2003-03-22

Interaction energy surfaces of small hydrocarbon molecules

Richard L. Rowley
rowley@byu.edu

Jukka-Pekka Jalkanen

Tapani A. Pakkanen

Yan Yang

Interaction energy surfaces of small hydrocarbon molecules

Jukka-Pekka Jalkanen and Tapani A. Pakkanen[a]
Department of Chemistry, University of Joensuu, FIN-80101 Joensuu, Finland

Yan Yang and Richard L. Rowley
Department of Chemical Engineering, Brigham Young University, Provo, Utah 84602

Nonbonding interactions of small alkane molecules were studied with *ab initio* methods. Previously reported energy data for ethane and propane dimers were supplemented with 1747 new configurations of dimers involving slightly larger hydrocarbons. The completed work provides interaction energy surfaces for all combinations of dimer pairs involving ethane, propane, isobutane, and neopentane and thus contains information of all chemical groups found in acyclic alkanes. The strongest attraction of the studied molecule pairs was encountered in isobutane C_{2h} dimer, where an energy minimum of -1.784 kcal mol^{-1} at 4.28 Å separation of centermost carbon atoms was observed. The composite data set was fitted with a modified Morse pair potential energy function representing each interatomic C–C, C–H, and H–H interaction for easy transfer to molecular dynamics simulation programs. The new generic parameter set was shown to describe the *ab initio* data for these small alkane molecules with good accuracy. Qualitative comparisons with previously reported potential models were also made, and the relative capability of the models to reproduce quantum-chemical potential energy surfaces was investigated. © *2003 American Institute of Physics.* [DOI: 10.1063/1.1540106]

I. INTRODUCTION

Accurate description of dispersion interactions is challenging because their effects are often masked by stronger Coulombic forces. Noble gases and saturated hydrocarbons are good examples of molecules in which there are no strong polar interactions permitting direct determination of the weaker van der Waals forces. Electron correlation between these molecules produces a temporary fluctuation of charge distribution, which is primarily responsible for the attraction of hydrocarbon molecules. This spatial distribution of electrons changes as other atoms come in close proximity, and it depends on the number and type of interacting atoms. The dispersion energy is therefore dependent on the relative orientation of the interacting molecules, as illustrated in our earlier work[1] on propane dimers. This orientational dependence is missed if the molecular potential surface is probed with a single atom, e.g., with a noble gas atom. While a probe–atom approach reduces the spatial dimensionality of the problem, there is no guarantee that the resulting description of the interactions represents those between two hydrocarbon molecules. We choose instead to use pairs of the test molecules in order to avoid assumptions about the nature of the probe–molecule interaction and its relationship to the true dimer interaction.

Intermolecular phenomena can be studied experimentally, for example, by studying crystal structures[2,3] or using spectroscopic methods (see, e.g., Ref. 4, and references therein). A popular computational approach to this problem has been to use Monte Carlo[5–10] or molecular dynamics simulations[11] to compute macroscopic properties based on pair potential models. In both of these methods, the pair potential parameters are adjusted to reproduce experimental measurements of thermophysical properties or vapor–liquid coexistence data. However, these methods produce averaged or indirect information of effective multibody potentials that can be density dependent. Building an accurate model describing the multidimensional energy surfaces is difficult if based solely on this data. *Ab initio* methods offer a detailed, rigidly controlled approach to study these effects. Using quantum chemistry, one can study the potential energy surfaces of varying sizes at an arbitrary orientation yielding a more comprehensive picture of phenomena arising from intermolecular interactions. In contrast to simulation methods, the quantum-chemistry approach yields a true pair potential consisting only of interactions between two molecules. This provides a clear opportunity to eventually add multibody interactions in a systematic way and eliminate effective potentials that are state dependent.

In the last few years, *ab initio* studies of weak, nonbonding intermolecular potentials of various hydrocarbon molecules have been published. The rapid increase in computer power has facilitated the study of various hydrocarbon systems ranging from small alkanes[1,12–27] and acyclic π-electron systems[19–21,28–32] to aromatic rings up to anthracene,[28,32–39] but most studies of larger molecules are still restricted to highly symmetric orientations in which the use of symmetry decreases the computational cost significantly. In this article we present potential energy surfaces of isobutane and neopentane dimers and their combinations with each other as well as ethane and propane. We extend the work done previously with ethane[25] and propane[1] dimers to cover ten different molecule pairs. Our aim is to create a

[a]Electronic mail: tapani.pakkanen@joensuu.fi

5474

TABLE I. Isobutane structural parameters.

Bond	Bond length (Å)	Angle	Angle (deg)	Dihedral	Dihedral angle (deg)
C3–C1	1.5233	C3–C1–Ha	108.3016	Hb–C3–C1–Ha	60.409
Hc–C3	1.0915	Hb–C3–C1	111.3894	Hc–C3–C1–Ha	180.0
Hb–C3	1.0898	Hc–C3–C1	110.0835	Point group	C3v
Ha–C1	1.0935				

general potential model, which is suitable for description of nonbonding interactions of alkanes. We report 1747 new configurations and regress a simple pair potential model that is able to reproduce our *ab initio* results with good accuracy.

II. COMPUTATIONAL DETAILS

A. Selected data set

Our work concentrated on interactions between small saturated hydrocarbons. In particular, we have computed energies for all the pair combinations of neopentane, isobutane, propane, and ethane (NIPE) molecules as a function of position and relative orientation. We do this with the intent of representing the resultant spatially complex potential energy surface as a sum of simple, spherically symmetric, interatomic potential functions. Each atomic potential center is assumed to be located at the center of the atom and to be unique for each bond environment. Thus, in a $-CH_x$ group (x ranging from 0 to 3) there can be three types of H centers and four types of C centers based on the bond environment. If one excludes induction effects from neighboring atoms, the ten NIPE dimer combinations represent the smallest collection of hydrocarbons that contain complete information about all of the interatomic interactions found in alkanes. A more detailed description of the types of different interactions is described in Fig. 1 of the Supplementary Data section of Ref. 40.

B. Method and basis set

The selection of computational method is crucial for obtaining the most reliable description of intermolecular inter-
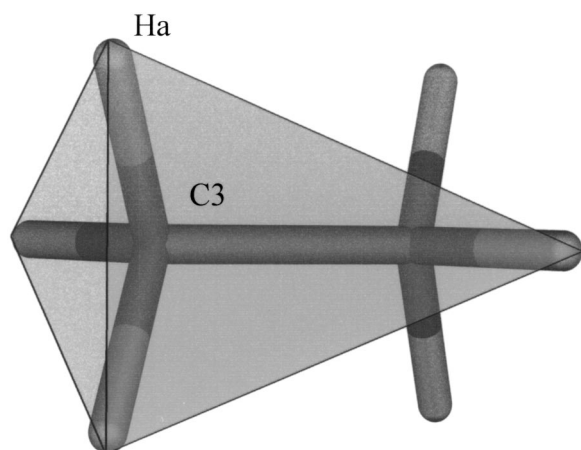

