Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

Molecular Electrostatic Potential as a General and Versatile Indicator for Electronic Substituent Effects: Statistical Analysis and Applications

Donghan Shin and YounJoon Jung*

Department of Chemistry, Seoul National University,
1 Gwanak-ro, Gwanak-gu,
Seoul 08826, Republic of Korea

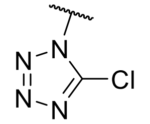
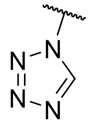
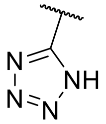
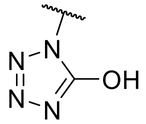
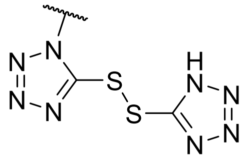
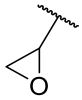
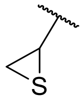
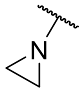
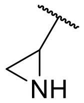
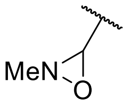
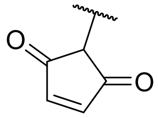
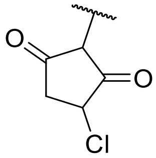
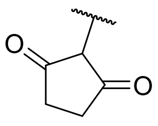
E-mail: yjjung@snu.ac.kr
Phone: +82 (2) 880-4369

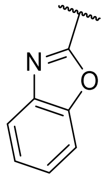
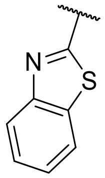
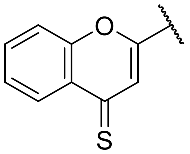
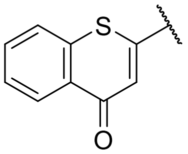
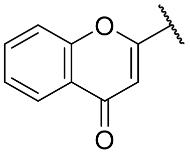
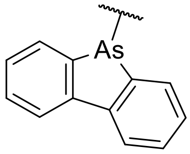
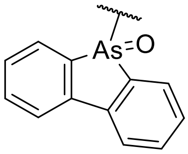
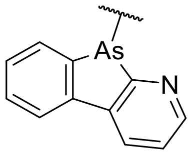
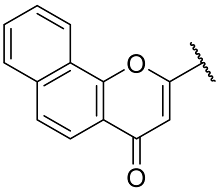
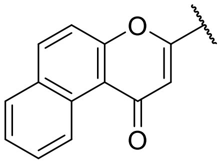
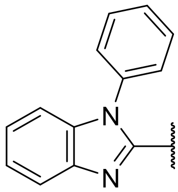
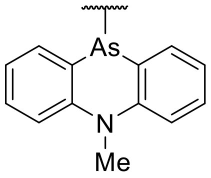
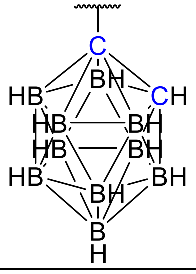
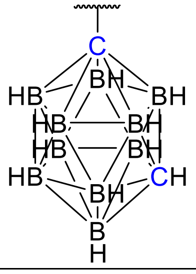
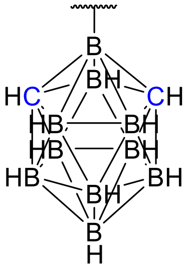
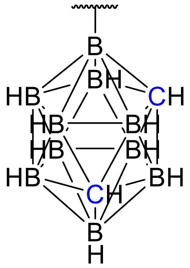
Table of Contents

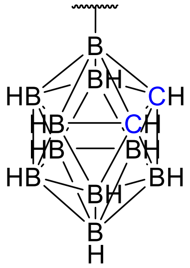
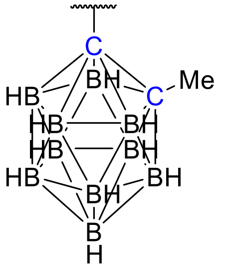
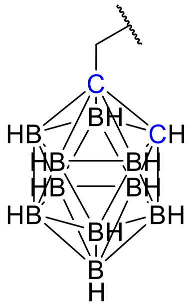
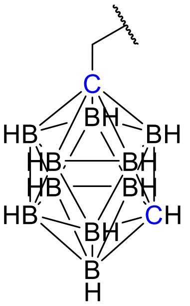
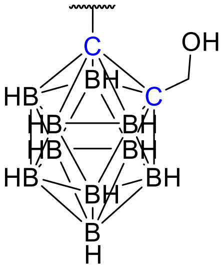
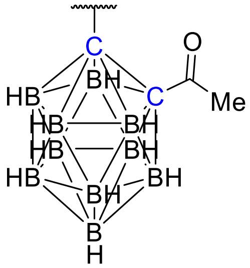
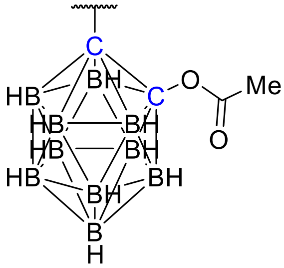
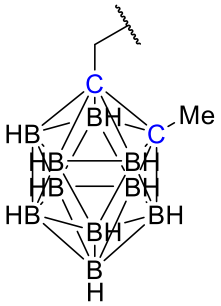
Functional Groups Used in This Research: Description of Supplementary_Data.xlsx

Essentially, we use 423 substituent constants from the Hammett equation (σ) as well as from the Swain-Lupton equation (F and R) well-organised in Ref 1. All of the information concerning them is described in the “Supplementary_Data.xlsx” file. We list key quantities and abbreviations used in the .xlsx file. If there is no additional description for Pr and Bu, they will refer to n-propyl group and n-butyl group respectively.

σ_m: σ value which describes *meta*-directing effect of a functional group
σ_p: σ value which describes *para*-directing effect of a functional group
F: inductive parameters from the Swain-Lupton equation
R: resonance parameters the Swain-Lupton equation
Δq_{C,m}: averages of (unless there are additional marks) HPA (Hirshfeld population analysis) charge shifts of carbons on two meta positions.
Δq_{C,p}: HPA charge shifts of carbons on para positions.
ΔV_{m}: averages of electrostatic potential shifts on two meta positions.
ΔV_{p}: electrostatic potential shifts on para positions.
(All of the reference values of the shifts come from an unsubstituted benzene)

Functional Group Names	Structures	Functional Group Names	Structures
5-chloro-1-tetrazolyl		1-(1H)-tetrazolyl	
5-(1H)-tetrazolyl		5-hydroxy-1-tetrazolyl	
C2N8S2H			
oxiranyl		2-thiacylopropyl	
1-aziridinyl		2-aziridinyl	
N-methyl-3-oxaziridinyl		1-pyrroline-2,5-dione	
3-chloro-1-pyrrolidine-2,5-dione		1-pyrrolidine-2,5-dione	

2-benzoxazolyl		2-benzthiazolyl	
2-benzo-4-thiopyronyl		2-benzothiopyronyl	
2-(benzo-1,4-pyronyl)		1-dibenzarsenyl	
1-dibenzoarsoxyl		1-dibenzoarsazinyl	
pyron-comp1		pyron-comp2	
imidazole-comp1		As-comp1	
1-(1,2(BH)10-C2H)		1-(1,7(BH)10-C2H)	
2-(1,7(BH)10-C2H)		4-(1,7(BH)10-C2H)	

3-(1,2(BH)10-C2H)		1-(1,2(BH)10-C2Me)	
CH2-1-(1,2(BH)10-C2H)		CH2-1-(1,7(BH)10-C2H)	
2-(hydroxymethyl) carboran-1-yl		2-(methylcarbonyl) carboran-1-yl	
2-((carbonyloxy)methyl) carboran-1-yl		CH2-1-(1,7(BH)10-C2Me)	

Δq_{C,p/m,NPA}: Charge shifts from natural population analysis[2,3]

Δq_{C,p/m,AIM}: Charge shifts from AIM (atoms-in-molecules) population analysis proposed by Bader, R. W. F. (implemented in Multiwfn 3.7[4,5])

Δq_{C,p/m,LPA}: Charge shifts from Löwdin population analysis (LPA) [6,7]

Δq_{C,p/m,VDD}: Charge shifts from Voronoi deformation density[8,9] (Multiwfn 3.7[5])

Δq_{C,p/m,CM5}: Charge shifts from CM5 (Charge Model 5), one of the variations of HPA to reproduce the reference dipole moments, proposed by Marenich, A. V. et al.

Δq_{C,p/m,HPA-I}: Charge shifts from HPA-I (Iterative), a method to make weighing factors from HPA adjust the molecular environment via iterative self-consistent field scheme[11–13] (Multiwfn 3.7[5])

Δq_{C,p/m,ADCH}: Charge shifts from ADCH (atomic dipole moment corrected Hirshfeld population method[14]) (Multiwfn 3.7[5])

Δq_{C,p/m,Becke}: Charge shifts from Becke population analysis[15] (Multiwfn 3.7[5])

Δq_{C,p/m,MPA}: Charge shifts from Mulliken population analysis (MPA)[16,17]

Δq_{C,p/m,PEOE}: Charge shifts from PEOE[18–19] (partial equalisation of orbital electronegativity, also known as Gasteiger charges, Multiwfn 3.7[5])
Δq_{C,p/m,CHELPG}: Charge shifts from CHELPG[20] (CHarges from ELectrostatic Potentials using a Grid-based method)
Δq_{C,p/m,CHELP}: Charge shifts from CHELP[21] (CHarges from ELectrostatic Potentials)
Δq_{C,p/m,MK}: Charge shifts from MK scheme[22,23]
Δq_{C,m,GC}: shifts of Gasteiger charge[24]

Linear Regressions Not Shown in the Main Text

In the main text, we show R^2 values and, for some particular cases, *p* values of being normal distribution obtained by Shapiro-Wilk test of linear regressions of representative values on experimental constants (Table 2, Table 3 and Table 4).
In this section, we will show all of figures of linear regressions used in the main text.

§ Plots for Shifts of HPA Charges

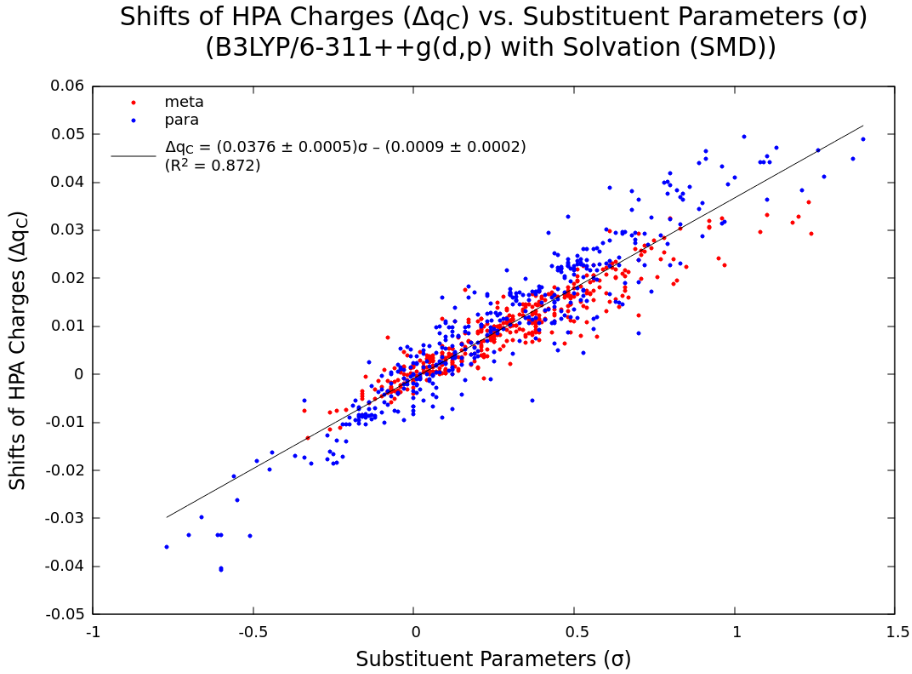


Figure S1: Shifts of Charges from HPA in B3LYP/6-311++g(d,p) and SMD Calculation

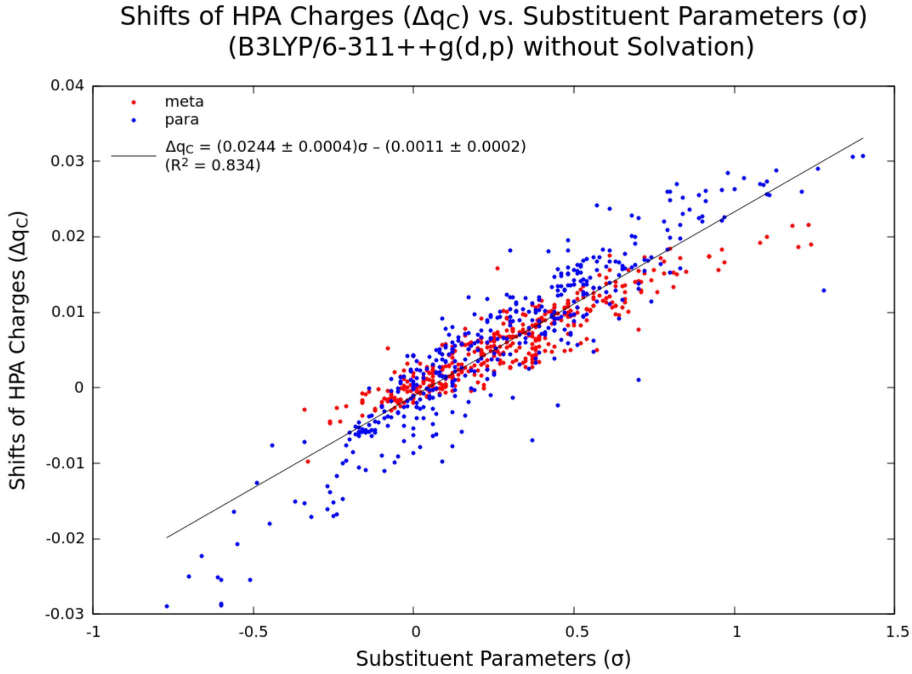


Figure S2: Shifts of Charges from HPA in B3LYP/6-311++g(d,p) Calculation (No Solvation)

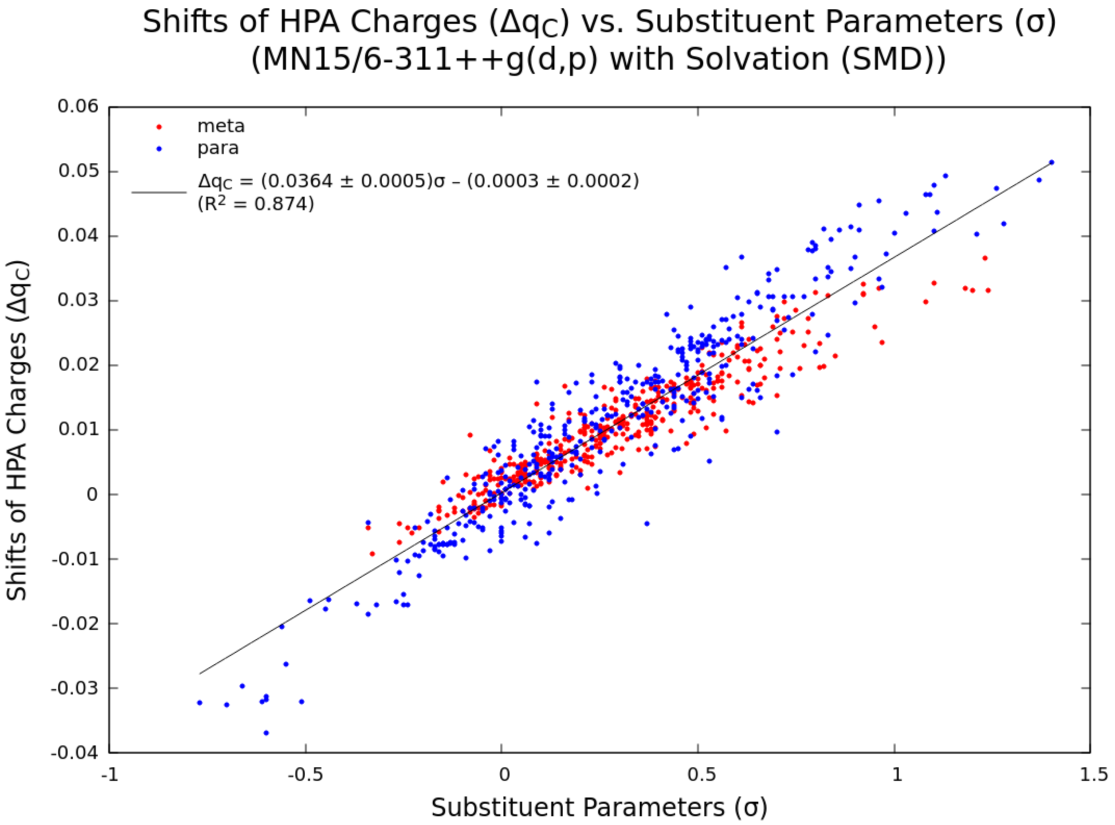


Figure S3: Shifts of Charges from HPA in MN15/6-311++g(d,p) and SMD Calculation

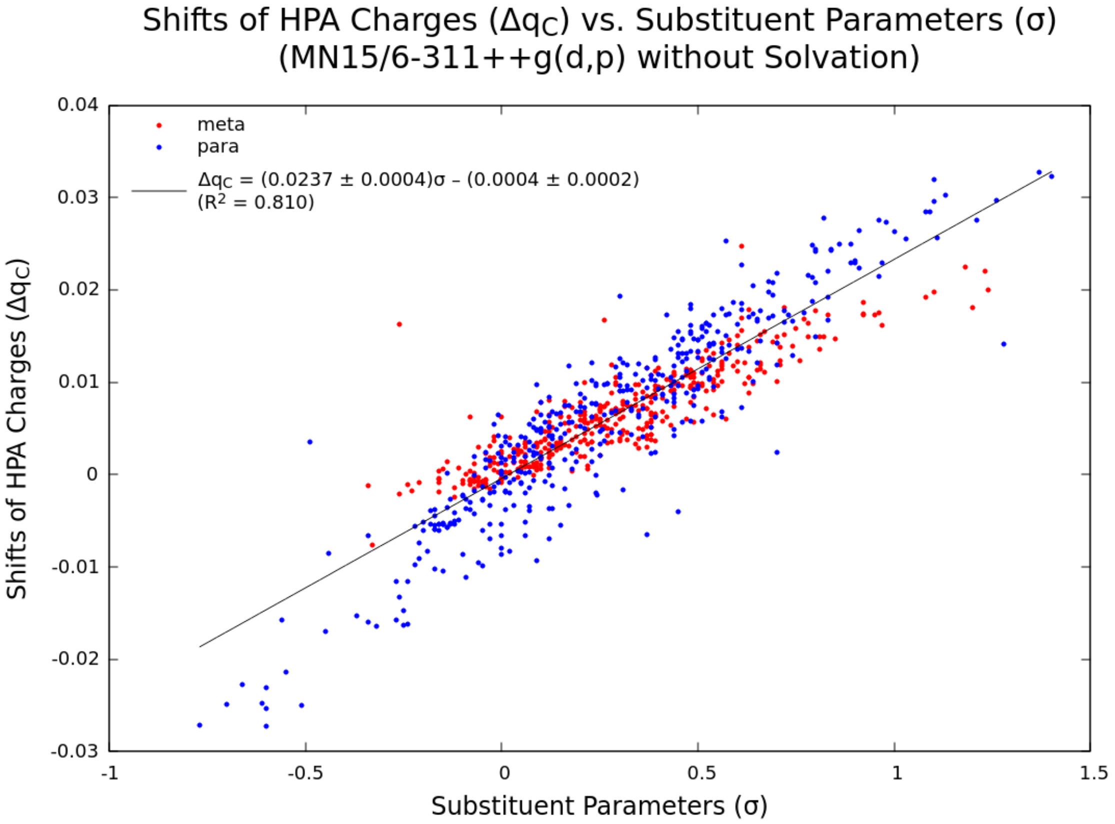


Figure S4: Shifts of Charges from HPA in MN15/6-311++g(d,p) Calculation (No Solvation)

