Supporting information for

Efficient Diffuse Basis Sets for Density Functional Theory

Ewa Papajak and Donald G. Truhlar
Department of Chemistry and Supercomputing Institute, University of Minnesota, 207 Pleasant Street S.E., Minneapolis, MN 55455-0431, USA

Date of finalization: Jan. 23, 2009

Number of pages, including this page: 8

For statistical completeness, in addition to the MSE and MUE values listed in the main part of the paper, we report here the root mean square error (RMSE). These values were calculated using

$$\text{RMSE} = \text{RMS error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \quad \text{(1-A)}$$

where e_i is the error in the atomization energy, and n is the number of atoms or molecules in the database, except for Table S-6, for which we used

$$\text{RMSE} = \text{RMS error per bond} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \Big/ M \quad \text{(1-A)}$$

where M is the average number of bonds per molecule in the test set (M is 4.83).

We also present here the results of the tests of the aug-cc-pVxZ, maug-cc-pVxZ, and cc-pVxZ basis sets for their accuracies in predicting barrier heights, electron affinities, ionization potentials, and atomization energies. The conclusions about the importance of diffuse basis functions and the cost savings for these xZ basis sets are qualitatively the same as for the (x+d)Z basis sets.

Table S-7 presents additional results for metal hydrides.

Table S-1. Expanded version of table 1, showing the sums of the number of basis functions in each of the databases, as well as their total, which is *N*

Basis set	Heavy atoms	H-He atoms	BH24	HB6	EA13/3	IP13/3	AE6	Total *N*	$(N^4)_{Nor}$	$(N^4)_{Rel}$
aug-cc-pV(Q+d)Z	spdfg	spdf	4565	4068	3218	3218	2528	17597	934	1.00
maug-cc-pV(Q+d)Z	sp	-	3238	2851	2386	2386	1804	12665	251	0.27
cc-pV(Q+d)Z	-	-	3098	2735	2258	2258	1724	12073	207	0.22
aug-cc-pV(T+d)Z	spdf	spd	2616	2208	1900	1900	1411	10035	99	1.00
maug-cc-pV(T+d)Z	sp	-	1839	1518	1426	1426	1000	7209	26	0.27
cc-pV(T+d)Z	-	-	1691	1402	1298	1298	920	6609	19	0.19
aug-cc-pV(D+d)Z	spd	sp	1247	1009	1024	1024	685	4989	6	1.00
maug-cc-pV(D+d)Z	sp	-	914	712	824	824	509	3783	2	0.33
cc-pV(D+d)Z	-	-	766	596	696	696	429	3183	1	0.17

Table S-2. Errors in predictions for the barrier height (kcal/mol per bond) in the BH24/08 database.

	B3LYP			M06-2X		
	MSE	MUE	RMSE	MSE	MUE	RMSE
aug-cc-pV(Q+d)Z	-3.99	4.06	4.86	0.03	0.93	1.14
maug-cc-pV(Q+d)Z	-3.97	4.06	4.84	0.09	0.92	1.14
cc-pV(Q+d)Z	-4.79	4.85	5.80	-0.50	1.38	2.09
aug-cc-pV(T+d)Z	-4.08	4.14	4.91	-0.06	0.88	1.16
maug-cc-pV(T+d)Z	-4.02	4.10	4.86	0.08	0.91	1.19
cc-pV(T+d)Z	-5.26	5.42	6.74	-0.87	2.06	3.34
aug-cc-pV(D+d)Z	-4.83	4.90	5.66	-0.60	1.20	1.56
maug-cc-pV(D+d)Z	-4.69	4.72	5.66	-0.39	1.33	1.95
cc-pV(D+d)Z	-7.01	7.54	9.67	-2.32	3.75	6.04
aug-cc-pVQZ	-4.07	4.13	4.90	0.03	0.91	1.16
maug-cc-pVQZ	-3.96	4.05	4.83	0.09	0.92	1.16
cc-pVQZ	-4.78	4.84	5.80	-0.50	1.43	2.12
aug-cc-pVTZ	-4.07	4.13	4.90	-0.05	0.92	1.23
maug-cc-pVTZ	-4.00	4.09	4.84	0.10	0.96	1.26
cc-pVTZ	-5.25	5.42	6.75	-0.86	2.17	3.40
aug-cc-pVDZ	-4.81	4.89	5.66	-0.59	1.31	1.70
maug-cc-pVDZ	-4.67	4.70	5.67	-0.38	1.44	2.10
cc-pVDZ	-7.00	7.55	9.73	-2.32	3.90	6.15

Table S-3. Errors in predictions for the hydrogen bonding energy (kcal/mol) in the HB6 database.

