 | 0.827 | 5 | 3 | 0.0056 |
| 208 | 123 | C...(...3... | 1.334 | 1.398 | 0.959 | 5 | 1 | 0.0033 |
| 209 | 129 | C.../....... | 3.555 | 2.890 | 3.297 | 5 | 0 | 1 |
| 210 | 175 | C...O...1... | 3.396 | 3.401 | 3.404 | 5 | 3 | 0.0056 |
| 211 | 187 | BOND10100000 | 1.947 | 2.149 | 1.652 | 5 | 0 | 1 |
| 212 | 276 | EC2-N...24.. | 3.299 | 3.198 | 3.098 | 5 | 1 | 0.0033 |
| 213 | 306 | EC2-P...28.. | 2.197 | 2.196 | 2.500 | 5 | 4 | 0.0085 |
| 214 | 352 | Cl..(...=... | 5.202 | 5.447 | 5.396 | 5 | 3 | 0.0056 |
| 215 | 364 | Cl..C....... | 6.702 | 6.698 | 7.005 | 5 | 4 | 0.0085 |
| 216 | 424 | O...(...2... | 0.823 | 0.399 | 0.274 | 5 | 1 | 0.0033 |
| 217 | 550 | VS2-C...21.. | 4.900 | 4.905 | 5.080 | 5 | 3 | 0.0056 |
| 218 | 599 | VS2-N...15.. | 3.300 | 3.104 | 3.397 | 5 | 1 | 0.0033 |
| 219 | 630 | VS2-P...12.. | 2.198 | 2.196 | 2.197 | 5 | 4 | 0.0085 |
| 220 | 23 | ++++CL--P=== | 1.929 | 1.548 | 1.345 | 4 | 4 | 0.0109 |
| 221 | 84 | 3...C...(... | 2.399 | 2.899 | 2.325 | 4 | 1 | 0.0019 |
| 222 | 166 | C...Cl..(... | 6.800 | 6.654 | 6.486 | 4 | 3 | 0.0079 |
| 223 | 176 | C...O...2... | 1.148 | 0.350 | 0.485 | 4 | 0 | 1 |
| 224 | 224 | EC2-C...37.. | 1.852 | 1.789 | 2.021 | 4 | 1 | 0.0019 |
| 225 | 231 | EC2-C...8... | 0.484 | 0.248 | 0.101 | 4 | 0 | 1 |
| 226 | 251 | EC2-Br..7... | 1.324 | 0.966 | 0.848 | 4 | 2 | 0.0038 |
| 227 | 259 | EC2-Cl..5... | 3.596 | 3.149 | 3.501 | 4 | 0 | 1 |
| 228 | 275 | EC2-N...23.. | 4.305 | 4.303 | 3.997 | 4 | 1 | 0.0019 |
| 229 | 311 | EC2-S...14.. | 2.595 | 1.976 | 2.377 | 4 | 3 | 0.0079 |
| 230 | 349 | Cl..(...2... | 3.996 | 3.749 | 3.697 | 4 | 1 | 0.0019 |
| 231 | 363 | Cl..C...(... | 6.723 | 6.900 | 7.003 | 4 | 3 | 0.0079 |
| 232 | 380 | N.../....... | 1.496 | 1.328 | 1.803 | 4 | 1 | 0.0019 |
| 233 | 402 | N...C...2... | 3.401 | 3.299 | 3.095 | 4 | 1 | 0.0019 |
| 234 | 439 | O...2...(... | 0.702 | 1.074 | 1.120 | 4 | 1 | 0.0019 |
| 235 | 464 | NNC-C...202. | 2.354 | 2.554 | 2.703 | 4 | 2 | 0.0038 |
| 236 | 477 | NNC-Br..101. | 4.525 | 4.045 | 4.008 | 4 | 1 | 0.0019 |
| 237 | 496 | NNC-O...220. | 2.795 | 2.999 | 2.804 | 4 | 0 | 1 |
| 238 | 506 | S...(...(... | 2.503 | 3.048 | 2.653 | 4 | 3 | 0.0079 |
| 239 | 525 | S...C...C... | 1.498 | 1.597 | 1.298 | 4 | 2 | 0.0038 |
| 240 | 581 | VS2-Br..6... | 0.880 | 1.187 | 0.888 | 4 | 2 | 0.0038 |
| 241 | 589 | VS2-Cl..4... | 3.173 | 3.446 | 3.222 | 4 | 0 | 1 |
| 242 | 633 | VS2-S...10.. | 2.410 | 2.454 | 2.305 | 4 | 3 | 0.0079 |
| 243 | 671 | [...=....... | 2.547 | 2.846 | 2.402 | 4 | 0 | 1 |
| 244 | 709 | \...C....... | 0.761 | 0.746 | 0.871 | 4 | 1 | 0.0019 |
| 245 | 101 | =...C.../... | 6.449 | 6.996 | 6.996 | 3 | 0 | 1 |
| 246 | 107 | =...N.../... | 4.303 | 4.595 | 4.898 | 3 | 1 | 0.0003 |
| 247 | 117 | C...#...C... | 2.050 | 2.295 | 2.105 | 3 | 0 | 1 |
| 248 | 144 | C...3...C... | 4.052 | 4.347 | 3.424 | 3 | 0 | 1 |
| 249 | 157 | C...C...#... | 2.379 | 1.929 | 2.422 | 3 | 0 | 1 |
| 250 | 162 | C...C...3... | 1.485 | 0.999 | 1.551 | 3 | 1 | 0.0003 |
| 251 | 169 | C...N...1... | 2.502 | 2.503 | 2.503 | 3 | 0 | 1 |
| 252 | 182 | C...[...(... | 1.179 | 1.197 | 0.822 | 3 | 0 | 1 |
| 253 | 188 | BOND11000000 | 5.847 | 6.100 | 5.799 | 3 | 2 | 0.0067 |
| 254 | 198 | EC2-C...10.. | 2.147 | 2.328 | 2.391 | 3 | 2 | 0.0067 |
| 255 | 201 | EC2-C...14.. | 5.598 | 6.073 | 5.628 | 3 | 0 | 1 |
| 256 | 254 | EC2-Cl..10.. | 1.000 | 1.429 | 1.500 | 3 | 0 | 1 |
| 257 | 256 | EC2-Cl..12.. | 1.422 | 0.492 | 1.898 | 3 | 0 | 1 |
| 258 | 278 | EC2-N...26.. | 0.101 | 0.095 | 0.400 | 3 | 0 | 1 |
| 259 | 296 | EC2-O...20.. | 5.002 | 4.697 | 4.800 | 3 | 0 | 1 |
| 260 | 299 | EC2-O...5... | 5.272 | 5.104 | 5.228 | 3 | 1 | 0.0003 |
| 261 | 315 | EC2-S...18.. | 5.351 | 5.398 | 5.922 | 3 | 0 | 1 |
| 262 | 328 | H...[...1... | 1.397 | 1.425 | 1.276 | 3 | 0 | 1 |
| 263 | 376 | N...(...Cl.. | 3.200 | 3.601 | 3.025 | 3 | 1 | 0.0003 |
| 264 | 418 | N...[...C... | 4.503 | 4.403 | 5.026 | 3 | 2 | 0.0067 |
| 265 | 419 | N...[...N... | 4.304 | 3.755 | 4.002 | 3 | 0 | 1 |
| 266 | 423 | O...(...1... | 1.347 | 1.127 | 1.646 | 3 | 2 | 0.0067 |
| 267 | 446 | O...=...C... | 2.200 | 2.200 | 2.703 | 3 | 2 | 0.0067 |
| 268 | 454 | O...C...N... | 1.899 | 1.727 | 1.200 | 3 | 0 | 1 |
| 269 | 486 | NNC-N...220. | 1.079 | 1.248 | 0.999 | 3 | 0 | 1 |
| 270 | 536 | S...[...P... | 2.697 | 2.302 | 2.199 | 3 | 3 | 0.0109 |