§ Plots for Table 2

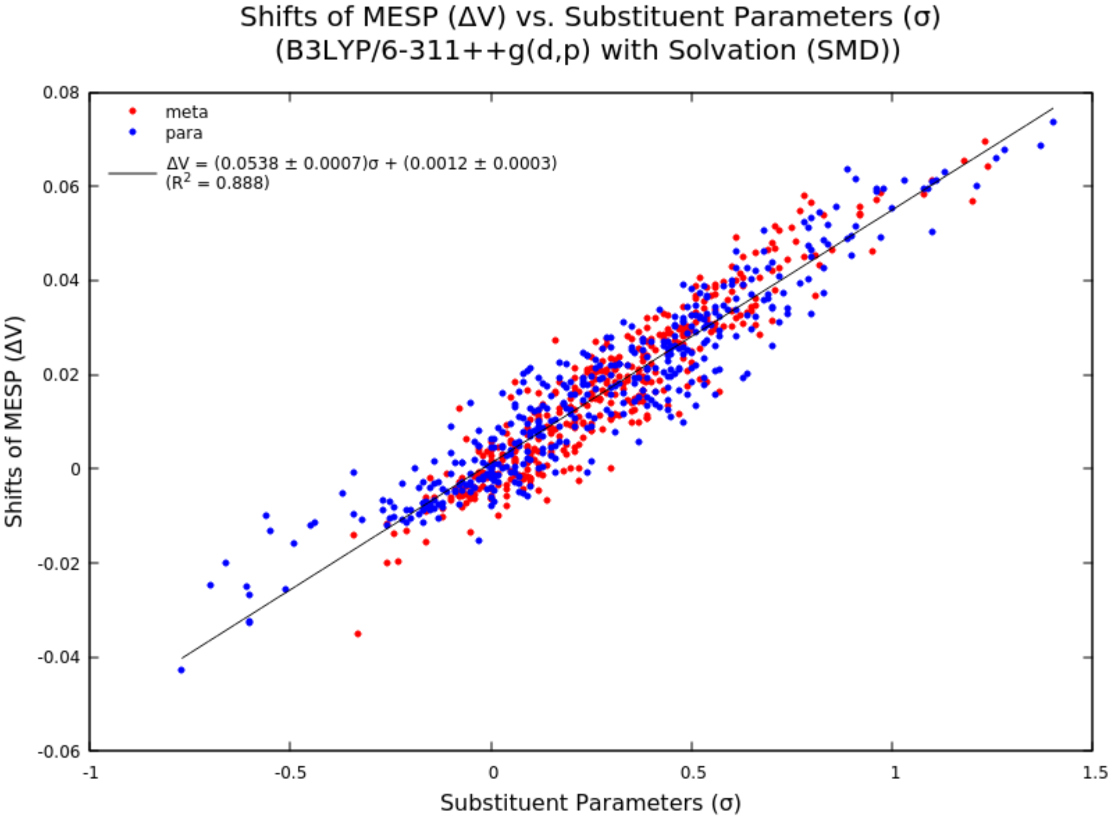


Figure S5: Shifts of MESP in B3LYP/6-311++g(d,p) and SMD Calculation

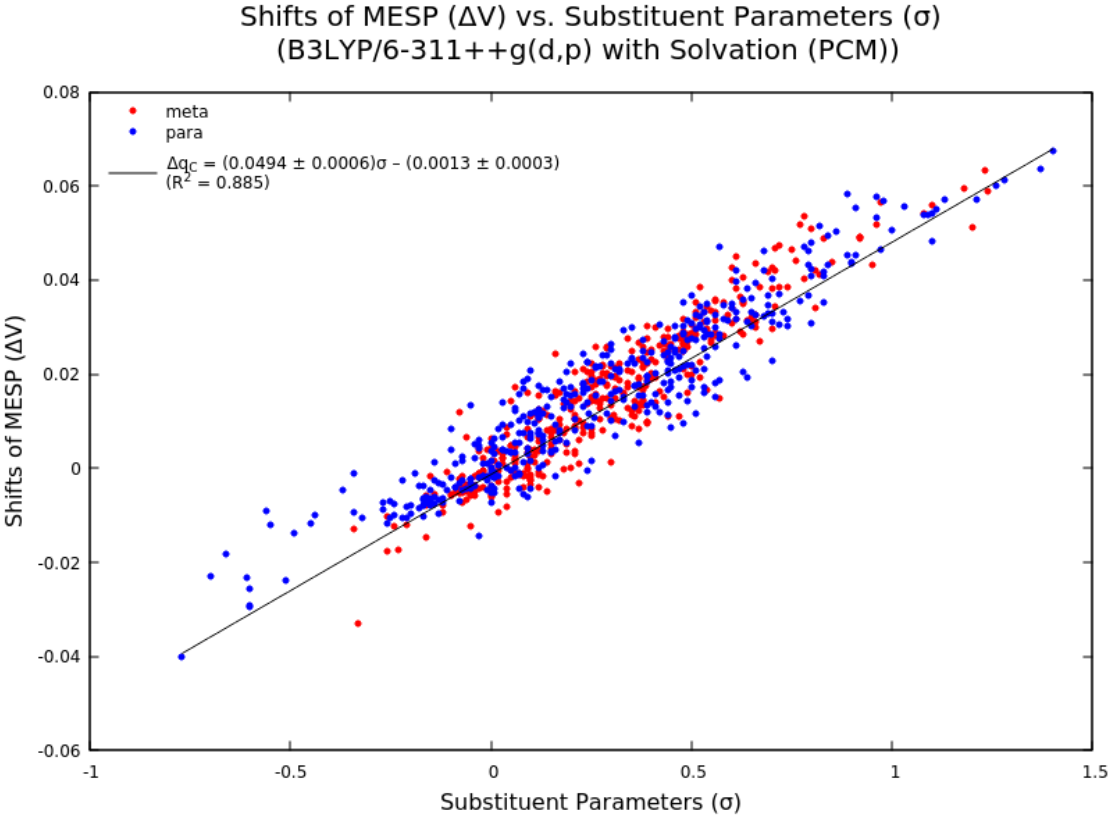


Figure S6: Shifts of MESP in B3LYP/6-311++g(d,p) and PCM Calculation

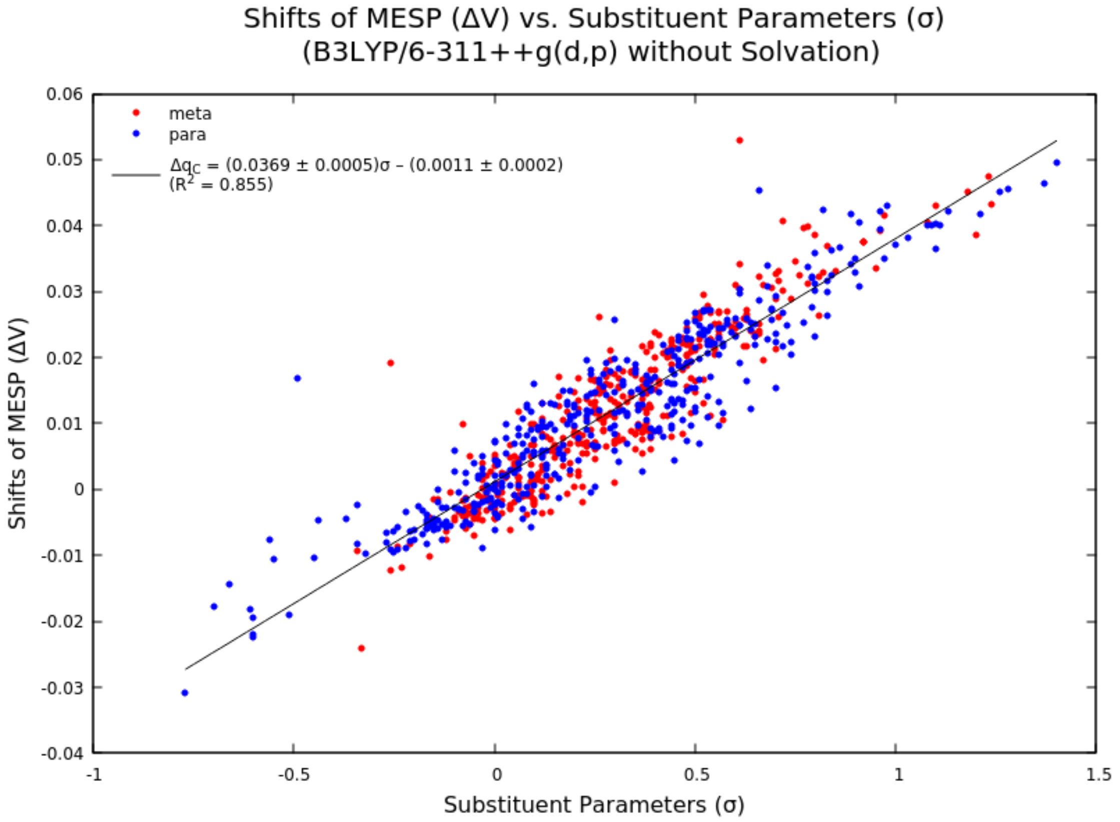


Figure S7: Shifts of MESP in B3LYP/6-311++g(d,p) Calculation (No Solvation)

