Supporting Information

Predicting the Magnitude of σ-holes Using VmaxPred, a Fast and Efficient Tool Supporting the Application of Halogen Bonds in Drug Discovery

Johannes Heidrich[†], Thomas E. Exner^{†,‡} and Frank M. Boeckler^{†,‡,*}

+ Molecular Design and Pharmaceutical Biophysics, Institute of Pharmaceutical Sciences, Eberhard Karls University Tuebingen, Auf der Morgenstelle 8, 72076 Tuebingen, Germany

‡ Center for Bioinformatics Tuebingen (ZBIT), Eberhard Karls University Tuebingen, Sand 1, 72076 Tuebingen, Germany

* frank.boeckler@uni-tuebingen.de

- I. QM Calculations
- II. Feature Vector Property Calculation
- III. Illustration of the Treatment of Annulated Ring Systems
- IV. SVM Hyper Parameter Optimization
- V. Dataset Creation
- VI. Deduction of a Charge Distance V_{max} Relationship
- VII. Adduct Formation Energy Calculation
- VIII. Outlier of Correlation of V_{max} and Adduct Formation Energy
- IX. Correlation of Predicted V_{max} and Adduct Formation Energy
- X. Electron Isodensity-dependent Results of Fit of Adduct Formation Energy
- XI. References
- XII. Appendix: Electron Isodensity-dependence of Adduct Formation Energy

I. QM Calculations

Quantum mechanical calculations were performed using TURBOMOLE Version 6.4, 7.1 and 7.2¹⁻⁴. Atomic information of xyz files was transformed into coord files using TURBOMOLE x2t script. Input files were generated with define. Redundant internal coordinates were generated (ired) followed by a single-point calculation with universal force filed (UFF) using ff with default parameters. For charged structures charge was set to the externally determined formal charge (c). Furthermore, bond analysis was carried out (banal). Atomic basis set was assigned to def2-TZVPP for all atoms (b all def2-TZVPP)^{5,6}. Molecular orbitals and occupation numbers were taken from extended Hückel guess with default parameters, except for charged structures, where charge was set to formal charge of the molecule (eht). SCF iteration limit (iter) was set to 1000 with an accuracy of SCF energy (conv) of 10⁻⁸ hartree. RICC2 defaults were modified by treating orbitals with eigenvalues smaller than -3.0 as inert using cc2 and freeze keywords. Furthermore auxiliary basis sets were assigned using cbas^{7,8}. Maximum core memory was set to 75 % of maximum available system random access memory (memory). Geometry optimization (or energy minimization) was done using jobex level cc2 with a maximum of cycles of 1,000 and the convergence criteria of the Cartesian gradient norm set to 10⁻⁴. GRIDPARM was used to generate vector definitions of a surrounding box of a given molecule with an overlap of 5 au in each Cartesian direction. The equidistant grid point distance was 0.2 au. The grid vector definitions and the origin were inserted into control file: For electron density calculations into *spointval* and for ESP calculations into *spointval pot* section, respectively. Single point energy calculations and electrostatic potential calculations were done by calling dscf^{1,9,10} followed by ricc2¹¹⁻¹³.

Molecular transformations for visualization purpose, i.e. creation standardized orientation, were done with coord data by an in-house Java application in atomic units with double precision. MOLCAD II v1.4 was used for visualization purposes of electrostatic potentials and electron densities¹⁴.

Adduct formation energies of adducts E_{ad} were calculated as follows:

$$E_{ad} = E_{Complex} - (E_{BB} + E_{Lig})$$

where $E_{complex}$ is the total single point energy of the backbone model system (*N*-methylacetamide) and the ligand, E_{BB} and E_{Lig} is the total energy of the same backbone or the ligand geometry, respectively. A distance scan was applied in order to calculate a close-to-global-minimum adduct formation energy while avoiding secondary interactions. The X-C axis of (hetero)aryl halides was oriented 120° to the carbonyl function of *N*-methylacetamide. The dihedral angle defined by the plane of the amide of the backbone model system and the plane of the aromatic ring of the σ -hole donor was set to 90°. The distance between the carbon of C-X and the oxygen of the backbone model system was increased from 4 Å to 6.8 Å by 0.1 Å (Figure S8). Global minima were approximated by calculating the minimum of a fitted polynomial of order 2 of the smallest energy values in R.

II. Feature Vector Property Calculation

Group electronegativity¹⁵ was calculated for each atom in central aromatic ring:

$${}_{\chi}G = \left\{ \underbrace{\frac{1}{n_{1l}} \sum_{l=1}^{n_{1l}} \frac{1}{n_{2l}}}_{\substack{l=1 \\ \underbrace{\sum_{l=1}^{n_{2l}} \dots \left(\underbrace{\frac{1}{n_{kl}} \sum_{l=1}^{n_{kl}} X_{kl}}_{equilibrium \ level \ k} \right)}_{equilibrium \ level \ 1}} \dots \right\}$$

where X_{kl} is the Pauling electronegativity for atom l in sphere k and $n_{1l}, n_{2l}, ..., n_{kl}$ are the sum of atoms that are directly attached to the k-th root atom. Annulated rings were treated as substructures with two or more roots.