| 271 | 586 | VS2-Cl..11.. | 1.948 | 1.297 | 1.609 | 3 | 0 | 1 |
| 272 | 593 | VS2-Cl..9... | 1.650 | 1.196 | 1.110 | 3 | 0 | 1 |
| 273 | 601 | VS2-N...17.. | 0.095 | 0.098 | 0.099 | 3 | 0 | 1 |
| 274 | 609 | VS2-N...5... | 1.298 | 0.997 | 1.297 | 3 | 1 | 0.0003 |
| 275 | 617 | VS2-O...13.. | 0.097 | 0.100 | 0.096 | 3 | 1 | 0.0003 |
| 276 | 620 | VS2-O...16.. | 5.702 | 4.729 | 5.075 | 3 | 0 | 1 |
| 277 | 637 | VS2-S...14.. | 5.746 | 5.073 | 5.099 | 3 | 0 | 1 |
| 278 | 642 | VS2-S...7... | 4.671 | 4.497 | 4.795 | 3 | 1 | 0.0003 |
| 279 | 654 | [...(...[... | 2.803 | 2.805 | 2.946 | 3 | 1 | 0.0003 |
| 280 | 660 | [...1...(... | 1.197 | 1.053 | 0.953 | 3 | 0 | 1 |
| 281 | 672 | [...=...O... | 1.501 | 1.001 | 0.945 | 3 | 0 | 1 |
| 282 | 698 | [...S...(... | 2.353 | 2.352 | 2.200 | 3 | 2 | 0.0067 |
| 283 | 701 | [...S...[... | 1.304 | 1.303 | 1.300 | 3 | 0 | 1 |
| 284 | 708 | \...C...(... | 3.048 | 2.801 | 3.048 | 3 | 1 | 0.0003 |
| 285 | 14 | ++++F---B3== | 0.902 | 1.301 | 0.703 | 2 | 0 | 1 |
| 286 | 15 | ++++F---Br== | 4.429 | 4.903 | 5.203 | 2 | 0 | 1 |
| 287 | 20 | ++++CL--Br== | 5.003 | 5.696 | 5.602 | 2 | 0 | 1 |
| 288 | 26 | ++++Br--B3== | 3.902 | 3.530 | 3.347 | 2 | 2 | 0.0109 |
| 289 | 34 | ++++Cl--B3== | 0.705 | 0.695 | 1.304 | 2 | 1 | 0.0038 |
| 290 | 56 | 1...(...(... | 5.802 | 5.798 | 5.796 | 2 | 0 | 1 |
| 291 | 77 | 2...O...(... | 4.196 | 4.900 | 4.664 | 2 | 1 | 0.0038 |
| 292 | 85 | 3...C...1... | 3.704 | 3.151 | 4.096 | 2 | 0 | 1 |
| 293 | 87 | 3...Cl..(... | 3.176 | 3.796 | 3.515 | 2 | 0 | 1 |
| 294 | 91 | =...(...(... | 3.396 | 3.701 | 3.100 | 2 | 0 | 1 |
| 295 | 130 | C.../...C... | 6.613 | 6.801 | 7.000 | 2 | 0 | 1 |
| 296 | 174 | C...O.../... | 4.896 | 4.901 | 4.902 | 2 | 1 | 0.0038 |
| 297 | 179 | C...S...1... | 2.799 | 2.796 | 2.796 | 2 | 0 | 1 |
| 298 | 184 | C...\...C... | 3.602 | 3.997 | 3.898 | 2 | 1 | 0.0038 |
| 299 | 232 | EC2-C...9... | 0.871 | 0.805 | 0.949 | 2 | 1 | 0.0038 |
| 300 | 235 | EC2-F...7... | 0.199 | 0.605 | 0.197 | 2 | 0 | 1 |
| 301 | 282 | EC2-N...29.. | 1.901 | 1.897 | 1.901 | 2 | 1 | 0.0038 |
| 302 | 297 | EC2-O...22.. | 6.995 | 7.003 | 7.000 | 2 | 1 | 0.0038 |
| 303 | 307 | EC2-P...31.. | 6.697 | 6.097 | 5.854 | 2 | 0 | 1 |
| 304 | 308 | EC2-S...11.. | 4.197 | 4.601 | 4.597 | 2 | 1 | 0.0038 |
| 305 | 309 | EC2-S...12.. | 2.501 | 2.496 | 2.497 | 2 | 1 | 0.0038 |
| 306 | 314 | EC2-S...17.. | 4.597 | 4.902 | 5.103 | 2 | 0 | 1 |
| 307 | 318 | EC2-S...24.. | 0.798 | 0.997 | 0.398 | 2 | 0 | 1 |
| 308 | 334 | Br..(...Br.. | 1.423 | 1.754 | 1.574 | 2 | 1 | 0.0038 |
| 309 | 350 | Cl..(...3... | 3.702 | 3.097 | 3.453 | 2 | 2 | 0.0109 |
| 310 | 361 | Cl..3....... | 3.803 | 3.700 | 3.522 | 2 | 0 | 1 |
| 311 | 362 | Cl..3...C... | 3.697 | 3.255 | 3.552 | 2 | 0 | 1 |
| 312 | 372 | N...(...1... | 2.502 | 2.796 | 2.498 | 2 | 2 | 0.0109 |
| 313 | 373 | N...(...2... | 4.004 | 3.779 | 3.696 | 2 | 0 | 1 |
| 314 | 377 | N...(...N... | 1.898 | 2.221 | 2.204 | 2 | 1 | 0.0038 |
| 315 | 395 | N...=...2... | 0.596 | 0.099 | 0.225 | 2 | 0 | 1 |
| 316 | 434 | O.../....... | 4.900 | 4.599 | 4.900 | 2 | 1 | 0.0038 |
| 317 | 435 | O.../...N... | 4.896 | 4.950 | 4.899 | 2 | 1 | 0.0038 |
| 318 | 436 | O...1...(... | 5.752 | 5.797 | 5.149 | 2 | 0 | 1 |
| 319 | 441 | O...2...C... | 1.596 | 2.053 | 1.128 | 2 | 0 | 1 |
| 320 | 451 | O...C...3... | 2.200 | 1.898 | 1.597 | 2 | 0 | 1 |
| 321 | 452 | O...C...=... | 4.901 | 4.903 | 5.200 | 2 | 0 | 1 |
| 322 | 502 | NNC-S...220. | 6.998 | 6.996 | 6.998 | 2 | 2 | 0.0109 |
| 323 | 534 | S...[...2... | 3.896 | 3.549 | 3.297 | 2 | 0 | 1 |
| 324 | 535 | S...[...=... | 3.276 | 3.058 | 3.224 | 2 | 0 | 1 |
| 325 | 566 | VS2-F...6... | 0.048 | 0.314 | 0.301 | 2 | 0 | 1 |
| 326 | 605 | VS2-N...20.. | 1.899 | 1.895 | 1.601 | 2 | 1 | 0.0038 |
| 327 | 608 | VS2-N...4... | 2.099 | 2.097 | 2.096 | 2 | 2 | 0.0109 |
| 328 | 621 | VS2-O...18.. | 7.000 | 7.003 | 6.998 | 2 | 1 | 0.0038 |
| 329 | 631 | VS2-P...15.. | 5.320 | 6.052 | 6.246 | 2 | 0 | 1 |
| 330 | 650 | [...(...Cl.. | 2.204 | 1.802 | 1.600 | 2 | 2 | 0.0109 |
| 331 | 676 | [...C...(... | 4.728 | 4.896 | 4.896 | 2 | 2 | 0.0109 |
| 332 | 688 | [...N...=... | 2.201 | 2.103 | 2.203 | 2 | 0 | 1 |
| 333 | 694 | [...O...1... | 2.987 | 3.097 | 3.802 | 2 | 1 | 0.0038 |