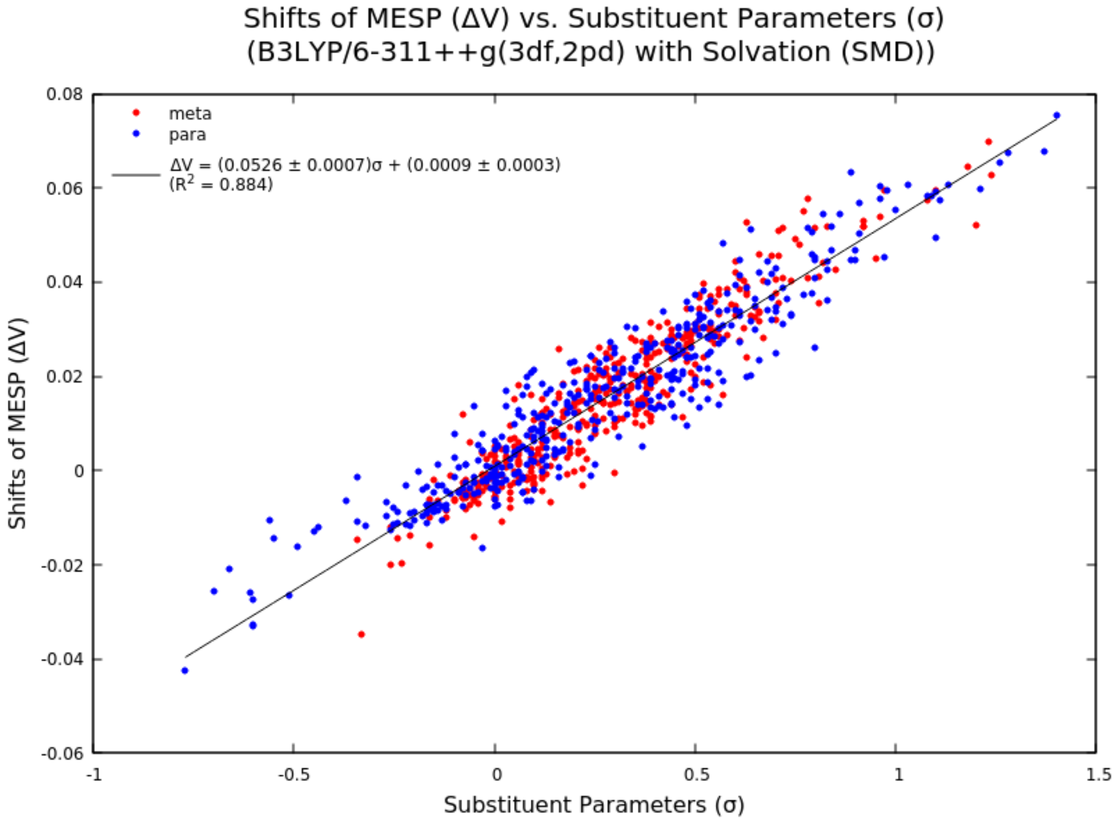


Figure S8: Shifts of MESP in B3LYP/6-311++g(3df,2pd) and SMD Calculation

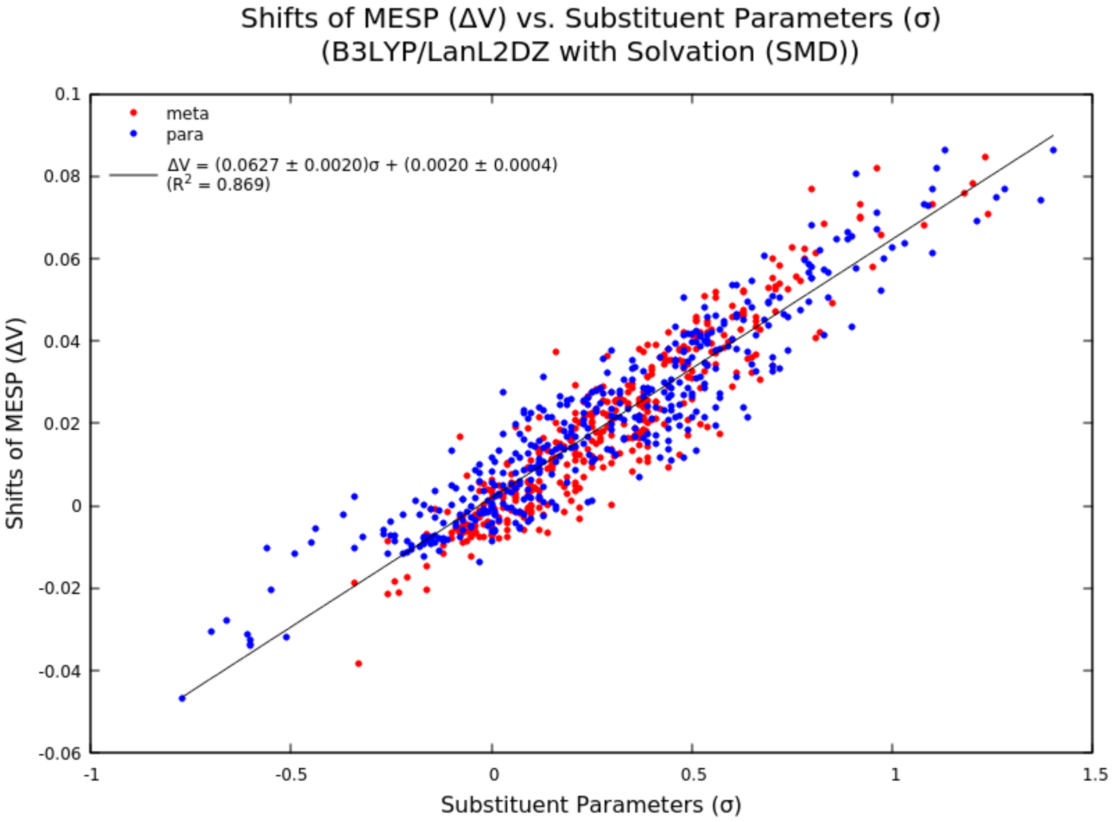


Figure S9: Shifts of MESP in B3LYP/LanL2DZ and SMD Calculation

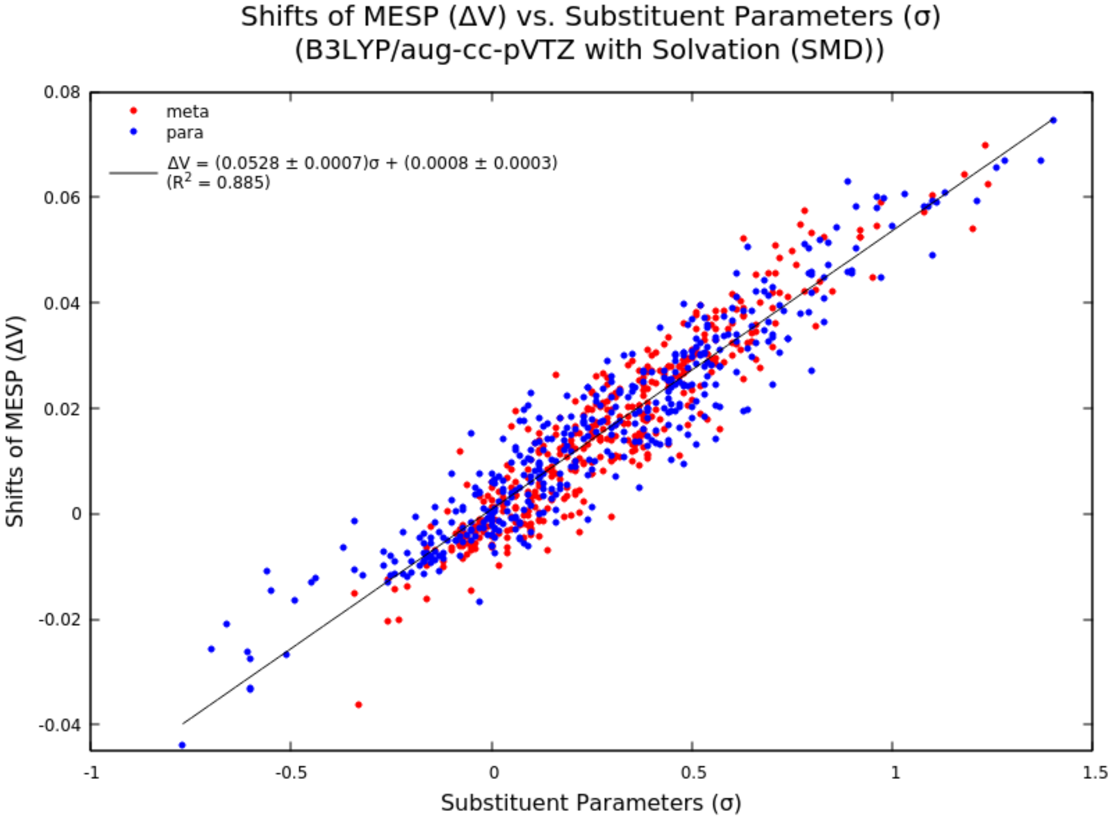


Figure S10: Shifts of MESP in B3LYP/aug-cc-pVTZ and SMD Calculation

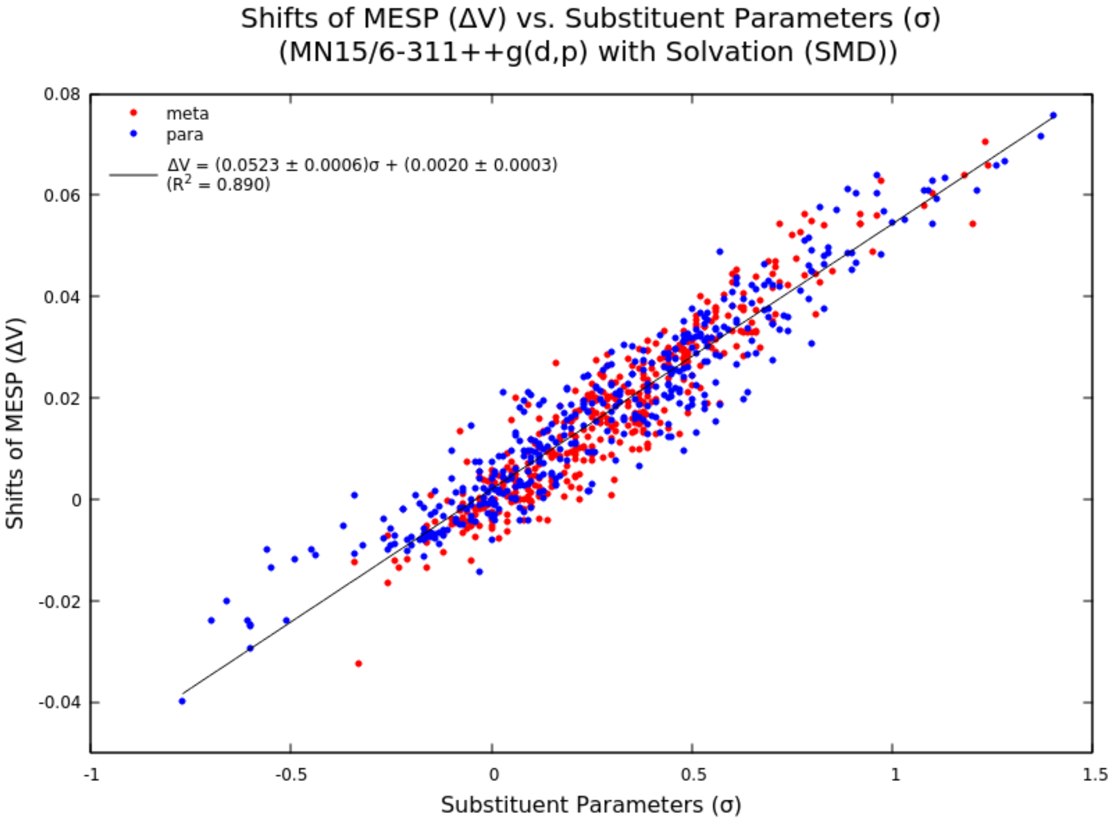


Figure S11: Shifts of MESP in MN15/6-311++g(d,p) and SMD Calculation

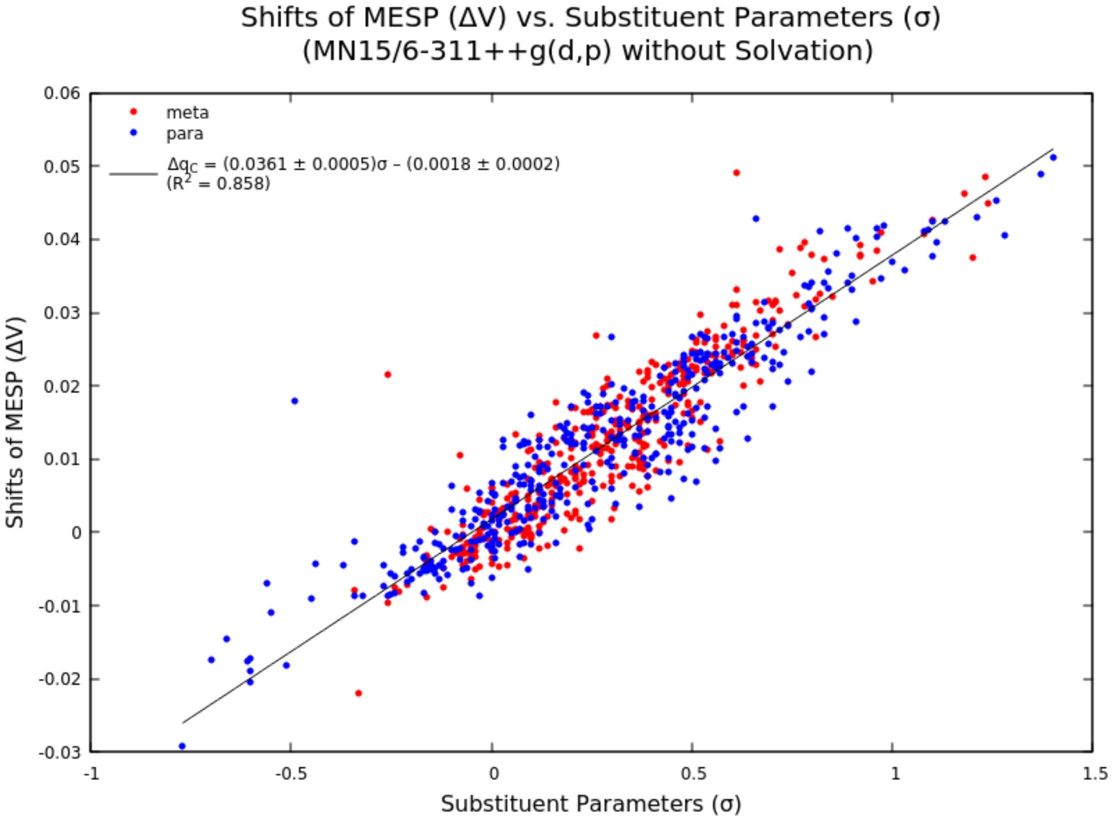


Figure S12: Shifts of MESP in MN15/6-311++g(d,p) Calculation (No Solvation)

