Supporting Information (PART II)

Quantitative Antidiabetic Activity Prediction for the Class of Guanidino- and Aminoguanidinopropionic Acid Analogs Based on Electron-Conformational Studies

Aleksandr V. Marenich, Pei-Han Yong, Isaac B. Bersuker, James E. Boggs *

Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry,

the University of Texas at Austin, 1 University Station A5300, Austin, TX 78712, USA

Contents: **page**

Table S1. Molecular structures of guanidino- and aminoguanidinopropionic acid analogs **1–154** **S2**

Table S2. Descriptors for prediction of antidiabetic activities **S18**

Table S3. The electron-conformational submatrix of activity and antidiabetic activities predicted by leave-one-out cross-validation **S25**

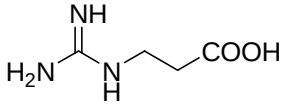
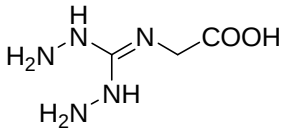
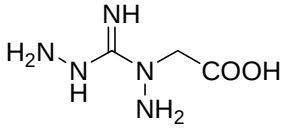
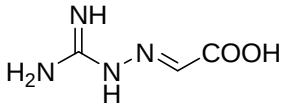
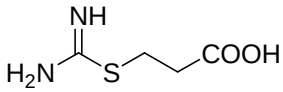
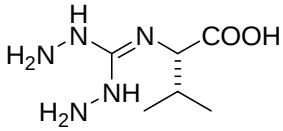
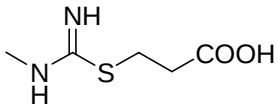
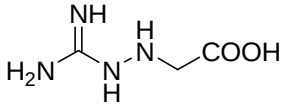
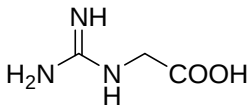
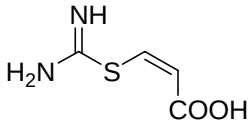
Table S4. Interaction indices *versus* Mulliken partial atomic charges in compound **82** **S27**

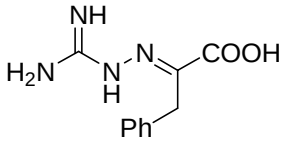
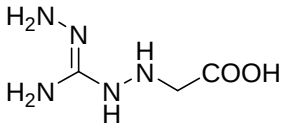
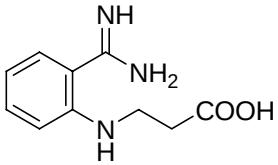
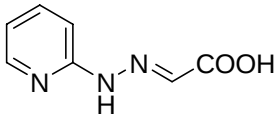
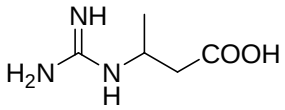
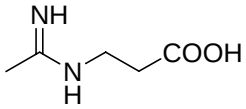
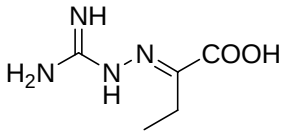
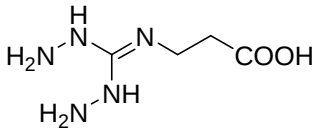
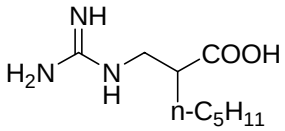
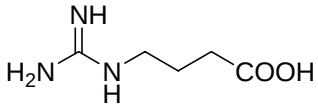
Total number of pages: 27

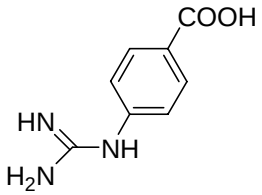
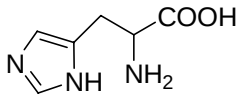
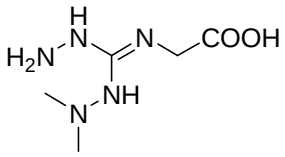
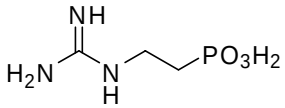
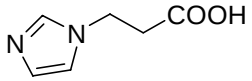
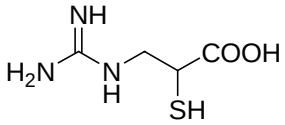
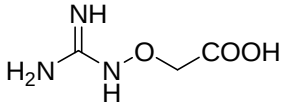
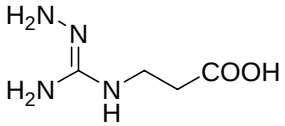
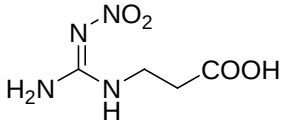
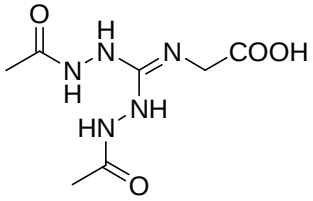
* Corresponding author. Tel: +1 512 466 9145. Fax: +1 512 471 8696.

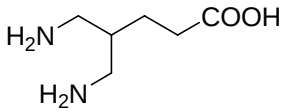
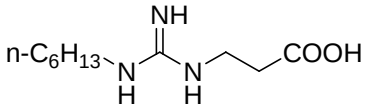
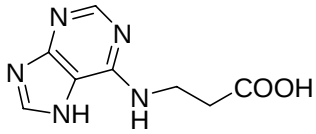
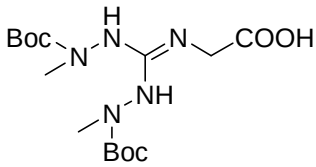
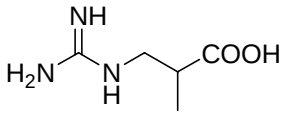
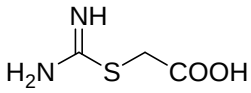
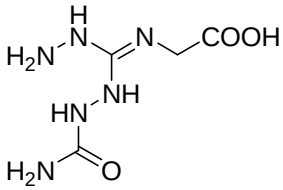
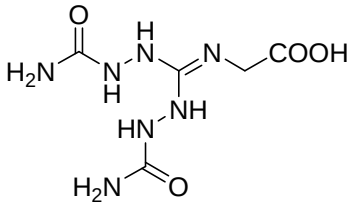
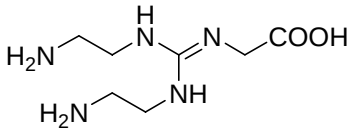
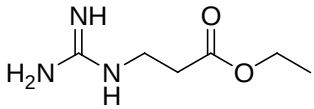
E-mail address: james.boggs@mail.utexas.edu (J.E. Boggs).