| 334 | 695 | [...O...C... | 4.285 | 3.598 | 3.374 | 2 | 0 | 1 |
| 1 | 127 | C........... | -1.226 | -1.202 | -1.349 | 96 | 31 | 0 |
| 2 | 3 | (........... | -1.101 | -1.072 | -0.986 | 94 | 31 | 0.0002 |
| 3 | 475 | NNC-H...101. | -0.771 | -0.609 | -0.733 | 94 | 30 | 0.0001 |
| 4 | 119 | C...(....... | -1.321 | -1.429 | -1.376 | 93 | 31 | 0.0003 |
| 5 | 95 | =........... | -1.479 | -1.313 | -1.358 | 90 | 31 | 0.0005 |
| 6 | 470 | NNC-C...404. | -0.499 | -0.328 | -0.377 | 85 | 24 | 0.001 |
| 7 | 433 | O........... | -0.215 | -0.346 | -0.396 | 84 | 29 | 0.0005 |
| 8 | 58 | 1........... | -1.896 | -1.900 | -1.953 | 82 | 25 | 0.0004 |
| 9 | 132 | C...1....... | -1.546 | -1.551 | -1.677 | 82 | 25 | 0.0004 |
| 10 | 159 | C...C....... | -0.976 | -0.890 | -1.039 | 75 | 22 | 0.0007 |
| 11 | 125 | C...(...=... | -0.495 | -0.604 | -0.277 | 72 | 23 | 0.0001 |
| 12 | 421 | O...(....... | -0.423 | -0.572 | -0.298 | 72 | 27 | 0.0012 |
| 13 | 379 | N........... | -1.513 | -1.636 | -1.708 | 68 | 23 | 0.0004 |
| 14 | 5 | (...C...(... | -1.427 | -1.154 | -1.466 | 66 | 19 | 0.0009 |
| 15 | 37 | ++++N---O=== | -0.449 | -0.111 | -0.065 | 62 | 21 | 0.0004 |
| 16 | 96 | =...1....... | -0.701 | -0.504 | -0.746 | 58 | 23 | 0.0017 |
| 17 | 134 | C...1...=... | -0.424 | -0.150 | -0.148 | 58 | 23 | 0.0017 |
| 18 | 540 | VS2-C...11.. | -0.746 | -0.559 | -0.810 | 56 | 16 | 0.0009 |
| 19 | 563 | VS2-C...9... | -0.172 | -0.176 | -0.203 | 52 | 18 | 0.0006 |
| 20 | 209 | EC2-C...22.. | -0.796 | -0.697 | -0.724 | 50 | 15 | 0.0006 |
| 21 | 471 | NNC-C...413. | -1.522 | -1.527 | -1.521 | 50 | 16 | 0.0001 |
| 22 | 371 | N...(....... | -0.888 | -1.002 | -0.712 | 48 | 16 | 0.0003 |
| 23 | 204 | EC2-C...17.. | -0.775 | -0.788 | -0.821 | 44 | 14 | 0.0001 |
| 24 | 163 | C...C...=... | -2.002 | -1.953 | -1.999 | 43 | 15 | 0.0006 |
| 25 | 206 | EC2-C...19.. | -0.453 | -0.340 | -0.329 | 42 | 18 | 0.0024 |
| 26 | 33 | ++++Cl--B2== | -1.752 | -1.813 | -1.998 | 41 | 15 | 0.001 |
| 27 | 659 | [........... | -0.314 | -0.316 | -0.250 | 41 | 16 | 0.0016 |
| 28 | 487 | NNC-N...303. | -1.408 | -1.538 | -1.654 | 40 | 12 | 0.0006 |
| 29 | 205 | EC2-C...18.. | -1.772 | -1.554 | -1.508 | 39 | 17 | 0.0025 |
| 30 | 374 | N...(...C... | -1.623 | -1.666 | -1.488 | 37 | 8 | 0.0028 |
| 31 | 22 | ++++CL--O=== | -1.676 | -1.327 | -1.599 | 36 | 15 | 0.0021 |
| 32 | 71 | 2........... | -0.395 | -0.340 | -0.878 | 36 | 10 | 0.0011 |
| 33 | 542 | VS2-C...13.. | -1.951 | -1.954 | -1.903 | 33 | 12 | 0.001 |
| 34 | 646 | [...(....... | -0.217 | -0.112 | -0.241 | 33 | 14 | 0.0023 |
| 35 | 624 | VS2-O...5... | -1.348 | -0.912 | -1.199 | 31 | 10 | 0 |
| 36 | 393 | N...=....... | -1.458 | -1.226 | -1.359 | 28 | 11 | 0.0016 |
| 37 | 213 | EC2-C...26.. | -1.297 | -1.525 | -1.279 | 27 | 5 | 0.0037 |
| 38 | 21 | ++++CL--N=== | -1.597 | -1.754 | -1.601 | 23 | 9 | 0.0016 |
| 39 | 353 | Cl..(...C... | -0.578 | -0.401 | -1.045 | 21 | 6 | 0.0009 |
| 40 | 215 | EC2-C...28.. | -0.592 | -0.614 | -1.139 | 19 | 4 | 0.003 |
| 41 | 121 | C...(...1... | -0.225 | -0.951 | -0.545 | 16 | 4 | 0.0019 |
| 42 | 216 | EC2-C...29.. | -1.949 | -1.996 | -2.000 | 16 | 5 | 0.0003 |
| 43 | 218 | EC2-C...31.. | -0.471 | -0.598 | -1.423 | 13 | 1 | 0.0074 |
| 44 | 596 | VS2-N...12.. | -1.954 | -1.201 | -1.647 | 11 | 2 | 0.0039 |
| 45 | 613 | VS2-N...9... | -0.697 | -0.550 | -0.496 | 11 | 4 | 0.001 |
| 46 | 238 | EC2-H...11.. | -1.524 | -1.308 | -1.851 | 10 | 5 | 0.0038 |
| 47 | 472 | NNC-C...422. | -1.500 | -1.702 | -1.851 | 10 | 3 | 0.0006 |
| 48 | 568 | VS2-H...10.. | -1.433 | -1.555 | -1.395 | 10 | 5 | 0.0038 |
| 49 | 66 | 1...N...(... | -2.004 | -1.999 | -1.916 | 9 | 1 | 0.0061 |
| 50 | 80 | 3...(....... | -1.048 | -0.704 | -0.675 | 9 | 3 | 0.0003 |
| 51 | 161 | C...C...2... | -0.725 | -1.495 | -0.403 | 9 | 3 | 0.0003 |
| 52 | 50 | -........... | -0.673 | -0.521 | -0.149 | 8 | 4 | 0.0038 |
| 53 | 240 | EC2-H...13.. | -0.701 | -0.804 | -1.075 | 8 | 2 | 0.0019 |
| 54 | 265 | EC2-N...13.. | -1.150 | -1.324 | -1.179 | 8 | 4 | 0.0038 |
| 55 | 289 | EC2-O...13.. | -0.714 | -0.802 | -0.745 | 8 | 3 | 0.0012 |
| 56 | 294 | EC2-O...18.. | -1.735 | -1.536 | -1.822 | 8 | 2 | 0.0019 |
| 57 | 429 | O...(...Cl.. | -1.504 | -1.400 | -1.947 | 8 | 3 | 0.0012 |
| 58 | 456 | O...[...(... | -0.098 | -0.886 | -0.199 | 8 | 4 | 0.0038 |