Lone-pair electrons index^{16,17} was calculated for each residue attached to the central aromatic ring as follows:

$$LEI = \frac{1}{k} \sum_{i=1}^{P} \sum_{j=1}^{K} \frac{LE_{i}}{D_{ij}^{2}}$$

where *K* is the number of vertices in heavy atom graph, *P* is the number of heteroatoms and D_{ij} is the topological distance between vertex *i* and *j*. *LE*_{*i*} was used to calculate the lone-pair electrostatic interaction as follows:

$$LE_i = \frac{\sqrt{L_i} * (Z_i^v - Z_i^b)}{X_i^{PA}}$$

Where X_i^{PA} is the Pauling electronegativity of the i-th atom, Z_i^{ν} is the number of valence electrons of the ith atom and Z_i^{b} is the number of bonding electrons and L_i the principle quantum number.^{16,17}

Electrotopological state indices were calculated as follows:

$$S_i = I_i + \sum_{j=1}^{A} \frac{I_i - I_j}{(d_{ij} + 1)^2}$$

Where d_{ij} is the topological distance of the ith and jth atom, A the number of atoms and the instrinsic state I_i :

$$I_i = \frac{\left(\frac{2}{L_i}\right)^2 * d_i^v + 1}{d_i}$$

Where L_i is the principle quantum number, d_i^v is the number of valence electrons and d_i is the number of sigma electrons of the i-th atom.^{16,18–20}

III. Illustration of the Treatment of Annulated Ring Systems



Figure S1. Exemplified decomposition of heteroaromatic ring in order to calculate substituent attributes of group electronegativity (G) and lone-pair electrostatic interaction (Li), as well as Pauling electronegativity and lone-pair electrons index of atoms of central aromatic ring in ortho (Het_O) and meta (Het_M) to the halogen atom (I).

IV. SVM Hyper Parameter Optimization

Weka 3.8.2 was used with SMOreg for regression using regression optimizer RegSMOImproved^{21,22}.



Figure S2. Grid search of SVR complexity (C) and RBF Kernel parameter (γ) evaluated in 2-fold CV by correlation coefficient (A and C) and RMSE (B and D) for calculating V_{max} at 0.001 au. A and B: Grid search for training set, C and D: Grid search for complete data set (set 1-3).



Figure S3. Grid search of SVR complexity (C) and RBF Kernel parameter (γ) evaluated in 2-fold CV by correlation coefficient (A and C) and RMSE (B and D) for calculating V_{max} at 0.020 au. A and B: Grid search for training set, C and D: Grid search for complete data set (set 1-3).

Three dimensional surface plots were created with Python 3 and Matplotlib²³.

5

V. Dataset Creation

Table S1. Overview of heterocycle composition in training set for V_{max} prediction. #HA number of heteroatoms in five- or six-membered ring and their annulated equivalents. #Structures number of selected structures in order to fulfil the target distribution defined by portion.

#HA	#Structures	Portion
1	10	
2	85	2/3
3	171	
4	67	1/6
5	67	1/6
> 5	0	0
	<u>400</u>	<u>1</u>

Structures used as test set are available as separate SMILES file.



Negatively Charged

Figure S4. Influence of negative and positive charges to V_{max} at 0.02 au in halo benzene derivatives with quaternary amine or carboxylate side chains with respect to the topological distance. Annotated linear curve fitting equation and coefficient of determination with standard error as grey shades.

Negatively Charged





Figure S5. Influence of negative and positive charges to V_{max} at 0.001 au in halo benzene derivatives with quaternary amine or carboxylate side chains with respect to the topological distance. Annotated linear curve fitting equation and coefficient of determination with standard error as grey shades.

Α



Figure S6. Euclidean distances of formal charged atoms to halogen after unconstraint geometry optimization. A 4-(3-bromophenyl)-N,N,N-trimethylbutan-1-aminium with positively charged quaternary amine in 9.4 Å Euclidean distance and a topological distance of 8 to bromine. B 4-(3-bromophenyl)butanoate with 6.5 and 4.5 Å Euclidean distance and a topological distance of 9 of formally charged oxygen atoms to bromine.



Figure S7. Structural formulas of exemplarily selected ligands of test set (**set 2**) for machine learning (**A**) and their alignment (**B**) for the distance scan approach (**C-E**). Angle between carbonyl group of protein backbone model system *N*-methylacetamide was set to 120° and dihedral angle between aromatic ring system of σ -hole donor to Lewis base' amide plane was set to 90°. Minimal energy was determined by regression analysis of global minimum of energies of distance scan approximated by quadratic equation in R.



Figure S8. Correlation of Adduct Formation Energy and calculated V_{max} at electron isodensity level of 0.020 au (A) and 0.001 au (B). Annotated the polynomial function of degree 3 of non-linear regression and the resulting RMSE in kJ/mol. Outlier that was removed for final determination of correlation depicted in C. Secondary interaction shown as yellow dashed line.





Figure S9. Correlation of Adduct Formation Energy and predicted V_{max} at electron isodensity level of 0.020 au (**A**) and 0.001 au (**B**). Annotated the polynomial function of degree 3 of non-linear regression and the resulting RMSE in kJ/mol. A single outlier was removed (see Supporting Information, Figure S8, C).



--- CI-Linear --- Br-Linear --- I-Linear --- Linear --- Quadratic --- Cubic

Figure S10. Selected electron isodensities of 0.001 au, 0.018 au and 0.050 au (from outer shell to inner shell) and electrostatic potential in plane from -0.01 au (blue) via 0 au (white) to 0.01 au (red) of chloro-, bromo- and iodobenzene. Regression analysis of correlation of V_{max} , evaluated on electron isodensity levels from 0.001 au to 0.050 au in steps of 0.001 au, to adduct formation energy by linear fit of each halogen atom type separately (Cl-Linear, Br-Linear, I-Linear) and over all atom types (Linear, Quadratic and Cubic). Goodness of fit evaluated by R².



Figure S11. Selected electron isodensities of 0.001 au, 0.018 au and 0.050 au (from outer shell to inner shell) and electrostatic potential in plane from -0.01 au (blue) via o au (white) to 0.01 au (red) of chloro-, bromo- and iodobenzene. Regression analysis of correlation of V_{max} , evaluated on electron isodensity levels from 0.001 au to 0.050 au in steps of 0.001 au, to adduct formation energy by linear fit of each halogen atom type separately (Cl-Linear, Br-Linear, I-Linear) and over all atom types (Linear, Quadratic and Cubic). Goodness of fit evaluated by RMSE.

