JCTC Journal of Chemical Theory and Computation

Thermochemistry of Hydrocarbons. Back to Extended Hückel Theory

Alexander A. Voityuk*

Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain, and Institute of Computational Chemistry, Universitat de Girona, 17071 Girona, Spain

Abstract: A modified Extended Hückel method that provides accurate values of heats of formation and structural parameters of hydrocarbons is described. The results are reported for an extensive set of molecules and radicals belonging to different classes. The calculated heats of formation for 120 molecules and 26 radicals are close to the experimental data with the mean absolute error of 1.90 kcal/mol. The internal consistency of the calculated data allows reliable prediction of the reaction enthalpy for various hydrocarbon transformations. The proposed scheme is computationally very efficient, and the calculation of a large system requires only a few seconds on a PC. A computer program for the calculation is provided in the Supporting Information.

1. Introduction

The heat (or enthalpy) of formation $\Delta H^0_{f,298}$ of chemical compounds is a key thermodynamic parameter allowing for quantitative description of chemical reactions. Nowadays, experimental heats of formation are available for many hydrocarbons (see NIST Chemistry WEB book[1] and other compilations[2]); however, for molecules that include more than 6 carbon atoms, experimental data are often missing, especially if the compounds include unsaturated cyclic structures. For instance, experimental values of $\Delta H^0_{f,298}$ are available only for 6 isomers of C_6H_6,[1] though more than 200 chemically reasonable structures are possible. Because experimental measurements of accurate heats of formation are quite difficult, elaboration of computational methods to predict thermochemical data has attracted considerable attention. The simplest approach is based on the group additivity scheme;[3] however, this method can be applied only to certain types of molecules and radicals because for many groups the parameters have been undetermined.

Unlike the group additivity scheme, quantum chemical (QC) methods do not have such limitations. A dramatic increase in the computer power makes the QC techniques a good alternative to experiment. The developments and results in this area have been considered in several reviews.[4−6] Various QC approaches are used to derive $\Delta H^0_{f,298}$ and related

data (atomization energies, thermochemical stability of isomers, bond dissociation energies, etc.). The composite *ab initio* schemes G2, G3, and G4 and their variants provide thermochemical data to chemical accuracy (± 1 kcal/mol);[7−9] however, such calculations can be performed only for relatively small systems containing up to 15 heavy atoms.

Density functional theory (DFT) has become the method of choice for computational studies of medium-sized molecules.[10] The B3LYP scheme appears to be the most popular approach to calculate the energetics of chemical reactions; however, recent critical examinations of this and other popular functionals have shown that errors in heats of formation and reaction energies increase significantly with the size of systems.[11−14] For instance, B3LYP performs poorly in predicting the reaction energy of C−C and C−H bond splitting.[15,16] Probably, many difficulties can be overcome by using new exchange-correlation functionals.[17]

The *ab initio* and DFT techniques become inapplicable for extended hydrocarbons because of huge computational demands. QC calculations of large molecular systems can be carried out by using more efficient semiempirical methods.[18] New models[19−21] based on the neglect diatomic differential overlap (NDDO) scheme are enabled to obtain accurate data on the ground state energetics and structure. In many cases, the semiempirical estimates are of the same quality as the results of DFT calculations.[22]

---

* Corresponding author e-mail:alexander.voityuk@icrea.es.

Table 1. Parameters of the MTB/2 Model

Atomic Parameters		
parameter	H	C
U_s^A (eV)	−13.605	−21.559
U_p^A (eV)		−13.507

Bond-Type Parameters			
parameter	H−H	C−H	C−C
λss	0.280	0.275	0.086
λsp		0.218	0.180
$\lambda pp\sigma$			0.186
$\lambda pp\pi$			0.282
βss (eV)	−4.442	−8.574	−5.969
βsp (eV)		−6.813	−6.160
$\beta pp\sigma$ (eV)			−8.420
$\beta pp\pi$ (eV)			−7.403
α_{AB} (Å^{-1})	2.823	2.831	3.401
γ_{AB} (eV)	12.612	99.370	658.659
ω_{AB} (eV)	−0.0791	−0.0340	0.0312
r_{AB} (Å)	2.279	2.843	3.044

Table 2. Experimental and Calculated $\Delta H_{f,298}^0$ for H_2 and Alkanes[a]

molecule	exp.	MTB/2	Δ
hydrogen	0.0	0.0	0.0
methane	−17.8	−14.4	3.4
ethane	−20.0	−18.2	1.8
propane	−25.0	−23.9	1.1
butane	−30.1	−28.9	1.2
pentane	−35.1	−33.9	1.2
n-hexane	−39.9	−38.9	1.0
heptane	−44.9	−43.8	1.1
octane	−49.8	−48.8	1.0
nonane	−54.5	−53.8	0.7
decane	−59.8	−58.7	1.1
undecane	−64.8	−63.7	1.1
dodecane	−69.2	−68.6	0.6
tridecane	−74.5	−73.6	0.9
tetradecane	−79.4	−78.6	0.8
pentadecane	−84.8	−83.5	1.3
hexadecane	−89.6	−88.5	1.1
2-methylpropane	−32.3	−30.7	1.6
2-methylbutane	−36.7	−35.9	0.8
2,2-dimethylpropane	−40.1	−38.1	2.0
3-methylpentane	−41.1	−39.8	1.3
2,3-dimethylbutane	−42.0	−41.7	0.3
2,2-dimethylbutane	−44.0	−43.0	1.0
2-methylpentane	−41.7	−40.9	0.8
3-ethylpentane	−45.4	−47.8	−2.4
3-methylhexane	−45.7	−44.9	0.8
3,3-dimethylpentane	−48.2	−48.3	−0.1
2-methylhexane	−46.6	−45.9	0.7
2,2-dimethylpentane	−49.3	−47.9	1.4
2,4-dimethylpentane	−48.3	−48.0	0.3
2,2,3-trimethylbutane	−48.9	−49.6	−0.7
3,3-dimethylhexane	−52.6	−53.1	−0.5
3,4-dimethylhexane	−50.9	−51.4	−0.5
3-ethyl-2-methylpentane	−50.4	−53.3	−2.9
2,2,4-trimethylpentane	−53.5	−53.0	0.5
2,3,3-trimethylpentane	−51.6	−54.0	−2.4
2,3,4-trimethylpentane	−52.0	−54.1	−2.1
2,5-dimethylhexane	−53.2	−53.5	−0.3
2,3-dimethylhexane	−51.1	−52.3	−1.2
2,4-dimethylhexane	−52.3	−50.1	2.2
2,2-dimethylhexane	−53.7	−52.8	0.9
2,2,3,3-tetramethylbutane	−54.0	−57.3	−3.3
2,2,3-trimethylpentane	−52.6	−54.9	−2.3
3,3-diethylpentane	−55.5	−56.1	−0.6
2,2,4,4-tetramethylpentane	−57.7	−57.9	−0.2
2,2,3,3-tetramethylpentane	−56.7	−61.4	−4.7
tri-t-butylmethane	−47.2	−56.2	9.0
mean error (47 comparisons)			0.40
mean abs. error			1.42
standard deviation			2.06

