

**CHEMISTRY 5021/8021
MIDTERM EXAM KEY — SPRING 2003**

1. ENTHALPY (150 points)

A commonly tabulated molecular thermodynamic quantity is the 298 K heat of formation. Explain in as detailed a fashion as possible how this quantity is typically computed at the below levels of theory. In addition, comment on roughly how good or bad the given approach is for the calculation of this property and provide some reason for why it is particularly good or bad. You may find choosing a specific example or examples to helpful at various points, and you should feel free to use them, although they are not required.

a) Directly using the MMX force field (30 points).

Force fields compute energies from classical expressions having forms like

$$E_{\text{total}} = \sum_{\text{bonds}} K_r (r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_{\theta} (\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \sum_{n=1}^3 \frac{V_n}{2} [1 + \cos(n\omega)] \\ + \sum_{i < j}^{\text{atoms}} \left(\frac{a_{ij}}{r_{ij}^{12}} - \frac{b_{ij}}{r_{ij}^6} \right) + \sum_{i < j}^{\text{atoms}} \frac{q_i q_j}{\epsilon r_{ij}}$$

For the first three terms, if all molecular parameters were to be at their equilibrium values, the energy contribution would be zero (so-called “strain free”). The remaining terms can make negative or positive energetic contributions.

A force-field heat of formation involves adding the force-field energy to the sum of the heats of formation of all of the atomic types found in the molecule (recall that different atoms can exist as many different types in molecular mechanics, defined by their hybridization, local environment, etc.) The atom-type heats of formation are determined empirically from comparison of heats of formation computed using this protocol to known experimental data.

The quality of force-field heats of formation is highly variable. For simple molecules whose atom types appeared many times in the dataset that was used for parameterizing fragment heats of formation, one would expect results to be reasonable (one is, in some sense, interpolating within well characterized data). For other cases, however, the uncertainty in the atom type’s heat of formation (and indeed, the possibility that it has not even been assigned one) makes the predicted value of questionable utility.

Note that no attempt is made here to account specifically for different contributions to enthalpy. Instead, it is simply assumed that the force-field energy term has the same status as enthalpy, i.e., we ignore the specifics of translation, rotation, vibration, and zero-point vibrational energy.

b) Directly using AM1 (30 points).

AM1 heats of formation are computed from adding atomization energies (*not* atomization *enthalpies*) to the sum of the enthalpies of formation of all of the constituent atoms in a molecule. Thus, in benzene we would compute the AM1 energy of benzene, subtract from that six times the AM1 energy of a carbon atom and six times the energy of a hydrogen atom (thereby obtaining the AM1 atomization energy) and add this quantity to the experimental heats of formation of 6 isolated H atoms and 6 isolated C atoms.

AM1 was specifically parameterized to minimize errors in heats of formation obtained by following this procedure. Thus, as part of the parameterization process, electronic energies were equated with enthalpies. As such, AM1 heats of formation are fair to good for cases where molecules resemble other molecules used in the original training set — a mean unsigned error of about 7 kcal/mol is observed over a large variety of organic molecules. Errors increase substantially outside this training set, and can be particularly large for hypervalent molecules.

One flaw of the AM1 procedure is the same as that noted above for force fields: no explicit accounting for ZPVE or thermal contributions is made in computing the so-called heat of formation. Thus, since parameterization was specifically carried out using enthalpy, these components are somehow absorbed into final parameters in an average way. An example to make this clear might be to consider acetone and oxetane. Both molecules are C_3H_6O species, so their heats of formation will both be computed from adding their atomization energies to the heat of formation of 3 carbon atoms, 6 hydrogen atoms, and one oxygen atom. Thus, the predicted difference in heats of formation for the two will simply be their difference in atomization energies. However, we might also expect the ZPVEs for these two molecules to be quite different (acetone contains a high-frequency C=O stretch that is not present in oxetane, for example) and this factor is not taken account of. Of course, one can *do* a frequency calculation at the AM1 level to compute ZPVE and thermal contributions, but since the electronic energies were parameterized to reproduce heats of formation, it is not obvious that adding in additional terms will not result in some double counting that may leave the results no better than before.

c) Directly at the MP2/6-31+G(d) level (30 points).

