

Guide to Transition State Modeling with CAChe

David A. Gallagher, CAChe Group, Fujitsu

Transition states can be useful in exploring reaction pathways and calculating activation energy barriers and hence, reaction rates. There are no limits on the number of molecules that can be involved in the transition state, except that MOPAC does limit the number of atoms depending on the specific method being used. Starting with the fastest method, four different strategies for locating and characterizing transition states are summarized below. It is assumed that the reader is familiar with the general operation of the Windows version of CAChe. The Macintosh version will be only slightly different from what is described here.

1 Modify an analogous T-state

The Fragment library in CAChe includes a collection of common transition state geometries. If you can find a transition state that is similar to one you are trying to characterize (i.e. all the atoms involved in bond making and breaking are identical in both transition states), then follow the steps below:

- a.) Make a copy of the analogous transition state from the Fragment library.
- b.) Select all the atoms directly involved in bond making and breaking, then use the "Adjust: Lock" menu item to lock all these atoms rigidly in space.
- c.) Make all the necessary changes to the side chains and other atoms, deleting or adding atoms as necessary. Be careful to select only the atoms that have been changed or added, then "Beautify" hybridization and geometry, or as necessary.
- d.) Optimize the whole structure with Molecular Mechanics and then MOPAC (AM1 or PM3). Note that some procedures automatically include both steps.
- e.) Select all the atoms and use the "Adjust: Unlock" command. Then, save a copy in case the final steps don't work first time.

2 Guess the transition state

This strategy usually takes a little practice and a reasonable knowledge of the probable geometry of the transition state.

- a.) Construct and "Beautify - Comprehensive" the product or reactant molecules, then manually move the atoms into an approximate transition state geometry.
- b.) Using the "Adjust - Define geometry label" menu item, check to see all bond lengths are reasonable. Also, with the space-filling atom display, check that there are no steric problems.
- c.) Save a copy in case the final steps don't work first time.

3 Saddle Calculation

This method automatically locates a transition state between the reactant and product geometries that you present. It is particularly useful for finding complex transition states such as, when more than two bonds are breaking or making simultaneously.

- a.) Construct and optimize the reactant molecule(s), then move them into a reasonable geometry for reaction to occur. Save the file as "Reactant".
- b.) Make a copy of "Reactant" and name it "Product".
- c.) In this new "Product" file, move the atoms, add and delete bonds to approximate the new expected product geometry. Do NOT add or delete any atoms, even temporarily, and do

- NOT use "Beautify Comprehensive" or "Beautify Valence" commands. These steps can destroy the atom numbering consistency, which is critical for Saddle calculations.
- d.) Check that the atom numbering (View: Atom attributes...: Label) on both the product and reactant are precisely consistent for ALL atoms. For example, make sure that Hydrogen numbers do not migrate to different positions.
 - e.) Optimize the geometry via any method.
 - f.) Make a second copy of "Reactant" and name it "T-state".
 - g.) Use the "T-state" file to run a Saddle calculation in MOPAC. When asked for a product, enter the "Product" file. This "T-state" file should become the final transition state geometry. (If the calculation does not run, note that the most common cause of failure for the Saddle calculation is inconsistent atom numbering between the reactant and product files.)
 - h.) Save a copy in case the final steps don't work first time.

4 Reaction Coordinate or Grid

When only one or two bonds are making or breaking during a reaction (but not more than two), it is possible to create atom-distance search labels and plot energy as a function of the one or two atom distance labels. This plot can be used to identify and save any saddle points or transition states for further exploration.

- a.) Construct the product molecule and set up atom distance search labels along the bond(s) that are making or breaking. Use steps of 0.1 angstroms or less.
- b.) Use MOPAC to run an optimized reaction coordinate or grid (Properties of Reactions and Transition states: Map reaction).
- c.) View the resulting map file and move the cursor as close as possible to the apparent saddle position on the 3D surface, or the transition point on a 2D plot. Save this molecule geometry as a new chemical sample file and call it "T-state".
- d.) Save a copy in case the final steps don't work first time.