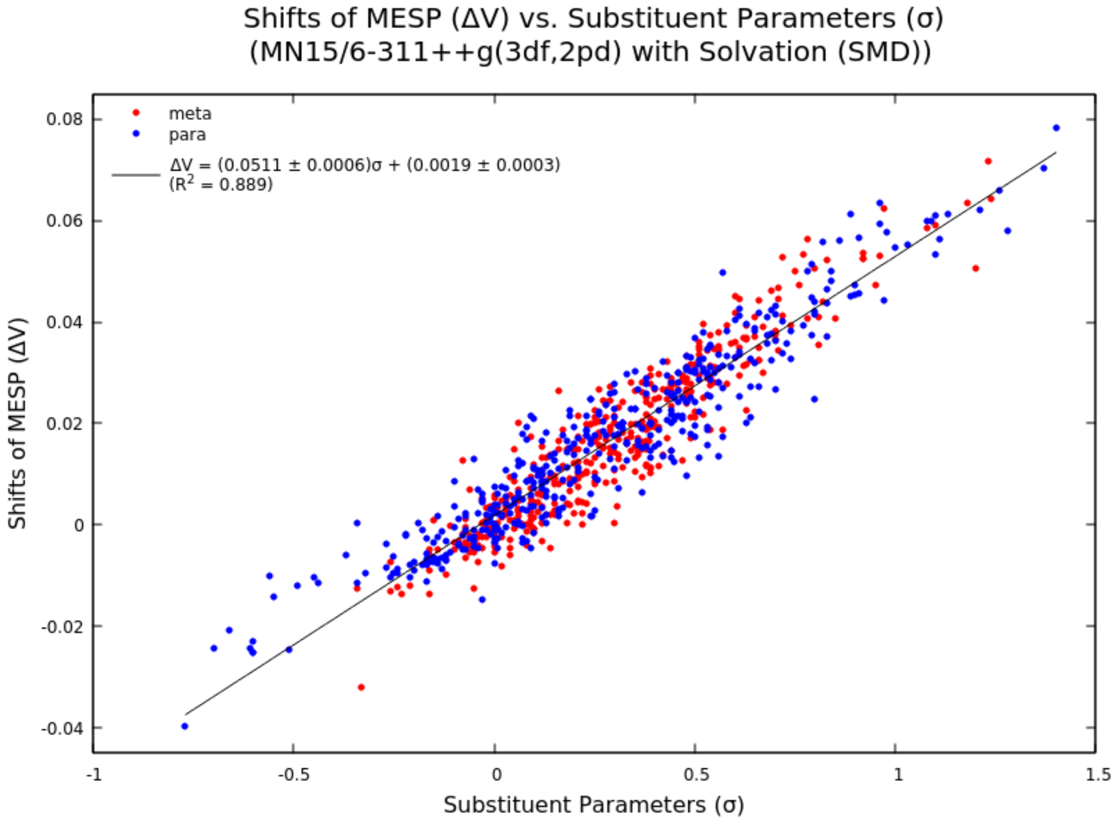


Figure S13: Shifts of MESP in MN15/6-311++g(3df,2pd) and SMD Calculation

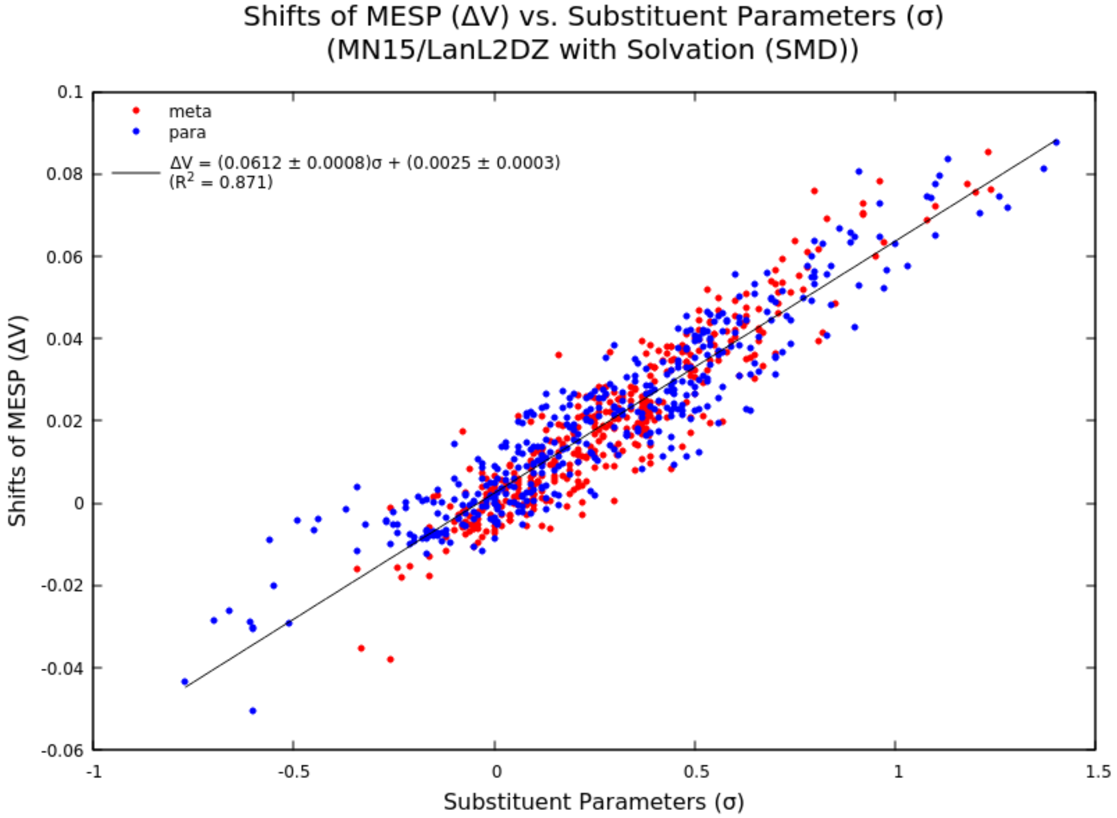


Figure S14: Shifts of MESP in MN15/LanL2DZ and SMD Calculation

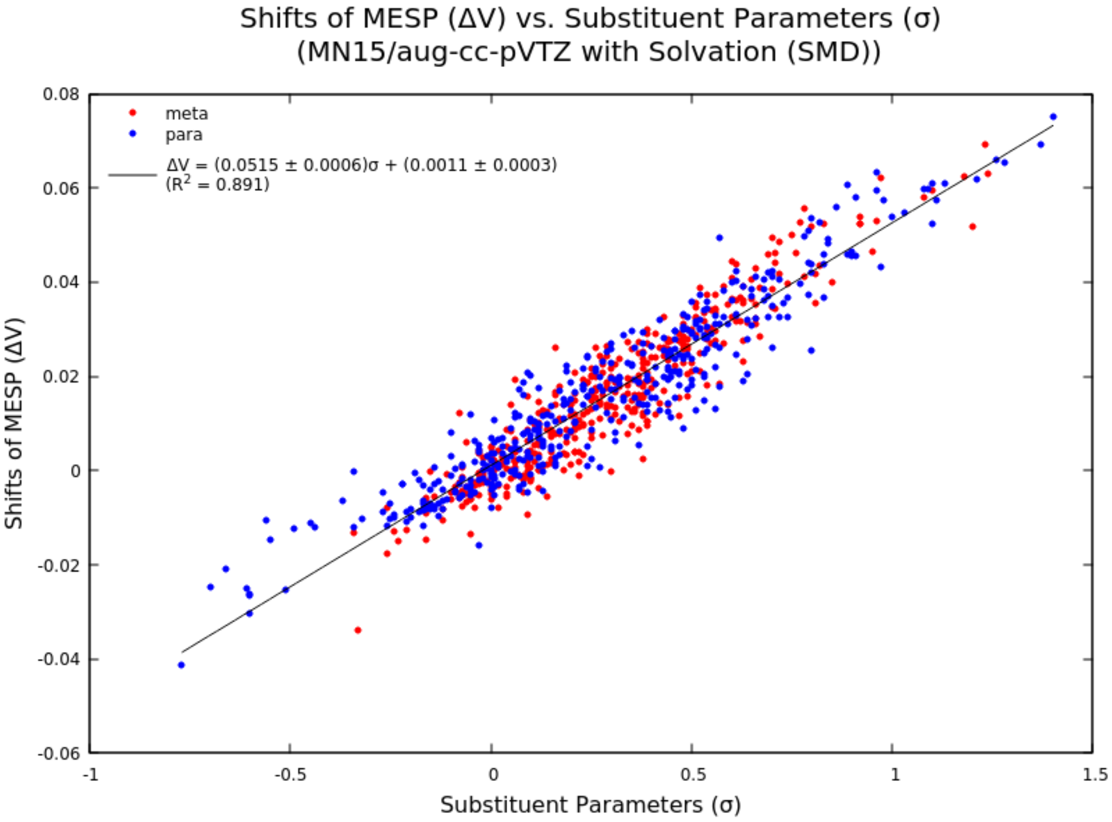


Figure S15: Shifts of MESP in MN15/aug-cc-pVTZ and SMD Calculation

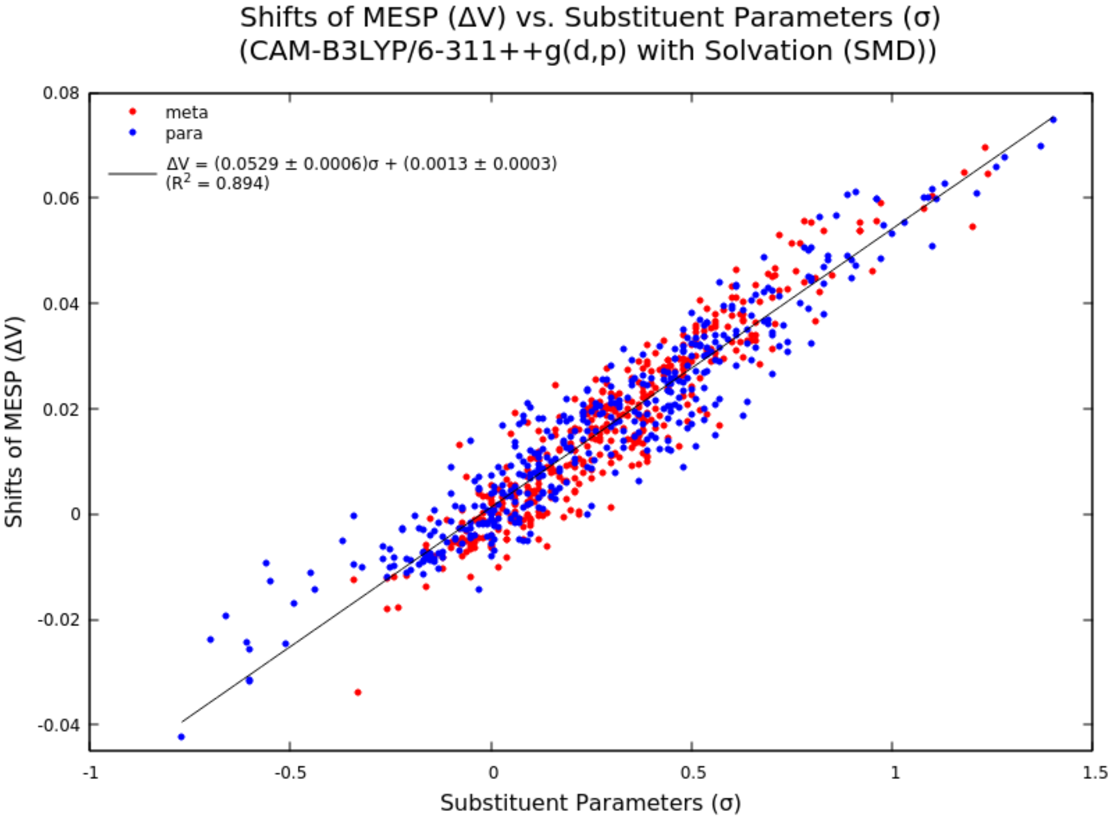


Figure S16: Shifts of MESP in CAM-B3LYP/6-311++g(d,p) and SMD Calculation

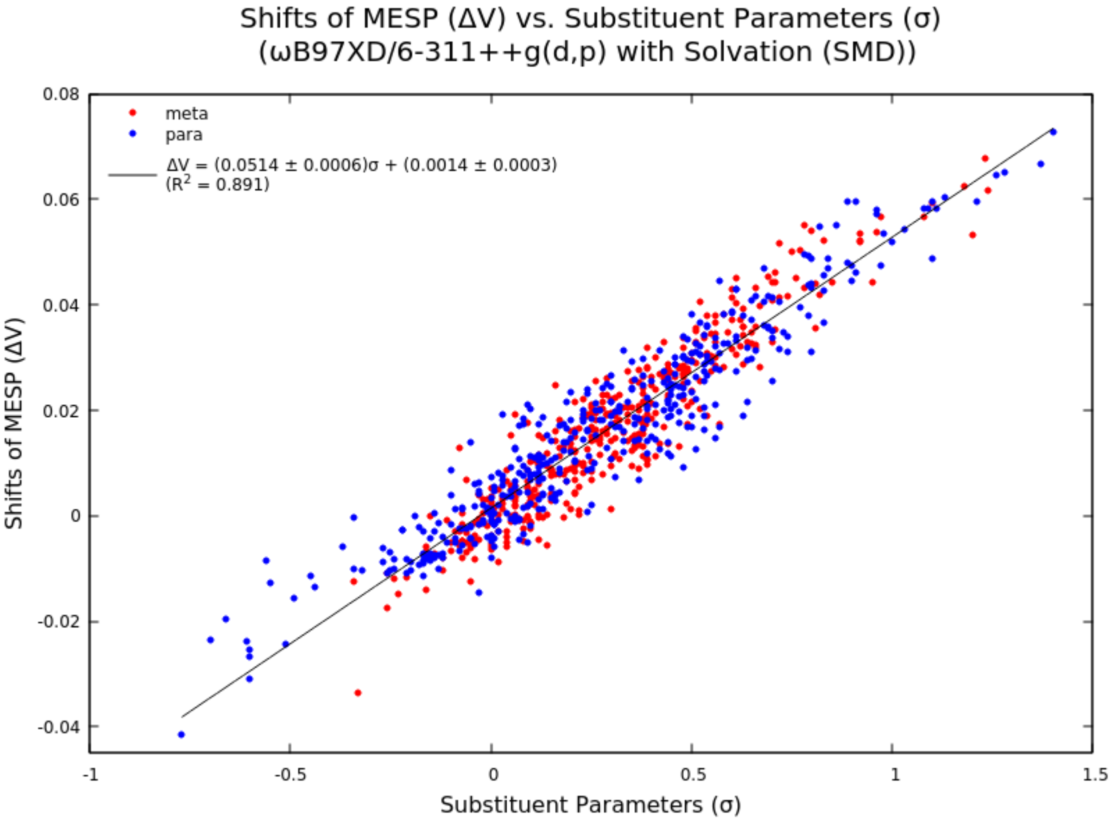


Figure S17: Shifts of MESP in ωB97XD/6-311++g(d,p) and SMD Calculation

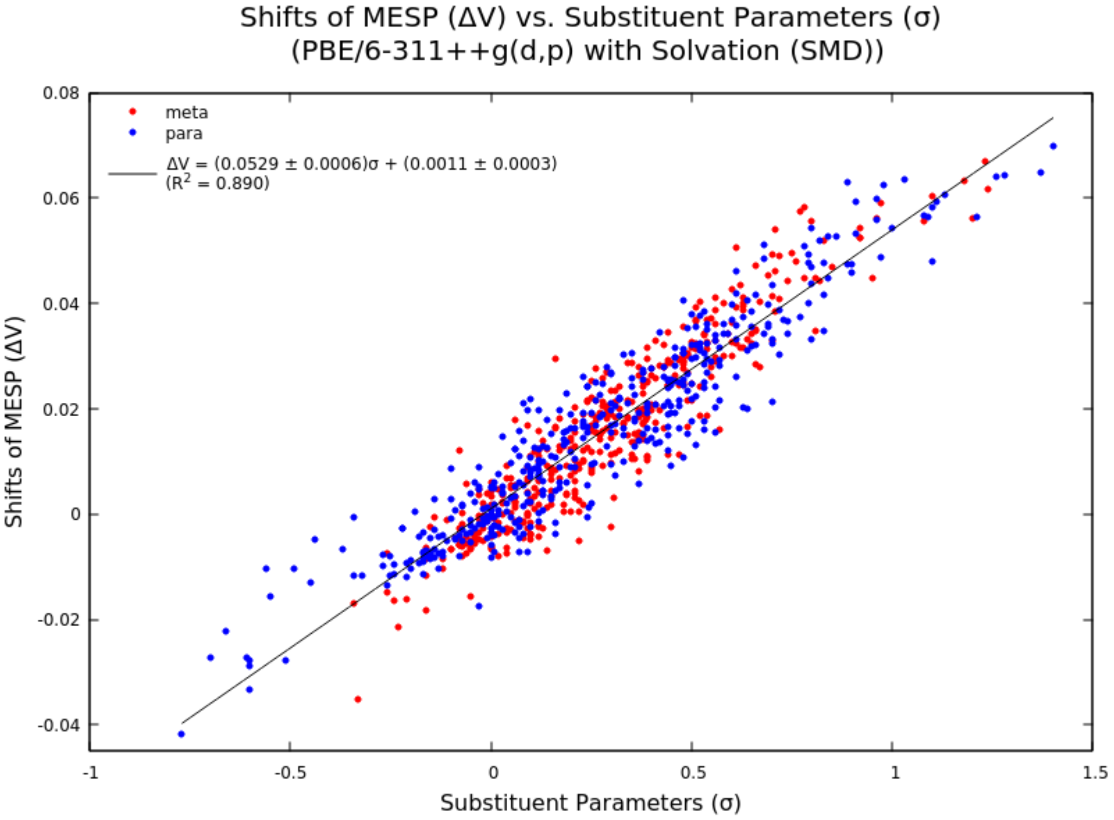


Figure S18: Shifts of MESP in PBE/6-311++g(d,p) and SMD Calculation

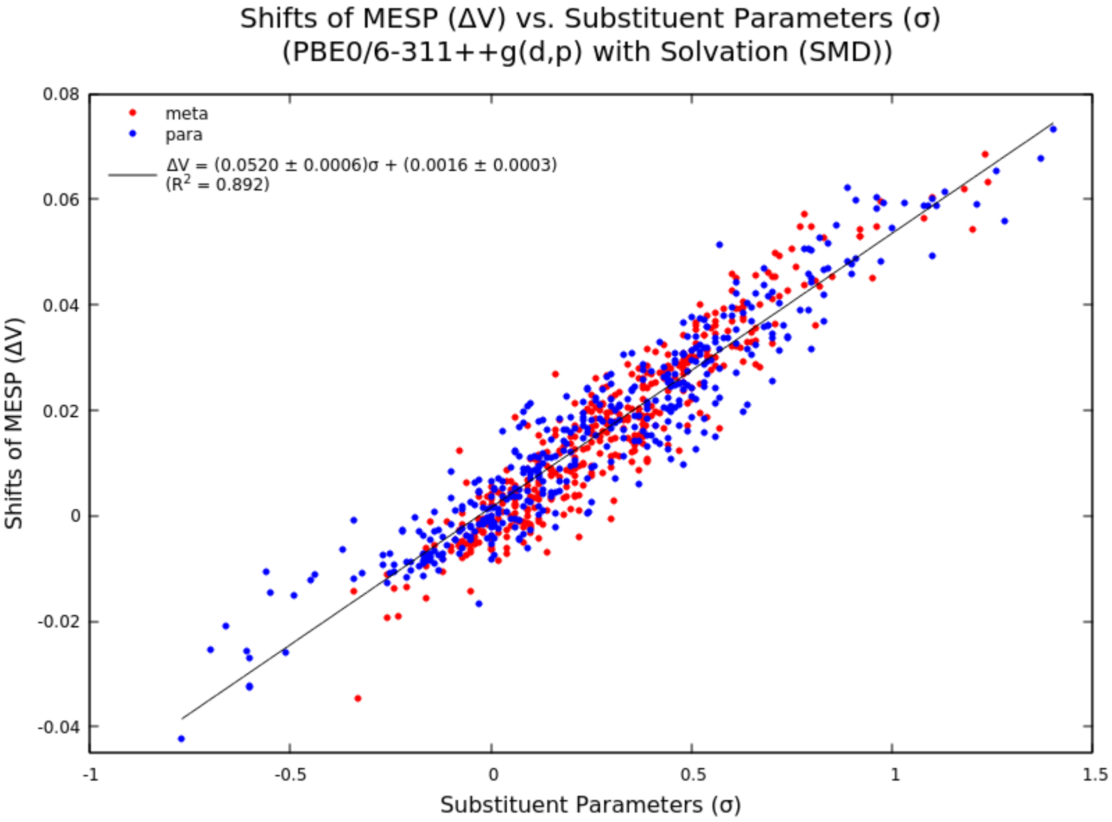


Figure S19: Shifts of MESP in PBE0/6-311++g(d,p) and SMD Calculation

§ Plots for Table 3

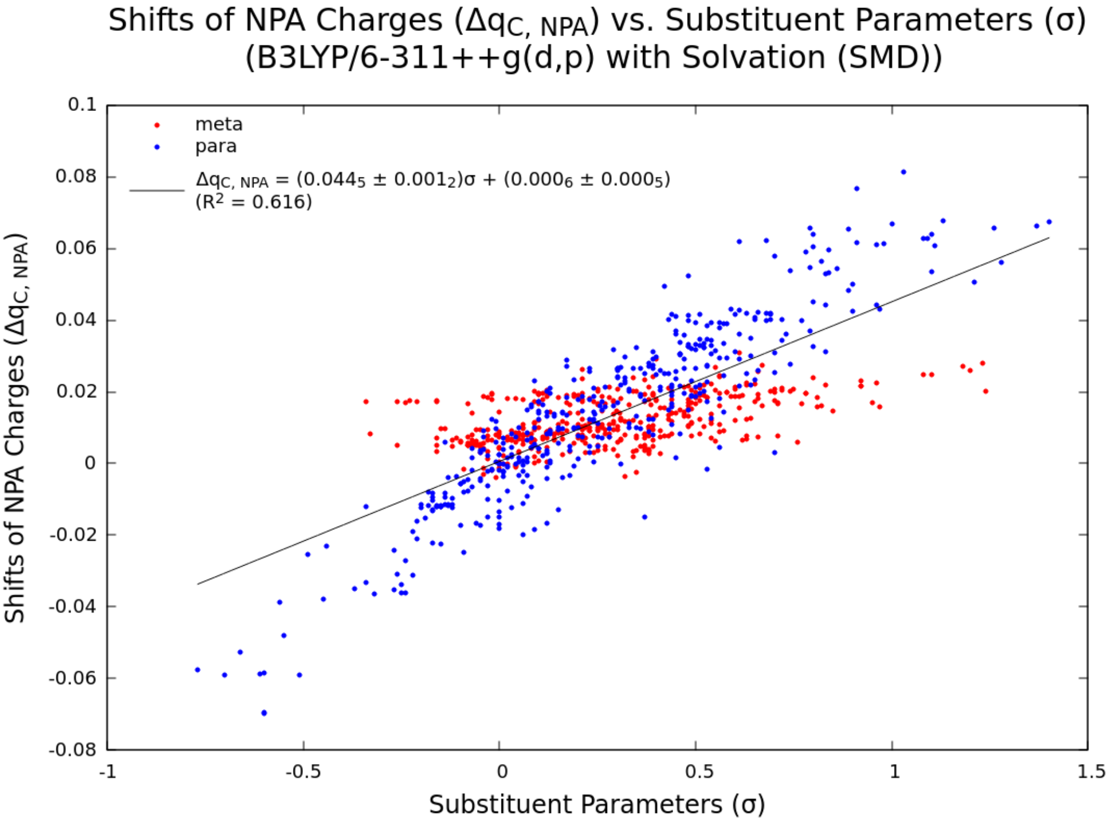


Figure S20: Shifts of Charges from NPA in B3LYP/6-311++g(d,p) and SMD Calculation

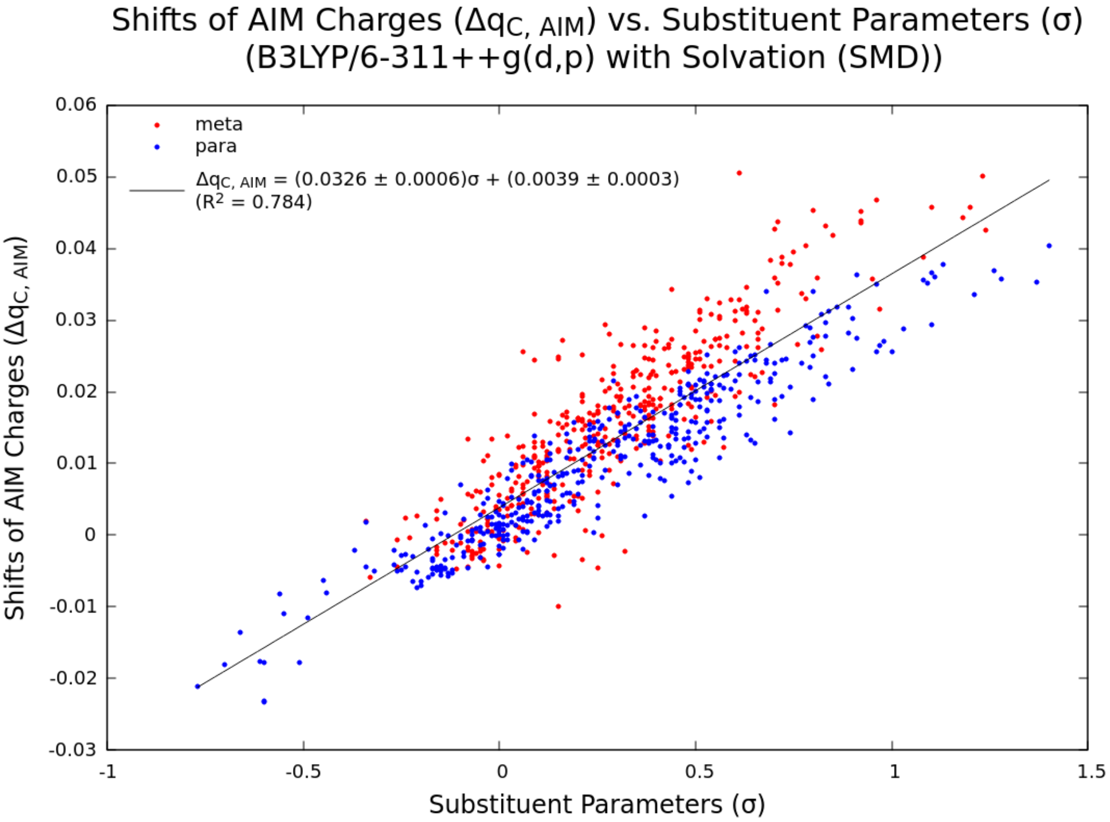


Figure S21: Shifts of Charges from AIM population analysis in B3LYP/6-311++g(d,p) and SMD Calculation

