An accurate analytic He-H_2 potential energy surface from a greatly expanded set of ab initio energies

Arnold I. Boothroyd and Peter G. Martin
Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario, M5S 3H8, Canada

Michael R. Peterson
Department of Computing and Networking Services, University of Toronto, Toronto, Ontario, M5S 3J1, Canada.
(November 1, 2004: J. Chem. Phys., in press; preprint CITA-2003-19)

Abstract

The interaction potential energy surface (PES) of He-H_2 is of great importance for quantum chemistry, as the simplest test case for interactions between a molecule and a closed-shell atom. It is also required for a detailed understanding of certain astrophysical processes, namely collisional excitation and dissociation of H_2 in molecular clouds, at densities too low to be accessible experimentally. A new set of 23 703 *ab initio* energies was computed, for He-H_2 geometries where the interaction energy was expected to be non-negligible. These have an estimated rms "random" error of ~ 0.2 millihartree and a systematic error of ~ 0.6 millihartree (0.4 kcal/mol). A new analytic He-H_2 PES, with 112 parameters, was fitted to 20 203 of these new *ab initio* energies (and to an additional 4862 points generated at large separations). This yielded an improvement by better than an order of magnitude in the fit to the interaction region, relative to the best previous surfaces (which were accurate only for near-equilibrium H_2 molecule sizes). This new PES has an rms error of 0.95 millihartree (0.60 kcal/mole) relative to the the 14 585 *ab initio* energies that lie below twice the H_2 dissociation energy, and 2.97 millihartree (1.87 kcal/mole) relative to the full set of 20 203 *ab initio* energies (the fitting procedure used a reduced weight for high energies, yielding a weighted rms error of 1.42 millihartree, i.e., 0.89 kcal/mole). These rms errors are comparable to the estimated error in the *ab initio* energies themselves; the conical intersection between the ground state and the first excited state is the largest source of error in the PES.

I. INTRODUCTION

The interaction potential energy surface (PES) of He-H_2 (i.e., two hydrogen atoms plus one helium atom) is of great importance for quantum chemistry. He-H_2 is the simplest test case for theories of the interaction of a molecule with a closed-shell atom. It is highly desirable to have a PES for He-H_2 that attains or at least approaches the "chemical accuracy" (of order one millihartree, i.e., better than one kcal/mol) required for reaction dynamics — cf. available *ab initio* potential energy surfaces for H_3 (molecule plus open-shell atom)[1–5] or H_4 (two molecules)[6–9]. Previous analytic He-H_2 surfaces[10–24] were based on only a few dozen *ab initio* energies, which provided highly inadequate coverage of all but a very restricted region in the conformation space of He–H_2. This paper reports a greatly-expanded set of *ab initio* energies, and a greatly improved He-H_2 PES fitted to these energies.

The He-H_2 PES is also of particular astrophysical interest for studying He + H_2 interactions in physical conditions not accessible to experiment, namely the low densities characteristic of giant molecular clouds in the interstellar medium, where star formation occurs. Helium is the second-most-common collision partner for H_2 in these molecular clouds, which contain roughly one helium atom for every ten hydrogen atoms. Heating of these clouds by strong shock waves causes rotational and vibrational excitation of the H_2 molecules, and can lead to collision-induced dissociation of H_2 into free H atoms. The collision rates in molecular clouds can be so low that the (observed) forbidden (quadrupole) infrared emission of excited H_2 molecules can induce highly non-thermal distributions over the internal states of H_2 (see, e.g., Refs. 25, 26, and 27). Because the mean free paths of molecules are thousands of kilometers at such low densities, these processes will remain inaccessible to laboratory experiment. Computer simulation is therefore a *sine qua non* in the study of the physics and chemistry of star forming regions.

A. Notation

Atomic units are used in this paper unless otherwise specified, i.e., distances are in bohrs (a_0) and energies are in hartrees (E_h), millihartrees (mE_h), or microhartrees (μE_h). Recall that 1 a_0 = 0.529177 Å, while 1 mE_h = 0.0272114 eV = 0.62751 kcal/mol.

Unless otherwise stated, energy values are measured relative to the energy of two isolated hydrogen atoms plus an isolated helium atom. Thus an energy $E = 0.0\ E_h$ corresponds to the H_2 dissociation energy, while an isolated ground-state H_2 molecule (plus an isolated helium atom) lies at $E = -164.6\ mE_h = -103.3$ kcal/mol. The minimum of the H_2 potential energy curve of Schwenke[28] is $E = -174.496\ mE_h = -109.50$ kcal/mol; the lowest possible energy on the He-H_2 PES lies about 0.05 mE_h (0.03 kcal/mol) below this, due to the He + H_2 van der Waals well (discussed in § II A 4).

Figure 1 illustrates the notation used in this paper for the interatomic distances and angles. The (r, R, γ) coordinate system is the natural one to use when considering low-energy interactions of a He atom with a near-equilibrium H_2 molecule, e.g., the van der Waals well. It is less suitable for higher-energy interactions, where the He atom may be relatively close to one of the H atoms; He-H_2 surfaces that attempt to fit this interaction region typically use the (r, R_A, R_B) coordinate system.

B. Previous surfaces

Many of the early He-H_2 surfaces were intended only to describe the $He + H_2$ van der Waals potential well and the outer (low-energy) part of the He-H_2 repulsive wall[10–16,18,19,23], and were valid only for equilibrium H_2 molecules. However, several analytic He-H_2 surfaces were intended to be valid in the higher-energy interaction region.

In 1974, Wilson, Kapral, & Burns[17] presented a more general He-H_2 PES, albeit with an extremely simple functional form (comprising the three pair-wise potentials plus an interaction term with only four free parameters). Shortly thereafter, Dove & Raynor[20] used a modified form of this surface (with somewhat improved pair-wise potentials) to investigate the collisional dissociation of H_2 by He. Both of these surfaces[17,20] contain an unphysical "hole" at relatively short distances where the PES dropped abruptly to large negative values. For these surfaces, the lowest point on the rim of this hole lies at 53 mE_h (33 kcal/mol) and 87 mE_h (55 kcal/mol), respectively (where zero energy corresponds to separated *atoms*: see § I A above); i.e., unphysical effects due to this hole could be encountered at less than twice the dissociation energy relative to equilibrium H_2 + He. In addition, neither of these surfaces[17,20] included any $He + H_2$ van der Waals well.

The 1982 surface of Russek & Garcia[21] was intended to fit only the repulsive wall; their *ab initio* energies had been calculated for relatively small $He - H_2$ separations for near-equilibrium H_2 molecules (of size $r = 1.2$, 1.4, and 1.6 a_0).

The 1985 rigid-rotor He-H_2 PES of Schaefer & Köhler[22] was fitted to the relatively accurate *ab initio* energies of Meyer, Hariharan, & Kutzelnigg[29]; these had been calculated for relatively large $He - H_2$ separations for near-equilibrium H_2 molecules ($r = 1.28$, 1.449, and 1.618 a_0), supplemented by less-accurate calculations over a wider range ($0.9\ a_0 \leq r \leq 2.0\ a_0$). The PES fitted the van der Waals well and the $He + H_2$ repulsive wall, but was defined only for a limited range of H_2 molecule sizes (differing by no more than 40% from equilibrium H_2), and thus could not be used for cases where the H_2 molecule might be dissociated or highly excited.

In 1994, Muchnick & Russek[24] presented a general He-H_2 PES, with 19 fitted parameters. This surface was fitted to a combination of the Meyer, Hariharan, & Kutzelnigg[29] *ab initio* energies (the less-comprehensive *ab initio* energies of Senff & Burton[30] were used as comparison values but not actually fitted), and the somewhat less accurate *ab initio* energies of Russek & Garcia[21]. This Muchnick & Russek He-H_2 PES was a great improvement over previous surfaces: it was designed to be accurate in the van der Waals well and in the repulsive wall, and to behave reasonably in regions not constrained by *ab initio* data.

Recently, Tao[31] presented improved *ab initio* He-H_2 energies in the van der Waals well and repulsive wall, for near-equilibrium H_2 molecules ($r = 1.28$, 1.449, and 1.618 a_0). These *ab initio* energies should be significantly more accurate than the earlier energies of Meyer, Hariharan, & Kutzelnigg[29] or Senff & Burton[30] (and in fact show a van der Waals well that is several percent deeper than indicated by the earlier energies). However, Tao[31] did not fit an analytic surface to the *ab initio* energies.

C. Goals of the present work

An analytic general He-H_2 PES was desired that would not only accurately represent the van der Waals well, but also fit the interaction region at the "chemical accuracy" level required for reaction dynamics (i.e., of order one millihartree). In order to do this, 20 203 new *ab initio* energies were computed in the interaction region. Both the ground state energy and the first few excited state energies were computed, and the conical intersection of the ground state with the first excited state was mapped out approximately. A new analytic He-H_2 PES was fitted to these 20 203 ground state *ab initio* energies (and to an additional 4862 points generated to constrain the fit at large separations), yielding an order-of-magnitude improvement over previous He-H_2 surfaces. These *ab initio* energies and a Fortran program for this analytic PES (including analytic first derivatives) are available from EPAPS[32] or from the authors[33].

II. METHODS

A. The grid of conformations to be fitted

1. The main grid

The main set of He-H_2 *ab initio* energies was computed for a grid of 16 703 conformations, defined in terms of the H_2 molecule size r, the distance R_A of the He atom from one of the H atoms, the angle θ between these two, and in some cases the distance z_3 of the He atom from the midplane of the H_2 (see Fig. 1). The set of R_A grid values comprised {0.6, 0.7, 0.85, 1.0, 1.1, 1.2, 1.3, 1.4, 1.525, 1.65, 1.775, 1.9, 2.05, 2.2, 2.35, 2.55, 2.75, 2.95, 3.2, 3.45, 3.7, 4.0, 4.3, 4.6, 4.95, 5.3, and 5.7 a_0}. The H_2-size grid values r were taken from the same set as R_A, supplemented by a set of extra values at short distances {0.5, 0.55, 0.65, 0.775, and 0.925 a_0}, and a set at long distances {6.1, 6.55, 7.0, 7.45, 7.9, 8.35, 8.8, 9.25, 9.7, 10.15, and 10.6 a_0} — Figure 2 shows a typical grid example (a slice of the grid, at $r = 2.35\ a_0$). For large H − H separations (namely, $r \geq 6.1\ a_0$), only conformations having $R_B \leq 5.9\ a_0$ were computed (the He atom lying more or less between the two H atoms) — in other words, the shortest two of the three interatomic distances were never allowed to exceed 5.9 a_0. An example of this is shown in Figure 3 (a grid-slice at $r = 9.25\ a_0$).

The angle θ was in general taken at intervals $\Delta\theta = 7.5°$ (in $0° \leq \theta \leq 180°$); for cases with $R_A = 5.7\ a_0$ or with $r = 0.5$, 0.55, or 0.65 a_0, intervals $\Delta\theta = 15°$ were used instead, and cases with $\theta > 90°$ were treated specially. Where the ray $\theta = 120°$ reached the y-axis of Figure 1 (i.e., the He atom equidistant from the two H atoms), it was "bent" to follow the y-axis (since the grid needed to cover only the first quadrant of the y-z plane). Similarly, the rays $\theta = 102.5°$, 105°, and 97.5° were "bent" to follow the lines $z_3 = r/8$, $r/4$, and $3r/8$, respectively; but in these cases, the rays were terminated when they reached the point where the usual angular separation $R_A\Delta\theta$ reached twice the separation between the "bent" rays (so as to avoid excessively closely spaced angles at large R_A). For $\theta > 120°$ (i.e., $\max\{R_A, R_B\} < r$, with the He atom lying between the H atoms), a tighter θ-interval was

used if necessary to yield $R_A \Delta\theta \lesssim \Delta R_A$; also, extra points were interspersed on the y-axis if necessary, so that $\Delta y_3 \approx \Delta R_A$ on the y-axis. An example of this is shown in Figure 2a.

2. Set of "random" conformations

Another set of He-H_2 *ab initio* energies was computed for 3500 random conformations. The distances were chosen in the ranges 0.6 $a_0 < r <$ 9 a_0 and 0.9 $a_0 < R_A <$ 5.7 a_0 using non-uniform probability distributions that had broad peaks near 1.5 a_0. The angle θ was chosen using a uniform distribution, subject to the constraints that the the He atom must lie in the first quadrant of the y-z plane and that no atom could lie further than 5.7 a_0 from its nearest neighbor. Also, "near-duplicate" conformations were not permitted (i.e., where all cartesian coordinates agreed within 0.01 a_0 with the coordinates of another conformation). Figure 2a shows those random conformations that lie near the grid-slice $r = 2.35\ a_0$.

3. Unfitted "random" conformations

The same probability distributions were used to obtain a second set of 3500 random He-H_2 conformations, for testing purposes. The resulting *ab initio* energies were used only to check the rms error at "interstitial" conformations; they were *not* used in the surface fitting process.

4. Accurate van der Waals He + H_2 energies

A set of 2145 accurate energies was generated for conformations in the van der Waals well and outer repulsive wall, for H_2 molecules not too far from equilibrium. These energies were generated from the accurate *ab initio* energies and fitted surfaces of other authors[22,24,31], as described below. These accurate van der Waals energies were given high weight in the fit, with lower weight for smaller R or for r further from equilibrium (as discussed in § II D 1).

Unlike the main grid, these conformations were defined in terms of the distance and angle of the He atom relative to the H_2 *molecule*, i.e., in terms of $\{r, R, \gamma\}$ (rather than $\{r, R_A, \theta\}$). The set of r values comprised {0.6, 0.7, 0.8, 0.9, 1.1, 1.28, 1.449, 1.618, 1.8, 2.0, 2.4, 3.0, and 4.0 a_0}; the set of R values comprised {4.0, 4.5, 5.0, 5.25, 5.5, 5.75, 6.0, 6.25, 6.5, 6.75, 7.0, 7.25, 7.5, 7.75, 8.0, 8.5, 9.0, 10.0, 11.0, 12.0, 15.0, and 20.0 a_0}. The angle γ was generally taken at 15° intervals (in $0° \le \gamma \le 90°$); for $R = 20.0\ a_0$, and for roughly half of the R values at non-equilibrium r, intervals of 30° were used instead. Conformations generated at relatively small R were discarded if they had an interaction energy of more than a few mE_h between the He atom and the H_2 molecule (i.e., only the van der Waals well and the very outermost part of the "repulsive wall" were included). An example of this grid is shown by the squares in Figure 2b (a grid-slice at $r = 2.4\ a_0$). Note that there is an overlap with the main *ab initio* grid. In the outer part of this overlap region, the interaction energy of He with the H_2 molecule is comparable to the uncertainty in the main-grid *ab initio* energies, so the fit depends largely on the generated van der Waals $H_2 + He$ energies. In the inner part of the overlap region, the interaction energies are larger and the generated van

der Waals $H_2 + He$ energies are given fairly low weight (see § II D 1), so both sets of energies contribute to the fit.