Table S2. AICc values of tested polynomial functions (from linear to cubic) as regression fit of electron isodensities from 0.001 au to 0.050 au of the correlation of adduct formation energy to V_{max}.

	AICc values													
							$a \cdot x^3 + b \cdot x^2 + c \cdot x$	$a \cdot x^3 + b \cdot x^2 +$						
ISODENSITY	a·x + b	a∙x	a·x² + b·x + c	a·x² + b·x	a·x² + c	a∙x²	+ d	c∙x	$a \cdot x^3 + b \cdot x^2 + d$	a∙x³ + b∙x²	a ⋅ x³ + c ⋅ x + d	a·x³ + c·x	a ⋅x³ + d	a∙x³
0.001	2755.717199	3690.797391	2667.029525	3556.778936	3564.122441	3916.728126	2664.825649	3468.082076	3880.434216	3312.538254	2738.001587	3692.3972	3311.204043	3957.647161
0.002	2725.07417	3506.04627	2618.110299	3383.125227	3494.900276	3838.979275	2616.310065	3255.839112	3825.773433	3304.207883	2690.552996	3507.269816	3302.211649	3924.391671
0.003	2702.969194	3334.318384	2583.842275	3223.907044	3425.936036	3762.676435	2582.084279	3078.961789	3763.037038	3296.842007	2653.714472	3336.003449	3295.780204	3895.944649
0.004	2685.529918	3171.773952	2557.101144	3071.185871	3357.769534	3686.237041	2555.268742	2926.569591	3686.050447	3282.966598	2623.381249	3173.786644	3287.886907	3868.655006
0.005	2669.76749	3023.20835	2535.063316	2929.358741	3292.811737	3612.5491	2532.915036	2798.957436	3595.791503	3260.533209	2597.5279	3024.065341	3280.519441	3843.623834
0.006	2656.176141	2888.412177	2515.04509	2795.144816	3228.547446	3537.113084	2512.88011	2690.080766	3489.359824	3221.943986	2573.678787	2884.548683	3268.66474	3816.237322
0.007	2646.091233	2778.698494	2501.456361	2681.60388	3169.496571	3464.906276	2498.96082	2605.495469	3375.265538	3171.528841	2556.641481	2765.006264	3258.330181	3790.454113
0.008	2633.950713	2693.476266	2485.369734	2586.700191	3111.747464	3392.921839	2482.805556	2538.542292	3254.908964	3106.319322	2537.138958	2662.726483	3245.314488	3763.515752
0.009	2624.395268	2639.03341	2472.040091	2516.594645	3056.273189	3320.39662	2469.572726	2491.073333	3134.356479	3031.378947	2520.477137	2583.330886	3229.937496	3734.438668
0.01	2617.95618	2616.195608	2463.407644	24/4.//43/9	3004.99659	3249.651679	2461.019863	2464.016909	3020.67445	2953.643255	2508.601793	2530.563708	3214.441918	3704.883303
0.011	2610.120004	2615.497203	2453.2696	2451.955081	2956.237751	3180.78439	2451.025328	2448.996558	2916.1/3855	28/5./93535	2495.487311	2497.39055	3198.009619	3675.050159
0.012	2602.885646	2632.195978	2444.47217	2444.999617	2910.572458	3113.708523	2442.417155	2445.362764	2823.457863	2802.180949	2483.790444	2482.20375	3181.197483	3644.708818
0.013	2590.008924	2000.390238	2437.144716	2446.00721	2807.812701	3046.247068	2435.529166	2450.115018	2745.692946	2/35.41655	24/3.064514	2460.645061	3105.915251	2591 40995
0.014	2592.200015	2090.207492	2432.052290	2459.249075	2020.14924	2964.600462	2430.529790	2400.901871	2076.506505	20/7.515457	2405.005005	2469.751607	2127 246.051297	3501.49005
0.015	2505.22514	2733.341323	2420.047030	2472.930490	2751.155542	2925.425065	2427.510075	2474.402031	2020.955570	2028.80333	2455.702065	2504.550555	2109 666120	2514 20521
0.010	2585 018078	2774.552107	2420.440002	2501 636206	2737.3713	2803.385707	2425.705575	2407.550552	2560 165443	2567 389604	2454.52057	25/1 00/735	3000 507351	3480 000433
0.017	2583.018078	2849 009277	2420.040041	2513 915116	2600 800170	2764 535091	2425.005407	2510 050516	2540 982997	2537.389004	2431.373413	2560 690806	3072 375052	3480.330433
0.019	2582 269567	2883 138327	2426 103592	2524 673189	2674 965857	2720 75199	2425.003548	2516 862643	2529 203358	2512 846286	2445.151178	2578 49274	3054 246796	3413 566757
0.013	2581 985126	2005.150527	2420.103332	2534 268339	2652 876192	2682 430498	2420.483341	2521 57241	2523 548667	2498 639924	2447.314353	2595 408004	3036 492109	3379 820589
0.021	2582 537947	2945 433326	2430 456369	2542 896057	2633 651111	2650 070466	2431 496304	2524 733127	2522 846351	2498.852888	2440.721102	2611 298936	3019 210185	3346 239016
0.021	2583 805723	2973 633992	2430.430303	2550 497456	2616 951389	2623 649312	2435 466908	2526 742959	2525 844514	2482 624738	2449 179252	2625 844649	3002 228101	3312 679275
0.022	2585 75882	3000 023345	2438 581527	2557 018348	2602 305969	2602 855754	2440 170085	2528 026683	2531 49349	2479 26878	2451 682901	2638 742776	2985 068776	3278 598218
0.024	2588.31461	3024.650452	2443.984375	2563.030703	2590,244483	2588,296843	2445.776177	2529,246973	2539,284171	2478,455206	2455.270057	2650.598196	2968.628624	3245.086096
