Tetrahedron Letters, Vol.31, No.35, pp 5085-5088, 1990
Printed in Great Britain

0040-4039/90 $3.00 + .00
Pergamon Press plc

EMPIRICALLY OPTIMIZED "TRANSITION STATE MODELS"

Michael J. Sherrod

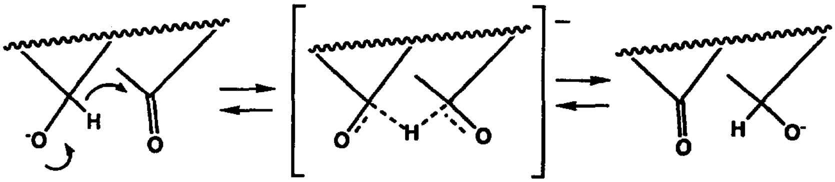
Department of Chemistry, Emory University, Atlanta, Georgia 30322, USA.
Current Address: University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England.

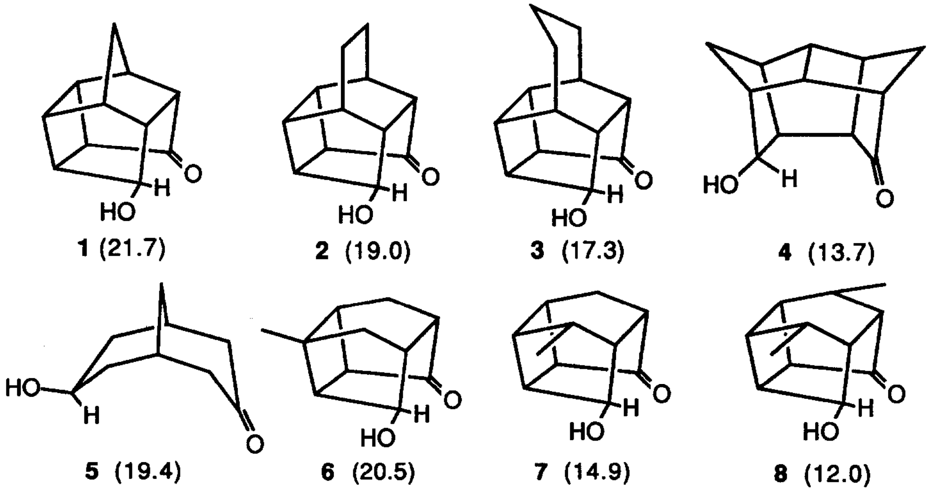
Summary. The predictive ability of molecular mechanics-based "transition state models" can be empirically optimized by systematic adjustment of the related geometric parameters and force constants. A previously developed force field for hydride transfers serves as an example.

In recent years, a major goal of theoretical organic chemistry has been the prediction of reactivity, whether in terms of rate or product ratios.[1,2] In particular, the molecular mechanics-based "transition state modeling" methodology of Houk[2] has recieved a great deal of attention. Unfortunately, careful analyses of the method[3,4] indicate that the high correlations achieved by such models may be an artifact of the arbitrary and approximate procedure of converting *ab initio* structures into parameter sets for molecular mechanics (MM) software.

The debate over the physical relevance of TS models is here considered secondary to their use as practical tools for the prediction of organic reactivity. Indeed, discarding any claim to mechanistic significance would allow us to treat transition substructures just as if they were ground state functional groups for which MM parameters have not yet been derived.[5] Consequently, reasonable guesses can be made about parameters for bond lengths and angles, as well as their related force constants. These initial values could then be iteratively adjusted until experimental observations are adequately duplicated.[6] In this hypothetical case, the experimental observation would be relative rates or product ratios, rather than x-ray structures or conformer distributions.

In order to demonstrate that such an approach is practical as well as possible, we have applied this methodology to our previously developed "standard" transition state model for intramolecular hydride transfers,[4a] observed for the alkoxides of compounds **1-8** (Figure 1). The barriers for the rearrangement, described schematically below, are known from dynamic NMR experiments.[7]





1 (21.7) 2 (19.0) 3 (17.3) 4 (13.7)

5 (19.4) 6 (20.5) 7 (14.9) 8 (12.0)

Figure 1. Hydroxyketone Substrates. Numbers in parentheses indicate experimentally measured activation energies for the hydride transfer (in kcal/mol), ref 7.