At an ab initio level, enthalpy is computed for any given system from summing the electronic energy for the optimized structure with its ZPVE and thermal enthalpy contributions computed from an analytic frequency calculation carried out for the optimized structure. Under the approximations of assuming the molecule to be an ideal gas-phase rigid rotator with all vibrational frequencies well described as harmonic oscillators, the ZPVE is $(1/2 \sum h\nu)$ where the sum is over all vibrational frequencies, and the thermal enthalpy contributions for translation are $3/2 RT$, for rotation $1/2 RT$ per unique rotational axis (3 for nonlinear molecules, 2 for linear molecules), and the vibrational contribution can be computed from the individual frequencies. There is one additional RT term from the PV contribution to enthalpy.

From this enthalpy is subtracted the enthalpy of all of the constituent atoms computed at the same level of theory (in simple cases, just electronic energy + $5/2 RT$ since atoms have neither rotational nor vibrational degrees of freedom). This ΔH of atomization is then added to the heat of formation for all of the individual constituent atoms as determined from experiment (i.e., referenced to the usual experimental standard state). Figure 10.1 of the course text illustrates the above process in graphical form.

At the MP2/6-31+G(d) level, one would expect extremely poor heats of formation. MP2 is a relatively crude way to compute correlation energy. Since there is so much more electron correlation energy in the molecule (with all its electrons) than in any of the individual atoms (with their much smaller numbers of electrons) the error in the molecular calculation will be much larger than the error in the atoms, and predicted heats of formation will be much too high at this level of electronic structure theory.

d) Directly at the G3 level (30 points).

The G3 level follows the same protocol for summing computed atomization enthalpies with atomic reference data as that listed above for the MP2 level, but the enthalpy itself is computed differently at the G3 level than at the MP2 level. Instead of a single electronic structure level being used for geometry optimization and a frequency calculation, multiple levels are used and the final enthalpy is cobbled together from the different components of the calculation. Table 7.5 of the text provides complete details (and one is not expected to be able to regurgitate them on this exam), but particularly salient points include using vibrational frequencies from simple HF/6-31G(d) calculations, electronic energies from hybrid summation of estimated effects associated with basis set completeness and electron correlation level, and an empirical correction for the numbers of paired and unpaired electrons.

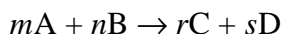
The computed electronic energies at the G3 level are sufficiently accurate (after the empirical correction) that so-called chemical accuracy, which is to say accuracy within 1 kcal/mol, can often be achieved at this level of theory.

The only significant downside to the G3 level is that it may be prohibitively expensive for moderate to large size molecules.

e) Using an isodesmic equation in conjunction with an arbitrarily chosen level (30 points).

An isodesmic equation defines a reaction where the number and types of bonds found on both sides of a chemical reaction are as close to equivalent as possible. In the favorable instance where the heats of formation of all species but one are known experimentally, it is then possible to compute the enthalpy of reaction — which should be reasonably well predicted even at modest levels of theory because errors should cancel out on both sides of the equation — and to subtract from that the heats of formation of known products and add the heats of formation of known reactants to get the heat of formation of the unknown product.

More generally, for reaction

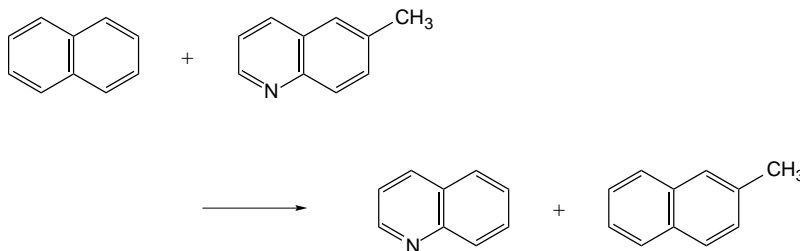


the heat of formation of unknown B can be determined given known heats of formation for A, C, and D as

$$\Delta H_{f,298}^{\circ}(B) = -\frac{1}{n} \left\{ \begin{array}{l} [rH_{298}(C) + sH_{298}(D)] - [mH_{298}(A) + nH_{298}(B)] \\ - [r\Delta H_{f,298}^{\circ}(C) + s\Delta H_{f,298}^{\circ}(D)] + m\Delta H_{f,298}^{\circ}(A) \end{array} \right\}$$

where ΔH values are experimental heats of formation and H values are computed enthalpies relative to separated electrons and nuclei. Note that the upper line in the braces above simply defines the computed enthalpy of reaction.

An example isodesmic equation from the text is



Isodesmic equations can also achieve chemical accuracy in many instances, and they can sometimes do so at much less complete levels of electronic structure theory (e.g., MP2/6-31G(d)) because errors cancel so well from construction of the isodesmic equation.