0.025	2591,503028	3047.643239	2450.354466	2568,74981	2580,695109	2579.691435	2452,289607	2530,856259	2548,707582	2479.877265	2459.940613	2661.520501	2952,984607	3212,292741
0.026	2595.346686	3069.133534	2457.731219	2574.435402	2573.643965	2576.657237	2459.749311	2533.228787	2559.408886	2483.299709	2465.730291	2671.691713	2938.268719	3180,429204
0.027	2599.807602	3089,259179	2466.011563	2580.218103	2568.924049	2578,565218	2468.052727	2536.526631	2571.034288	2488.403122	2472.545172	2681.229841	2924.546814	3149.62204
0.028	2604.842302	3108.144398	2475.13739	2586.226629	2566.397344	2584.759558	2477.144285	2540.839396	2583.328483	2494.952833	2480.325969	2690.264399	2911.879728	3120.027041
0.029	2610.421085	3125.860663	2485.000138	2592.502376	2565.856103	2594.501577	2486.918507	2546.160674	2596.036675	2502.71247	2488.959519	2698.839935	2900.273488	3091.752955
0.03	2616.527767	3142.501851	2495.519239	2599.087715	2567.102383	2607.105292	2497.297316	2552.459051	2608.982005	2511.499204	2498.358123	2707.025889	2889.726303	3064.879496
0.031	2623.108854	3158.14661	2506.583986	2605.990665	2569.928129	2621.927096	2508.172668	2559.646055	2622.007786	2521.122774	2508.405973	2714.873823	2880.228886	3039.49449
0.032	2630.116606	3172.842451	2518.103861	2613.196061	2574.13793	2638.373206	2519.452505	2567.639461	2634.970284	2531.442626	2518.99604	2722.402815	2871.74613	3015.680586
0.033	2637.618547	3186.708315	2530.205696	2620.874755	2579.74924	2656.12858	2531.262414	2576.525763	2647.953948	2542.53202	2530.244371	2729.808574	2864.348039	2993.558902
0.034	2645.65794	3199.83098	2542.889456	2629.086297	2586.673356	2674.853037	2543.608025	2586.266574	2660.964963	2554.339932	2542.15805	2737.204739	2858.081656	2973.248097
0.035	2654.229022	3212.259158	2556.072511	2637.812048	2594.751258	2694.208589	2556.409298	2596.772739	2673.952873	2566.757201	2554.65101	2744.620188	2852.927659	2954.792573
0.036	2663.159835	3224.062936	2569.582369	2646.946665	2603.754263	2713.87974	2569.500359	2607.821604	2686.831338	2579.582033	2567.549751	2752.062927	2848.829723	2938.196073
0.037	2672.450184	3235.256561	2583.343745	2656.441074	2613.531469	2733.624583	2582.808998	2619.334798	2699.547192	2592.726862	2580.773563	2759.515849	2845.72139	2923.434756
0.038	2682.048434	3245.881745	2597.306494	2666.31366	2624.002745	2753.294879	2596.297119	2631.249763	2712.136015	2606.130168	2594.274202	2767.070839	2843.660833	2910.62241
0.039	2691.837	3256.000479	2611.322183	2676.449656	2634.963444	2772.742133	2609.830589	2643.384261	2724.545045	2619.628155	2607.903754	2774.707792	2842.558156	2899.668168
0.04	2701.720042	3265.660439	2625.299491	2686.772668	2646.271941	2791.87075	2623.321337	2655.626755	2736.750212	2633.125462	2621.561347	2782.415443	2842.331367	2890.479473
0.041	2711.552981	3274.888785	2639.068725	2697.142143	2657.719898	2810.562639	2636.605009	2667.815893	2748.665008	2646.45376	2635.070087	2790.10985	2842.834526	2882.889168
0.042	2721.285316	3283.697065	2652.625502	2707.575024	2669.282405	2828.765328	2649.675577	2679.949488	2760.324969	2659.614054	2648.416116	2797.831629	2844.067698	2876.907531
0.043	2731.007224	3292.132535	2666.10326	2718.167198	2681.045958	2846.521106	2662.660752	2692.13868	2771.833574	2672.737596	2661.724091	2805.66967	2846.054978	2872.537371
0.044	2740.7476	3300.231297	2679.480285	2728.869905	2692.935831	2863.832056	2675.548634	2704.340375	2783.182973	2685.792711	2674.971279	2813.598814	2848.705006	2869.621642
0.045	2750.516894	3307.984308	2692.796127	2739.734466	2704.981424	2880.659431	2688.389125	2716.580975	2794.439758	2698.818293	2688.197958	2821.700253	2852.08399	2868.221583
0.046	2760.383567	3315.393758	2706.117162	2750.790813	2717.212893	2896.990569	2701.238916	2728.921603	2805.63482	2711.881079	2701.460408	2829.993287	2856.167528	2868.265382
0.047	2770.411022	3322.498722	2719.570292	2762.175131	2729.753521	2912.892903	2714.235057	2741.482434	2816.92093	2725.109571	2714.887724	2838.629503	2861.072955	2869.820771
0.048	2780.700717	3329.326466	2733.186428	2773.833901	2742.570047	2928.397332	2727.375715	2754.283117	2828.236092	2738.528037	2728.493082	2847.497413	2866.60798	2872.591732
0.049	2791.087892	3335.857613	2746.729868	2785.530454	2755.404332	2943.385335	2740.427046	2767.079094	2839.406133	2751.894608	2742.037773	2856.423536	2872.56889	2876.308184
0.05	2801.494923	3342.11074	2760.120629	2797.187348	2768.167241	2957.844705	2753.31747	2779.780015	2850.393394	2765.125668	2755.443067	2865.361751	2878.875059	2880.829461

Table S₃. Relative AICc values of each tested electron isodensity.