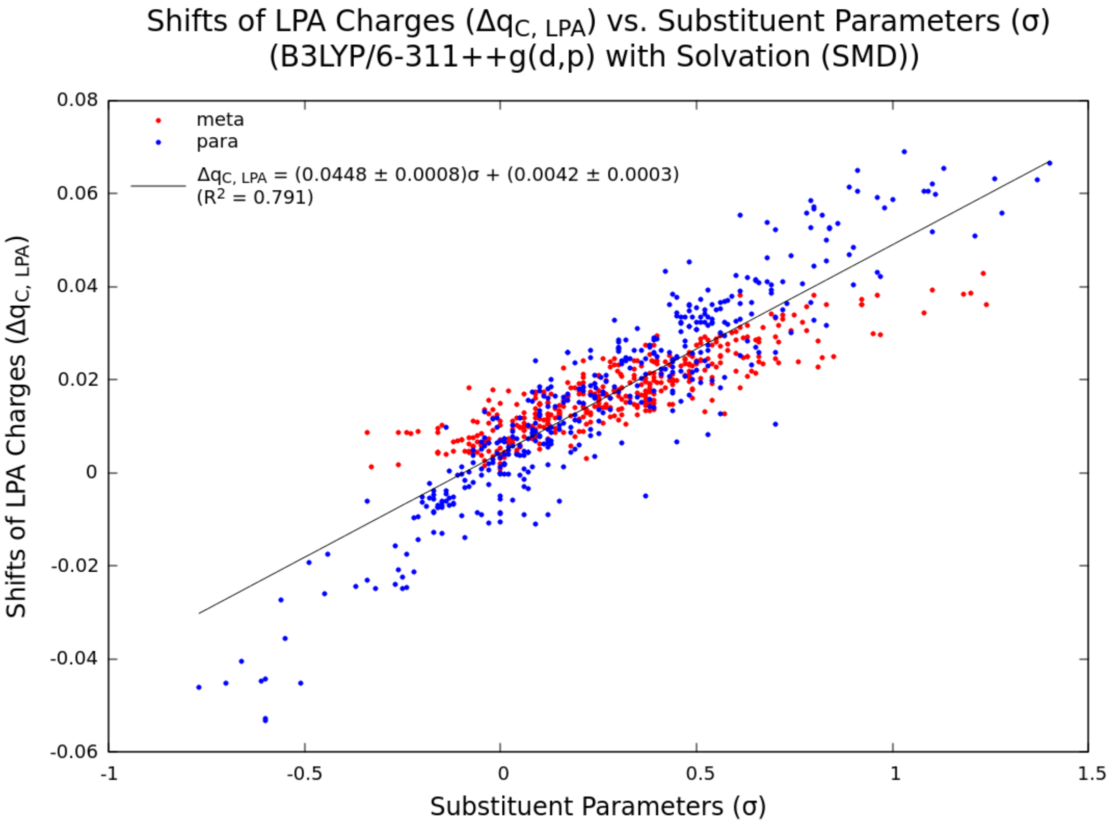


Figure S22: Shifts of Charges from LPA in B3LYP/6-311++g(d,p) and SMD Calculation

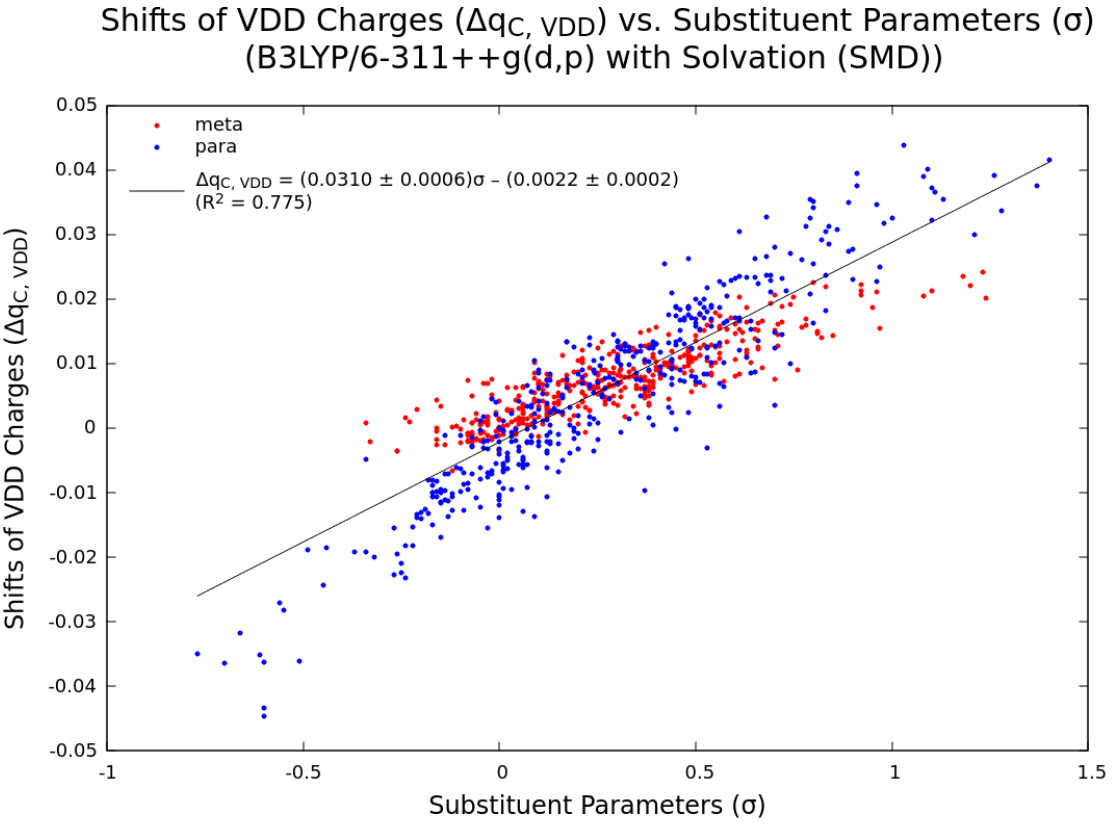


Figure S23: Shifts of Charges from VDD in B3LYP/6-311++g(d,p) and SMD Calculation

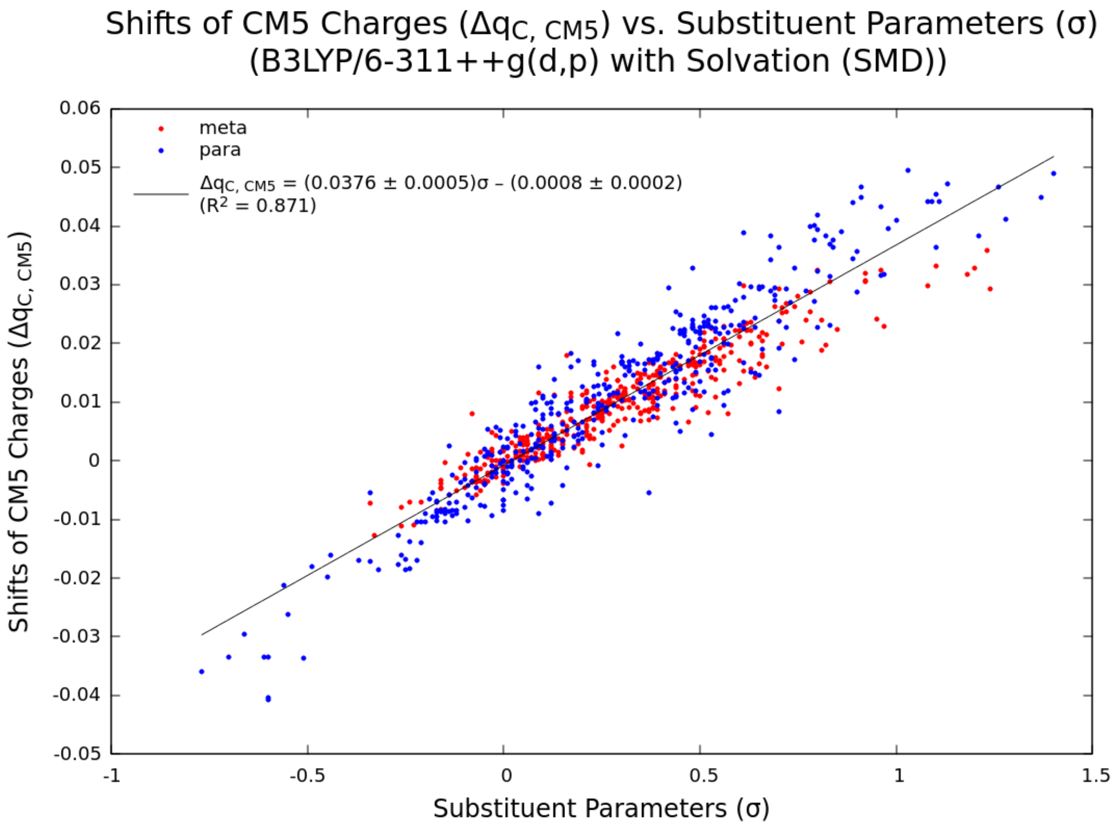


Figure S24: Shifts of Charges from CM5 in B3LYP/6-311++g(d,p) and SMD Calculation

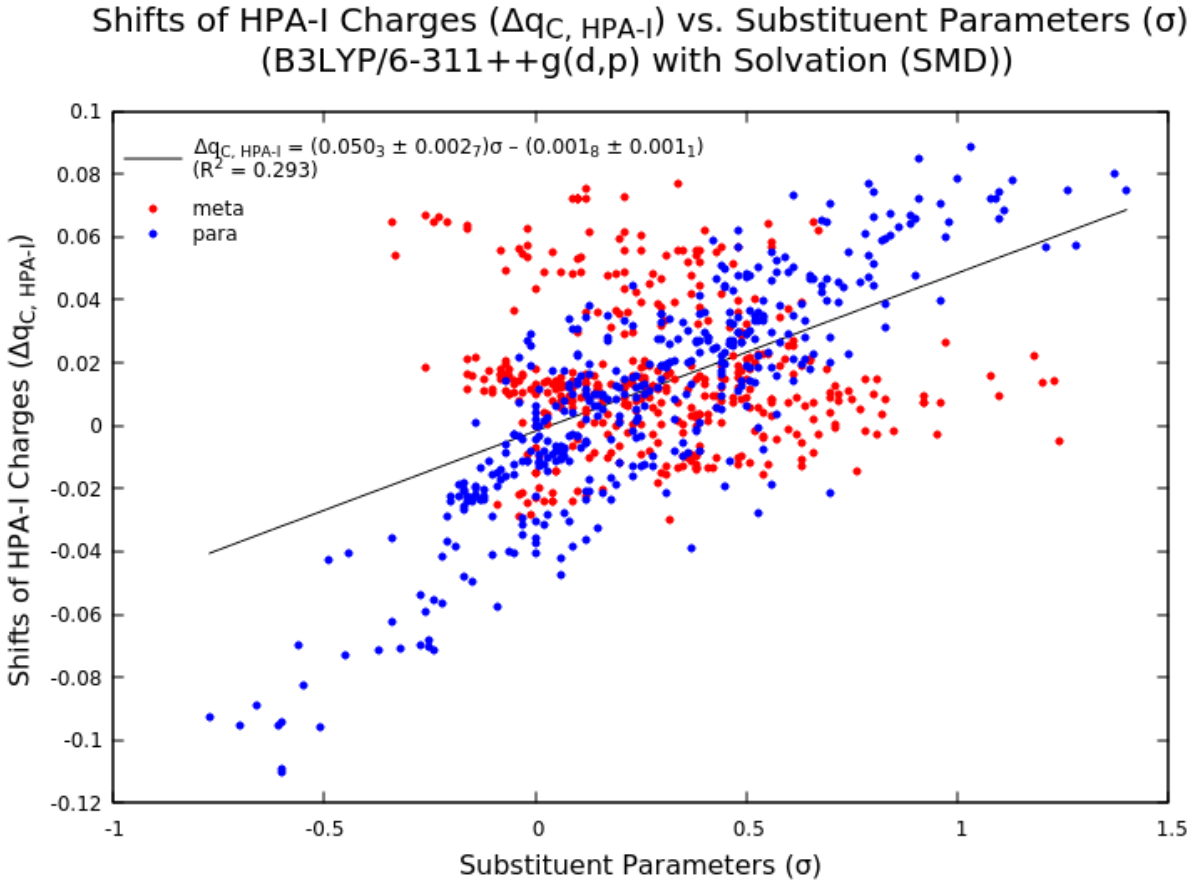


Figure S25: Shifts of Charges from HPA-I in B3LYP/6-311++g(d,p) and SMD Calculation

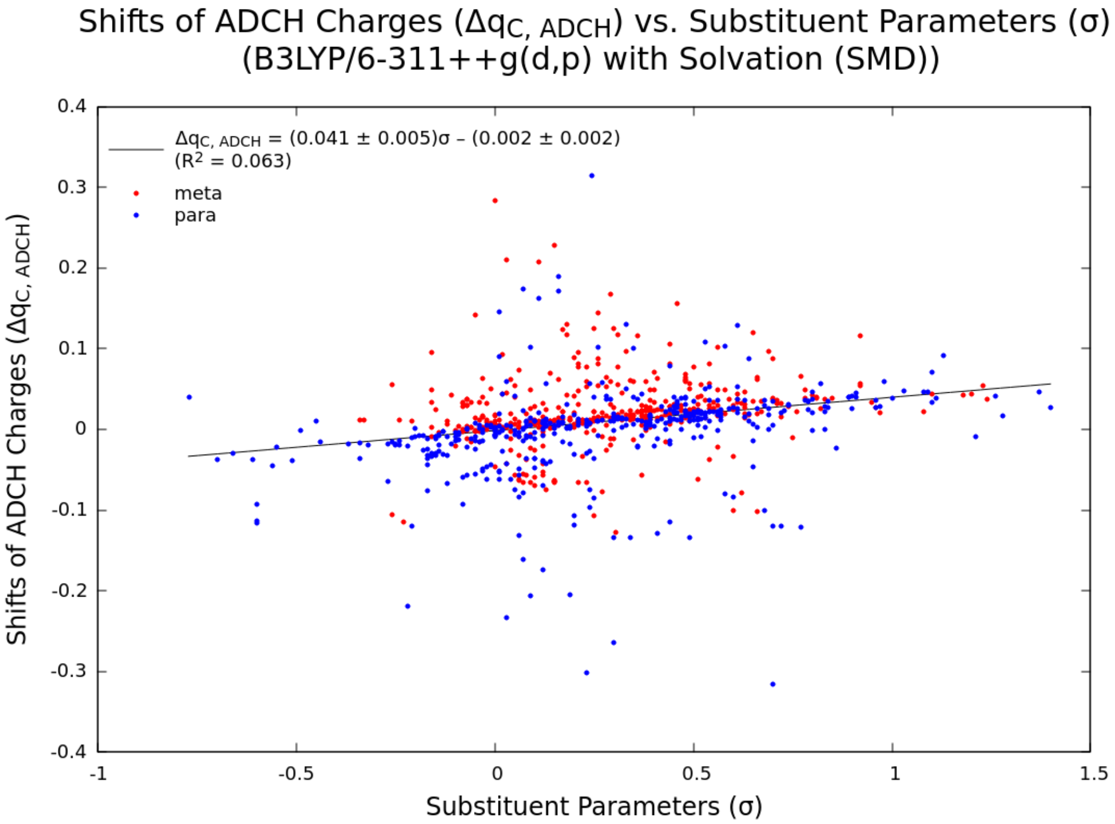


Figure S26: Shifts of Charges from ADCH in B3LYP/6-311++g(d,p) and SMD Calculation

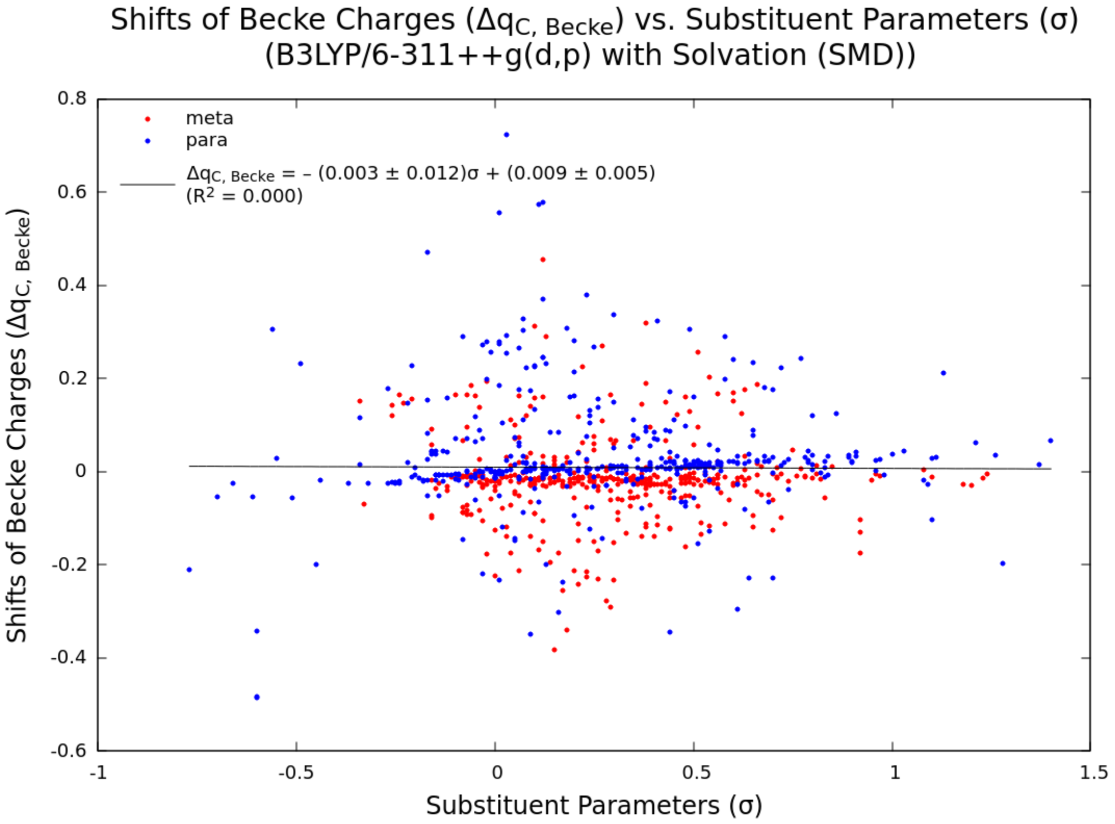


Figure S27: Shifts of Charges from Becke Population Analysis in B3LYP/6-311++g(d,p) and SMD Calculation

