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

Spring Semester 2014

(Due 2 / 26 / 14)

Using PC Model, answer the questions below. If you have questions/issues working on this Problem Set, do please consider using Piazza to address them.

1. What are the parameters for the force constant (mdyne/Å) and equilibrium bond length (Å) for the bond between a carbonyl carbon and a carbonyl oxygen in each of the MM3, MMX, and MMFF94 force fields? What is the parameter for the equilibrium bond length between two sp3 carbon atoms in the MMX force field? If you were to pick a "canonical" value for a C–C single bond between two sp^3 carbon atoms, what would it be to the nearest hundredth of an angstrom (explain how you made your choice)? How does that compare to the MMX parameter? Run a geometry optimization of ethane with the MMX force field. Is the optimized C–C bond length equal to the MMX equilibrium bond length parameter? If not, explain why not.

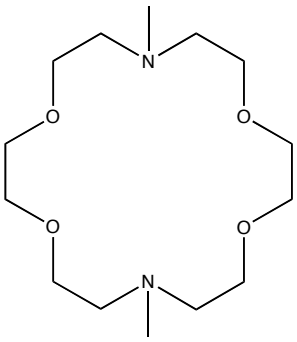
   C=O parameters: MM3, 10.1 mdyne/Å, 1.208 Å; MMX, 10.8 mdyne/Å, 1.208 Å; MMFF94, 12.95 mdyne/Å, 1.222 Å.

   MMX C–C equilibrium bond length: 1.523 Å. The usual value in textbooks for a C–C single bond between two sp^3 carbon atoms is 1.54 Å, which is longer. However, the geometry optimization with MMX leads to a value of 1.532 Å. The reason it is longer than the parameter is that there are other strain contributors that must be minimized other than just bond stretch (primarily the repulsive non-bonded van der Waals interactions between the H atoms on the two different methyl groups, which contribute 0.679 kcal/mol worth to the total strain of 0.816 kcal/mol).

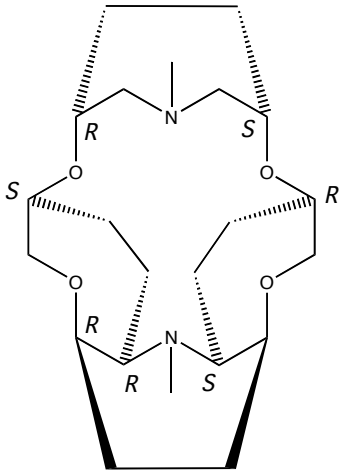
2. Oh no! Someone spilled a solution of plutonium salts onto the Chemistry Department's softball trophy from that famous 1934 championship that saw Mathematics and Chemistry go 73 scoreless innings over the course of four days before I. M.

Kolthoff himself powered one over the left field fence, shattering the windshield of the Dean's Cadillac V16 Aerodynamic Coupe.

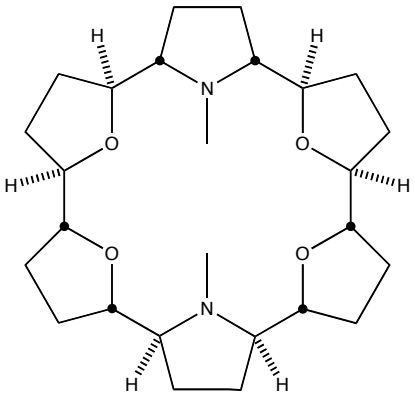
You probably know that plutonium (Pu) is one of the deadliest substances on Earth in addition to being highly radioactive. But, we can't throw away the trophy! The only remedy will be to soak the wooden base, into which the Pu has leached, in a solution containing a sequestering agent that will extract the Pu from the wood until the remaining amount of Pu in the wooden base falls below femtogram levels. The organic and inorganic chemists have huddled together, and propose the 3 molecules shown on the next page, each of which can be synthesized for the indicated cost per gram. Given that this money has to come out of the seminar donut fund, they want to choose the most cost effective option. Which molecule should they make? Justify your answer on the basis of molecular mechanics calculations, explaining precisely and in detail what you did in order to come to some conclusion. (As it is mildly tricky to interpret stereochemistry for the final structure, absolute assignments at each position are provided for completeness.) For purposes of this exercise, let's assume the speciation of Pu is entirely high-spin Pu(IV).