Tao[31] used the complete fourth-order Møller-Plesset approximation (MP4)[34,35] with a large basis set to compute 69 accurate *ab initio* He-H_2 energies in the van der Waals well and repulsive wall (15 R values in 2 $a_0 \leq R \leq 15$ a_0), for three orientations ($\gamma = 0°$, 45°, and 90°); separations R over the full range were reported only for the equilibrium H_2 molecule ($r = 1.449$ a_0), but a few ($R = 3.0$, 5.0, 6.5, and 8.0 a_0) were also reported at $r = 1.28$ and 1.618 a_0. (Note that Tao reported a maximum van der Waals well depth of 47 μE_h for He plus equilibrium H_2, at $r = 1.449$ a_0, $R = 6.5$ a_0, $\gamma = 0°$.) This subset of *ab initio* energies was referred to as the "Tao vdW" points.

For $r = 1.28$, 1.449, and 1.618 a_0, a relatively simple formula was fitted (for $\gamma = 0°$, 45°, and 90°) to the difference between the Tao[31] *ab initio* energies and the analytic surface of Schaefer & Köhler[22], and used to interpolate corrected energies at R and γ values not computed by Tao. These generated energies were referred to as "gen-Tao vdW" points.

For $r \leq 1.1$ a_0 and $r \geq 1.8$ a_0, energies were obtained from a version of the Muchnick & Russek[24] surface where the van der Waals term in their potential had been modified slightly to better fit the newer, more accurate Tao[31] *ab initio* energies (this surface is the one described in § II D 2). These generated energies were referred to as "gen-MR vdW" points.

For $r = 0.9$, 1.1, 1.8, and 2.0 a_0, energies from the Schaefer & Köhler[22] surface were also used (since it might give energies of comparable accuracy in this region). These generated energies were referred to as "gen-SK vdW" points.

5. Accurate "H-He" energies

A set of 1887 energies was generated for conformations consisting of a relatively close-together $H - He$ pair (of size R_A) with a very-distant H atom (a distance R' from the nearest point on the line segment R_A). Since the contribution of the isolated H atom was completely negligible, this subset was referred to as "H-He" points.

For $R_A \leq 4$ a_0, the H-He energies were generated from 31 *ab initio* $He + H$ energies that we computed at $R_A = \{0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.35, 1.4, 1.45, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 3.6, 3.8, \text{and } 4.0\ a_0\}$ (note that we had also computed 9 other *ab initio* energies at larger R_A values, which were not used). For $R_A > 4$ a_0, the H-He energies were generated from the $H - He$ pair-wise potential of Dove & Raynor[20] (which used slightly-modified functions from Gengenbach, Hahn, & Toennies[36]), with values chosen at $R_A = \{4.25, 4.5, 4.75, 5.0, 5.25, 5.5, 5.75, 6.0, 6.25, 6.5, 6.75, 7.0, 7.25, 7.5, 7.75, 8.0, 8.5, 9.0, 10.0, 11.0, 12.0, 15.0, \text{and } 20.0\ a_0\}$. Distances to the isolated H atom comprised $R' = \{7.1, 7.7, 8.3, 9.3, 10.3, 11.4, 13.1, 14.8, 16.5, 19.0, 22.0, \text{and } 25.0\ a_0\}$. Conformations were discarded whenever pair-wise $H - He$ or $H - H$ energies from the isolated H atom would be non-negligible compared to the energy at separation R_A of the relatively close-together $H - He$ pair. (Although they were negligibly small for all retained conformations, pair-wise interaction energies of the isolated H atom with the other two atoms were nonetheless added to the main H-He energy, for consistency with the "H-He plus H" energies of § II A 6 below). For each pair $\{R_A, R'\}$, eight angular grid-points were

chosen, four with $r = R'$ at $\varphi = 90°$, 120°, 150°, and 180°, and four with $R_B = R'$ at $\Psi = 90°$, 120°, 150°, and 180° (see Fig. 1) — a "race-track" shaped grid. For $R' > 2\,R_A$, the two points at $\varphi = 90°$ and $\Psi = 90°$ were replaced by a single point at a distance R' from the midpoint of R_A. Higher weight was given to conformations with large R_A values (as discussed in § II D 1).

6. Approximate "H-He + H" energies

A set of 830 approximate energies was generated to constrain the fit in regions not covered by the above grids, where the two H atoms were fairly far apart ($5\ a_0 \le r \le 19\ a_0$), and at least one of the H atoms was well separated from the He atom so that the interaction energy of this isolated H atom would be small compared to the uncertainty in our *ab initio* energies (but might not be small compared to the van der Waals well depth). Since the contribution of the isolated H atom was small but not completely negligible, this subset was referred to as "H-He + H" points.

The grid was similar to the main grid, but with $r = \{5.0, 6.55, 7.2, 7.9, 8.5, 9.25, 9.9, 10.6, 11.4, 14.8,$ and $19.0\ a_0\}$ and $R_A = \{0.8, 1.1, 1.4, 1.7, 2.1, 2.5, 2.9, 3.3, 3.7, 4.2, 4.7, 5.2, 5.7, 6.2, 6.7, 7.3, 8.05, 9.0, 10.0, 12.0, 15.0,$ and $20.0\ a_0\}$; the minimum angular spacing was 15° in θ. Many parts of this grid were in fact more coarsely spaced than this in R_A and/or θ, and for $r = 7.2$, 8.5, 9.9, and 11.4 almost all gridpoints with $\theta < 105°$ were omitted. In addition, conformations relatively close to areas covered by the other grids were omitted; this restriction eliminated cases where more than one of the three pair-wise potentials would exceed $\sim 0.1\ mE_h$. Approximate energies were thus computed by simply summing the three pair-wise potentials. Cases where this computed energy would be expected to be quite accurate (i.e., large R_A) were given increased weight (as discussed in § II D 1).

B. Ab initio computations and analysis

1. Computational methods and CPU-time

The *ab initio* computations and analysis of errors largely follow the methods described in our previous papers[37,38,3,4,8,9]. Energies were computed using the workstation version[39] of Buenker's MRD-CI program[40]. At the position of each of the two hydrogen atoms, we used the $(9s3p2d)/[4s3p2d]$ Gaussian basis set from Boothroyd *et al.*[37,38] (with d-functions optimized for H_2). At the position of the helium atom, we used a $(10s4p2d)/[5s4p2d]$ Gaussian basis set, with s-functions taken from a (partially-decontracted) set from Huzinga[41], and the p- and d-functions chosen based on hints from several papers[41–43]; the basis set error for an isolated He atom is only 1.028 mE_h with this basis set. These two basis sets are given in Table I; they comprise the largest combination of basis sets allowed by array sizes in the MRD-CI program[39]. For most conformations, molecular orbitals were obtained from closed shell SCF; if closed shell SCF iterations were slow to converge, open shell SCF or mixed (1 closed, 2 open) shell SCF was automatically used instead. A configuration selection threshold of $T = 2\ \mu E_h$ was used for most points, with $T = 10\ \mu E_h$ used for 3006 high-energy cases at short distances (Buenker's MRD-CI program automatically extrapolates to

zero threshold, using truncated-CI energies at T and at $2T$ plus a parameterization of the estimated effects of neglected configurations). Multiple CI roots were computed, to ensure that the ground state energy was always among the energies reported by the program. A fairly extensive set of reference configurations was used for the CI calculation: the minimum size of C_R^2 (the sum of reference configuration C^2 values) was 0.969, and the *average* size was 0.982. (Note that additional reference configurations were included automatically if the program estimated that they would increase C_R^2 by more than 0.001 individually and by more than 0.003 collectively).

The above relatively small basis set and relatively large selection threshold were required by CPU-time considerations. So long as the *ab initio* energies are of accuracy comparable to the desired "chemical accuracy," the most important factor in obtaining a good fitted analytic surface is adequate coverage in the 3-dimensional conformation space of He-H_2, which requires large numbers of *ab initio* energies. In total, over a year of CPU-time plus I/O-time was required for the total of 23 406 *ab initio* computations (this includes 3203 points that were initially computed with only a single CI root, or with $T = 10\ \mu E_h$, and which were subsequently re-computed more accurately). These computations were distributed among about a dozen different computers, mostly workstations of various types, but including a couple of higher-performance machines. For the faster computers the I/O-time tended to dominate (due to large temporary files), but there was one exception to this: over two thirds of the energies were computed on a single higher-end SGI with a fast disk, which allowed computation of a typical *ab initio* energy in about 10 minutes. The other computers typically required 30 to 60 minutes per point, with some older machines requiring up to several hours per point.

2. Small corrections applied to the MRD-CI energies

Extrapolation to zero threshold (performed automatically by the MRD-CI program) and the Davidson correction to full CI are both standard procedures. Although the MRD-CI program uses only the energy at thresholds T and $2T$ to extrapolate to zero threshold, it also calculates the energy at $3T$ and $4T$; we therefore used these latter to make a very small (rms size 0.05 mE_h) "curvature correction" to the extrapolation (as Boothroyd *et al.*[37,38,8,9] did for H_4 — but unlike H_4, this was the *only* modification of the extrapolation). The average size of the extrapolation was 1.73 mE_h for the threshold $T = 2\ \mu E_h$, and about 3 times as large for $T = 10\ \mu E_h$. For the 17 025 cases where energies were obtained with both $T = 10$ and 2 μE_h, the rms difference between extrapolated energies was 0.17 mE_h. This should be an overestimate of the extrapolation error of either, suggesting that the rms error in the extrapolation to zero threshold is about 0.15 mE_h for $T = 10\ \mu E_h$ (used for 3006 high-energy points) and about 0.1 mE_h for $T = 2\ \mu E_h$ (used for the other 17 197 points). This error estimate is also supported by the fact that, for 996 points where a single-root $T = 0.4\ \mu E_h$ case was also computed, the rms difference was 0.12 mE_h relative to the 5-root $T = 2\ \mu E_h$ cases.

As in Boothroyd *et al.*[37,38,8,9], we made a parameterized Davidson-type correction to full CI, namely,

$$\Delta_{\text{full-CI}} = \lambda_{DC}^{(\text{SCF})} \Delta E_{DC} = \lambda_{DC}^{(\text{SCF})} (1 - C_R^2) \Delta E_{sd} / C_R^2 \quad , \tag{1}$$

where ΔE_{sd} is the contribution of single and double excitations to the energy, and ΔE_{DC} is the standard Davidson correction. For single-reference closed-shell configurations with at least four electrons, if the effects of quadruple excitations are assumed to be much larger than triple excitations, then the above formula for the Davidson correction (with an expected value of $\lambda_{DC} \sim 0.5$ for closed-shell He-H_2) can be derived theoretically[44–47]. We not only used multiple references, we also used open and mixed shell SCF cases; we therefore obtained values for the parameter λ_{DC} by minimizing the average differences between MRD-CI energy values computed for the same point with different reference sets (which thus had different sizes for the Davidson correction). We used $\lambda_{DC}^{(\mathrm{SCF})} = 0.52$, 0.16, and 0.32 for closed, open, and mixed SCF cases, respectively; the average size of ΔE_{DC} was 1.31, 1.90, and 0.60 mE_h, respectively, yielding full-CI corrections averaging 0.68, 0.30, and 0.19 mE_h, respectively. For the 12 077 cases where the program obtained improved reference sets, and thus two different-sized Davidson corrections were available, the rms difference between the resulting energies was 0.16 mE_h. This may be a slight underestimate of the energy uncertainty resulting from the Davidson correction, since on average the initial reference set yielded a Davidson correction only 1.75 times as large as the improved reference set, and we have no information from cases where only one value of the Davidson correction was available. There are suggestions of a systematic uncertainty of order 0.1 in the preferred value of $\lambda_{DC}^{(\mathrm{SCF})}$, leading to a systematic uncertainty in the energy correction of order 0.13 mE_h. We estimate the total uncertainty in the Davidson-type correction to full-CI to be about 0.3 mE_h.

The final energy value E_{final} was obtained from the estimated full-CI energy using a London-type basis correction similar to that for H_4 described in Boothroyd *et al.*[37,38,8,9]. The H_4 London-type basis correction is obtained by taking the difference between the H_4 London energy with the accurate H_2 singlet curve and the H_4 London energy with the H_2 singlet curve computed using the finite basis set. For He-H_2, a London-type basis correction was obtained by taking the H_4 basis correction, letting one distance go to zero (to yield the He atom), and replacing the contribution from the zero distance by the isolated He atom basis correction of 1.028 mE_h. (After doing the algebra, this turns out to mean taking the H_2 singlet basis correction for the H_2 molecule, adding the isolated He atom basis correction, and adding half of the H_2 singlet basis correction for each of the $\mathrm{H} - \mathrm{He}$ distances.) This yields the correct basis correction in the limit of a separated He atom plus H_2 molecule. The average size of this correction was 2.32 mE_h (or 1.29 mE_h if one excludes the He atom basis correction). Boothroyd *et al.*[8,9] concluded in the case of H_4 that the systematic error in the basis correction might be as much as 1 mE_h. For He-H_2, the error should be somewhat smaller; one might estimate that the error could be comparable to the non-H_2, non-He part of the basis correction, i.e., about 0.6 mE_h. Comparison of a number of energies that were also computed using smaller basis sets suggests that this error estimate is reasonable, as does comparison with energies computed by Tao[31] using a larger basis set. The error in the basis correction will be larger for geometries where the He atom is close to one (or both) of the H atoms, and smaller for geometries where the He atom is relatively far from the H atoms.