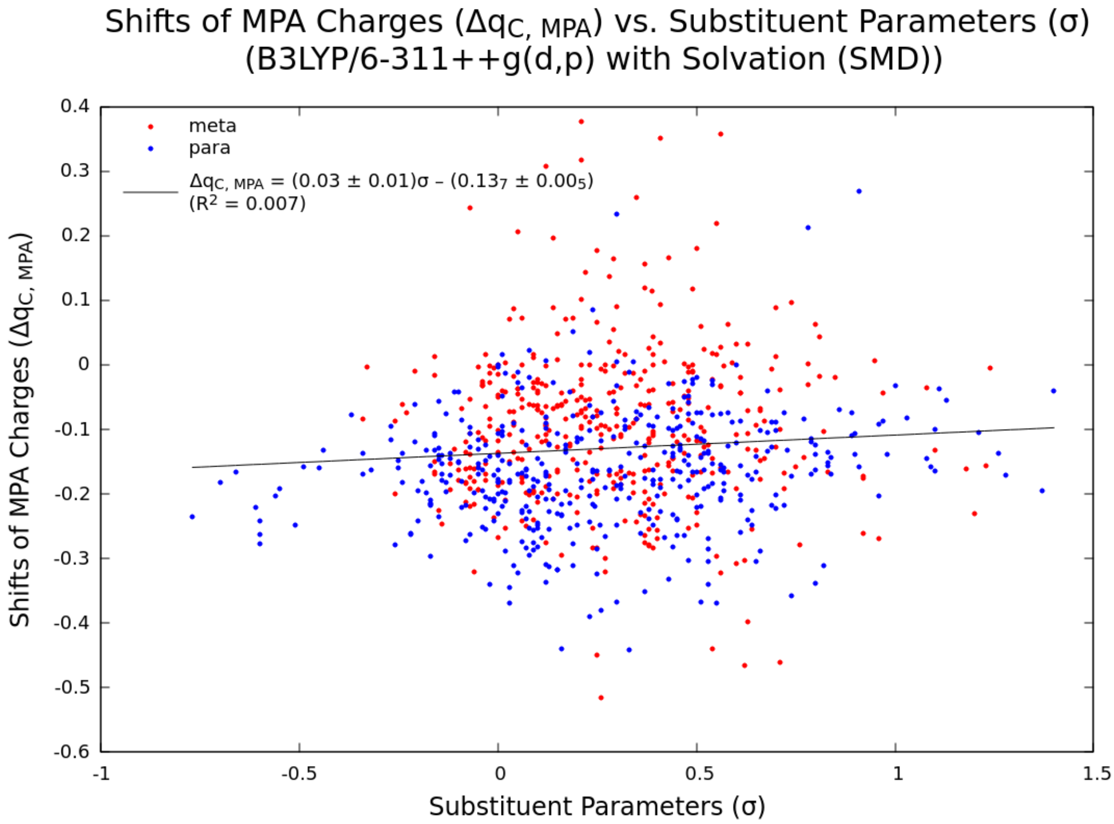

Shifts of MPA Charges ($\Delta q_{C,\ MPA}$) vs. Substituent Parameters (σ)
(B3LYP/6-311++g(d,p) with Solvation (SMD))

Figure S28: Shifts of Charges from MPA in B3LYP/6-311++g(d,p) and SMD Calculation

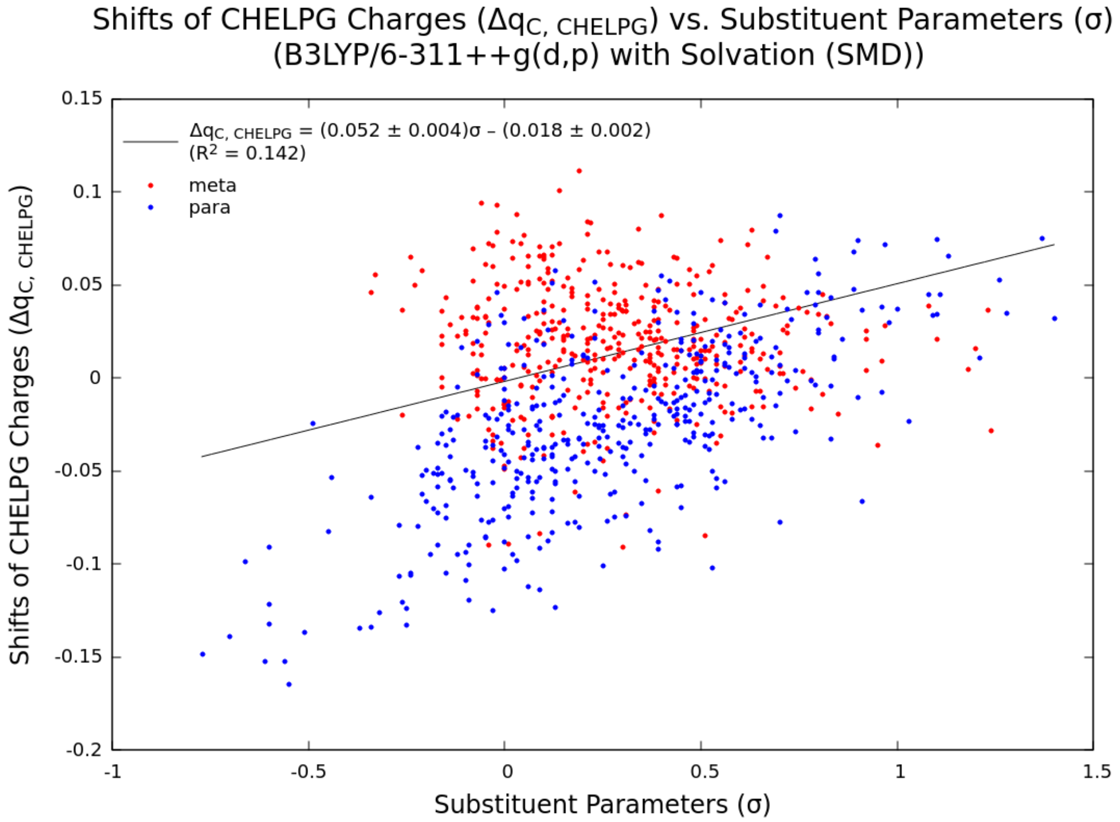

Shifts of CHELPG Charges ($\Delta q_{C,\ CHELPG}$) vs. Substituent Parameters (σ)
(B3LYP/6-311++g(d,p) with Solvation (SMD))

Figure S29: Shifts of Charges from CHELPG in B3LYP/6-311++g(d,p) and SMD Calculation

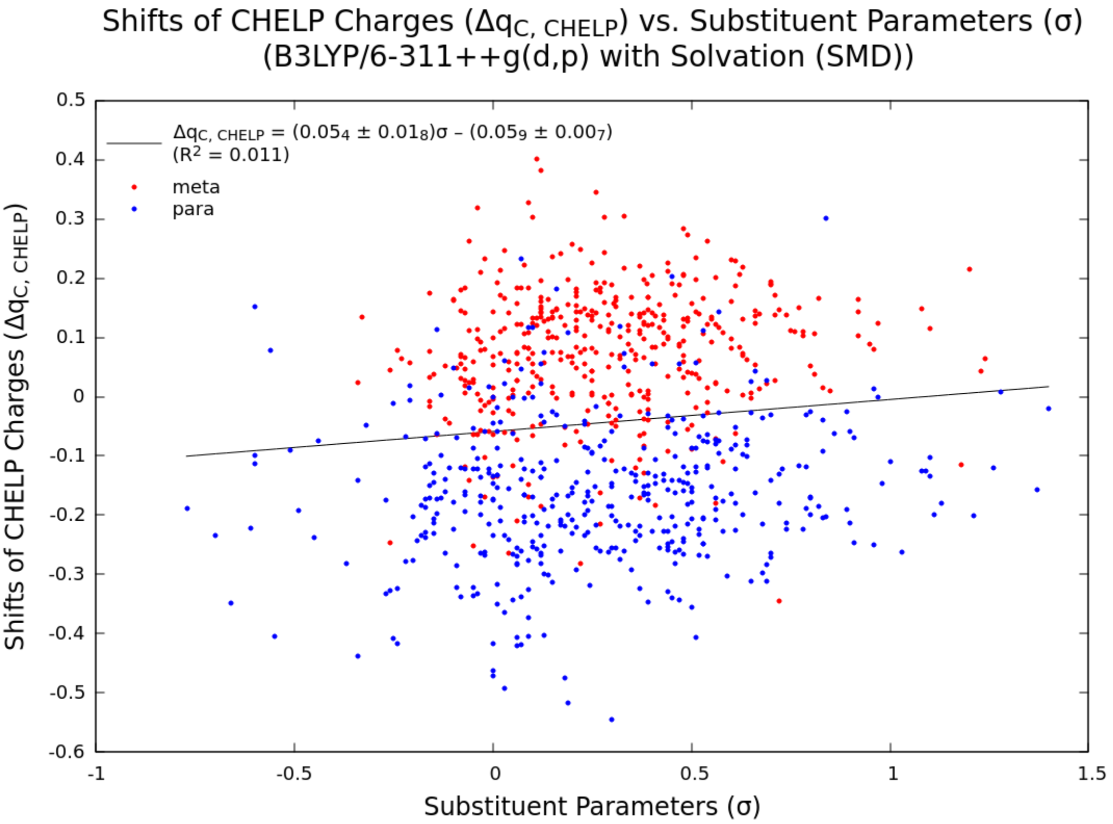


Figure S30: Shifts of Charges from CHELP in B3LYP/6-311++g(d,p) and SMD Calculation

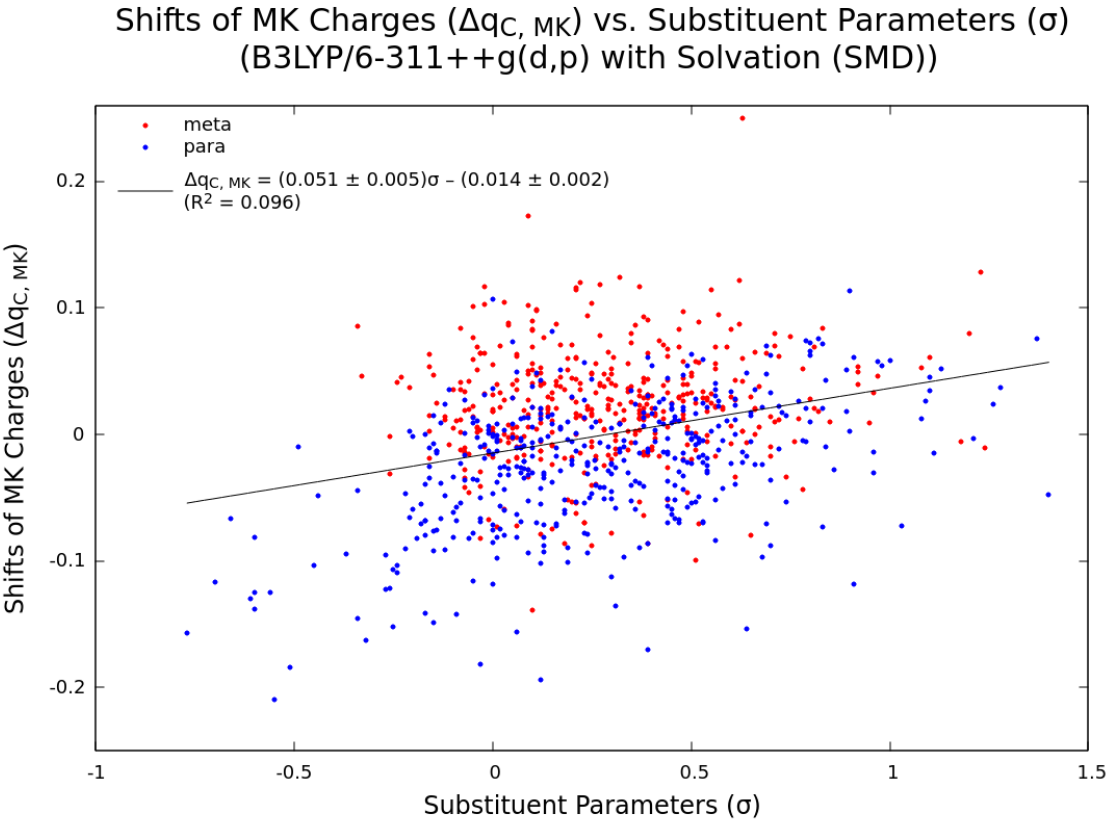


Figure S31: Shifts of Charges from MK Scheme in B3LYP/6-311++g(d,p) and SMD Calculation

§ Plots for Table 4

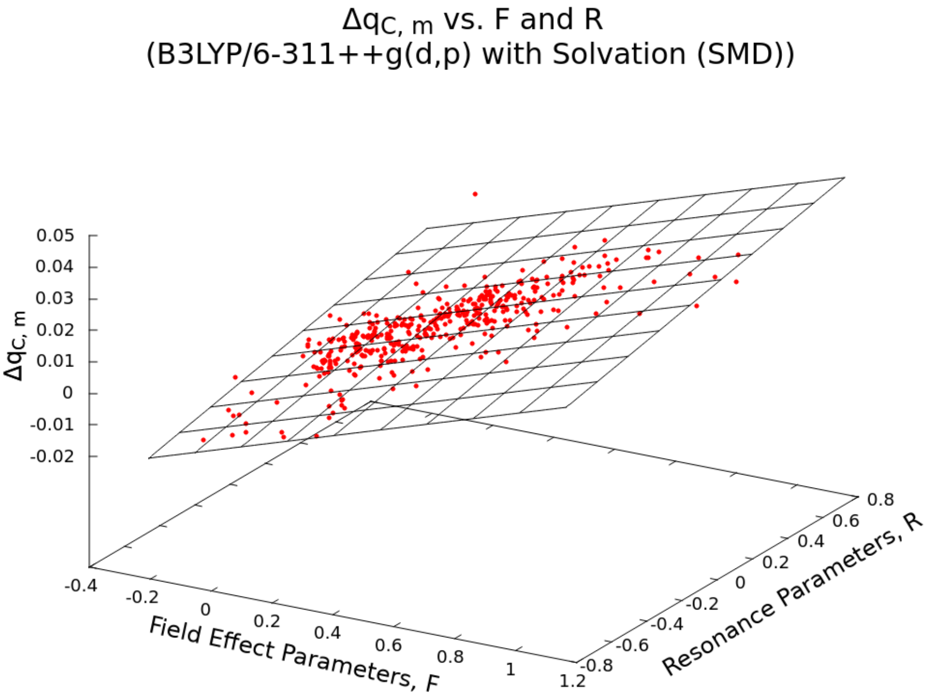

$\Delta q_{C,\,m}$ vs. F and R (B3LYP/6-311++g(d,p) with Solvation (SMD))

Figure S32: Double Linear Regression of Shifts of HPA Charges (of *meta* carbons, $\Delta q_{C,\,m}$) on Swain-Lupton parameters (F and R) in B3LYP/6-311++g(d,p) and SMD Calculation
Equation: $\Delta q_{C,\,m} = (0.0307 \pm 0.0007) * F + (0.0132 \pm 0.0008) * R + (0.0004 \pm 0.0003)$ ($R^2 = 0.894$)

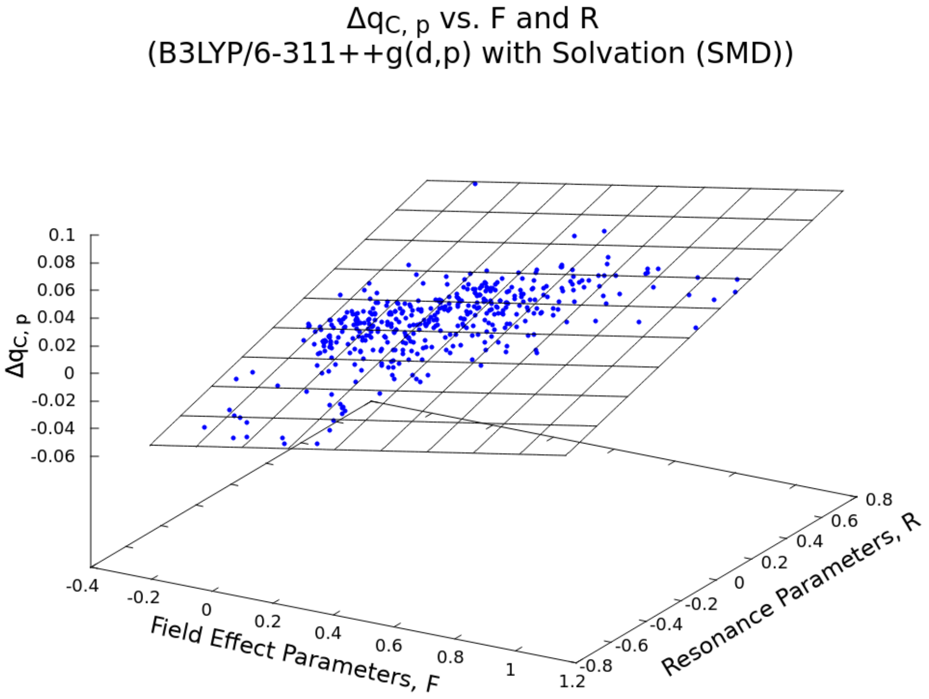

$\Delta q_{C,\,p}$ vs. F and R (B3LYP/6-311++g(d,p) with Solvation (SMD))

Figure S33: Double Linear Regression of Shifts of HPA Charges (of *para* carbons, $\Delta q_{C,\,p}$) on Swain-Lupton parameters (F and R) in B3LYP/6-311++g(d,p) and SMD Calculation
Equation: $\Delta q_{C,\,p} = (0.037_9 \pm 0.001_2) * F + (0.046_0 \pm 0.001_3) * R + (0.000_1 \pm 0.000_4)$ ($R^2 = 0.897$)

