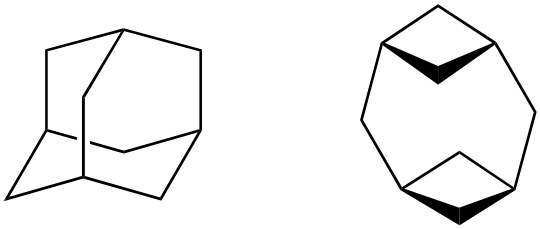
Chemistry 4021/8021 **Computational Chemistry** 3/4 Credits

Spring Semester 2007

(Key)

1. Below are two isomeric geometries that we previously examined in Problem Set 1 as both $C_{10}H_{16}$ and $Si_{10}H_{16}$. To what symmetry point group does each structure belong? Compute energies for all 4 cases for geometries optimized at the AM1, PM3, HF/MIDI!, and HF/6-31G(d). For the final set of geometries, compute single point energies at the MP2/6-311G(2df,p)//HF/6-31G(d) level (*do not optimize at this level*). Report the relative energies of the two isomeric forms at all levels of theory for the two elemental compositions. How have things changed compared to PCModel's predictions and how do your numbers influence your assessment of what may be the "right" answer?



1 **2**

Level of Theory	E(**1**)[a]	E(**2**)[a]	E(**2**) – E(**1**), kcal mol^{-1}
Carbon frame			
AM1	–43.2	18.0	61.2
PM3	–34.6	6.3	40.9
HF/MIDI!	–385.688 27	–385.572 59	72.6
HF/6-31G(d)	–388.026 48	–387.928 14	61.7
MP2/6-311G(2df,p)// HF/6-31G(d)	–389.781 69	–389.680 53	63.5
Silicon frame			
AM1	70.4	91.2	20.8
PM3	6.7	16.8	10.1
HF/MIDI!	–2883.725 42	–2883.667 12	36.6
HF/6-31G(d)	–2898.529 99	–2898.480 58	31.0
MP2/6-311G(2df,p)// HF/6-31G(d)	–2899.944 92	–2899.894 31	31.8

[a] Units are kcal mol^{-1} for semiempirical levels of theory and E_h for ab initio levels of theory.

A few observations:

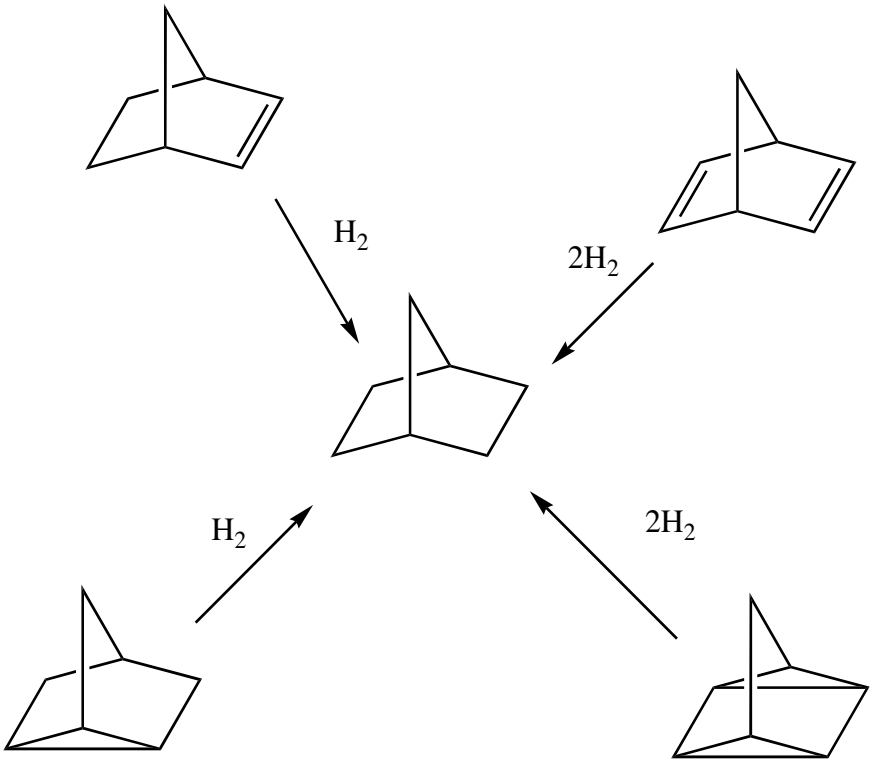
*(1) In the carbon case, AM1, HF/6-31G(d), and MP2/6-311G(2df,p)//HF/6-31G(d) all agree reasonably well with one another that **1** is favored over **2** by about 63 kcal/mol. The small change on going from HF/6-31G(d) to MP2/6-311G(2df,p)//HF/6-31G(d) suggests that there is limited sensitivity to basis set size or correlation. Hence, this number seems fairly solid. MMX predicts 57.8, which is not bad (within about 10% of MP2). PM3 does surprisingly poorly here, perhaps overstabilizing 4-membered rings or near H–H contacts in **2**.*

