

# Scaled Effective One-Electron Method Based on G2 Theory: Results for Aliphatic Alkane Molecules

Gyula Tasi\*,†,‡ and Fujio Mizukami†

Department of Surface Chemistry, National Institute of Materials and Chemical Research, 1-1 Higashi, Tsukuba, Ibaraki 305, Japan, and Applied Chemistry Department, József Attila University, Rerrich B. tér 1, H-6720 Szeged, Hungary

Received January 5, 1998

An effective one-electron model was constructed and parameterized using G2 molecular total energies. For aliphatic alkane molecules, the calculated gas-phase standard heats of formation closely matched the experimental values. For 63 molecules, the average absolute deviation and maximum absolute deviation were 0.83 and 3.70 kcal/mol, respectively. The results show that it is important to perform full conformational analysis and to treat the torsional vibrations as free internal rotations at higher temperatures. The calculated standard heats of formation were also compared with those obtained from revised group additivity values and via the AM1 method. The optimized equilibrium molecular geometries were even better than the RHF/6-31G\* ones.

## INTRODUCTION

The Extended Hückel Molecular Orbital<sup>1</sup> (EHMO) and Atom Superposition and Electron Delocalization<sup>2</sup> (ASED) effective one-electron methods have long been used to study the electronic structures of molecules, clusters, and extended systems. Several recent papers have dealt with the applicability of these methods to molecular geometry optimization.<sup>3–5</sup> The popularity of the EHMO-ASED methods arises from their simplicity.

For the most widely-used semiempirical methods (SINDO1,<sup>6</sup> MNDO,<sup>7</sup> AM1,<sup>8</sup> PM3,<sup>9</sup> and SAM1<sup>10</sup>), a time-consuming parameterization was performed. In general, this was not so for the EHMO-ASED methods: simple rules were used to determine the values of the parameters. It is not surprising at all, therefore, that these methods were not suitable for molecular modeling.<sup>11</sup>

With the help of the recently developed Gaussian-2 (G2) composite *ab initio* method, precise total energies can be obtained for the local minima of potential energy surfaces of neutral molecules in their ground states.<sup>12</sup> Its modified version can also be applied to determine transition state structures and total energies.<sup>13</sup> Curtiss *et al.* studied the performance of the G2 method, its modified versions, and several DFT methods with respect to the calculation of enthalpies of formation of various molecules.<sup>14</sup> For the G2 neutral test set of 148 molecules, it was found that, of the Gaussian methods, the G2 theory affords the most precise theoretical results. The average absolute deviation and the maximum absolute deviation were 1.58 and 8.2 kcal/mol, respectively. The G2 neutral test set or the G2 theory can be applied to develop new DFT functionals,<sup>15</sup> to parameterize new quantum chemical methods, and in several cases to check the reliability of experimental data. Unfortunately, the application of this high-level *ab initio* quantum chemical

theory is very expensive. At present, it cannot be used to study the thermochemical properties of large molecules (ca. 10 heavy atoms or more) with conformational flexibility, which are common in everyday practice.

The present work is the first step in the comprehensive parameterization of an effective one-electron model. The results obtained for aliphatic alkane molecules are reported and compared with those of other quantum chemical calculations and with the experimental values.

## THEORETICAL BACKGROUND

The total energy ( $E_{\text{tot}}$ ) of a molecule is calculated from eqs 1–3 within the framework of the EHMO-ASED methods:

$$E_{\text{tot}} = E_{\text{elec}} + E_{\text{rep}} \quad (1)$$

$$E_{\text{elec}} = \sum_i \sigma_i \epsilon_i = \sum_{\mu} \sum_{\nu} P_{\mu\nu} H_{\mu\nu} \quad (2)$$

$$E_{\text{rep}} = \frac{1}{2} \sum_A \sum_{B \neq A} E_{AB} \quad (3)$$

In eqs 1 and 2,  $E_{\text{elec}}$  is the electronic energy,  $\sigma_i$  is the occupation number ( $\sigma_i \in \{0, 1, 2\}$ ), and  $\epsilon_i$  is the orbital energy of the  $i$ th MO;  $\mathbf{P}$  is the first-order density matrix defined by eq 4

$$\mathbf{P}_{\mu\nu} = \sum_i^{\text{occ}} \sigma_i C_{\mu i} C_{\nu i} \quad (4)$$

In eq 4, occ means that the summation is carried out over the occupied MOs. For calculation of the repulsion energy  $E_{\text{rep}}$  via eq 3, only the two-body repulsion potential function  $E_{AB}(R_{AB})$  needs to be defined.

The first derivative of the total energy with respect to the coordinate  $x_A$  of atom A is given by eq 5:

\* National Institute of Materials and Chemical Research.

† József Attila University.

$$\frac{\partial E_{\text{tot}}}{\partial x_A} = \sum_{\mu} \sum_{\nu} \mathbf{P}_{\mu\nu} \frac{\partial H_{\mu\nu}}{\partial x_A} - \sum_{\mu} \sum_{\nu} \mathbf{W}_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial x_A} + \frac{\partial E_{\text{rep}}}{\partial x_A} \quad (5)$$

The matrix  $\mathbf{W}$  in eq 5 is the energy-weighted density matrix defined by eq 6:

$$\mathbf{W}_{\mu\nu} = \sum_i^{\text{occ}} \sigma_i \epsilon_i c_{\mu i} c_{\nu i} \quad (6)$$

In most cases, the first derivatives of  $E_{\text{rep}}$  with respect to the atomic coordinates can readily be calculated, and thus they do not require further discussion. With the usual approximation ( $\mathbf{H} = k\mathbf{S}$ , where  $k$  is a scalar), only the matrix  $\partial\mathbf{S}/\partial\mathbf{R}$  is necessary for calculation of the matrix  $\partial\mathbf{H}/\partial\mathbf{R}$ .

To transform eq 7 (a generalized matrix eigenvalue equation) into eq 8 (a standard matrix eigenvalue equation), the neat procedure<sup>16</sup> provided by Masamura can be used. This procedure is appreciably faster than the standard Löwdin and Gram–Schmidt orthogonalization processes.

$$\mathbf{H}\mathbf{c} = \epsilon\mathbf{S}\mathbf{c} \quad (7)$$

$$\mathbf{H}\mathbf{c}' = \epsilon\mathbf{c}' \quad (8)$$

Below, we define an effective one-electron model and investigate its ability to provide realistic molecular geometries, permanent electric dipole moments, total energies, and standard heats of formation for aliphatic alkane molecules.