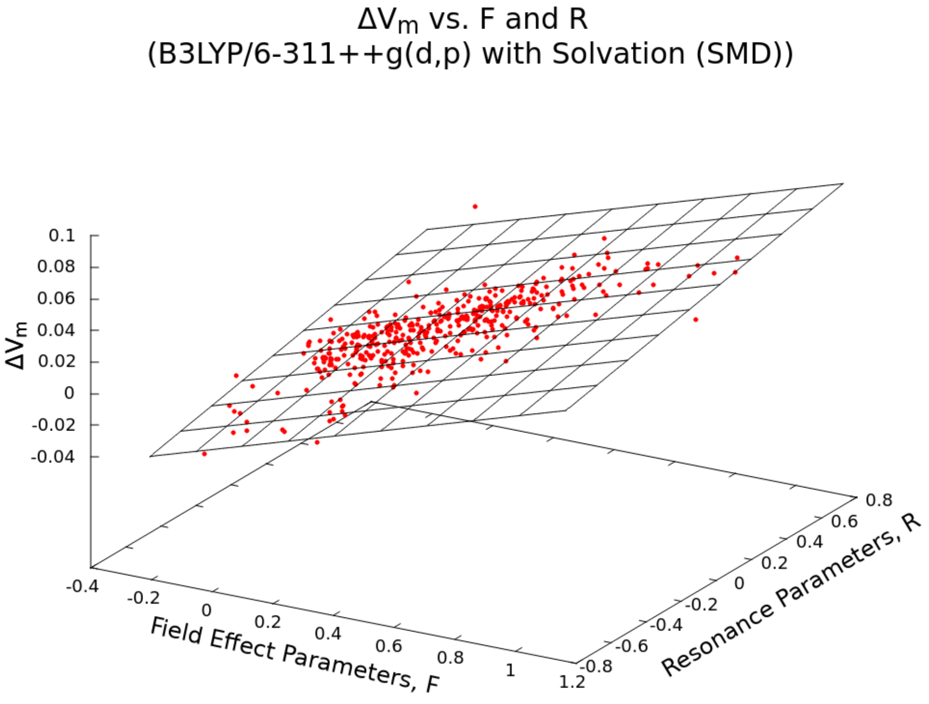


Figure S34: Double Linear Regression of Shifts of MESP (of *meta* carbons, ΔV_m) on Swain-Lupton parameters (F and R) in B3LYP/6-311++g(d,p) and SMD Calculation
Equation: $\Delta V_m = (0.059_1 \pm 0.001_3) * F + (0.025_2 \pm 0.001_4) * R - (0.000_9 \pm 0.000_5)$ ($R^2 = 0.896$)

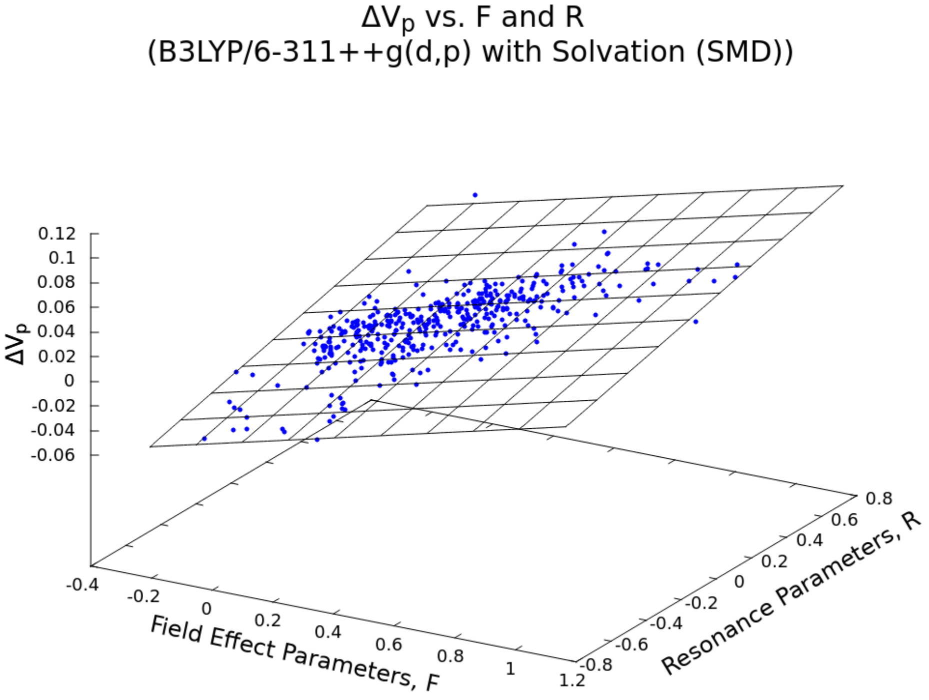


Figure S35: Double Linear Regression of Shifts of MESP (of *para* carbons, ΔV_p) on Swain-Lupton parameters (F and R) in B3LYP/6-311++g(d,p) and SMD Calculation
Equation: $\Delta V_p = (0.060_5 \pm 0.001_3) * F + (0.039_5 \pm 0.001_5) * R - (0.000_7 \pm 0.000_5)$ ($R^2 = 0.913$)

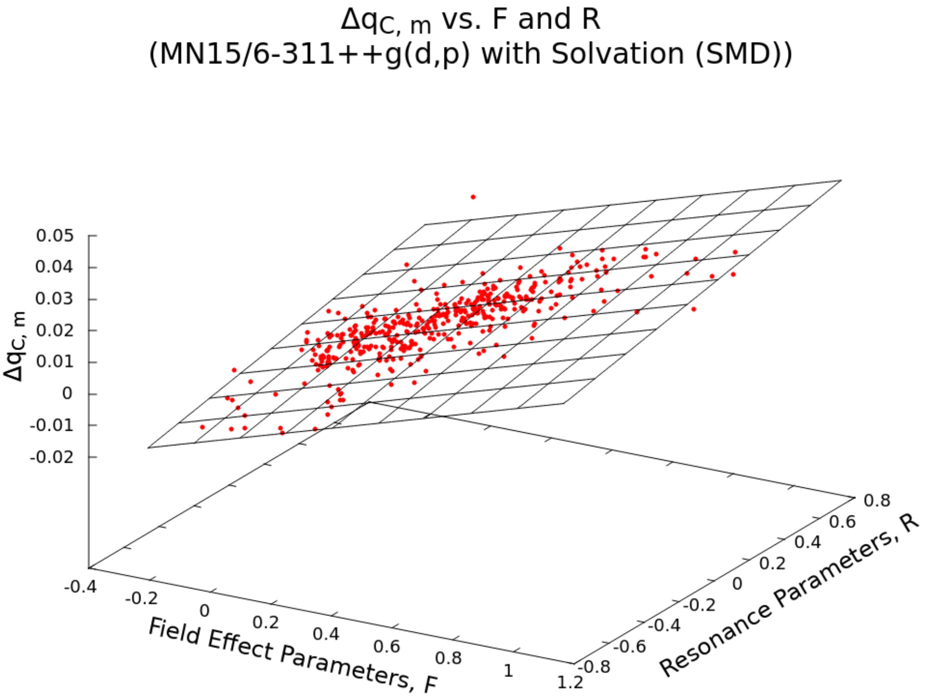


Figure S36: Double Linear Regression of Shifts of HPA Charges (of *meta* carbons, $\Delta q_{C,\,m}$) on Swain-Lupton parameters (F and R) in MN15/6-311++g(d,p) and SMD Calculation
Equation: $\Delta q_{C,\,m} = (0.0292 \pm 0.0007) * F + (0.0118 \pm 0.0007) * R + (0.0017 \pm 0.0002)$ ($R^2 = 0.896$)

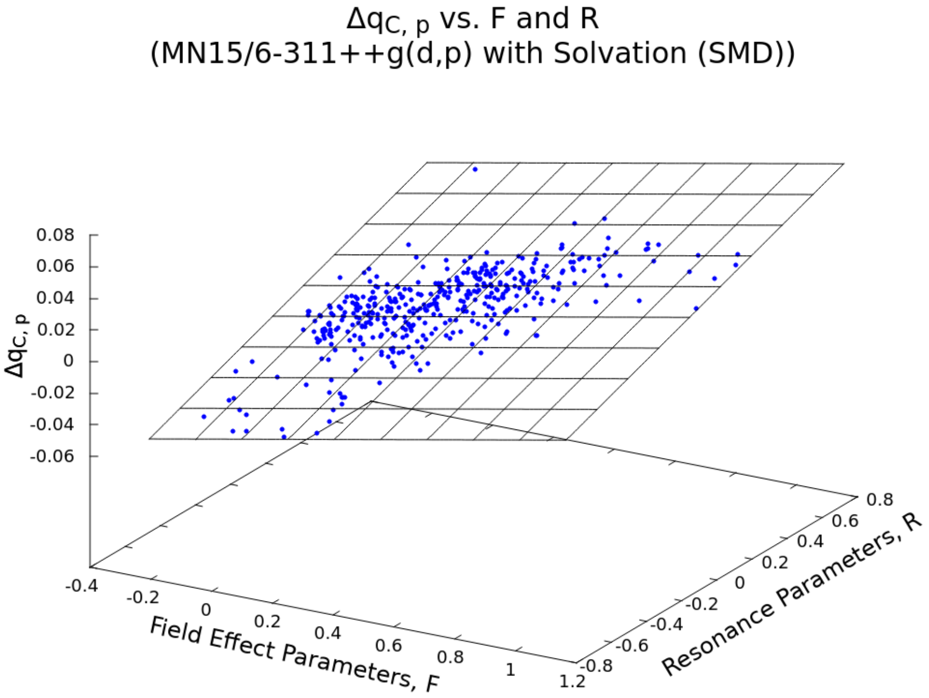


Figure S37: Double Linear Regression of Shifts of HPA Charges (of *para* carbons, $\Delta q_{C,\,p}$) on Swain-Lupton parameters (F and R) in MN15/6-311++g(d,p) and SMD Calculation
Equation: $\Delta q_{C,\,p} = (0.037_3 \pm 0.001_1) * F + (0.045_0 \pm 0.001_2) * R + (0.000_9 \pm 0.000_4)$ ($R^2 = 0.904$)

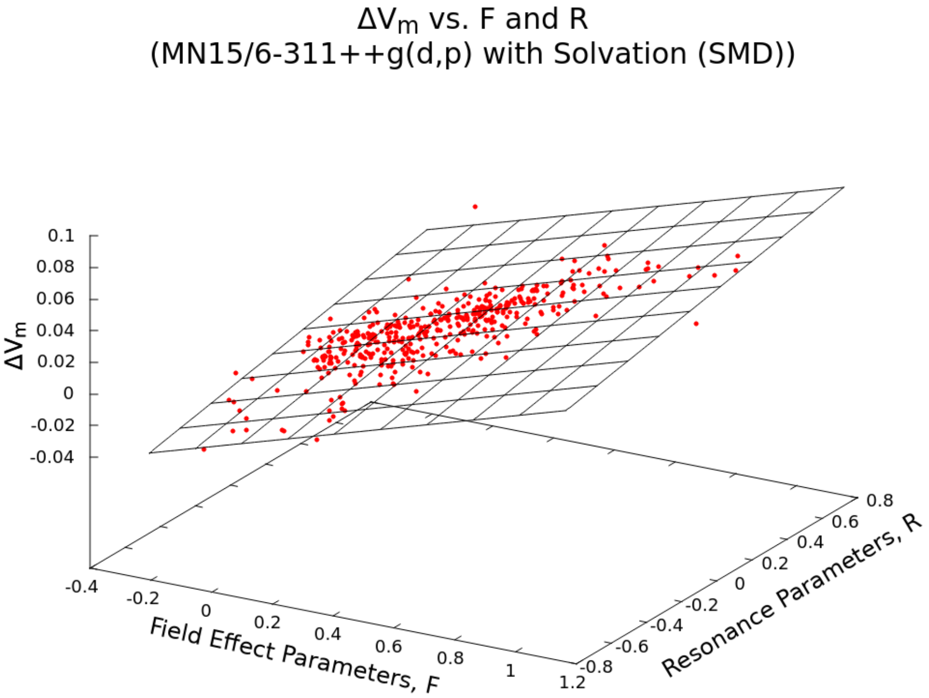


Figure S38: Double Linear Regression of Shifts of MESP (of *meta* carbons, ΔV_m) on Swain-Lupton parameters (F and R) in MN15/6-311++g(d,p) and SMD Calculation
Equation: $\Delta V_m = (0.057_4 \pm 0.001_2) * F + (0.023_7 \pm 0.001_4) * R - (0.000_0 \pm 0.000_5)$ ($R^2 = 0.895$)

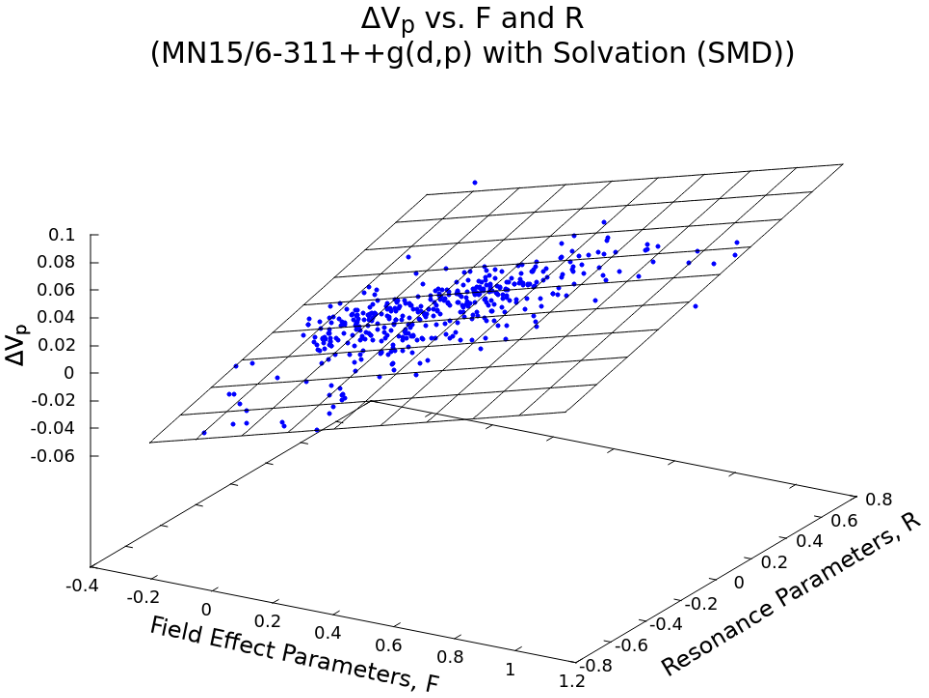


Figure S39: Double Linear Regression of Shifts of MESP (of *para* carbons, ΔV_p) on Swain-Lupton parameters (F and R) in B3LYP/6-311++g(d,p) and SMD Calculation
Equation: $\Delta V_p = (0.059_3 \pm 0.001_3) * F + (0.038_3 \pm 0.001_4) * R - (0.000_0 \pm 0.000_5)$ ($R^2 = 0.917$)

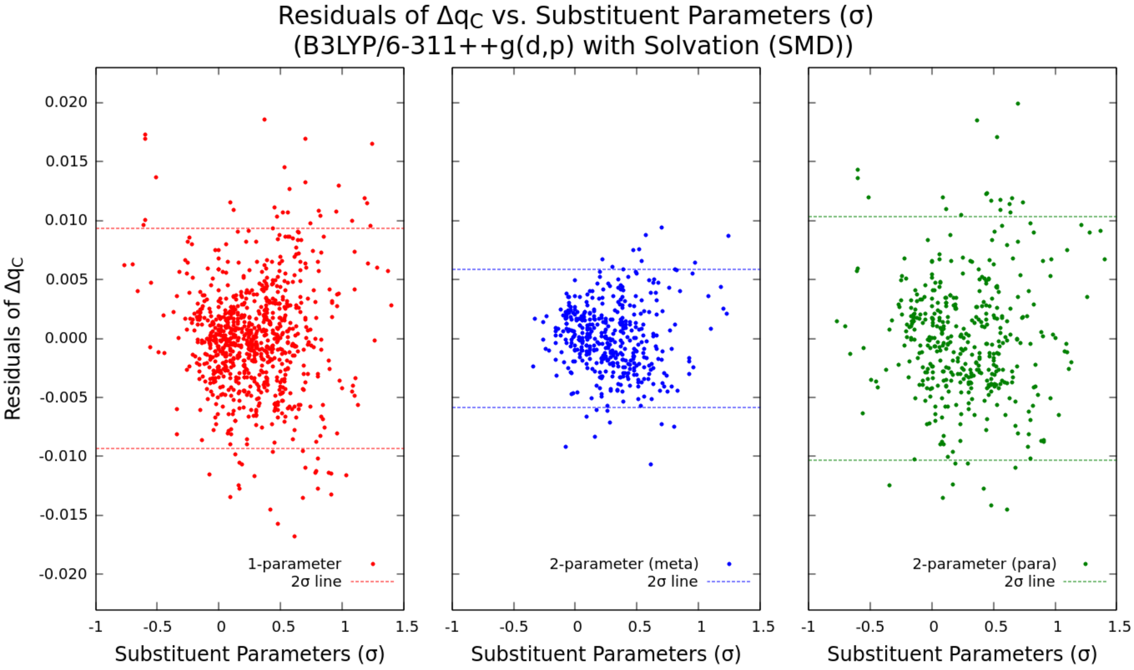


Figure S40: Distributions of Residuals from One-Parameter Plot and Two-Parameter Plots of HPA Charges (in B3LYP/6-311++g(d,p) and SMD Calculation)