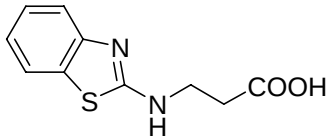
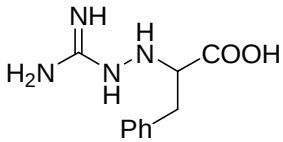
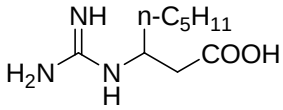
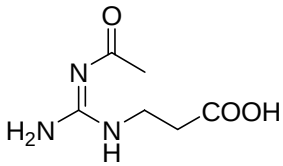
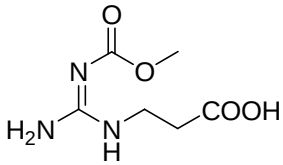
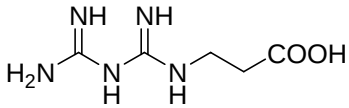
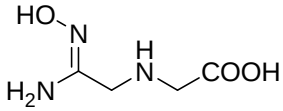
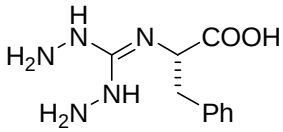
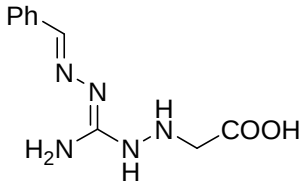
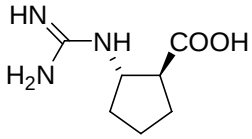
Table S1. Molecular Structures of Guanidino- and Aminoguanidinopropionic Acid Analogs **1**–**154** [a,b]

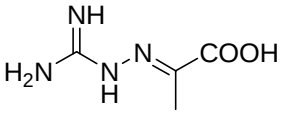
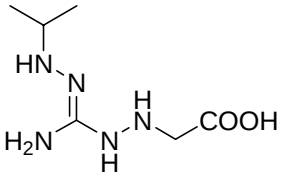
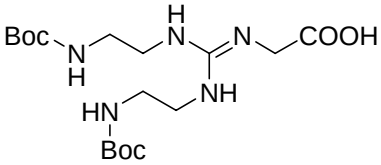
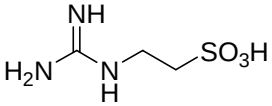
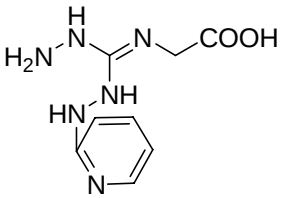
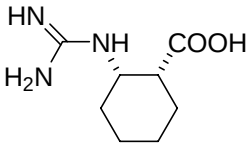
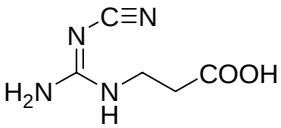
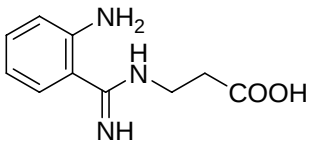
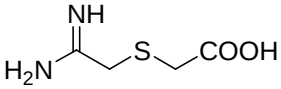
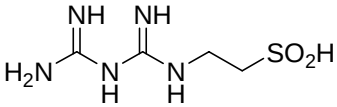
N		molecular structure	n_{conf}	MISS T/C (exp)	MISS T/C (calc)
1	I-1		2	0.52 ± 0.24	0.52
2	II-52		1	0.23 ± 0.05	0.39
3	II-53		1	0.43 ± 0.20	0.55
4	II-18		2	0.52 ± 0.24	0.77
5	I-67		2	0.53 ± 0.34	0.79
6	II-56		1	0.56 ± 0.21	0.78
7	I-76		1	0.59 ± 0.33	0.43
8	I-69		2	0.61 ± 0.29	0.81
9	I-6		2	0.62 ± 0.24	0.55
10	I-73		2	0.68 ± 0.33	0.79

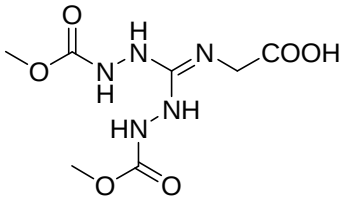
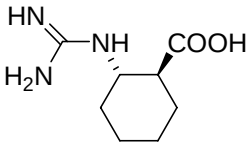
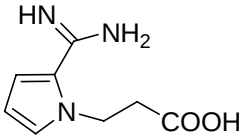
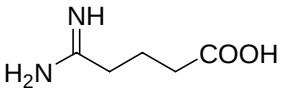
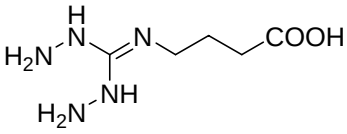
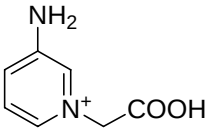
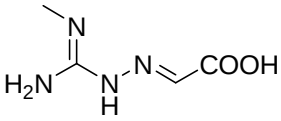
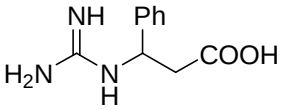
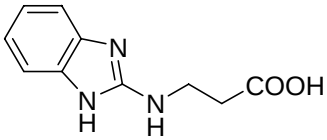
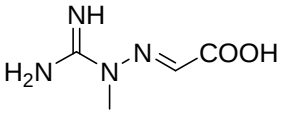
11	II-24		1	0.68 ± 0.34	0.99
12	II-7		1	0.69 ± 0.22	0.93
13	I-60		2	0.73	1.00
14	II-28		1	0.73 ± 0.21	0.72
15	I-9		4	0.77	0.76
16	I-63		1	0.77	0.83
17	II-21		2	0.77 ± 0.31	0.98
18	II-59		1	0.78 ± 0.14	0.71
19	I-15		1	0.79	1.00
20*	I-7		1	0.80	1.00

