Supplementary Information for:

Hydrogen Bonding Motifs in a Hydroxy-Bisphosphonate Moiety: Revisiting the Problem of Hydrogen Bond Identification

Mitra Ashouri[†], Ali Maghari[†], Mohammad Hossein Karimi-Jafari*[,‡]

† Department of Physical Chemistry, School of Chemistry, College of Science, University of Tehran, Tehran, Iran.

‡ Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran.

Table S1. H-bond motifs, populations, relative conformational energy terms and equilibrium constants of all conformers in **HL** form of **HEDP**. All energetic values are in kcal/mol.

	motif	Population (%)	E_{tot}	E_{gas}	ΔG_{solv}	E_{ZPE}	G_{therm}	K	K^{elec}	K^{vib}	K^{rot}
C1	**M1**	50.8	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00
C2	**M2**	34.4	0.62	3.94	-2.48	-0.84	-1.23	0.68	0.35	1.94	1.00
C3	**M1**	14.2	0.87	0.81	0.04	0.02	-0.10	0.28	0.23	1.21	1.00
C4	**M3**	0.2	3.70	6.50	-2.54	-0.26	-0.66	0.00	0.00	1.97	0.99
C5	**M4**	0.3	3.94	7.58	-3.11	-0.54	-1.40	0.01	0.00	4.25	1.00
C6	**M4**	0.0	5.07	9.63	-4.07	-0.49	-1.05	0.00	0.00	2.59	1.00
C7	**M4**	0.0	5.52	14.17	-8.58	-0.07	-0.25	0.00	0.00	1.37	1.00
C8	**M5**	0.0	9.84	16.45	-6.57	-0.03	-0.49	0.00	0.00	2.09	1.03
C9	**M5**	0.0	10.16	14.24	-3.74	-0.33	-1.08	0.00	0.00	3.45	1.03
C10	**M6**	0.0	10.85	19.80	-8.98	0.03	-0.29	0.00	0.00	1.64	1.04
C11	**M6**	0.0	11.01	22.60	-11.36	-0.24	-0.90	0.00	0.00	2.93	1.04
C12	**M6**	0.0	11.24	21.61	-10.30	-0.06	-0.42	0.00	0.00	1.76	1.05
C13	**M6**	0.0	11.91	17.13	-4.76	-0.46	-1.33	0.00	0.00	4.21	1.03
C14	**M6**	0.0	12.33	22.19	-9.61	-0.24	-0.71	0.00	0.00	2.13	1.03
C15	**M6**	0.0	12.85	22.13	-9.20	-0.09	-0.55	0.00	0.00	2.09	1.04
C16	**M7**	0.0	14.38	28.25	-13.57	-0.31	-1.10	0.00	0.00	3.65	1.04
C17	**M7**	0.0	15.41	32.69	-17.36	0.09	-0.34	0.00	0.00	1.96	1.04
C18	**M8**	0.0	16.22	36.15	-19.79	-0.14	-0.68	0.00	0.00	2.38	1.04
C19	**M7**	0.0	16.37	31.48	-14.90	-0.21	-0.64	0.00	0.00	1.98	1.04
C20	**M8**	0.0	16.43	36.68	-20.32	0.07	-0.25	0.00	0.00	1.64	1.04
C21	**M7**	0.0	16.46	32.44	-15.53	-0.45	-1.20	0.00	0.00	3.39	1.04
C22	**M8**	0.0	17.78	38.50	-20.26	-0.45	-1.17	0.00	0.00	3.22	1.04

Table S2. H-bond motifs, populations, relative conformational energy terms and equilibrium constants of all conformers in **H_2L** form of **HEDP**. All energetic values are in kcal/mol.