Our effective one-electron model takes all the electrons of the system into consideration (all-electron method); for the calculation of molecular integrals and their derivatives, Gaussian functions are used<sup>5</sup> (STO-3G). The elements of the effective one-electron matrix  $\mathbf{H}$  are given by the well-known eqs 9 and 10:<sup>1,2</sup>

$$H_{\mu\mu} = -I_{\mu} \quad (9)$$

$$H_{\mu\nu} = \frac{1}{2}K_{\mu\nu}(H_{\mu\mu} + H_{\nu\nu})S_{\mu\nu} \quad (\mu \neq \nu) \quad (10)$$

The ionization potentials  $I_{\mu}$  in eq 9 are regarded as parameters. For simplicity, the parameters  $K_{\mu\nu}$  are assumed to be bonding parameters: H–H, C(sp<sup>3</sup>)–H, and C(sp<sup>3</sup>)–C(sp<sup>3</sup>) for aliphatic alkanes. Equation 11 defines the distance-dependency<sup>2</sup> of the parameter  $K_{\mu\nu}$ :

$$K_{\mu\nu} = \kappa_{\mu\nu} \exp(-\alpha_{\mu\nu} R_{AB}^{\beta_{\mu\nu}}) \quad (11)$$

In eq 11,  $R_{AB}$  is the distance between atoms A and B. The parameters  $\kappa_{\mu\nu}$ ,  $\alpha_{\mu\nu}$ , and  $\beta_{\mu\nu}$ , are to be determined via parameter estimation. The exponents of the GTO basis functions play a role in the calculation of the matrix  $\mathbf{S}$ . They determine the size of the AOs. The scaling factors ( $\zeta_{\text{ns}}$  and  $\zeta_{\text{np}}$ ) of the GTO basis functions are taken as parameters.

For the two-body repulsion potential  $E_{AB}$  in eq 3, eq 12 suggested by Jurs and Dixon is used:<sup>4</sup>

$$E_{AB} = \frac{Z_A Z_B}{R_{AB}} \exp[-(\delta_A + \delta_B) R_{AB}^{(\epsilon_A + \epsilon_B)}] \quad (12)$$

In eq 12,  $Z_A$  and  $Z_B$  are the atomic numbers (nuclear charges) of atoms A and B, respectively. The atomic parameters  $\delta_A$ ,  $\delta_B$ ,  $\epsilon_A$ , and  $\epsilon_B$  should be determined via parameter estimation.

**Table 1.** Reference Molecular Properties and Calculated Enthalpies of Formation<sup>a</sup>

molecules	$\mu(\text{ref})$	$E_{\text{tot}}^0(\text{G2})$	ZPVE (G2)	$\Delta H_f^0$ (G2)	$\Delta H_f^{298}$ (G2)
methane	0	−40.410 883	0.042665	−69.8	−77.7
ethane	0	−79.630 875	0.071222	−70.4	−86.1
propane	0.0853	−118.855 809	0.098723	−84.0	−106.3
butane (t)	0	−158.081 177	0.126046	−98.7	−127.1
butane (g <sup>+</sup> )	0.0907	−158.080 150	0.126192	−96.0	−124.5
isobutane	0.1269	−158.084 302	0.125694	−106.9	−135.5
pentane (tt)	0.0867	−197.306 578	0.153322	−113.5	−147.9
pentane (g <sup>−</sup> t)	0.0631	−197.305 540	0.153476	−110.8	−145.3
pentane (g <sup>−</sup> g <sup>−</sup> )	0.0598	−197.304 914	0.153697	−109.2	−144.0
pentane (g <sup>−</sup> g <sup>+</sup> )	0.1232	−197.301 908	0.153647	−101.3	−135.9
isopentane	0.0924	−197.308 956	0.153048	−119.8	−154.3
neopentane	0	−197.314 775	0.152377	−135.0	−169.8

<sup>a</sup>  $\mu(\text{ref})$  in Debye,  $E_{\text{tot}}^0(\text{G2})$  and ZPVE(G2) in au,  $\Delta H_f^0(\text{G2})$  and  $\Delta H_f^{298}(\text{G2})$  in kJ/mol.

This model can be called the *scaled* effective one-electron method (SEOEM) because every element of the matrix  $\mathbf{H}$  is to be “scaled”. On the basis of eqs 9–12, the number of parameters to be adjusted is 21.

The gradient vector  $\nabla_A E_{\text{tot}}$  with respect to the coordinates of atom A can be calculated via eq 13, derived from eq 5:

$$\begin{aligned} \nabla_A E_{\text{tot}} = & \sum_{B \neq A} \left[ \sum_{\mu \in A} \sum_{\nu \in B} \{ [K_{\mu\nu}(H_{\mu\mu} + H_{\nu\nu})P_{\mu\nu} - 2W_{\mu\nu}] \nabla_A S_{\mu\nu} - \right. \\ & \alpha_{\mu\nu} \beta_{\mu\nu} K_{\mu\nu} (H_{\mu\mu} + H_{\nu\nu}) R_{AB}^{\beta_{\mu\nu}-1} P_{\mu\nu} S_{\mu\nu} \nabla_A R_{AB} \} - \\ & \frac{Z_A Z_B}{R_{AB}^2} \exp[-(\delta_A + \delta_B) R_{AB}^{(\epsilon_A + \epsilon_B)}] [1 + \\ & \left. (\delta_A + \delta_B)(\epsilon_A + \epsilon_B) R_{AB}^{(\epsilon_A + \epsilon_B)}] \nabla_A R_{AB} \right] \quad (13) \end{aligned}$$

In eq 13, the notation  $\mu \in A$  means that the summation is carried out over the AOs centered on atom A.

## SOFTWARE PACKAGES

The *ab initio* calculations were performed with the Gaussian 94 package<sup>17</sup> on a CRAY C90 computer. The initial geometries were prepared, and the results were evaluated on a PC with the Pcmol package.<sup>18</sup>

The SEHMO (Scaled Extended Hückel MO) program<sup>5</sup> was used for the effective one-electron quantum chemical calculations. In the program, one or more contracted Gaussian basis functions can be used to approximate the atomic orbitals, and split-valence basis sets can also be applied. For an element, any number of electrons can be taken into account: the inner shell electrons can therefore be treated as well.