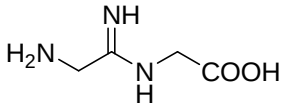
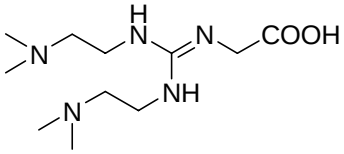
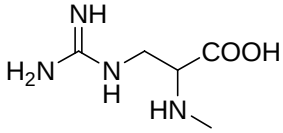
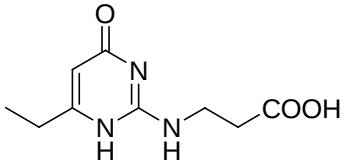
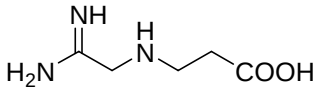
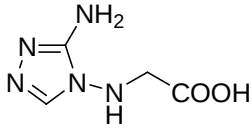
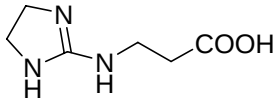
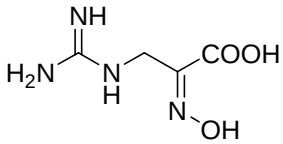
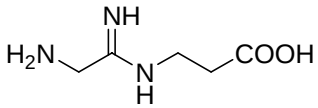
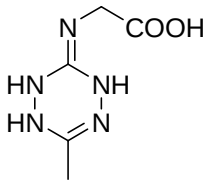
21*	I-24		1	0.80	1.00
22	I-46		3	0.80	0.82
23	II-63		1	0.80	0.96
24	I-37		3	0.81	0.79
25*	I-44		3	0.81	1.00
26	I-34		1	0.83	0.96
27	I-68		1	0.83	0.97
28	II-14		2	0.83	0.99
29	I-27		1	0.84	0.98
30	II-71		1	0.84	0.93

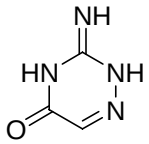
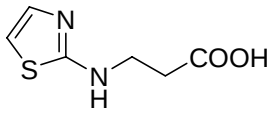
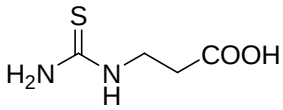
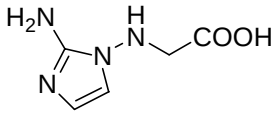
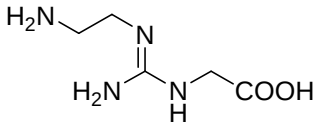
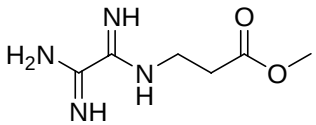
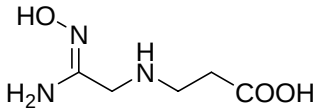
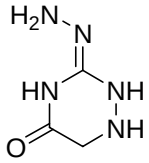
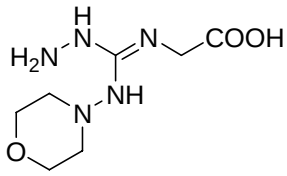
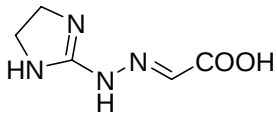
31*	II-78		1	0.84	1.00
32	I-17		4	0.85	1.00
33	I-57		1	0.85	0.98
34	II-66		2	0.85	0.97
35	I-8		1	0.86	0.95
36	I-70		1	0.86	0.90
37	II-64		1	0.86	0.93
38	II-70		1	0.86	0.96
39	II-73		1	0.86	0.96
40	I-41		7	0.87	0.87

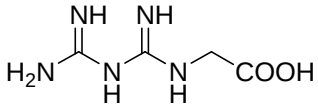
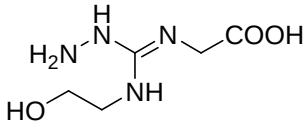
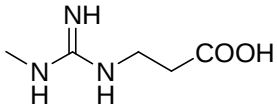
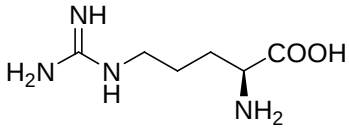
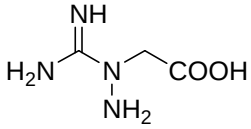
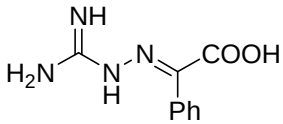
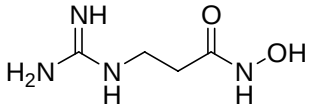
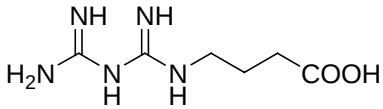
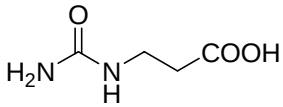
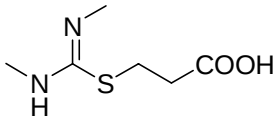
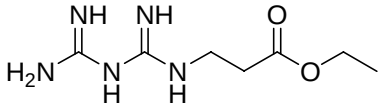
41	I-53		1	0.87	0.95
42	II-5		2	0.87	0.99
43	I-16		2	0.88	1.00
44	I-25		2	0.88	0.99
45	I-26		4	0.88	0.91
46	II-34		1	0.88	0.97
47*	II-43		1	0.88	1.00
48	II-57		1	0.88	0.98
49	II-48		2	0.89	0.99
50	I-22		1	0.90	0.99

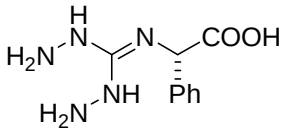
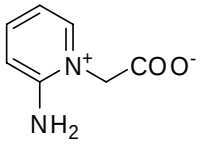
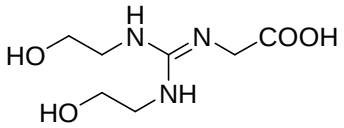
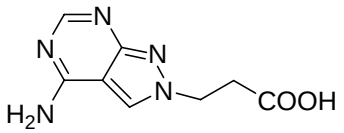
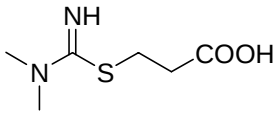
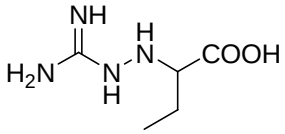
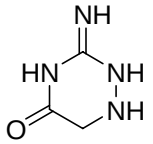
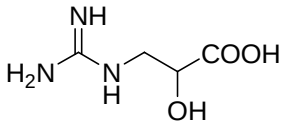
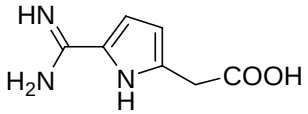
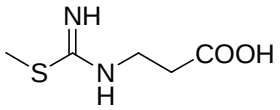
51	II-20		1	0.90	0.97
52	II-49		4	0.90	0.99
53	II-67		1	0.90	1.00
54	I-36		1	0.91	0.99
55	II-76		1	0.91	0.98
56	I-21		1	0.92	0.97
57	I-28		1	0.92	0.99
58	I-59		2	0.92	0.96
59	I-74		1	0.92	0.97
60	II-35		1	0.92	0.96