| 59 | 485 | NNC-N...211. | -0.778 | -0.999 | -1.051 | 8 | 4 | 0.0038 |
| 60 | 570 | VS2-H...12.. | -0.038 | -1.441 | -0.264 | 8 | 2 | 0.0019 |
| 61 | 618 | VS2-O...14.. | -1.402 | -1.426 | -1.278 | 8 | 2 | 0.0019 |
| 62 | 628 | VS2-O...9... | -0.699 | -0.600 | -0.504 | 8 | 3 | 0.0012 |
| 63 | 657 | [...-....... | -1.042 | -0.479 | -0.405 | 8 | 4 | 0.0038 |
| 64 | 658 | [...-...O... | -0.829 | -0.267 | -0.974 | 8 | 4 | 0.0038 |
| 65 | 681 | [...H....... | -0.085 | -0.328 | -0.458 | 8 | 1 | 0.0057 |
| 66 | 692 | [...O...-... | -0.799 | -0.074 | -0.639 | 8 | 4 | 0.0038 |
| 67 | 13 | ++++F---B2== | -1.996 | -1.349 | -1.998 | 7 | 2 | 0.0009 |
| 68 | 17 | ++++F---N=== | -1.783 | -1.602 | -1.620 | 7 | 2 | 0.0009 |
| 69 | 48 | ........... | -0.749 | -0.891 | -0.574 | 7 | 4 | 0.0051 |
| 70 | 49 | ...[...(... | -1.235 | -1.095 | -0.754 | 7 | 4 | 0.0051 |
| 71 | 191 | F...(....... | -0.026 | -0.117 | -0.233 | 7 | 2 | 0.0009 |
| 72 | 193 | F...(...C... | -0.496 | -0.573 | -0.501 | 7 | 2 | 0.0009 |
| 73 | 195 | F........... | -0.724 | -0.779 | -1.338 | 7 | 2 | 0.0009 |
| 74 | 378 | N.......... | -1.170 | -0.897 | -1.264 | 7 | 4 | 0.0051 |
| 75 | 655 | [.......... | -1.024 | -1.301 | -0.849 | 7 | 4 | 0.0051 |
| 76 | 656 | [......N... | -1.223 | -1.245 | -0.701 | 7 | 4 | 0.0051 |
| 77 | 686 | [...N...... | -0.903 | -1.033 | -1.276 | 7 | 4 | 0.0051 |
| 78 | 177 | C...O...C... | -0.201 | -0.503 | -0.174 | 6 | 3 | 0.0038 |
| 79 | 327 | H...[...(... | -1.799 | -1.621 | -1.228 | 6 | 1 | 0.0043 |
| 80 | 348 | Cl..(...1... | -1.597 | -1.102 | -1.497 | 6 | 0 | 1 |
| 81 | 515 | S...1...(... | -1.629 | -1.705 | -1.659 | 6 | 2 | 0.0003 |
| 82 | 516 | S...1....... | -1.800 | -1.366 | -1.852 | 6 | 2 | 0.0003 |
| 83 | 36 | ++++N---B3== | -1.221 | -0.403 | -0.749 | 5 | 3 | 0.0056 |
| 84 | 98 | =...3....... | -0.990 | -1.048 | -1.185 | 5 | 3 | 0.0056 |
| 85 | 143 | C...3...=... | -1.624 | -0.895 | -1.237 | 5 | 3 | 0.0056 |
| 86 | 152 | C...=...3... | -1.150 | -0.295 | -0.955 | 5 | 2 | 0.0018 |
| 87 | 221 | EC2-C...34.. | -1.202 | -1.673 | -0.624 | 5 | 1 | 0.0033 |
| 88 | 287 | EC2-O...11.. | -1.904 | -1.672 | -1.948 | 5 | 2 | 0.0018 |
| 89 | 394 | N...=...1... | -2.002 | -1.999 | -2.001 | 5 | 1 | 0.0033 |
| 90 | 444 | O...=...1... | -1.874 | -1.999 | -1.947 | 5 | 0 | 1 |
| 91 | 643 | VS2-S...8... | -0.698 | -0.602 | -0.204 | 5 | 1 | 0.0033 |
| 92 | 707 | \........... | -1.921 | -1.605 | -1.627 | 5 | 1 | 0.0033 |
| 93 | 4 | (...C...#... | -1.871 | -1.904 | -1.804 | 4 | 2 | 0.0038 |
| 94 | 16 | ++++F---Cl== | -0.205 | -0.503 | -1.047 | 4 | 2 | 0.0038 |
| 95 | 227 | EC2-C...40.. | -0.098 | -1.411 | -1.078 | 4 | 1 | 0.0019 |
| 96 | 234 | EC2-F...6... | -1.247 | -1.166 | -1.211 | 4 | 2 | 0.0038 |
| 97 | 263 | EC2-N...11.. | -1.701 | -1.996 | -1.997 | 4 | 2 | 0.0038 |
| 98 | 271 | EC2-N...19.. | -1.800 | -1.802 | -1.996 | 4 | 3 | 0.0079 |
| 99 | 312 | EC2-S...15.. | -1.326 | -0.597 | -0.754 | 4 | 2 | 0.0038 |
| 100 | 331 | Br..(....... | -1.090 | -0.822 | -0.724 | 4 | 2 | 0.0038 |
| 101 | 401 | N...C...1... | -2.004 | -1.997 | -2.004 | 4 | 2 | 0.0038 |
| 102 | 549 | VS2-C...20.. | -1.998 | -1.997 | -2.001 | 4 | 2 | 0.0038 |
| 103 | 553 | VS2-C...24.. | -1.026 | -0.687 | -0.646 | 4 | 1 | 0.0019 |
| 104 | 565 | VS2-F...5... | -1.240 | -1.162 | -1.604 | 4 | 2 | 0.0038 |
| 105 | 634 | VS2-S...11.. | -1.150 | -1.077 | -1.022 | 4 | 2 | 0.0038 |
| 106 | 651 | [...(...N... | -0.802 | -1.096 | -0.802 | 4 | 3 | 0.0079 |
| 107 | 53 | /...(....... | -0.796 | -0.749 | -0.503 | 3 | 0 | 1 |
| 108 | 55 | /...C...(... | -0.595 | -0.748 | -0.817 | 3 | 0 | 1 |
| 109 | 94 | =...(...2... | -1.604 | -1.491 | -1.775 | 3 | 1 | 0.0003 |
| 110 | 115 | @@.......... | -1.698 | -1.996 | -1.575 | 3 | 1 | 0.0003 |
| 111 | 128 | C.../...(... | -0.797 | -0.797 | -0.798 | 3 | 0 | 1 |
| 112 | 156 | C...@@...... | -1.827 | -1.872 | -1.789 | 3 | 1 | 0.0003 |
| 113 | 170 | C...N...2... | -0.970 | -1.150 | -1.217 | 3 | 0 | 1 |
| 114 | 183 | C...\...(... | -1.871 | -1.878 | -1.903 | 3 | 0 | 1 |
| 115 | 189 | BOND11100000 | -1.226 | -1.521 | -1.928 | 3 | 1 | 0.0003 |
| 116 | 220 | EC2-C...33.. | -2.004 | -1.996 | -2.003 | 3 | 0 | 1 |
| 117 | 225 | EC2-C...38.. | -1.897 | -1.995 | -1.999 | 3 | 0 | 1 |
| 118 | 226 | EC2-C...39.. | -0.012 | -1.614 | -0.671 | 3 | 0 | 1 |
| 119 | 262 | EC2-Cl..8... | -0.753 | -1.050 | -0.962 | 3 | 1 | 0.0003 |