3. Best MRD-CI energies and estimated errors

Where more than one energy value had been computed for a given conformation, a weighted average of the "best" values was obtained for use in the fitting procedure. (In general, multiple-root cases were preferred to single-root cases, closed shell cases to open and mixed shell cases, lower thresholds to higher ones, and smaller Davidson corrections to larger ones.) The extrapolation to zero threshold and the Davidson-type correction to full CI appear to yield largely "random" errors; the systematic components of these, though not completely negligible, are much smaller than the systematic error in the basis correction. Thus the total uncertainty in the *ab initio* energies probably comprises a "random" error of about 0.2 mE_h and a systematic error of order 0.6 mE_h. The *ab initio* energies are available from EPAPS[32] or from the authors[33].

C. Functional Representation of the He-H_2 Surface

This section lays out the equations which underlie our fits to the He-H_2 *ab initio* data. Since the Muchnick & Russek[24] surface was the best of the previous surfaces, we based our analytic formulae on theirs, but added terms to get more flexibility. The surface is defined in terms of the three interatomic distances $\vec{\mathcal{R}} \equiv \{r, R_A, R_B\}$, where r is the $\mathrm{H}-\mathrm{H}$ distance and R_A and R_B are the two $\mathrm{He}-\mathrm{H}$ distances (see Fig. 1). As in Muchnick & Russek[24], new (basically spheroidal) variables are defined as

$$\bar{R} \equiv \frac{R_A + R_B}{2} , \qquad \eta \equiv \frac{R_A - R_B}{r} . \tag{2}$$

Note that $\bar{R} \to R$ and $\eta \to \cos\gamma$ for $R \gg r$, i.e., the usual Legendre function expansion terms $P_L(\cos\gamma)/R^n$ in the van der Waals well can be replaced by $P_L(\eta)/\bar{R}^n$. However, the functional form of the fitted surface does *not* explicitly contain either R or γ.

1. Functional form of our fitted surfaces

The full He-H_2 potential energy surface is given by

$$V_{\mathrm{He\,H_2}}(\vec{\mathcal{R}}) = V_{\mathrm{H_2}}(r) + V_{int}(\vec{\mathcal{R}}) , \tag{3}$$

where

$$V_{int}(\vec{\mathcal{R}}) = V_C(\vec{\mathcal{R}}) + V_H(\vec{\mathcal{R}}) + V_{cb}(\vec{\mathcal{R}}) + V_d(\vec{\mathcal{R}}) + V_D(\vec{\mathcal{R}}) . \tag{4}$$

The $V_{\mathrm{H_2}}$ term is the isolated-H_2-molecule potential (which is of course a function only of the distance r between the two H atoms); for this we used the accurate H_2-molecule potential of Schwenke[28]. The interaction energy V_{int} between the H_2 "molecule" and the He atom was fitted by several terms, functions of all three interatomic distances (all terms except V_D being generalizations of terms used by Muchnick & Russek[24]). The basic forms of these terms correspond to Coulomb, Pauli, and dispersion terms, as discussed by Muchnick &

Russek[24]. However, the added flexibility (provided by the summations in the equations below) largely blurs the distinctions between these types of terms in our fitted surfaces.

The V_C term has the form of a Coulomb term (omitting the H − H contribution):

$$V_C(\vec{\mathcal{R}}) = A_C(\vec{\mathcal{R}}) \left[1 + \left(\frac{2}{A_C(\vec{\mathcal{R}})} - 1\right) P_2(\eta)\right] \left[F_C(R_A) + F_C(R_B)\right] , \tag{5}$$

where

$$F_C(R_\nu) = \frac{e^{-\lambda_C R_\nu}}{R_\nu} \left[1 + \sum_{j=1}^{J_C} \frac{b_C(j)}{{R_\nu}^j}\right] \qquad \text{for} \quad \nu = A, B \tag{6}$$

and

$$A_C(\vec{\mathcal{R}}) = 2 - e^{-\beta_C r} \sum_{m=1}^{M_C} r^m \sum_{\substack{L=0 \\ L \text{ even}}}^{L_C} P_L(\eta) \sum_{i=0}^{I_C} \frac{a_C(m, L, i)}{\bar{R}^i} \; ; \tag{7}$$

note that the lower limit $m = 1$ on the summation (rather than $m = 0$) was chosen in order that $A_C(\vec{\mathcal{R}}) \to 2$ for $r \to 0$ as well as for $r \to \infty$ (as in the similar term of Muchnick & Russek[24]). Parameters to be fitted include λ_C, β_C, and all of the $a_C(m, L, i)$ and $b_C(j)$. Our final adopted surface had $\{M_C, L_C, I_C, J_C\} = \{2, 2, 1, 5\}$, but we tested fits with ranges as high as $\{3, 4, 2, 5\}$. Note that Muchnick & Russek[24] used this V_C term with only the parameters λ_C, β_C, and $a_C(1, 0, 0)$.

The V_H term looks like a Pauli-type interaction between the He atom and each of the H atoms:

$$V_H(\vec{\mathcal{R}}) = A_H \left[1 - e^{-\beta_H r} \sum_{m=0}^{M_H} r^m \sum_{\substack{L=0 \\ L \text{ even}}}^{L_H} P_L(\eta) \sum_{i=0}^{I_H} \frac{a_H(m, L, i)}{\bar{R}^i}\right] \left[F_H(R_A) + F_H(R_B)\right] , \tag{8}$$

where $a_H(0, 0, 0) \equiv 1$ is a constant (in order that the lowest-order form of V_H vanish at $r = 0$, for consistency with the similar term of Muchnick & Russek[24]), and

$$F_H(R_\nu) = e^{-\lambda_H R_\nu} \left[1 + \sum_{j=1}^{J_H} \frac{b_H(j)}{{R_\nu}^j}\right] \qquad \text{for} \quad \nu = A, B \; . \tag{9}$$

Parameters to be fitted include A_H, λ_H, β_H, and all of the $b_C(j)$ and $a_C(m, L, i)$ except for $a_H(0, 0, 0)$. Our final adopted surface had $\{M_H, L_H, I_H, J_H\} = \{2, 4, 2, 5\}$, but we tested fits with ranges as high as $\{3, 6, 3, 5\}$. Note that Muchnick & Russek[24] used this V_H term with only the parameters A_H, λ_H, and β_H.

The V_{cb} term looks like a Pauli-type interaction between the He atom and that part of the electronic distribution drawn in to form the H − H covalent bond:

$$V_{cb}(\vec{\mathcal{R}}) = e^{-\beta_{cb} r - \Lambda_{cb}(r)\bar{R}} \sum_{m=0}^{M_{cb}} r^m \sum_{\substack{L=0 \\ L \text{ even}}}^{L_{cb}} P_L(\eta) \sum_{i=I_{cb}}^{J_{cb}} a_{cb}(m, L, i) \bar{R}^i \; , \tag{10}$$

where

$$\Lambda_{cb}(r) = \varepsilon_{cb} + \zeta_{cb} e^{-\delta_{cb} r} \ . \tag{11}$$

Parameters to be fitted include β_{cb}, ε_{cb}, ζ_{cb}, δ_{cb}, and all of the $a_{cb}(m,L,i)$. Our final adopted surface had $\{M_{cb}, L_{cb}, I_{cb}, J_{cb}\} = \{2, 4, -1, 2\}$, but we tested fits with ranges as high as $\{3, 6, -2, 2\}$. Note that Muchnick & Russek[24] used this V_{cb} term with only the parameters β_{cb}, ε_{cb}, ζ_{cb}, δ_{cb}, $a_{cb}(0,0,0)$, and $a_{cb}(0,2,0)$.

The V_d term is a dispersion (long-range) interaction between the He atom and the H_2 molecule:

$$\begin{aligned} V_d(\vec{\mathcal{R}}) = - f_{Ad}\, A_d(r) \Bigg\{ & \frac{1}{\bar{R}^6 + {C_1}^6} \left[1 + f_{\alpha d}\, \alpha_d(r) P_2(\eta) + \sum_{\substack{L=4 \\ L \text{ even}}}^{L_d} e^{\varepsilon(6,L)\, r} P_L(\eta) \sum_{m=1}^{M_d} a_d(6,L,m)\, r^m \right] \\ & + \frac{C_3 \left[1 + f_{\alpha d}\, \alpha_d(r) P_2(\eta) \right]}{\bar{R}^8 + {C_2}^8} \\ & + \sum_{\substack{n=8 \\ n \text{ even}}}^{N_d} \frac{1}{\bar{R}^n + [C_d(n)]^n} \sum_{\substack{L=0 \\ L \text{ even}}}^{L_d} e^{\varepsilon(n,L)\, r} P_L(\eta) \sum_{m=\min\{L,1\}}^{M_d} a_d(n,L,m)\, r^m \Bigg\} \ , \end{aligned} \tag{12}$$

where

$$A_d(r) = \left(4.790521\, r^2 - 2.494444\, r + 3.008850 \right) e^{-0.5891\, r} \tag{13}$$

and

$$\alpha_d(r) = \left(0.238248\, r^2 - 0.126241\, r \right) e^{-0.8075\, r} \tag{14}$$

(where r is in a_0). Note that the values of $\varepsilon(n,L)$ were close to -1.0 in all fits, so that the term containing C_2 and C_3 (retained for consistency with the form of Muchnick & Russek[24]) has a significantly different behavior from the $n = 8$ term in the last of the summations. Note also that only the $L = 0$ terms are allowed to have $m = 0$, since the angular variation in V_d must vanish for $r \to 0$. For cases (such as the final fitted surface) with an upper limit $L_d < 4$, the last term in the large square brackets is omitted (since $L = 4$ is the lowest L value allowed in that term).

The 7 parameters quoted in the terms $A_d(r)$ and $\alpha_d(r)$ above were obtained by fitting (with an accuracy of a few percent) to the accurate values calculated by Thakkar *et al.*[23] as a function of r for these asymptotic terms, as shown in Figure 4. The remaining parameters to be fitted include f_{Ad}, $f_{\alpha d}$, C_1, C_2, C_3, and all of the $C_d(n)$, $\varepsilon(n,L)$, and $a_d(n,L,m)$. Since the parameters f_{Ad} and $f_{\alpha d}$ multiply a version of the asymptotic long-range form (although "softened" at small R by the parameter C_1), these parameters might be expected to remain within a few percent of unity in a fit; this did indeed prove to be the case, particularly for f_{Ad} (note that α_d is much smaller than A_d, and thus $f_{\alpha d}$ is determined with less precision by the fitting process).

Our final adopted surface had $\{N_d, L_d, M_d\} = \{8, 2, 1\}$, but we tested fits with ranges as high as $\{10, 4, 2\}$. Note that Muchnick & Russek[24] used this V_d term with only the parameters C_1, C_2, and C_3; they also replaced $P_2(\eta)$ with η^2, and set $A_d(r) = 0.81 + 1.92\, r$ and $\alpha_d(r) = 0.112 + 0.06\, r$ (although they noted that these latter two functions really

should have used a decreasing exponential in r). Their versions of $A_d(r)$ and $\alpha_d(r)$ are also illustrated in Figure 4, although their $\alpha_d(r)$ is not strictly comparable, since it multiplies η^2 rather than $P_2(\eta)$. (In spite of the fact that their A_d and α_d increase without limit with increasing r, the V_d term used by Muchnick & Russek[24] remains bounded, since by definition $\bar{R} \geq r/2$, and thus their form of V_d is proportional to r^{-4} for $r \gg C_1$.)

The V_D term is a long-range interaction between the He atom and separated H atoms:

$$V_D(\vec{\mathcal{R}}) = S_D(r)\Big[F_D(R_A) + F_D(R_B)\Big] \,, \tag{15}$$

where

$$F_D(R_\nu) = \sum_{\substack{n=6 \\ n \text{ even}}}^{N_D} \frac{a_D(n)}{{R_\nu}^n + [C_D(n)]^n} \qquad \text{for} \quad \nu = A, B \tag{16}$$

and $S_D(r)$ is a switch function to turn V_D on with increasing r:

$$S_D(r) = 1 - \frac{B_D(r)}{\Big\{6^2 + [B_D(r)]^2\Big\}^{1/2}} \,, \tag{17}$$

with

$$B_D(r) = 4.790521\left(r^2 + 6\right) e^{-0.5891\, r} \tag{18}$$

(where r is in a_0). The softening parameter of 6 in the switch function $S_D(r)$ above (which regulates the switchover at relatively small r) may be considered to be a surface parameter, but the other constants in the switch function were taken from $A_d(r)$ in Equation (13), so that the V_D term would be turned on with increasing r as the V_d term was being turned off by $A_d(r)$. Other parameters to be fitted include all of the $C_D(n)$ and $a_D(n)$. Our final adopted surface had $N_D = 8$, but we also tested fits with $N_D = 10$. Note that Muchnick & Russek[24] did *not* include any such V_D term.

2. Other types of terms tested

Various index ranges were tried for the summations in the above equations. In addition, it was tested whether a better fit could be obtained if the terms V_C, V_H, and V_{cb} were multiplied by an exponential cutoff in the overall size of the He-H_2 conformation, namely, by a factor

$$e^{-\beta\rho^\alpha} \,, \qquad \text{where} \quad \rho = \left(r^2 + {R_A}^2 + {R_B}^2\right)^{1/2} \tag{19}$$

(with β and α being parameters that could be fitted). It turned out that such a cutoff did not improve the fit, and it was discarded.

It was tested whether adding a many body expansion term V_M to the He-H_2 PES would improve the fit:

$$V_M(\vec{\mathcal{R}}) = \sum_{k=1}^{K_M} \sum_{j=0}^{\min\{k, M_M-1-k\}} \left[{p_A}^k {p_B}^j + {p_A}^j {p_B}^k \right] \sum_{i=1}^{\min\{I_M, M_M-j-k\}} a_M(i,j,k)\, {p_r}^i \;, \qquad (20)$$

where

$$p_r = r e^{-\beta_M r} \;, \qquad \text{and} \qquad p_\nu = R_\nu e^{-\beta_M R_\nu} \qquad \text{for} \quad \nu = A, B \;, \qquad (21)$$

with parameters to be fitted including β_M and all of the $a_M(i,j,k)$. This V_M term could also be multiplied by an exponential cutoff as given by Equation 19. Various orders M_M and individual index limits K_M and I_M were tested, but V_M always turned out to be of very little help in improving the fit (compared to the other terms of § II C 1), so it was discarded.