FIG. 1. Ethane (E) monomer labeling. Faces and edges are named after their vertices. The Eaas edge is between the two faces, the Eaal-edge is between hydrogen atoms on the different methyl group.

actions of alkanes.[40] Various authors have reported information on the performance of different computational methods and basis sets. Description of dispersion interactions with Hartree–Fock and density functional methods has been shown to be insufficient,[15,19,20,41,42] and a more rigorous approach is necessary. Capturing the dispersion interactions of hydrocarbons requires the inclusion of electron correlation, the genesis of dispersion attractions. The present work was done with the $6\text{-}311+G(2df,2pd)$ basis set and second order Møller–Plesset perturbation theory (MP2) using both GAUSSIAN 94[43] and GAUSSIAN 98.[44] At the MP2 level, most of the correlation effects are accounted for and introduction of computationally heavier methods (MP4(SDTQ), CCSD(T), CISD) do not significantly affect the intermolecular attraction of alkanes.[19,20,23,26,41] While MP2 produces a good description of alkane intermolecular behavior it reportedly overestimates the attraction of some aromatic systems.[28,38,45,46] However, use of electron correlation introduces a strong dependency between the resultant interaction energy and the basis set size. Several papers address the effect of basis set size on interaction energy used in conjunction with various electron correlation methods.[19–21,23,26,41] Counterpoise correction[47] was applied to all computed dimer interaction energies to account for basis set superposition error.

C. Molecular geometries

Molecular geometries were obtained by running MP2/$6\text{-}311+G(2df,2pd)$ optimization of all hydrocarbon monomers studied. These fixed geometries, electron correlation method, and basis set were used throughout this work for all dimer calculations. Structural properties of isobutane and neopentane are given in Tables I and II. The same ethane and propane structures were used as described in Refs. 1 and 25. For ease of discussion of different dimer orientations, each molecule is divided into vertices, edges, and faces. These geometrical constructs, defined in Figs. 1–4, involve points, lines, and planes defined by the outermost hydrogen atoms in the molecules. The different combinations of faces, edges, and vertices are designated as main routes.

TABLE II. Neopentane structural parameters.

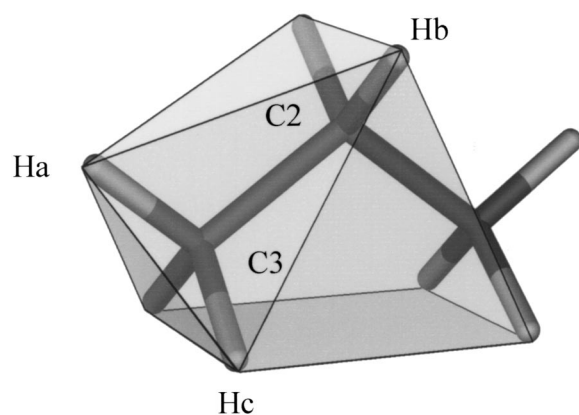| Bond	Bond length (Å)	Angle	Angle (deg)	Dihedral	Dihedral angle (deg)
C3–C0	1.5257	C3–C0–C3	109.4712	C3–C0–C3–C3	120.0
C3–Ha	1.0908	Ha–C3–C0	110.8753	Ha–C3–C0–C3	60.0
				Point group	Td

FIG. 2. Propane (P) monomer labeling. Faces and edges are named after their vertices. Pab edge is between Pabc and Pabb faces, Pac edge between Pabc and Pacc faces, Pbc edge between Pbcc and Pabc faces, Pbb edge is between the two Hb atoms, Pccs-edge is between the two Hc atoms on the same methyl group, and Pccl-edge is between the two Hc atoms on different methyl groups.

FIG. 4. Neopentane (N) monomer labeling. Faces and edges are named after their vertices. Naas-edge is between Ha atoms on the same methyl group, Naal edge is between Ha atoms on different methyl groups.

The first letter of any route name designates the molecule and subsequent letters identify the orientation in terms of these geometrical constructs. In some cases, there is ambiguity between two different planes or edges consisting of the same type of hydrogen atoms. These are differentiated by introducing a letter s for the small plane (or short in the case of an edge) or 1 for the large plane (or long in the case of an edge) (see Figs. 1–4). A complete scan of the potential energy surface would also include intermolecular rotation of one monomer relative to the other along the approach axis of the two molecules. Rotation of monomer about the approach axis has a large effect on dimer interaction energy for some of the routes studied, but little effect on others. Rotation is included on route names with a number, signaling the intermolecular dihedral angle. As an example of the naming system used in this work we consider the route Iabbs-Iabbs 180. This name corresponds to two small isobutane abb-faces interacting, rotated 180° about the intermolecular axis. Intermolecular approach axes intercept vertices along the C–H bond, edges at the center points of the edge, and faces at the
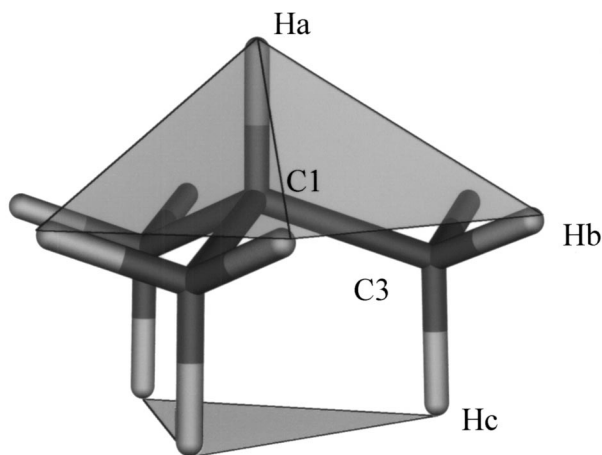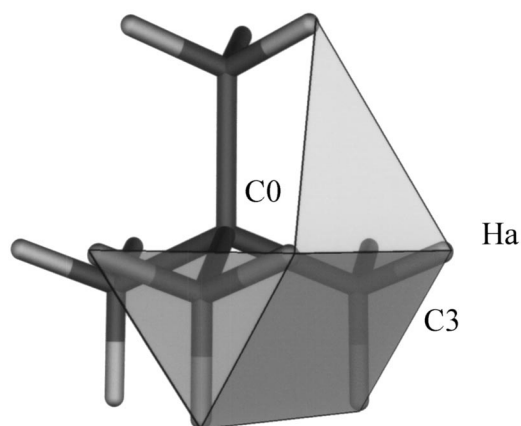
midpoint of the vertices forming the face. Faces and edges are held perpendicular to each other. Intermolecular distances in orientations for propane, isobutane, and neopentane are always measured as a separation of centermost carbons. In the case of ethane, distances are measured relative to the center point of the carbon–carbon bond.