[a] In kcal/mol.

Another computationally efficient approach is based on a tight-binding (TB) model,[23,24] which bears close similarity to the extended Hückel method (EHM). Almost 50 years ago in his seminal work, Hoffmann employed the EHM method to calculate the structure and energetics of hydrocarbons.[25] While EHM was displaced from computational organic chemistry by NDDO-based semiempirical methods,[18] the TB schemes have been widely employed in computational physics.[23,24,26] The most accurate TB model is based on DFT.[24,27,28] Its performance in predicting the structure and energetics of organic molecules is comparable to that of the conventional semiempirical methods.[29−31]

Several TB schemes have been proposed for hydrocarbons;[32−40] however, most of them are not accurate enough for thermochemical calculations. The mean absolute error in $\Delta H_{f,298}^0$ and atomization energies is often essentially larger than 10 kcal /mol. In addition, the calculated structural parameters of hydrocarbons are less accurate than the results of semiempirical calculations (MNDO, AM1, PM3, and their variants). Recently, we developed two TB schemes that provide more accurate estimates of $\Delta H_{f,298}^0$ and molecular geometries for hydrocarbons.[40,41] The first scheme[40] is based on the conventional extended Hückel method and explicitly includes the overlap matrix of basis atomic orbitals (AOs). In the second approach,[41] the AOs are assumed to be orthogonal, and the overlap matrix is replaced by the unit matrix. This *orthogonal* model is simpler and computationally more efficient than EHM. While TB methods that employ nonorthogonal AOs are commonly considered to be more accurate and transferable,[23,24] the opposite has been found by Jasper et al. for aluminium clusters.[42] We have shown that also for hydrocarbons the *orthogonal* model MTB[41] performs better than the nonorthogonal scheme.[40] In particular, the mean absolute error (MAE) in $\Delta H_{f,298}^0$ calculated with MTB is only half as much as that of the *nonorthogonal model.*[41]

The purpose of the present work is to describe a new variant of the MTB,[41] hereafter referred to as MTB/2, and to demonstrate its performance. Like its predecessor, the

MTB/2 method is based on the orthogonal model and thus is even simpler than EHM; however, the performance of MTB/2 is impressive. The calculated MAE of 1.90 kcal/mol in $\Delta H_{f,298}^0$ for 120 molecules and 26 radicals is considerably smaller than MAEs of the standard semiempirical methods AM1 and PM3 (13.0 and 8.2 kcal/mol, respectively). Notice that MTB/2 is at least an order of magnitude faster than AM1 and PM3. Like the previous schemes,[41] MTB/2 provides quite accurate geometries of hydrocarbons, and the MAEs in bond lengths and bond angles are 0.013 Å and 1.3°. Therefore, we believe that MTB/2 may be helpful for fast estimation of $\Delta H_{f,298}^0$, reaction energetics, and structural parameters for hydrocarbons.

Thermochemistry of Hydrocarbons

J. Chem. Theory Comput., Vol. 4, No. 11, 2008 **1879**

Table 3. Experimental and Calculated $\Delta H^0_{f,298}$ for Unsaturated Aliphatic Hydrocarbons[a]

molecule	MTB/2	exp.	Δ
ethylene	15.7	12.5	3.2
acetylene	53.5	54.5	−1.0
propene	6.7	4.9	1.8
propyne	45.2	44.2	1.0
allene	46.8	45.5	1.3
1-butene	2.8	0.0	2.8
trans-2-butene	−1.6	−2.7	1.1
cis-2-butene	−2.2	−1.7	−0.5
isobutene	−2.7	−4.0	1.3
1,2-butadiene	38.7	38.8	−0.1
1,3-trans-butadiene	28.1	26.3	1.8
1-butyne	39.9	39.5	0.4
2-butyne	37.1	34.8	2.3
vinylacetylene	67.2	70.4	−3.2
diacetylene	107.1	111.0	−3.9
trimethylethene	−10.2	−9.9	−0.3
tetramethylethene	−19.4	−16.8	−2.6
trans-1,3-pentadiene	19.8	18.2	1.6
cis-1,3-pentadiene	20.0	19.5	0.5
isoprene	19.0	18.0	1.0
1,4-pentadiene	29.4	25.2	4.2
mean error (21 comparisons)			0.59
mean abs. error			1.71
standard deviation			2.13

[a] In kcal/mol.