D. Our Approach to Fitting

This section outlines in general the steps followed in developing the terms and optimizing the parameters in our analytical representation of the He-H_2 surface. The details of the equations are given in § II C 1 above.

1. Weights applied to fitted energies

In this section, the weights referred to are those that were used to multiply the deviations between the fit and the data, before squaring these weighted deviations to calculate the rms of the fit. Note that, frequently, the "weight" in a fit is defined as that applied to the square of the (unweighted) deviation, i.e., with this latter definition the weight would be the *square* of the values reported below.

Several different criteria were used to determine weight values. If more than one applied, the final combined weight used was the *product* of the individual weight factors from the following separate criteria.

High-energy portions of the surface are less likely to be accessed in collisions of a hydrogen molecule with a helium atom; also, higher *ab initio* energies have larger uncertainties than lower energies. Thus *all* points with high energy E were given reduced weight, namely,

$$w_E(E) = \begin{cases} 1 & , \; E \le 0.2\ E_h \\ (0.2\ E_h)/E & , \; E > 0.2\ E_h \end{cases} \qquad (22)$$

i.e., a weight inversely proportional to the energy E for cases with energies more than about 2.2 times the H_2 dissociation energy above that of an equilibrium H_2 molecule plus a separated He atom.

For non-compact conformations where the He atom was relatively far from either H atom, the (absolute) size of the uncertainties in our *ab initio* energies was expected to be smaller than for more compact conformations (especially for the error in the basis correction: see § II B 2). Thus our *ab initio* energies (described in § II A 1 and § II A 2) were given an additional weight factor

$$w_{nc}(R_m) = \begin{cases} 1 & , \; R_m \leq 3\ a_0 \\ (R_m - 2) & , \; 3\ a_0 < R_m < 4\ a_0 \\ 2 & , \; R_m \geq 4\ a_0 \end{cases} \quad , \qquad \text{where} \quad R_m = \min\{R_A, R_B\} \tag{23}$$

(i.e., a weight increasing linearly from 1 to 2 as the shortest He − H distance R_m is increased from 3 to 4 a_0).

The van der Waals He + H_2 points described in § II A 4 are expected to have smaller (absolute) uncertainties as the distance R between the H_2 molecule and the He atom increases; they should also be most accurate for near-equilibrium H_2 molecule sizes (i.e., $r \sim 1.4\ a_0$). For these "vdW" points, two weight factors w_R and w_r were applied:

$$w_R(R) = \begin{cases} 0.3 & , \; R \leq 3\ a_0 \\ 1 & , \; R = 4\ a_0 \\ 10 & , \; R = 5\ a_0 \\ 300 & , \; R = 6\ a_0 \\ 1200 & , \; R = 6.5\ a_0 \\ 3000 & , \; R \geq 10\ a_0 \end{cases} \quad , \qquad w_r(r) = \begin{cases} 0.1 & , \; r \leq 0.6\ a_0 \\ 0.3 & , \; r = 0.9\ a_0 \\ 1 & , \; r = 1.2\ a_0 \\ 1 & , \; r = 1.62\ a_0 \\ 0.3 & , \; r = 2.0\ a_0 \\ 0.1 & , \; r = 2.4\ a_0 \\ 0.01 & , \; r \geq 4.0\ a_0 \end{cases} \tag{24}$$

(where w_R and w_r were interpolated linearly in between the distance values specified above). The "Tao vdW" points (*ab initio* energies of Tao[31]) were given an additional weight factor $w_T = 3$, and the "gen-Tao vdW" points were given an additional weight factor $w_{gT} = 2$. Note that quite a large number of fits were performed where w_R was either increased or decreased relative to Equation (24) by about a factor of 3 for $R \gtrsim 5\ a_0$, but these fits were discarded — the weight w_R of Equation (24) appeared to work the best, neither over-emphasizing nor under-emphasizing the van der Waals well energies relative to the *ab initio* energies.

The "H-He" points described in § II A 5 should be as accurate as any of our other *ab initio* energies for $R_A \lesssim 5\ a_0$ (where R_A is the *shorter* H − He distance for these points: $R_A \leq R_B$); they should be rather more accurate in the region of the H − He van der Waals well and further out. Thus, a weight factor $w_{HHe}(R_A) = \max\{w_R(R_A), 1.0\}$ was applied to them, where w_R has the form given in Equation (24).

The "H-He + H" points described in § II A 6 are expected to be somewhat less accurate; a weight factor $w_{HHe+H}(R_A) = \max\{0.01\, w_R(R_A), w_{nc}(R_A)\}$ was applied to them, where w_{nc} has the form given in Equation (23).

2. Modified Muchnick & Russek surface

The first step was to get a very slightly modified version of the Muchnick & Russek[24] surface, which we refer to as the "modMR" surface. This differed from the surface of Muchnick & Russek[24] in that we used the lowest-order version of our functional form for V_d, including use of our $A_d(r)$ and $\alpha_d(r)$ from Equations 13 and 14, and the fact that $\alpha_d(r)$ multiplied $P_2(\eta)$ rather than η^2. We also added the lowest order V_D term (i.e., $N_D = 6$) from Equation (15). Since we used our already-fitted versions (8 parameters) of $A_d(r)$, $\alpha_d(r)$ and $S_D(r)$ from Equations (13), (14), and (17), this initial "modMR"

surface had 19 other basic parameters that could be fitted. There were 9 short-range non-linear parameters $\{\lambda_C, \beta_C, a_C(1,0,0), \lambda_H, \beta_H, \beta_{cb}, \varepsilon_{cb}, \zeta_{cb}, \delta_{cb}\}$, 3 short-range linear parameters $\{A_H, a_{cb}(0,0,0), a_{cb}(0,2,0)\}$, 3 long-range non-linear parameters $\{C_1, C_2, C_D(6)\}$, and 4 long-range linear parameters $\{f_{Ad}, f_{\alpha d}, C_3, a_D(6)\}$. For this first surface, we obtained $a_D(6)$ from Gengenbach, Hahn, & Toennies[36], and set $C_D(6) = 6.7876662396\ a_0$ (actually, $2^{1/6} \times 3.2$ Å — a reasonable softening value to keep the V_D term from contributing significantly at relatively compact geometries). We set the other parameters equal to the values used by Muchnick & Russek[24], and fitted only the parameters f_{Ad} and $f_{\alpha d}$, using only the "Tao vdW" and "gen-Tao vdW" points described in § II A 4. This procedure yielded values of $f_{Ad} = 0.98234$ and $f_{\alpha d} = 1.261$, reasonably close to unity (i.e., at large $H_2 - He$ separations R, this surface still had a form very close to that calculated by Thakkar *et al.*[23]). The surface was almost identical to that of Muchnick & Russek[24], but was a significantly better fit to the more recent (and more accurate) van der Waals *ab initio* energies of Tao[31] (unweighted rms error of 1 μE_h for these points, rather than 3 μE_h). This "modMR" surface was the one used to generate the "gen-MR vdW" points described in § II A 4.

3. Optimization of fitted parameters

Several hundred fits were subsequently performed to the full set of 20 203 *ab initio* energies and 4862 generated points, with various ranges for the indices in the summations in the surface equations of § II C 1; the number of parameters actually fitted varied from a couple of dozen to about 150. Fitting was performed using the NAG Fortran library non-linear fitting routine E04FDF to minimize the weighted rms error with respect to the 20 203 *ab initio* plus 4862 generated energies, using the weights of § II D 1 (the covariance matrix was then produced by the NAG routine E04YCF) — a typical fit required a few hours on a higher-end computer. The first fit described in the previous paragraph supplied initial values for the basic parameters in the next fits, often with values shifted randomly; the added parameters in the summations were either initialized to zero or to random values. A reasonably good fit was frequently used as the basis for subsequent larger fits (sometimes with random shifts applied to the parameters before fitting). There were some exceptions to this: the 3 basic long-range non-linear parameters $\{C_1, C_2, C_D(6)\}$ were not re-fitted, and other long-range non-linear parameters were generally not fitted, but just given a reasonable initial value. Also, 3 of the basic long-range linear parameters, namely, $\{f_{Ad}, f_{\alpha d}, C_3\}$, were not re-fitted until near the end of the fitting process. Note that in our final adopted fit the values of $f_{Ad} \approx 0.9552$ and $f_{\alpha d} \approx 1.149$ were near unity; since these parameters multiply forms $A_d(r)$ and $\alpha_d(r)$ that had been fitted with an accuracy of a few percent to the values of Thakkar *et al.*[23] (as discussed in § II C 1), this final adopted fit yielded an asymptotic interaction between the H_2 molecule and the He atom very close to that calculated by Thakkar *et al.*[23]

4. Selection of a good fitted PES

For a quick comparison of the quality of all the different fits, we considered the full weighted rms minimized by the fitting program with respect to all 24 804 fitted points, and

also the energy-weighted rms for 9 subsets of these points, where only the weight w_E from Equation (22) had been applied. These 9 subsets comprised:

1. all 20 203 *ab initio* energies,
2. the 14 585 *ab initio* energies with $E < 0.174\ E_h$ (i.e., below about twice the H_2 dissociation energy),
3. the 7177 *ab initio* energies below the H_2 dissociation energy ($E < 0$),
4. 931 *ab initio* energies with near-equilibrium H_2 ($1.2\ a_0 \le r \le 1.62\ a_0$) and a not-too-distant He ($2\ a_0 < R \le 4\ a_0$),
5. 468 fairly-close "H-He" points (with $2\ a_0 < R_A \le 4\ a_0$),
6. 143 van der Waals "H-He" points (with $6\ a_0 \le R_A < 10\ a_0$),
7. 231 "Tao vdw" and "gen-Tao vdw" H_2+He van der Waals points with near-equilibrium H_2 (i.e., $1.28\ a_0 \le r \le 1.62\ a_0$, and $6\ a_0 \le R < 10\ a_0$),
8. 434 "gen-MR vdW" and "gen-SK vdW" van der Waals points with small-H_2 ($r < 1.2\ a_0$ and $6\ a_0 \le R < 10\ a_0$),
9. 714 "gen-MR vdW," "gen-SK vdW," and "H-He + H" van der Waals points with large-H_2 ($r > 1.62\ a_0$ and $6\ a_0 \le R < 10\ a_0$).

For a few cases where the above rms values suggested that the surface was among the best of the fits, a collection of weighted, energy-weighted, and unweighted rms values was considered for a more comprehensive list of about 100 subsets.

The fitting program produced estimated uncertainties of the fitted parameters (from the NAG covariance routine E04YCF). When an enlarged fit did not significantly reduce the rms error and yielded new parameters whose values were not significant (i.e., comparable to or smaller than the estimated errors in these parameters), this suggested that the point of diminishing returns had been reached — larger fits would presumably yield spurious "wiggles" rather than an improved fit. This criterion was used to obtain a near-optimum set of index ranges for the summations in the surface equations of § II C 1 (although a number of larger fits were performed, with different initial parameter values, to make certain that no significant improvement was indeed possible). Finally, for a few of the "best" surfaces with near-optimum index ranges (and also a few very large fits), scatterplots were made of the error vs. the *ab initio* energy, and a couple of dozen contour plots of each surface were examined, as well as over a hundred plots of energy vs. one of the He-H_2 distances or angles (in these latter plots, the surfaces could be directly compared to the *ab initio* energies along some cut of the PES). Our adopted He-H_2 PES was chosen in this manner from among the "best" surfaces with near-optimum index ranges. A Fortran program to compute this analytic PES (including analytic first derivatives) is available from EPAPS[32] or from the authors[33].

We have used a relatively large number of linear parameters in order to fit our He-H_2 data. Our final BMP surface uses 112 parameters, with a few of the non-linear parameters

being given reasonable constant values as discussed in § II C 1, and a few others being set by early (small) fits and held constant thereafter (such as the parameters in A_d and α_d); in the final fitted surface, the 9 non-linear parameters that were refitted at that point were all significant (at better than the 10-σ level), and only 2 of the 87 linear parameters had values smaller than their estimated errors, with only 3 others having less than a 3-σ significance level. These 112 parameters of our BMP surface were used to fit a total of 25 065 points (20 203 of which are *ab initio* energies), a 224:1 ratio of points to parameters (180:1 if one considers *ab initio* points only). For comparison, Truhlar and Horowitz[1] fitted 287 *ab initio* H_3 points with about 23 parameters, a 12:1 ratio. Our BKMP2 H_3 surface[4] fitted 8559 H_3 points (7591 *ab initio* points) with about 120 parameters, a 71:1 ratio (63:1 for *ab initio*). Aguado *et al.*[7] fitted 6101 *ab initio* H_4 energies with 865 linear parameters, a 7:1 ratio. Our BMKP H_4 surface[8] uses 400 parameters to fit 61547 points (48180 *ab initio* points), giving a ratio of 154:1 (120:1 for *ab initio*). Thus the final fitted BMP He-H_2 surface of the present work actually uses relatively few fitting parameters for the number of points fitted, minimizing the risk of "overfitting," i.e., of spurious wiggles *between* fitted points.

Added confidence in our error estimates is provided by the recent extremely accurate H_3 surface of Wu *et al.*[5] Their *ab initio* energies and fitted surface are an order of magnitude more accurate than our earlier BKMP2 H_3 surface[4]; they found that our H_3 error estimates were essentially correct, both for our H_3 *ab initio* energies and our fitted H_3 surface. They report no spurious wiggles in our BKMP2 H_3 surface with sizes larger than its quoted accuracy. This suggests that our error estimates should be reasonable for the similarly-computed He-H_2 *ab initio* energies of the present work, and that our fitted BMP He-H_2 surface should likewise be free of large spurious wiggles between fitted points. Finally, random "interstitial" unfitted conformations were also tested directly, as described in § III A 1.

III. DISCUSSION

A. Accuracy of analytic He-H_2 surfaces

Our adopted surface (the "BMP" He-H_2 PES) contains a total of 112 parameters; even much larger surfaces (up to 175 parameters) yielded no significant improvement (having rms values only a few percent smaller). Figure 5 shows a scatterplot of the deviations relative to our *ab initio* energies. The surface fits best at low energy ($E \lesssim 0.15\ E_h$, i.e., energies below about twice the H_2-molecule dissociation energy, relative to the energy of an equilibrium H_2 molecule plus a distant He atom). However, even up to quite high energies, the surface lies within a few mE_h of most of the *ab initio* points; it turns out that most of the outliers are due to the fact that the surface can only produce an approximate fit to the conical intersection with the first excited state (this is discussed in more detail below, in § III A 3).