The PSEHMO (parameterization for SEHMO) program<sup>5</sup> was applied for parameterization of the SEOEM model. Only the following steps were necessary: (i) definition of the quantum chemical model, then (ii) selection of the reference molecules and properties, and finally (iii) provision of a trial parameter set. The SEHMO and PSEHMO programs were run on the PC and on the CRAY machine.

**Table 2.** Atomic Parameters (in au) of the SEOEM Model

atoms	$\zeta_{1s}$	$I_{1s}$	$\zeta_{2s}$	$I_{2s}$	$\zeta_{2p}$	$I_{2p}$	$\delta$	$\epsilon$
H	1.324	0.486153					0.62165	1.21604
C(sp <sup>3</sup> )	5.366	17.804347	1.820	0.774008	1.600	0.397845	0.48858	1.08583

**Table 3.** Bonding Parameters (in au) of the SEOEM Model

interaction types	$\kappa$	$\alpha$	$\beta$
H-H	0.064796	0.0569	0.8597
C(sp <sup>3</sup> )-H	0.064270	0.0447	0.9219
C(sp <sup>3</sup> )-C(sp <sup>3</sup> )	0.069577	0.0518	1.1389

**Table 4.** SEOEM Calculation Results for Molecules Listed in Table 1<sup>a</sup>

molecules	$\mu$	$E_{\text{tot}}^0$	ZPVE	$\Delta H_f^0$	$\Delta H_f^{298}$
methane	0	-40.408305	0.04456	-63.0	-70.9
ethane	0	-79.630682	0.07224	-69.9	-84.7
propane	0.0736	-118.856040	0.09922	-84.6	-105.3
butane (t)	0	-158.081298	0.12603	-99.0	-125.2
butane (g <sup>+</sup> )	0.0942	-158.080996	0.12622	-98.2	-124.6
isobutane	0.1177	-158.084327	0.12575	-107.0	-133.2
pentane (tt)	0.0752	-197.306574	0.15281	-113.5	-145.1
pentane (g <sup>-</sup> t)	0.0607	-197.306228	0.15298	-112.6	-144.4
pentane (g <sup>-</sup> g <sup>-</sup> )	0.0681	-197.305825	0.15330	-111.6	-143.7
pentane (g <sup>-</sup> g <sup>+</sup> )	0.1114	-197.304292	0.15322	-107.5	-139.6
isopentane	0.0858	-197.309010	0.15264	-119.9	-151.3
neopentane	0	-197.315455	0.15192	-136.8	-168.4

<sup>a</sup>  $\mu$  in Debye,  $E_{\text{tot}}^0$  and ZPVE in au,  $\Delta H_f^0$  and  $\Delta H_f^{298}$  in kJ/mol.

**Table 5.** Statistical Data Obtained by Linear Regression Analyses between the Reference and SEOEM Data

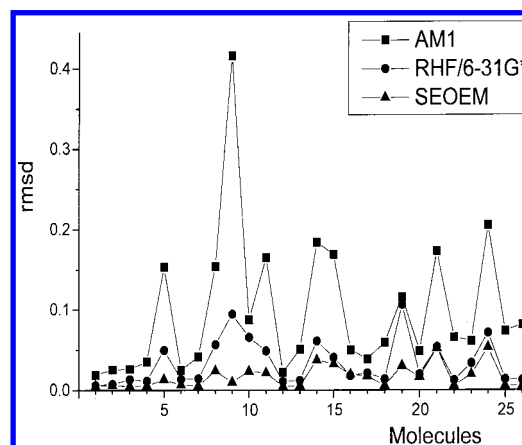
quantity	$\gamma$	$\rho$	$\sigma$
$\mu$	1.06366	0.99735	0.00557
$E_{\text{tot}}^0$	1.00000	1.00000	0.00100
ZPVE	1.00096	1.00000	0.00070
$\Delta H_f^0$	0.99044	0.99964	2.74818
$\Delta H_f^{298}$	1.00946	0.99983	2.43889

The AM1 calculations were performed with the MOPAC 6.0 package<sup>19</sup> on the PC and with the Gaussian 94 package on the CRAY computer.

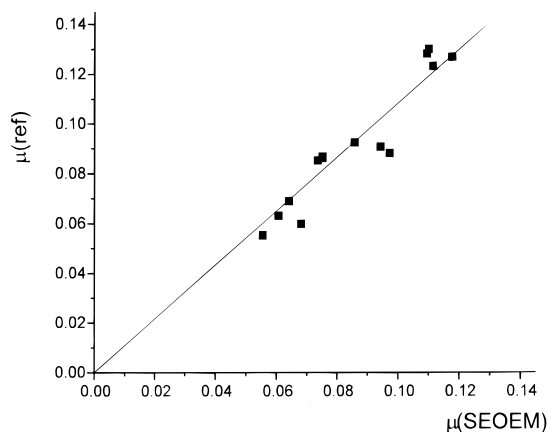
#### CALCULATIONS ON ALIPHATIC ALKANE MOLECULES

Several papers have recently dealt with the structures and electronic properties of aliphatic alkane molecules.<sup>20-22</sup> In general, the introduction of alkyl groups to form branching in the chain has a significant influence on the molecular geometries and electron distributions. Consequently, an adequate quantum chemical treatment is of utmost importance.

It has been shown that the RHF/6-311++G\*\*//RMP2-(FC)/6-311G\*\* *ab initio* quantum chemistry model can be successfully applied to calculate the structural and electronic properties of aliphatic alkane molecules.<sup>22</sup> (FC (frozen core) in parentheses means that the inner shell electrons were excluded from the electron correlation calculations.) The computed equilibrium molecular geometries are close to the experimental gas-phase geometries, and the same is true for the permanent electric dipole moment vectors. To acquire reliable reference molecular total energies, G2 calculations were performed on several aliphatic alkane molecules. Curtiss *et al.* have already published G2 results on the



**Figure 1.** The rmsd deviations (Å) of the equilibrium molecular geometries determined by various methods from the reference geometries (1: methane, 2: ethane, 3: propane, 4: butane (t), 5: butane (g<sup>+</sup>), 6: isobutane, 7: pentane (tt), 8: pentane (g<sup>-</sup>t), 9: pentane (g<sup>-</sup>g<sup>-</sup>), 10: pentane (g<sup>-</sup>g<sup>+</sup>), 11: isopentane, 12: neopentane, 13: hexane (ttt), 14: 2-methylpentane, 15: 3-methylpentane, 16: 2,3-dimethylbutane, 17: 2,2-dimethylbutane, 18: heptane (tttt), 19: 2,3-dimethylpentane, 20: 2,2-dimethylpentane, 21: 3-methylhexane, 22: octane (ttttt), 23: 2,2,3,3-tetramethylbutane, 24: 2-methylheptane, 25: nonane (tttttt), 26: decane (ttttttt).)



