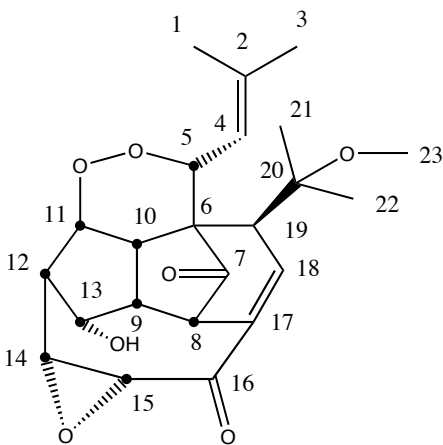
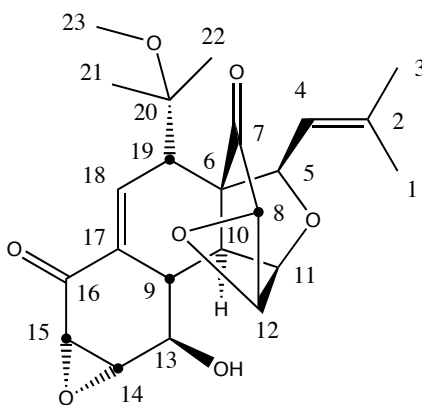


1. We begin by following up on some work from the first class problem set. In particular, in `~cm8021pr/templates` I have placed PCModel text output files for the lowest energy structures of both **A** and **B** appearing in problem 3 of Problem set 1 (named `A_pcm.txt` and `B_pcm.txt`). For each of these structures (without reoptimization of the geometry), compute the ^{13}C NMR spectrum using the WC04 density functional together with the 6-31G(d) basis set (see Wiitala et al. *J. Chem. Theory Comput.* **2006**, 2, 1085 for details of the WC04 functional). You will find shieldings computed at this level of theory for tetramethylsilane (TMS) in `~cm8021pr/templates`, which you may want to use in order to compute the proper deshielding values δ for comparison to experiment. To assist in this comparison, report your deshieldings for the carbon atoms as numbered in the figure on the next page.

**A****B****Experimental data**

Carbon	δ , ppm
1	18.6
2	142.2
3	26.1
4	120.7
5	75.8
6	60.5
7	202.9
8	53.1
9	40.4
10	47.8
11	71.5
12	54.5
13	72.7
14	61.0
15	53.2
16	192.8
17	132.5
18	139.6
19	40.9
20	77.3
21	26.6
22	24.7
23	49.1

The experimental data are provided in the table above. They were originally assigned to structure **A** (see Schlegel et al. *J. Antibiot.* **2002**, 55, 814), but is structure **B** plausible? In addition to providing an answer to this question, together with justification for your response, outline what additional steps you could take to increase your confidence in making a final assignment.

Computed data (from subtraction of computed shieldings from TMS ^{13}C shielding) are:

Carbon	Expt	A	B	A^a	B^a
1	18.6	22.5	23.0	24.8	27.8
2	142.2	128.5	134.0	134.7	143.6
3	26.1	29.9	29.2	32.5	34.3
4	120.7	124.7	118.4	130.8	127.3
5	75.8	79.3	62.3	83.7	68.8
6	60.5	41.7	45.5	44.7	51.3
7	202.9	184.9	192.6	193.2	204.7
8	53.1	48.5	37.9	51.8	43.4
9	40.4	37.2	33.7	40.1	39.0
10	47.8	55.3	41.5	58.8	47.1
11	71.5	77.9	64.3	82.3	70.9
12	54.5	38.0	46.3	40.9	52.1
13	72.7	67.4	67.8	71.4	74.5
14	61.0	43.2	49.3	46.3	55.2
15	53.2	45.2	39.8	48.4	45.3
16	192.8	186.8	175.8	195.2	187.1
17	132.5	115.6	119.9	121.3	128.9
18	139.6	146.1	138.8	153.0	148.6
19	40.9	44.9	32.0	48.0	37.2
20	77.3	62.6	62.8	66.4	69.3
21	26.6	27.9	31.2	30.4	36.4
22	24.7	30.7	31.6	33.3	36.8
23	49.1	49.3	49.1	52.6	55.0
MUE ^b		8.3	8.5	7.9	5.7
RMS ^c		10.1	9.8	9.1	6.6
m^d		1.0368	1.0427		
b^e		1.4868	3.835		
R^{2f}		0.9686	0.9831		

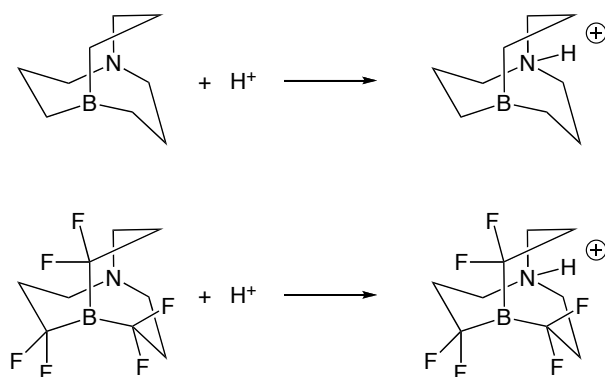
^a Corrected predictions from regression of raw computed data on experiment (computed as $m \times d + b$). ^b Mean unsigned error. ^c Root mean square error. ^d Slope from regression of computed data on experiment. ^e Intercept from regression of computed data on experiment. ^f Correlation from regression of computed data on experiment.