Table II compares the rms errors of four earlier analytic He-H_2 surfaces with rms errors of the adopted (BMP) surface of this work, for various subsets of *ab initio* and generated energies; the modified Muchnick & Russek (modMR) surface of § II D 2 (which was used to generate some of the constraining van der Waals points for non-equilibrium H_2) is also shown. "Energy-weighted" deviations were used to obtain the rms values of Table II, i.e., energies above about two H_2-dissociation energies were given reduced weight w_E according to the

formula of Equation (22). The only exception to this is the last line of Table II (subset 44), where the *fully weighted* rms values are reported for *all points* used in our He-H_2 fit (for our BMP surface, this is the rms value that the fitting program tried to minimize, as described in § II D 1 and § II D 3).

It should be noted that the earlier analytic He-H_2 surfaces were fitted to *ab initio* points in relatively restricted regions, compared to the region covered by the *ab initio* and generated points of the present work. The rms values given in *italics* in Table II indicate subsets of points that cover regions where the corresponding surface was not fitted, i.e., where energies predicted by these previous surfaces are being tested by our new *ab initio* energies. (The only unconstrained region of our fitted BMP surface is at very small interatomic separations, i.e., very high energies; this is discussed in § III A 6.)

The Wilson, Kapral, & Burns[17] (WKB) surface was fitted to a set of *ab initio* energies in the region $0.75\ a_0 \le r \le 5.0\ a_0$, $0.0\ a_0 \le R \le 5.0\ a_0$, plus the three pair-wise potentials. The Dove & Raynor[20] (DR) surface was essentially the same, with improved versions of the pair-wise potentials. Since the interaction part of these surfaces was fitted using only four free parameters, they lack sufficient flexibility for accuracy even in their fitting region.

The Schaefer & Köhler[22] (SK) surface was fitted to a set of *ab initio* energies in the region $1.28\ a_0 \le r \le 1.618\ a_0$, $3.0\ a_0 \le R \le 15.0\ a_0$ plus a somewhat less accurate set in the region $0.9\ a_0 \le r \le 2.0\ a_0$, $1.5\ a_0 \le R \le 8.0\ a_0$; it is quite accurate in these regions. However, it is designed to be interpolated on a grid in the region $0.9\ a_0 \le r \le 2.0\ a_0$, $R \ge 1.6\ a_0$, and therefore can be *highly* inaccurate when extrapolated outside this region (just how inaccurate depends on the extrapolation formula used — we just used an spline in r, even for extrapolation, but fitted exponentials to extrapolate the R spline to small R).

The Muchnick & Russek[24] (MR) surface was fitted to a set of energies in the region $1.2\ a_0 \le r \le 1.6\ a_0$, $0.0\ a_0 \le R \le 15.0\ a_0$; the very similar modMR surface of § II D 2 differs from the MR surface only in the parameters describing the $H_2 - He$ and $H - He$ van der Waals wells. These surfaces are accurate in the part of their fitting regions that lies at $R \gtrsim 3\ a_0$. However, although the MR surface was fitted to *ab initio* energies down to $R = 0$, it is a relatively poor fit at small R (subset 9 in Table II). This is due both to the fact that Muchnick & Russek[24] had only a few *ab initio* points in this region and to the fact that a good fit in this region requires a good deal of flexibility in the fitted surface (we found in our first, smallest fits that significant improvements at small R could be obtained only at the expense of worsening the fit in the van der Waals well, unless more parameters were added). Even at intermediate R ($2\ a_0 < R \le 3\ a_0$: subsets 6, 10, and 14 in Table II), the SK surface does somewhat better than the MR surface, provided that r is not too large, i.e., $r \le 2\ a_0$ (our BMP surface is of course significantly better than either, in these regions of small-to-intermediate R).

Table II demonstrates that, overall, the MR surface[24] is the best of the previous surfaces considered here. Nonetheless, even the MR surface does quite poorly overall, with an rms error of 28.6 mE_h relative to our *ab initio* energies (subset 1 in Table II). The present BMP surface is an order of magnitude improvement, with an rms error of 1.42 mE_h relative to the *ab initio* energies — of the same order as the estimated *ab initio* error of $\sim 0.6\ mE_h$. The advantage of the BMP surface is largest in the region that can be sampled by dissociative collisions, but even below the dissociation energy (i.e., $E < 0.0\ E_h$: subset 3 in Table II) the present BMP surface (rms of 0.48 mE_h) remains a significant improvement over the MR

surface (rms of 4.48 mE_h).

For the restricted region of not-too-large H_2-molecules ($r \leq 2\ a_0$) plus a not-too-close He atom ($R \gtrsim 3\ a_0$), subsets 7, 11, 15, and 31 in Table II show that both the SK and MR surfaces[22,24] do nearly as well as the present BMP surface, but the earlier WKB and DR surfaces[17,20] are an order of magnitude worse even in this region (and have no van der Waals well at all). As one might expect, both the SK and MR surfaces do better when extrapolated to small r (subsets 4 to 7 in Table II) than when extrapolated to large r (subsets 16 to 19), due presumably to the fact that reducing r tends to reduce the anisotropy with respect to γ at reasonable R values, while for large r the anisotropy can become very large (in fact, for $R \sim r/2$, the He atom approaches one of the H atoms for $\gamma \to 0$).

We performed a number of further tests to assess in more detail the quality of our BMP surface (and of the previous surfaces), as discussed in the following subsections.

1. Test for "interstitial" wiggles in the BMP surface

The rms errors of the 3500 unfitted "random" conformations of § II A 3 (subsets 23 to 25 in Table II) were compared to the rms errors of the 3500 fitted "random" conformations of § II A 2 (subsets 20 to 22), which had been generated with the same probability distribution. The similarity of these two sets of rms values is evidence that our BMP surface does not have large spurious wiggles between fitted *ab initio* points. The fact that the rms values for these sets of "random" conformations are also similar to the overall *ab initio* rms values (subsets 1 to 3, dominated by the main grid of § II A 1) yields added reassurance.

These "random" points of these subsets (20 to 25 in Table II) avoid duplicate conformations by a relatively small "avoidance margin" of $\sim 0.01\ a_0$ (see § II A 2), in order to keep the fraction of "avoided" conformation space small even in regions where the main grid of conformations was dense. We also considered rms errors of our BMP surface for "maximally interstitial" cases with larger avoidance margins: e.g., for a case where points from fitted subset 20 were discarded if they lay within a 3% avoidance margin, the remaining 2949 points had an (energy-weighted) rms error of 1.356 mE_h (0.528 mE_h for the 1140 points with $E < 0$), while the corresponding 2742 points from unfitted subset 23 had an rms error of 1.374 mE_h (0.536 mE_h for the 1063 points with $E < 0$). (Recall that the typical main-grid spacing was $\sim 7\%$ in r and R_A, $\sim 7\%$ in θ for $\theta > 120°$, and $\sim 13\%$ in θ for $\theta < 120°$: see § II A 1.) As expected, there was good agreement between rms values of fitted and unfitted "random" subsets independent of the size of the avoidance margin.

It would of course have been very surprising if any interstitial wiggles had showed up; as discussed at the end of § II D 4, the density of fitted *ab initio* points is high enough that there is no room for interstitial wiggles (unless we had used a great many more parameters in the fit). However, there are nonetheless certain regions where the BMP surface *does* have errors significantly larger than its typical ones, as discussed further below.

2. Contour plots of the BMP surface

The contour plots of Figure 6 show the overall shape of our fitted BMP surface for four orientations γ of the $H_2 - He$ separation R relative to the $H - H$ separation r. Note that

the total potential $V_{\mathrm{He\,H_2}}$ is represented in Figures 6 and 7 (not the interaction energy V_{int}), and thus its gradient gives the total force on the H and He atoms.

In Figure 6a, with $\gamma = 0°$, the He atom lies very close to one of the H atoms for $R \approx 0.5\,r$, yielding a double-sided "wall" in the contour plot. In Figure 6b, this "wall" is replaced by a ridge, since for $\gamma = 30°$ the He atom is closest to the H atom when $R = 0.5\,r$. The ridge-like structure at lower left in Figure 6c ($\gamma = 60°$) is probably due to the attempt by the surface to fit the nearby conical intersection with the first excited state, which is discussed in more detail in § III A 3 below.

From the slopes of the contour lines in Figure 6d (a T-shaped orientation, i.e., $\gamma = 90°$), one can see that, as R decreases, the He atom does not begin to "push apart" the H atoms until it is quite close to the H_2 molecule: the partial derivative $(\partial V/\partial r)_{R,\gamma}$ remains positive for $r \gtrsim 1.5\ a_0$ until R reaches $\sim 1.5\ a_0$, between the 100 and 200 mE_h contours. (There is a large repulsive force between the He atom and the H_2 molecule that begins much further out, as may be seen from the significant reduction in depth of the H_2 potential well as R is reduced below $\sim 4\ a_0$). The cuts at constant r in Figure 7 also illustrate this effect — the curves for larger r values lie above those for $r = 1.4\ a_0$, except at small R. This illustrates an effective barrier to dissociation of H_2 by He: at least in the T-shaped geometry, the He atom must have a high enough energy to sample the part of the surface where $(\partial V/\partial r) < 0$, i.e., considerably more than just the the H_2 dissociation energy.

For an equilibrium H_2 ($r = 1.4\ a_0$) and $\gamma = 90°$, Muchnick & Russek[24] found that, with their surface, the presence of the He atom actually caused an inward force on the H_2 molecule for $2.36\ a_0 \leq R \lesssim 6.5\ a_0$, i.e., $(\partial V_{int}/\partial r)_{R,\gamma} > 0$ in this region; for the BMP surface of the present work, the corresponding region is $2.03\ a_0 \leq R \leq 6.29\ a_0$. As Muchnick & Russek[24] pointed out, this "force reversal" is a real effect, even visible in the *ab initio* energies. Nonetheless, this "force reversal" is only in the interaction energy, and is a minor effect compared to the stability provided by H_2-molecule potential; as may be seen from Figure 6, the equilibrium H_2 size is almost independent of R, except at small R.

Figure 7 also illustrates that the nearly-vertical parts of the contours at the lower left in Figure 6d ($\gamma = 90°$) are a real reflection of the *ab initio* data (and not an artifact in the fit arising from fitting the smaller-γ region where the He atom passes close to one of the H atoms). For $\gamma = 90°$ and $r \sim 3\ a_0$, there really is a (well-fit) plateau in the energy as a function of R for $R \lesssim 1.5\ a_0$. For the $r \leq 2.55\ a_0$ cases plotted in Figure 7, this plateau is more tilted and lies at higher energy; the "wiggles" of up to $\sim 10\ mE_h$ visible in the fitted BMP surface there are probably due partly to the reduced weight given to high-energy points, but are probably also connected to the relatively nearby conical intersection with the first excited state (see § III A 3 below).

Note that in Figure 6 the $\sim 0.3\ mE_h$ "basin" near $\{r = 10.5\ a_0,\ R = 0\}$, and also the $\sim 0.3\ mE_h$ "ridge" in Figures 6c and 6d near $\{r = 12\ a_0,\ R = 2\ a_0\}$, are probably spurious features, where our fitted BMP surface overcompensates slightly due to having to fit the nearby repulsive walls (the spurious features being comparable in size to the expected errors in the *ab initio* energies in this region). Figure 8 illustrates that there is no real evidence for this "basin" and "ridge" in the *ab initio* energies or generated points. These (presumably) spurious features in our BMP fit are smaller even than the rms error with which the surface fits the low-energy points, and furthermore lie in a region of the surface not likely to be sampled, i.e., by the time the H atoms are that far apart in a dissociative collision, the

He atom is not likely to lie half-way between them. Consequently, there seemed to be little point in attempting to smooth out the "ridge" and "basin" by increasing the weight on the points in this region (especially since this might lead to a poorer fit elsewhere).

Figure 8 also shows that the previous He-H_2 surfaces have a larger spurious "bulge" or "basin" than our BMP surface in this region. Through the use of $A_d(r)$ from Equation (13), which has an exponential cutoff at large r, the modMR surface of § II D 2 largely eliminates the spurious "basin," but, like the MR surface itself, is a poor fit to the base of the repulsive "wall" in the $\gamma = 0°$ direction. (In fact, the limit $r \to \infty$ at constant R_A was not constrained at all in the Muchnick & Russek[24] fit, and thus it is not surprising that their resulting MR surface and the similar modMR surface yield a very poor fit to the pair-wise H-He potential.) The WKB and DR surfaces did include a pair-wise H-He potential, yielding a reasonable fit to the base of the repulsive H-He "wall" in the $\gamma = 0°$ direction (although their H-He potentials were less accurate at higher energies). However, their three-body terms did not have a van der Waals well, but rather a relatively long positive "tail" at large separations, resulting in the "bulge" (in the $\gamma = 90°$ direction) shown by these potentials at small R in Figure 8 (where $r = 10.6\ a_0$ and $R_A = R_B \geq 5.3\ a_0$).

In the following subsections, we consider the quality of our BMP fit in various other regions of He-H_2 conformation space.

3. Conical intersection with the first excited state

The conical intersection of the ground state with the first (electronic) excited state forms a curved (1-dimensional) line in the 3-dimensional conformation space of He-H_2. Since we used the MRD-CI program to calculate the first few excited states as well as the ground state, we were able to estimate the position of this conical intersection by considering the energy difference between the ground state and first excited state in the *ab initio* energies. Figure 9 shows this estimated position: the squares, with ground state and first excited state lying at nearly the same energy, should lie closest to the conical intersection. The position of the conical intersection is thus determined to an accuracy of $\sim 0.03\ a_0$ in R_A and $\sim 3°$ in θ, for $1\ a_0 \lesssim r \lesssim 6\ a_0$. Figure 9d shows that, in contrast to the H_3 surface, the conical intersection for He-H_2 does not in general occur at high-symmetry conformations (i.e., conformations having $\gamma = 90°$).