Table 4. Experimental and Calculated $\Delta H^0_{f,298}$ for Cyclic Compounds[a]

molecule	MTB/2	exp.	Δ
cyclopropane	16.7	12.7	4.0
cis-dimethylcyclopropane	−0.7	1.3	−2.0
cyclopropene	62.5	66.2	−3.7
1-methylcyclopropene	53.4	58.2	−4.8
1,2-dimethylcyclopropene	44.2	46.4	−2.2
methylenecyclopropane	46.5	47.9	−1.4
cyclobutane	5.2	6.8	−1.6
cyclobutene	43.3	37.5	5.8
1,2-dimethylcyclobutene	24.3	19.8	4.5
methylenecyclobutane	28.4	29.0	−0.6
cyclopentane	−20.6	−18.3	−2.3
cyclopentene	7.3	8.1	−0.8
cyclopentadiene	34.2	32.1	2.1
3,4−dimethylenecyclobutene	83.7	80.4	3.3
fulvene	51.1	53.5	−2.4
cyclohexane, chair	−29.2	−29.5	0.3
cyclohexene, half-chair	−2.1	−1.2	−0.9
1,3-cyclohexadiene	20.6	25.0	−4.4
1,4-cyclohexadiene	23.0	24.0	−1.0
cycloheptatriene	42.3	43.2	−0.9
cyclooctatetraene	66.9	70.7	−3.8
bicyclobutane	46.6	51.9	−5.3
trans-bicyclopropyl	33.9	30.9	3.0
bicycle[2.1.0]pentane	38.5	37.8	0.7
spirocyclopentane	46.8	44.3	2.5
norbornane	−5.4	−13.1	7.7
norbornadiene	58.3	58.8	−0.5
cyclo[2,2,2]-octane	−23.7	−23.6	−0.1
adamantane	−28.1	−32.2	4.1
cubane	149.8	148.7	1.1
mean error (30 comparisons)			0.01
mean abs. error			2.59
standard deviation			3.24

[a] In kcal/mol.

The paper is organized as follows. In the next section, the method is outlined. Then the performance of the proposed model is considered by comparing MTB/2 estimates with the experimental data and results of G3 and DFT calculations. Finally, conclusions are given and a further possible elaboration of the model is noted. A computer program for MTB/2 calculations is provided in the Supporting Information.

2. Method

The total energy of the system is calculated as a sum of the electronic energy and short-range repulsion energy:

$$E = \sum_i n_i \varepsilon_i + E_{rep} \qquad (1)$$

In eq 1, ε_i is an eigenvalue of the effective one-electron Hamiltonian H, $\varepsilon_i = \langle \psi_i | H | \psi_i \rangle$, and n_i is the occupation number of MO ψ_i. The repulsion energy E_{rep} is approximated by a sum of two-center potentials G_{AB} depending only on the distance between atoms A and B:

$$E_{rep} = \sum_{A>B} G_{AB} \qquad (2)$$

The potential G_{AB} includes implicitly terms of different physical nature (the core-core repulsion, correction due to double counting of the two-electron interactions, the exchange and correlation energy). It is determined by fitting to reference data. Accounting for E_{rep} suggested by Chadi in 1979[43] is significant to get an accurate and transferable TB model.[23,24] Note that much earlier Hoffmann discussed the inclusion of such a term into the EHM total energy.[25]

In the general case (nonorthogonal model), the orbital energies ε_i are obtained by solving the eigenvalue equations

$$\sum_i (H_{\mu\nu} - \varepsilon_i S_{\mu\nu}) c_{\nu i} = 0 \qquad (3)$$

As already mentioned, the scheme becomes faster and probably more accurate if an orthogonal AO basis is employed (it means that the overlap integrals in eq 3 are set to zero, $S_{\mu\nu} = \delta_{\mu\nu}$).[41] The MTB/2 method is based on the orthogonal scheme $(H-\varepsilon) C = 0$. A minimal valence basis of atomic orbitals (one 1s AO for H and four AOs 2s, $2p_x$, $2p_y$, and $2p_z$ for C) is used. The basis set is not explicitly constructed but just has the same symmetry properties as the atomic orbitals. The Hamiltonian operator is replaced with a matrix H, which is defined through semiempirical parameters.

In MTB/2, the diagonal matrix elements $H_{\mu\mu}$ is determined by the parameter U_μ, corresponding to the energy of an electron in AO ϕ_μ ($\mu = s$ and p); while off-diagonal one-center matrix elements are set to zero:

$$H_{\mu\nu} = U_\mu(A) \cdot \delta_{\mu\nu} (\mu, \nu \in A) \qquad (4)$$

Two-center matrix elements (in the diatomic coordinate system) are defined using an empirical formula suggested by Kolb and Thiel[44]

$$H_{\mu\nu} = \pm \beta_{\mu\nu}^{AB} (R_{AB}/a_0)^{1/2} \exp(-\lambda_{\mu\nu}^{AB} R_{AB}^2 / a_0^2) \qquad (5)$$

where $\beta_{\mu\nu}^{AB}$ and $\lambda_{\mu\nu}^{AB}$ ($\mu \in A$, $\nu \in B$) are adjustable parameters. A suitable phase factor ± 1 has to be used to retain the rotational invariance of the method. R_{AB} is the interatomic distance in Å, and the Bohr radius $a_0 = 0.52917$ Å. As the Hamiltonian does not explicitly depend on the density matrix,

1880 *J. Chem. Theory Comput., Vol. 4, No. 11, 2008*

Voityuk

Table 5. Experimental and Calculated $\Delta H^0_{f,298}$ for Aromatic Compounds and Fullerenes[a]

molecule	MTB/2	exp.	Δ
benzene	20.5	19.7	0.8
toluene	12.2	12.0	0.2
ethylbenzene	8.3	7.1	1.2
styrene	34.9	35.3	−0.4
indene	40.4	39.1	1.3
mesitylene	−3.4	−3.8	0.4
naphthalene	35.0	35.9	−0.9
anthracene	52.2	55.2	−3.0
phenanthrene	49.3	49.0	0.3
azulene	66.0	69.1	−3.1
biphenylene	98.2	100.5	−2.3
pyrene	54.3	53.9	0.4
triphenylene	69.2	65.5	3.7
benzo[c]phenanthrene	69.6	69.6	0.0
benzo[a]anthracene	66.1	68.9	−2.8
chrysene	65.5	63.0	2.5
naphthacene	70.3	73.3	−3.0
perylene	73.4	76.3	−2.9
coronene	75.3	73.5	1.8
corannulene	110.8	110.1	0.7
fullerene C_{60}	603.9	604.8	−0.9
fullerene C_{70}	653.1	653.8	−0.7
mean error (22 comparisons)			−0.30
mean abs. error			1.52
standard deviation			1.94

[a] In kcal/mol.