III. RESULTS AND DISCUSSION

Previously reported data for ethane and propane dimer potential energy surfaces were used in conjunction with the results reported here for configurations of larger alkanes. Propane and ethane dimer data reported in Refs. 1 and 25 were used in conjunction with energy data of other combinations of the NIPE set. To simplify the presentation of a large amount of numerical data describing the calculated potential energy surfaces, each dimer route was fitted to the modified Morse potential,

$$E = -\epsilon[1 - \{1 - \exp^{-A(r_{ij} - r^*)}\}^2], \qquad (1)$$

independently of each other. This procedure gives an accurate description of the interaction energy curve as a function of intermolecular distance. The complete numerical data set consisting of all calculated dimer energies at various intermolecular distances is available from the authors upon request. Parameters of this fitting are given in Table III. Epsilon in the second column of Table III indicates the depth of the attractive potential well. A large epsilon indicates a large attractive interaction. Parameter A (third column) characterizes the steepness of the potential energy curve where a large value denotes a steep curve. Parameter r^* (fourth column) identifies the separation distance at the energy minimum of two molecules at that specific orientation. The fifth column shows the quality of the least-squares fit (kcal mol^{-1})^2 of the MP2 data points.

Metzger *et al.*[14] used AM1 and PM3 semiempirical methods to study neopentane and methane dimer interaction energies and compared these to MP2/6-311G($2d$,$2p$) *ab initio* results. Some of the semiempirical methods were found to yield unrealistic behavior as part of the energy curves turned over to less repulsive energies at shorter intermolecu-

FIG. 3. Isobutane (I) monomer labeling. Faces and edges are named after their vertices. Iab-edge is between the Iabbl and Iabbs faces.

TABLE III. Parameters for *ab initio* curves, obtained by doing a least-squares fit to Eq. (1). SSR stands for sum of squared residuals in $(\text{kcal mol}^{-1})^2$.

Route	ϵ (kcal mol^{-1})	A (Å^{-1})	r^* (Å)	SSR (kcal mol^{-1})^2	Route	ϵ (kcal mol^{-1})	A (Å^{-1})	r^* (Å)	SSR (kcal mol^{-1})^2	Route	ϵ (kcal mol^{-1})	A (Å^{-1})	r^* (Å)	SSR (kcal mol^{-1})^2
Pb-Ea	0.317	1.459	5.026	0.002	Naal-Eaal	1.226	1.347	4.606	0.006	Naaaa-Pbc	1.116	1.448	5.077	0.020
Pb-Eaal	0.747	1.484	4.090	0.022	Naal-Eaaas	0.751	1.433	5.497	0.005	Naaaa-Pbb	1.052	1.343	4.601	0.008
Pb-Eaaas	0.518	1.453	4.845	0.002	Naal-Eaaal	0.933	1.409	4.897	0.007	Naaaa-Pb	1.054	1.297	4.774	0.006
Pab-Ea	0.596	1.405	4.486	0.001	Naaal-Eaal	0.946	1.412	4.833	0.009					
Pab-Eaal	1.042	1.477	3.969	0.004	Naaal-Eaaas	0.928	1.358	5.253	0.004	Ia-Ia	0.685	1.401	4.499	0.008
Pab-Eaaas	0.651	1.520	4.756	0.004	Naaal-Eaaal	1.124	1.365	4.695	0.007	Ia-Ia 180	0.694	1.399	4.495	0.008
Pbb-Ea	0.442	1.459	4.624	0.001	Naaaa-Eaal	1.076	1.369	4.729	0.008	Iccc-Iccc 0	0.969	1.443	5.691	0.026
Pbb-Eaal	0.581	1.411	4.136	0.003	Naaaa-Eaaas	0.752	1.439	5.449	0.004	Iccc-Iccc 180	1.298	1.240	5.334	0.004
Pbb-Eaaas	0.522	1.504	4.631	0.003	Naaaa-Eaaal	0.976	1.408	4.875	0.007	Ia-Iccc 0	1.379	1.264	4.541	0.008
Pabb-Ea	0.568	1.358	4.483	0.001						Ia-Iccc 180	1.335	1.224	4.533	0.006
Pabb-Eaal	1.062	1.466	3.910	0.004	Pb-Ia	0.540	1.422	4.562	0.009	Iab-Iab 0	1.135	1.440	4.356	0.017
Pabb-Eaaas	0.681	1.510	4.700	0.004	Pbb-Ia	0.707	1.416	4.182	0.010	Iab-Iab 90	1.458	1.462	4.238	0.024
					Pab-Ia	1.163	1.441	4.112	0.013	Iabbs-Iabbs 0	0.968	1.473	4.558	0.014
Ea-Ia	0.429	1.498	4.931	0.008	Pabb-Ia	0.984	1.393	3.973	0.010	Iabbs-Iabbs 180	1.309	1.570	4.725	0.024
Ea-Iab	0.759	1.405	4.500	0.005	Pb-Iab	0.987	1.407	4.206	0.011	Iabbl-Iabbl 0	1.240	1.483	4.301	0.017
Ea-Iabbs	0.706	1.445	4.655	0.007	Pbb-Iab	0.849	1.408	4.186	0.012	Iabbl-Iabbl 180	1.784	1.493	4.282	0.033
Ea-Iabbl	0.908	1.360	4.335	0.005	Pab-Iab	1.025	1.538	4.578	0.023					
Ea-Iccc	0.723	1.279	5.085	0.002	Pabb-Iab	1.021	1.546	4.257	0.021	Naaaa-Ia	1.348	1.281	4.702	0.009
Eaas-Ia	0.559	1.427	4.871	0.004	Pb-Iabbs	0.858	1.429	4.441	0.011	Naaaa-Iab	1.183	1.410	5.114	0.016
Eaas-Iab	0.687	1.506	4.864	0.010	Pbb-Iabbs	0.838	1.462	4.307	0.010	Naaaa-Iabbs	1.114	1.451	5.207	0.019
Eaas-Iabbs	0.681	1.502	4.970	0.007	Pab-Iabbs	0.954	1.476	4.354	0.014	Naaaa-Iabbl	1.381	1.411	4.985	0.024
Eaas-Iabbl	0.865	1.453	4.689	0.009	Pabb-Iabbs	0.863	1.591	4.510	0.019	Naaal-Ia	1.446	1.268	4.556	0.010
Eaas-Iccc	0.761	1.350	5.320	0.005						Naaal-Iab	1.290	1.368	4.953	0.015
Eaal-Ia	0.910	1.393	4.064	0.007	Naal-Pabc	0.952	1.447	5.274	0.014	Naaal-Iabbs	1.112	1.391	5.197	0.015
Eaal-Iab	1.231	1.558	2.801	0.014	Naal-Pbcc	1.186	1.458	5.053	0.023	Naaal-Iabbl	1.138	1.399	5.110	0.017
Eaal-Iabbs	0.901	1.525	4.430	0.014	Naal-Pabb	1.125	1.360	4.683	0.011	Naal-Ia	1.103	1.262	4.854	0.007
Eaal-Iabbl	1.070	1.505	4.184	0.023	Naal-Pab	0.772	1.458	5.240	0.014	Naal-Iab	1.018	1.430	5.244	0.018
Eaal-Iccc	0.933	1.426	4.816	0.016	Naal-Pbc	0.871	1.437	5.249	0.015	Naal-Iabbs	1.065	1.431	5.285	0.016
Eaaas-Ia	0.655	1.453	4.789	0.006	Naal-Pbb	0.766	1.394	4.947	0.010	Naal-Iabbl	1.301	1.402	5.019	0.021
Eaaas-Iab	0.795	1.481	4.789	0.007	Naal-Pb	0.862	1.285	4.926	0.004					
Eaaas-Iabbs	0.703	1.558	4.969	0.010	Naaal-Pabc	1.018	1.460	5.159	0.020	Naal-Naal 0	0.878	1.351	6.034	0.005
Eaaas-Iabbl	0.840	1.482	4.777	0.011	Naaal-Pbcc	1.133	1.444	4.996	0.023	Naal-Naal 90	1.456	1.228	5.417	
Eaaas-Iccc	0.761	1.383	5.392	0.008	Naaal-Pabb	1.016	1.437	4.777	0.019	Naas-Naas 0	0.603	1.477	6.765	0.005
Eaaal-Ia	1.002	1.409	4.024	0.010	Naaal-Pab	1.040	1.419	4.905	0.015	Naas-Naas 90	0.669	1.449	6.694	0.005
Eaaal-Iab	0.877	1.471	4.345	0.009	Naaal-Pbc	1.197	1.429	4.925	0.022	Naaal-Naaal 0	1.042	1.392	5.741	0.033
Eaaal-Iabbs	0.793	1.540	4.506	0.012	Naaal-Pbb	1.053	1.341	4.602	0.011	Naaal-Naaal 60	1.441	1.234	5.351	0.014
Eaaal-Iabbl	1.009	1.498	4.248	0.018	Naaal-Pb	1.049	1.235	4.650	0.004	Naaas-Naaas 0	0.635	1.520	6.776	0.008
Eaaal-Iccc	0.892	1.434	4.878	0.014	Naaaa-Pabc	0.989	1.428	5.228	0.011	Naaas-Naaas 60	0.683	1.482	6.742	0.021
					Naaaa-Pbcc	1.260	1.404	4.966	0.017	Naaaa-Naaaa 0	1.288	1.329	5.679	0.023
					Naaaa-Pabb	1.008	1.445	4.830	0.016	Naaaa-Naaaa 90	1.014	1.389	5.890	0.026
					Naaaa-Pab	1.170	1.360	4.797	0.010	Na-Na	0.692	1.286	5.333	0.009

