2001-04-08

Determination of an ethane intermolecular potential model for use in molecular simulations from ab initio calculations

Richard L. Rowley
rowley@byu.edu

Tapani A. Pakkanen

See next page for additional authors

Original Publication Citation
R.L. Rowley, Y. Yang, and T.A. Pakkanen, "Determination of an ethane intermolecular potential model for use in molecular simulations from ab initio calculations", J. Chem. Phys. 114, 659 (21)

Determination of an ethane intermolecular potential model for use in molecular simulations from *ab initio* calculations

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

Tapani A. Pakkanen
Department of Chemistry, University of Joensuu, FIN-80101 Joensuu, Finland

(Received 8 December 2000; accepted 25 January 2001)

Counterpoise-corrected, supermolecule, *ab initio* energies obtained at the MP2/6-311+G($2df$,$2pd$) level were computed for 22 different relative orientations of two ethane molecules as a function of the separation distance between the molecular centers. These energies were used to regress the parameters in several simple, analytical, interatomic or site–site models that can be used for implementation in molecular simulations. Sensitivity analysis indicates that the intermolecular potential surface is insensitive to C–C interactions and that the parameters in the C–C model are coupled and unobtainable from the dimer energies. Representation of the potential surface can be made in terms of C–H and H–H interatomic potentials if the C–C interactions are treated as shielded. Simple Lennard-Jones and exp-6 models do not adequately represent the potential surface using these shielded models, nor do they produce the anticipated physics for the interatomic potentials. The exp-6 model with a damping function and the modified-Morse interatomic potentials both reproduce the intermolecular potential surface well with physically realistic intersite potentials suitable for use in molecular dynamics simulations. © *2001 American Institute of Physics.*
[DOI: 10.1063/1.1356003]

I. INTRODUCTION

The accuracy of molecular dynamics (MD) simulations for real fluids is primarily limited by the efficacy of the potential models used to model the fluid. Current MD models are generally of the force-field variety with the potential represented as a sum of intra- and intermolecular potentials.

Two major assumptions are commonly used to simplify the total potential: pairwise additivity and the use of site–site interactions. Pairwise additivity assumes that the potential energy of molecule m is adequately approximated by a sum of isolated pair energies. Thus,

$$U_m = \sum_{n \neq m}^{N} U_{mn}, \tag{1}$$

where N is the number of molecules. This assumption permits parameterization of the potential in terms of the relative coordinates of only two molecules, but it neglects multibody effects.

Neglect of multibody effects is usually partially compensated for by the use of empirical parameters in the pair-potential model. Therefore, even though multibody effects may be important for condensed-phase simulations, errors due to multibody effects may not be apparent if the pair parameters have been tuned with experimental data at about the same density. While the use of empirical parameters permits prediction accuracy exceeding the inherent limitations of the model, it may also restrict the efficacious use of the model to densities and properties that are similarly affected by this compensation of model inadequacy with adjusted parameters.

The second common assumption, site–site additivity, assumes that the molecular pair can be further represented as a sum of potentials between interacting sites, often atomic centers, located within the molecules. Within this approximation, the isolated pair potential between molecules m and n can be represented by

$$U_{mn}(r,\omega) = \sum_{i=1}^{I} \sum_{j=1}^{J} u_{mn}^{ij}(r), \tag{2}$$

where u_{mn}^{ij} is the potential energy between site i on molecule m and site j on molecule n and I and J are the total number of sites on m and n, respectively. We use here a lower case u for interatomic or site interactions and an upper case for molecular interactions. Such potential models are particularly convenient for molecular simulations because the angle dependence of the model is included implicitly through the intersite distances and their distribution within the molecules. This permits retention of mathematically simple, spherically-symmetrical models for the intersite potentials.

The site–site assumption also gives rise to a powerful concept of transferrable intersite potentials,[1,2] wherein model parameters are tuned for specific atomic or group (e.g., CH_x) interactions based on limited experimental data (e.g., densities, heats of vaporization, dipole moment, etc.) for a training set of compounds that contain the specific sites. These site parameters are then assumed to be transferrable to all molecules that contain the site. The power of the transferrable site potential approach is that tabulated site parameters obtained from a training set of compounds can be used in predictive simulations for compounds not included in the train-

[a]Author to whom correspondence should be addressed.

ing set. Limitations of the approach include those previously mentioned regarding the use of experimentally regressed parameters as well as inherent lack of transferability due to different electronic environments for bonded sites with different neighboring sites.

The use of experimental data to regress model parameters, while improving the agreement between simulated and experimental properties, generally provides little insight as to how the model inadequacies can be improved and may even confuse the issue as to how rigorous model corrections may be applied. An alternative approach is to obtain parameters for the true isolated pair potential. Even though condensed-phase simulations using true pair potentials are not expected to be as accurate as those using potentials tuned with experimental data, there are numerous advantages to this approach. Foremost is consistency with theory, thereby facilitating model improvement. Equation (1) can be viewed as a truncation of a multibody expansion. If true pair parameters are utilized, then additional terms in the expansion can be included as needed. For example, Rowley and Pakkanen[3] (RP) used *ab initio* calculations to evaluate three-, four-, and five-body interactions for condensed-phase methane. Secondly, the use of true pair potentials may give better consistency between simulated properties. Thirdly, because of the more rigorous tie to theory, it is hoped that site–site pair potentials will be more transferrable than their empirically deduced counterparts. Finally, pair potentials can be determined directly from *ab initio* potentials, avoiding the difficulties associated with the inverse problem of regressing potential parameters from macroscopic property data.

We report here a continuation of the work reported in RP. In RP, counterpoise corrected (CPC) methane dimer potentials calculated using MP2/6-311G($2df$,$2pd$) were obtained using the supermolecule approach. We report here similar calculations for the dimer potential of ethane. We plan similar calculations for *n*-propane, isobutane, and neopentane to examine the transferability of the atomic site potentials to different molecules and to obtain a complete set of atomic intersite potentials for different CH_x environments.

II. *AB INITIO* CALCULATION OF INTERMOLECULAR POTENTIAL

A. Background

Determination of intermolecular potentials that include dispersion potentials directly from *ab initio* calculations on a supermolecule has become more common due to software and hardware capabilities in handling electron correlation with perturbation theory and large basis sets. Woon[4] showed that CPC supermolecule potentials calculated with MP4/aug-cc-pVQZ were in excellent agreement with experimental data for noble gases. The effect of basis set size and level of theory were examined, and it was found that MP2/6-311G($2df$,$2pd$) still produced reasonably good results. *Ab initio* calculations of potential surfaces between noble gases and a few multiatomic molecules were also reported. Tao *et al.*[5] calculated the potential surface of H_2O–He; Hu and Thakker[6] calculated the potential energy surface for interactions between N_2 and He; Hill[7] calculated the Ne–CH_4 potential for several orientations; and Marshall *et al.*[8] calculated the CO_2–Ar potential. Most of these calculations were done at the MP2 or MP4 levels with correlation-consistent basis sets. A few other more complex intermolecular potentials have also been studied. The CO_2 dimer was calculated by Tsuzuki *et al.*[9] using MP2/6-311+G($2df$); Shen *et al.*[10] calculated potentials for CO_2–benzene using MP2/6-31G*, and Soetens *et al.*[11] developed a potential model for CCl_4 by obtaining Coulombic and induction terms from monomer calculations and dispersion terms from MP2/aug-cc-pVDZ calculations for the dimer. Tsuzuki and co-workers have been particularly active in studying the intermolecular interaction potentials between hydrocarbons,[12–16] benzene,[17] hydrogen-bonding complexes,[18] and even larger molecules.[19] The methane dimer potential was calculated for four different orientations by Metzger *et al.*[20] using MP2/6-311G($2d$,$2p$); Novoa *et al.*[21] used MP2 with various smaller- to moderately-sized basis sets; RP used MP2/6-311G($2d$,$2p$) to calculate 11 different approach routes for the dimer; and Tsuzuki derived a methane dimer potential based on MP2/6-31G* calculations.[15] Benzene dimers have also been studied recently.[22,23] Several studies included regression of simple model potential parameters from the *ab initio* results;[3,6,8,10,24,25] others have used *ab initio* derived potentials in molecular simulations.[6,9,11,15,26–28]