Table 6. Experimental and Calculated $\Delta H^0_{f,298}$ for Radicals[a]

radical	MTB/2	exp.	deviation
methyl	39.8	35.0	4.8
ethyl	29.7	28.9	0.8
vinyl	67.5	64.0	3.5
ethynyl	113.0	130.0	−17.0
1-propyl	23.3	23.3	0.0
i-propyl	18.3	19.1	−0.8
allyl	40.0	39.3	0.7
propargyl	82.1	82.0	0.1
cyclopropyl	64.4	66.9	−2.5
cyclopropenyl	102.7	105.0	−2.3
n-butyl	18.4	19.1	−0.7
s-butyl	13.9	17.0	−3.1
t-butyl	7.4	9.5	−2.1
cyclobutyl	50.2	51.1	−0.9
1-pentyl	13.4	10.1	3.3
2-pentyl	8.0	7.4	0.6
2-me,2-butyl	1.0	3.1	−2.1
neo-pentyl	9.6	10.1	−0.5
1,4-pentadiene-3-yl	47.5	49.8	−2.3
cyclopentyl	20.6	25.6	−5.0
spiropentyl	90.8	90.9	−0.1
cyclopentadienyl	61.7	58.0	3.8
2-hexyl	3.0	3.0	0.0
cyclohexyl	12.2	13.9	−1.7
phenyl	72.5	75.8	−3.3
benzyl	46.5	49.0	−2.5
mean error (26 comparisons)			−1.12
mean abs. error			2.47
standard deviation			4.16

[a] In kcal/mol.

only a single diagonalization of the H matrix is required to calculate the total energy; also, there is no difference by the treatment of closed- and open-shell systems.

The short-range potential G_{AB} consists of two terms:

$$G_{AB} = \gamma_{AB}\exp(-\alpha_{AB}R_{AB}) + \omega_{AB}\exp[-6(R_{AB} - r_{AB})^2] \quad (6)$$

The first term approximates the repulsion potential between atoms A and B, whereas the last one corrects the repulsion at distances close to the van der Waals contact r_{AB}.

The standard heat of formation at 298 K, $\Delta H^0_{f,298}$, is computed using the same procedure as employed in semiempirical methods

$$\Delta H^0_{f,298} = E - \sum_A E_{isol}(A) + \sum_A H^0_{f,298}(A) \quad (7)$$

where E is the total energy of the molecule defined by eq 1, and $E_{isol}(A)$ is the energy of an isolated atom A

$$E_{isol}(A) = n_s U_s(A) + n_p U_p(A) \quad (8)$$

Here, n_s and n_p denote the number of valence s- and p-electrons in atom A. $\Delta H^0_{f,298}(A)$ is the experimental heat of formation of atom A, and for hydrogen and carbon $\Delta H^0_{f,298}(A)$ is 52.1 kcal/mol and 171.29 kcal/mol, respectively.[1] Note that in the previous version of the MTB method[41] as well as in several semiempirical programs an obsolete value of $\Delta H^0_{f,298}(C) = 170.89$ kcal/mol is employed to retain the consistency with previous results. Obviously, this difference of 0.4 kcal/mol does not affect calculated reaction enthalpies, e.g. the relative stability of isomers; however, the absolute values of $\Delta H^0_{f,298}$ of a system will change by $0.4 \times N_C$ kcal/mol with the number of carbon atoms N_C in the molecule. For compounds with $N_C \geq 10$, this difference is remarkable, and reparametrization of the computational scheme is required when using the new value of $\Delta H^0_{f,298}(C)$.
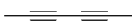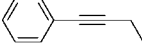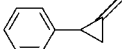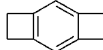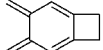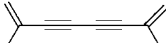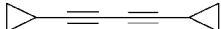
Notice that the zero-point vibrational energy and the enthalpy to heat the molecule from T=0 K to T=298 K are implicitly incorporated into the total energy E through the parametrization (these terms are absorbed in some way into the parameters). Therefore, unlike *ab initio* and DFT, the calculations of the zero-point vibrational and heating energies are not required within MTB/2 to estimate $\Delta H^0_{f,298}$ and reaction enthalpies.

The parameters of the effective Hamiltonian and the repulsive potential G_{AB} were derived as follows. The parameter U_s for hydrogen was set to −13.605 eV (the negative of the ionization potential of the atom). Note that a shift of U_s and U_p parameters for H and C by a constant δ does not affect $\Delta H^0_{f,298}$ of neutral species, although it changes absolute values of orbital energies $\varepsilon_i \rightarrow \varepsilon_i + \delta$. All other parameters were fitted using experimental heats of formation[1,2] and structural parameters for a training set of molecules and radicals. Unlike the MTB scheme,[41] where the parameter r_{AB} (eq 6) is determined as the sum of the van der Waals radii of atoms A and B, in MTB/2, r_{AB} is an adjustable parameter.