	B3LYP			M06-2X		
	MSE	MUE	RMSE	MSE	MUE	RMSE
aug-cc-pV(Q+d)Z	-0.76	0.76	0.94	0.11	0.29	0.13
maug-cc-pV(Q+d)Z	-0.73	0.73	0.91	0.10	0.28	0.11
cc-pV(Q+d)Z	-0.23	0.55	0.52	0.42	0.46	0.29
aug-cc-pV(T+d)Z	-0.73	0.73	0.86	0.17	0.31	0.14
maug-cc-pV(T+d)Z	-0.67	0.67	0.90	0.17	0.34	0.14
cc-pV(T+d)Z	0.51	0.74	0.66	1.02	1.02	1.18
aug-cc-pV(D+d)Z	-0.39	0.40	0.31	0.37	0.37	0.18
maug-cc-pV(D+d)Z	-0.27	0.64	0.74	0.43	0.61	0.43
cc-pV(D+d)Z	2.97	2.97	9.54	2.91	2.91	9.20

Table S-4. Errors in predictions for the electron affinity (kcal/mol) in the EA13/3 database.

	B3LYP			M06-2X		
	MSE	MUE	RMSE	MSE	MUE	RMSE
aug-cc-pV(Q+d)Z	-1.99	2.33	3.29	0.78	1.47	1.88
maug-cc-pV(Q+d)Z	-1.96	2.31	3.27	0.82	1.49	1.89
cc-pV(Q+d)Z	4.35	5.30	6.51	5.30	5.30	6.68
aug-cc-pV(T+d)Z	-2.07	2.37	3.37	0.67	1.46	1.92
maug-cc-pV(T+d)Z	-2.04	2.36	3.36	0.70	1.49	1.94
cc-pV(T+d)Z	9.49	10.13	11.74	9.99	9.99	11.70
aug-cc-pV(D+d)Z	-2.45	2.75	4.10	0.87	2.21	2.89
maug-cc-pV(D+d)Z	-2.53	2.92	4.40	0.71	2.28	3.05
cc-pV(D+d)Z	20.59	20.85	24.19	20.37	20.37	23.62
aug-cc-pVQZ	-2.11	2.39	3.45	0.68	1.48	1.91
maug-cc-pVQZ	-2.07	2.36	3.43	0.72	1.50	1.91
cc-pVQZ	4.24	5.31	6.55	5.20	5.27	6.66
aug-cc-pVTZ	-2.26	2.52	3.64	0.50	1.57	2.03
maug-cc-pVTZ	-2.22	2.50	3.63	0.53	1.59	2.05
cc-pVTZ	9.30	10.14	11.74	9.82	9.83	11.66
aug-cc-pVDZ	-2.79	3.04	4.67	0.55	2.40	3.26
maug-cc-pVDZ	-2.86	3.20	4.95	0.40	2.49	3.44
cc-pVDZ	20.26	20.85	24.12	20.06	20.06	23.53

Table S-5. Errors in predictions for the ionization potential (kcal/mol) in the IP13/3 database.

	B3LYP			M06-2X		
	MSE	MUE	RMSE	MSE	MUE	RMSE
aug-cc-pV(Q+d)Z	3.42	4.62	5.72	0.66	2.26	3.02
maug-cc-pV(Q+d)Z	3.41	4.61	5.71	0.64	2.25	3.00
cc-pV(Q+d)Z	3.19	4.42	5.44	0.51	2.14	2.88
aug-cc-pV(T+d)Z	3.54	4.65	5.79	0.99	2.63	3.53
maug-cc-pV(T+d)Z	3.53	4.63	5.76	0.96	2.62	3.54
cc-pV(T+d)Z	2.94	4.14	5.09	0.57	2.31	3.20
aug-cc-pV(D+d)Z	4.01	4.72	6.03	1.08	2.81	3.58
maug-cc-pV(D+d)Z	4.14	4.67	6.14	1.26	2.99	3.90
cc-pV(D+d)Z	1.67	3.15	3.93	-0.67	2.84	3.43
aug-cc-pVQZ	3.51	4.62	5.72	0.74	2.26	3.05
maug-cc-pVQZ	3.50	4.61	5.72	0.72	2.25	3.04
cc-pVQZ	3.28	4.42	5.44	0.59	2.13	2.91
aug-cc-pVTZ	3.68	4.67	5.80	1.12	2.71	3.60
maug-cc-pVTZ	3.67	4.64	5.78	1.10	2.71	3.61
cc-pVTZ	3.09	4.15	5.10	0.71	2.40	3.27
aug-cc-pVDZ	4.28	4.77	6.13	1.35	2.98	3.78
maug-cc-pVDZ	4.39	4.89	6.27	1.51	3.15	4.13
cc-pVDZ	1.93	3.14	4.02	-0.41	2.99	3.62