The purpose of this work is to determine the *ab initio* potential energy surface for ethane dimers consistent with the work done in RP. The ability of spherical atom–atom interactions to reproduce this surface under the assumption of pairwise additivity is examined. We also examine the performance in this regard of several simple, intersite potential models for the C–C, H–H and C–H interactions. These models are examined in terms of parameter coupling and any resultant deterioration of the physical meaning of the parameterized potential. The results of this study in conjunction with RP also contributes to an overall effort to find a complete set of C–C, C–H, and H–H interactions for each different type of CH_x group in small alkanes.

TABLE I. Ethane optimized MP4/6-311+G($2df$,$2pd$) geometry.

Parameter	Definition	Value
b_{CC}	C–C bond length	1.5227 Å
b_{CH}	C–H bond length	1.0883 Å
$\angle_{HCH}$	HCH bond angle	120°
$\angle_{HCC}$	HCC bond angle	111.252°
d_{HCCH}	dihedral angles	60°, −60°

B. *Ab initio* calculations

GAUSSIAN 98© (Ref. 29) was used to perform all of the calculations for this study. The equilibrium geometry for a single, isolated ethane molecule was optimized with MP4/6-311+G($2df$,$2pd$). The geometry obtained is detailed in Table I.

All dimer calculations were performed using the isolated, optimized molecular geometry for the monomer in the $D3d$ staggered configuration without relaxation. Although

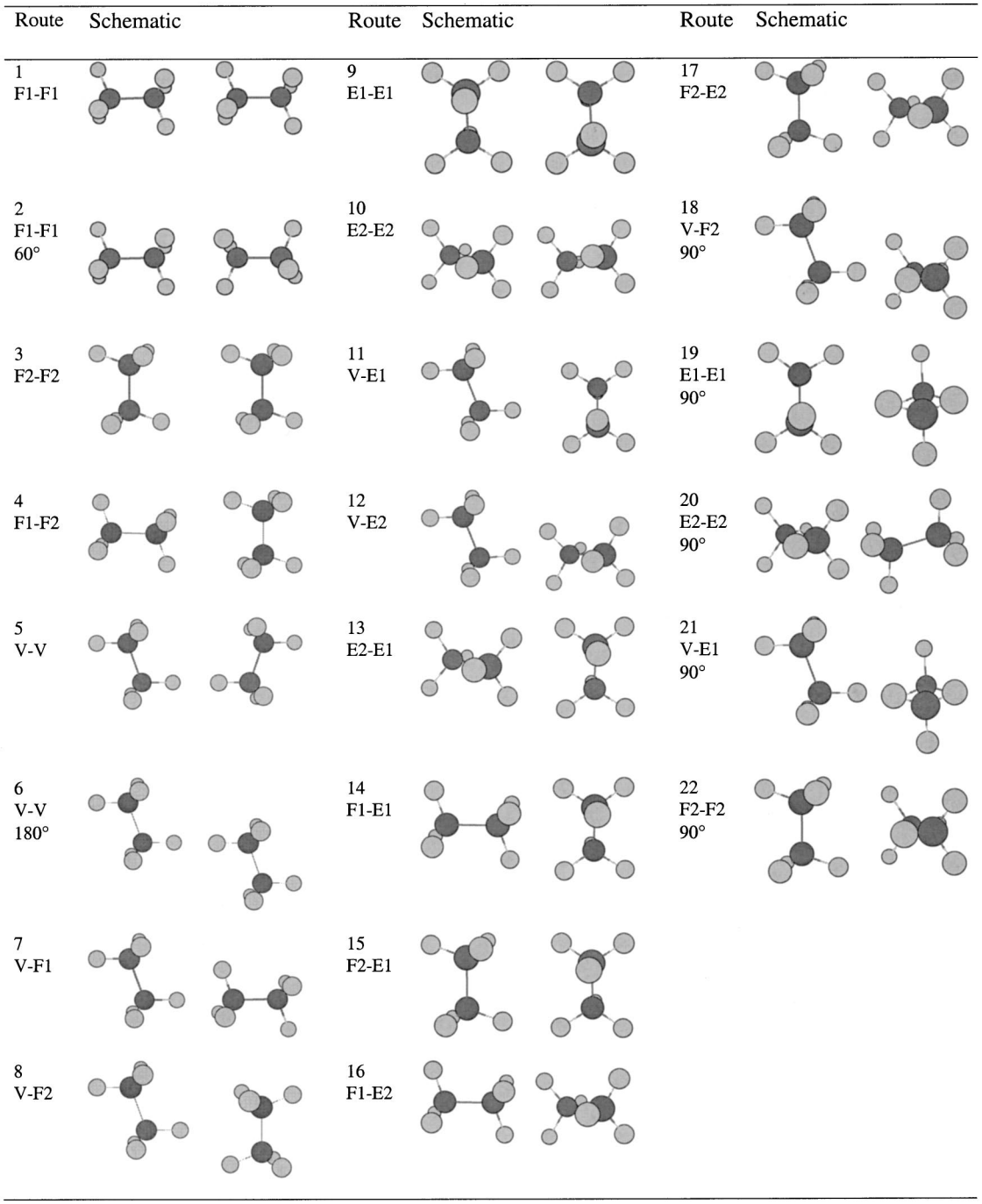

FIG. 1. Relative orientations (routes) used to sample the dimer potential surface. A route involves varying distances between the centers of the two molecules along particular lines of approach involving the faces, edges, and lines shown in Fig. 2.

geometry relaxation could be included in the dimer calculations, we are interested here in obtaining potential parameters for the rigid ethane model commonly used in MD simulations. It is clear that in the actual condensed-phase environment, torsional, angle, and bond strain will occur. These effects can be approximately included in MD simulations through additional internal potentials; what we seek here is a parameterization of intersite potentials from the most stable rigid ethane structure. All dimer energies included CPCs to eliminate basis set size differences between the monomer and dimer calculations.