ΔAICc (for each isodensity)															
															R
							$a \cdot x^3 + b \cdot x^2 +$	a·x³ + b·x² +	a·x³ + b·x² +						$\sum e^{-\frac{\Delta AIC_r}{2}}$
ISODENSITY	a·x + b	a∙x	a⋅x² + b⋅x + c	a∙x² + b∙x	a·x² + c	a∙x²	c·x + d	c∙x	d	a·x³ + b·x²	$a \cdot x^3 + c \cdot x + d$	a·x³ + c·x	a ⋅ x³ + d	a∙x³	r=1
0.001	90.8915501	1025.97174	2.20387656	891.953287	899.296792	1251.90248	0	803.256427	1215.60857	647.712605	73.1759384	1027.57155	646.378394	1292.82151	1.332226512
0.002	108.764105	889.736206	1.80023443	766.815162	878.590211	1222.66921	0	639.529047	1209.46337	687.897818	74.2429309	890.959751	685.901584	1308.08161	1.406522007
0.003	120.884916	752.234105	1.75799615	641.822765	843.851758	1180.59216	0	496.87751	1180.95276	714.757728	71.6301934	753.91917	713.695925	1313.86037	1.415198702
0.004	130.261176	616.505209	1.83240234	515.917129	802.500792	1130.9683	0	371.300849	1130.7817	727.697856	68.1125073	618.517902	732.618165	1313.38626	1.400035827
0.005	136.852454	490.293315	2.1482804	396.443706	759.896702	1079.63406	0	266.0424	1062.87647	727.618174	64.6128646	491.150306	747.604405	1310.7088	1.34159133
0.006	143.296031	375.532067	2.16497939	282.264706	715.667336	1024.23297	0	177.200656	976.479714	709.063876	60.798677	371.668572	755.784629	1303.35721	1.338751088
0.007	147.130413	279.737674	2.49554054	182.64306	670.535751	965.945456	0	106.534649	876.304718	672.56802	57.6806609	266.045444	759.369361	1291.49329	1.287144338
0.008	151.145158	210.670711	2.56417825	103.894635	628.941908	910.116283	0	55./36/366	772.103408	623.513767	54.3334024	1/9.920928	762.508932	1280.7102	1.277457052
0.009	154.822542	169.460684	2.46736495	47.0219183	586.700463	850.823894	0	21.5006069	664.783753	561.806221	50.9044109	113.75816	760.36477	1264.86594	1.291239636
0.01	150.930317	155.1/5/45	2.38778102	13./545101	543.970727	721 707022	2 02976029	2.99704622	229.024280	492.023392	47.5819299	09.543845	753.422055	1243.80344	1.52/53089/
0.011	161.123446	100.500044	4.27304179	2.9585227	507.241193	731.787832	2.02876928	2.04560920	467.177297	420.790977	40.4907531	48.3939917	749.013061	1220.0530	1.708496292
0.012	161 220726	225 26107	2.03501428	2.36240200	400.155505	612 01799	0	2.94360829	309 56276	200 080262	41.3732009	15 512902	730.700327	1179 29005	1.002111011
0.013	161 676219	225.20107	1.51332361	29 7102772	432.483374	554 270666	0	20 2720742	247 078507	246 082661	25 25/0602	43.313833	715 5215	1150 06005	1.403303382
0.014	161 706267	203.737050	1.32243373	45 4206220	262 679460	405 00921	0	AC 005050	100 /19702	240.383001	22 2452162	77 0240659	600 72061	1120 42221	1 56912502
0.015	160 940097	3/18 828793	0.74550912	61 5908422	331 868127	493.90821	0	62 2552186	162 150868	163 209716	29 2171966	96 6934771	682 962756	1088 60104	1 688834731
0.010	159 348671	387 142176	0.37663411	75 9667985	301 867474	387 52569	0	74 9302069	134 496035	131 720197	26 3040081	116 321327	664 927943	1055 32103	1 828353973
0.017	157 709059	473 339866	0.57005411	88 2467052	274 230768	338 86667	0.00153775	84 3821057	115 314586	106 508259	23 5227671	135 022396	646 706641	1021 68015	1 999239218
0.019	156 165974	457 034735	0	98 5695972	248 862265	294 648398	0 37994871	90 7590508	103 099766	86 742694	21 2109674	152 389148	628 143204	987 463165	1.827005123
0.02	154 301177	487 595537	0	106 58439	225 192243	254 746549	0 72740948	93 8884602	95 8647173	70 955974	19 0372121	167 724054	608 80816	952 13664	1 695169872
0.021	152 081578	514 976956	0	112 439688	203 194741	219 614097	1 03993424	94 2767574	92 3899814	58 3965182	16 988634	180 842567	588 753816	915 782646	1 594744724
0.022	149.662912	539,491182	0	116.354646	182,808579	189.506501	1.32409789	92,600148	91,7017034	48.4819272	15.0364415	191,701839	568.08529	878.536465	1.516336516
0.023	147.177293	561.441818	0	118.436821	163.724442	164.274226	1.58855754	89.4451558	92.9119624	40.6872531	13.1013741	200.161249	546.487249	840.016691	1.453336181
0.024	144.330235	580.666077	0	119.046327	146.260107	144.312468	1.79180189	85.2625971	95.2997958	34,4708311	11.2856817	206.613821	524.644248	801.101721	1.411782454