	motif	Population (%)	E_{tot}	E_{gas}	ΔG_{solv}	E_{ZPE}	G_{therm}	K	K^{elec}	K^{vib}	K^{rot}
C1	**M1**	56.8	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00
C2	**M1**	26.4	0.40	2.36	-1.95	-0.01	0.05	0.46	0.51	0.91	1.00
C3	**M1**	16.7	0.71	3.59	-2.74	-0.14	-0.12	0.29	0.30	0.97	1.00
C4	**M2**	0.1	4.51	8.02	-3.18	-0.33	-0.66	0.00	0.00	1.64	1.05
C5	**M2**	0.0	5.66	10.11	-4.26	-0.18	-0.62	0.00	0.00	2.00	1.04
C6	**M2**	0.0	5.81	11.25	-5.00	-0.44	-0.81	0.00	0.00	1.80	1.05
C7	**M3**	0.0	5.89	13.25	-7.26	-0.10	-0.41	0.00	0.00	1.62	1.05
C8	**M4**	0.0	5.97	10.92	-4.79	-0.16	-0.52	0.00	0.00	1.78	1.04
C9	**M3**	0.0	5.99	14.36	-8.31	-0.06	-0.18	0.00	0.00	1.17	1.04
C10	**M3**	0.0	6.09	12.78	-6.65	-0.03	-0.21	0.00	0.00	1.28	1.05
C11	**M2**	0.0	6.20	12.29	-5.75	-0.35	-0.61	0.00	0.00	1.49	1.05
C12	**M5**	0.0	6.64	6.30	0.26	0.09	0.04	0.00	0.00	1.05	1.04
C13	**M6**	0.0	7.05	6.03	0.89	0.13	0.24	0.00	0.00	0.81	1.03
C14	**M4**	0.0	7.58	15.23	-7.48	-0.18	-0.43	0.00	0.00	1.47	1.04
C15	**M7**	0.0	7.58	18.41	-10.54	-0.29	-0.68	0.00	0.00	1.84	1.04
C16	**M4**	0.0	7.70	14.30	-6.28	-0.32	-0.77	0.00	0.00	2.07	1.04
C17	**M7**	0.0	8.11	15.83	-7.22	-0.50	-1.11	0.00	0.00	2.67	1.04
C18	**M7**	0.0	9.23	20.00	-10.31	-0.46	-1.01	0.00	0.00	2.40	1.04
C19	**M8**	0.0	9.64	12.87	-3.21	-0.02	-0.14	0.00	0.00	1.16	1.05
C20	**M7**	0.0	10.09	23.84	-13.31	-0.44	-0.98	0.00	0.00	2.37	1.05
C21	**M9**	0.0	10.77	13.41	-2.52	-0.12	-0.60	0.00	0.00	2.19	1.03
C22	**M10**	0.0	11.00	16.97	-5.81	-0.17	-0.80	0.00	0.00	2.76	1.06
C23	**M11**	0.0	11.16	21.51	-10.02	-0.33	-0.77	0.00	0.00	1.94	1.08
C24	**M8**	0.0	11.31	14.82	-3.18	-0.33	-0.95	0.00	0.00	2.73	1.04
C25	**M11**	0.0	12.18	18.64	-5.76	-0.70	-1.64	0.00	0.00	4.52	1.07
C26	**M11**	0.0	12.52	24.84	-11.60	-0.73	-1.80	0.00	0.00	5.65	1.09
C27	**M12**	0.0	12.87	23.56	-10.17	-0.52	-1.23	0.00	0.00	3.12	1.05
C28	**M11**	0.0	13.16	25.65	-11.93	-0.56	-1.34	0.00	0.00	3.47	1.08
C29	**M13**	0.0	13.16	23.83	-10.36	-0.30	-0.70	0.00	0.00	1.82	1.07
C30	**M7**	0.0	13.18	21.98	-8.58	-0.22	-0.93	0.00	0.00	3.11	1.06
C31	**M11**	0.0	13.68	25.13	-10.76	-0.70	-1.88	0.00	0.00	6.86	1.08
C32	**M14**	0.0	14.03	25.90	-11.19	-0.68	-1.39	0.00	0.00	3.11	1.07
C33	**M12**	0.0	14.28	26.37	-11.66	-0.44	-0.98	0.00	0.00	2.37	1.06
C34	**M13**	0.0	14.67	30.29	-14.77	-0.84	-1.80	0.00	0.00	4.65	1.09
C35	**M13**	0.0	14.68	26.73	-11.42	-0.63	-1.24	0.00	0.00	2.62	1.07
C36	**M13**	0.0	15.38	27.47	-11.30	-0.80	-1.68	0.00	0.00	4.13	1.07
C37	**M13**	0.0	15.93	29.52	-13.00	-0.59	-1.23	0.00	0.00	2.76	1.08

Table S3. Hydrogen bond properties in different conformers of **HL** form of **HEDP**. All values are in atomic units except E_{HB} that is in kcal/mol.