C. Results

Supermolecule calculations of CPC energies were obtained as a function of distance between the centers of two ethane molecules, relative to infinite separation, for each of the 22 different relative orientations shown in Fig. 1. These routes were selected so as to sample the primary, unique, relative orientations of two trigonal solid objects defined by passing planes through sets of three hydrogen atoms on the ethane molecule as shown in Fig. 2. The planes through the hydrogen nuclei define ethane orientations in terms of two

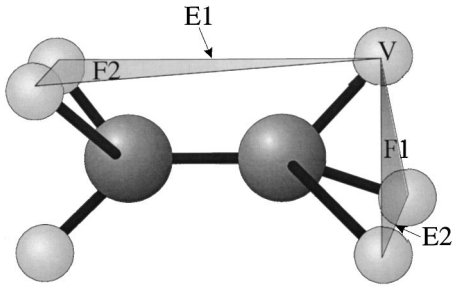

FIG. 2. Geometry used to define relative orientations of dimer pairs.

types of faces (F1 and F2), two unique edges (E1 and E2), and one structurally unique vertex (V); i.e., two planes, two lines, and one point. Dimer energies were calculated at approximately 18 different separation distances (between ethane centers) along the different approach routes. Routes are defined in terms of the geometrical features using lines of approach that pass through F, E, or V in the following ways: (1) for F1 the line passes along the C–C bond and through the center of the F1 face at right angles, (2) for F2 the line bisects the C–C bond and is perpendicular to that bond, (3) for E1 and E2 the line bisects at right angles the edge line (note that in the case of E2 this line also bisects the H–C–H angle), and (4) for V the line passes along the C–H bond. Thus, the F1–F1 route is defined by a line that passes through all four carbon atoms; a V–V route is defined by a line along a C–H bond in both monomers; and so forth. In addition to the 15 routes formed from unique combinations of the five defined structural identifiers, a variation on seven of the routes was formed by rotation of one molecule about the line of approach to set up a different configuration for the hydrogen atoms on the two approaching molecules [see, for example, the V–V and V–V (180°) configurations in Fig. 1]. Table II shows a matrix of the 22 routes in terms of these structural identifiers.

In conjunction with the inherent symmetry of the pairs, these 395 CPC energies provide a relatively complete potential energy surface for ethane dimers in the region where attraction can be important. We call this set of data the ''attraction data set'' even though some of the data are for distances where the potential is positive. The results for the attraction data set are given in Table III. An additional 128 points (approximately six per route) were calculated at separation distances closer than those for the attraction data set in order to more clearly define the repulsive region of the potential surface. We call these data the ''repulsion data set.'' These latter results are not given in the paper, but can be obtained from the authors.

TABLE II. Number of routes of each type used to characterize the dimer potential surface.

	F1	F2	E1	E2	V
F1	2				
F2	1	2			
E1	1	1	2		
E2	1	1	1	2	
V	1	2	2	1	2

III. ANALYTICAL SITE–SITE PAIR POTENTIALS

The reduction of the *ab initio* pair potentials into site–site potentials is not trivial. For each distance between molecular centers along each route there are 36 H–H, 24 C–H, and 4 C–C pair interactions. Because bond distances are short compared to the effective range of dispersion, all 64 pair interactions may contribute to the sum for each orientation as shown in Eq. (1). Thus,

$$U(r,\omega)=\sum_{i=1}^{2}\sum_{j=1}^{2}u_{\mathrm{CC}}(r_{\mathrm{C}_i,\mathrm{C}_j})+\sum_{i=1}^{2}\sum_{j=1}^{6}u_{\mathrm{CH}}(r_{\mathrm{C}_i,\mathrm{H}_j})+\sum_{i=1}^{6}\sum_{j=1}^{2}u_{\mathrm{CH}}(r_{\mathrm{H}_i,\mathrm{C}_j})+\sum_{i=1}^{6}\sum_{j=1}^{6}u_{\mathrm{HH}}(r_{\mathrm{H}_i,\mathrm{H}_j}), \quad (3)$$

where again the indices i and j refer to sites on molecules 1 and 2, respectively. Likewise parameters in the potential models for C–C, C–H, and H–H must, in theory, be regressed simultaneously. However, as in PR we found that the parameters in the C–C model were difficult to obtain despite the large quantity of data with different spatial orientations. This difficulty results primarily from the much larger sensitivity of the intermolecular potential to H–H and C–H interactions. In an effort to understand the sensitivity of the total potential to these site–site interactions, we numerically calculated the sensitivity coefficients, S_i,

$$S_i=\left(\frac{\partial U}{\partial p_i}\right)_{p_j\neq i}, \quad (4)$$

for each parameter in the site–site regression, where p_i represents parameter i in the site–site model. For the purpose of calculating these sensitivity coefficients, the modified-Morse site–site model,

$$u_{ij}=-\epsilon(1-\{1-\exp[-A(r-r^*)]\}^2), \quad (5)$$

introduced in PR, was used with values of the parameters obtained from regression of the potential surface (see Sec. III C below).

Three key aspects are illuminated by the plot of sensitivity coefficients for the F1–F1 route shown in Fig. 3. First, the potential is insensitive to the ϵ parameter in the H–H potential over the whole approach distance. Second, the identical shapes and sensitivity of the r^* and A parameters for the C–C potential shows that these two parameters are coupled and can not be regressed independently from the energies for this route. Third, all of the other parameters are sensitive and uncoupled and should be obtainable from the regression. Sensitivity coefficients for other routes yield essentially the same message. The F2–F2 route which from geometry considerations should enhance the relative C–C contributions is shown in Fig. 4. However, even for this route the r^* and A parameters are coupled over most of the range. While there is some decoupling of the parameters at very short distances, their sensitivity is actually lower than in Fig. 3, and in this region the sensitivity coefficients of the other model parameters rise much faster. This sensitivity analysis suggests that it is impossible to get the C–C spatial information (r^* and A) from the dimer energies and that the H–H ϵ parameter will be rather uncertain due to the lack of sensitivity of the dimer energies to it.

TABLE III. Calculated CPC dimer potential energies, *U*, for the attraction data set.