$1.00



$1000.00



$10.00

This is by no means a trivial problem — many reasonable answers could be offered. Here's what I did:

First, the issue of synthetic cost is balanced by binding efficiency. If the third macrocycle costs 100 times the second, but binds 10,000 times stronger, it will be more cost effective to synthesize the more expensive one because it will require much less of it to leach the Pu to an acceptable level. Put more chemically, we care about the equilibrium

$$\text{trophy} \bullet \text{Pu} + \text{macrocycle} \Leftrightarrow \text{trophy} + \text{macrocycle}\bullet\text{Pu}$$

which has an associated equilibrium constant

$$K = \frac{[\text{trophy}][\text{macrocycle}\bullet\text{Pu}]}{[\text{trophy}\bullet\text{Pu}][\text{macrocycle}]}$$

where we want the ratio of trophy to trophy•Pu to exceed 10^{15}, which means we will want to maximize the binding of Pu to macrocycle. We can

see this by expressing the above equilibrium as the sum of two others, namely

$$\text{trophy} \bullet \text{Pu} \Leftrightarrow \text{trophy} + \text{Pu}$$
$$\text{Pu} + \text{macrocycle} \Leftrightarrow \text{macrocycle} \bullet \text{Pu}$$

the first is a constant about which we can do nothing, but the second is what we have some hope of adjusting. So, how can we compute the free energy of binding? In practice, that's a lot of work, because free energy will require a careful sampling over phase space. But, a quick estimate can be had from replacing free energy with potential energy, and further assuming that we can deal just with lowest energy isomers. And, we can recognize that we aren't trying to get an accurate *absolute* binding energy, we just want a *relative* binding energy.

So, what controls binding energy? Well, let's assume that all 3 crowns will complex the Pu atom equally well once they adopt a nice geometry (see figure on next page for the second crown), so then the question becomes, how much energy must they lose in order to *adopt* that optimal structure compared to their lowest energy *uncomplexed* structure?

To estimate that, I chose the MMX force field. I generated the best structure I could for the simplest crown with a metal, Pu, chosen to be *bonded* (not metal coordinated) to the six heteroatoms. I chose bonded because the choice of "metal coordination" led to only 4 heteroatoms being less than 3 Å from the metal, and I did not consider this reasonable. I might be making a mistake, but, hey, the trophy is glowing and I needed to make some progress, so there's no point sitting around existentially paralyzed about the whole thing.