Somewhat better estimates of the position of the conical intersection (as well as estimates of its opening angle in the coordinate directions orthogonal to the line of the conical intersection, as a function of position along that line) could be obtained by plotting the energies of the ground and lowest excited state(s) as a function of position for gridpoints near the estimated conical intersection. Even better accuracy could be obtained by the computation of new *ab initio* energies to home in on the position of the conical intersection (e.g., as was done for H_4 in a limited region by Boothroyd *et al.*[8]). Procedures such as these would be required if one wished to include the conical intersection explicitly in a fitted surface; however, they are beyond the scope of the present work.

As may be seen from Figure 9e, the conical intersection occurs at energies $E \gtrsim 0.2\ E_h$, i.e., above twice the H_2 dissociation energy relative to H_2 + He. Figure 10 shows various cuts through points on the conical intersection (note that Figs. 10a, 10b, and 10c show

three different cuts through the same conical intersection point, while Fig. 10d shows a fourth cut through two different conical intersection points). The BMP surface can be seen to "round off" the conical intersection, missing by as much as 20 mE_h (although it still does much better than the previous surfaces in this region of the He-H_2 surface). Figure 9f, showing errors of the BMP surface relative to the *ab initio* energies at points near the conical intersection, also shows the surface lying quite far below the cusp of the conical intersection (black squares) for $r \gtrsim 1\ a_0$, and (at least in some cases) overcompensating at points slightly further away (black dots). As a result, the conical intersection is the source of many, perhaps most, of the outliers at $E \gtrsim 0.2\ E_h$ in the scatterplot of Figure 5.

4. The H_2 + He van der Waals well

Figure 11 compares fitted He-H_2 surfaces to *ab initio* van der Waals energies — note that the unusual "shifted-logarithmic" energy scale emphasizes effects at the bottom of the van der Waals well, while also showing a significant portion of the "repulsive wall" at smaller R. One can see that our BMP He-H_2 surface fits the depth of the van der Waals well to within about 1 μE_h, when compared to the accurate *ab initio* energies of Tao[31] (which have $1.28\ a_0 \leq r \leq 1.618\ a_0$); from subset 31 in Table II, the rms difference is 0.8 μE_h over the range $6\ a_0 \leq R < 10\ a_0$. The previous SK and MR surfaces[22,24] agree nearly as well (i.e., within a few μE_h; rms values of 1.8 and 2.6 μE_h, respectively), since they were fitted to the fairly accurate *ab initio* energies of Meyer *et al.*[29] However, the earlier WKB and DR surfaces[17,20] do not contain a van der Waals well at all — they are visible only at the upper right of Figure 11.

Note that the "torque reversal" at $R \approx 6.1\ a_0$ noted by Muchnick & Russek[24] is also visible in Figure 11. For small angles γ, the van der Waals well is slightly deeper but the "repulsive wall" is slightly further out than for γ near 90°. Thus the $\gamma = 0°$ (linear) orientation is energetically favored at $R > 6.1\ a_0$, while the $\gamma = 90°$ (T-shaped) orientation is favored at smaller distances R.

Figure 12 considers the variation of the van der Waals well with the size r of the H_2 molecule (using the same energy scale as Fig. 11). For $r = 1.28$, 1.449, and 1.618 a_0, accurate *ab initio* energies from Tao[31] are available, which our BMP surface fits very well (rms of 0.8 μE_h: see subset 31 in Table II). For the more extreme distances $r = 0.9$, 1.1, 1.8, and 2.0 a_0, only the much-less-accurate "gen-SK vdw" and "gen-MR vdW" generated points of § II A 4 were available. The former were obtained from the SK surface[22], which at those r values was fitted to the lower-accuracy extreme-position *ab initio* energies of Meyer *et al.*[29]; the latter were obtained from the modMR surface of § II D 2, fitted to the accurate near-H_2-equilibrium *ab initio* energies from Tao[31] with the r-dependence taken from Thakkar *et al.*[23] These two sets of generated points agree with each other to within about 5 μE_h in the van der Waals well, and are fitted to roughly this accuracy by our BMP surface (rms of ~ 3 μE_h over the range $6\ a_0 \leq R < 10\ a_0$, from subsets 30 and 32 in Table II; largest difference ~ 10 μE_h). As the size r of the H_2 molecule is increased, the van der Waals well moves slightly further out; the depth also changes slightly, but this is less well constrained by the *ab initio* energies.

At $r > 2\ a_0$, there are no accurate *ab initio* energies available in the region of the van der

Waals well (some of our *ab initio* energies may lie nearby, but these have estimated errors larger than the typical depth of the van der Waals well). The generated "vdW" energies in this region are thus expected to be quite inaccurate — the SK and modMR surfaces differ from each other by $\sim 0.1\ mE_h$ in this region. These generated points were therefore given quite low weight, only enough to keep the fitted surface from developing large spurious features in this region. This is the cause of the relatively large rms error of $0.1233\ mE_h$ of the BMP surface shown by Table II (subset 33) for the "vdW" at $r = 2.4, 3.0, 4.0\ a_0$ and $6\ a_0 \le R < 10\ a_0$.

5. The interaction region

As mentioned above, Table II shows that the earlier SK and MR surfaces[22,24] do relatively well for not-too-large H_2-molecules ($r < 2\ a_0$), even up to relatively large interaction energies. An example of this, with $r = 1.4\ a_0$, is illustrated in Figure 13a (note the logarithmic energy scale). Even at $R \sim 1.3$ to $2\ a_0$, these two surfaces have errors of only a few percent for near-equilibrium H_2 — the BMP surface of the present work is an order of magnitude more accurate, with errors of a fraction of a percent in this region, but nonetheless all three of these surfaces are almost indistinguishable in Figure 13a. Even the much-earlier WKB and DR surfaces[17,20] are not so very much worse in the outer part of this region, but, as may be seen in Figure 13a, they have an unphysical "hole" at $R \lesssim 1.5\ a_0$ for small angles γ (i.e., the "hole" is at small $H - He$ separations R_A).

Table II shows that, for all the fitted surfaces considered there, errors at small r (i.e., $r < 1.2\ a_0$: subsets 4 to 7) are very similar to the errors at $r \sim 1.4\ a_0$ (subsets 8 to 11) that are illustrated in Figure 13a. In contrast, at larger $H - H$ separations r (such as would arise from dissociative collisions and highly-excited H_2 molecules), the SK surface[22] is undefined, and even the MR surface[24] does quite poorly, with errors of order 50% visible in Figure 13b, c, and d (for $r = 4\ a_0$). At these larger r values, the WKB and DR surfaces[17,20] have errors of order a factor of 2, and Figure 13b shows that the "hole" in these surfaces appears at lower interaction energy (and, in fact, at a lower total energy, though this is less obvious from the figure). A systematic numerical search found that the lowest part of the lip of this "hole" lies at energies $E = 53$ and $87\ mE_h$ for the WKB and DR surfaces, respectively — less than two H_2 dissociation energies above equilibrium $H_2 + He$. In this region, the BMP surface of the present work is much more accurate than any previous surface, with typical errors only slightly larger than for equilibrium H_2 molecules. This is also illustrated by the other cuts through this region shown in the previous Figure 10.

6. Extrapolation to very short distances

The SK surface[22] contains no repulsive Coulomb terms, and thus does not have the correct form at small $H - He$ separations — an example can be seen in Figure 13a. The WKB and DR surfaces[17,20] do even worse: they turn over and become negative at small $H - He$ separations (R_A or $R_B \lesssim 1\ a_0$: the "hole" in these surfaces), as may be seen in Figure 13a and b. The MR surface[24] behaves roughly as an exponentially damped Coulomb repulsion for small R_A, yielding reasonable behavior at short distances.

No constraints were placed on our fitted BMP surface at H − He distances less than those of the most compact *ab initio* geometries (namely $R_A = 0.6\ a_0$), which lie at energies $E \gtrsim 1\ E_h$. Some of our earlier fitted surfaces turned over into an unphysical "hole" at shorter distances than this, but our final BMP He-H_2 surface becomes strongly repulsive there instead. An example of this behavior may be seen in Figure 13a and b: above $\sim 1\ E_h$ the BMP curve (solid line) rises very steeply. Extrapolating from the *ab initio* energies, the short-distance behavior at H − He distances $R_A < 0.6\ a_0$ ought to be intermediate between the BMP curve and that of the MR surface[24] (long-dashed line), but closer to the latter for $R_A \lesssim 0.55\ a_0$. For cases where an extremely hard repulsive core at $R_A < 0.6\ a_0$ is undesirable, the Fortran program of our BMP analytic surface[32,33] contains an option allowing the user to switch over at small R_A or R_B (gradually, with continuous first derivatives) to alternate forms, which behave more like an exponentially damped Coulomb repulsion there.

All of the analytic surfaces include the H_2-molecule potential V_{H_2} as a separate term in the surface formulae. The WKB and MR papers[17,24] left the choice of V_{H_2} formula unspecified, while the Fortran program for the DR surface[20] used the spline fit of Kolos & Wolniewicz[48] for V_{H_2} (which is almost as good as Schwenke's[28] V_{H_2} formula). For consistency, we used the accurate H_2-molecule potential of Schwenke[28] for *all* the surfaces. This formula behaves well at most H − H separations r, but begins to behave poorly at $r \lesssim 0.1\ a_0$. Our Fortran program therefore switches over at $r \approx 0.12\ a_0$ to a polynomial in $1/r$, which in turn becomes a pure $1/r$ repulsive potential for $r \lesssim 0.045\ a_0$. (The first derivative with respect to r is continuous at these switchover points $r \approx 0.12\ a_0$ and $r \approx 0.045\ a_0$.)

B. Prospects for further improvement

Significant improvements over the BMP surface of this paper are possible, but would require a major effort.

Improvements in the van der Waals well for strongly non-equilibrium H_2 molecules ($r < 1.2\ a_0$ and $r > 1.6\ a_0$) would require very accurate *ab initio* calculations in this region (such as those performed for equilibrium H_2 molecule sizes by Tao[31]). Once these *ab initio* energies were available, increased flexibility (i.e., larger index ranges) in the V_d and V_D terms should allow the van der Waals well to be well represented over a wide range of H_2 molecule sizes.

Improvements in the interaction region would require, in the analytic functional form, explicit inclusion of the conical intersection with the first excited state. To do this, one would first have to pin down its location and its opening angle in the perpendicular directions (it is possible that these could be determined sufficiently well from the existing *ab initio* energies of the present work; the alternative would be to use an explicit search via new *ab initio* computations, as was done for H_4 in a limited region by Boothroyd *et al.*[8]). The position and opening angle would then have to be fitted to an analytic form, which could then be added to the fitted surface — the magnitude of this new term (and perhaps its extent in the directions perpendicular to the locus of the conical intersection) would need to be parameterized and fitted as a function of position along the locus of the conical intersection. The other terms in the fitted surface would then not have to try to fit the sharp peak very near the conical intersection, allowing a considerable improvement in accuracy there.

IV. CONCLUSIONS

A new set of 23 703 *ab initio* energies was computed for He-H_2 geometries where the interaction energy is expected to be non-negligible, using Buenker's multiple reference (single and) double excitation configuration interaction (MRD-CI) program[40,39]; the lowest excited states were computed as well as the ground state energy. These new *ab initio* energies have an estimated rms "random" error of $\sim 0.2\ mE_h$ and a systematic error of $\sim 0.6\ mE_h$ (0.4 kcal/mol). The position (in the 3-dimensional conformation space of He-H_2) of the conical intersection between the ground state and the first excited state has been roughly mapped out; unlike H_3, this conical intersection for He-H_2 does *not* lie at high-symmetry conformations, but rather along a curved line in conformation space.

These new *ab initio* energies were used to test previous analytic He-H_2 surfaces. Even the best of the previous surfaces, that of Muchnick & Russek[24], does quite poorly for very large H_2-molecule sizes (such as would be encountered in dissociative collisions or highly excited H_2 molecules), although both this surface and that of Schaefer & Köhler[22] are fairly accurate for not-too-large H_2 molecules ($r \lesssim 2\ a_0$) with a not-too-close He atom ($R \gtrsim 3\ a_0$).

A new analytic He-H_2 surface (the BMP surface) was fitted to 20 203 of the new *ab initio* energies (and to an additional 4862 points generated at large separations) — the other 3500 new *ab initio* energies were used only to test "interstitial" conformations. This BMP surface fits the interaction region at the "chemical accuracy" level ($\sim 1\ mE_h$) required for reaction dynamics; the overall energy-weighted rms error of $1.42\ mE_h$ (0.89 kcal/mole) is comparable to the accuracy of the *ab initio* energies (note that this energy-weighting specifies lower weight for high-energy points, i.e., $E > 0.2\ E_h$, but a weight of unity everywhere else: see § II D 1). For the 14 585 *ab initio* energies that lie below twice the H_2 dissociation energy, the new BMP He-H_2 surface has an rms error of $0.95\ mE_h$ (0.60 kcal/mole). This surface is an order of magnitude better than previous surfaces in the interaction region, and also yields a slight improvement in the fit to the recent van der Waals energies of Tao[31].

For relatively compact conformations (i.e., with energies above twice the H_2 dissociation energy), the conical intersection between the ground state and the first excited state is the largest source of error in the analytic surface. The BMP surface "rounds off" the conical intersection, yielding errors of up to $\sim 20\ mE_h$ there. The position of this conical intersection forms a curved 1-dimensional locus in the 3-dimensional conformation space of He-H_2; its approximate position has been mapped out, but trying to include the conical intersection explicitly in an analytic surface would require a more accurate map of its position, and is beyond the scope of the present paper.

The *ab initio* energies and a Fortran program for the analytic BMP He-H_2 surface of the present work are available from EPAPS[32] or from the authors[33].

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

[1] D. G. Truhlar and C. J. Horowitz, J. Chem. Phys. **68**, 2466 (1978); J. Chem. Phys. **79**, 1514 (1979: Errata).

[2] A. J. C. Varandas, F. B. Brown, C. A. Mead, D. G. Truhlar, and N. C. Blais, J. Chem. Phys. **86**, 6258 (1987).

[3] A. I. Boothroyd, W. J. Keogh, P. G. Martin, and M. R. Peterson, J. Chem. Phys. **95**, 4343 (1991).

[4] A. I. Boothroyd, W. J. Keogh, P. G. Martin, and M. R. Peterson, J. Chem. Phys. **104**, 7139 (1996).

[5] Y.-S. Wu, A. Kuppermann, and J. B. Anderson, Phys. Chem. Chem. Phys., **1**, 929 (1999).