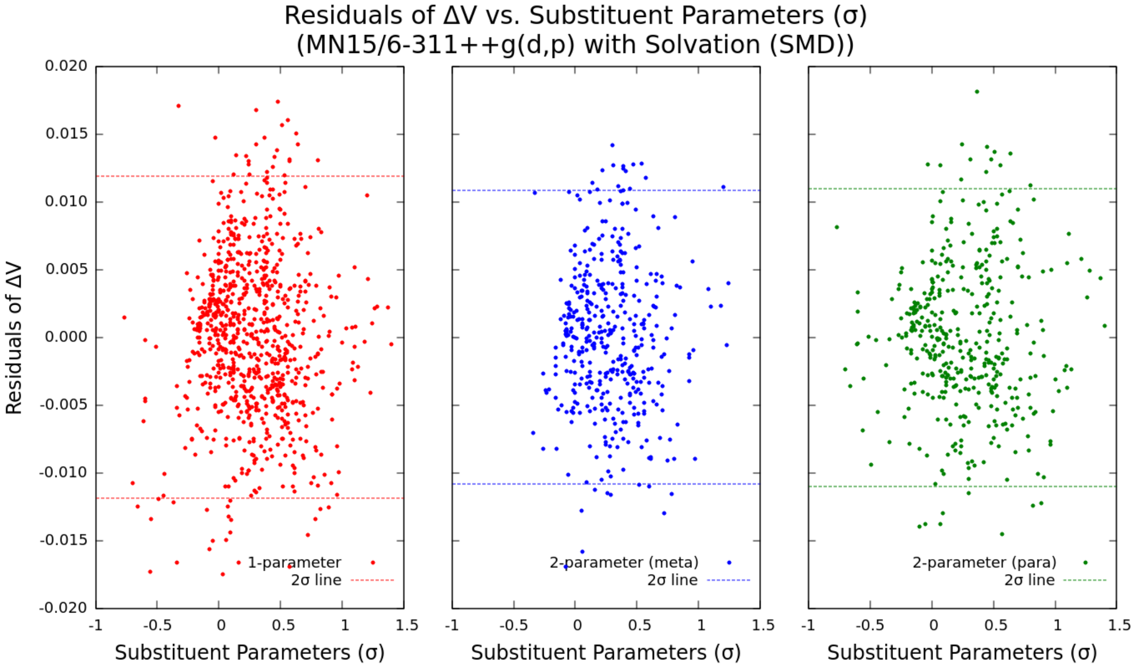


Figure S41: Distributions of Residuals from One-Parameter Plot and Two-Parameter Plots of HPA Charges (in MN15/6-311++g(d,p) and SMD Calculation)

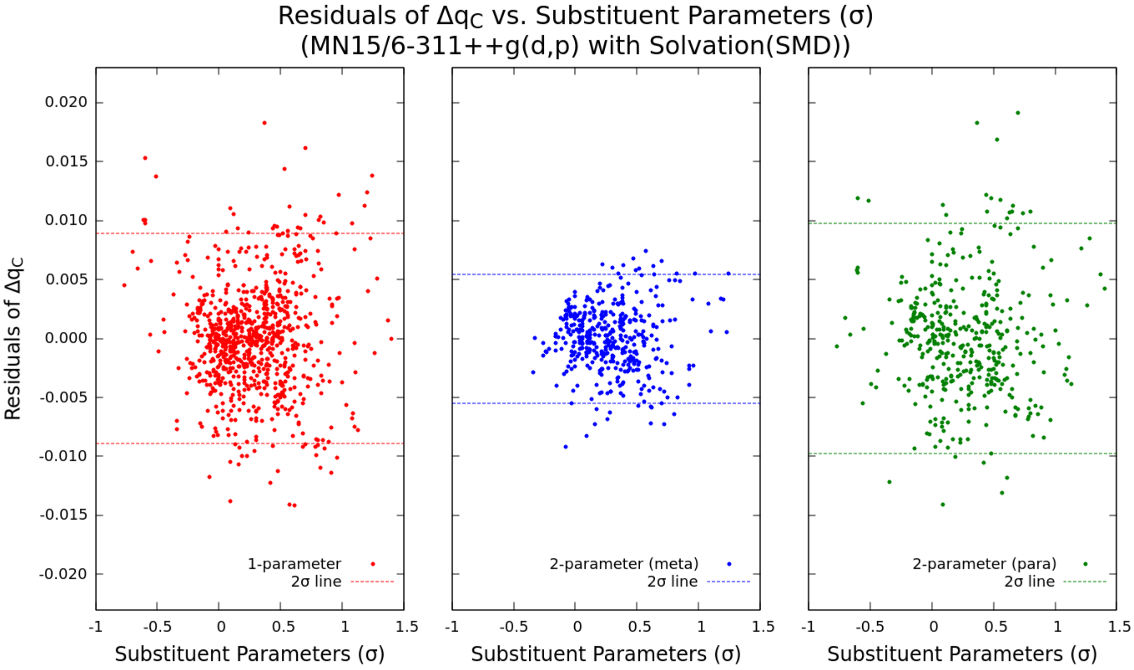


Figure S42: Distributions of Residuals from One-Parameter Plot and Two-Parameter Plots of HPA Charges (in MN15/6-311++g(d,p) and SMD Calculation)

Corrected Parameters from Ref 1

During the work, we found that some values from the Ref 1 should be replaced with values from other references[25–27], which better explicate the chemistry of corresponding functional groups. Hence, we list replaced values from the Ref 1 in “Replaced_Values” tab of the excel file, “Supplementary_Data.xlsx”.

Excluded Functional Groups for Calculations with LanL2DZ

During the study, we find that LanL2DZ basis set cannot generate appropriate structures for some (particularly, hypervalent) compounds. Therefore, it is inevitable to exclude some functional groups depicted below for the regression.

§ SF_5: compared to structures optimised with B3LYP/6-311++g(d,p), C–S bond length (1.97 Å vs. 1.85 Å) and S–F bond lengths (1.80 Å vs. 1.65 Å) are much exaggerated so that it implies much weaker bonds rather than expected.

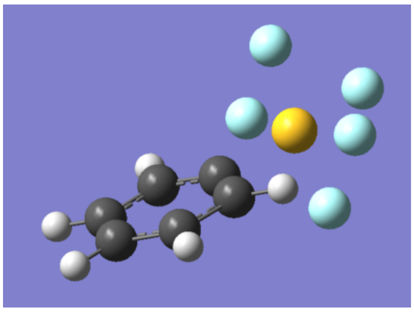


Figure S43: Structure of $PhSF_5$ optimised using B3LYP/LanL2DZ

§ OSO_2CF_3, $NHSO_2CF_3$, $N(SO_2CF_3)_2$, $CH_2SO_2CF_3$ and $N(Me)SO_2CF_3$: disruption of structures when attempt to make them optimised. LanL2DZ basis set has a fatal flaw that it cannot stably form structure of sulfonyl group.

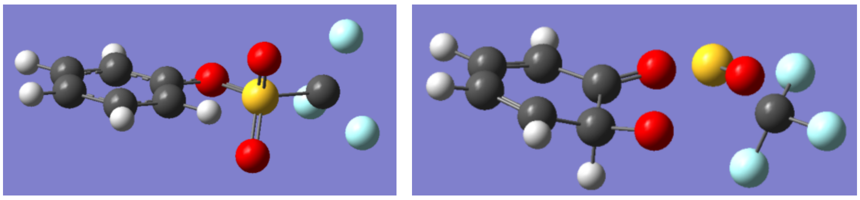


Figure S44: Structures of $PhOSO_2CF_3$ before optimisation (Left) and after the optimisation using B3LYP/LanL2DZ (Right)

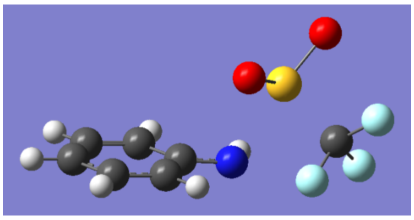


Figure S45: Structure of $PhNHSO_2CF_3$ optimised using B3LYP/LanL2DZ

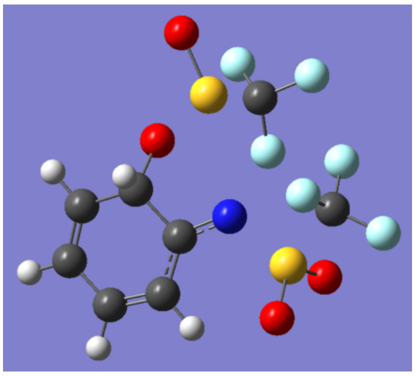


Figure S46: Structure of $PhN(SO_2CF_3)_2$ optimised using B3LYP/LanL2DZ

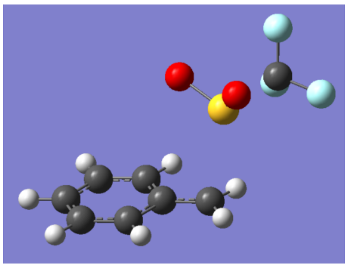


Figure S47: Structure of $PhCH_2SO_2CF_3$ optimised using B3LYP/LanL2DZ

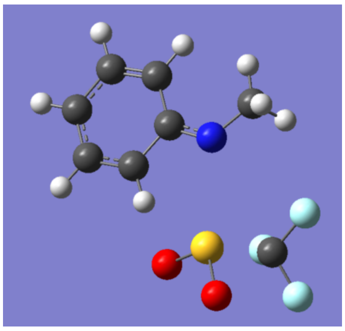


Figure S48: Structure of $PhN(Me)SO_2CF_3$ optimised using B3LYP/LanL2DZ

§ $N{=}C(CF_3)_2$: According to a previous study investigating IR spectrum of $PhN{=}C(CH_3)_2$, this kind of N-phenylimine structure should be strongly twisted at Ph–N bond[28] whilst another paper with MP2 calculation also suggested the dihedral angle of about 73 degrees at Ph–N bond.[29] Therefore, optimised structure with LanL2DZ in Figure S48 is not a feasible one so that we cannot obtain proper MESP shifts from the calculation with LanL2DZ.

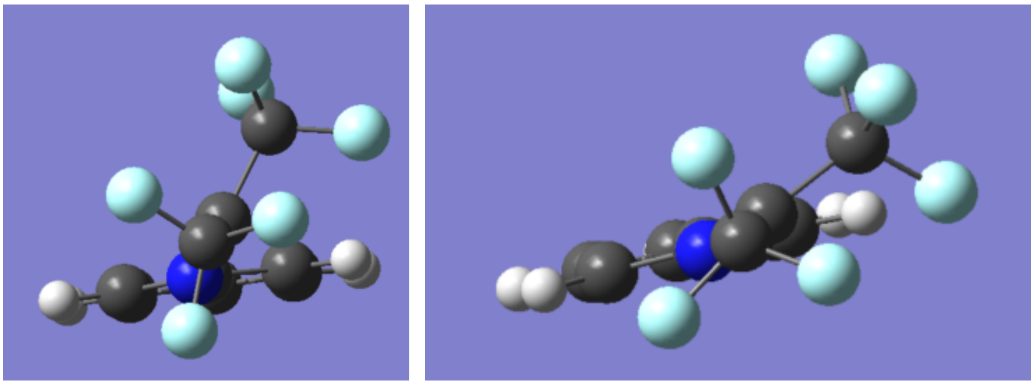


Figure S49: Structures of $PhN{=}C(CF_3)_2$ optimised using B3LYP/6-311++g(d,p) (Left) and optimised using B3LYP/LanL2DZ (Right)

Linear Regression on Gasteiger Charge Shifts with Hammett Parameters

Since parameters required to compute Gasteiger charges have been defined only for some cases, namely carbon, nitrogen, hydrogen, oxygen, sulfur and halogens, it is not possible to obtain them from the majority of the functional groups used in our manuscript, such as $-SF_5$. Therefore, we tried to calculate the shifts of Gasteiger charges from 45 selected functional groups listed in the Supplementary_Data.xlsx file with B3LYP/6-311++g(d,p) along with SMD solvation model of ethanol. Figure S50 clearly shows that there is no linear relationship between the charge shifts and the Hammett parameters.

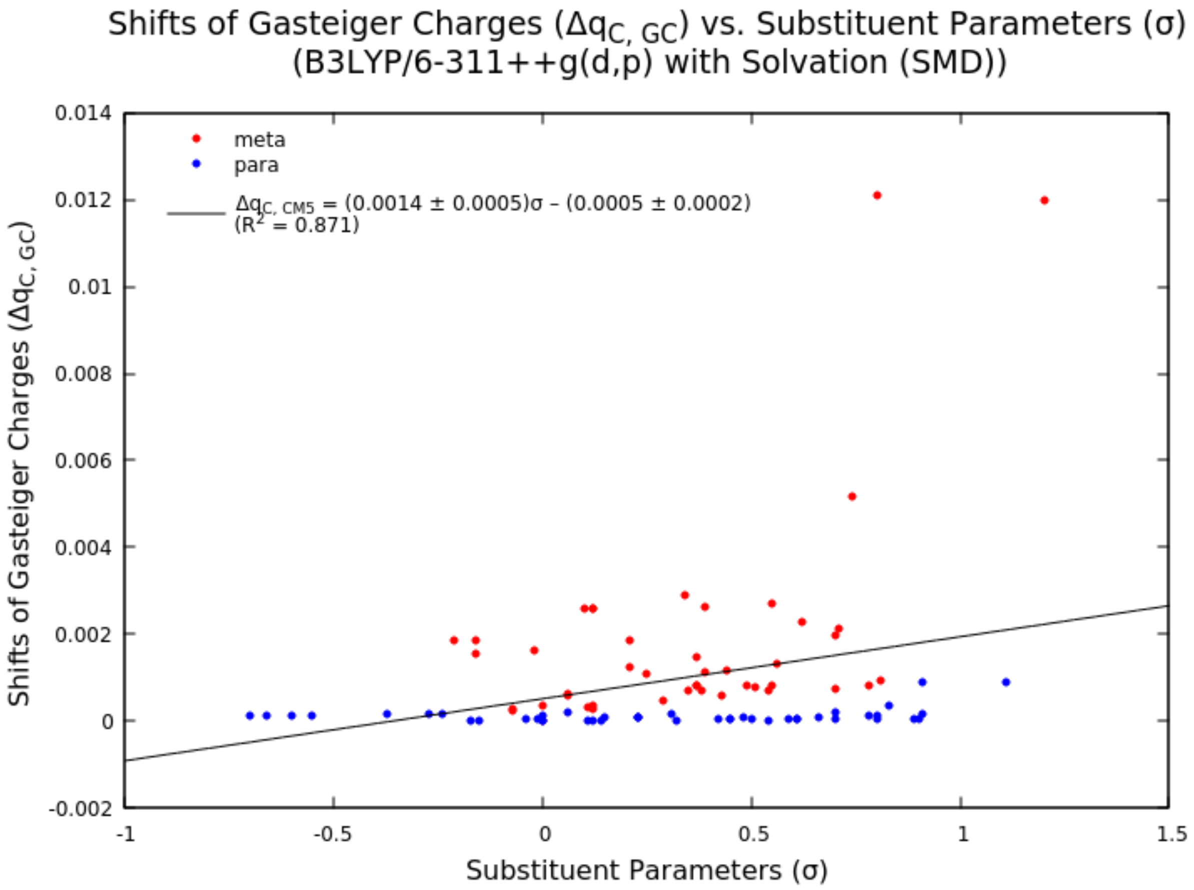


Figure S50 Shifts of Gasteiger charges for 45 selected functional groups using B3LYP/6-311++g(d,p) and SMD Calculation

References

(1) Hansch, C.; Leo, A.; Taft, R. *Chem. Rev.* **1991**, 91, 165–195.

(2) Foster, J.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, 102, 7211–7218.

(3) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, 83, 735–746.

(4) Bader, R. *Atoms in Molecules: A Quantum Theory;* International series of monographs on chemistry; Clarendon Press, **1990**.

(5) Lu, T.; Chen, F. *J. Comput. Chem.* **2012**, 33, 580–592

(6) Löwdin, P.-O. *J. Chem. Phys.* **1950**, 18, 365–375

(7) Löwdin, P.-O. *Advances in quantum chemistry*; Elsevier, **1970**; Vol. 5; pp 185–199.

(8) Bickelhaupt, F. M.; van Eikema Hommes, N. J.; Fonseca Guerra, C.; Baerends, E. J. *Organometallics*. 1996, 15, 2923–2931.

(9) Fonseca Guerra, C.; Handgraaf, J.-W.; Baerends, E. J.; Bickelhaupt, F. M. *J. Comput. Chem.* 2004, 25, 189–210.

(10) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2012**, 8, 527–541

(11) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. *J. Chem. Phys.* **2007**, 126, 144111.

(12) Bultinck, P.; Ayers, P. W.; Fias, S.; Tiels, K.; Van Alsenoy, C. *Chem. Phys. Lett.* **2007**, 444, 205–208.

(13) Van Damme, S.; Bultinck, P.; Fias, S. *J. Chem. Theory Comput.* **2009**, 5, 334–340.

(14) Lu, T.; Chen, F. *J. Theor. Comput. Chem.* **2012**, 11, 163-183.

(15) Becke, A. D. *J. Chem. Phys.* **1988**, 88, 2547–2553.

(16) Mulliken, R. S. *J. Chem. Phys.* **1955**, 23, 1833–1840.

(17) Mulliken, R. S. *J. Chem. Phys.* **1955**, 23, 1841–1846.

(18) Gasteiger, J.; Marsili, M. *Tetrahedron*. **1978**, 34, 3181–3184.

(19) Gasteiger, J.; Marsili, M. *Tetrahedron*. **1980**, 36, 3219–3228.

(20) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, 11, 361–373.

(21) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, 8, 894–905.

(22) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, 5, 129–145.

(23) Besler, B. H.; Merz Jr, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, 11, 431–439.

(24) Gasteiger, J.; Marsili, M. *Tetrahedron*, **1980**, 36, 3219–3228.

(25) Jaffe, H. H. *Chem. Rev.* **1953**, 53, 191–261

(26) Romanenko, V. D.; Kalibabchuk, N. N.; Rositskii, A. A.; Zalesskaya, V. G.; Didkovskii, V. E.; Iksanova, S. V. *Chem Heterocycl Compd*, **1976**, 12, 753–757.

(27) Exner, O.; Lakomý, J. *Collect. Czechoslov. Chem. Commun.*, **1970**, 35, 1371–1386.

(28) Krauss, W.; Wagner-Bartak, C. *Tetrahedron Lett.*, **1968**, 9, 4799-4800.

(29) Afonin, A. V.; Pavlov, D. V.; Ushakov, I. A.; Levanova, E. P.; Levkovskaya, G. G. *Russ. J. Org. Chem.,* **2013**, 49, 1117-1121.