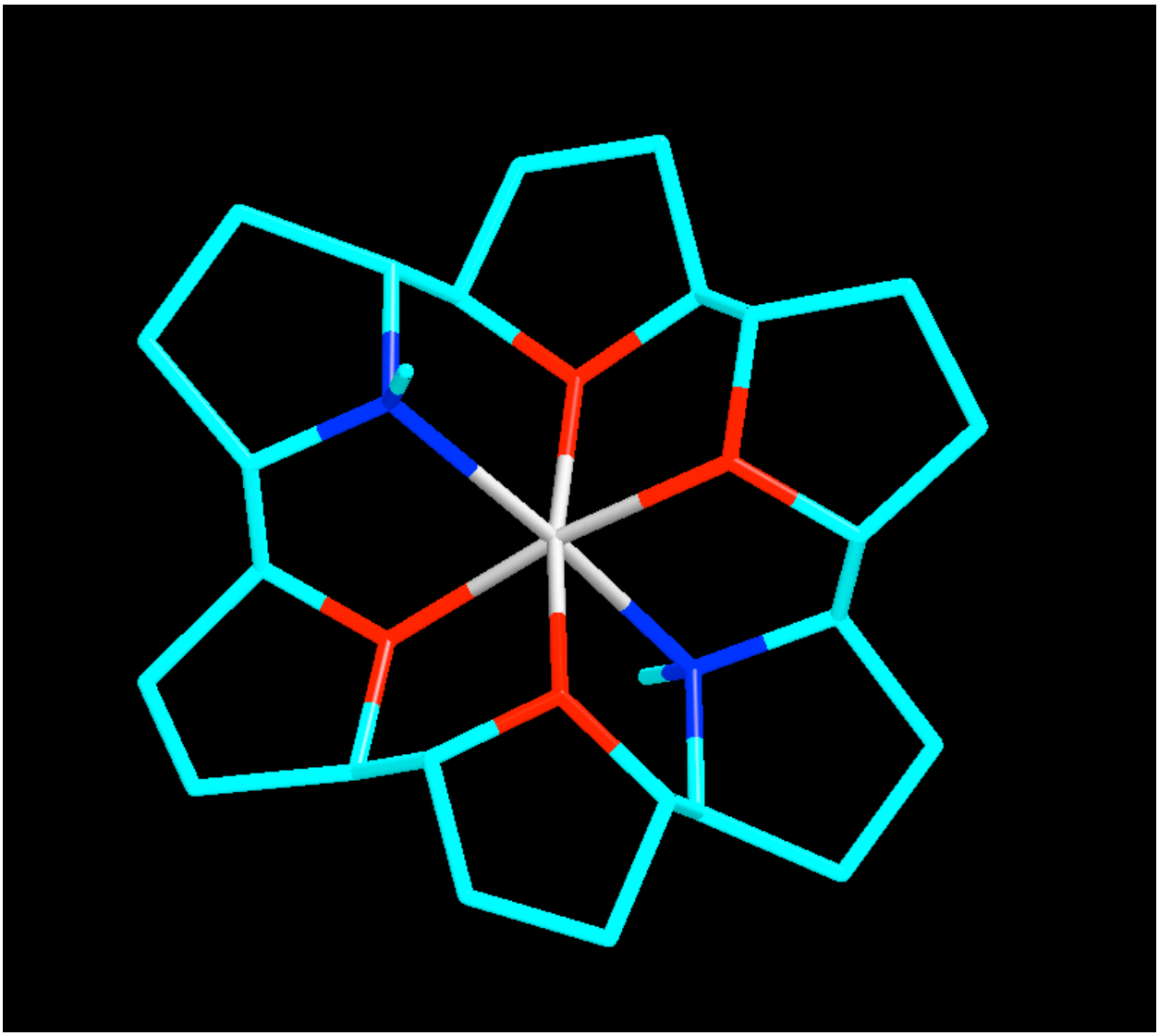


Figure. Tube representation of optimal crown•Pu complex for second macrocycle; hydrogen atoms have been removed for clarity. Carbon, nitrogen, oxygen, and plutonium atoms are cyan, blue, red, and gray, respectively.

Based on this geometry, I then generated analogous geometries for the other two crowns. In each case, I then deleted the Pu atom and computed single point energies for the crowns. Then, I minimized them and looked at how much the energy dropped; that is, what did it cost them to adopt that geometry. To be thorough, I then did GMMX searches for the *global* minimum, taking 1000 steps in each case. As a technical point, I did all of this *without* lone pairs in the MMX calculations. However, for two cases, I checked whether including lone pairs made a difference, and they did not

seem to affect my energy differences, so this point may have been moot. My results are shown in a table on the next page.

Table. Steric energies (kcal/mol) for different crowns.

Cost ($)	Frozen[a]	Relaxed[b]	Fully relaxed[c]	Distortion cost[d]
1.	85.6	44.4	35.0	50.6
10.	123.0	95.1	89.2	33.8
1,000.	135.6	104.3	97.1	38.5

[a] Computed for empty crown at the frozen geometry of the optimized crown•Pu complex. [b] Energy for structure that derives from direct geometry optimization of the frozen, uncomplexed crown. [c] Lowest energy structure found from 1,000 steps of GMMX optimization. [d] Computed as difference between columns 2 and 4.

As can be seen from the table, the $10 crown has the lowest distortion cost to form the proper pocket to complex Pu. As that cost is much more than 1.4 kcal/mol better than that for the $1 crown, the improved binding is much better than the factor of 10 difference in the cost. So, the chemists should synthesize the $10 crown.

There are many other things that could be discussed with respect to this problem, and there are certainly other reasonable approaches that could be taken -- credit will be awarded based on logical thoughts/procedures, not based on absolute similarity to the solution proposed here.

3. Consider the interaction between two molecules of pyrazine (i.e., the pyrazine dimer). What geometries do pyrazine dimers adopt and what are their associated complexation energies as calculated with the MMX, MM3, and MMFF94 force fields? Compare/contrast the different force-field results. From the components that contribute to your results, what can you say about the nature of the intermolecular interactions in this dimer (i.e., what type of interaction(s) is/are dominant?)