**Figure 2.** Graphical representation of the relation  $\mu(\text{ref})$  vs  $\mu(\text{SEOEM})$  for aliphatic alkanes (Debye). (propane, butane (g<sup>+</sup>), isobutane, pentane (tt), pentane (g<sup>-</sup>t), pentane (g<sup>-</sup>g<sup>-</sup>), pentane (g<sup>-</sup>g<sup>+</sup>), isopentane, 2-methylpentane, 3-methylpentane, 2,2-dimethylbutane, heptane (ttt), 3-methylhexane, 2-methylheptane, nonane (ttttt).)

methane, ethane, propane, *trans*-butane, *gauche*-butane, and isobutane molecules.<sup>14</sup>

The parameterization was performed on the following basis set of molecules: methane, ethane, propane, butane (t), butane (g<sup>+</sup>) ( $\equiv$  *gauche*(+)-butane), isobutane, pentane (tt), pentane (g<sup>-</sup>t), pentane (g<sup>-</sup>g<sup>-</sup>), pentane (g<sup>-</sup>g<sup>+</sup>), isopentane, and neopentane. The reference properties were the RMP2-(FC)/6-311G\*\* equilibrium molecular geometries, the RHF/6-311++G\*\*//RMP2(FC)/6-311G\*\* permanent electric dipole moment vectors ( $\mu(\text{ref})$ ) and the G2 total energies at 0 K ( $E_{\text{tot}}^0(\text{G2})$ ), including the zero-point vibrational energy (ZPVE(G2)) corrections. Within the G2 approximation, the scaled HF/6-31G\* harmonic vibrational frequencies are used

**Table 6.** Experimental and Calculated Standard Heats of Formation of Aliphatic Alkanes (kJ/mol)

molecules	exptl	SEOEM			AM1
		I	II	GAV	
methane	-74.4	-70.9	-70.9		-36.7
ethane	-83.8	-84.7	-84.9	-83.7	-72.9
propane	-104.7	-105.3	-106.1	-104.6	-101.5
butane	-125.6	-125.2	-126.4	-125.5	-130.2
isobutane	-134.2	-133.3	-134.4	-135.6	-122.9
pentane	-146.9	-145.1	-146.8	-146.4	-158.9
isopentane	-153.7	-151.6	-153.7	-153.1	-148.1
neopentane	-168.1	-168.4	-169.7	-167.8	-137.1
hexane	-167.1	-165.0	-167.0	-167.4	-187.5
2-methylpentane	-174.8	-171.2	-173.7	-174.1	-176.6
3-methylpentane	-172.1	-169.4	-172.7	-170.7	-172.7
2,3-dimethylbutane	-178.3	-176.4	-179.1	-180.7	-162.5
2,2-dimethylbutane	-186.1	-185.0	-187.5	-182.0	-157.5
heptane	-187.7	-184.8	-187.4	-188.3	-216.2
2-methylhexane	-194.6	-191.1	-194.4	-195.0	-205.2
3-methylhexane	-191.3	-189.1	-192.9	-191.6	-201.1
3-ethylpentane	-189.6	-187.1	-190.0	-188.3	-194.3
2,2-dimethylpentane	-205.9	-204.5	-206.6	-202.9	-185.7
2,3-dimethylpentane	-198.9	-194.4	-197.6	-198.3	-186.9
2,4-dimethylpentane	-201.7	-197.1	-200.5	-201.7	-191.0
3,3-dimethylpentane	-201.2	-200.3	-204.5	-196.2	-176.9
2,2,3-trimethylbutane	-204.5	-207.1	-210.2	-206.3	-168.3
octane	-208.6	-204.7	-207.8	-209.2	-244.8
2-methylheptane	-215.4	-210.9	-215.8	-215.9	-233.9
3-methylheptane	-212.5	-208.9	-212.8	-212.5	-229.8
4-methylheptane	-212.0	-208.7	-213.3	-212.5	-229.6
3-ethylhexane	-210.7	-206.8	-210.4	-209.2	-222.8
2,2-dimethylhexane	-224.6	-224.3	-227.7	-223.8	-214.4
2,3-dimethylhexane	-213.8	-214.1	-217.3	-219.2	-214.8
2,4-dimethylhexane	-219.2	-214.9	-217.8	-219.2	-215.6
2,5-dimethylhexane	-222.5	-217.4	-221.0	-222.6	-222.9
3,3-dimethylhexane	-220.0	-219.7	-223.6	-217.1	-205.1
3,4-dimethylhexane	-212.8	-212.4	-215.2	-212.5	-209.1
3-ethyl-2-methylpentane	-211.0	-210.5	-213.2	-215.9	-211.6
3-ethyl-3-methylpentane	-214.8	-217.2	-220.3	-210.5	-197.1
2,2,3-trimethylpentane	-220.0	-224.2	-227.1	-223.8	-192.6
2,2,4-trimethylpentane	-224.0	-225.5	-228.8	-230.5	-199.6
2,3,3-trimethylpentane	-216.3	-223.2	-225.9	-220.5	-187.4
2,3,4-trimethylpentane	-217.3	-216.9	-219.6	-222.6	-199.8
2,2,3,3-tetramethylbutane	-225.6	-235.6	-238.2	-231.8	-170.7
nonane	-228.2	-224.5	-228.2	-230.1	-273.5
3,3-diethylpentane	-232.3	-232.0	-235.4	-224.7	-214.5
2,2,3,3-tetramethylpentane	-237.1	-250.2	-252.6	-246.0	-189.1
2,2,3,4-tetramethylpentane	-236.9	-244.1	-247.2	-244.8	-201.5
2,2,4,4-tetramethylpentane	-241.6	-249.5	-253.3	-246.0	-193.5
2,3,3,4-tetramethylpentane	-236.1	-243.4	-246.0	-244.8	-196.9
2,2,5-trimethylhexane	-253.1	-250.2	-255.7	-251.5	-232.0
2,3,5-trimethylhexane	-242.7	-239.6	-245.2	-243.5	-232.7
decane	-249.5	-244.4	-248.7	-251.0	-302.1
2-methylnonane	-259.9	-250.6	-255.8	-257.7	-291.2
5-methylnonane	-258.6	-248.4	-254.1	-254.4	-286.9
undecane	-270.9	-264.2	-269.8	-272.0	-330.8
2,2,5,5-tetramethylheptane	-302.4	-299.4	-305.7	-294.6	-260.9
3,3,5,5-tetramethylheptane	-276.7	-281.5	-285.7	-274.5	-233.1
2,2,4,4,5-pentamethylhexane	-281.0	-286.4	-290.2	-284.5	-222.7
dodecane	-289.7	-284.1	-290.5	-292.9	-359.4
3,3,6,6-tetramethyloctane	-318.2	-313.8	-322.0	-308.8	-279.8
4,4,6,6-tetramethylnonane	-313.2	-316.7	-325.8	-323.0	-285.1
3,5-diethyl-3,5-dimethylheptane	-301.3	-310.6	-316.1	-302.9	-268.6
5,5,7,7-tetramethylundecane	-366.6	-356.3	-368.1	-371.5	-342.4
4,6-diethyl-4,6-dimethylnonane	-347.1	-336.1	-345.9	-344.8	-310.3
hexadecane	-374.8	-363.5	-376.8	-376.6	-474.0
octadecane	-414.6	-403.2	-418.5	-418.4	-531.3