lar distances, but the *ab initio* curves did not have this problem. Metzger *et al.* reported an energy minimum of -1.361 kcal mol^{-1} for a neopentane dimer corresponding to our Naaal-Naaal 60 route. Our calculations show a slightly stronger interaction, -1.441 kcal mol^{-1}, due to the larger basis set used in our study. The equilibrium distance of ~ 5.35 Å does agree with our results for this dimer.

The most favorable orientations of each molecule pair studied in this work are illustrated in Figs. 5(a)–5(h), and the interaction energies of these configurations are given in Table IV. The strongest attractions are found for routes, where molecules avoid head-on approaches of atoms. The least attraction occurs for routes, where the closest hydrogen atoms are directly facing each other, as in vertex–vertex routes. The strongest attraction of all the data was found for the C_{2h} symmetric Iabbl-Iabbl 180 route, 1.784 kcal mol^{-1}. According to our calculations, this minimum is reached at intermolecular separation of 4.28 Å. The most attractive orientations for ethane, propane, isobutane, and neopentane paired with neopentane reveal an interesting trend. The deepest minima are -1.226 kcal mol^{-1} (Naal-Eaal), -1.260 kcal mol^{-1} (Naaaa-Pbcc), -1.446 kcal mol^{-1} (Naaal-Ia), and -1.456 kcal mol^{-1} (Naal-Naal 90), respectively. The attraction increases with molecular size, but the increment from isobutane–neopentane to the neopentane dimer is not large, only ~ 0.01 kcal mol^{-1}. The globular shaped neopentane with its protruding methyl groups prevents the closer, more energetically favorable interactions of more open molecules. Although all the neopentane dimer routes studied show attractive potential wells, isobutane interactions with neopentane are more favorable.

A. Fitting the potential energy surfaces with interatomic pair potentials

A major objective of this and our previous calculations of dimer energy surfaces is the representation of alkane intermolecular potentials as a sum of atomic pair potentials. The hope is that these interatomic pair potentials are transferable and can then be used to adequately describe the intermolecular interactions in molecular dynamics simulations. Several different types of potential energy functions have been investigated previously ranging from a simple Lennard-Jones 12-6 model to more complicated functions. These previous studies have shown that a Morse-type function can be used to accurately describe the interactions at the repulsive region as well as the bottom of the potential well.[12,25,48,49] A Morse-type function has good flexibility as the slope of the repulsion can be adjusted, instead of having a fixed repulsive exponent as in the case of the Lennard-Jones-type function.

To obtain a generic parameter set for alkanes, a sum of pair potentials each in the form of Eq. (1) were fitted to our *ab initio* data with an unweighted least-squares method. A simulated annealing algorithm by Goffe *et al.* was used for this purpose. Details of the algorithm can be found elsewhere.[1,50] Five different fitting runs were conducted, and each of these was independent of the others using random starting points and different initial search directions. The generic parameter set obtained for small alkanes is given in Table V, along with previously reported ethane[25] and

propane[1] parameter sets. Repulsive interaction energies larger than ~ 3 kcal mol^{-1} were not included in the regression to avoid fitting the repulsive side of the potential at the expense of the attractive well. Although simulated annealing is a stochastic global optimization method and covers only a part of the total parameter space, all five runs converged to the same solution. Only slight variations, $\sim 1 \times 10^{-5}$ kcal mol^{-1} in ϵ_{CC} and $\sim 2 \times 10^{-5}$ Å^{-1} in A_{CC}, were observed between the different fitting runs. These variations of parameters produced a sum of squared residuals (SSR) of $\sim 3 \times 10^{-8}$ (kcal mol^{-1})^2 for the overall fit. We also found that a slight change in the carbon–carbon interaction parameters had a relatively small change in overall quality of the fit, because of the small number of interactions between carbon atoms in the dimers compared to the number of other type interactions. For example, the neopentane dimer has 25 carbon–carbon, 144 hydrogen–hydrogen, and 120 carbon–hydrogen interactions. The parameter ϵ was constrained to be greater than zero to ensure that potentials would not turn over at short distances.

The overall accuracy of the generic parameter set is excellent, yielding a total error of 50.354 (kcal mol^{-1})^2 and an average of 0.015 (kcal mol^{-1})^2 per data point. This result for the ten different pairs of molecules studied is remarkable since all of the potential energy surfaces have been reduced to three interatomic pair potentials: C–C, C–H, and H–H. (A more detailed description of the error sum on any particular main route studied is presented in Table I of the Supplementary Data section of Ref. 40). Table VI shows the fitting error for each pair and illustrates how the error is distributed between different combinations of molecules. A slight increase in the error is observed for larger systems. This seems consistent, since even the smallest errors in the regressed parameters are bound to accumulate with increasing numbers of interactions. Furthermore, it can be seen from Table VI that the error for the repulsive data points constitute almost two-thirds of the overall SSR, even though the majority of the data points are attractive.