The molecules and radicals in training sets were chosen to represent the most common bonding situations in hydrocarbons. Despite remarkable disparities among experimental heats of formation of C_{60} (the measured values of $\Delta H^0_{f,298}$ range from 599.1 to 634.9 kcal/mol),[1,45] we included C_{60} in the training set. As a reference value, we adopted $\Delta H^0_{f,298} = 604.6$ kcal/mol employed by Green et al. for parameterization of a group additivity method for polycyclic aromatic hydrocarbons.[45] A nonlinear least-squares method was used to

Thermochemistry of Hydrocarbons

J. Chem. Theory Comput., Vol. 4, No. 11, 2008 **1881**

Table 7. Comparison of Experimental and Calculated Values of $\Delta H^0_{f,298}$ and Relative Energies of C_6H_6 and $C_{10}H_{10}$ Isomers[a]

C_6H_6	$\Delta H^0_{f,298}$, in *kcal/mol*		Relative energies, in *kcal/mol*			
	Exp	MTB/2	Exp	MTB/2	B3LYP	MP2
	19.8	21.3	0.0	0.0	0.0	0.0
	53.6	51.3	33.8	30.0	34.7	36.0
	80.4	82.6	60.6	61.4	62.9	67.9
	90.2	91.8	70.4	70.5	68.9	69.5
	94.6	93.1	74.8	71.8	82.1	91.7
$C_{10}H_{10}$	Exp	MTB/2	Exp	MTB/2	B3LYP	MP2
	53.5	56.2	0.0	0.0	0.0	0.0
	59.4	61.4	5.9	5.3	-8.5	1.2
	70.0	67.8	16.5	11.6	0.7	12.4
	74.0	77.0	20.5	20.9	3.1	10.1
	79.9	74.8	26.3	18.6	20.3	22.8
	85.8	78.8	32.3	22.6	17.6	31.1
	118.1	116.8	64.6	60.6	48.8	63.0
	134.3.1	138.5	80.8	82.3	71.2	81.0

[a] B3LYP/6-31G** and MP2/6-31G** data were taken from the paper by Schreiner et al.[14]

optimize the parameters. Several parameterization runs were carried out starting from different parameter values and using different training sets. To choose the parameter set that yields most balanced results, extensive survey calculations were performed with several sets of parameters. The final MTB/2 parameters are listed in Table 1. A simple computer program for MTB/2 calculations of hydrocarbons under the MS Windows environment is provided in the Supporting Information.

3. Results and Discussion

3.1. Heats of Formation. Tables 2−6 contain the calculated and experimental values of $\Delta H^0_{f,298}$ for 120 molecules and 26 radicals that belong to different classes of hydrocarbons. The experimental data are taken from the standard sources[1,2] unless otherwise specified. The MAE of 1.90 kcal/mol is considerably smaller than those of the conventional semiempirical methods. For the same set of molecules, the

MAE of AM1 and PM3 is 13.0 and 8.2 kcal/mol, respectively. Excluding fullerenes C_{60} and C_{70} from the statistics (the $\Delta H^0_{f,298}$ values for these compounds are substantially overestimated by AM1 and PM3), one gets the MAEs to be 7.3 and 5.0 kcal/mol, which are still essentially larger than that of MTB/2.

For 16 medium-sized hydrocarbons of the G3 subset (see Table S1 in the Supporting Information), the MTB/2 MAE of 1.65 kcal/mol is considerably smaller than that of B3LYP (9.64 kcal/mol), AM1 (6.45 kcal/mol), and PM3 (3.44 kcal/mol). Various computational schemes of the G3 family provide very accurate results with the MAE ranging from 0.7 to 1.0 kcal/mol;[12] however, such estimations are much more time-consuming (by a factor of 10^5) than the MTB/2 calculations.

Alkanes. As seen from Table 2, small deviations of the calculated values from the experimental $\Delta H^0_{f,298}$ are found both for short and long alkanes. The MAE for 47 molecules including normal and branched compounds is 1.42 kcal/mol. MTB/2 consistently overestimates $\Delta H^0_{f,298}$ n-alkanes C_nH_{2n+2} by 0.40 kcal/mol. The CH_2 increment is well reproduced. Notice that this is not a common mark of QC methods. For instance, the error in the G3 enthalpies increases with chain length ranging from 0.25 to 1.93 kcal/mol.[16] The B3LYP error in $\Delta H^0_{f,298}$ amounts to -1.5 kcal/mol for propane and -30.3 kcal/mol for hexadecane. Note that the small effect of conformational averaging on the enthalpies of the n-alkanes can be neglected.[11]

The $\Delta H^0_{f,298}$ values of branched hydrocarbons are also well reproduced (Table 2). Thus, MTB/2 shows a good performance in estimating the energy differences between sterically crowded hydrocarbons and their linear isomers. For instance, the conversion enthalpy of neopentane to pentane is predicted to be 4.2 kcal/mol in good agreement with the experimental value of 5.1 kcal/mol. In contrast, B3LYP shows rather large branching errors (10−15 kcal/mol).[14] MTB/2 overestimates the stability of compounds with neighboring methyl groups, the deviation of -4.7 kcal/mol is found for 2,2,3,3-tetramethylpentane (Table 2), whereas $\Delta H^0_{f,298}$ of the 2,2,4,4-isomer is well reproduced (the error is only -0.2 kcal/mol).

Unsaturated Aliphatic Hydrocarbons. Table 3 compares the calculated and experimental heats of formation for 21 compounds with double and triple bonds. The MAE of MTB/2 is found to be 1.71 kcal/mol. On average, the scheme underestimates the stability of these compounds, and the mean deviation is 0.59 kcal/mol. The largest deviation of 4.18 kcal/mol is found for 1,4-pentadiene.

Cyclic Hydrocarbons. Analysis of the data for 30 cyclic molecules (Table 4) shows that the MAE for these species is somewhat larger, 2.59 kcal/mol. On average, MTB/2 overestimates the stability of 21 *monocyclic* molecules by 0.61 kcal/mol. The largest deviation of 5.80 kcal/mol is calculated for cyclobutene. Table 4 lists also heats of formation for nine *bi-* and *polycyclic* hydrocarbons. Such molecules are known to be quite difficult for semiempirical methods. MTB/2 gives surprisingly good estimates of $\Delta H^0_{f,298}$ for most compounds. A relatively large deviation is found for norbornane (7.7 kcal/mol).
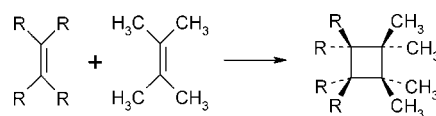
Aromatic Compounds and Fullerenes. MTB/2 provides reliable heats of formation for 22 molecules listed in Table 5. The MAE is only 1.52 kcal/mol. The largest deviation with experiment of 3.7 kcal/mol is obtained for triphenylene. Experimental values of $\Delta H^0_{f,298}$ are available only for two fullerenes, C_{60} and C_{70}. Large disparity among the experimental heats of formation has been already mentioned. For $\Delta H^0_{f,298}$ of C_{60} and C_{70}, MTB/2 gives 604.8 and 653.1 kcal/mol which are in excellent agreement with the most reliable experimental values of 604.8 and 653.8 kcal/mol (Table 5).[45] Note that the standard and new semiempirical NDDO methods do not provide acceptable values of $\Delta H^0_{f,298}$ for fullerenes. For instance, the AM1 and PM3 errors in the heat of formation of C_{70} are 407.3 and 229.5 kcal/mol, respectively.