	D	A	E_{HB}	R_{DH}	θ_{DHA}	R_{HA}	R_{DA}	ρ_{bc}	L_{bc}	V_{bc}	ρ_{rc}	L_{rc}	V_{rc}
C1	O11	O22	19.0	1.926	165.5	2.970	4.859	0.0593	0.1528	0.0605	0.0158	0.0742	0.0150
C1	OB	O23	9.5	1.859	136.5	3.545	5.058	0.0328	0.1210	0.0303	0.0268	0.1332	0.0286
C2	O11	O22	27.4	1.978	170.3	2.776	4.737	0.0783	0.1434	0.0874	0.0168	0.0809	0.0162
C2	OB	O13	6.9	1.848	131.0	3.827	5.229	0.0250	0.0991	0.0219	0.0236	0.1098	0.0235
C3	O11	O22	18.7	1.923	165.5	2.981	4.867	0.0586	0.1525	0.0595	0.0156	0.0728	0.0147
C3	OB	O23	6.5	1.849	130.7	3.855	5.252	0.0238	0.0944	0.0206	0.0224	0.1054	0.0220
C4	O11	O23	18.7	1.927	167.1	2.990	4.888	0.0590	0.1466	0.0595	0.0164	0.0795	0.0158
C4	*OB*	*O23*		*1.838*	*122.3*	*4.340*	*5.544*						
C5	O11	O22	19.6	1.931	165.5	2.957	4.851	0.0609	0.1519	0.0623	0.0161	0.0774	0.0155
C6	O11	O22	16.5	1.919	168.1	3.060	4.954	0.0538	0.1397	0.0525	0.0160	0.0762	0.0151
C7	O11	O22	20.0	1.937	165.9	2.942	4.844	0.0620	0.1516	0.0637	0.0160	0.0756	0.0153
C8	OB	O22	10.1	1.866	137.5	3.485	5.021	0.0347	0.1245	0.0322	0.0271	0.1360	0.0290
C8	O11	OB	4.9	1.839	117.4	4.125	5.232	0.0180	0.0806	0.0156	0.0180	0.0826	0.0161
C9	OB	O23	9.5	1.864	135.9	3.549	5.057	0.0329	0.1191	0.0302	0.0270	0.1328	0.0286
C9	O11	OB	5.2	1.842	124.3	3.983	5.247	0.0196	0.0848	0.0166	0.0190	0.0915	0.0178
C10	OB	O22	11.8	1.869	140.6	3.366	4.955	0.0393	0.1354	0.0377	0.0281	0.1438	0.0307
C11	OB	O22	12.4	1.870	141.1	3.334	4.932	0.0408	0.1392	0.0395	0.0286	0.1470	0.0315
C12	OB	O22	12.5	1.870	141.0	3.330	4.926	0.0410	0.1404	0.0398	0.0288	0.1477	0.0317
C13	OB	O23	13.4	1.876	142.2	3.282	4.901	0.0434	0.1445	0.0427	0.0294	0.1505	0.0325
C14	OB	O22	11.9	1.869	141.0	3.359	4.954	0.0396	0.1364	0.0381	0.0282	0.1444	0.0309
C15	OB	O23	13.4	1.875	141.8	3.278	4.891	0.0435	0.1446	0.0427	0.0292	0.1508	0.0324
C16	*OB*	*O13*		*1.840*	*118.8*	*4.189*	*5.327*						
C17	OB	O13	6.8	1.847	127.6	3.861	5.198	0.0243	0.0994	0.0215	0.0234	0.1108	0.0234
C18	OB	O11	5.0	1.838	117.9	4.166	5.281	0.0180	0.0795	0.0158	0.0180	0.0827	0.0165
C19	OB	O13	7.3	1.848	131.2	3.757	5.165	0.0266	0.1044	0.0234	0.0244	0.1162	0.0247
C20	OB	O11	5.4	1.839	122.1	3.994	5.209	0.0200	0.0871	0.0172	0.0193	0.0947	0.0187
C21	OB	O13	10.0	1.852	136.9	3.510	5.024	0.0336	0.1281	0.0318	0.0270	0.1364	0.0290
C22	*OB*	*O11*		*1.838*	*119.9*	*4.287*	*5.441*						

Table S4. Hydrogen bond properties in different conformers of **H_2L** form of **HEDP**. All values are in atomic units except E_{HB} that is in kcal/mol.