Table S-6. Errors in predictions for the atomization energy (kcal/mol) in the AE6 database.

	B3LYP			M06-2X		
	MSE	MUE	RMSE	MSE	MUE	RMSE
aug-cc-pV(Q+d)Z	-0.42	0.59	0.71	0.01	0.22	0.29
maug-cc-pV(Q+d)Z	-0.43	0.61	0.72	0.01	0.22	0.29
cc-pV(Q+d)Z	-0.38	0.56	0.67	0.03	0.22	0.28
aug-cc-pV(T+d)Z	-0.68	0.76	0.95	-0.19	0.33	0.35
maug-cc-pV(T+d)Z	-0.70	0.79	0.97	-0.23	0.36	0.39
cc-pV(T+d)Z	-0.59	0.69	0.85	-0.16	0.29	0.32
aug-cc-pV(D+d)Z	-2.31	2.31	2.71	-1.65	1.73	2.00
maug-cc-pV(D+d)Z	-2.67	2.67	3.13	-1.94	1.94	2.29
cc-pV(D+d)Z	-2.29	2.29	2.70	-1.59	1.65	1.90
aug-cc-pVQZ	-0.54	0.60	0.75	-0.10	0.24	0.29
maug-cc-pVQZ	-0.56	0.62	0.77	-0.11	0.24	0.29
cc-pVQZ	-0.51	0.57	0.71	-0.08	0.24	0.29
aug-cc-pVTZ	-0.92	0.92	1.08	-0.40	0.40	0.46
maug-cc-pVTZ	-0.94	0.94	1.11	-0.45	0.45	0.50
cc-pVTZ	-0.83	0.83	0.99	-0.38	0.38	0.44
aug-cc-pVDZ	-2.80	2.80	3.01	-2.12	2.12	2.28
maug-cc-pVDZ	-3.16	3.16	3.43	-2.40	2.40	2.58
cc-pVDZ	-2.78	2.78	3.03	-2.06	2.06	2.22

Table S-7. Electron affinity (kcal/mol) of beryllium and magnesium hydrides.

	BeH		MgH	
basis set	B3LYP	M06-2X	B3LYP	M06-2X
aug-cc-pV(Q+d)Z	12.16	9.57	19.32	16.14
jul-cc-pV(Q+d)Z	12.00	9.37	19.27	16.11
maug-cc-pV(Q+d)Z	11.68	8.90	18.94	15.67
cc-pV(Q+d)Z	8.04	6.00	18.08	15.00
aug-cc-pV(T+d)Z	12.17	9.49	19.37	16.29
jul-cc-pV(T+d)Z	11.84	9.05	19.21	16.10
maug-cc-pV(T+d)Z	11.30	8.58	18.67	15.53
cc-pV(T+d)Z	6.65	4.60	16.96	14.08
aug-cc-pV(D+d)Z	12.24	9.66	19.50	16.42
jul-cc-pV(D+d)Z	11.26	8.75	18.91	15.89
maug-cc-pV(D+d)Z	10.17	7.59	17.97	14.84
cc-pV(D+d)Z	0.53	-1.03	14.18	11.52

Table S-8. Errors in predictions for the electron affinity (kcal/mol) in the EA13/3 database in the wave function theory (WFT) calculations employing second-order perturbation theory

	MP2		
	MSE	MUE	RMS
aug-cc-pV(T+d)Z	1.44	2.26	2.99
maug-cc-pV(T+d)Z	4.01	4.31	4.81
cc-pV(T+d)Z	14.36	14.36	15.56