Radicals. Even for open-shell systems, the scheme gives quite accurate estimates of $\Delta H^0_{f,298}$ (Table 6). The calculated heats of formation are in good agreement with experiment.[46] For 26 radicals, the MAE of MTB/2 is 2.47 kcal/mol (Table 6). The largest deviation of 17 kcal/mol is found for the ethynyl radical in which one carbon has the only adjacent atom. It suggests that MTB/2 is of limited use for treatment of R-C≡C• radicals. If the ethynyl radical is excluded from statistics, we obtain the MAE=1.89 kcal/mol, which is similar to that found for closed-shell systems.

3.2. Isomerization Reactions. The ability of carbon atoms to form single, double, and triple bonds leads to the great variety of hydrocarbons. As the CC bond energy ranges from 80 kcal/mol for C−C to 200 kcal/mol for C≡C, the enthalpy of isomerization reactions (*i.e.* relative values of $\Delta H^0_{f,298}$ for isomers) may be very helpful for assessing the performance and inherent consistency of a computational method. In Table 7, we compare experimental and calculated heats of formation of C_6H_6 and $C_{10}H_{10}$ isomers. B3LYP and MP2 results for the isomers were recently reported by Schreiner et al.[14] Overall, the MTB/2 scheme provides the most accurate estimates. For C_6H_6, the three methods show a comparable accuracy; however, the stability of the last isomer is remarkably underestimated by B3LYP and MP2 (the errors are 7.3 and 16.9 kcal/mol), whereas MTB/2 gives an accurate estimate.

For $C_{10}H_{10}$ isomers, the MAE of MTB/2 is found to be 3.4 kcal/mol. The MP2 data are also in good agreement with experiment. In contrast, the relative energies of isomers derived from B3LYP seem to be unreliable. The data presented in Table 7 suggest that MTB/2 provides internally consistent estimates of $\Delta H^0_{f,298}$ across different types of hydrocarbons.

3.3. [2+2] Cyclization of Tetramethylethene and Various Alkenes. Table 8 compares the energy for the reaction calculated with different QC approaches. As refer-

ence data for the reaction energy, we use G3MP2 results.[16] G3MP2 predicts all the reactions to be exothermic, although the formation of crowded cyclobutanes has slightly smaller reaction energy. The MTB/2 estimates are in very good

Thermochemistry of Hydrocarbons

J. Chem. Theory Comput., Vol. 4, No. 11, 2008 **1883**

Table 8. Predicted Energies (kcal/mol) for the [2+2] Cyclization of Tetramethylethene and Various Alkenes[a]

alkene	G3MP2	MP2	B3LYP	MTB/2
ethene	−**19.3**	−27.2	−7.2	−20.1
propene	−**17.4**	−26.9	−4.4	−18.6
isobutene	−**15.7**	−23.4	2.2	−16.1
Z-butene	−**18.0**	−25.8	−0.1	−17.7
E-butene	−**16.7**	−27.7	−2.0	−17.3
trimethyethene	−**15.5**	−25.5	4.1	−14.9
tetramethylethene	−**15.8**	−23.3	9.4	−12.0

[a] The 6-31++G** basis set was used for MP2 and B3LYP calculations. The G3MP2, MP2, and B3LYP data were taken from the paper by Check and Gilbert.[16]

agreement with the G3MP2 data (MAE is of 1.1 kcal/mol), whereas MP2 calculations provide less accurate results (MAE is 8.9 kcal/mol);[16] B3LYP performs poorly predicting the reaction energies with MAE of 17.2 kcal/mol.[16]

3.4. Bond Dissociation Energies. The C−C and C−H bond energies are often required for the discrimination of reaction mechanisms for complex processes such as hydrocarbon cracking and combustion. Despite the expanding availability of the experimental measurements, the database for the accurate bond energetics is still not comprehensive. Since MTB/2 provides quite accurate values of $\Delta H^0_{f,298}$ for molecules and radicals, it can be used to estimate the enthalpy of bond-breaking and bond-forming reactions. In Table 9, we compare the enthalpies for 14 C−H and 18 C−C bond dissociation reactions. As can be seen, the experimental data are well reproduced by MTB/2. The mean absolute error is 2.45 kcal/mol. It has been shown that B3LYP provides unreliable results for C−C bond energies in branched hydrocarbons,[16] in particular, the bond dissociation energy in $(CH_3)_3C−C(CH_3)_3$ is underestimated by *ca.* 20 kcal/mol.

On average, MTB/2 overestimates the energy of C−H bond (the mean deviation is 1.61 kcal/mol) and underestimates the strength of C−C bond (the mean deviation is −1.63 kcal/mol). Thus, the scheme can be helpful for computer simulations of processes that involve bond-breaking or/and bond-forming reactions in hydrocarbons.

3.5. Geometries. Calculated molecular geometries of hydrocarbons and fullerenes agree well with experiment (see Table S2 in the Supporting Information). MTB/2 reproduces accurately bond lengths and bond angles of different types, the corresponding MAEs are 0.013 Å (110 comparisons) and 1.3° (38 comparisons). The standard semiempirical methods show similar accuracy: 0.013 Å and 1.4° (AM1) and 0.011 Å and 1.6° (PM3). Such errors seem to be not of great chemical significance, and more accurate estimation of molecular geometries than this is not an important requirement for approximate QC approaches.