| 120 | 279 | EC2-N...27.. | -0.998 | -0.197 | -0.950 | 3 | 1 | 0.0003 |
| 121 | 325 | H...@@...... | -1.952 | -1.548 | -1.601 | 3 | 1 | 0.0003 |
| 122 | 326 | H...@@..C... | -1.700 | -1.903 | -1.904 | 3 | 1 | 0.0003 |
| 123 | 388 | N...2...(... | -0.821 | -0.601 | -0.678 | 3 | 2 | 0.0067 |
| 124 | 450 | O...C...2... | -1.996 | -1.874 | -1.999 | 3 | 2 | 0.0067 |
| 125 | 465 | NNC-C...211. | -1.002 | -1.100 | -0.798 | 3 | 1 | 0.0003 |
| 126 | 478 | NNC-Br..110. | -0.300 | -0.902 | -0.901 | 3 | 1 | 0.0003 |
| 127 | 501 | NNC-S...211. | -2.001 | -2.001 | -1.996 | 3 | 2 | 0.0067 |
| 128 | 511 | S...(...N... | -1.997 | -2.002 | -2.001 | 3 | 0 | 1 |
| 129 | 551 | VS2-C...22.. | -1.652 | -1.999 | -2.000 | 3 | 0 | 1 |
| 130 | 552 | VS2-C...23.. | -1.950 | -1.954 | -1.898 | 3 | 0 | 1 |
| 131 | 592 | VS2-Cl..7... | -1.046 | -1.050 | -0.900 | 3 | 1 | 0.0003 |
| 132 | 603 | VS2-N...18.. | -0.760 | -0.003 | -1.200 | 3 | 1 | 0.0003 |
| 133 | 680 | [...C...@@.. | -1.805 | -1.353 | -1.999 | 3 | 1 | 0.0003 |
| 134 | 683 | [...H...@@.. | -1.649 | -1.470 | -1.802 | 3 | 1 | 0.0003 |
| 135 | 704 | \...(....... | -1.950 | -1.650 | -1.848 | 3 | 0 | 1 |
| 136 | 67 | 1...O...(... | -1.104 | -1.199 | -1.396 | 2 | 0 | 1 |
| 137 | 68 | 1...S...(... | -1.998 | -2.004 | -2.004 | 2 | 2 | 0.0109 |
| 138 | 76 | 2...Cl..(... | -2.005 | -1.652 | -2.003 | 2 | 1 | 0.0038 |
| 139 | 88 | 4...(....... | -1.001 | -0.650 | -0.200 | 2 | 0 | 1 |
| 140 | 105 | =...C...4... | -0.804 | -1.029 | -0.852 | 2 | 0 | 1 |
| 141 | 168 | C...N.../... | -0.903 | -1.100 | -1.095 | 2 | 0 | 1 |
| 142 | 172 | C...N...C... | -1.729 | -2.001 | -2.003 | 2 | 0 | 1 |
| 143 | 242 | EC2-H...3... | -1.948 | -1.998 | -1.804 | 2 | 0 | 1 |
| 144 | 295 | EC2-O...19.. | -2.004 | -1.999 | -2.001 | 2 | 1 | 0.0038 |
| 145 | 359 | Cl..2....... | -1.995 | -2.005 | -2.002 | 2 | 1 | 0.0038 |
| 146 | 360 | Cl..2...C... | -1.998 | -2.002 | -1.998 | 2 | 1 | 0.0038 |
| 147 | 415 | N...[...2... | -1.997 | -2.001 | -2.003 | 2 | 1 | 0.0038 |
| 148 | 457 | O...[...C... | -1.625 | -1.648 | -1.405 | 2 | 1 | 0.0038 |
| 149 | 474 | NNC-F...110. | -1.950 | -1.748 | -1.925 | 2 | 0 | 1 |
| 150 | 518 | S...2....... | -0.600 | -0.297 | -0.722 | 2 | 1 | 0.0038 |
| 151 | 528 | S...N....... | -1.010 | -1.047 | -1.397 | 2 | 1 | 0.0038 |
| 152 | 572 | VS2-H...2... | -1.904 | -1.952 | -1.999 | 2 | 0 | 1 |
| 153 | 619 | VS2-O...15.. | -2.003 | -1.998 | -2.000 | 2 | 1 | 0.0038 |
| 154 | 706 | \...(...[... | -1.046 | -1.697 | -0.348 | 2 | 0 | 1 |
| 155 | 711 | \...C...=... | -1.761 | -1.759 | -1.850 | 2 | 0 | 1 |

**Table S5** Attributes with undefined contributions in the best model

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | ID | SAk | Probe 1 | Probe 2 | Probe 2 | NSs | NSv | Defect[SAk] |
| 1 | 466 | NNC-C...303. | -0.048 | -0.173 | 0.052 | 87 | 26 | 0.0006 |
| 2 | 92 | =...(....... | 0.123 | 0.251 | -0.097 | 85 | 28 | 0.0002 |
| 3 | 208 | EC2-C...21.. | 0.397 | 0.410 | -0.235 | 50 | 16 | 0.0001 |
| 4 | 427 | O...(...C... | 0.303 | -0.096 | 0.241 | 49 | 14 | 0.0009 |
| 5 | 247 | EC2-H...8... | 0.309 | -0.022 | 0.235 | 47 | 11 | 0.0023 |
| 6 | 577 | VS2-H...7... | -0.160 | 0.285 | 0.284 | 47 | 11 | 0.0023 |
| 7 | 301 | EC2-O...7... | 0.215 | -0.245 | 0.026 | 29 | 8 | 0.0012 |
| 8 | 495 | NNC-O...211. | 0.002 | -0.346 | -0.016 | 27 | 6 | 0.0027 |
| 9 | 626 | VS2-O...7... | -0.327 | -0.585 | 0.352 | 22 | 6 | 0.0013 |
| 10 | 559 | VS2-C...5... | 0.073 | -0.247 | 0.238 | 20 | 6 | 0.0006 |
| 11 | 453 | O...C...C... | -0.052 | -0.205 | 0.302 | 19 | 7 | 0.0011 |
| 12 | 171 | C...N...=... | 0.087 | -0.078 | -0.286 | 18 | 7 | 0.0015 |
| 13 | 10 | (...N...(... | 0.346 | 0.296 | -0.559 | 17 | 6 | 0.0007 |
| 14 | 392 | N...=...(... | -0.024 | -0.313 | 0.111 | 14 | 4 | 0.0009 |
| 15 | 160 | C...C...1... | -0.913 | 0.200 | 0.090 | 12 | 2 | 0.0043 |
| 16 | 142 | C...3....... | -0.033 | 0.348 | 0.254 | 11 | 4 | 0.001 |
| 17 | 141 | C...3...(... | -1.225 | 0.053 | -0.815 | 9 | 2 | 0.0027 |
| 18 | 24 | ++++CL--S=== | 0.097 | -0.198 | 0.105 | 8 | 3 | 0.0012 |
| 19 | 322 | H........... | -0.822 | 0.048 | -0.626 | 8 | 1 | 0.0057 |
| 20 | 357 | Cl..1....... | -0.196 | 0.254 | -0.003 | 8 | 4 | 0.0038 |
| 21 | 358 | Cl..1...C... | -0.004 | 0.396 | 0.204 | 8 | 3 | 0.0012 |
