

**Physical Chemistry Final Take Home
Fall 2003**

Do *one* of the following questions. These projects are worth 30 points (i.e. equivalent to about two problems on the final). Each of the computational problems below took me less than 20 minutes to do. Under no circumstances should your report be more than one written page (and perhaps a scatter plot). Answer the specific questions. Save all your molecule structures. If you use an SGI, use your own account. Give your SGI account name if you used an SGI or the PC number if you used a PC and give the names of all the files that you generated. You don't need to type your report; hand written is fine. Your reports are due at the **beginning** of the final.

Problem 1 uses ISIS/Base pharmacophore searches for ACE inhibitors to explore the importance of several options within the cfs algorithm.

Problem 2 uses MOE Gibbs Free energy of solvation descriptors in a QSAR study of anticancer compounds.

Problem 3 uses MOE and electrostatics descriptors in a QSAR study of anticancer compounds.

Problem 4 uses MOE and continuum solvation calculations to look at solvent effects on tautomeric equilibria.

You can use any written source, as long as that source does not involve the active efforts of any other individual, other than yourself. For example, you can use the tutorials, your book, your notes, and references in the library, but you cannot use e-mail, fax, letter, teleconference, voice, or video communications. You may of course ask questions of me (x3315, 872-2956), but I may not be available Wednesday night. If you wish to contact me before you start, so I can check in for questions- I'll be happy to do so. I will be more than happy to help you with software related issues, for example how to add or delete descriptors, use the Builder, or switching between ISIS/Base forms. See the note at the end for help with Printing issues from MOE.

1. Use the pharmacophore from Figure 8 in the Henry, Güner paper to do five cfs searches that explore the importance of the options in the cfs procedure. These searches will be just like the ones we did in class except for changing the options in the cfs search (as listed below). These options concern the molecular mechanics portion of the cfs procedure. This pharmacophore is already built and is available in the "D:/ISIS/ace" folder, so you don't need to build it. Just open the q4.skf file in ISIS/Draw. Use the CMC3D database (password "isis/3d"). Remember that 3D queries need to be pasted into the Model box from the Model tab of the search form. (Pasting your query in the Structure box will give an error for a cfs search).

In evaluating the hits from a pharmacophore search it is important to determine how many of the known active compounds are recovered from the database along with the new compounds. There are 362 compounds in the database that are listed as "antihypertensive." Of course not all of these are ACE inhibitors, but the recovery rate of these compounds will none-the-less help us to evaluate the cfs procedure. To determine the number of "antihypertensive" compounds in your hit list do the following:

1. After your cfs search is complete, pull down the Search menu, slide right on Set Search Domain to..., and choose Current List.

2. Switch to the Query mode by clicking on the Query button in the top button bar. Switch to the Structure form by clicking on the Structure tab on the "index card deck."

3. In the Class query box enter the text:

like "%antihyper%"

- Pull down the Search menu and choose By Form. Record the number of hits.
- Remember to return the search to the full database by pulling down the Search menu, slide right on Set Search Domain to..., and choose Entire Database.
- Remember to do Step 5 above: choose Entire Database.

Start with the cfs options set to the values we used in class. These options were chosen to ensure rapid search times:

Fitter selection: Automatic Simplex Derivative

Fitting at search time Max. no. of attempted fittings

Ring flexing Min. ring size Max. ring size

Rotate single bonds between unsat. atoms Global relaxation

Bump checking at: Search time View time

	Torsional degrees of freedom[TDF]	RMS deviation	van der Waals energy difference
Calculate for search/sort	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Cutoff for search/view	<input type="text" value="6"/>	<input type="text" value="0.7"/>	<input type="text" value="5.0"/>
Suggested value	10 rotatable bonds	0.7 RMS units	5.0 Kcals

Record the results in the table below. Use the following sets of cfs options listed in the table. You may get a couple errors along the way. These won't have an effect on the final outcome.