to calculate the ZPVEs. The value of the scale factor<sup>23</sup> is 0.8929. Table 1 shows the reference data, together with the G2 enthalpies of formation at 0 K ( $\Delta H_f^0(\text{G2})$ ) and at 298.15 K ( $\Delta H_f^{298}(\text{G2})$ ), for the 12 molecules. The procedure for calculation of the enthalpy of formation within the G2 approximation is fully described in the literature.<sup>14</sup> The thermal correction to the enthalpy of molecules was calcu-

**Table 7.** Conformational Analysis of the Hexane Molecule Performed by the SEOEM Model

no.	conformers	$x_i$	$\Delta H_f^{298}/\text{kJ mol}^{-1}$
1	tg <sup>+</sup> t	0.0619	-167.9
2	g <sup>+</sup> tt	0.0656	-168.1
3	ttt	0.0935	-168.9
4	g <sup>+</sup> tg <sup>+</sup>	0.0484	-167.3
5	tg <sup>-</sup> t	0.0619	-167.9
6	g <sup>+</sup> g <sup>+</sup> t	0.0421	-167.0
7	g <sup>-</sup> g <sup>-</sup> g <sup>-</sup>	0.0300	-166.1
8	g <sup>+</sup> g <sup>+</sup> t	0.0082	-162.9
9	g <sup>-</sup> g <sup>+</sup> g <sup>+</sup>	0.0057	-162.0
10	g <sup>-</sup> g <sup>-</sup> t	0.0421	-167.0
11	g <sup>+</sup> g <sup>+</sup> g <sup>+</sup>	0.0049	-161.6
12	tg <sup>+</sup> g <sup>+</sup>	0.0421	-167.0
13	g <sup>+</sup> g <sup>-</sup> g <sup>-</sup>	0.0049	-161.6
14	tg <sup>+</sup> g <sup>-</sup>	0.0082	-162.9
15	g <sup>-</sup> tg <sup>+</sup>	0.0424	-167.0
16	g <sup>-</sup> tg <sup>-</sup>	0.0484	-167.3
17	ttg <sup>-</sup>	0.0656	-168.1
18	g <sup>+</sup> g <sup>+</sup> g <sup>-</sup>	0.0p49	-161.6
19	tg <sup>-</sup> g <sup>+</sup>	0.0082	-162.9
20	g <sup>+</sup> g <sup>+</sup> g <sup>+</sup>	0.0300	-166.1
21	tg <sup>-</sup> g <sup>-</sup>	0.0421	-167.0
22	tg <sup>-</sup> g <sup>-</sup>	0.0085	-163.0
23	g <sup>+</sup> tg	0.0424	-167.0
24	g <sup>-</sup> tt	0.0656	-168.1
25	g <sup>-</sup> g <sup>-</sup> g <sup>+</sup>	0.0057	-162.0
26	g <sup>+</sup> g <sup>+</sup> t	0.0085	-163.0
27	g <sup>+</sup> g <sup>+</sup> g <sup>-</sup>	0.0057	-162.0
28	ttg <sup>+</sup>	0.0656	-168.1
29	g <sup>-</sup> g <sup>+</sup> t	0.0082	-162.9
30	g <sup>+</sup> g <sup>-</sup> t	0.0085	-163.0
31	g <sup>+</sup> g <sup>-</sup> g <sup>-</sup>	0.0057	-162.0
32	g <sup>-</sup> g <sup>+</sup> g <sup>-</sup>	0.0007	-157.0
33	g <sup>-</sup> gg <sup>+</sup>	0.0049	-161.6
34	tg <sup>-</sup> g <sup>+</sup>	0.0085	-163.0
35	g <sup>+</sup> g <sup>-</sup> g <sup>+</sup>	0.0007	-157.0

lated as follows:  $3/2RT$  for translation,  $3/2RT$  for rotation (nonlinear molecules), the vibrational term from the harmonic vibrational frequencies, and the pV term ( $RT$ ) which converts energy to enthalpy. The electronic and nuclear partition functions, as usual, were set to unity, so their contributions were ignored. For calculation of the data  $\Delta H_f^{298}$  in Table 1, the low-frequency torsional vibrations were regarded as harmonic oscillations.

In general, the coincidence of the *ab initio* and the semiempirical total energies cannot be expected, since the latter considerably smaller. For the SEOEM model, the coincidence of the total energies was accomplished by increasing and parameterizing the ionization potential  $I_{1s}$  of the inner shell electrons of the carbon atom.

The initial values of the parameters were taken from the literature,<sup>1,2,24</sup> and for the trial shape of the scaling function of  $E_{AB}$  in eq 12, a simple Gaussian function was chosen ( $\delta_H = \delta_C = 0.5$ ,  $\epsilon_H = \epsilon_C = 1.0$ ). The optimal values of the parameters were determined with the help of the PSEHMO program on the CRAY computer. The values obtained are given in Tables 2 and 3.