[6] W. J. Keogh, Ph. D. Thesis, University of Toronto (1992).

[7] A. Aguado, C. Suarez, and M. Paniagua, J. Chem. Phys. **101**, 4004 (1994).

[8] A. I. Boothroyd, P. G. Martin, W. J. Keogh, and M. J. Peterson, J. Chem. Phys. **116**, 666 (2002).

[9] A. I. Boothroyd, P. G. Martin, W. J. Keogh, and M. J. Peterson, EPAPS Document No. E-JCPSA6-115-304140 (2002).

[10] C. S. Roberts, J. Chem. Phys. **131**, 203 (1963).

[11] M. Karplus and H. J. Kolker, J. Chem. Phys. **41**, 3955 (1964).

[12] M. Kraus and F. H. Mies, J. Chem. Phys. **42**, 2703 (1965).

[13] M. D. Gordon and D. Secrest, J. Chem. Phys. **52**, 120 (1970).

[14] R. Gengenbach and Ch. Hanh, Chem. Phys. Lett. **15**, 604 (1972).

[15] B. Tsapline and W. Kutzelnigg, Chem. Phys. Lett. **23**, 173 (1973).

[16] P. J. M. Guerts, P. E. S. Woermer, and A. van der Avoird, Chem. Phys. Lett. **35**, 444 (1975).

[17] C. W. Wilson Jr., R. Kapral, and G. Burns, Chem. Phys. Lett. **24**, 488 (1974).

[18] K. T. Tang and J. P. Toennies, J. Chem. Phys. **66**, 1496 (1977).

[19] K. T. Tang and J. P. Toennies, J. Chem. Phys. **68**, 5501 (1978).

[20] J. E. Dove and S. Raynor, Chem. Phys. **28**, 113 (1978).

[21] A. Russek and R. Garcia, Phys. Rev. A **26**, 1924 (1982).

[22] J. Schaefer and W. E. Köhler, Physica **129A**, 469 (1985).

[23] A. J. Thakkar, Z.-M. Hu, C. E. Chauqui, J. S. Carley, and R. J. LeRoy, Theor. Chim. Acta **82**, 57 (1992).

[24] P. Muchnick and A. Russek, J. Chem. Phys. **100**, 4336 (1994).

[25] J. E. Dove, A. C. M. Rusk, P. H. Cribb, and P. G. Martin, Astrophys. J. **318**, 379 (1987).

[26] C. A. Chang and P. G. Martin, Astrophys. J. **378**, 202 (1991).

[27] P. G. Martin, D. H. Schwarz, and M. E. Mandy, Astrophys. J. **461**, 265 (1996).

[28] D. W. Schwenke, J. Chem. Phys. **89**, 2076 (1988).

[29] W. Meyer, P. C. Hariharan, and W. Kutzelnigg, J. Chem. Phys. **73**, 1880 (1980).

[30] U. E. Senff and P. G. Burton, J. Phys. Chem. **89**, 797 (1985)

[31] F.-M. Tao, J. Chem. Phys. **100**, 4947 (1994).

[32] See EPAPS Document No. _____ for Fortran programs to compute the He-H_2 surface (including first derivatives), and files of *ab initio* He-H_2 energies. This document may be retrieved 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.
[33] A Fortran program of the adopted surface, or files of *ab initio* energies, can also be obtained by contacting P. G. Martin, email: pgmartin@cita.utoronto.ca and/or visiting the website http://www.cita.utoronto.ca/~pgmartin/heh2pes/.
[34] C. Møller & M. S. Plesset, Phys. Rev. **46**, 618 (1934).
[35] J. A. Pople, P. Krishnan, H. B. Schegel, & J. S. Binkley, Int. J. Quantum Chem. Symp. **13**, 325 (1979).
[36] R, Gengenbach, Ch. Hahn, and J. P. Toennies, Phys. Rev. A **7**, 98 (1973).
[37] A. I. Boothroyd, W. J. Keogh, P. G. Martin, and M. J. Peterson, J. Chem. Phys. **95**, 4331 (1991).
[38] A. I. Boothroyd, W. J. Keogh, P. G. Martin, and M. J. Peterson, AIP document no. PAPSJCPSA-95-4331-179 (1991).
[39] R. J. Buenker and P. Funke, private communication (1992).
[40] R. J. Buenker, private communication (1987).
[41] S. Huzinga, J. Chem. Phys. **42**, 1293 (1964).
[42] M. J. Frisch, J. A. Pople, and J. S. Brinkley, J. Chem. Phys. **80**, 3265 (1984).
[43] G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris, J. Chem. Phys. **89**, 2193 (1988).
[44] P. G. Burton, S. D. Peyerimhoff, and R. J. Buenker, Chem. Phys. **73**, 83 (1982).
[45] P. G. Burton, Int. J. Quantum Chem. **23**, 613 (1983).
[46] P. G. Burton, R. J. Buenker, P. J. Bruna, and S. D. Peyerimhoff, Chem. Phys. Lett. **95**, 379 (1983).
[47] I. Shavitt, F. B. Brown, and P. G. Burton, Int. J. Quantum Chem. **31**, 507 (1987).
[48] W. Kolos and L. Wolniewicz, J. Chem. Phys. **43**, 2429 (1965).

TABLES

TABLE I. Gaussian basis sets

Basis set	Type	Exponent ($a_0{}^{-2}$)	Coefficient
H: $(9s3p2d)/[4s3p2d]$	s	887.22	0.000112
		123.524	0.000895
		27.7042	0.004737
		7.82599	0.019518
		2.56504	0.065862
		0.938258	0.178008
	s	0.372145	1.0
	s	0.155838	1.0
	s	0.066180	1.0
	p	2.1175	1.0
	p	0.77	1.0
	p	0.28	1.0
	d	1.76	1.0
	d	0.62	1.0
He: $(10s4p2d)/[5s4p2d]$	s	3293.694	0.00010
		488.8941	0.00076
		108.7723	0.00412
		30.17990	0.01721
		9.789053	0.05709
		3.522261	0.14909
	s	1.352436	1.0
	s	0.552610	1.0
	s	0.240920	1.0
	s	0.107951	1.0
	p	5.823125	1.0
	p	2.1175	1.0
	p	0.77	1.0
	p	0.28	1.0
	d	2.83871	1.0
	d	1.0	1.0

TABLE II. Energy-weighted[a] rms errors of analytic He-H_2 surfaces, in mE_h

Surface I.D.:[b]		WKB	DR	SK	MR	modMR	BMP
No. of parameters:		4	4	[750][c]	19	27	112
Subset	N_{pts}	rms	rms	rms	rms	rms	rms
1: All fitted *ab initio*	20 203	*100.90*	*100.90*	*5.1E+5*	*28.60*	*28.35*	1.417
2: ... $E < 0.174\ E_h$[d]	14 585	*37.80*	*40.40*	*6.0E+5*	*20.24*	*19.89*	0.955
3: ... $E < 0.0\ E_h$[e]	7177	*11.62*	*10.58*	*1.1E+4*	*4.48*	*4.44*	0.481
4: ... $0.5\ a_0 \le r < 1.2\ a_0$	3744	26.75	20.95	*21.35*	*7.80*	*7.80*	0.875
5: ... ... $R \le 2\ a_0$	1646	37.73	29.57	*32.18*	*10.87*	*10.87*	1.205
6: ... ... $2\ a_0 < R \le 3\ a_0$	857	17.89	14.18	1.323	*6.20*	*6.21*	0.624
7: ... ... $R > 3\ a_0$	1241	7.11	4.97	0.449	*0.418*	*0.421*	0.336
8: ... $1.2\ a_0 \le r \le 1.62\ a_0$	2311	48.62	47.82	*22.14*	15.79	15.79	1.003
9: ... ... $R \le 2\ a_0$	907	75.84	74.80	*35.31*	24.69	24.68	1.463
10: ... ... $2\ a_0 < R \le 3\ a_0$	575	18.72	18.04	1.497	6.39	6.40	0.690
11: ... ... $R > 3\ a_0$	829	7.39	5.19	0.433	0.445	0.450	0.365
12: ... $1.62\ a_0 < r \le 2\ a_0$	1918	74.01	76.16	*116.17*	*25.34*	*25.33*	1.351
13: ... ... $R \le 2\ a_0$	735	117.85	121.07	*187.64*	*39.59*	*39.58*	2.068
14: ... ... $2\ a_0 < R \le 3\ a_0$	518	22.31	25.29	3.42	*12.36*	*12.34*	0.723
15: ... ... $R > 3\ a_0$	665	7.71	5.51	0.458	*0.865*	*0.889*	0.358
16: ... $2\ a_0 < r \le 4\ a_0$	7136	130.27	133.13	*1.1E+4*	*37.63*	*37.64*	1.797
17: ... ... $R > 3\ a_0$	2806	10.16	8.30	*24.75*	*7.64*	*7.44*	0.456
18: ... $4\ a_0 < r \le 10.6\ a_0$	5094	*113.78*	*109.56*	*1.0E+6*	*29.32*	*28.34*	1.318
19: ... ... $R > 4\ a_0$	1244	*8.44*	*6.40*	*84.99*	*18.47*	*17.87*	0.588
20: Fitted "random" *ab initio*[f]	3500	*57.82*	*64.25*	*2.9E+5*	*27.36*	*26.95*	1.411
21: ... $E < 0.174\ E_h$[d]	2962	*36.35*	*39.61*	*3.1E+5*	*20.81*	*20.32*	0.994
22: ... $E < 0.0\ E_h$[e]	1367	*13.92*	*13.03*	*5.77*	*5.07*	*5.04*	0.527
23: Unfitted "random" *ab initio*[g]	3500	*57.07*	*63.35*	*2.6E+5*	*27.74*	*27.31*	1.400
24: ... $E < 0.174\ E_h$[d]	2943	*35.58*	*38.41*	*2.8E+5*	*21.05*	*20.52*	0.964
25: ... $E < 0.0\ E_h$[e]	1349	*14.10*	*13.12*	*14.50*	*4.93*	*4.89*	0.538
26: All "vdW"	2145	*2.037*	*1.698*	*0.134*	0.609	0.609	0.129
27: ... $R < 6\ a_0$	587	*3.67*	*3.03*	*0.251*	1.164	1.164	0.237
28: ... $6\ a_0 \le R < 10\ a_0$	1098	*0.930*	*0.830*	*0.0379*	[0.0114]	[0.0013]	0.0507
29: ... ... $r = 0.6, 0.7, 0.8\ a_0$[h]	186	*0.491*	*0.404*	*0.0023*	[0.0020]	[0.0000]	0.0031
30: ... ... $r = 0.9, 1.1\ a_0$[i]	248	*0.731*	*0.640*	[0.0011]	[0.0014]	[0.0011]	0.0017
31: ... ... $r = 1.28, 1.449, 1.618\ a_0$[j]	231	*1.004*	*0.906*	0.0018	0.0026	0.0011	0.0008
32: ... ... $r = 1.8, 2.0\ a_0$[i]	248	*1.139*	*1.038*	[0.0024]	[0.0042]	[0.0024]	0.0037
33: ... ... $r = 2.4, 3.0, 4.0\ a_0$[h]	185	*1.084*	*0.952*	*0.0922*	[0.0271]	[0.0000]	0.1233
34: ... $R \ge 10\ a_0$	460	*0.273*	*0.270*	0.0004	0.0005	0.0001	0.0001
35: All "H-He"	1887	*114.69*	*106.25*	*4.9E+4*	*44.54*	*44.00*	1.122
36: ... $R_A \le 2\ a_0$	1008	*156.90*	*145.35*	*1.2E+3*	*59.88*	*59.18*	1.455
37: ... $2\ a_0 < R_A < 4\ a_0$	468	3.65	2.75	*4.5E+3*	*16.59*	*16.19*	0.714

38: ... 4 $a_0 < R_A < 6\ a_0$	228	0.416	0.0959	*3.8E+3*	*0.680*	*0.639*	0.0978
39: ... $R_A \geq 6\ a_0$	183	0.0431	0.0008	*1.6E+5*	*0.0068*	0.0038	0.0014
40: All "H-He + H"	830	*26.03*	*22.19*	*7.6E+6*	*19.82*	*19.34*	0.874
41: ... $R_A \leq 3\ a_0$	141	*63.15*	*53.84*	*2.4E+4*	*48.03*	*46.88*	2.096
42: ... 3 $a_0 < R_A < 6\ a_0$	201	0.950	0.602	*8.9E+5*	*2.02*	*1.63*	0.266
43: ... $R_A \geq 6\ a_0$	488	*0.184*	*0.156*	*9.9E+6*	*0.0813*	0.0051	0.0341
44: All fitted points; full-weight[a]	25 065	*329.02*	*310.75*	*5.2E+7*	*28.70*	*28.41*	1.383

[a] For energy-weighted rms errors, deviations get only weight w_E from equation (22) (unity for $E \leq 0.2\ E_h$, reduced weight at higher energy), as opposed to full-weight rms errors in last line of table (subset 44). Note that the rms value for a surface is *italicized* for subsets containing conformations in regions where that surface was not fitted.

[b] WKB: Wilson, Kapral, & Burns (Ref. 17), DR: Dove & Raynor (Ref. 20), SK: Schaefer & Köhler (Ref. 22), MR: Muchnick & Russek (Ref. 24), modMR: slightly modified Muchnick & Russek surface of § II D 2, BMP: adopted surface of the present work.

[c] Schaefer & Köhler present 3 Legendre coefficients at 48 R-values for each of 5 r-values (namely, $r = 0.9$, 1.28, 1.449, 1.618, and 2.0 a_0), plus formulae for very large R.

[d] $E < 0.174\ E_h$ corresponds to points lying below about twice the H_2 dissociation energy, relative to a H_2 molecule plus a distant He atom.

[e] $E < 0.0\ E_h$ corresponds to points lying below the H_2 dissociation energy, relative to a H_2 molecule plus a distant He atom.

[f] These subsets 20 to 22 are also contained in subsets 1 to 3, respectively.

[g] These subsets 23 to 25 were obtained using the same probability distribution as subsets 20 to 22, but were *not* included in the actual fitting process for the BMP surface of the present work.

[h] The accuracy of these generated points is unknown. They were intended only to yield "reasonable" values and prevent wild excursions in the fitted surfaces, and were given relatively low weight in the fit, resulting in the relatively large "BMP" rms for these points at large r. The "modMR" surface was used to generate them, and thus by definition has zero rms here (likewise, the very similar "MR" surface has a small rms here); this is indicated by the square brackets enclosing their rms values.