In general, the previously reported ethane dimer data[25] is reproduced better with the generic set than with parameters regressed from the propane data. An error of 8.19 (kcal mol^{-1})^2 was obtained when the propane-specific parameters were used to predict the ethane data, but the error was reduced to 4.5 (kcal mol^{-1})^2 when the generic parameter set was used. These results correspond to 0.021 and 0.011 (kcal mol^{-1})^2 error per data point, respectively. Only three routes out of 240 give a SSR per data point over 0.1 (kcal mol^{-1})^2. The largest error per data point is encountered in the propane dimer route Pccs-Pcccc (0.165 kcal mol^{-1})^2. In this case, the predicted potential well is deeper and the repulsion starts at slightly closer distances than the corresponding *ab initio* results. A similar behavior is observed also for routes Pcccc-Pcccc and Naaaa-Naaaa 0, i.e., the predicted attraction is stronger and the repulsion gentler than the *ab initio* data. It seems that the largest errors are concentrated on routes, where non-CH$_3$ groups interact closely. The

J. Chem. Phys., Vol. 118, No. 12, 22 March 2003

Energy surfaces of hydrocarbons   5479

(a) Pabb-Eaal

(b) Eaal-Iab

(c) Naal-Eaal

(d) Pab-Ia

(e) Iabbl-Iabbl

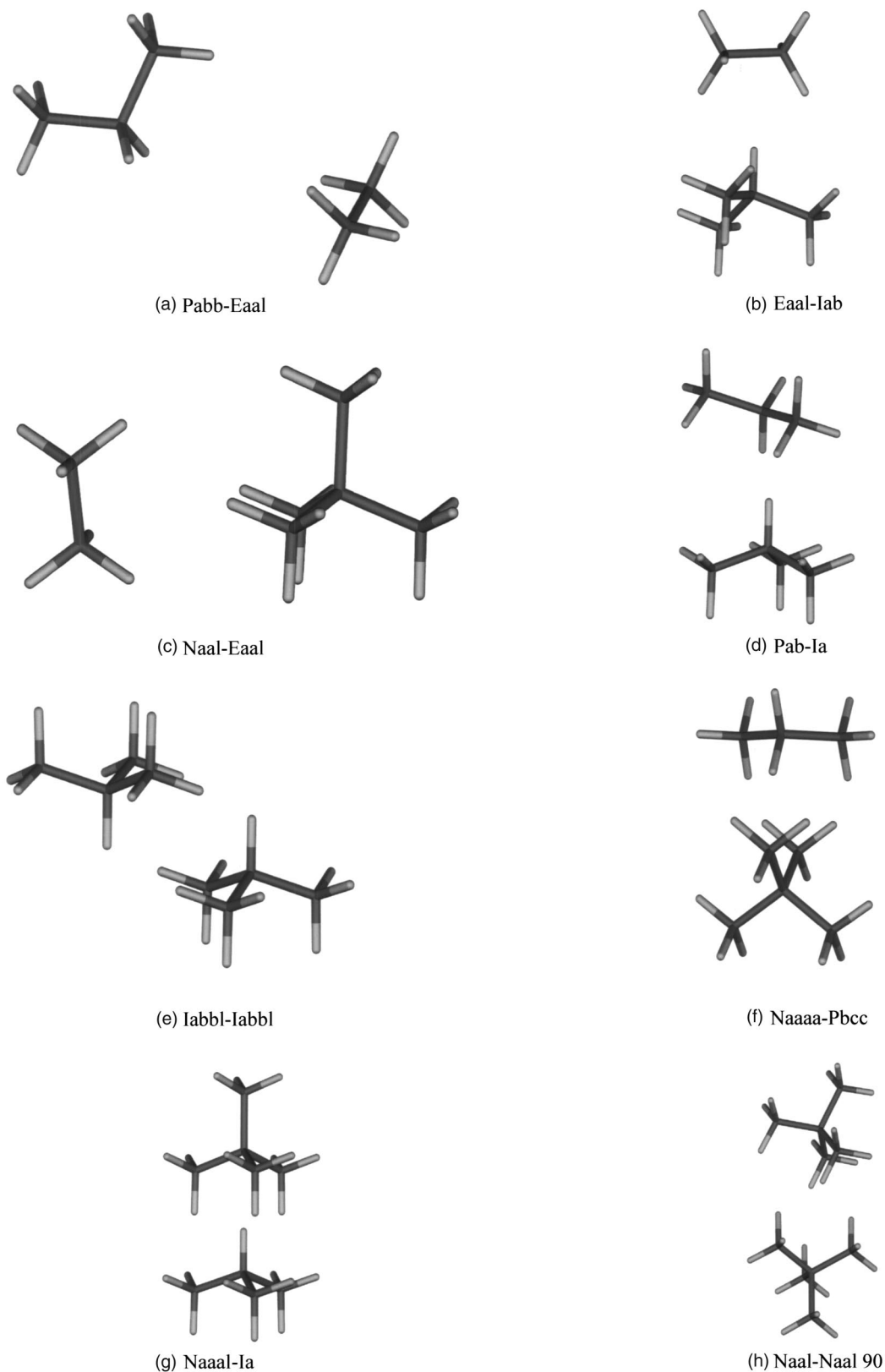(f) Naaaa-Pbcc

(g) Naaal-Ia

(h) Naal-Naal 90

FIG. 5. (a)–(h) Most attractive orientations of the molecule pairs. The first letter identifies the molecule and the subsequent letters identify the vertex, edge, or face that is held perpendicular to the approach axis.

worst, average, and best-fitted cases of all the calculated al-kane routes are depicted in Fig. 6. As can be seen from Fig. 6, the overall accuracy of the fit is good. The generic parameter set was also used to predict ethane dimer repulsive en-ergies up to several hundred kilocalories. In over half of the repulsive ethane dimer routes it correctly described the dimer behavior, even though energies this high were intentionally excluded from the fitting process.

TABLE IV. Most favorable orientations of the studied configurations. Well depth in kcal mol^{-1}. Capital letters designate molecule pair; E=ethane, P=propane, I=isobutane, N=neopentane. Route column identifies orientation of molecules.

Molecule pair	Route	Well depth (kcal mol^{-1})
EE[a]	Route 19	1.046
EI	Eaal-Iab	1.231
EP	Pabb-Eaal 90	1.062
EN	Naal-Eaal	1.226
IP	Pab-Ia	1.163
II	Iabbl-Iabbl	1.784
IN	Naaal-Ia	1.446
PP[b]	Pbcc-Pbcc 180	1.625
PN	Naaaa-Pbcc	1.260
NN	Naal-Naal 90	1.456

[a]Reference 25.
[b]Reference 1.

TABLE VI. Fitting error ((in kcal mol^{-1})^2) for all molecule pairs. Capital letters identify the interacting molecule pair (E=ethane, P=propane, I=isobutane, N=neopentane).