Final steps for all four strategies:

- x.) Refine the Transition state: Using MOPAC, minimize the gradient. Afterwards, view the calculated bond order (View: Partial Chg. & Calc. Bond Order). If the bonds that are making and breaking appear to be from approximately one-third to two-thirds of a normal bond thickness, then it is probably a transition state. Next, save a copy in case the last step doesn't work first time.
- y.) Verify Transition State: Run a Force (Infra-red) calculation. When viewing the vibrational spectrum from +4,000 to -4,000 wave-numbers, if there is only a single negative frequency, then it is a true transition state. Clicking on the negative vibration should place arrows on the atoms that are moving in the reaction. If there are no negative frequencies, then it is a ground state minimum (see below). If there is more than one negative frequency, then it is a "hill-top". (Note that a possible cause of multiple negative vibrations may be a methyl group that has not fully rotated to the lowest energy position.)

There are no negative frequencies...what now?

Assuming that a transition state really does exist, then there are two likely causes for this problem.

First, the approximate transition state starting geometry was not close enough to the real one. In this case the "Refine the Transition state" option (minimize gradient) would search for the nearest zero-gradient structure which might be a stable minimum (i.e. product or reactant). Try again with a new transition state starting geometry for the final "Refine the Transition state" step.

Secondly, the MOPAC minimize gradient settings may have been too "aggressive". Using the Procedure Editor, change the MOPAC settings under the "Optimization Tab" from "Minimize gradient by EF" to "Minimize gradient by NLLSQ". Then, try the two final steps again on the previously saved copy of the approximate transition state.

Is it the only transition state? Is it a multi-step reaction?

Find Reaction Paths: This "Experiment" option uses MOPAC's intrinsic reaction coordinate (IRC) to test that the transition state found is the only transition state, and the possibility of a multi-step reaction. The +1 and -1 modes track forwards and backwards to the closest minima (reactants and products). If these structures are not identical to the expected reactants and products, then you have probably found an intermediate. Using this intermediate structure as a new "reactant" file, you will need to repeat the whole process to locate a second transition state between the intermediate and the expected end-point.

Solvents?

The use of a solvent field such as COSMO slows calculations considerably. Hence, it is usually much faster to leave the COSMO calculations to the very last step. After the "Refine the transition state" and "Verify the transition state" steps, save copies of the file and repeat these two steps again using the appropriate COSMO solvent field.

An alternative solvation approach is to use explicit solvent molecules. The new MOZYME function in MOPAC 2000 in CAChe 4.4 and 4.5 makes it possible to calculate up to 1,000 atoms. In any case, for systems larger than 50 to 100 atoms, the MOZYME option in MOPAC is much faster, as it scales almost linearly with the number of atoms.

How to speed up calculations on large systems?

If possible, start with a pared down version of your reacting system. Delete all peripheral chains and rings that are not directly involved in the reaction before finding the transition state. This will significantly speed up the calculations. As traditional semiempirical (MOPAC) calculations scale in time as the cube of the number of atoms, removing half the atoms can increase the calculation speed by approximately an order of magnitude. Once the transition state has been found, then you can modify it as an analogous transition state in strategy one described above. For systems larger than 50 to 100 atoms, the MOZYME option in MOPAC is usually faster, as it scales almost linearly with the number of atoms.

Ions and radicals?

The above strategies work with charged systems. As long as the total net charge in the chemical sample file is correct, the charge can be placed on any atom or atoms, as MOPAC will decide how to distribute the electrons based on proximity of atoms. Calculated bond order and partial charges can be viewed graphically after the calculation.

Most of the MOPAC procedures in CAChe will automatically adjust for radicals. Alternatively, the user can invoke UHF (unrestricted Hartree-Fock) in the MOPAC control panel with the Procedure Editor. UHF treats all the electrons individually rather than pairs and, typically, doubles the calculation time.

MOPAC AM1 or PM3?

Both these methods should give similar results. However, AM1 is usually slightly more accurate for partial charges, while PM3 is usually more accurate for heats of formation and covers a wider range of elements. Whichever method you choose, it is important to use the same method for all calculations as most errors are systematic within a method. Hence, although individual absolute numbers may have significant errors, relative differences are much more accurate provided that a consistent method has been used.

For further information visit www.cachesoftware.com or email: support@cachesoftware.com