The SEOEM results for the molecules comprising the basis set of molecules are to be seen in Table 4. To calculate the ZPVE corrections within the SEOEM approximation, the harmonic vibrational frequencies are used without scaling. For calculation of the data  $\Delta H_f^{298}$  in Table 4, the low-frequency torsional vibrations, as a first approximation, were again regarded as harmonic oscillations. Table 5 shows the statistical data ( $\gamma$ , regression coefficient;  $\rho$ , linear correlation



coefficient; and  $\sigma$ , root mean-square deviation) obtained by linear regression analyses between the reference and SEOEM data. It is to be seen that the goodness of fit is acceptable in every case, and the SEOEM enthalpies of formation are very close to the G2 values even at 298.15 K. The latter means that the SEOEM harmonic vibrational frequencies can be applied to calculate vibrational thermal corrections too.

For some aliphatic alkane molecules, Figure 1 shows the deviations of the equilibrium molecular geometries obtained by various quantum chemical methods from the RMP2(FC)/6311G\*\* geometries. The rmsd (root mean-square distance) data were calculated via eq 14:

$$\text{rmsd} = \left( \frac{\sum_{i=1}^n \Delta_i^2}{n} \right)^{1/2} \quad (14)$$

In eq 14, the data  $\Delta_i$  give the minimum distances between the corresponding fitted atoms in the moving and in the reference molecule;  $n$  is the number of atomic pairs fitted. Every atom in each molecule was selected for fitting. It is to be seen that the SEOEM geometries are even better than the RHF/6-31G\* ones.

Figure 2 shows the correlation between the RHF/6-311++G\*\*//RMP2(FC)/6-311G\*\* and the SEOEM dipole moments. It can be seen that the correlation is acceptably good.

The following question now arises. Can we apply the SEOEM model to large molecules to obtain reliable molecular total energies? Does it afford, for instance, good thermochemical data (e.g., standard heats of formation) for both small and large molecules, as the G2 theory does for small ones?

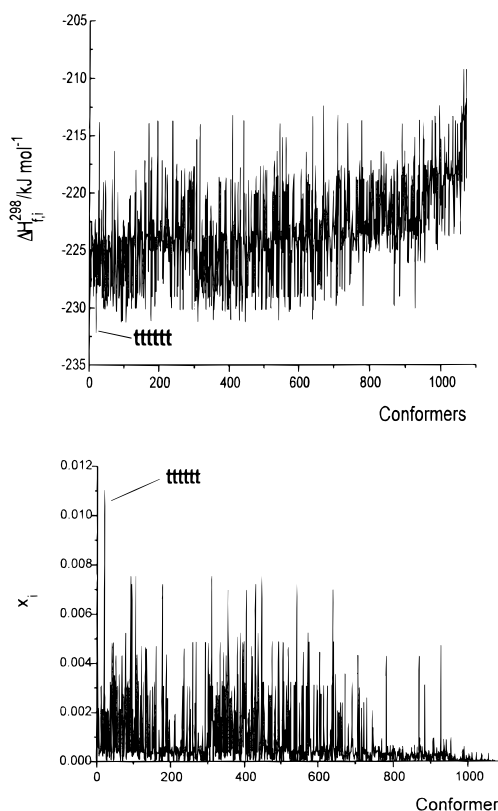
Table 6 shows gas-phase experimental and calculated standard heats of formation for 63 aliphatic alkane molecules. The experimental values were obtained from a recent compilation.<sup>25</sup> For comparison, the standard heats of formation calculated from revised group additivity values<sup>25</sup> (GAV) and via the AM1 method are also presented.

For SEOEM I data, the calculation was performed on the global minimum of the potential energy surface (PES), and the low-frequency torsional vibrations were regarded as harmonic oscillations.

As far as the SEOEM II data are concerned, we performed conformational analysis for each molecule to collect as many local minima of the PES as possible. Firstly, a large number of trial geometries (between 1000 and 100 000) were generated via the Monte-Carlo method. For larger unbranched molecules, of course, more initial geometries were considered. Secondly, from every initial geometry, geometry optimization was started, and the distinct local minima obtained were collected.<sup>26</sup> When the local minima of the PES were available and vibrational analysis was performed in each case, the standard heat of formation was calculated via eq 15, with summation of the contributions of different conformers:

$$\Delta H_f^{298} = \sum_i x_i \Delta H_{f,i}^{298} \quad (15)$$

In eq 15,  $x_i$  is the fraction of conformer  $i$  in the gas mixture



**Figure 3.** SEOEM II results of the conformational analysis of nonane.

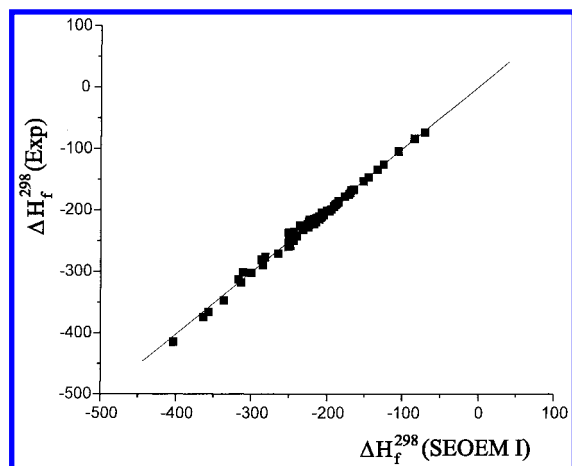
at 298.15 K and at 1 atm. These fractions were calculated via eq 16 by assuming that Boltzmann statistics can be applied to determine the populations of different conformers:

$$x_i = \frac{\exp(-\Delta E_i/kT)}{\sum_i \exp(-\Delta E_i/kT)} \quad (16)$$

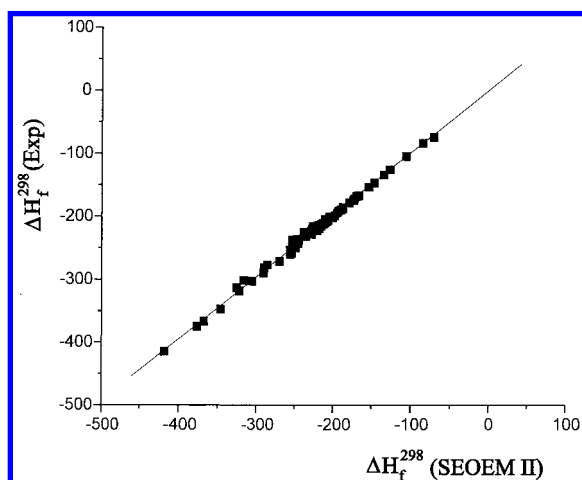
In eq 16,  $\Delta E_i$  is the difference between the energies of conformer  $i$  and the conformer with the lowest energy (global minimum). For calculation of the data  $\Delta H_{f,i}^{298}$ , the low-frequency torsional vibrations were regarded as free internal rotations. The cut-off value used was 260  $\text{cm}^{-1}$ , as suggested by Radom *et al.*<sup>27</sup> Table 7 and Figure 3 show the results for the hexane and nonane molecules, respectively.