	D	A	E_{HB}	R_{DH}	θ_{DHA}	R_{HA}	R_{DA}	ρ_{bc}	L_{bc}	V_{bc}	ρ_{rc}	L_{rc}	V_{rc}
C1	O21	O12	10.7	1.878	155.9	3.334	5.106	0.0372	0.1225	0.0341	0.0144	0.0646	0.0133
C1	O11	O22	10.8	1.880	154.1	3.335	5.092	0.0375	0.1226	0.0344	0.0148	0.0674	0.0139
C2	O11	O22	11.5	1.882	158.1	3.282	5.077	0.0396	0.1268	0.0366	0.0149	0.0672	0.0138
C2	O21	O12	9.9	1.875	156.4	3.390	5.164	0.0349	0.1152	0.0314	0.0143	0.0643	0.0131
C3	O11	O22	11.2	1.881	157.4	3.299	5.088	0.0387	0.1253	0.0357	0.0148	0.0670	0.0138
C3	O21	O12	10.1	1.877	154.7	3.373	5.133	0.0356	0.1178	0.0322	0.0145	0.0646	0.0133
C4	O11	O23	13.9	1.889	162.1	3.151	4.983	0.0462	0.1414	0.0444	0.0150	0.0701	0.0141
C4	*OB*	*O13*		*1.844*	*125.7*	*4.155*	*5.442*						
C5	O11	O23	13.3	1.885	160.6	3.186	5.003	0.0443	0.1387	0.0423	0.0145	0.0672	0.0135
C5	*OB*	*O13*		*1.841*	*118.4*	*4.469*	*5.584*						
C6	O11	O22	13.4	1.886	161.3	3.176	4.999	0.0447	0.1392	0.0428	0.0148	0.0686	0.0138
C6	*OB*	*O13*		*1.845*	*127.0*	*4.083*	*5.398*						
C7	O11	O22	10.6	1.872	156.7	3.339	5.112	0.0366	0.1234	0.0336	0.0140	0.0638	0.0129
C7	*OB*	*O23*		*1.841*	*122.3*	*4.176*	*5.389*						
C8	O11	O23	10.6	1.878	154.7	3.342	5.103	0.0369	0.1222	0.0338	0.0145	0.0678	0.0137
C8	*OB*	*O11*		*1.838*	*121.4*	*4.315*	*5.500*						
C9	O11	O22	11.3	1.874	157.4	3.293	5.074	0.0387	0.1286	0.0360	0.0143	0.0665	0.0135
C9	*OB*	*O23*		*1.844*	*124.5*	*4.062*	*5.328*						
C10	O11	O22	10.5	1.873	156.8	3.342	5.117	0.0365	0.1228	0.0335	0.0140	0.0637	0.0129
C10	*OB*	*O23*		*1.841*	*122.3*	*4.178*	*5.390*						
C11	O11	O23	13.1	1.885	162.0	3.190	5.017	0.0440	0.1366	0.0418	0.0149	0.0699	0.0140
C11	*OB*	*O13*		*1.845*	*127.0*	*4.065*	*5.382*						
C12	O21	O12	9.6	1.871	151.9	3.432	5.158	0.0341	0.1145	0.0307	0.0148	0.0658	0.0137
C12	OB	O22	7.9	1.855	134.0	3.679	5.143	0.0282	0.1089	0.0252	0.0248	0.1210	0.0256
C12	O11	OB	5.9	1.844	122.0	3.909	5.130	0.0218	0.0937	0.0187	0.0204	0.1003	0.0199
C13	O21	O12	8.4	1.865	141.6	3.566	5.159	0.0304	0.1079	0.0269	0.0154	0.0677	0.0145
C13	OB	O22	7.7	1.849	134.1	3.700	5.160	0.0275	0.1078	0.0247	0.0243	0.1190	0.0251
C13	O11	O21	3.9	1.840	131.4	4.160	5.551	0.0157	0.0617	0.0123	0.0119	0.0548	0.0110
C14	O11	O23	10.2	1.875	154.1	3.368	5.121	0.0358	0.1198	0.0326	0.0144	0.0671	0.0136
C14	*OB*	*O11*		*1.836*	*119.4*	*4.364*	*5.502*						
C15	O11	O23	10.5	1.874	157.4	3.337	5.117	0.0367	0.1227	0.0336	0.0142	0.0651	0.0131
C16	O11	O22	11.2	1.878	155.1	3.306	5.072	0.0384	0.1270	0.0356	0.0146	0.0678	0.0137
C16	*OB*	*O11*		*1.836*	*119.4*	*4.432*	*5.569*						
C17	O21	O12	10.4	1.875	156.5	3.348	5.122	0.0363	0.1214	0.0332	0.0140	0.0639	0.0129
C18	O21	O13	10.0	1.873	155.1	3.378	5.138	0.0351	0.1185	0.0318	0.0141	0.0639	0.0130
C19	O21	O11	8.5	1.854	154.3	3.476	5.208	0.0306	0.1117	0.0271	0.0126	0.0589	0.0113
C19	OB	O22	8.8	1.855	134.2	3.606	5.076	0.0306	0.1172	0.0280	0.0261	0.1300	0.0278
C20	O21	O13	11.1	1.876	157.3	3.310	5.093	0.0381	0.1256	0.0353	0.0143	0.0658	0.0133