| 22 | 413 | N...[...(... | -0.203 | 0.400 | 0.100 | 8 | 4 | 0.0038 |
| 23 | 432 | O...-....... | 0.125 | -0.951 | -0.801 | 8 | 4 | 0.0038 |
| 24 | 108 | =...N...1... | 0.178 | -0.197 | 0.297 | 7 | 2 | 0.0009 |
| 25 | 546 | VS2-C...17.. | -0.152 | 0.400 | 0.229 | 7 | 0 | 1 |
| 26 | 18 | ++++F---O=== | 0.646 | -0.496 | 0.067 | 6 | 2 | 0.0003 |
| 27 | 389 | N...2....... | -0.104 | -0.296 | 0.698 | 6 | 4 | 0.0067 |
| 28 | 7 | (...Br..(... | -0.562 | 0.021 | 0.110 | 4 | 2 | 0.0038 |
| 29 | 222 | EC2-C...35.. | -0.249 | -0.073 | 0.224 | 4 | 1 | 0.0019 |
| 30 | 332 | Br..(...C... | 0.352 | 0.249 | -0.199 | 4 | 0 | 1 |
| 31 | 548 | VS2-C...19.. | -0.178 | 0.023 | -0.500 | 4 | 1 | 0.0019 |
| 32 | 665 | [...2....... | 0.302 | -0.095 | 0.605 | 4 | 1 | 0.0019 |
| 33 | 9 | (...N...#... | 0.301 | 0.655 | -0.050 | 3 | 3 | 0.0109 |
| 34 | 109 | =...N...2... | -0.200 | -0.145 | 0.534 | 3 | 3 | 0.0109 |
| 35 | 293 | EC2-O...17.. | 0.103 | 0.100 | -0.502 | 3 | 1 | 0.0003 |
| 36 | 368 | N...#....... | -0.354 | 0.125 | -0.272 | 3 | 3 | 0.0109 |
| 37 | 369 | N...#...C... | -0.146 | 0.402 | -0.497 | 3 | 3 | 0.0109 |
| 38 | 482 | NNC-N...101. | -0.164 | -0.062 | 0.104 | 3 | 1 | 0.0003 |
| 39 | 89 | 4........... | 0.250 | -0.563 | -1.196 | 2 | 0 | 1 |
| 40 | 99 | =...4....... | -0.698 | 0.497 | -0.800 | 2 | 0 | 1 |
| 41 | 145 | C...4...(... | 0.302 | -0.547 | -1.097 | 2 | 0 | 1 |
| 42 | 146 | C...4....... | 0.323 | -0.102 | 0.311 | 2 | 0 | 1 |
| 43 | 147 | C...4...=... | -1.041 | -0.073 | 0.504 | 2 | 0 | 1 |
| 44 | 153 | C...=...4... | -0.564 | 0.465 | -0.398 | 2 | 0 | 1 |
| 45 | 390 | N...2...C... | -0.122 | 0.600 | 0.799 | 2 | 1 | 0.0038 |
| 1 | 19 | ++++F---S=== | 0 | 0 | 0 | 1 | 0 | 0 |
| 2 | 29 | ++++I---B2== | 0 | 0 | 0 | 1 | 0 | 0 |
| 3 | 30 | ++++I---B3== | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 31 | ++++I---N=== | 0 | 0 | 0 | 1 | 0 | 0 |
| 5 | 32 | ++++I---O=== | 0 | 0 | 0 | 1 | 0 | 0 |
| 6 | 46 | ++++S---B3== | 0 | 0 | 0 | 1 | 0 | 0 |
| 7 | 52 | /...(...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 8 | 59 | 1.../....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 9 | 60 | 1...3...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 10 | 63 | 1...C.../... | 0 | 0 | 0 | 1 | 0 | 0 |
| 11 | 64 | 1...F...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 12 | 74 | 2...C...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 13 | 75 | 2...Br..(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 14 | 78 | 2...S...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 15 | 79 | 3...(...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 16 | 82 | 3...1....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 17 | 86 | 3...C...2... | 0 | 0 | 0 | 1 | 0 | 0 |
| 18 | 90 | 4...C...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 19 | 114 | @...[...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 20 | 120 | C...(.../... | 0 | 0 | 0 | 1 | 0 | 0 |
| 21 | 124 | C...(...4... | 0 | 0 | 0 | 1 | 0 | 0 |
| 22 | 133 | C...1.../... | 0 | 0 | 0 | 1 | 0 | 0 |
| 23 | 165 | C...I...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 24 | 180 | C...S...2... | 0 | 0 | 0 | 1 | 0 | 0 |
| 25 | 181 | C...S...C... | 0 | 0 | 0 | 1 | 1 | 0 |
| 26 | 192 | F...(...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 27 | 196 | F...1....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 28 | 197 | F...1...C... | 0 | 0 | 0 | 1 | 0 | 0 |
| 29 | 199 | EC2-C...12.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 30 | 200 | EC2-C...13.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 31 | 223 | EC2-C...36.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 32 | 229 | EC2-C...42.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 33 | 230 | EC2-C...43.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 34 | 233 | EC2-F...5... | 0 | 0 | 0 | 1 | 0 | 0 |
| 35 | 236 | EC2-F...8... | 0 | 0 | 0 | 1 | 0 | 0 |
| 36 | 241 | EC2-H...14.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 37 | 249 | EC2-Br..10.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 38 | 250 | EC2-Br..6... | 0 | 0 | 0 | 1 | 0 | 0 |
| 39 | 252 | EC2-Br..8... | 0 | 0 | 0 | 1 | 0 | 0 |
| 40 | 253 | EC2-I...3... | 0 | 0 | 0 | 1 | 0 | 0 |
| 41 | 255 | EC2-Cl..11.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 42 | 257 | EC2-Cl..13.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 43 | 258 | EC2-Cl..4... | 0 | 0 | 0 | 1 | 0 | 0 |