0.025	141.148562	597.288773	0	118.395344	130.340643	129.336969	1.93514052	80.5017931	98.353116	29.5227991	9.58614697	211.166035	502.630141	761.938275	1.388292568
0.026	137.615467	611.402315	0	116.704182	115.912746	118.926018	2.01809141	75.4975681	101.677667	25.5684895	7.99907139	213.960494	480.5375	722.697985	1.382893668
0.027	133.79604	623.247616	0	114.20654	102.912487	112.553656	2.04116434	70.5150687	105.022725	22.3915596	6.53360998	215.218278	458.535252	683.610477	1.398526858
0.028	129.704912	633.007008	0	111.089239	91.2599542	109.622168	2.00689526	65.7020063	108.191093	19.8154433	5.18857918	215.127009	436.742338	644.889651	1.441362026
0.029	125.420948	640.860525	0	107.502238	80.8559652	109.50144	1.91836984	61.1605364	111.036538	17.7123326	3.95938109	213.839797	415.273351	606.752817	1.521459572
0.03	121.008529	646.982612	0	103.568476	71.583144	111.586053	1.77807776	56.9398127	113.462767	15.9799651	2.83888451	211.506651	394.207064	569.360257	1.653238342
0.031	116.524868	651.562623	0	99.4066781	63.3441428	115.34311	1.58868144	53.062069	115.423799	14.5387878	1.82198698	208.289836	373.6449	532.910503	1.854700104
0.032	112.012745	654.73859	0	95.0922	56.0340693	120.269345	1.34864381	49.5356	116.866423	13.3387645	0.89217852	204.298954	353.642269	497.576725	2.150897609
0.033	107.412851	656.502619	0	90.6690598	49.5435443	125.922884	1.05671809	46.3200677	117.748252	12.3263244	0.03867581	199.602878	334.142344	463.353206	2.572525095
0.034	103.499889	657.67293	0.73140548	86.9282467	44.5153052	132.694986	1.44997468	44.108523	118.806913	12.181882	0	195.046689	315.923606	431.090047	2.180302962
0.035	99.5780121	657.608148	1.42150102	83.1610383	40.1002477	139.557579	1.75828841	42.1217287	119.301863	12.1061909	0	189.969178	298.276649	400.141563	1.908763962
0.036	95.6100843	656.513185	2.03261822	79.3969137	36.2045126	146.329989	1.95060831	40.2718529	119.281587	12.0322818	0	184.513176	281.279973	370.646323	1.74144504
0.037	91.6766208	654.482998	2.57018162	75.6675109	32.7579063	152.851021	2.03543465	38.5612349	118.773629	11.9532988	0	178.742286	264.947827	342.661193	1.640581859
0.038	87.7742317	651.607543	3.03229183	72.0394581	29.7285425	159.020676	2.02291642	36.9755609	117.861813	11.8559656	0	172.796637	249.386631	316.348208	1.585908917
0.039	83.9332458	648.096726	3.41842869	68.5459017	27.0596904	164.838379	1.92683513	35.4805072	116.641291	11.7244007	0	166.804038	234.654402	291.764414	1.565440832
0.04	80.158695	644.099092	3.73814389	65.211321	24.7105942	170.309402	1.75998932	34.0654073	115.188864	11.5641146	0	160.854096	220.77002	268.918125	1.572138605
0.041	76.4828946	639.818698	3.99863838	62.0720566	22.649811	175.492552	1.53492172	32.7458065	113.594921	11.3836729	0	155.039763	207.76444	247.819081	1.603003208
0.042	72.8692005	635.280949	4.20938624	59.1589087	20.8662891	180.349212	1.25946149	31.5333727	111.908854	11.197938	0	149.415513	195.651582	228.491415	1.658349556
0.043	69.283132	630.408443	4.37916819	56.4431061	19.3218662	184.797014	0.93666055	30.4145881	110.109482	11.0135046	0	143.945579	184.330886	210.813279	1.74213327
0.044	65.776321	625.260018	4.50900551	53.8986263	17.9645516	188.860777	0.57735446	29.3690954	108.211694	10.8214321	0	138.627535	173.733726	194.650363	1.858774183
0.045	62.3189361	619.786351	4.59816933	51.5365085	16.783466	192.461473	0.19116757	28.3830173	106.241801	10.6203352	0	133.502295	163.886032	180.023625	2.014361387
0.046	59.1446508	614.154841	4.87824509	49.5518964	15.9739769	195.751653	0	27.6826862	104.395903	10.6421628	0.22149193	128.75437	154.928611	167.026465	1.987631785
0.047	56.1759648	608.263665	5.33523443	47.9400739	15.5184634	198.657846	0	27.2473772	102.685873	10.8745142	0.65266678	124.394445	146.837898	155.585714	1.795761419
0.048	53.3250022	601.950751	5.81071303	46.458186	15.1943314	201.021617	0	26.9074016	100.860376	11.1523216	1.11736687	120.121698	139.232265	145.216017	1.630981252
0.049	50.6608461	595.430566	6.30282171	45.1034078	14.9772861	202.958289	0	26.6520475	98.9790868	11.4675621	1.61072684	115.996489	132.141844	135.881137	1.493513044
0.05	48.1774523	588.79327	6.80315886	43.8698772	14.8497702	204.527234	0	26.4625448	97.0759235	11.8081981	2.12559626	112.044281	125.557589	127.51199	1.382134601