Table S4. (Continued)

	D	A	E_{HB}	R_{DH}	θ_{DHA}	R_{HA}	R_{DA}	ρ_{bc}	L_{bc}	V_{bc}	ρ_{rc}	L_{rc}	V_{rc}
C21	O21	O11	5.2	1.842	134.8	3.919	5.378	0.0200	0.0789	0.0164	0.0126	0.0588	0.0119
C21	O11	O23	5.0	1.851	133.0	3.975	5.408	0.0202	0.0756	0.0161	0.0138	0.0609	0.0128
C22	O11	O21	7.2	1.848	146.7	3.601	5.245	0.0265	0.1023	0.0229	0.0113	0.0525	0.0100
C22	*OB*	*O22*		*1.843*	*122.0*	*4.132*	*5.342*						
C23	OB	O12	6.4	1.847	128.7	3.878	5.236	0.0234	0.0959	0.0205	0.0225	0.1057	0.0221
C24	OB	O22	7.2	1.849	132.0	3.769	5.193	0.0258	0.1025	0.0229	0.0238	0.1143	0.0241
C24	O21	O11	4.2	1.842	133.9	4.029	5.469	0.0170	0.0688	0.0135	0.0117	0.0520	0.0103
C25	*OB*	*O13*		*1.844*	*124.4*	*4.052*	*5.317*						
C26	*OB*	*O23*		*1.843*	*122.8*	*4.095*	*5.324*						
C27	O21	O11	5.3	1.846	141.9	3.831	5.405	0.0206	0.0805	0.0167	0.0108	0.0496	0.0094
C28	OB	O23	6.0	1.845	127.5	3.960	5.290	0.0220	0.0906	0.0192	0.0217	0.0980	0.0206
C29	*OB*	*O11*		*1.839*	*118.6*	*4.180*	*5.312*						
C30	O11	O23	8.4	1.862	146.7	3.549	5.206	0.0303	0.1076	0.0267	0.0152	0.0667	0.0140
C31	*OB*	*O23*		*1.839*	*116.1*	*4.594*	*5.648*						
C32	OB	O21	5.1	1.840	122.3	4.056	5.273	0.0192	0.0824	0.0164	0.0189	0.0893	0.0177
C32	O11	OB	4.7	1.839	119.6	4.068	5.228	0.0177	0.0805	0.0150	0.0175	0.0851	0.0160
C33	O21	O11	7.1	1.850	151.2	3.596	5.293	0.0265	0.0997	0.0227	0.0111	0.0531	0.0099
C34	*OB*	*O11*		*1.837*	*115.2*	*4.284*	*5.332*						
C35	*OB*	*O11*		*1.838*	*118.7*	*4.161*	*5.296*						
C36	OB	O11	5.3	1.838	122.4	4.031	5.251	0.0198	0.0852	0.0170	0.0193	0.0909	0.0182
C37	*OB*	*O21*		*1.838*	*115.2*	*4.297*	*5.345*						

Table S5. Fitted functional forms, fit parameters and regression quality obtained between the best correlated features of H-bond. All values are in atomic units except the frequency shift ($\Delta\nu$) which is in cm^{-1} and the H-bond energy (E_{HB}) that is in kcal/mol.