r/Å	U/kcal	r/Å	U/kcal	r/Å	U/kcal	r/Å	U/kcal	r/Å	U/kcal
Route 1		Route 2		Route 3		Route 4		Route 5	
4.523	1.841	4.623	1.262	3.500	1.160	3.961	1.411	4.529	0.925
4.723	0.374	4.723	0.587	3.700	−0.166	4.161	0.070	4.729	0.201
4.923	−0.226	4.823	0.152	3.900	−0.657	4.361	−0.466	4.929	−0.104
5.123	−0.421	4.923	−0.119	4.100	−0.766	4.561	−0.620	5.129	−0.209
5.323	−0.441	5.023	−0.304	4.300	−0.717	4.761	−0.605	5.329	−0.228
5.523	−0.394	5.123	−0.388	4.500	−0.613	4.961	−0.529	5.529	−0.211
5.723	−0.329	5.223	−0.423	4.700	−0.503	5.161	−0.439	5.729	−0.182
5.923	−0.265	5.323	−0.429	4.800	−0.452	5.361	−0.355	5.929	−0.153
6.123	−0.211	5.423	−0.416	5.000	−0.361	5.561	−0.283	6.129	−0.126
6.323	−0.167	5.523	−0.393	5.200	−0.287	5.761	−0.226	6.329	−0.103
6.523	−0.133	5.623	−0.363	5.400	−0.228	5.961	−0.180	6.529	−0.084
6.723	−0.106	5.723	−0.331	6.000	−0.118	6.161	−0.144	6.729	−0.069
6.923	−0.085	6.123	−0.207	6.800	−0.054	6.361	−0.116	6.929	−0.057
7.523	−0.045	6.323	−0.165	7.600	−0.027	6.761	−0.077	7.129	−0.047
8.323	−0.022	6.723	−0.104			7.561	−0.036	7.329	−0.039
		7.523	−0.045			8.361	−0.019	7.529	−0.032
		8.323	−0.021					7.729	−0.027
								7.929	−0.023
								8.129	−0.020
								8.329	−0.017
Route 6		Route 7		Route 8		Route 9		Route 10	
4.746	0.939	4.581	0.662	3.831	1.262	3.757	0.524	4.580	0.775
4.937	0.213	4.779	−0.065	4.027	0.116	3.957	−0.257	4.774	0.051
5.129	−0.094	4.976	−0.341	4.224	−0.372	4.157	−0.522	4.967	−0.235
5.321	−0.202	5.175	−0.409	4.422	−0.528	4.357	−0.560	5.162	−0.317
5.514	−0.222	5.373	−0.388	4.619	−0.531	4.557	−0.510	5.356	−0.312
5.708	−0.206	5.571	−0.336	4.817	−0.473	4.757	−0.433	5.552	−0.275
5.902	−0.179	5.769	−0.278	5.015	−0.398	4.957	−0.355	5.747	−0.230
6.096	−0.150	5.968	−0.225	5.213	−0.326	5.157	−0.287	5.943	−0.189
6.291	−0.124	6.167	−0.181	5.411	−0.263	5.357	−0.231	6.139	−0.153
6.486	−0.101	6.365	−0.145	5.609	−0.211	5.557	−0.186	6.335	−0.124
6.681	−0.083	6.564	−0.117	5.808	−0.169	5.757	−0.149	6.531	−0.101
6.877	−0.068	6.763	−0.094	6.006	−0.136	5.957	−0.121	6.728	−0.082
7.072	−0.056	6.962	−0.076	6.205	−0.110	6.157	−0.098	6.925	−0.067
7.268	−0.046	7.161	−0.062	6.404	−0.089	6.357	−0.081	7.122	−0.055
7.661	−0.032	7.559	−0.042	6.603	−0.073	6.557	−0.066	7.319	−0.046
8.055	−0.023	7.957	−0.030	6.801	−0.060	6.757	−0.055	7.517	−0.038
				7.000	−0.049	6.957	−0.046	7.714	−0.032
				7.398	−0.034	7.357	−0.032	7.912	−0.027
								8.307	−0.019
Route 11		Route 12		Route 13		Route 14		Route 15	
3.908	0.585	4.374	1.557	4.033	0.832	4.040	0.968	3.678	0.955
4.104	−0.196	4.570	0.392	4.231	−0.053	4.240	−0.106	3.878	−0.098
4.302	−0.494	4.766	−0.118	4.429	−0.389	4.440	−0.509	4.078	−0.489
4.499	−0.558	4.962	−0.306	4.628	−0.472	4.640	−0.602	4.278	−0.579
4.697	−0.522	5.159	−0.347	4.826	−0.450	4.840	−0.565	4.478	−0.544
4.894	−0.450	5.356	−0.324	5.024	−0.390	5.040	−0.485	4.678	−0.469
5.092	−0.372	5.553	−0.280	5.223	−0.324	5.240	−0.398	4.878	−0.387
5.290	−0.302	5.750	−0.232	5.422	−0.263	5.440	−0.320	5.078	−0.314
5.489	−0.243	5.947	−0.189	5.621	−0.212	5.640	−0.256	5.278	−0.252
5.687	−0.195	6.145	−0.153	5.819	−0.171	5.840	−0.204	5.478	−0.202
5.886	−0.156	6.343	−0.123	6.018	−0.138	6.040	−0.163	5.678	−0.163
6.084	−0.126	6.541	−0.099	6.217	−0.111	6.240	−0.131	5.878	−0.131
6.283	−0.102	6.739	−0.081	6.417	−0.091	6.440	−0.106	6.078	−0.107
6.482	−0.083	6.937	−0.066	6.616	−0.074	6.640	−0.086	6.278	−0.087
6.680	−0.068	7.135	−0.054	6.815	−0.061	6.840	−0.070	6.678	−0.059
6.879	−0.056	7.333	−0.045	7.014	−0.051	7.040	−0.058	7.078	−0.041
7.078	−0.046	7.532	−0.037	7.213	−0.042	7.240	−0.048		
7.476	−0.032	7.730	−0.031	7.413	−0.035	7.440	−0.040		
7.875	−0.023	8.128	−0.022			7.640	−0.034		
8.273	−0.017	8.525	−0.016			7.840	−0.028		

TABLE III. (*Continued*.)

r/Å	U/kcal	r/Å	U/kcal	r/Å	U/kcal	r/Å	U/kcal	r/Å	U/kcal
Route 16		Route 17		Route 18		Route 19		Route 20	
4.512	0.928	3.956	1.374	3.831	1.202	3.157	1.596	4.305	1.234
4.710	0.034	4.154	0.161	4.027	0.091	3.357	−0.165	4.501	0.210
4.908	−0.313	4.352	−0.330	4.224	−0.379	3.557	−0.860	4.697	−0.226
5.107	−0.407	4.550	−0.481	4.422	−0.527	3.757	−1.038	4.894	−0.374
5.305	−0.393	4.748	−0.482	4.619	−0.528	3.957	−0.988	5.091	−0.389
5.504	−0.342	4.947	−0.428	4.817	−0.470	4.157	−0.854	5.289	−0.351
5.703	−0.283	5.145	−0.359	5.015	−0.396	4.357	−0.704	5.486	−0.297
5.902	−0.228	5.344	−0.293	5.213	−0.324	4.557	−0.567	5.684	−0.244
6.101	−0.183	5.543	−0.237	5.411	−0.262	4.757	−0.451	5.882	−0.197
6.300	−0.146	5.741	−0.190	5.609	−0.210	4.957	−0.357	6.080	−0.159
6.499	−0.117	5.940	−0.153	5.808	−0.169	5.157	−0.283	6.278	−0.128
6.698	−0.094	6.139	−0.123	6.006	−0.136	5.357	−0.225	6.476	−0.103
6.898	−0.076	6.339	−0.100	6.205	−0.110	5.557	−0.180	6.674	−0.084
7.097	−0.062	6.538	−0.082	6.404	−0.089	5.757	−0.145	6.873	−0.069
7.296	−0.051	6.737	−0.067	6.603	−0.073	5.957	−0.117	7.071	−0.056
7.495	−0.042	6.936	−0.055	6.801	−0.060	6.157	−0.096	7.270	−0.047
7.894	−0.029	7.135	−0.046	7.000	−0.050	6.557	−0.065	7.468	−0.039
				7.199	−0.041	6.957	−0.045	7.667	−0.032
						7.357	−0.032	7.866	−0.027
								8.065	−0.023