   Now, assume that you would like to create a force field *specific* to the pyrazine dimer. And, assume that you have access to supercomputing resources that allow you to compute "exact" interaction energies for any geometry, but the timeline of the

project limits you to the calculation of no more than 100 single-point energies. The ultimate goal of the project is to design a force field that will approximate the exact interaction energies in a coarse-grained fashion; that is, think about how you might reduce the number of variables from 54 ($3N - 6$ for two pyrazine molecules) to a much smaller number, and then how would you sample over those variables and come up with an overall function fitting the relevant energies. How would you choose your 100 points?

Again, there are many possible ways to approach this problem, but a first step is recognizing the complexation energy to be:

$$E_{\text{complex}} = E_{\text{dimer}} - 2 * \left(E_{\text{monomer}} \right)$$

where E_{dimer} and E_{monomer} are the optimized pyrazine dimer and optimized pyrazine molecule at their respective geometries (that is, intramolecular distortion is taken into account).

With the three force fields tested, the monomer geometry is unambiguous, and the computed force field energies are reported in Table 1. The dimer energy, however, is dependent on the geometric orientation, for which there are a number of local minima. One could describe these dimers as planar hydrogen bonded (C-H—N), Pi stacked, or some other orientation (most commonly found as something out of plane or T-Shaped). The lowest energy dimer geometries are pictured in Figure 1, and their force field energies are reported in Table 2. The complexation energies corresponding to the various orientations are given in Table 3, as computed by the E_{complex} equation above.

Table 1: Monomer computed force field energies in kcal/mol.

	MMX	MM3	MMFF94
Monomer E	29.7	18.2	52.5

Table 2: Dimer computed force field energies in kcal/mol. The (a,b,c) indices refer to Figure 1.

Dimer E	MMX	MM3	MMFF94
Planar (CH—N)	59.4 (b)		101.7 (c)
Pi Stacked	56.4 (a)	32.5 (a) / 32.6 (b)	100.2 (a)
Out of Plane	60.0 (c)		101.1 (b)

Table 3: Interaction energies in kcal/mol

Interaction E	MMX	MM3	MMFF94
Planar (C-H—N)	0.0		-3.3
Pi Stacked	-3.0	-3.9/-3.8	-4.8
Out of Plane	+0.6		-3.9