As far as the AM1 data are concerned, Table 6 presents the standard heats of formation calculated for the conformers obtained from the global minima of the SEOEM potential energy surfaces via geometry optimization. For unbranched aliphatic alkane molecules, the AM1 method tends to overestimate the stability of the all-*trans* conformer.<sup>21</sup>

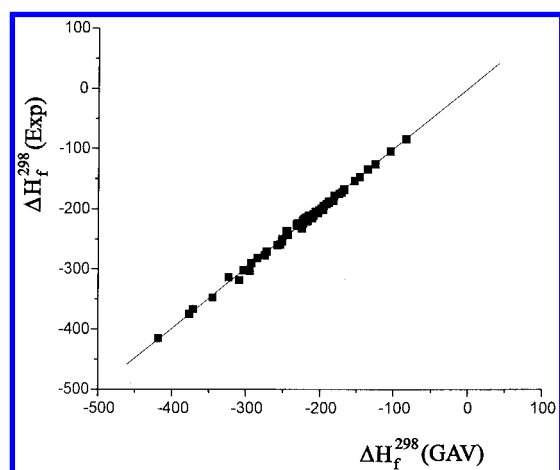
Figures 4–7 show the correlations between the experimental and calculated values. Table 8 contains the statistical data ( $\gamma$ , regression coefficient;  $\rho$ , linear correlation coefficient;  $\sigma$ , root mean-square deviation; AAD, average absolute deviation; and MAD, maximum absolute deviation) obtained by linear regression analysis. It can be seen that the GAV and SEOEM II standard heats of formation are the best among the calculated values. It is worth noting, however, that the SEOEM model can be applied at any temperature without the introduction of new parameters. A comparison of the SEOEM II and I data leads to the



**Figure 4.** Graphical representation of the relation  $\Delta H_f^{298}(\text{Exp})$  vs  $\Delta H_f^{298}(\text{SEOEM I})$  for aliphatic alkanes (kJ/mol).



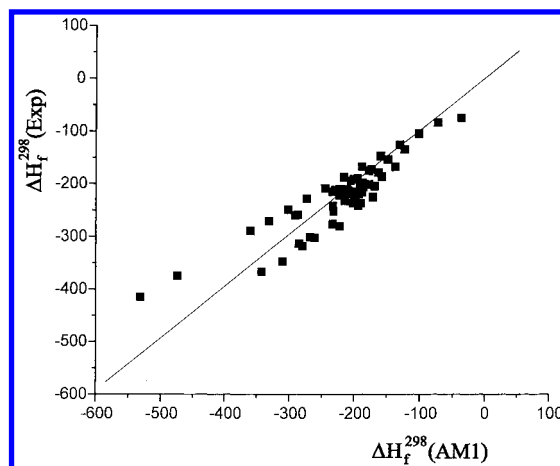
**Figure 5.** Graphical representation of the relation  $\Delta H_f^{298}(\text{Exp})$  vs  $\Delta H_f^{298}(\text{SEOEM II})$  for aliphatic alkanes (kJ/mol).



**Figure 6.** Graphical representation of the relation  $\Delta H_f^{298}(\text{Exp})$  vs  $\Delta H_f^{298}(\text{GAV})$  for aliphatic alkanes (kJ/mol).

conclusion that it is important to perform full conformational analysis and to treat the hindered torsional vibrations as free internal rotations at higher temperatures.

It is concluded that, with the help of the SEOEM model, it is possible to extend the capability of the G2 theory to large molecules which cannot be handled at this level even with present-day supercomputers. The SEHMO program is capable of performing a full conformational analysis on large



**Figure 7.** Graphical representation of the relation  $\Delta H_f^{298}(\text{Exp})$  vs  $\Delta H_f^{298}(\text{AM1})$  for aliphatic alkanes (kJ/mol).

**Table 8.** Statistical Data Obtained by Linear Regression Analyses between the Experimental and Calculated Standard Heats of Formation

methods	$\gamma$	$\rho$	$\sigma/\text{kJ mol}^{-1}$	AAD/kJ mol <sup>-1</sup>	MAD/kJ mol <sup>-1</sup>
SEOEM I	1.00650	0.99976	5.06	4.12	13.10
SEOEM II	0.98706	0.99983	4.25	3.47	15.50
GAV	0.99638	0.99987	3.80	2.80	9.80
AM1	0.98737	0.98889	34.51	26.57	116.70

alkane molecules, and the resulting geometries and total energies of the conformers are reliable.

The source code of the SEHMO program is available from DecaGen Ltd. (decagen@tisznet.hu).

#### ACKNOWLEDGMENT

This work was sponsored by the Agency of Industrial Science and Technology (AIST, MITI). Gyula Tasi performed part of this work as an AIST guest researcher. The authors would like to thank the AIST Computer Center in Tsukuba for the use of their computing facilities.

#### REFERENCES AND NOTES

- (1) Hoffmann R. An Extended Hückel Theory. I. Hydrocarbons. *J. Chem. Phys.* **1963**, 39, 1397–1412.
- (2) Anderson A. B.; Hoffmann R. Description of Diatomic Molecules Using Electron Configuration Energies with Two-body Interactions. *J. Chem. Phys.* **1974**, 60, 4271–4273. Anderson A. B. Derivation of the Extended Hückel Method with Corrections: One Electron Molecular Orbital Theory for Energy Level and Structure Determinations. *J. Chem. Phys.* **1975**, 62, 1187–1188. Anderson, A. B. Vibrational Potentials and Structures in Molecular and Solid Carbon, Silicon, Germanium, and Tin. *J. Chem. Phys.* **1975**, 63, 4430–4436.
- (3) Calzaferri, G.; Forss, L.; Kamber, I. Molecular Geometries by the Extended Hückel Orbital Method. *J. Phys. Chem.* **1989**, 93, 5366–5371. Tupper, K. J.; Counts, R. W.; Gajewski, J. J. Geometry Optimization Within a Modified Extended Hückel Formalism: Modifications to the ASED Program. *Comput. Chem.* **1991**, 15, 157–160. Savary, F.; Weber, J.; Calzaferri, G. Geometry Optimization of Organometallic Compounds Using a Modified Extended-Hückel Formalism. *J. Phys. Chem.* **1993**, 97, 3722–3727. Savary, F.; Furet, E.; Weber, J. Application of a Modified EHMO-ASED Formalism to the Determination of the Structural Parameters of Organometallics. *J. Mol. Struct. (Theochem)* **1995**, 330, 191–195.
- (4) Dixon, S. L.; Jurs, P. C. Fast Geometry Optimization Using a Modified Extended Hückel Method: Results for Molecules Containing H, C, N, O, and F. *J. Comput. Chem.* **1994**, 15, 733–746.
- (5) Tasi, G.; Mizukami, F.; Pálkó, I. A New Program for Effective One-Electron (EHMO-ASED) Calculations. *Comput. Chem.* **1997**, 21, 319–325.