Route 21		Route 22	
3.908	0.666	3.400	0.112
4.104	−0.144	3.600	−0.662
4.302	−0.461	3.800	−0.901
4.499	−0.538	4.000	−0.893
4.697	−0.508	4.200	−0.788
4.894	−0.441	4.400	−0.658
5.092	−0.367	4.600	−0.534
5.290	−0.298	4.800	−0.428
5.489	−0.240	5.000	−0.341
5.687	−0.192	5.200	−0.271
5.886	−0.155	5.400	−0.216
6.084	−0.124	5.600	−0.173
6.283	−0.101	5.800	−0.140
6.482	−0.082	6.000	−0.113
6.680	−0.067	6.200	−0.092
6.879	−0.055	6.400	−0.076
7.078	−0.046	6.600	−0.063
7.277	−0.038	6.800	−0.052
7.476	−0.032		
7.676	−0.027		
7.875	−0.023		
8.074	−0.020		

This coupling between parameters in the C–C model is a characteristic of the geometry and pairwise additive calculation, not the particular interatomic model chosen. We have used the modified-Morse potential to illustrate the problem, but the same problem occurred for all of the other models tested as well. The geometry of the molecules with the C atoms interior to the H atoms results in a shielding of the C–C interactions in the sense that the C–H and H–H interactions dominate because of their closer proximity to each other than the C–C interactions. This effect is compounded by the fact that there are 15 times as many C–H and H–H pairs as there are C–C.

Parameter coupling was overcome in PR by eliminating one of the spatial parameters in the C–C model. In PR the separation distance at which the C–C energy in the modified-Morse model is a minimum, r^*_{CC}, was constrained to be related to the minimum H–H and C–H distances by

$$r^*_{CC}=2r^*_{CH}-r^*_{HH}. \tag{6}$$

The use of this approximation still leaves the C–C potential ill-defined because the other model parameters then depend upon this arbitrary definition. As the real problem is the relative insignificance or screening of the C–C terms in Eq. (3), we choose here to eliminate the C–C terms entirely from the summation to obtain,

$$U(r,\omega)=\sum_{i=1}^{2}\sum_{j=1}^{6}u_{CH}(r_{C_i,H_j})+\sum_{i=1}^{6}\sum_{j=1}^{2}u_{CH}(r_{H_i,C_j}) +\sum_{i=1}^{6}\sum_{j=1}^{6}u_{HH}(r_{H_i,H_j}). \tag{7}$$

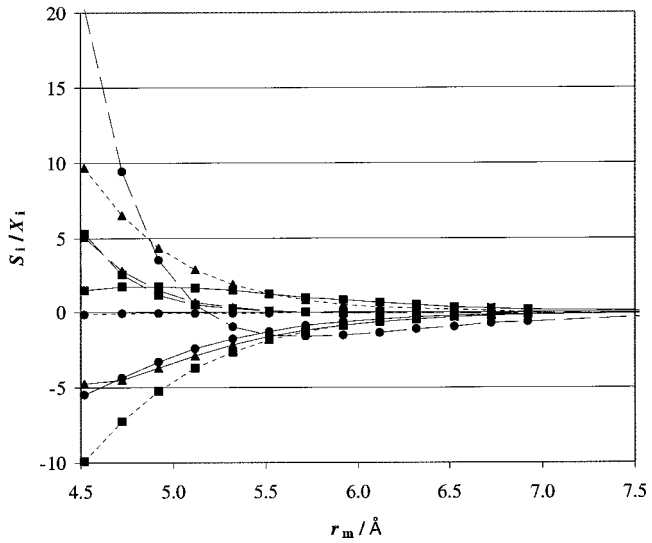

FIG. 3. Sensitivity coefficients for modified-Morse potential parameters for the F1–F1 dimer route. Pair model: C–C (long-dashed line), C–H (solid line), H–H (short-dashed line); parameters: ϵ (●), A (■), r^* (▲); r_m is center-to-center distance of dimer; $X_\epsilon = 1$, $X_A = 1$ kcal Å/mol, and $X_{r^*} = 1$ kcal/(mol Å).

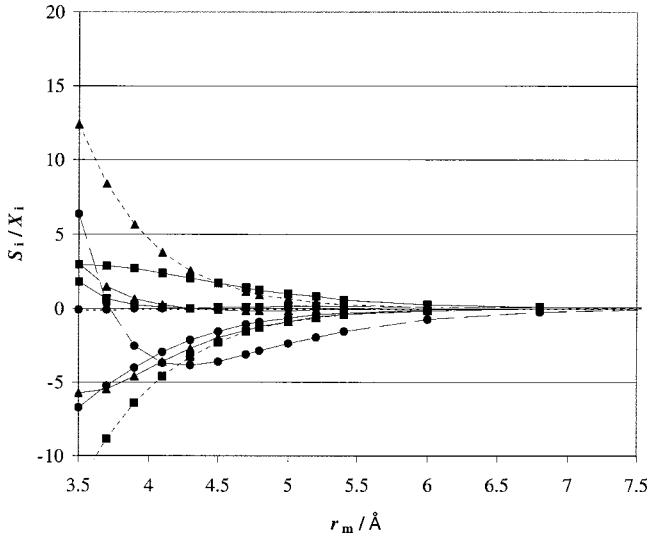

FIG. 4. Sensitivity coefficients for modified-Morse potential parameters for the F2–F2 dimer route. See Fig. 3 for legend.

We call this a screened pair-additive (SPA) potential in which the central C–C interactions do not contribute. We found this model to effectively represent the dimer energies while solving the parameter uniqueness problem. Any actual contributions to the summation by the C–C terms are included in the effective C–H and H–H terms. This SPA model is used in the studies discussed below except as noted.

A. Lennard-Jones potential

The Lennard-Jones (LJ) potential has been used widely to represent interatomic interactions in MD simulations. The LJ potential is often represented in two forms,

$$u_{ij} = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] = \frac{C_{12}}{r^{12}} - \frac{C_6}{r^6}, \tag{8}$$

where ϵ is the depth of the potential well and σ is the value of r at which u_{ij} becomes zero. The second form is related to the first by

TABLE IV. Parameters obtained for SPA interactions from various models, the sum of the squared residuals (SSR) for the regression, and the average absolute residual (AAR) per point.