System	No. data points	SSR, all (kcal mol^{-1})^2	SSR, attractive points only (kcal mol^{-1})^2	SSR all, per data point (kcal mol^{-1})^2
EE	395	4.502	1.465	0.011
EP	130	2.355	0.835	0.018
EI	397	3.032	1.045	0.008
EN	129	1.429	0.382	0.011
PP	1245	20.293	6.521	0.016
PI	206	1.078	0.267	0.005
PN	330	4.748	1.570	0.014
II	178	3.221	0.996	0.018
IN	175	3.677	1.361	0.021
NN	202	6.018	2.401	0.030
Sum	**3387**	**50.354**	**17.243**	**Average**: 0.015

B. Comparison to previous parameter sets

Parameter sets based only on ethane or propane dimer data were published recently by us.[1,25] In both of these cases it was found that pair potentials described by the modified Morse function were able to produce energy curves that were in good agreement with *ab initio* data. The data used to regress the generic parameter set presented in this work included data from our previous work on ethane and propane reported in Refs. 25 and 1. A graphical representation of the different parameter set results is given in Figs. 7(a)–7(c). Locations of potential energy minima on the *x* axis do not differ much between the Morse sets. Larger differences are observable for the *A* and ϵ parameters, which describe the repulsive slope and potential well depth, respectively. The generic carbon–carbon interaction is very similar to that obtained from ethane. The C–C potential obtained from propane data is the most attractive of the three Morse potentials. The slope of the H–H repulsion is also very similar for all three Morse parameter sets and the size parameter varies only slightly in the order Er$_{HH}$>Pr$_{HH}$>Gr$_{HH}$. The generic set H–H coincides with the optimized potentials for liquid simulations[8] (OPLS) H–H at ~2 Å separation, but the potential is less attractive at larger distances and less repulsive at shorter distances resulting in "softer" hydrogens. In general, the modified Morse model does not produce as steep a repulsion as does the Lennard-Jones 12-6 potential function.

Cross interactions between carbon and hydrogen atoms are the major source of attraction between alkanes according to all three Morse sets. The ethane parameter set shows the strongest C–H attraction, but the propane and generic parameters both show considerable attraction. The most striking difference between the models compared here is the difference between the Morse models and the OPLS set which uses a geometric average of the energy parameter and an arithmetic average of the size parameter of the C–C and H–H interactions to obtain the C–H parameters in the Lennard-Jones 12-6 potential. As can be seen in Figs. 7(a)–7(c), the generic modified Morse CH is closer to the OPLS C–H potential than the ethane or propane dimer set, but it is still much more attractive at shorter equilibrium distances than the OPLS model potential. Introduction of the CH$_2$ group into the regression through the propane dimer slightly decreases the depth of the attractive well. This might be due to the increased appearance of non-CH$_3$ features in the *ab initio* data and suggests that even better agreement between the sum of atomic pair interactions and the molecular dimer potential energy surface might be achieved if cross C–H interactions between different CH$_x$ groups were to be treated differently, i.e, if the methyl–methylene C–H interaction was considered to be different than the methyl–methyl C–H interaction.

C. Comparison to some other potential models

Some qualitative comparisons of potential energy surfaces produced by the OPLS and *ab initio* calculations were made to give a general idea of the performance of modified Morse model. Energies of all alkane data points were calculated with the OPLS parameter sets. Both united atom (UA)

TABLE V. Parameter sets for different alkanes. Set labeled as generic includes *ab initio* data from ethane, propane, isobutane, neopentane, and all their combinations.

Set	ϵ_{CC} (kcal mol^{-1})	A_{CC} (Å^{-1})	r^*_{CC} (Å)	ϵ_{CH} (kcal mol^{-1})	A_{CH} (Å^{-1})	r^*_{CH} (Å)	ϵ_{HH} (kcal mol^{-1})	A_{HH} (Å^{-1})	r^*_{HH} (Å)
Ethane dimer[a]	0.068 9	1.308 3	4.350 0	0.740 9	1.828 7	2.491	−40.65	2.223 9	0.2456
Propane dimer[b]	0.161 05	1.265 5	4.184 4	0.551 62	2.274 4	2.544	4.5284×10^{-5}	1.255	6.1543
Generic	0.051 33	1.459 85	4.341 17	0.355 62	2.111 74	2.602 11	0.010 48	1.260 72	3.975 36

[a]Reference 25.
[b]Reference 1.

J. Chem. Phys., Vol. 118, No. 12, 22 March 2003

Energy surfaces of hydrocarbons   5481
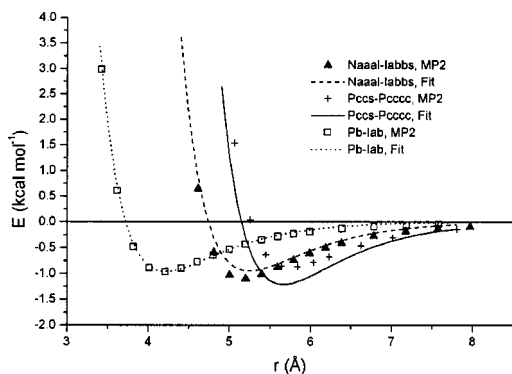

FIG. 6. Quality of fitted potential energy curves. The propane dimer Pccs-Pcccc shows the largest fitting error [SSR=0.165 (kcal mol^{-1})^2 per data point]. The neopentane–isobutane route Naaal-Iabbs represents an average [SSR=0.015 (kcal mol^{-1})^2 per data point] error case and the propane–isobutane route Pb-Iab is the most accurately (SSR=0.000 47 kcal mol^{-1})^2 predicted route. Symbols (diamonds=Pccs-Pcccc, triangles=Naaal-Iabbs, squares=Pb-Iab) represent *ab initio* MP2 energies and lines (dashed =Naaal-Iabbs, solid=Pccs-Pcccc, dotted=Pb-Iab) represent energies calculated with the modified-Morse potential model for pairwise interatomic interactions with the generic parameter set.

and all atom (AA) models are based on Monte Carlo simulations of various hydrocarbon properties. The heterogeneous OPLS UA model by Jorgensen *et al.*[5] was calibrated to reproduce heats of vaporization and density data for 15 different hydrocarbon liquids consisting of normal and branched alkanes as well as some aromatic and unsaturated molecules. A more detailed OPLS AA model by Kaminski *et al.*[8] was regressed from Monte Carlo simulation data of methane, ethane, propane, and butane. Simulated properties included heat of vaporization, molecular volume, heat capacity, the coefficient of thermal expansion, isothermal compressibility, and the C–C radial distribution function yielding a good overall estimate of nonbonding properties. Both OPLS models use combining rules to describe the cross interaction between unlike atoms or sites. Both of these parameter sets represent effective pair potentials owing to the use of liquid-phase simulations in regression of these sets.