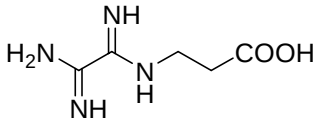
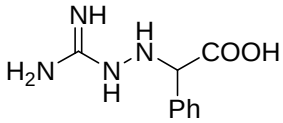
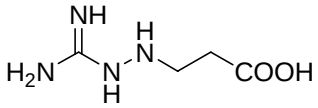
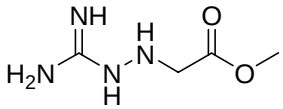
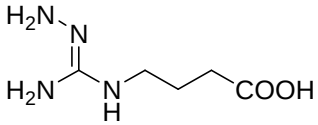
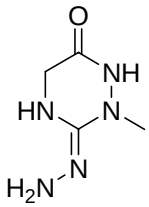
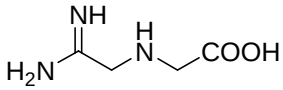
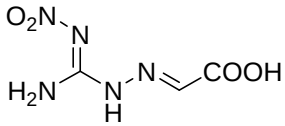
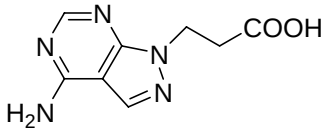
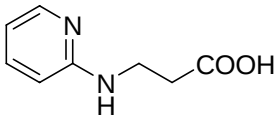
61	II-72		1	0.92	0.99
62	I-20		3	0.93	0.99
63	I-61		1	0.93	1.00
64	I-64		1	0.93	0.96
65	II-60		1	0.93	0.99
66*	I-51		1	0.94	1.00
67	II-26		1	0.94	0.90
68	I-13		1	0.95	0.99
69	I-54		1	0.95	0.99
70	II-25		1	0.95	0.96

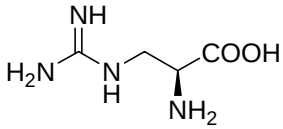
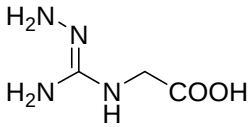
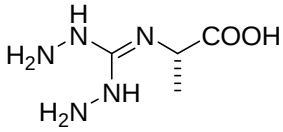
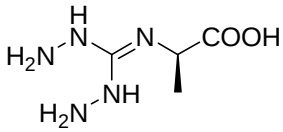
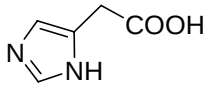
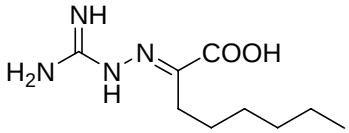
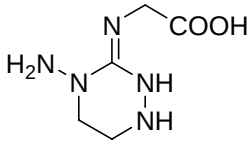
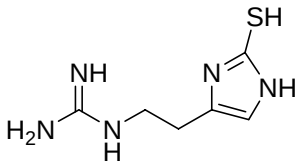
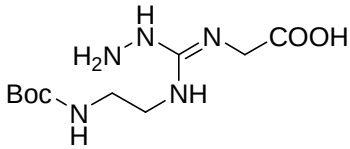
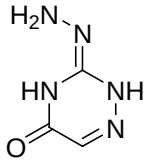
71	II-41		1	0.95	0.92
72	II-68		1	0.95	0.99
73	I-31		1	0.96	0.99
74	I-58		3	0.96	1.00
75	II-46		2	0.96	0.99
76*	II-51		1	0.96	1.00
77	I-12		2	0.97	0.95
78	I-32		1	0.97	0.96
79	II-42		1	0.97	0.96
80	II-80		1	0.97	0.92

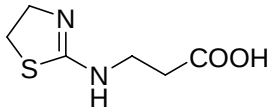
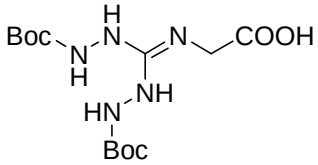
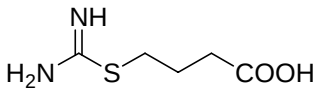
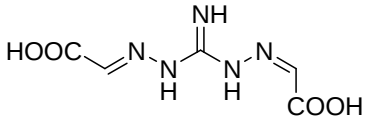
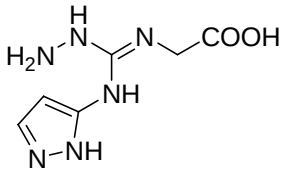
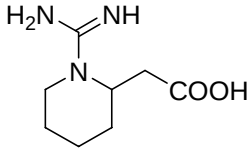
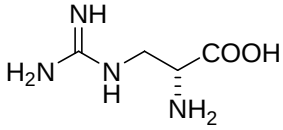
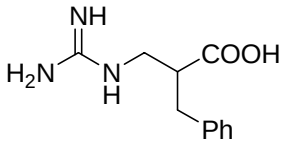
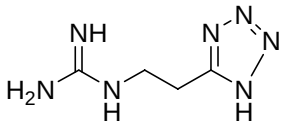
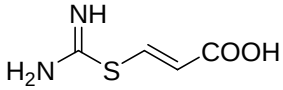
81*	II-31		1	0.98	1.00
82	I-52		1	0.99	0.90
83	I-66		2	0.99	0.77
84*	II-12		1	0.99	1.00
85	II-16		2	0.99	0.97
86	II-40		1	1.00	0.96
87*	II-44		2	1.00	1.00
88*	II-50		1	1.00	1.00
89	II-75		1	1.00	0.94
90*	II-29		1	1.01	1.00

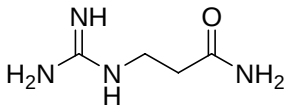
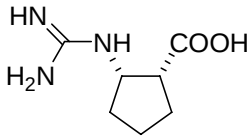
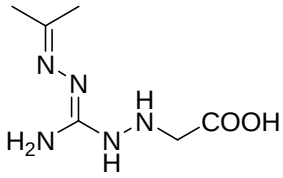
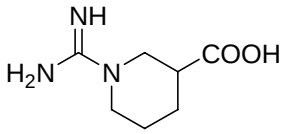
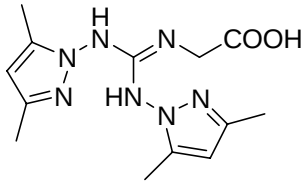
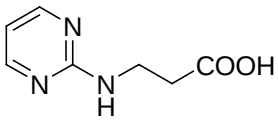
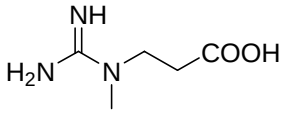
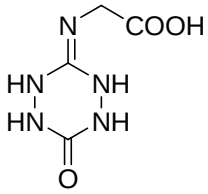
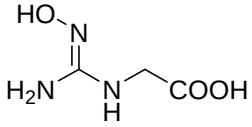
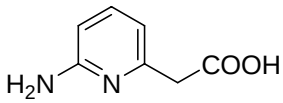
91	II-33		1	1.01	0.98
92	II-61		1	1.01	0.93
93	I-11		1	1.02	0.96
94*	I-35		2	1.02	1.00
95	II-10		1	1.04	0.85
96	II-22		1	1.04	0.99
97*	I-38		1	1.05	1.00
98	II-36		1	1.05	0.97
99	I-65		1	1.06	0.93
100	I-77		2	1.06	0.88
101	II-37		1	1.06	0.96