Functional form		RMSE	r^2
$\ln(\rho_{bc}) = 0.2383 - 1.0456 R_{HA}$		0.0563	0.9770
$V_{bc} = -0.0065 + 1.1206 \rho_{bc}$		0.0012	0.9931
$V_{bc} = -0.0003 + 0.7731 \rho_{bc} + 4.2851 \rho_{bc}^2$		0.0005	0.9986
$V_{rc} = -0.0041 + 1.2094 \rho_{rc}$		0.0006	0.9920
$V_{rc} = 0.0035 + 0.3850 \rho_{rc} + 20.4688 \rho_{rc}^2$		0.0003	0.9977
$\ln(V_{bc}) = -5.4372 + 16.3928 L_{bc}$		0.0481	0.9820
$R_{DH} = 1.8242 + 70.0533 R_{HA}^{-6}$		0.0041	0.9782
$V_{rc} = -0.0016 + 0.2240 L_{rc}$		0.0004	0.9959
$R_{DH} = 1.83286 - 0.0001 \Delta\nu$		0.0012	0.9980
$E_{HB} = 4.2387 - 0.0143 \Delta\nu$		0.9797	0.9534
$E_{HB} = 2671.08 R_{HA}^{-4.539}$	(all data)	0.6917	0.9961
$E_{HB} = 4431.07 R_{HA}^{-5.001}$	(inter-phosphonate data)	0.2562	0.9996
$E_{HB} = 2691.97 R_{HA}^{-4.466}$	(hydroxyl-phosphonate data)	0.1706	0.9996

Table S6. Relative MP2 gas-phase energies (E_{gas}) of some selected conformers of MBP obtained with different basis sets and extrapolated to CBS limit. Bold values are at the same level as those reported in Tables 1, S1 and S2. All values are in kcal/mol relative to the corresponding **C1** conformer.

	H_2L-C2	**H_2L-C5**	**H_2L-C8**	**HL-C4**
cc-pVDZ	16.53	11.44	31.79	29.15
cc-pVTZ	14.32	10.08	28.91	26.21
cc-pVQZ	13.48	9.54	27.69	24.75
aug-cc-pVDZ	12.73	8.74	25.90	22.29
aug-cc-pVTZ	12.97	9.20	26.83	23.08
aug-cc-pVQZ	12.96	9.25	26.89	23.06
6-311++G(2df,2p)	**13.32**	**9.28**	**27.61**	**23.54**
CBS1	13.12	9.27	27.04	23.10
CBS2	13.24	9.48	27.51	23.54