*(2) In the silicon case, HF/6-31G(d) and MP2/6-311G(2df,p)//HF/6-31G(d) agree quite well with one another that **1** is favored over **2** by about 31 kcal/mol. AM1 is now in only fair agreement with the more complete levels of electronic structure theory, and PM3 continues to do poorly. The MMX force field prediction from the last problem set of 57 kcal/mol is also quite poor, indicating that the parameters for Si are evidently much less good than those for C.*

*(3) From a chemistry standpoint, **2** is closer to **1** in stability for Si compared to C, and this is consistent with second-row atoms preferring to use more p character than first-row atoms in their bonding hybrid orbitals, thus stabilizing smaller bond angles (like those in the 4-membered rings of **2**).*

(4) A more thorough comparison would require us to look at structures, and possible changes in both structure and energy were we to optimize at correlated ab initio levels of theory, and additional correlation effects beyond MP2.

2.



Enthalpies of Formation of the Norbornadiene Cycle

Compound	MMX	MNDO	AM1	PM3	Expt.
Norbornadiene	55.5	62.7	67.7	58.8	57.4
Norbornene	19.5	25.3	26.0	22.0	21.4
Norbornane	-12.8	-10.4	-14.4	-13.7	-12.4
Nortricyclane	19.5	27.1	33.8	26.0	20.2
Quadricyclane	79.4	79.1	104.4	86.3	79.5

*units are in kcal/mol

Rogers, D.W. et al., 1992, *Structural Chemistry*, **3(1), 53.

Looking at the above table of heats of formation, does the accuracy of any of the methods surprise you? Is it possible to rationalize the difference in accuracy of the molecular mechanics method and the semi-empirical methods?

The molecular mechanics force field does a surprisingly good job at calculating the various heats of formation. One suspects that many of the subject molecules are in the training set, and/or that the atom types are fairly specific to the bicyclic systems. The semi-empirical methods have an accuracy ordering of roughly AM1 < MNDO < PM3 (somewhat unexpected since AM1 usually outperforms MNDO). The errors in the semiempirical methods are slightly higher than what one might like, but are not bad given their enormous speed.

Now construct a Z-matrix for norbornadiene (of the windshield wiper mechanism fame) and include it in your answer. Make sure to check the structure in Chem3D to see if it is reasonable.

Many correct answers are possible.

Compute the $\Delta H_{f,298}$ for norbornadiene (using its atomization energy and Table 10.2 in the text) at the B3LYP, HF, and MP2 levels with both the 6-31G and 6-311G(2df,p) basis sets. Also compute $\Delta H_{f,298}$ using the G3 method (there is a G3 keyword in Gaussian 03). For any one of the methods, demonstrate how you did one of the calculations of the heat of formation.

Method	Basis Set	Norbornadiene ΔH_f(298K)
B3LYP	6-31G	108.4
	6-311G(2df,p)	77.9
HF	6-31G	545.6
	6-311G(2df,p)	489.3
MP2	6-31G	250.2
	6-311G(2df,p)	73.5
G3		58.0

$$\Delta H_f(298K)\ of\ \ norbornadiene =$$
$$\Delta H_f(G3,298K)norbornadiene$$
$$-8\Delta H_f(G3,298K)H - 7\Delta H_f(G3,298K)C$$
$$+8\Delta H_f(\text{expt},298K)H + 7\Delta H_f(\text{expt},298K)C$$

Which of the higher level methods does best compared to experiment? Which does the worst? How sensitive are the methods to basis set incompleteness? How does the G3 composite method perform compared to the single level methods? Why do the single level methods do so poorly compared to molecular mechanics and the semi-empirical methods (which are vastly cheaper), and G3 (the composite method)?

The G3 method does by far the best, then MP2 followed closely by B3LYP, and then far worse is Hartree-Fock. For HF and MP2, there are huge basis-set effects. The heat of formation drops by ~55 kcal/mol for HF and ~125 kcal/mol for MP2. B3LYP drops by 30 kcal/mol. The G3 method is within 1 kcal/mol of the experimental heat of formation, compared to ~15 kcal/mol and ~20 kcal/mol for MP2 and B3LYP with the largest basis set. G3's accuracy is consistent with the method's focus on the accurate computation of this property.

Single level methods are known to do poorly in calculating heats of formation mostly due to their inability to adequately capture correlation energy. Molecules have vastly more correlation energy compared to their constituent atoms (since they have many more electrons than do each of the individual atoms), so if the model cannot accurately compute the correlation energy, this leads to large errors. Composite methods are designed to correct for incompleteness in the correlation energy and basis set size. They also sometimes have empirical terms to improve accuracy. As mentioned above, MM and SE models are parameterized for heats of formation, so their accuracy is founded on their training sets.

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/C4H6SO2/`

Full credit for sensible data.