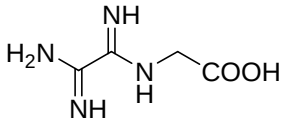
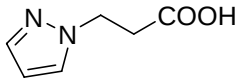
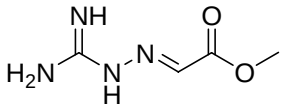
102	II-58		1	1.06	0.96
103*	I-50		1	1.07	1.00
104	II-69		1	1.07	0.98
105*	I-56		1	1.08	1.00
106	I-78		1	1.08	0.84
107	II-4		2	1.08	0.84
108*	II-11		1	1.08	1.00
109	I-33		1	1.10	0.96
110*	I-62		2	1.10	1.00
111	I-75		1	1.10	0.97

112	II-39		1	1.10	0.96
113	II-6		1	1.11	0.99
114	II-8		2	1.11	0.96
115	II-3		1	1.12	0.79
116	II-15		1	1.14	0.98
117*	II-82		1	1.16	1.00
118	II-45		1	1.18	0.96
119	II-27		1	1.23	0.96
120	I-55		2	1.27	1.00
121	I-47		2	0.20 ± 0.17	0.43

122	I-29		1	0.35 ± 0.20	0.96
123	II-9		3	0.51 ± 0.24	0.54
124	II-54		1	0.57 ± 0.28	0.30
125	II-55		1	0.66 ± 0.29	0.30
126*	I-45		1	0.68 ± 0.35	1.00
127	II-23		1	0.73 ± 0.27	0.99
128	II-79		1	0.77 ± 0.38	0.89
129	I-40		3	0.80	0.93
130	II-62		1	0.81	0.96
131*	II-32		1	0.84	1.00

132	I-79		1	0.85	0.99
133	II-65		1	0.86	0.99
134	I-71		1	0.87	0.72
135	II-30		1	0.88	1.00
136*	II-74		1	0.89	1.00
137	I-19		1	0.91	0.99
138	I-30		1	0.92	0.96
139	I-14		1	0.93	1.00
140*	I-39		1	0.94	1.00
141	I-72		2	0.95	0.79

142	I-42		2	0.96	0.82
143	I-23		1	0.97	0.99
144	II-47		1	0.98	0.98
145	I-18		1	1.00	0.97
146	II-77		1	1.00	0.98
147	I-48		1	1.01	0.42
148	I-10		1	1.02	0.89
149	II-81		1	1.04	0.92
150	II-17		1	1.06	0.81
151	I-49		1	1.08	0.94

152	II-38		1	1.08	0.97
153*	I-43		1	1.11	1.00
154	II-19		1	1.16	0.72

[a] The first column refers to the order of appearance of these compounds in the present paper where compounds **1**–**120** present the training set and compounds **121**–**154** present the test set used in our study. The second column refers to the order of appearance of these compounds in the experimental papers: Larsen et al., *J. Med. Chem.* **2001,** *44*, 1217 (paper I); Vaillancourt et al., *J. Med. Chem.* **2001,** *44*, 1231 (paper II). The values of n_{conf} show the numbers of conformers for each compounds that are lying below 1 kcal/mol according to the conformational analysis performed in the present study. Experimental data on mouse insulin sensitizing screen test/control MISS T/C (exp) blood glucose levels are quoted from the above-mentioned papers and given along with the corresponding theoretical values MISS T/C (calc) obtained in the present study. The asterisk refers to those compounds that do not have the pharmacophore of antidiabetic activity revealed in the present study for any conformation below 1 kcal/mol. The tertiary-butoxycarbonyl $(CH_3)_3C–O–CO$ and phenyl C_6H_5 groups were abbreviated to *Boc* and *Ph*, respectively.

[b] Note that we corrected the structure **100** (I-77) accordingly with the description of this compound in the experimental section of the original paper (Larsen et al., *J. Med. Chem.* **2001,** *44*, 1217); the structure I-77 was confused with the structure I-76 in Table 7 of this article.

Table S2. Descriptors for Prediction of Antidiabetic Activities [a]

N	pharmacophore	E	R_c	a_{24}	ΔG_S°
1	**O_1 O_2 N_1 N_2**	**0**	**0**	**5.755**	**–13.79**
2	O_1 O_2 N_2 N_4	0	0	5.600	–10.73
3	O_1 O_2 N_4 N_3	0	2.627	5.191	–10.15
4	N_2 N_4 O_2 O_1	0	0	6.473	–12.58
5	S_1 N_1 O_1 O_2	0	0	6.364	–17.67
6	O_1 O_2 N_2 N_4	0	4.235	5.211	–13.97
7	S_1 N_2 O_1 O_2	0	0	5.701	–10.10
8	O_1 O_2 N_2 N_3	0	0	6.408	–18.45
9	N_3 N_2 O_1 O_2	0	0	5.724	–16.27
10	S_1 N_2 O_1 O_2	0	0	6.581	–11.98
11	N_2 N_4 O_2 O_1	0	6.181	5.976	–40.89
12	N_2 N_3 O_2 O_1	0	0	6.641	–33.56
13	N_1 N_2 O_2 O_1	0	5.302	6.867	–40.63
14	O_2 O_1 N_1 N_2	0	4.564	3.954	–39.85
15	O_1 O_2 N_1 N_2	0	2.945	5.610	–11.68
16	N_2 N_1 O_2 O_1	0	4.028	4.721	–34.66
17	N_2 N_4 O_2 O_1	0	4.105	5.948	–40.80
18	O_1 O_2 N_4 N_2	0	2.207	5.729	–10.19
19	N_1 N_3 O_2 O_1	0	7.883	6.570	–39.35
22	O_1 O_2 N_2 N_3	0.44	2.773	5.128	–18.15
23	N_3 N_2 O_2 O_1	0	3.945	5.571	–39.95
24	O_1 O_3 N_2 N_1	0.97	0	5.058	–19.86
26	N_1 N_3 O_2 O_1	0	0	6.560	–45.95