Briefly, our "standard" TS model was obtained by following the methodology of Dorigo and Houk:[2] The geometry for the symmetrical transition substructure was taken from the C_{2v} *3-21G* transition structure for the transfer of hydride from methoxide to formaldehyde. Bond stretching and angle bending force constants for the O and H atoms about the reactive carbon center were set to one-half their normal sp3 values. These carbons were also treated as sp3 centers for the purpose of assigning torsional constants.[4a]

The transition structures for compounds **1-8** were then fully optimized with the newly defined parameter set using MODEL,[8] and the resulting steric energies compared to those of their alkoxide ground states[4a] to give calculated "energies of activation." As shown in Figure 2a, the correlation between the calculated and experimental values is poor, with r=0.71.

The parameters listed in the first half of Table I were therefore selected for optimization. As the iterative procedure of refining parameters can be quite tedious and time consuming,[5] we devised a series of VAX/VMS command procedures and FORTRAN programs, collectively given the name FUDGIT,[4b,9] to automatically optimize the force field in the batch mode. (The FUDGIT procedure formally corresponds to a trial and error, steepest ascents search for the parameter set which gives the highest correlation coeficient. Thus, a small change is made in a geometric parameter or force constant, the transition structures reoptimized, and the correlation coefficient re-evaluated. If that change resulted in an increase in the correlation coefficient, that change is retained and the program moves to the next parameter selected for manipulation. This cyclic process is repeated until increasingly small changes in the selected parameters yields no further increase in the correlation coefficient. *The resulting parameter set may therefore be neither unique nor fully optimal in its ability to predict reactivity.*)

The optimization process took ~20 CPU hours on a VAX 8550. The final, optimized parameters are listed in the second half of Table I, and the greatly improved correlation (r=0.96) between computed and experimental energies of activation shown in Figure 2b.

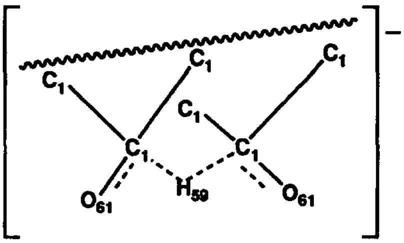


Table I. Optimized Parameters[a,b]

	Before Optimization			After Optimization		
Bonds	R_0	k_r	Dipole Moment	R_0	k_r	Dipole Moment
1-59	1.463	2.3	0.270	1.430	2.392	0.262
1-61	1.261	7.5	4.303	1.424	7.200	4.371
Angles	Θ_0		k_Θ	Θ_0		k_Θ
1-1-59	90.47		0.18	85.12		0.19
1-1-61	120.59		0.70	115.81		0.66
59-1-61	118.16		0.27	104.93		0.31
1-59-1	151.41		0.20	175.67		0.23

a) The atom numbering in the diagram corresponds to the assigned MODEL atom types.
b)Bond lengths are in Å, bond stretching constants in mD/Å, dipoles in DeBye, bond angles in degrees, and bending constants in mD-Å/rad2.

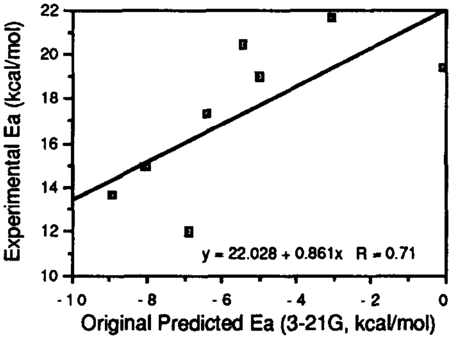


Figure 2a. Correlation between Ea(exp) and Ea(calc) with original parameters (ref. 4).

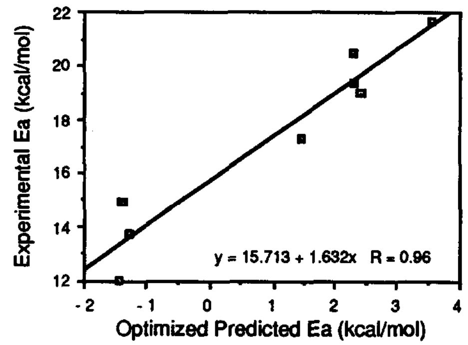


Figure 2b. Correlation between Ea(exp) and Ea(calc) with optimized parameters.