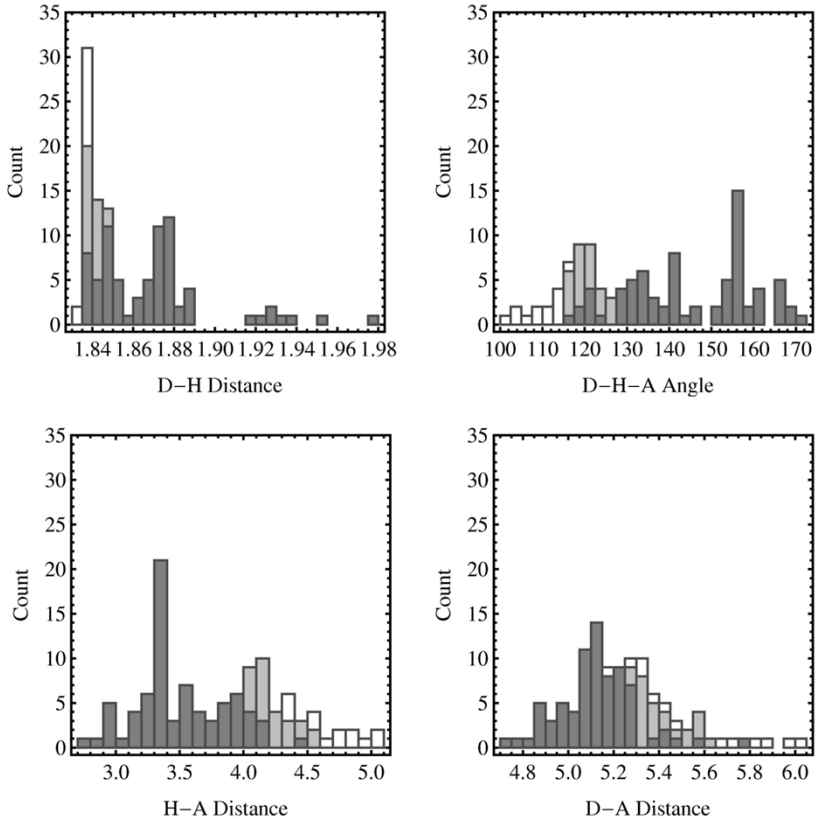

Figure S1. Distribution of all H-bonds in **MBP** and **HEDP** over geometrical parameters. All distances and angles are in atomic units and degrees, respectively. Dark gray: topologically approved H-bonds (S^{top}); Light gray: topologically unapproved H-bonds considered in motif analysis ($S^{tight} - S^{top}$); White: remaining weak H-bonds ($S^{loose} - S^{tight}$).

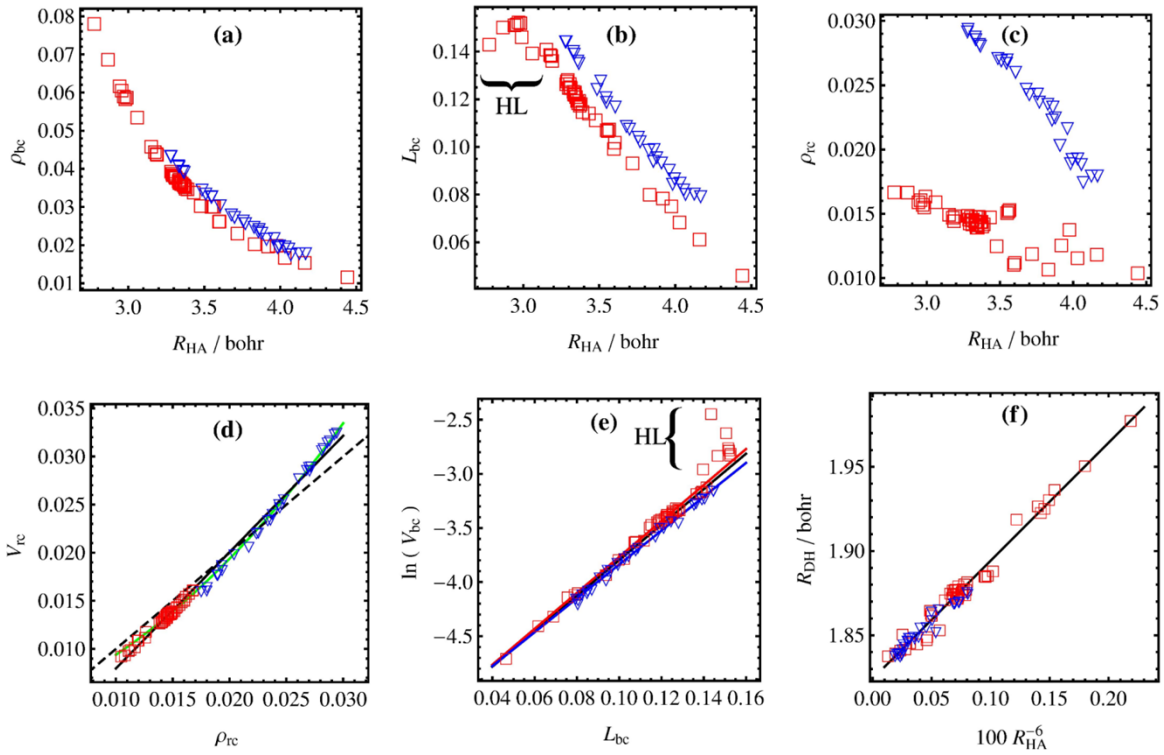

Fig.S2 Important correlations between different features of hydrogen bonds. Red squares: inter-phosphonate H-bonds in S^{top}; Blue triangles: hydroxyl-phosphonate H-bonds in S^{top}; Black circles: H-bonds in S^{tight} but not in S^{top}; Gray circles: H-bonds in S^{loose} but not in S^{tight}; Black, red and blue lines represent a linear fit of whole, inter-phosphonate and hydroxyl-phosphonate data. The green curve represents a quadratic fit of whole data. All values are in atomic units.