27	N_2	N_3	O_2	O_1	0	2.040	6.409	–42.65
28	N_1	N_2	O_2	O_1	0.62	2.354	6.428	–43.62
29	N_1	N_4	O_2	O_1	0	3.556	6.501	–37.90
30	O_4	O_2	N_5	N_2	0	5.798	4.836	–35.56
32	N_1	N_3	O_2	O_1	0	9.278	6.632	–41.11
33	N_1	N_5	O_2	O_1	0	6.200	4.730	–59.46
34	O_5	O_6	N_1	N_3	0	8.073	5.511	–14.46
35	N_1	N_3	O_2	O_1	0	0	6.573	–41.29
36	N_1	N_2	C_2	O_1	0	0	6.036	–41.59
37	O_3	N_6	N_2	N_1	0	5.086	5.263	–30.30
38	O_1	O_2	N_5	N_2	0	4.964	5.679	–33.16
39	N_1	N_2	O_2	O_1	0	6.096	5.556	–26.96
40	N_1	N_2	O_1	O_2	0	1.714	6.348	–15.44
41	N_2	N_1	O_2	O_1	0	6.022	4.720	–44.57
42	N_2	N_4	O_2	O_1	0	5.968	6.512	–39.00
43	N_1	N_3	O_2	O_1	0	7.739	6.525	–39.88
44	N_1	N_3	O_2	O_1	0	3.731	6.480	–45.40
45	N_1	N_2	O_1	O_2	0	6.017	5.322	–16.59
46	N_1	N_3	O_2	O_1	0	0	6.645	–50.17
48	N_2	N_4	O_2	O_1	0	6.585	5.530	–39.71
49	N_2	N_5	O_2	O_1	0	6.810	6.043	–38.30
50	N_1	N_3	O_2	O_1	0	3.658	6.553	–40.20
51	N_2	N_4	O_2	O_1	0	3.427	5.949	–41.55
52	N_2	N_5	O_2	O_1	0	4.099	6.505	–36.45
53	N_2	N_3	N_5	O_6	0	8.777	6.555	–28.64

54	N_1	N_3	O_1	O_2	0	0	6.697	–64.08
55	N_5	N_3	O_2	O_1	0	5.892	5.572	–44.71
56	N_1	N_3	O_2	O_1	0	3.965	6.052	–35.82
57	N_1	N_4	O_2	O_1	0	4.128	6.521	–46.10
58	N_2	N_1	O_2	O_1	0	6.649	4.764	–41.76
59	C_3	N_2	O_2	O_1	0	2.078	6.376	–42.80
60	C_1	O_2	N_3	N_4	0	1.325	5.610	–60.13
61	O_3	N_3	O_2	O_1	0	7.805	6.081	–38.61
62	N_1	N_3	O_2	O_1	0.24	3.908	6.539	–39.33
63	N_2	N_3	O_2	O_1	0	4.915	6.954	–41.70
64	N_1	N_2	O_2	O_1	0	2.196	6.085	–39.66
65	N_1	N_2	O_2	O_1	0	2.163	6.915	–39.63
67	N_2	N_4	O_2	O_1	0	0	6.049	–41.14
68	N_1	N_3	O_2	O_1	0	5.848	6.510	–41.87
69	N_3	N_2	O_2	O_1	0	6.033	6.029	–42.78
70	N_3	N_4	O_2	O_1	0	2.138	6.201	–41.79
71	N_1	N_2	O_2	O_1	0	2.763	5.310	–45.39
72	N_2	N_3	O_2	O_1	0	7.674	5.364	–38.75
73	N_2	N_4	O_2	O_1	0	3.988	6.556	–42.89
74	N_3	N_1	O_2	O_3	0	6.731	6.602	–47.92
75	N_2	N_3	O_2	O_1	0.29	2.427	6.835	–43.75
77	N_1	N_3	O_2	O_1	0	0	6.457	–43.89
78	N_2	N_4	O_2	O_1	0	0	6.550	–47.65
79	N_1	N_2	O_2	O_1	0	2.603	5.899	–44.89
80	O_1	O_2	N_2	N_3	0	5.316	5.631	–14.82

82	N_1	N_2	O_2	O_1	0	4.290	4.721	–42.79
83	N_1	N_2	O_1	O_2	0	0	6.190	–19.63
85	N_2	N_4	O_2	O_1	0	4.069	5.600	–45.08
86	N_3	N_4	O_1	O_2	0	2.368	6.672	–24.95
89	O_2	O_3	N_2	N_4	0	5.698	5.725	–16.52
91	N_1	N_4	O_2	O_1	0	2.401	5.985	–54.70
92	N_2	N_3	O_2	O_1	0	3.307	5.208	–46.02
93	N_1	N_3	O_2	O_1	0	0	6.639	–42.33
95	N_3	N_4	O_2	O_1	0	0	5.742	–41.04
96	N_2	N_4	O_2	O_1	0	6.201	5.879	–41.67
98	N_1	N_3	O_2	O_1	0	0	6.687	–48.35
99	N_1	N_2	O_2	O_1	0	0	6.537	–35.71
100	S_1	N_2	O_1	O_2	0	4.816	5.743	–8.68
101	N_1	N_2	O_2	O_1	0	3.965	6.172	–24.91
102	N_2	N_4	O_2	O_1	0	5.091	5.619	–34.01
104	N_1	N_3	O_2	O_1	0	3.858	5.276	–61.44
106	S_1	N_1	O_1	O_2	0	5.329	5.343	–9.50
107	N_2	N_4	O_1	O_2	0	4.114	5.617	–11.52
109	N_1	N_3	O_2	O_1	0	0	6.556	–46.50
111	N_1	S_1	O_2	O_1	0	0	7.012	–40.11
112	N_1	N_2	O_2	O_1	0	2.540	5.951	–45.31
113	N_2	N_3	O_2	O_1	0	5.918	6.624	–34.23
114	N_2	N_3	O_2	O_1	0	0	6.822	–41.50
115	N_2	N_3	O_1	O_2	0	0	6.432	–15.53
116	N_2	N_4	O_2	O_1	0	2.703	6.665	–37.78