Model/Parameters	C–H	H–H	SSR/(kcal mol^{-1})^2	$10^3\times$AAR (kcal mol^{-1})
1. LJ: 6 parameter			10.9	8.4
C_{12}/kcal mol^{-1} Å^{12}	95 724	−1526.4		
C_6/kcal mol^{-1} Å^6	398.95	−112.9		
2. exp-6:			6.42	6.4
A/kcal mol^{-1}	3203.7	2293		
B/Å^{-1}	3.015	1657		
C_6/kcal mol^{-1} Å^6	385.91	−57.924		
3. exp-6: 2 damping functions			6.18	6.3
A/kcal mol^{-1}	1987.93	2308.5		
B/Å^{-1}	2.6575	5.4472		
C_6/kcal mol^{-1} Å^6	172.45	−52.738		
b/Å^{-1}	0.498 43	23.318		
t	1.474	6.732		
4. exp-6: 1 damping function			6.18	6.3
A/kcal mol^{-1}	1987.93	2308.5		
B/Å^{-1}	2.6575	5.4472		
C_6/kcal mol^{-1} Å^6	191.47	−53.315		
b/Å^{-1}	0.520 33	⋯		
t	1.444	⋯		
5. Modified Morse:			4.19	5.2
ϵ/kcal mol^{-1}	0.5853	−38.612		
A/Å^{-1}	1.4873	1.7274		
r^*/Å	2.7484	−0.2308		

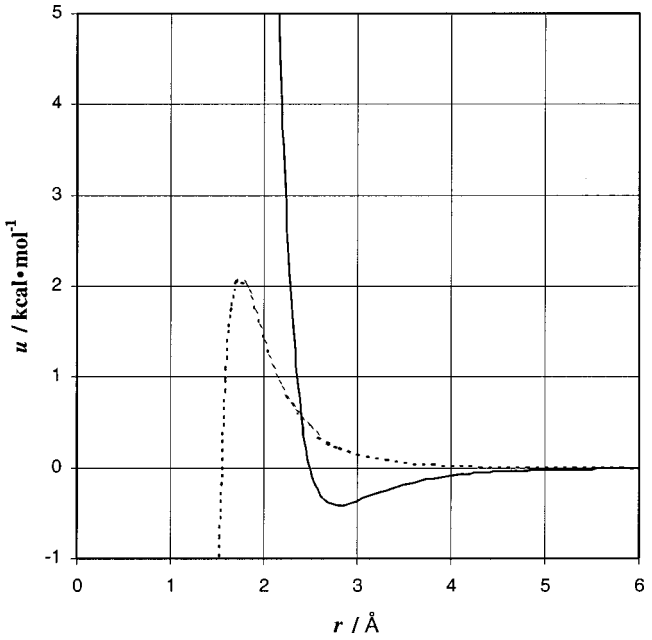

FIG. 5. LJ interatomic potential models for C–H (solid line) and H–H (dotted line) regressed from the *ab initio* potential surface.

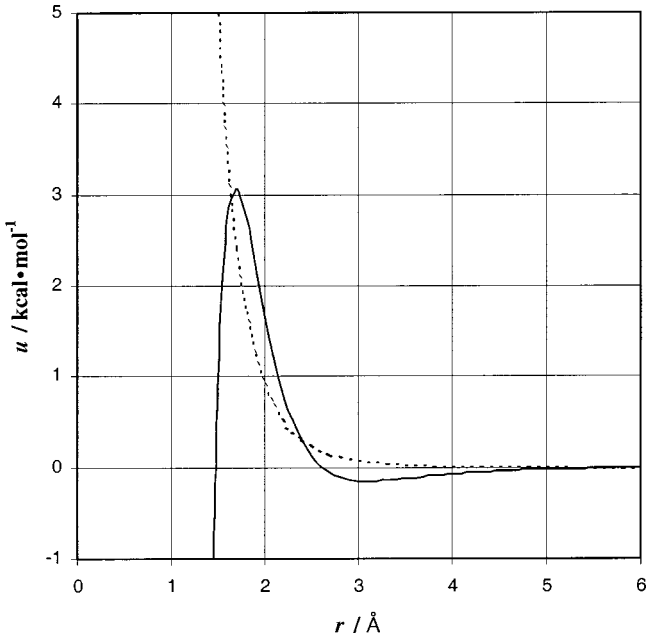

FIG. 6. Exp-6 interatomic potential models for C–H (solid line) and H–H (dotted line) regressed from the *ab initio* potential surface.

$$\sigma=\left(\frac{C_{12}}{C_6}\right)^{1/6} \quad \text{and} \quad \epsilon=\frac{C_6^2}{4C_{12}}. \tag{9}$$

The parameters C_{12} and C_6 for the C–H and H–H potentials were regressed from the *ab initio* data with a resultant sum of squared residuals (SSR) equal to 10.9 $(\text{kcal/mol})^2$ and an average absolute residual per point (AAR) of 8.4×10^{-3} kcal/mol. Values obtained for the parameters are listed in Table IV. The relatively high SSR indicates that the LJ interatomic model is not very effective in modeling the potential surface generated from *ab initio* calculations. In addition, the model was not able to simultaneously describe both the attraction and repulsion data sets. (The results shown in Table IV are for the attraction set.) Moreover, the resultant interatomic potentials for the H–H interactions appear nonphysical. As shown in Fig. 5, the regressed H–H potential is repulsive at longer distances and attractive at very short distances.

B. Exp-6 model

The exp-6 model can be written as

$$u_{ij}=Ae^{-Br}-f(r)\frac{C_6}{r^6}, \tag{10}$$

where the damping factor, $f(r)$, provides additional flexibility beyond the original equation (with $f=1$) in switching between repulsion and dispersion. We have tested the exp-6 model with $f=1$ for both C–H and H–H potentials. The fit was significantly better than the LJ model with SSR=6.42 $(\text{kcal/mol})^2$ and AAR=6.4×10^{-3} kcal/mol. In this case, the parameters A and B were regressed from the repulsion data set, and then the C_6 parameter was obtained from the attraction data set. In spite of the better fit to the *ab initio* values, the physics of the resultant interatomic models are incorrect as can be seen in Fig. 6. The C–H potential shows a second attractive region at short distances.

We should point out that we have also regressed the potential surface using this model with a separation of partial charges on the C and H nuclei into an additional Coulombic term. In this case the partial charges were determined from both electrostatic potential and Mulliken population analyses. Unfortunately, the two methods gave different signs for the partial charges assigned to the C and H atoms. But, in both cases, the results were very similar to the exp-6 without the charge separation.

Recently, Hodges and Stone[30] proposed several damping functions, the simplest (fewest adjustable parameters) of which was examined in this study. The form used here was

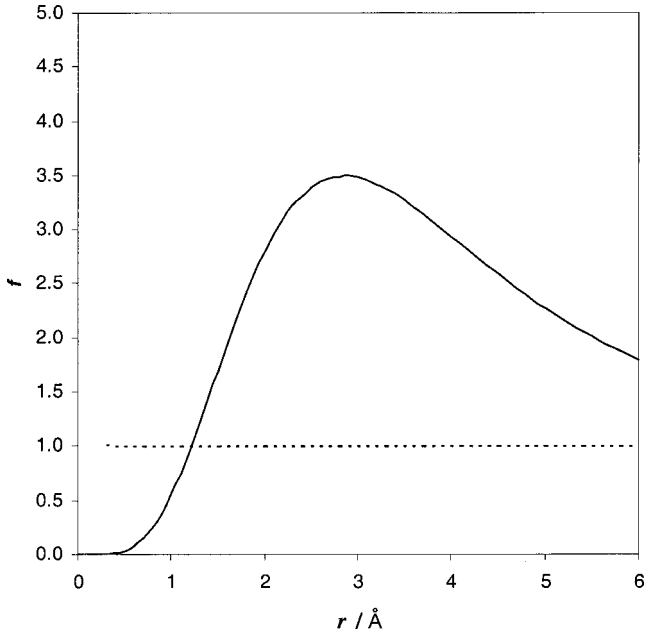

FIG. 7. Damping function for the exp-6 model for C–H (solid line) and H–H (dotted line) interactions.