- (6) Nanda, D. N.; Jug, K. SINDO1. A Semiempirical SCF MO Method for Molecular Binding Energy and Geometry. I. Approximations and Parametrization. *Theoret. Chim. Acta (Berl.)* **1980**, *57*, 95–106. Jug, K.; Iffert, R.; Schulz, J. Development of SINDO1 for Second-row Elements. *Int. J. Quantum Chem.* **1987**, *32*, 265–277.
- (7) Dewar, M. J. S.; Thiel, W. J. Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
- (8) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (9) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods. I. Method. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (10) Dewar, M. J. S.; Jie, C.; Yu, J. SAM1: The First of a New Series of General Purpose Quantum Mechanical Molecular Models. *Tetrahedron* **1993**, *49*, 5003–5038.
- (11) Levine, I. N. *Quantum Chemistry*; Allyn and Bacon: Boston, 1983; p 499.
- (12) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 Theory for Molecular Energies of First- and Second-Row Compounds. *J. Chem. Phys.* **1991**, *94*, 7221–7230.
- (13) Durant, J. L.; Rohlfing, C. M. Transition State Structures and Energetics Using Gaussian-2 Theory. *J. Chem. Phys.* **1993**, *98*, 8031–8036.
- (14) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Enthalpies of Formation. *J. Chem. Phys.* **1997**, *106*, 1063–1079.
- (15) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (16) Masamura, M. An Efficient Algorithm for Solving Eigenvalue Problems of the Type  $HC = S C e$  or  $FC = S C e$ . *J. Comput. Chem.* **1988**, *9*, 257–265.
- (17) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision D.4*; Gaussian, Inc.: Pittsburgh, PA, 1995.
- (18) Tasi, G.; Pálkó, I.; Halász, J.; Náray-Szabó, G. *Semiempirical Calculations on Microcomputers*; CheMicro Limited: Budapest, 1992.
- (19) Stewart, J. J. P. MOPAC: A Semiempirical Molecular Orbital Program. *J. Computer-Aided Mol. Des.* **1990**, *4*, 1–105.
- (20) Tsuzuki, S.; Schäfer, L.; Gotō, H.; Jemmis, E. D.; Hosoya, H.; Siam, K.; Tanabe, K.; Ōsawa, E. Investigation of Intramolecular Interactions in n-Alkanes. Cooperative Energy Increments Associated with GG and GTG<sup>+</sup> Sequences. *J. Am. Chem. Soc.* **1991**, *113*, 4665–4671. Frey, R. F.; Cao, M.; Newton, S. Q.; Schäfer, L. Electron Correlation Effects in Aliphatic Non-bonded Interactions: Comparison of n-Alkane MP2 and HF Geometries. *J. Mol. Struct. (Theochem)* **1993**, *285*, 99–113. Smith, G. D.; Jaffe, R. L. Quantum Chemistry Study of Conformational Energies and Rotational Energy Barriers in n-Alkanes. *J. Phys. Chem.* **1996**, *100*, 18718–18724. Gang, J.; Pilling, M. J.; Robertson, S. H. Partition Functions and Densities of States for Butane and Pentane. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 3509–3518. Lahtela, M.; Pakkanen, T. A.; Nissfolk, F. Molecular Dynamics Study of the Conformational Properties of Branched Alkanes. *J. Phys. Chem. A* **1997**, *101*, 5949–5952.
- (21) Goodman, J. M. What is the Longest Unbranched Alkane with a Linear Global Minimum Conformation? *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 876–878.
- (22) Tasi, G.; Mizukami, F.; Pálkó, I. Analysis of Permanent Electric Dipole Moments of Aliphatic Hydrocarbon Molecules. *J. Mol. Struct. (Theochem)* **1997**, *401*, 21–27. Tasi, G.; Mizukami, F.; Pálkó, I. Ab Initio Quantum Chemical Calculations on Aliphatic Hydrocarbon Molecules I: Analysis of Hartree-Fock SCF Permanent Electric Dipole Moments. *Magy. Kem. Foly.* **1997**, *103*, 373–380. Tasi, G.; Mizukami, F. Analysis of Permanent Electric Dipole Moments of Aliphatic Hydrocarbon Molecules, II. DFT Results. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 313–316.
- (23) Pople, J. A.; Scott, A. P.; Wong, M. W.; Radom, L. Scaling Factors for Obtaining Fundamental Vibrational Frequencies and Zero-Point Energies from HF/6-31G\* and MP2/6-31G\* Harmonic Frequencies. *Isr. J. Chem.* **1993**, *33*, 345–350.
- (24) Clementi, E.; Raimondi, D. L. Atomic Screening Constants from SCF Functions. *J. Chem. Phys.* **1963**, *38*, 2686–2689.
- (25) Cohen, N. Revised Group Additivity Values for Enthalpies of Formation (at 298 K) of Carbon-Hydrogen and Carbon-Hydrogen-Oxygen Compounds. *J. Phys. Chem. Ref. Data* **1996**, *25*, 1411–1481.
- (26) Tasi, G.; Mizukami, F.; Pálkó, I.; Csontos, J.; Györfy, W.; Nair, P.; Maeda, K.; Toba, M.; Niwa, S.; Kiyozumi, Y.; Kiricsi, I. Does the Rule 3<sup>rd</sup> Govern the Conformational Behavior of Unbranched Aliphatic Alkanes? Submitted for publication to *Science*.
- (27) Nicolaides, A.; Rauk, A.; Glukhovtsev, M. N.; Radom, L. Heats of Formation from G2, G2(MP2), and G2(MP2,SVP) Total Energies. *J. Phys. Chem.* **1996**, *100*, 17460–17464.

CI980001Z