| 44 | 280 | EC2-N...3... | 0 | 0 | 0 | 1 | 0 | 0 |
| 45 | 283 | EC2-N...4... | 0 | 0 | 0 | 1 | 1 | 0 |
| 46 | 284 | EC2-N...5... | 0 | 0 | 0 | 1 | 2 | 0 |
| 47 | 285 | EC2-N...54.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 48 | 298 | EC2-O...23.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 49 | 313 | EC2-S...16.. | 0 | 0 | 0 | 1 | 1 | 0 |
| 50 | 316 | EC2-S...19.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 51 | 317 | EC2-S...20.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 52 | 320 | EC2-S...8... | 0 | 0 | 0 | 1 | 0 | 0 |
| 53 | 321 | EC2-S...9... | 0 | 0 | 0 | 1 | 0 | 0 |
| 54 | 329 | H...[...3... | 0 | 0 | 0 | 1 | 0 | 0 |
| 55 | 330 | Br..(...(... | 0 | 0 | 0 | 1 | 1 | 0 |
| 56 | 333 | Br..(...F... | 0 | 0 | 0 | 1 | 0 | 0 |
| 57 | 336 | Br..2....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 58 | 337 | Br..2...C... | 0 | 0 | 0 | 1 | 0 | 0 |
| 59 | 338 | Br..C....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 60 | 339 | Br..C...C... | 0 | 0 | 0 | 1 | 0 | 0 |
| 61 | 340 | Br..N...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 62 | 341 | I...(....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 63 | 342 | I...(...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 64 | 343 | I........... | 0 | 0 | 0 | 1 | 0 | 0 |
| 65 | 344 | I...C...#... | 0 | 0 | 0 | 1 | 0 | 0 |
| 66 | 345 | I...C....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 67 | 351 | Cl..(...4... | 0 | 0 | 0 | 1 | 0 | 0 |
| 68 | 365 | Cl..C...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 69 | 366 | Cl..C...C... | 0 | 0 | 0 | 1 | 1 | 0 |
| 70 | 367 | Cl..O...C... | 0 | 0 | 0 | 1 | 0 | 0 |
| 71 | 370 | N...(...(... | 0 | 0 | 0 | 1 | 1 | 0 |
| 72 | 381 | N.../...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 73 | 382 | N.../...N... | 0 | 0 | 0 | 1 | 0 | 0 |
| 74 | 387 | N...1...N... | 0 | 0 | 0 | 1 | 0 | 0 |
| 75 | 405 | N...H....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 76 | 406 | N...Br...... | 0 | 0 | 0 | 1 | 0 | 0 |
| 77 | 410 | N...S...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 78 | 411 | N...S...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 79 | 417 | N...[...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 80 | 422 | O...(.../... | 0 | 0 | 0 | 1 | 0 | 0 |
| 81 | 425 | O...(...3... | 0 | 0 | 0 | 1 | 0 | 0 |
| 82 | 445 | O...=...2... | 0 | 0 | 0 | 1 | 0 | 0 |
| 83 | 455 | O...Cl...... | 0 | 0 | 0 | 1 | 0 | 0 |
| 84 | 458 | O...\...N... | 0 | 0 | 0 | 1 | 0 | 0 |
| 85 | 459 | O...|...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 86 | 462 | P...[...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 87 | 469 | NNC-C...330. | 0 | 0 | 0 | 1 | 1 | 0 |
| 88 | 479 | NNC-I...101. | 0 | 0 | 0 | 1 | 0 | 0 |
| 89 | 490 | NNC-N...330. | 0 | 0 | 0 | 1 | 0 | 0 |
| 90 | 498 | NNC-P...413. | 0 | 0 | 0 | 1 | 1 | 0 |
| 91 | 503 | NNC-S...303. | 0 | 0 | 0 | 1 | 0 | 0 |
| 92 | 504 | NNC-S...404. | 0 | 0 | 0 | 1 | 0 | 0 |
| 93 | 505 | NNC-S...413. | 0 | 0 | 0 | 1 | 0 | 0 |
| 94 | 508 | S...(...1... | 0 | 0 | 0 | 1 | 1 | 0 |
| 95 | 510 | S...(...Cl.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 96 | 513 | S...(...S... | 0 | 0 | 0 | 1 | 1 | 0 |
| 97 | 521 | S...C...#... | 0 | 0 | 0 | 1 | 0 | 0 |
| 98 | 524 | S...C...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 99 | 526 | S...C...S... | 0 | 0 | 0 | 1 | 0 | 0 |
| 100 | 527 | S...N...(... | 0 | 0 | 0 | 1 | 1 | 0 |
| 101 | 529 | S...N...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 102 | 530 | S...S...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 103 | 531 | S...S....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 104 | 532 | S...S...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 105 | 533 | S...[...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 106 | 537 | Sn.......... | 0 | 0 | 0 | 1 | 0 | 0 |
| 107 | 538 | Sn..[...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 108 | 555 | VS2-C...26.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 109 | 556 | VS2-C...27.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 110 | 564 | VS2-F...4... | 0 | 0 | 0 | 1 | 0 | 0 |
| 111 | 567 | VS2-F...7... | 0 | 0 | 0 | 1 | 0 | 0 |