Max. no. of attempted fittings	Global relaxation	TDF	Van der Waals energy difference	Total hits	Antihypertensives recovered
5	off	6	5.0		
5	✓	6	5.0		
5	off	6	10.0		
10	off	6	5.0		
5	off	10	5.0		

Answer the following questions:

- Which change gave the best improvement in the recovery of known antihypertensives?
- Which change gave the best improvement in the totals number of hits?
- The global relaxation description in the ISIS documentation explains:

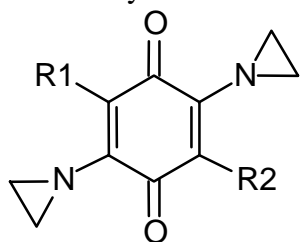
You can specify that ISIS rotate single bonds in other parts of the structure to reduce nonbonded interactions throughout the molecule. When global relaxation is active during bump checking, ISIS finds structures that might be eliminated if ISIS rotated single bonds

solely in the region that matched the query. Global relaxation increases the size of your hit list and produces more realistic 3D structures.

Global relaxation does a full molecular mechanics minimization at the chosen conformation. Does this rather time consuming step make a big difference in the hit rate and the recovery of known antihypertensives?

4. The maximum number of attempted fittings and the Torsional degrees of freedom options concern mostly statistical issues rather than chemical issues. The global relaxation and Van der Waals energy difference options are more "physical chemical" issues that concern strain energy and the thermodynamic accessibility of the strained conformations that fit the query. Are statistical or thermodynamic issues more important for good cfs searches as implemented in this algorithm?

2. A group of quinones (2,5-bis(1-aziridinyl)-p-benzoquinones) has been studied for their antitumor activity. You will find a selection of these in the benzo/ directory and the database file "benzo.mdb" in your SGI account or in D:/moe/ on the PC's. You will do a QSAR study similar to the study we did in lab. The base structure is:



See S. P. Gupta, "Quantitative Structure-Activity Relationship Studies on Anticancer Drugs," *Chem. Rev.* **1994**, *94*, 1517-1518 and Table 7, for more information if you would like (but not required)(See me for a copy).

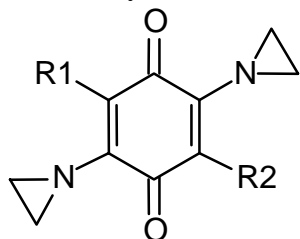
Use the instructions for the ACE QSAR study in the CAMD tutorial as the basis for this assignment. However, this time we want to do a better job of considering electrostatic effects. This database includes F_{H_2O} and F_{Oct} already entered in the descriptor list (see Problem 3 below). Remove these descriptors from the database. They won't be necessary for this exercise. For this exercise use the PEOE based electrostatic descriptors in the table below. These descriptors attempt to determine the effect of the electrostatic distribution in the molecule on the activity of the compounds in a more precise way than possible with "bulk" descriptors like the dipole moment.

PEOE_VSA+6	Sum of v_i where q_i is greater than 0.3.
PEOE_VSA+5	Sum of v_i where q_i is in the range [0.25,0.30).
PEOE_VSA+4	Sum of v_i where q_i is in the range [0.20,0.25).
PEOE_VSA+3	Sum of v_i where q_i is in the range [0.15,0.20).
PEOE_VSA+2	Sum of v_i where q_i is in the range [0.10,0.15).
PEOE_VSA+1	Sum of v_i where q_i is in the range [0.05,0.10).
PEOE_VSA+0	Sum of v_i where q_i is in the range [0.00,0.05).
PEOE_VSA-0	Sum of v_i where q_i is in the range [-0.05,0.00).
PEOE_VSA-1	Sum of v_i where q_i is in the range [-0.10,-0.05).
PEOE_VSA-2	Sum of v_i where q_i is in the range [-0.15,-0.10).
PEOE_VSA-3	Sum of v_i where q_i is in the range [-0.20,-0.15).
PEOE_VSA-4	Sum of v_i where q_i is in the range [-0.25,-0.20).
PEOE_VSA-5	Sum of v_i where q_i is in the range [-0.30,-0.25).
PEOE_VSA-6	Sum of v_i where q_i is less than -0.30.