The correlation between calculated and experimental values for the activation barriers could be further improved by the optimization of additional parameters. (As a case in point, the above procedure ignored the possibility of modifying the arbitrarily assigned torsional constants.)

However, there are caveats to be considered. For example, the rates for the hydride transfers considered here are known to be quite sensitive to concentration and counterion,[7] and as such the reported barriers may not be truly comparable with one another. (This is a potential problem with *all* MM-based TS models.) The ability of an empirically derived TS model to predict the outcome of an untried reaction is therefore directly dependent on the number and accuracy of experimental data points used in the optimization procedure. Empirical models, like the one presented here, can at least be reoptimized as additional or more accurate experimental information becomes available.

In summary, the ability of a molecular mechanics-based transition state model to predict reactivity appears to depend less on the modeling of transition states than it does upon the proper selection of parameters for the force field. Irrespective of whether this process is deliberate or fortuitous, the resulting structures and energies should be approached with a healthy skepticism regarding their resemblance to the events in a reaction flask. Models which are deliberately created and empirically optimized for predictive purposes have three advantages over "standard" transition state models: 1) They are unencumbered by assumptions about the nature of the transition state or of chemical reactivity; 2) they by definition provide superior predictive ability; and 3) the model can be refined and improved as additional experimental data become available.

ACKNOWLEDGEMENTS: This work was supported by the National Science Foundation. The author is grateful to Professors D. Liotta and F.M. Menger for discussions regarding this manuscript.

1) CAMEO. W. L. Jorgensen, Purdue University. For a leading reference see: Pageres, G. D.; Jorgensen, W. L. *J. Org. Chem.* **1989**, *54*, 2058.

2) Houk, K. N.; Tucker, J. A.; Dorigo, A. E. *Acc. Chem. Res.* **1990**, *23*, 107.

3) Cieplak, A. S.; Tait, B. D.; Johnson, C. R. *J. Am. Chem. Soc.* **1989**, *111*, 8447.

4) a) Sherrod, M. J.; Menger, F. M. *J. Am. Chem. Soc.* **1989**, *111*, 2611.
b) Sherrod, M. J.; Menger, F.M. submitted for publication.

5) Burkert, U.; Allinger, N. L. *Molecular Mechanics*, Am. Chem. Soc., Washington DC, 1982.

6) As this manuscript was being prepared, a force field for boron enolates, partially derived by "trial and error," was reported: Bernardi, A.; Capelli, A. M.; Gennari, C.; Goodman, J. M.; Paterson, I. *J. Org. Chem.* **1990**, *55*, 3576.

7) Hillier, I. H.; Smith, H.; Mason, S. C.; Whittleton, S. N.; Watt, C. I. F.; Willis, J. *J. Chem. Soc. Perkin Trans II*, **1988**, 1345. Field, M. J.; Hillier, I. H.; Smith, H.; Vincent, M. A.; Mason, S. C.; Whittleton, S. N.; Watt, C. F.; Guest, M. F. *J. Chem. Soc. Chem. Commun.* **1987**, 84. Cernik, R.; Craze, G.; Mills, O. S.; Watt, I. *J. Chem. Soc. Perkin Trans. II*, **1982**, 361. Craze, G.-A.; Watt, I. *J. Chem. Soc. Perkin Trans II*, **1981**, 175.

8) MODEL. Still, W. C., Department of Chemistry, Columbia University.
Version KS 2.94 (May 1989). Steliou, K., Department of Chemistry, University of Montreal. The calculations herein were repeated with MODEL Version KS 2.94 (October 1989) with identical results.

9) The FUDGIT procedure has also been used to locate several parameter sets for the prediction of rates for acid-catalyzed lactonizations.[4b] Although in this case we developed our own custom program for the purpose of optimizing predictive ability, software such as SIMPLEX (available from QCPE) may also be used.

(Received in UK 13 June 1990)