Note that the raw predictions give similar errors for **A** and **B**, but the errors for **B** are more systematic (as judged by the higher value of R^2). As the TMS value is uncertain (we

don't know where the geometry came from, after all), there is no reason to assume the raw computed shielding values are absolute. So, if we apply a linear correction (from regression), we see that the predictions for **A** are only rather slightly improved, but the predictions for **B** are *substantially* better, and moreover are much better than **A**, suggesting a mis-assignment of the original structure. For the full detective story, *and* an implication of rather serious scientific fraud (in a reported synthesis of **A**), you may be interested in reading Rychnovsky, S. D. *Org. Lett.* **2006**, *8*, 2895.

As for how we could improve the calculations, we could reoptimize the geometries at higher levels of theory, we could use multiple geometries (Boltzmann averaging their contributions to the total chemical shifts), we could include solvation effects (we'll deal with those later in the course), we could use a bigger basis set in the NMR calculations, and we could compare a variety of levels of theory to ensure no bias in our chosen protocol.

2. Consider the bicyclic borazane molecules on the next page, and in particular their basicities. Compute, at the levels indicated in the below tables, B–N bond lengths optimized at the indicated levels of theory (consider: are there possibilities for bond-stretch isomerism?) and proton affinities at the indicated levels of theory. In some detail, discuss your modeling strategy and results, addressing in particular the chemistry, but also considering how the semiempirical models compare to the more complete levels of wave function theory. *Nota bene:* sensible attention to job ordering and symmetry will help you to stay within the development queue time limits as you work on this problem. Imagine that this problem was motivated by a question from an experimental colleague, namely, “How does the basicity of the borazane compare to that of the hexafluoroborazane?”



B–N Bond Lengths (Å) in Borazanes

	Bicyclo[3.3.3]borazane		Hexafluorobicyclo[3.3.3]borazane	
	Unprotonated	Protonated	Unprotonated	Protonated
AM1	1.610 ^a	2.951	1.570	3.012
PM3	1.608 ^a	3.009	1.597 ^a	3.024
PM6	1.616 ^a	2.978	1.611	3.041
HF/3-21G	1.758	3.017	1.683	3.022
HF/6-31G(d)	1.767	3.025	1.694	3.062
MP2/6-31G(d)	1.719	3.009	1.664	3.064

^a Stable minima *without* a B–N bond (about 2.8 Å separation) are predicted at these levels of theory, but seem likely to be artifacts, as the MP2/6-31G(d) level does *not* predict such structures to be stationary.

Computed Energies (E_h) of Borazanes

	Bicyclo[3.3.3]borazane		Hexafluorobicyclo[3.3.3]borazane	
	Unprotonated	Protonated	Unprotonated	Protonated
AM1	-0.112 11	0.148 42	-0.550 13	-0.221 51
PM3	-0.109 26	0.188 43	-0.580 54	-0.215 21
PM6	-0.090 62	0.183 80	-0.609 27	-0.269 04
HF/3-21G	-428.094 03	-428.431 63	-1018.085 39	-1018.353 42
HF/6-31G(d)	-430.446 08	-430.794 62	-1023.597 19	-1023.887 83
MP2/6-31G(d)// HF/6-31G(d)	-431.878 76	-432.203 89	-1026.031 42	-1026.298 60
MP2/6-31G(d)	-431.881 13	-432.206 52	-1026.037 72	-1026.305 79

Computed Proton Affinities of Borazanes (kcal/mol)

	Proton Affinity		Fluoro effect
	Borazane	Fluoroborazane	
AM1	-203.7	-161.0	42.7
PM3	-180.4	-138.0	42.4
PM6	-195.0	-153.7	41.3
HF/3-21G	-211.8	-168.2	43.7
HF/6-31G(d)	-218.7	-182.4	36.3
MP2/6- 31G(d)// HF/6-31G(d)	-204.0	-167.7	36.4
MP2/6-31G(d)	-204.2	-168.2	36.0

Note that the semiempirical levels compute heats of formation, *not* electronic energies, so the proton affinities are computed as the protonated heat of formation, minus the neutral heat of formation, *minus the heat of formation of a bare proton* (an experimentally known quantity, 367.2 kcal/mol—not

computable as there are no electrons...) Of course, the same equation applies to the ab initio levels, but there the energy of a bare proton is defined to be zero.

If we assume that the MP2 level is the “best” in terms of accuracy, AM1 is quite good for borazane, but less good for the fluoro case. PM3 is very bad and PM6 is improved but not as good as AM1. However, all three semiempirical levels predict the fluoro effect to be about the same: about 42 kcal/mol. That’s 6 kcal/mol more than is computed at the MP2 level, which is as good as HF/3-21G (a more expensive model). If one is interested only in the fluoro effect, HF/6-31G(d) is as good as the best level at much lower cost. But, the absolute predictions at this level are in as great an error as the semiempirical levels. On the other hand, the MP2//HF prediction is good for both absolute proton affinities *and* the absolute affinities. Given the high cost of the MP2 optimization, this model is clearly the better choice.

As for geometries, the semiempirical models predict the B–N dative bonds to be quite a bit too short, the HF levels have them too long, and all models do fine with the protonated cases. Mind you, the error in the HF vs. the MP2 models does not have much *energetic* consequence, since the MP2//HF predictions of proton affinities are about as good as the MP2 optimized predictions.

3. Here begins a problem that will carry over to the third problem set and ultimately the final exam. Take a look at

pollux.chem.umn.edu/8021/PES/

Full credit for this problem is awarded for sensible data.