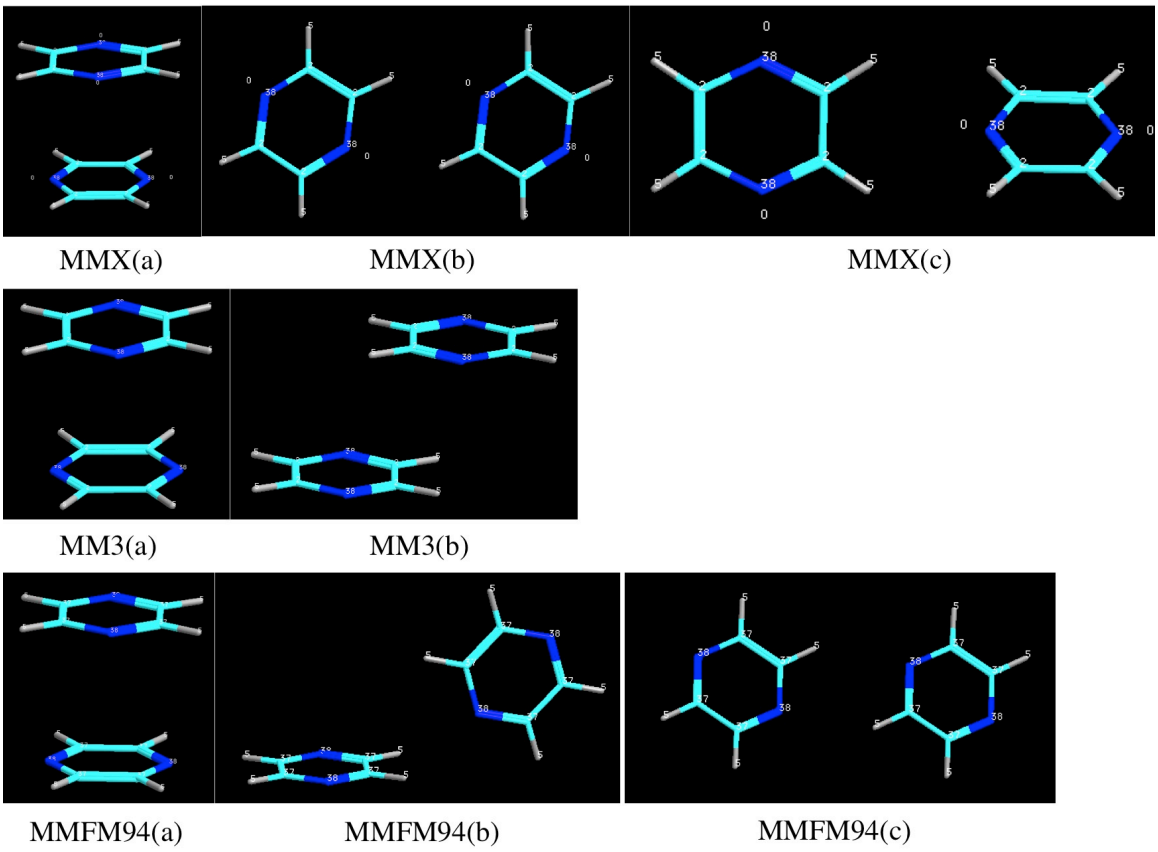


Figure 1: Dimer geometries found using the GMMX option in PC Model with the 3 force fields. Relative energetics associated with these geometries are given in Table 2.

All of the force fields find a pi-stacked pyrazine dimer as the lowest in energy. The magnitude of the interaction energy for this structure is quantitatively similar for all of the investigated force fields. However, while the other force fields find other dimer minima relatively close in energy, MM3 finds *only* pi-stacked minima. Additionally, MM3 is the only force field that finds a "slipped" pi-stacked dimer geometry as a minimum (albeit a slightly higher energy minimum), while the MMX and MMFF94 force

fields do not find the “slipped” pi-stacked minima, but rather a ‘sandwich-type’ stacked pi-stacked minimum.

These results suggest that the three force fields investigated in this work all take into account aromatic pi-pi interactions in a favorable way, while only some of the force fields take into account “Hydrogen-bonded” type C-H—N interactions in such a favorable way. In all cases, the force fields predict that aromatic pi-pi interactions in the pyrazine dimer yield more favorable energetic minima as compared to the hydrogen bonded or out-of-plane alternative minima.

So in general, one could say that aromatic pi-pi “dispersion-type” interactions are taken into account by the force fields employed here, and in this particular dimer, those interactions dominate over any hydrogen bonded C-H—N type of interactions.

Now, let us move on to the coarse graining and creation of a force field specific for this dimer.

I will start with the assumption that it is reasonable to approximate intermolecular interaction energies without the need for explicitly taking into account geometric distortions due to polarization of one pyrazine molecule due to the presence of the second pyrazine molecule (that is, I will only calculate intermolecular interaction energies at monomer optimized rigid geometries).

Considering the symmetry of the pyrazine molecule, I can decide to treat my atoms as ‘atom-types’ rather than atoms, in order to reduce the computational complexity of considering each C atom independently or, for example, treating all C’s as equivalent since they are equivalent by symmetry. Fortunately, I can say that all of my N atoms in a given pyrazine molecule are equivalent by symmetry, and the same is true for all of the C and H atoms of a given pyrazine molecule.

Considering the initial screening using the three force fields above, I chose to distribute my 100 points along potential energy coordinates corresponding to the minima shown above in Figure 1. If I would like to have these parameters be compatible with, let’s say, later parameterization or with an existing force field, I could choose to use the Lorentz-Berthelot combining rules with a standard Lennard-Jones 12-6 potential for non-bonded interactions between atoms in different rings, which means that I would have to fit, in total, 6 parameters (sigma and epsilon of each H, C, and N).

An alternative coarse-grained model might involve a focus on seven coordinates: the distance between the centers of mass of any two rigid pyrazine molecules, and the three rotation angles about each center of mass needed to define fully the orientation of an individual pyridine. The energy functions of this one distance and six angles would need to be fit to reasonable functional forms (e.g., a Lennard-Jones form for the distance and some suitable periodic form for the angles). The 100 points could be chosen by random displacements of these coordinates from the various equilibrium structures (since one wants the force field to be most accurate for the lowest energy structures, these being the most relevant at low temperatures).