118	N_2	N_3	O_2	O_1	0	2.188	6.115	–42.42
119	N_2	N_5	O_2	O_1	0	4.856	6.367	–15.41
120	N_3	N_1	O_2	O_1	0.21	4.433	7.015	–45.37
121	O_1	O_2	N_2	N_1	0	0	5.733	–9.29
122	N_3	N_2	O_2	O_1	0	0	6.551	–44.50
123	N_1	N_4	O_1	O_2	0	0	5.722	–15.99
124	O_1	O_2	N_1	N_2	0	0	5.506	–8.58
125	O_1	O_2	N_1	N_2	0	0	5.506	–8.58
127	N_2	N_4	O_2	O_1	0	7.653	5.930	–39.43
128	N_1	N_5	O_2	O_1	0	4.598	5.816	–10.86
129	N_2	N_1	N_3	N_4	0	1.463	6.814	–18.31
130	N_1	N_2	O_4	C_{10}	0	4.626	5.473	–37.25
132	N_1	S_1	O_2	O_1	0	2.137	6.944	–41.62
133	N_3	O_6	N_5	O_3	0	6.404	6.328	–35.28
134	N_1	S_1	O_2	O_1	0	0	5.454	–35.00
135	N_4	N_3	O_4	O_3	0	5.628	6.310	–82.71
137	N_1	N_3	O_2	O_1	0	4.043	6.608	–39.62
138	N_3	N_2	O_2	O_1	0	0	6.551	–44.50
139	N_1	N_3	O_2	O_1	0	6.003	6.565	–40.90
141	S_1	N_2	O_1	O_2	0	0	6.581	–11.98
142	O_1	N_1	N_2	N_3	0	0	6.273	–23.12
143	O_2	O_1	N_1	N_2	0	3.574	6.453	–43.15
144	N_2	N_5	O_2	O1	0	3.674	6.466	–39.09
145	N_1	N_3	O_2	O_1	0	3.004	5.953	–46.40
146	N_3	N_5	O_2	O_1	0	7.259	5.498	–33.62

147	O_1	O_2	N_3	N_1	0	0	5.614	–12.23
148	N_1	N_2	O_2	O_1	0	0	5.871	–43.78
149	O_1	O_2	N_2	N_3	0	5.077	5.632	–17.76
150	N_1	N_3	O_2	O_1	0	0	5.579	–40.98
151	N_1	N_2	O_2	O_1	0	3.919	6.165	–19.63
152	N_3	N_4	O_2	O_1	0	2.189	6.190	–47.23
154	N_2	N_4	O_2	O_1	0	0	6.419	–9.57
155	O_2	O_1	N_1	N_2	0	1.488	3.995	–17.68
156	O_1	O_2	N_2	N_1	0	0	5.384	–11.81
157	O_1	O_2	N_2	N_1	0	0	4.586	–13.06
158	O_1	O_2	N_1	N_2	0	0	5.570	–14.42
159	O_1	O_2	N_2	N_1	0	0	5.021	–12.57
160	O_1	O_2	N_1	N_2	0	0	5.721	–10.20
161	N_3	N_1	O_1	O_2	0	0	5.426	–13.53
162	O_1	O_2	N_2	N_1	0	0	5.540	–15.65
163	N_3	N_1	O_2	O_1	0	0	4.573	–11.11
164	O_1	O_2	N_2	N_1	0	0	5.020	–10.19
165	O_1	O_2	N_1	N_2	0	0	4.862	–9.20
166	O_1	O_2	N_3	N_2	0	0	5.705	–14.57

[a] The data are given only for those molecules that have the pharmacophore of antidiabetic activity as revealed in this study including the molecules in the training set (**1**–**120**), in the test set (**121**–**154**), and in the series of new compounds **155**–**166** (see the main text) predicted to be potentially active. The reference molecule is bold-faced. The atoms corresponding to the pharmacophore's sites are indicated for each compound. Notations of atoms in the molecules **1**–

166 are shown in Part I of the Supporting Information along with the Cartesian coordinates of the molecules. The relative energy of the lowest conformation (*E*) that has the pharmacophore for each of the compounds and the corresponding free energy of solvation in aqueous solution calculated by the SM5.4/P solvation model (ΔG_S°) are given in kcal/mol. The other quantities (R_c and a_{24}) are defined in angstroms. A value of R_c is the distance from the geometric center of the pharmacophore to the most distant non-hydrogen out-of-Pha atom. To apply the parameter R_c only for the description of large anti-Pha shielding groups with respect to the pharmacophore centre (C), we excluded all "small" groups by reducing R_c to zero when the distance from any of the Pha sites (A_i) to the outmost non-Pha atom in the molecule was shorter than the corresponding C–A_i distance (the index i runs over all the pharmacophore sites, i = 1 – 4). The quantity of a_{24} is the distance between the Pha sites A_2 and A_4 in a given molecule. See Figures 1 and 2, Tables 3 and 4, and the main text for more information.

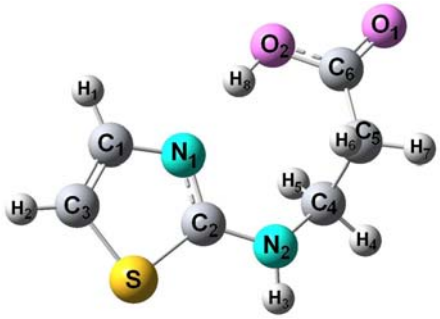
Table S3. Electron-Conformational Submatrix of Activity (ECSA) and Antidiabetic Activities MISS T/C Predicted by Leave-One-Out (LOO) Cross-Validation [a]