Remember to follow the guidelines concerning descriptor correlations and the maximum number of descriptors. Find a small set of descriptors that give a reasonable QSAR. Then add $\log P(o/w)$ and redo the QSAR to determine if solvation effects that are not correlated with the electrostatic distribution are important. Answer the following questions.

1. Give your best QSAR equation using the PEOE_VSA descriptors. Include the R^2 and R values and the RELATIVE IMPORTANCE for each of the descriptors.
2. Give your best QSAR equation using the same descriptors as in question 1 but including $\log P(o/w)$. Include the R^2 and R values and the RELATIVE IMPORTANCE for each of the descriptors.
3. Give the R^2 and R values for the QSAR using $\log P(o/w)$ alone.
4. Does inclusion of the PEOE_VSA descriptors markedly improve the QSAR over the classical $\log P(o/w)$ descriptor alone?
5. Does the inclusion of the $\log P(o/w)$ descriptor markedly improve the QSAR using only the PEOE_VSA descriptors? What does this imply about hydrophobic effects? (Note that PEOE_VSA+0 and PEOE_VSA-0 should also correlate with hydrophobic effects to some extent).
6. Give the name of the computer and the database file that you generated.

3. A group of quinones (2,5-bis(1-aziridinyl)-p-benzoquinones) has been studied for their antitumor activity. You will find a selection of these in the benzo/ directory and the database file "benzo.mdb" in your SGI account or in D:/moe/ on the PC's. You will do a QSAR study similar to the study we did in lab. The base structure is:



See S. P. Gupta, "Quantitative Structure-Activity Relationship Studies on Anticancer Drugs," *Chem. Rev.* **1994**, *94*, 1517-1518 and Table 7, for more information if you would like (but not required)(See me for a copy).

Use the instructions for the ACE QSAR study in the CAMD tutorial as the basis for this assignment. However, this time we want to do a better job of considering solvation effects. This database includes F_{H_2O} and F_{Oct} already entered in the descriptor list (see the CAMD Introduction). Make sure to use E_{solv} in the new descriptors that you add. E_{solv} uses the GB/SA solvation calculation for its values, so make sure to set up the MMFF94 force field with solvation in the same way as we did for the molecular mechanics solvation Gibbs Free Energy exercises (don't use the distance dependent dielectric).

F_{H_2O} and E_{solv} should give the same values. They are both the Gibbs Free Energy of solvation in water. However, the F_{H_2O} values are derived from empirical correlations with experimental data extracted into group additive substituent constants. And, of course the GB/SA calculations are based on the first principles approach using the Born Approximation.

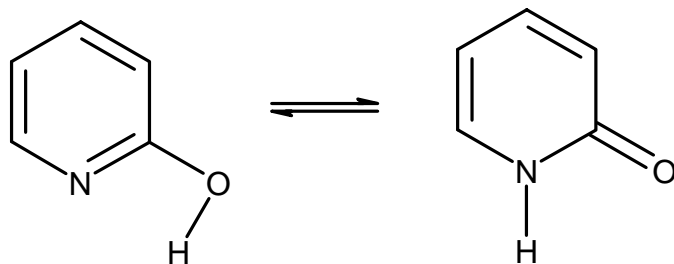
Remember to follow the guidelines concerning descriptor correlations and the maximum number of descriptors. This dataset is fun since many different combinations of descriptors do a good job. So don't spend a lot of time trying to find the "perfect" QSAR equation. A reasonable QSAR that follows the general guidelines is fine. However, do consider that F_{H_2O} , F_{Oct} , E_{solv} ,

and $\log P(o/w)$ are all closely related. It is best not to use more than two of this set in your final QSAR to avoid distortion of the final QSAR results (even though they are not necessarily all strongly correlated).