Akaike weights (from ΔAICc for each isodensity)														
							$a \cdot x^3 + b \cdot x^2 +$	a·x³ + b·x² +	$a \cdot x^3 + b \cdot x^2 +$					
ISODENSITY	a∙x+b	a∙x	$a \cdot x^2 + b \cdot x + c$	a·x² + b·x	a·x² + c	a∙x²	c·x+d	C·X	d	a·x³ + b·x²	$a \cdot x^3 + c \cdot x + d$	a·x³ + c·x	a x³ + d	a∙x³
0.001	0.0%	0.0%	24.9%	0.0%	0.0%	0.0%	75.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.002	0.0%	0.0%	28.9%	0.0%	0.0%	0.0%	71.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.003	0.0%	0.0%	29.3%	0.0%	0.0%	0.0%	70.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.004	0.0%	0.0%	28.6%	0.0%	0.0%	0.0%	71.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.005	0.0%	0.0%	25.5%	0.0%	0.0%	0.0%	74.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.006	0.0%	0.0%	25.3%	0.0%	0.0%	0.0%	74.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.007	0.0%	0.0%	22.3%	0.0%	0.0%	0.0%	77.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.008	0.0%	0.0%	21.7%	0.0%	0.0%	0.0%	/8.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.009	0.0%	0.0%	22.6%	0.0%	0.0%	0.0%	/7.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.01	0.0%	0.0%	19.8%	0.1%	0.0%	0.0%	65.5%	14.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.011	0.0%	0.0%	6.9%	13.3%	0.0%	0.0%	21.2%	58.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.012	0.0%	0.0%	19.2%	14.8%	0.0%	0.0%	53.7%	12.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.013	0.0%	0.0%	28.7%	0.1%	0.0%	0.0%	/1.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.014	0.0%	0.0%	31.8%	0.0%	0.0%	0.0%	68.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.015	0.0%	0.0%	36.2%	0.0%	0.0%	0.0%	63.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.016	0.0%	0.0%	40.8%	0.0%	0.0%	0.0%	59.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.017	0.0%	0.0%	45.3%	0.0%	0.0%	0.0%	54.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.018	0.0%	0.0%	50.0%	0.0%	0.0%	0.0%	50.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.019	0.0%	0.0%	54.7%	0.0%	0.0%	0.0%	45.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.02	0.0%	0.0%	59.0%	0.0%	0.0%	0.0%	41.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.021	0.0%	0.0%	62.7%	0.0%	0.0%	0.0%	37.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.022	0.0%	0.0%	65.9%	0.0%	0.0%	0.0%	34.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
0.023	0.0%	0.0%	68.8%	0.0%	0.0%	0.0%	31.1%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%
0.024	0.0%	0.0%	70.8%	0.0%	0.0%	0.0%	28.9%	0.0%	0.0%	0.0%	0.3%	0.0%	0.0%	0.0%
0.025	0.0%	0.0%	72.0%	0.0%	0.0%	0.0%	27.4%	0.0%	0.0%	0.0%	0.6%	0.0%	0.0%	0.0%
0.026	0.0%	0.0%	72.3%	0.0%	0.0%	0.0%	26.4%	0.0%	0.0%	0.0%	1.3%	0.0%	0.0%	0.0%
0.027	0.0%	0.0%	71.5%	0.0%	0.0%	0.0%	25.8%	0.0%	0.0%	0.0%	2.7%	0.0%	0.0%	0.0%
0.028	0.0%	0.0%	69.4%	0.0%	0.0%	0.0%	25.4%	0.0%	0.0%	0.0%	5.2%	0.0%	0.0%	0.0%
0.029	0.0%	0.0%	65.7%	0.0%	0.0%	0.0%	25.2%	0.0%	0.0%	0.0%	9.1%	0.0%	0.0%	0.0%
0.03	0.0%	0.0%	60.5%	0.0%	0.0%	0.0%	24.9%	0.0%	0.0%	0.0%	14.6%	0.0%	0.0%	0.0%
0.031	0.0%	0.0%	53.9%	0.0%	0.0%	0.0%	24.4%	0.0%	0.0%	0.0%	21.7%	0.0%	0.0%	0.0%
0.032	0.0%	0.0%	46.5%	0.0%	0.0%	0.0%	23.7%	0.0%	0.0%	0.1%	29.8%	0.0%	0.0%	0.0%
0.033	0.0%	0.0%	38.9%	0.0%	0.0%	0.0%	22.9%	0.0%	0.0%	0.1%	38.1%	0.0%	0.0%	0.0%
0.034	0.0%	0.0%	31.8%	0.0%	0.0%	0.0%	22.2%	0.0%	0.0%	0.1%	45.9%	0.0%	0.0%	0.0%
0.035	0.0%	0.0%	25.7%	0.0%	0.0%	0.0%	21.7%	0.0%	0.0%	0.1%	52.4%	0.0%	0.0%	0.0%
0.030	0.0%	0.0%	20.8%	0.0%	0.0%	0.0%	21.7%	0.0%	0.0%	0.1%	57.4%	0.0%	0.0%	0.0%
0.037	0.0%	0.0%	10.9%	0.0%	0.0%	0.0%	22.0%	0.0%	0.0%	0.2%	61.0%	0.0%	0.0%	0.0%
0.038	0.0%	0.0%	13.8%	0.0%	0.0%	0.0%	22.9%	0.0%	0.0%	0.2%	63.1%	0.0%	0.0%	0.0%
0.039	0.0%	0.0%	11.6%	0.0%	0.0%	0.0%	24.4%	0.0%	0.0%	0.2%	63.9%	0.0%	0.0%	0.0%
0.04	0.0%	0.0%	9.8%	0.0%	0.0%	0.0%	20.4%	0.0%	0.0%	0.2%	03.0%	0.0%	0.0%	0.0%
0.041	0.0%	0.0%	8.4%	0.0%	0.0%	0.0%	29.0%	0.0%	0.0%	0.2%	62.4%	0.0%	0.0%	0.0%
0.042	0.0%	0.0%	7.3% 6.4%	0.0%	0.0%	0.0%	32.1%	0.0%	0.0%	0.2%	60.3% E7.4%	0.0%	0.0%	0.0%
0.043	0.0%	0.0%	6.4% E.C%	0.0%	0.0%	0.0%	35.9%	0.0%	0.0%	0.2%	57.4%	0.0%	0.0%	0.0%
0.044	0.0%	0.0%	5.0%	0.0%	0.0%	0.0%	40.3%	0.0%	0.0%	0.2%	53.8%	0.0%	0.0%	0.0%
0.045	0.0%	0.0%	5.0%	0.0%	0.0%	0.0%	45.1%	0.0%	0.0%	0.2%	49.6%	0.0%	0.0%	0.0%
0.046	0.0%	0.0%	4.4%	0.0%	0.0%	0.0%	50.3%	0.0%	0.0%	0.2%	45.0%	0.0%	0.0%	0.0%
0.047	0.0%	0.0%	3.9%	0.0%	0.0%	0.0%	55.7%	0.0%	0.0%	0.2%	40.2%	0.0%	0.0%	0.0%
0.048	0.0%	0.0%	3.4%	0.0%	0.0%	0.0%	67.0%	0.0%	0.0%	0.2%	35.1%	0.0%	0.0%	0.0%
0.049	0.0%	0.0%	2.9%	0.0%	0.0%	0.0%	72.4%	0.0%	0.0%	0.2%	29.9%	0.0%	0.0%	0.0%
0.05	0.070	0.070	2.470	0.070	0.078	0.070	12.470	0.070	0.070	0.270	23.070	0.070	0.076	0.078