| 112 | 571 | VS2-H...13.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 113 | 580 | VS2-Br..5... | 0 | 0 | 0 | 1 | 0 | 0 |
| 114 | 582 | VS2-Br..7... | 0 | 0 | 0 | 1 | 0 | 0 |
| 115 | 583 | VS2-Br..9... | 0 | 0 | 0 | 1 | 0 | 0 |
| 116 | 584 | VS2-I...2... | 0 | 0 | 0 | 1 | 0 | 0 |
| 117 | 585 | VS2-Cl..10.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 118 | 587 | VS2-Cl..12.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 119 | 588 | VS2-Cl..3... | 0 | 0 | 0 | 1 | 0 | 0 |
| 120 | 602 | VS2-N...2... | 0 | 0 | 0 | 1 | 0 | 0 |
| 121 | 606 | VS2-N...3... | 0 | 0 | 0 | 1 | 1 | 0 |
| 122 | 607 | VS2-N...29.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 123 | 622 | VS2-O...19.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 124 | 635 | VS2-S...12.. | 0 | 0 | 0 | 1 | 1 | 0 |
| 125 | 636 | VS2-S...13.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 126 | 638 | VS2-S...15.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 127 | 639 | VS2-S...16.. | 0 | 0 | 0 | 1 | 0 | 0 |
| 128 | 640 | VS2-S...5... | 0 | 0 | 0 | 1 | 0 | 0 |
| 129 | 647 | [...(...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 130 | 653 | [...(...S... | 0 | 0 | 0 | 1 | 1 | 0 |
| 131 | 662 | [...1...3... | 0 | 0 | 0 | 1 | 0 | 0 |
| 132 | 664 | [...2...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 133 | 666 | [...2...C... | 0 | 0 | 0 | 1 | 0 | 0 |
| 134 | 667 | [...2...N... | 0 | 0 | 0 | 1 | 1 | 0 |
| 135 | 668 | [...2...O... | 0 | 0 | 0 | 1 | 0 | 0 |
| 136 | 669 | [...3....... | 0 | 0 | 0 | 1 | 1 | 0 |
| 137 | 673 | [...=...S... | 0 | 0 | 0 | 1 | 0 | 0 |
| 138 | 674 | [...@....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 139 | 675 | [...@...C... | 0 | 0 | 0 | 1 | 0 | 0 |
| 140 | 678 | [...C...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 141 | 684 | [...H...N... | 0 | 0 | 0 | 1 | 0 | 0 |
| 142 | 685 | [...N...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 143 | 689 | [...N...H... | 0 | 0 | 0 | 1 | 0 | 0 |
| 144 | 702 | [...Sn...... | 0 | 0 | 0 | 1 | 0 | 0 |
| 145 | 703 | [...Sn..[... | 0 | 0 | 0 | 1 | 0 | 0 |
| 146 | 705 | \...(...C... | 0 | 0 | 0 | 1 | 0 | 0 |
| 147 | 710 | \...C...1... | 0 | 0 | 0 | 1 | 0 | 0 |
| 148 | 712 | \...N....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 149 | 713 | \...N...=... | 0 | 0 | 0 | 1 | 0 | 0 |
| 150 | 714 | \...O...(... | 0 | 0 | 0 | 1 | 0 | 0 |
| 151 | 715 | \...O....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 152 | 716 | |...(....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 153 | 717 | |...(...3... | 0 | 0 | 0 | 1 | 0 | 0 |
| 154 | 718 | |........... | 0 | 0 | 0 | 1 | 0 | 0 |
| 155 | 719 | |...O....... | 0 | 0 | 0 | 1 | 0 | 0 |
| 156 | 61 | 1...C...#... | 0 | 0 | 0 | 0 | 1 | 0 |
| 157 | 72 | 2...3...(... | 0 | 0 | 0 | 0 | 1 | 0 |
| 158 | 83 | 3...2....... | 0 | 0 | 0 | 0 | 1 | 0 |
| 159 | 110 | =...N...3... | 0 | 0 | 0 | 0 | 1 | 0 |
| 160 | 138 | C...2...3... | 0 | 0 | 0 | 0 | 1 | 0 |
| 161 | 228 | EC2-C...41.. | 0 | 0 | 0 | 0 | 1 | 0 |
| 162 | 281 | EC2-N...28.. | 0 | 0 | 0 | 0 | 1 | 0 |
| 163 | 304 | EC2-P...24.. | 0 | 0 | 0 | 0 | 1 | 0 |
| 164 | 354 | Cl..(...F... | 0 | 0 | 0 | 0 | 1 | 0 |
| 165 | 375 | N...(...Br.. | 0 | 0 | 0 | 0 | 2 | 0 |
| 166 | 386 | N...1...Cl.. | 0 | 0 | 0 | 0 | 1 | 0 |
| 167 | 391 | N...3....... | 0 | 0 | 0 | 0 | 1 | 0 |
| 168 | 396 | N...=...3... | 0 | 0 | 0 | 0 | 1 | 0 |
| 169 | 398 | N...=...N... | 0 | 0 | 0 | 0 | 1 | 0 |
| 170 | 407 | N...N....... | 0 | 0 | 0 | 0 | 1 | 0 |
| 171 | 408 | N...N...2... | 0 | 0 | 0 | 0 | 1 | 0 |
| 172 | 409 | N...N...C... | 0 | 0 | 0 | 0 | 1 | 0 |
| 173 | 412 | N...S...2... | 0 | 0 | 0 | 0 | 1 | 0 |
| 174 | 416 | N...[...3... | 0 | 0 | 0 | 0 | 1 | 0 |
| 175 | 428 | O...(...Br.. | 0 | 0 | 0 | 0 | 1 | 0 |
| 176 | 483 | NNC-N...110. | 0 | 0 | 0 | 0 | 2 | 0 |
| 177 | 491 | NNC-N...514. | 0 | 0 | 0 | 0 | 0 | 0 |
| 178 | 512 | S...(...O... | 0 | 0 | 0 | 0 | 1 | 0 |
| 179 | 517 | S...2...(... | 0 | 0 | 0 | 0 | 1 | 0 |
| 180 | 554 | VS2-C...25.. | 0 | 0 | 0 | 0 | 1 | 0 |
| 181 | 604 | VS2-N...19.. | 0 | 0 | 0 | 0 | 1 | 0 |
| 182 | 632 | VS2-P...8... | 0 | 0 | 0 | 0 | 1 | 0 |
| 183 | 645 | [...(...(... | 0 | 0 | 0 | 0 | 1 | 0 |
| 184 | 670 | [...3...C... | 0 | 0 | 0 | 0 | 1 | 0 |
| 185 | 700 | [...S...C... | 0 | 0 | 0 | 0 | 1 | 0 |

**Figure S1.** Snapshot represenetation of model building criteria