Despite the different approach in potential model construction, potential well depths are reasonably well predicted with the AA model. In many cases, the UA model characterizes the attractive well as too broad and misplaced, although the rough strength of attraction is approximately correct. However, the repulsive side of the potential energy curves is problematic for both models, especially for the UA model. The OPLS-UA model seems to have problems with branched molecules, like isobutane and neopentane, despite the fact that these molecules were part of the original training set used by Jorgensen *et al.*[5] In these cases, some of the routes are correctly described, while the bottom of the potential well is too broad. This occurs especially on the routes where two hydrogen atoms are brought close together (e.g., Na-Na, Pa-Pa, or Ia-Ia). This is understandable, because by definition the UA model does not consider hydrogen atoms separately from the CH$_x$ group.
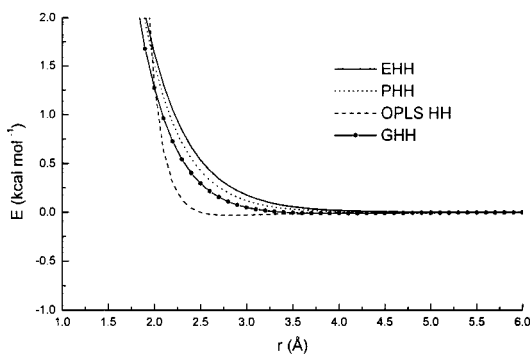
The AA model is definitely an improvement, but the OPLS AA model exhibits too strong a repulsion on some of the neopentane dimer routes, while the attractive potentials
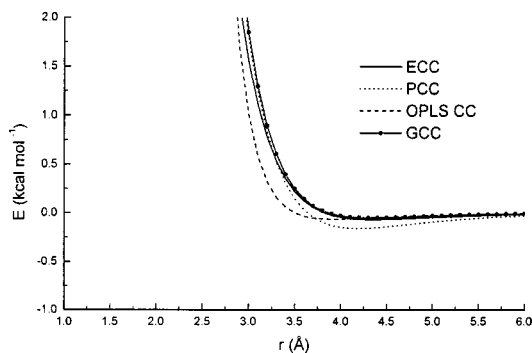
(a) Interatomic interactions for C-C
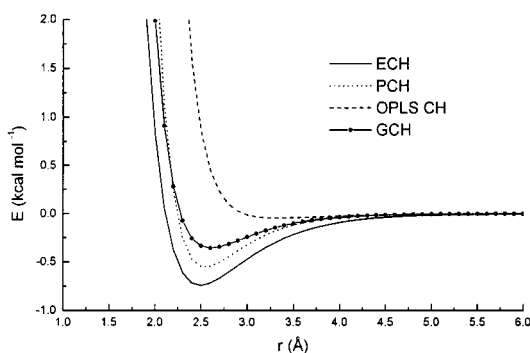
(b) Interatomic interactions for C-H

(c) Interatomic interactions for H-H

FIG. 7. Interatomic interactions for (a) C–C, (b) C–H, and (c) H–H from different potential models. E=ethane Morse, P=propane Morse, G=generic Morse (NIPE).

are approximately correct, as can be seen in Figs. 8(a)–8(d). Excessive repulsion was observed also with other molecule pairs. Our *ab initio* results are fairly close to the energy curves produced by the AA model. In general, potential well depths predicted by the AA model are in good agreement with the MP2 energies, but it is surprising to see that in some extreme cases well depths are 80% larger than the *ab initio* results. This is a considerably larger difference than the gap between the MP2$_{\text{limit}}$ and our calculated values.[26] Neverthe-
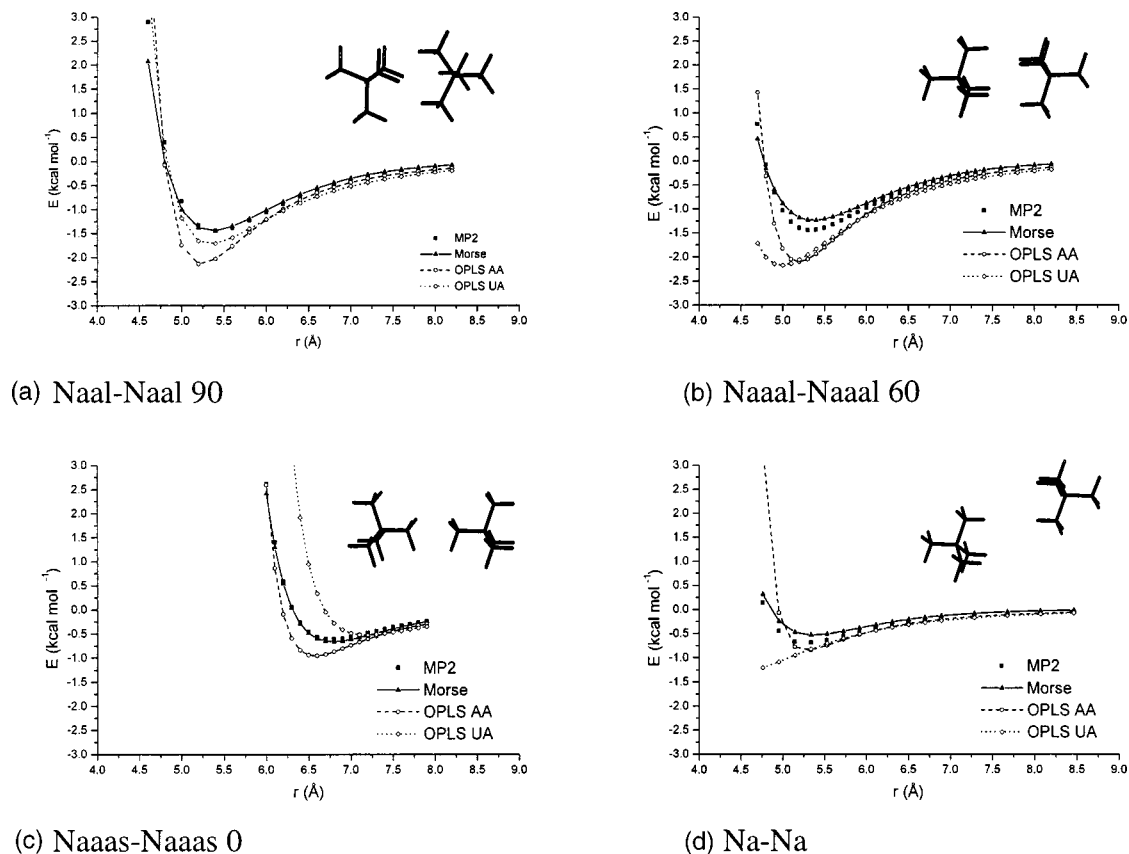
FIG. 8. (a)–(d) Neopentane dimer potential energy surfaces from various potential models. Symbols (diamonds) represent *ab initio* MP2/6-311 +G(2*df*,2*pd*) energies, solid lines represent the energies from the pairwise additive modified Morse potential model presented here, dashed lines represent the OPLS-UA model, and dotted lines represent the OPLS-AA model.

less, the relatively good agreement of our pair potentials with empirically regressed potential models from simulations suggests that our generic pair potential models should provide good representation of the macroscopic properties when used in simulations.

IV. CONCLUSIONS

We have shown that *ab initio* potential energy surfaces of various small alkanes can be fitted accurately with a sum of interatomic pair potentials using a modified Morse function for each of the three constituent types of interatomic interactions. While the H–H interactions are nearly independent of the dimer data used in the regression, the C–C and C–H interactions changed somewhat when data from different dimers were used in the regression. The potential energy surface for any molecule pair can be fitted more accurately independently, but simultaneous regression of parameters for the complete set of compounds studied here provides better transferability to other alkane systems. The most striking difference of the regressed modified-Morse potential models from the empirical OPLS-AA model is for the C–H cross interaction. This is because the modified-Morse potential model presented here does not use an *ad hoc* relationship between like and cross parameters.