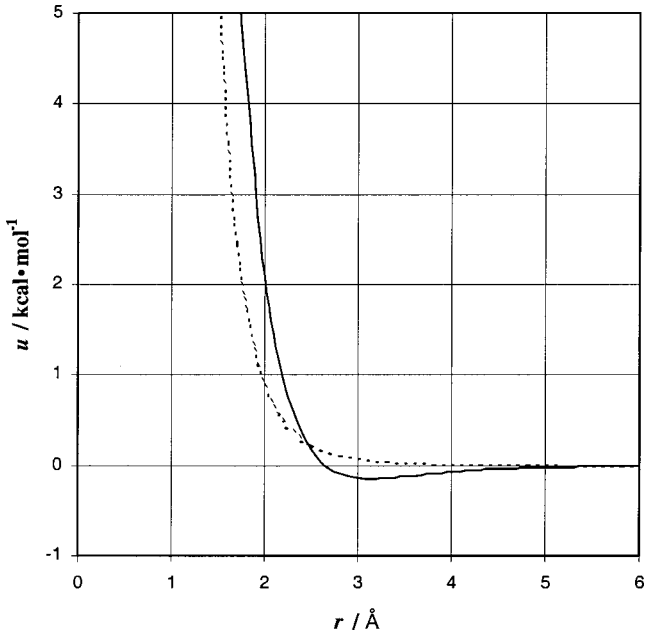

FIG. 8. Exp-6 (with damping functions) interatomic potential models for C–H (solid line) and H–H (dotted line) regressed from the *ab initio* potential surface.

$$f(r)=(1+3e^{-br}-4e^{-tbr})^6. \tag{11}$$

In this potential model there are ten parameters, five for each pair interaction. The decoupling and regression of these parameters was done using the method suggested by Hodge and Stone. The parameters A and B in Eq. (10) were regressed first using only the repulsion data set. The t, b, and C_6 parameters were then regressed simultaneously (six parameters) using both the attraction and repulsion data sets.

The exponential term with regressed values of A and B described the repulsion data set very well. The simultaneous regression of b, t, and C_6 was more difficult because of multiple local minima. In this case, a global minimum was found by performing the regression using 10^6 sets, chosen randomly, of different starting values for the parameters. The resultant SSR was 6.18 $(\text{kcal/mol})^2$ with $\text{ARR}=6.3\times10^{-3}$ kcal/mol. The damping functions obtained for the two potentials are shown in Fig. 7. As can be seen, the damping function for the H–H potential is essentially unity, suggesting that the regression might also be done with two fewer parameters by regressing b, t, and C_6 for the C–H interactions, but only C_6 for the H–H potential. The SSR when these four parameters were simultaneously regressed was equivalent to the previous case in which a damping function was used for both interactions. The resultant pair-potential for the latter regression is shown in Fig. 8. Note that with the damping function, the regression now yields physically reasonable interatomic potentials. The repulsive nature of the H–H interactions over the whole range of distances is consistent with the expected repulsion between the equal (partial) charges on the H atoms.

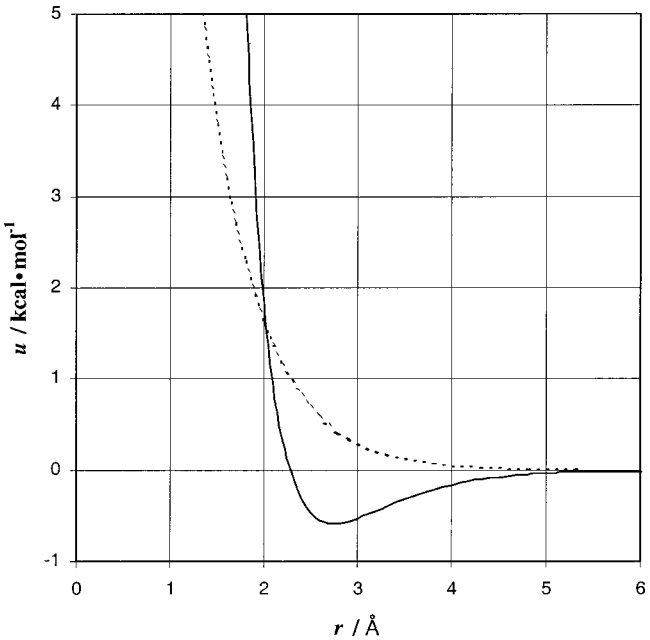

FIG. 9. Modified-Morse interatomic potential models for C–H (solid line) and H–H (dotted line) regressed from the *ab initio* potential surface.

TABLE V. Modified-Morse potential parameters obtained using full pairwise additivity while constraining the r^* parameter in the C–C potential.

	C–C	C–H	H–H	SSR/$(\text{kcal mol}^{-1})^2$
r^*_{CC} fixed by Eq. (6)				2.3
ϵ/kcal mol^{-1}	0.3319	0.3061	−5.1396	
$A/\text{Å}^{-1}$	1.3944	2.3427	2.8276	
r^*/Å	3.7845	2.5427	1.3009	
r^*_{CC} fixed at 4.35 Å				2.3
ϵ/kcal mol^{-1}	0.0689	0.7409	−40.65	
$A/\text{Å}^{-1}$	1.3083	1.8287	2.2239	
r^*/Å	4.35	2.491	0.2456	

C. Modified-Morse potential

We have also regressed the *ab initio* dimer energies to obtain six parameters (A, r^*, and ϵ for the C–H and H–H pairs) for Eq. (5). The resultant SSR was 4.19 $(\text{kcal/mol})^2$ ($\text{AAR}=5.2\times10^{-3}$ kcal/mol), producing a very good fit for all 22 routes. The physics of this model, like the exp-6 with damping functions, (1) are consistent with the partial charge considerations on the sites, (2) attribute dimer attractions to strong C–H attractions, and (3) exhibit repulsive H–H interactions over all separation distances. The resultant model parameters are listed in Table IV and the resultant interatomic potentials are shown in Fig. 9.

Additionally, we have used the unshielded pair-additive model, Eq. (3), with a constraint on the spatial parameters in the C–C model to regress the dimer energies. We have used the constraint employed in PR for methane dimers, Eq. (6), as well as simply fixing the value of r^*_{CC} at a reasonable value and regressing all remaining parameters. The results of the regression analysis with these models are given in Table V and the resultant model pair potentials are shown in Fig. 10. The value chosen for r^*_{CC} changes the strength of the C–H attraction considerably as well as the shapes of the pair interactions. As shown in Fig. 10, a nonphysical turnover in the H–H repulsions results when eight parameters are regressed in conjunction with the Eq. (6) combining rule.