3.6. Comparison of MTB/2 and MTB. The MTB/2 method shows remarkable improvement over its predecessor MTB.[41] The mean absolute error estimated for $\Delta H^0_{f,298}$ of 146 species of MTB/2 (1.90 kcal/mol) is essentially smaller than that of MTB (2.82 kcal/mol). The new scheme shares with MTB the ability to accurately reproduce heats of formation of alkanes and alkenes. A major improvement has been found in the treatment of polycyclic aromatic molecules, and the MAEs of MTB/2 and MTB are 1.52 and 4.07 kcal/

Table 9. Bond Dissociation Enthalpies[a]

reaction		exp.	MTB/2
C−H Bond Breaking			
methane	→ CH$_3$+H	104.9	106.3
ethane	→ C$_2$H$_5$+H	101.0	100.0
ethylene	→ C$_2$H$_3$+H	103.6	103.9
propane	→ n-C$_3$H$_7$+H	100.4	99.3
	→ i-C$_3$H$_7$+H	96.2	94.3
propene	→ C$_3$H$_5$+H	86.6	85.5
n-butane	→ n-C$_4$H$_9$+H	100.2	99.4
	→ s-C$_4$H$_9$+H	99.1	94.9
isobutane	→ t-C$_4$H$_9$+H	93.7	90.8
cycloropane	→ c-C$_3$H$_5$+H	106.3	99.8
cyclopropene	→ c-C$_3$H$_3$+H	90.9	92.3
cyclopentadiene	→ c-C$_5$H$_5$+H	78.0	79.7
benzene	→ Ph + H	108.2	104.0
toluene	→ PhCH$_2$ + H	89.1	86.4
mean error (14 comparisons)			1.54
mean abs. error			1.88
standard deviation			2.39
C−C Bond Breaking			
ethane	→ 2 CH$_3$	90.0	97.9
propane	→ C$_2$H$_5$+CH$_3$	88.9	93.4
propene	→ C$_2$H$_3$+CH$_3$	94.2	100.6
n-butane	→ 2C$_2$H$_5$	87.8	88.3
	→ C$_3$H$_7$+CH$_3$	88.3	92.1
isobutane	→ i-C$_3$H$_7$+CH$_3$	86.2	88.9
butene-1	→ C$_2$H$_3$+C$_2$H$_5$	92.9	94.3
	→ Allyl+CH$_3$	74.3	77.0
methylcyclopropene	→ c-C$_3$H$_3$+ CH$_3$	98.9	101.5
2-methylbutane	→ i-C$_3$H$_7$+C$_2$H$_5$	84.7	84.0
	→ s-C$_4$H$_9$+CH$_3$	88.7	89.7
neo-pentane	→ tC$_4$H$_9$+CH$_3$	84.7	86.0
2,3-dimethylbutane	→2 i-C$_3$H$_7$	80.7	78.5
2,2,3,3-tetramethylbutane	→ 2 t-C$_4$H$_9$	75.2	73.4
toluene	→ Ph + CH$_3$	98.8	101.1
ethylbenzene	→ PhCH$_2$ + CH$_3$	76.9	78.1
	→ Ph + C$_2$H$_5$	97.6	93.9
styrene	→ Ph + C$_2$H$_3$	104.5	105.1
mean error (18 comparisons)			1.69
mean abs. error			2.15
standard deviation			2.89

[a] In kcal/mol.

mol. MTB overestimates $\Delta H^0_{f,298}$ of aromatic compounds; for instance, $\Delta H^0_{f,298}$ of chrysene and coronene are predicted to be too high by ca. 10 kcal/mol. In contrast, MTB/2 performs very well for large aromatic molecules (see Table 5); in particular, the errors in $\Delta H^0_{f,298}$ of chrysene and coronene are 2.5 and 1.8 kcal/mol. The MTB/2 $\Delta H^0_{f,298}$ values for radicals are, overall, somewhat better than those from MTB; the corresponding MAEs are 2.47 and 2.84 kcal/mol. Both methods provide satisfactory molecular geometries of hydrocarbons.

3.7. Computational Performance. The limiting step of semiempirical calculations is the diagonalization of the Hamiltonian matrix. Usually a single-point calculation requires about 20 iterations. It means that 20 diagonalizations are needed for closed-shell systems, while twice as many operations are required for the spin-unrestricted scheme commonly applied to radicals. Moreover, the number of iterations remarkably increases with the size of the model, especially when treating radicals. In such situations, semiempirical calculations become time-consuming. In contrast, a single diagonalization of the Hamiltonian matrix is required within MTB/2, independent of whether a closed- or open-shell system is computed. This makes MTB/2 extremely

1884 *J. Chem. Theory Comput., Vol. 4, No. 11, 2008*

Voityuk

efficient for studying large species. For instance, a single-point MTB/2 calculation on a PC of the ground state of C_{60} and C_{70} takes 0.30 and 0.42 s, respectively, whereas AM1 calculations are found to be by a factor of 8 less efficient (for C_{60} and C_{70}, they require 2.20 and 3.34 s). As expected, a more considerable gain in computational time is found for open-shell systems. Our MTB/2 and AM1 calculations of the lowest triplet state of C_{60} take 0.30 and 12.40 s, respectively (a factor of 40!).

4. Conclusions

We described the MTB/2 approach that provides accurate thermochemical data and structural parameters of hydrocarbons and fullerenes. The mean absolute error of 1.90 kcal/mol is obtained for the $\Delta H^0_{f,298}$ values across various classes of molecules and radicals (146 comparisons). The proposed scheme gives more accurate estimates for reaction enthalpies than MP2, B3LYP, and the standard semiempirical methods. Molecular geometries are well reproduced by MTB/2; the MAEs of 0.013 Å and 1.3° are found for bond lengths and bond angles, respectively.