Our data set is, in its present state, heavily dominated by methyl group data and this is bound to influence the regressed parameter set. Adding more data for alkanes having

non-CH$_3$ groups in fact does slightly alter the strength of the C–H cross interaction in particular. Comparisons of potential energy surfaces generated by the OPLS AA model with those from the *ab initio* calculations reveal that OPLS AA potential energy surfaces are, for the most part, consistent with *ab initio* results. Performance of our potential model in condensed phase and fine-tuning of the potential model for larger molecules are left for future work. Nevertheless, the ability of the generic parameter set to describe all of the potential energy surfaces for the ten molecular pairs of small alkanes discussed here is very encouraging.

[1] J.-P. Jalkanen, R. Mahlanen, T. A. Pakkanen, and R. L. Rowley, J. Chem. Phys. **116**, 1303 (2002).
[2] D. E. Williams, J. Chem. Phys. **47**, 4680 (1967).
[3] F. A. Momany, L. M. Carruthers, R. F. McGuire, and H. A. Scheraga, J. Phys. Chem. **78**, 1595 (1974).
[4] A. J. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, New York, 1996), p. 185.
[5] W. L. Jorgensen, J. D. Madura, and C. J. Swenson, J. Am. Chem. Soc. **106**, 6638 (1984).
[6] A. Poncela, A. M. Rubio, and J. J. Freire, Mol. Phys. **91**, 189 (1997).
[7] J. I. Siepmann, M. G. Martin, C. J. Mundy, and M. L. Klein, Mol. Phys. **90**, 687 (1997).
[8] G. Kaminski, E. M. Duffy, T. Matsui, and W. L. Jorgensen, J. Phys. Chem. **98**, 13077 (1994).
[9] W. L. Jørgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).
[10] S. K. Nath and R. Khare, J. Chem. Phys. **115**, 10837 (2001).
[11] H. Sun, J. Phys. Chem. B **102**, 7338 (1998).
[12] R. L. Rowley and T. A. Pakkanen, J. Chem. Phys. **110**, 3368 (1999).

[13] D. H. Gay, H. Dai, and D. R. Beck, J. Chem. Phys. **95**, 9106 (1991).

[14] T. G. Metzger, D. M. Ferguson, and W. A. Glauser, J. Comput. Chem. **18**, 70 (1997).

[15] A. K. Rappe and E. R. Bernstein, J. Phys. Chem. A **104**, 6117 (2000).

[16] S. Tsuzuki, T. Uchimaru, and K. Tanabe, Chem. Phys. Lett. **287**, 327 (1998).

[17] E. Fraschini and A. J. Stone, J. Comput. Chem. **19**, 847 (1998).

[18] J. J. Novoya, M.-H. Whangbo, and J. M. Williams, J. Chem. Phys. **94**, 4835 (1991).

[19] S. Tsuzuki, T. Uchimaru, and K. Tanabe, Chem. Phys. Lett. **287**, 202 (1998).

[20] S. Tsuzuki, T. Uchimaru, and K. Tanabe, J. Mol. Struct.: THEOCHEM **307**, 107 (1994).

[21] S. Tsuzuki, T. Uchimaru, M. Mikami, and K. Tanabe, J. Phys. Chem. A **102**, 2091 (1998).

[22] S. Tsuzuki, T. Uchimaru, K. Tanabe, and S. Kuwajima, J. Phys. Chem. A **98**, 1830 (1994).

[23] S. Tsuzuki and K. Tanabe, J. Phys. Chem. **95**, 2272 (1991).

[24] D. E. Williams and D. J. Craycroft, J. Phys. Chem. **91**, 6365 (1987).

[25] R. L. Rowley, Y. Yang, and T. A. Pakkanen, J. Chem. Phys. **114**, 6058 (2001).

[26] S. Tsuzuki, T. Uchimaru, M. Mikami, and K. Tanabe, J. Phys. Chem. A **106**, 3867 (2002).

[27] S. Gupta, J.-A. Yang, and N. R. Kestner, J. Chem. Phys. **89**, 3733 (1988).

[28] S. Tsuzuki, T. Uchimaru, K. Matsumura, M. Mikami, and K. Tanabe, Chem. Phys. Lett. **319**, 547 (2000).

[29] S. Tsuzuki, T. Uchimaru, M. Mikami, and K. Tanabe, Chem. Phys. Lett. **252**, 206 (1996).

[30] S. Tsuzuki and K. Tanabe, J. Phys. Chem. **96**, 10804 (1992).

[31] K. Shuler and C. E. Dykstra, J. Phys. Chem. A **104**, 4562 (2000).

[32] K. S. Kim, P. Tarakeshwar, and J. Y. Lee, Chem. Rev. **100**, 4145 (2000).

[33] V. Spirko, O. Engkvist, P. Soldan, H. L. Selzle, E. W. Schlag, and P. Hobza, J. Chem. Phys. **111**, 572 (1999).

[34] P. Hobza, H. L. Selzle, and E. W. Schlag, J. Am. Chem. Soc. **116**, 3500 (1994).

[35] P. Hobza, H. L. Selzle, and E. W. Schlag, J. Phys. Chem. **97**, 3937 (1993).

[36] P. Hobza, H. L. Selzle, and E. W. Schlag, J. Chem. Phys. **93**, 5893 (1990).

[37] C. Gonzalez and E. C. Lim, J. Phys. Chem. A **104**, 2953 (2000).

[38] P. Hobza, H. L. Selzle, and E. W. Schlag, J. Phys. Chem. **100**, 18790 (1996).

[39] S. Tsuzuki, T. Uchimaru, M. Mikami, and K. Tanabe, Chem. Phys. Lett. **252**, 206 (1996).

[40] See EPAPS Document No. E-JCPSA6-118-001309 for a more detailed explanation for selected molecules and error analysis. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory epaps/. See the EPAPS homepage for more information.

[41] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, and K. Tanabe, J. Phys. Chem. A **103**, 8265 (1999).

[42] W. Koch, *Chemist's Guide to Density Functional Theory* (Wiley VCH, Weinheim, 2000).

[43] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 94, Revision E.2, Gaussian, Inc., Pittsburgh, PA, 1995.

[44] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 98, Revision A.7, Gaussian, Inc., Pittsburgh, PA, 1998.

[45] J. Šponer and P. Hobza, Chem. Phys. Lett. **267**, 263 (1997).

[46] R. L. Jaffe and G. D. Smith, J. Chem. Phys. **105**, 2780 (1996).

[47] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).

[48] P. Blowers and R. I. Masel, J. Phys. Chem. A **103**, 7725 (1999).

[49] J. R. Hart and A. K. Rappe, J. Chem. Phys. **97**, 1109 (1992).

[50] W. L. Goffe, G. D. Ferrier, and J. Rogers, J. Econometr. **60**, 65 (1994).