Table S4. Akaike weights for selecting model for each electron isodensity level.

XI. References

(1) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic Structure Calculations on Workstation Computers: The Program System Turbomole. *Chemical Physics Letters* **1989**, *1*62 (3), 165–169.

(2) TURBOMOLE V7.1 2016, a Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; 2016.

(3) TURBOMOLE V6.4 2012, a Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; 2012.

(4) TURBOMOLE V7.2 2017, a Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; 2017.

(5) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: Optimized Auxiliary Basis Sets and Demonstration of Efficiency. *Chemical Physics Letters* **1998**, *294* (1), 143–152.

(6) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305.

(7) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. Phys. Chem. Chem. Phys. 2006, 8 (9), 1057–1065.

(8) Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. Optimized Accurate Auxiliary Basis Sets for RI-MP2 and RI-CC2 Calculations for the Atoms Rb to Rn. *Theor Chem Acc* 2007, *11*7 (4), 587–597.

(9) Von Arnim Malte; Ahlrichs Reinhart. Performance of Parallel TURBOMOLE for Density Functional Calculations. *Journal of Computational Chemistry* **1999**, *19* (15), 1746–1757.

(10) Treutler, O.; Ahlrichs, R. Efficient Molecular Numerical Integration Schemes. *The Journal of Chemical Physics* **1995**, *102* (1), 346–354.

(11) Friese, D. H.; Winter, N. O. C.; Balzerowski, P.; Schwan, R.; Hättig, C. Large Scale Polarizability Calculations Using the Approximate Coupled Cluster Model CC2 and MP2 Combined with the Resolution-of-the-Identity Approximation. *The Journal of Chemical Physics* **2012**, *136* (17), 174106.

(12) Hättig, C.; Hellweg, A.; Köhn, A. Distributed Memory Parallel Implementation of Energies and Gradients for Second-Order Møller–Plesset Perturbation Theory with the Resolution-of-the-Identity Approximation. *Phys. Chem. Chem. Phys.* **2006**, 8 (10), 1159–1169.

(13) Hättig, C.; Weigend, F. CC₂ Excitation Energy Calculations on Large Molecules Using the Resolution of the Identity Approximation. *The Journal of Chemical Physics* **2000**, *1*(13), 5154–5161.

(14) Brickmann Jürgen; Exner Thomas. *Molcad II*; 2012.

(15) Zhou, C.; Nie, C.; Li, S.; Li, Z. A Novel Semi-Empirical Topological Descriptor Nt and the Application to Study on QSPR/QSAR. *J. Comput. Chem.* **2007**, *28* (15), 2413–2423.

(16) *Molecular Descriptors for Chemoinformatics: Volume I: Alphabetical Listing / Volume II: Appendices, References;* Todeschini, R., Consonni, V., Eds.; Methods and Principles in Medicinal Chemistry; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2009.

(17) Cheng, Y.-Y.; Yuan, H. Quantitative Study of Electrostatic and Steric Effects on Physicochemical Property and Biological Activity. *Journal of Molecular Graphics and Modelling* **2006**, *24* (4), 219–226.

(18) Kier, L. B.; Hall, L. H. An Electrotopological-State Index for Atoms in Molecules. *Pharm. Res.* 1990, 7 (8), 801–807.

(19) Kier, L. B.; Hall, L. H. Molecular Structure Description : The Electrotopological State; San Diego : Academic Press, 1999.

(20) Ivanciuc, O. Electrotopological State Indices. In Molecular Drug Properties; Wiley-Blackwell, 2008; pp 85-109.

(21) Shevade, S. K.; Keerthi, S. S.; Bhattacharyya, C.; Murthy, K. R. K. Improvements to the SMO Algorithm for SVM Regression. *IEEE Transactions on Neural Networks* 2000, *11* (5), 1188–1193.

(22) Smola, A. J.; Schölkopf, B. A Tutorial on Support Vector Regression. Statistics and Computing 2004, 14 (3), 199–222.

(23) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* 2007, 9 (3), 90–95.

XII. Appendix: Electron Isodensity-dependence of Adduct Formation Energy

0.001 au



0.002 au



0.003 au



0.004 au



0.005 au



0.006 au



0.007 au



0.008 au



0.009 au



0.01 au



0.011 au



0.012 au



0.013 au



0.014 au



0.015 au



0.016 au



0.017 au



• Br

• |

0.018 au



0.019 au



0.02 au



0.021 au



0.022 au



• Cl

• Br

• |

0.023 au



0.024 au



0.025 au



0.026 au



0.027 au



44

0.028 au



0.029 au



0.03 au



0.031 au



0.032 au



0.033 au



0.034 au



0.035 au



0.036 au



0.037 au



0.038 au



0.039 au



0.04 au



0.041 au



0.042 au



• Cl

• Br

• |

0.043 au



0.044 au



0.045 au



0.046 au



0.047 au



• Cl

• Br

• |

0.048 au



0.049 au



0.05 au