Answer the following questions:

1. How well correlated are F_{H_2O} and E_{solv} ? Since both are rough approximations, it is not easy to decide which will give the best approximation to the experimental solvation free energy.
2. Give your best QSAR equation. Include the R^2 and R values and the RELATIVE IMPORTANCE for each of the descriptors.
3. Give your best QSAR equation using the same descriptors as in question 2 but without F_{H_2O} , F_{oct} , or E_{solv} if you used them in question 2. Include the R^2 and R values and the RELATIVE IMPORTANCE for each of the descriptors.
4. Do inclusion of F_{H_2O} , F_{oct} , or E_{solv} markedly improve the QSAR over the classical descriptors (e. g. $\log P(o/w)$, mr , vdw_vol , ASA , $apol$, $dipole$)?
5. Give the name of name of the computer and the database file that you generated.

4. Solvent effects can have a large influence on tautomeric equilibria. Polar and nonpolar molecules are differentially stabilized in polar solvents.¹ Consider the tautomeric equilibrium in the system:



2-Hydroxypyridine is favored in the gas phase, but the pyridone is favored in aqueous solution. The molecular mechanics steric energies of these two molecules can't be directly compared, since they don't have the same number and types of chemical bonds. However, we can compare the GB/SA solvation energies. The experimental results require that the solvation energy of the pyridone should be more favorable (larger negative) than for 2-hydroxypyridine. Do the GB/SA solvation energies calculated using MOE reproduce the experimental trends? Use the procedure that we used in lab from the Molecular Mechanics tutorial to determine the solvation energy for these two compounds. Remember to consider conformational isomerism in determining the global minimum energy. In addition, why is the solvation energy of the pyridone more favorable? Create a database containing these two compounds and use the QuaSar-Descriptors application to calculate the dipole moment, ASA, and $\log P(o/w)$ for these two compounds. Use the following instructions to create the database:

1. With one of the two molecules in the MOE main window, pull down the File menu, slide right on New and choose Database. In the file librarian dialog box enter the database file name as "D:/moe/XXXpyridine.mdb", where XXX are your initials.
2. In the Database Viewer, pull down the Field menu and choose Create Field.

3. In the dialog bar at the top of spreadsheet window pull down the Create Field type menu and choose molecule. In the adjacent Name dialog box enter "mol". Press enter. A column label "mol" should appear.

4. Pull down the Entry menu and choose Add Entry... In the New Entry window click OK. The current molecule in the MOE window should be transferred into the database.

5. Close the current molecule in the MOE main window. Build or Load the file for the other of the two molecules. In the Database Viewer, once again pull down the Entry menu and choose Add Entry... In the New Entry window click OK. The current molecule in the MOE window should be transferred into the database as the second molecule.

6. You can now add your descriptors by pulling down the Compute menu and choosing Descriptors.

Answer the following questions:

1. Which molecule is better stabilized in aqueous solvent? Do the GB/SA solvation energies calculated using MOE reproduce the experimental trends?
2. Using the dipole moment and the ASA values, decide whether the charge distribution or the solvent accessible surface plays a bigger role in determining the relative solvation energy of the two compounds. Remember the surface tension term used in the GB/SA approach is $8.72 \text{ cal}/\text{\AA}^2$.
3. Given the answer to questions 1 and 2, why is this molecule better stabilized by aqueous solvation?
4. Does the logP(o/w) value correlate as expected with the solvation energy and the dipole moment?
5. Give the name of the computer and the database file that you generated.

Printing Plots from MOE on a PC

Use the Start menu to run the Paint program. Make the window that you want to print the active window. Press Alt-Print Screen. Switch to the Paint program, pull down the Edit menu and choose Paste. For scatter plots, the background will be black. To switch to a white background, pull down the Image menu and choose Invert Colors. Print the graphics from the Paint program as you would normally.

Literature Cited:

1. W. J. Hehre, L. D. Burke, A. J. Shusterman, W. J. Pietro, Experiments in Computational Organic Chemistry, Wavenfunction, Inc., Irvine, CA, 1993. Experiment 6: Substituent and Solvent Effects on Tautomeric Equilibria.