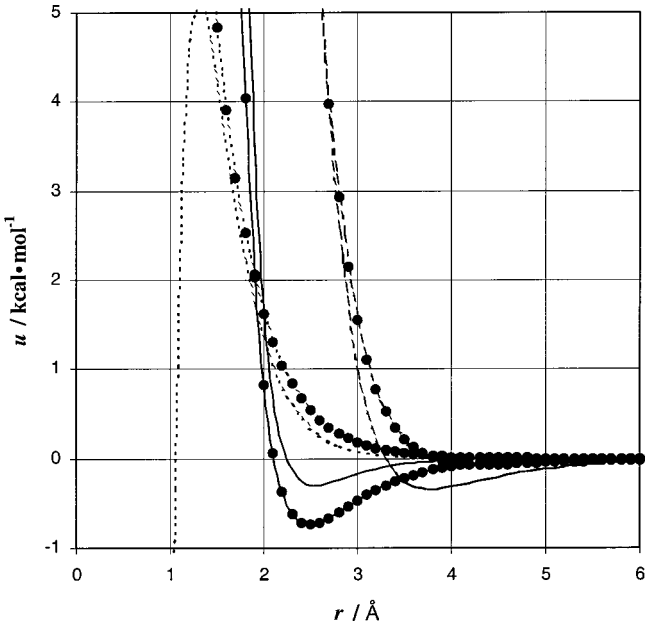

FIG. 10. Modified-Morse interatomic potentials for C–H (solid line), H–H (dotted line), and C–C (dashed line) when r^*_{CC} is fixed by Eq. (6) (no symbols) and when r^*_{CC}=4.35 Å (lines with symbols).

Specifying an r^*_{CC} value instead of using Eq. (6) allows one to remove this turnover in potential if the value of r^*_{CC} selected is adequately large. The value used in Fig. 10 and Table V is r^*_{CC}=4.35 Å.[7,31] We do not believe these potentials offer any advantage over the SPA potentials developed by excluding the C–C interactions.

IV. CONCLUSIONS

We have used MP2/6-311+G(2*df*,2*pd*) to calculate CPC energies of ethane dimers in 22 different relative orientations as a function of separation distance in order to generate an accurate representation of the potential energy surface of two ethane molecules. These energies were used to regress the parameters in several simple, analytical, site–site models that can be used for implementation in molecular simulations. The agreement of energies calculated from the site–site models with the *ab initio* calculations indicates that the use of such models is appropriate and does not limit the accuracy of simulations.

Regression results and sensitivity coefficients using the analytical site–site models suggest that care must be used in obtaining the model parameters and in attributing physical characteristics to the resultant interaction models. Much of this difficulty results from the lack of sensitivity of the calculated potential energy to the shielded C–C interactions. The number of regressed parameters must be reduced to those that are sensitive to the data and are not completely coupled if physically meaningful inter-site potentials are to be obtained. In this vein, we have used a SPA potential that omits the C–C interactions between these shielded ''interior sites.'' Using the SPA equation, the LJ model was unable to adequately fit both the repulsive and attraction data sets. The exp-6 potential fits the data better, but both models produced nonphysical H–H interactions. With a damping function added, the exp-6 model fit the data quite well and the H–H potential appears more physically reasonable. The modified-Morse potential fits the *ab initio* potentials best as well as producing interatomic models physically consistent with the charge distributions within the molecules.

The question of parameter transferability for these models yet remains. The results reported here in conjunction with *ab initio* calculations currently in progress for *n*-propane, isobutane, and neopentane dimers should help answer this question.

ACKNOWLEDGMENT

Support of this work by the Neste Foundation is gratefully acknowledged.

[1] W. L. Jorgensen, J. Am. Chem. Soc. **103**, 335 (1981).
[2] W. L. Jorgensen, J. D. Madura, and C. J. Swenson, J. Am. Chem. Soc. **106**, 6638 (1984).
[3] R. L. Rowley and T. Pakkanen, J. Chem. Phys. **110**, 3368 (1999).
[4] D. E. Woon, J. Chem. Phys. **100**, 2838 (1994).
[5] F. Tao, Z. Li, and Y. Pan, Chem. Phys. Lett. **255**, 179 (1996).
[6] C. Hu and A. J. Thakkar, J. Chem. Phys. **104**, 2541 (1996).
[7] J. Hill, J. Comput. Chem. **18**, 211 (1997).
[8] P. J. Marshall, M. M. Szczesniak, G. Chalasinski, J. Sadlej, M. A. ter Horst, and C. J. Jameson, J. Chem. Phys. **104**, 6569 (1996).
[9] S. Tsuzuki, T. Uchimaru, M. Mikami, K. Tanabe, T. Sako, and S. Kuwajima, Chem. Phys. Lett. **255**, 347 (1996).
[10] J. Shen, K. B. Domański, O. Kitao, and K. Nakanishi, Fluid Phase Equilibria **104**, 375 (1995).
[11] J. Soetens, G. Jansen, and C. Millot, Mol. Phys. **96**, 1003 (1999).
[12] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, and K. Tanabe, J. Am. Chem. Soc. **122**, 3746 (2000).
[13] S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, Masuhiro, and K. Tanabe, J. Phys. Chem. A **103**, 8265 (1999).
[14] S. Tsuzuki, T. Uchimaru, and K. Tanabe, Chem. Phys. Lett. **287**, 202 (1998).
[15] S. Tsuzuki, T. Uchimaru, and K. Tanabe, Chem. Phys. Lett. **287**, 327 (1998).
[16] S. Tsuzuki, T. Uchimaru, M. Mikami, and K. Tanabe, J. Phys. Chem. A **102**, 2091 (1998).
[17] S. Tsuzuki, T. Uchimaru, M. Mikami, and K. Tanabe, Chem. Phys. Lett. **252**, 206 (1996).
[18] S. Tsuzuki, T. Uchimaru, K. Matsumura, M. Mikami, and K. Tanabe, J. Chem. Phys. **110**, 11906 (1999).
[19] S. Tsuzuki, T. Uchimaru, and K. Tanabe, J. Phys. Chem. A **102**, 740 (1998).
[20] T. G. Metzger, D. M. Ferguson, and W. A. Glauser, J. Comput. Chem. **18**, 70 (1997).
[21] J. J. Novoa, M. Whangbo, and J. M. Williams, J. Chem. Phys. **94**, 4835 (1991).
[22] V. Spirko, O. Engkvist, P. Soldan, H. L. Selzle, E. W. Schlag, and P. Hobza, J. Chem. Phys. **111**, 572 (1999).
[23] O. Engkvist, P. Hobza, H. L. Selzle, and E. W. Schlag, J. Chem. Phys. **110**, 5758 (1999).
[24] W. J. Jorgensen, J. Chem. Phys. **70**, 5888 (1979).
[25] J. Shen, O. Kitao, and K. Nakanishi, Fluid Phase Equilibria **120**, 81 (1996).
[26] M. A. ter Horst and C. J. Jameson, J. Chem. Phys. **105**, 6787 (1996).
[27] D. P. Visco, Jr. and D. A. Kofke, J. Chem. Phys. **109**, 4015 (1998).
[28] H. Schindler, R. Vogelsang, V. Staemmler, M. A. Siddiqi, and P. Svejda, Mol. Phys. **80**, 1413 (1993).
[29] GAUSSIAN 98, Revision A.6, M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.* (Gaussian, Inc., Pittsburgh, PA, 1998).
[30] M. P. Hodges and A. J. Stone, Mol. Phys. **98**, 275 (2000).
[31] A. T. Hagler, E. Huler, and S. Lifson, J. Am. Chem. Soc. **96**, 5319 (1974).