As the MTB/2 method is based on the noniterative orthogonal scheme, it is computationally very efficient and allows for fast screening of $\Delta H^0_{f,298}$ and reaction enthalpies for large hydrocarbons and fullerenes. Because of its performance, the scheme appears very promising to study the reaction dynamics of hydrocarbons.

In MTB/2, we employ simple functions to construct the effective Hamiltonian matrix (eqs 4 and 5) and short-range interatomic potential (eq 6). The performance of the model is likely to be improved by using more flexible functions to approximate the distance dependence of these terms. Obviously, more accurate thermochemical data can be obtained for a certain class of hydrocarbons (at the cost of less accurate estimates for other compounds) by the corresponding tuning of the MTB/2 parameters.

Supporting Information Available: G3, B3LYP, AM1, PM3, and MTB/2 values of $\Delta H^0_{f,298}$ for 16 medium-sized hydrocarbons (Table S1), comparison of the MTB/2 molecular geometries of hydrocarbons with experimental data (Table S2), and a computer program for MTB/2 calculations on PC under the MS Windows environment along with a brief instruction and input examples. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) *NIST Chemistry WebBook*; NIST Standard Reference Database, No. 69; Mallard, W. G., Linstrom, P. J., Eds.; National Institute of Standards and Technology: Gaithersburg, MD.http://webbook.nist.gov (accessed Apr 12, 2008).

(2) Pedley, J. B.; Naylor, R. D.; Kirby, S. P. *Thermochemical Data of Organic Compounds*, 2nd ed.; Chapman and Hall: New York, 1986.

(3) Cohen, N.; Benson, S. W. *Chem. Rev.* **1993**, *93*, 2419.

(4) *Computational Thermochemistry*; Irikura, K. K., Frurip, D. J., Eds.; ACS Symposium Series No. 677, Washington, DC, 1998.

(5) Bond, D. *J. Org. Chem.* **2007**, *72*, 5555.

(6) *Quantum-Mechanical Prediction of Thermochemical Data*; Understanding Chemical Reactivity Vol. 22, Cioslowski, J., Ed.; Kluwer Academic: Dordrecht, 2001.

(7) Curtiss, L. A.; Raghavachari, K. *Theor. Chem. Acc.* **2002**, *108*, 61.

(8) Raghavachari, K.; Curtiss, L. In *Theory and Applications of Computational Chemistry: The First 40 Years*; Dykstra, C., Frenking, G., Kim, K., Scuseria, G., Eds.; Elsevier: 2005.

(9) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *127*, 124105.

(10) Koch, W.; Holthausen, M. C. *Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, 2001.

(11) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. *J. Phys. Chem. A* **2000**, *104*, 5850.

(12) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374.

(13) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2005**, *123*, 124107.

(14) Schreiner, P. R.; Fokin, A. A.; Pascal, R. A.; de Meijere, A. *Org. Lett.* **2006**, *8*, 3635.

(15) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. *Org. Lett.* **2006**, *8*, 3631.

(16) Check, C. E.; Gilbert, T. M. *J. Org. Chem.* **2005**, *70*, 9828.

(17) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(18) Thiel, W. *Adv. Chem. Phys.* **1996**, *93*, 703.

(19) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601.

(20) Giese, T. J.; Sherer, E. C.; Cramer, C. J.; York, D. M. *J. Chem. Theory Comput.* **2005**, *1*, 1275.

(21) Tuttle, T.; Thiel, W. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2159.

(22) Tirado-Rives, J.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2008**, *4*, 297.

(23) Goringe, C. M.; Bowler, D. R.; Hernandez, E. *Rep. Prog. Phys.* **1997**, *60*, 1447.

(24) Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316.

(25) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397.

(26) Selvam, P.; Tsuboi, H.; Koyama, M.; Kubo, M.; Miyamoto, A. *Catal. Today* **2005**, *100*, 11.

(27) Porezag, D.; Frauenheim, T.; Kohler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947.

(28) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.

(29) Krüger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T. *J. Chem. Phys.* **2005**, *122*, 114110.

(30) Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. A* **2006**, *110*, 13551.

(31) Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751.

(32) Xu, C. H.; Wang, C. Z.; Chan, C. T.; Ho, K. M. *J. Phys. Condens. Matter* **1992**, *4*, 6047.

(33) Wang, C. Z.; Chan, C. T.; Ho, K. M. *Phys. Rev. Lett.* **1991**, *66*, 189.

(34) Wang, Y.; Mak, C. H. *Chem. Phys. Lett.* **1995**, *235*, 37.

Thermochemistry of Hydrocarbons

J. Chem. Theory Comput., Vol. 4, No. 11, 2008 **1885**

(35) Horsfield, A. P.; Godwin, P. D.; Pettifor, D. G.; Sutton, A. P. *Phys. Rev. B* **1996**, *54*, 15773.

(36) Winn, M. D.; Rassinger, M.; Hafner, J. *Phys. Rev. B* **1997**, *55*, 5364.

(37) Pan, B. C. *Phys. Rev. B* **2001**, *64*, 155408.

(38) Zhao, J. J.; Lu, J. P. *Phys. Lett. A* **2003**, *319*, 523.

(39) Wang, C. Z.; Ho, K. M. *J. Comput. Theor. Nanosci.* **2004**, *1*, 3.

(40) Voityuk, A. A. *J. Chem. Theory Comput.* **2006**, *2*, 1038.

(41) Voityuk, A. A. *Chem. Phys. Lett.* **2006**, *433*, 216.

(42) Jasper, A. W.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 210.

(43) Chadi, D. J. *Phys. Rev. B* **1979**, *19*, 2074.

(44) Kolb, M.; Thiel, W. *J. Comput. Chem.* **1993**, *14*, 775.

(45) Yu, J.; Sumathi, R.; Green, W. H. *J. Am. Chem. Soc.* **2004**, *126*, 12685.

(46) Sablier, M.; Fujii, T. *Chem. Rev.* **2002**, *102*, 2855.