N	ECSA elements	absolute tolerances	MISS T/C (calc)	
			LOO	non-LOO
1	0.942	0.154	0.40	0.52
	2.399 1.135	0.401 0.363		
	3.388 4.801 0.985	1.315 1.011 0.176		
	4.475 5.838 2.246 0.899	1.093 1.885 0.260 0.201		
2	0.940	0.156	0.39	0.39
	2.399 1.137	0.401 0.364		
	3.410 4.801 0.988	1.292 1.011 0.179		
	4.495 5.852 2.243 0.907	1.113 1.898 0.263 0.192		
3	0.942	0.154	0.57	0.55
	2.399 1.136	0.401 0.363		
	3.382 4.771 0.993	1.321 0.981 0.184		
	4.503 5.874 2.251 0.900	1.121 1.921 0.255 0.199		
4	0.942	0.155	0.80	0.77
	2.390 1.131	0.409 0.358		
	3.406 4.787 0.989	1.297 0.998 0.174		
	4.491 5.805 2.257 0.905	1.108 1.851 0.249 0.194		
5	0.933	0.163	0.82	0.79
	2.366 1.134	0.430 0.362		
	3.329 4.775 0.998	1.372 0.986 0.189		
	4.463 5.830 2.261 0.895	1.080 1.876 0.245 0.204		
6	0.939	0.157	0.80	0.78
	2.399 1.137	0.400 0.365		
	3.409 4.805 0.989	1.294 1.016 0.180		
	4.494 5.873 2.243 0.907	1.112 1.920 0.259 0.192		
7	0.933	0.163	0.42	0.43
	2.365 1.139	0.435 0.366		
	3.303 4.755 0.998	1.400 0.965 0.189		
	4.452 5.846 2.261 0.895	1.069 1.892 0.245 0.205		
8	0.940	0.156	0.82	0.81
	2.391 1.131	0.408 0.358		
	3.382 4.787 0.997	1.321 0.998 0.188		
	4.485 5.822 2.257 0.897	1.102 1.868 0.249 0.202		
9	0.929	0.156	0.54	0.55
	2.385 1.125	0.415 0.353		
	3.382 4.796 0.991	1.321 1.007 0.182		
	4.505 5.848 2.263 0.903	1.122 1.894 0.244 0.197		
10	0.937	0.159	0.95	0.79
	2.364 1.128	0.435 0.355		
	3.306 4.762 0.989	1.396 0.973 0.180		
	4.459 5.798 2.261 0.901	1.076 1.844 0.245 0.185		

11	0.940				0.156				0.99	0.99
	2.383	1.126			0.417	0.353				
	3.378	4.785	0.984		1.325	0.996	0.175			
	4.527	5.831	2.264	0.911	1.144	1.877	0.242	0.188		
12	0.940				0.157				0.94	0.93
	2.384	1.124			0.416	0.352				
	3.360	4.770	0.982		1.343	0.981	0.173			
	4.482	5.794	2.265	0.911	1.099	1.840	0.241	0.189		
13	0.931				0.165				1.00	1.00
	2.386	1.127			0.414	0.355				
	3.411	4.773	0.982		1.291	0.984	0.173			
	4.480	5.784	2.265	0.910	1.097	1.830	0.241	0.189		
14	0.928				0.168				1.00	0.72
	2.401	1.152			0.398	0.172				
	3.412	4.803	0.995		1.291	1.013	0.186			
	4.553	5.943	2.249	0.906	0.737	1.222	0.257	0.193		
15	0.932				0.164				0.76	0.76
	2.388	1.128			0.412	0.355				
	3.380	4.788	0.987		1.323	0.999	0.178			
	4.501	5.816	2.258	0.903	1.118	1.862	0.248	0.196		
16	0.928				0.168				1.00	0.83
	2.389	1.138			0.410	0.366				
	3.410	4.840	0.981		1.292	0.509	0.173			
	4.486	5.900	2.264	0.910	1.103	1.947	0.242	0.189		
17	0.939				0.157				0.98	0.98
	2.383	1.126			0.417	0.353				
	3.377	4.785	0.983		1.326	0.996	0.174			
	4.528	5.834	2.264	0.912	1.146	1.880	0.242	0.188		
18	0.941				0.156				0.71	0.71
	2.399	1.135			0.401	0.363				
	3.379	4.764	0.991		1.324	0.975	0.182			
	4.484	5.844	2.244	0.907	1.101	1.891	0.262	0.192		
19	0.929				0.167				1.00	1.00
	2.385	1.125			0.415	0.353				
	3.349	4.769	0.981		1.353	0.980	0.172			
	4.478	5.798	2.265	0.910	1.095	1.844	0.242	0.189		

[a] The ECSA and the matrix of the ECSA tolerances for a given compound (out of the 19 compounds of the training set used in the pharmacophore identification) was obtained by removing this compound from the Pha identification. The new ECSA was then used for evaluation of the activity of the eliminated compound. We used the same three descriptors (Table S2) for all of the 19 quantitative models.

Table S4. Interaction Indices *versus* Mulliken Partial Atomic Charges in Compound **82**



	interaction index				Mulliken charge			
	MP2 [a]			RHF [b]	MP2 [a]			RHF [b]
atom	6-31G(d)	3-21G(d)	6-31+G(d,p)	6-31G(d)	6-31G(d)	3-21G(d)	6-31+G(d,p)	6-31G(d)
C_1	0.290	0.289	0.356	0.254	0.037	0.038	–0.078	0.099
C_2	0.189	0.160	0.412	0.137	0.238	0.319	–0.175	0.378
C_3	0.642	0.700	0.558	0.738	–0.398	–0.457	–0.289	–0.480
C_4	0.445	0.449	0.369	0.417	–0.195	–0.213	–0.084	–0.164
C_5	0.664	0.720	0.501	0.672	–0.421	–0.496	–0.236	–0.431
C_6	0.094	0.076	0.140	0.058	0.540	0.653	0.347	0.730
H_1	0.072	0.063	0.083	0.064	0.210	0.241	0.174	0.238
H_2	0.068	0.063	0.076	0.061	0.225	0.242	0.194	0.248
H_3	0.035	0.044	0.046	0.032	0.365	0.319	0.301	0.386
H_4	0.079	0.069	0.091	0.075	0.191	0.220	0.154	0.201
H_5	0.075	0.070	0.087	0.071	0.203	0.216	0.163	0.214
H_6	0.071	0.063	0.078	0.067	0.215	0.240	0.189	0.226
H_7	0.073	0.066	0.079	0.068	0.209	0.231	0.186	0.224
H_8	0.020	0.032	0.022	0.014	0.483	0.388	0.442	0.542
N_1	0.825	0.899	0.670	1.054	–0.496	–0.561	–0.353	–0.669
N_2	1.028	1.059	0.712	1.189	–0.631	–0.659	–0.357	–0.749
O_1	0.802	0.843	0.799	0.995	–0.431	–0.466	–0.439	–0.561
O_2	1.109	1.090	0.894	1.305	–0.630	–0.621	–0.493	–0.744
S	0.863	0.808	0.804	0.850	0.286	0.365	0.354	0.313
absolute MUE [c]		0.024	0.078	0.057		0.044	0.108	0.066
relative MUE [c]		0.117	0.233	0.139		0.134	0.468	0.253

[a] Second-order Møller-Plesset perturbation theory

[b] Restricted Hartree-Fock method

[c] Mean unsigned errors with respect to the corresponding MP2/6-31G(d) calculation