[i] The "SK" and "modMR" surfaces were used to generate these points, and thus they (and the "MR" surface) have small rms values here (again indicated by square brackets).

[j] These points comprise the *ab initio* energies of Tao (Ref. 31) and points generated from them, and thus these rms values provide a measure of the accuracy for *all* the surfaces in the van der Waals well.

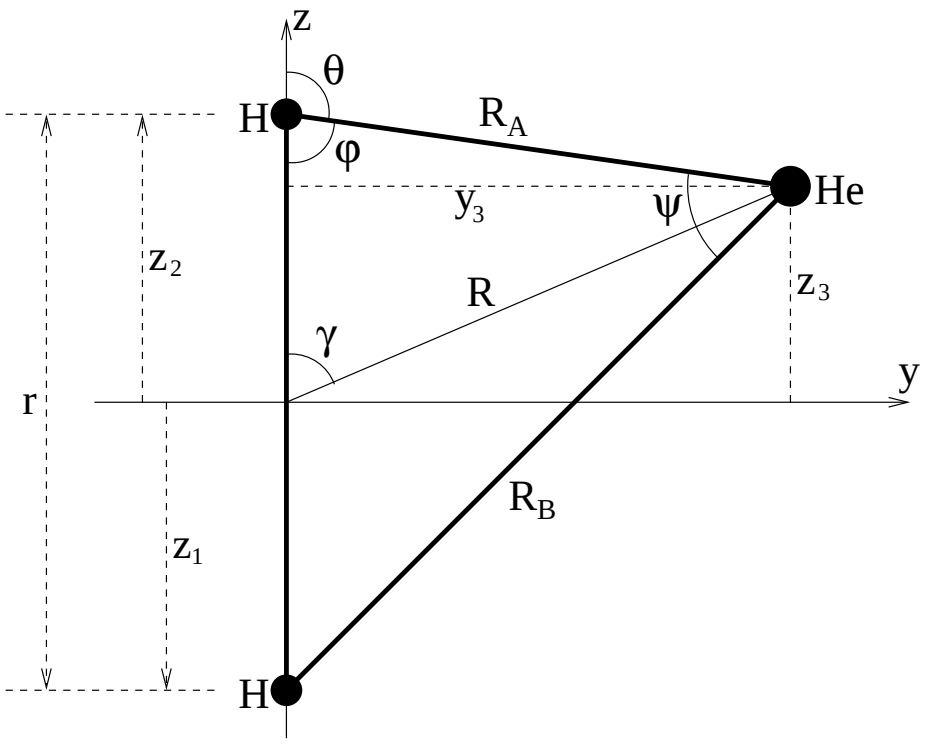

FIG. 1. Notation used in this paper for interatomic distances and angles of He-H_2. By definition, $|z_1| = z_2 = r/2$, with atom-3 being the helium atom.

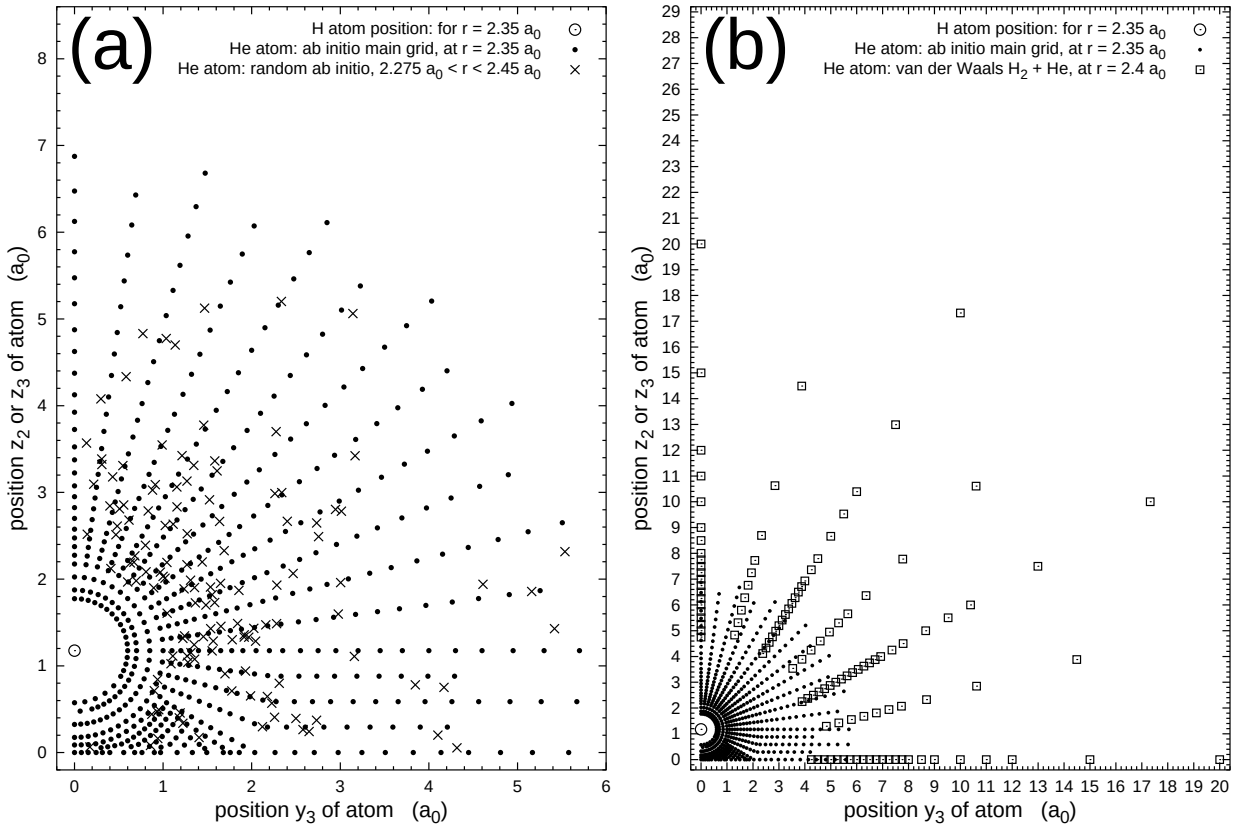

FIG. 2. The grid of fitted points for a given H_2-molecule size r, showing positions of the He atom relative to a molecule lying on the z-axis and centered at the origin. (a) The main grid of *ab initio* points (solid dots), for $r = 2.35\ a_0$ (the H-atom position is shown by the open circle); crosses show positions of "nearby" random *ab initio* points (in the same "r-bin", namely, with $2.275\ a_0 < r < 2.45\ a_0$). (b) The grid for the van der Waals $H_2 + He$ points (open squares), for $r = 2.4\ a_0$; the main grid of part (a) is also visible (solid dots).

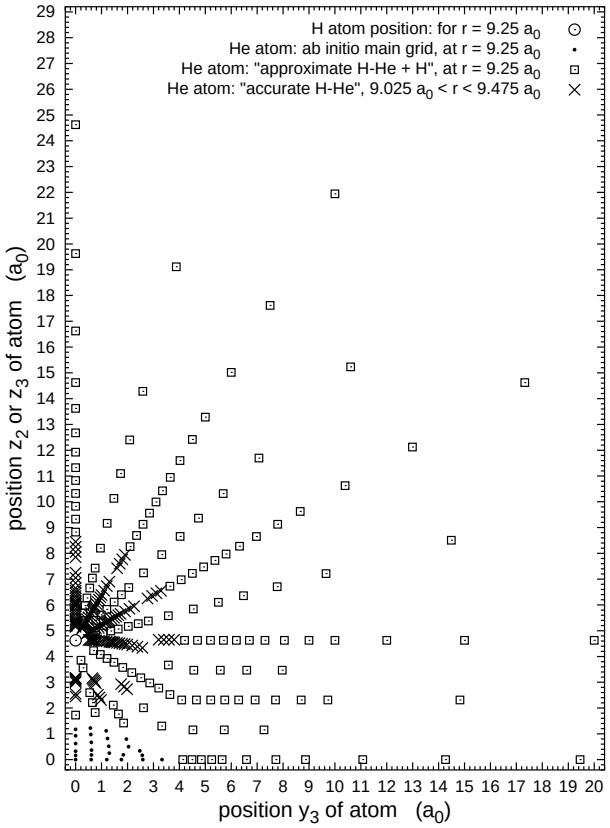

FIG. 3. The grid of fitted points for a large H_2 separation $r = 9.25\ a_0$; unlike Fig. 2, the *ab initio* points (solid dots) all lie relatively near the origin, between the H atoms. The "approximate H-He + H" grid (open squares) at $r = 9.25\ a_0$ was designed to avoid other grids, including any "nearby" points on the "accurate H-He" grid (crosses — with $9.025\ a_0 < r < 9.475\ a_0$).

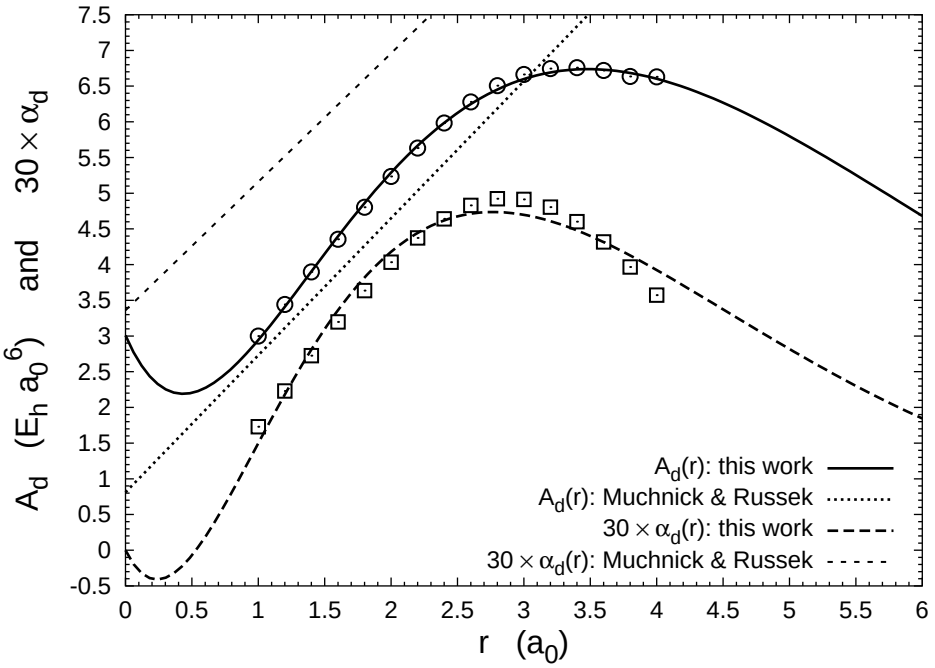

FIG. 4. Our fits $A_d(r)$ and $\alpha_d(r)$ (solid and dashed lines, respectively) to the asymptotic parameters A_d and α_d of Thakkar *et al.* (Ref. 23: circles and squares), as a function of the H_2 molecule size r (note that actual α_d values have been multiplied by a factor of 30 so as to be visible on the same plot as A_d). Values extrapolated to $r < 0.8\ a_0$ are of little practical relevance, as H_2 is unbound at the corresponding energies. The linear forms of $A_d(r)$ and $\alpha_d(r)$ (dotted and wide-dashed lines, respectively) that were used by Muchnick & Russek (Ref. 24) are also shown, although their $\alpha_d(r)$ is not strictly comparable, since they used it in a term with η^2 rather than $P_2(\eta)$.

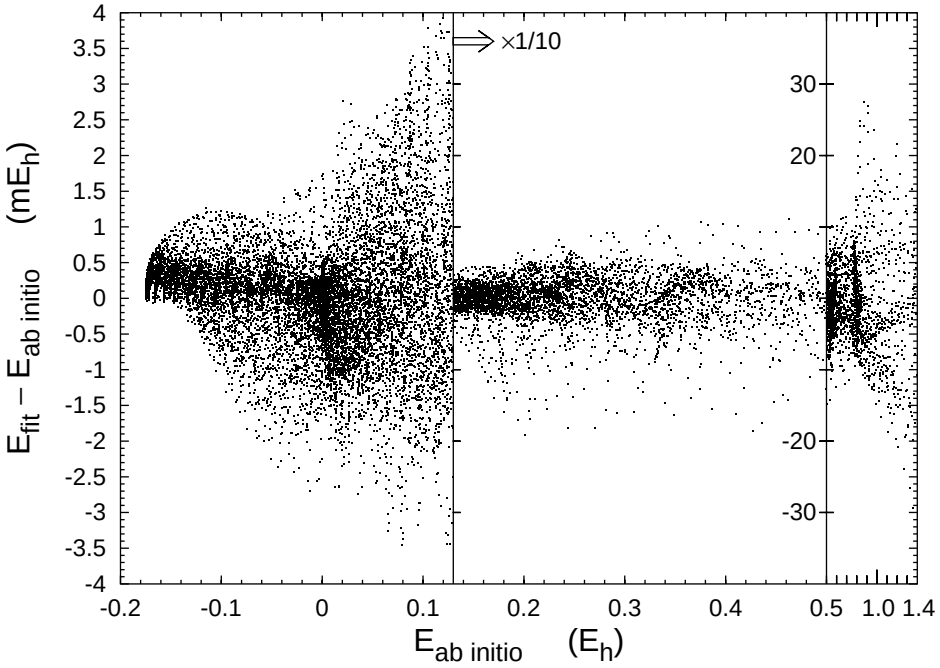

FIG. 5. Scatterplot of the deviations of the adopted "BMP" He-H_2 PES relative to the 20 203 *ab initio* energies, as a function of energy E. Note that for $E > 0.13\ E_h$ the vertical scale is compressed by a factor of 10; for $E > 0.5\ E_h$, the horizontal scale is compressed likewise. (There are 13 397, 4592, and 1991 points in these three parts of the plot, respectively; 223 points lie offscale to the right.) Denser bands of points (usually nearly vertical) result from the discrete r values of the main *ab initio* grid. Note that most of the extreme outliers for $E \gtrsim 0.2\ E_h$ arise from the conical intersection of the ground state with the first excited state.

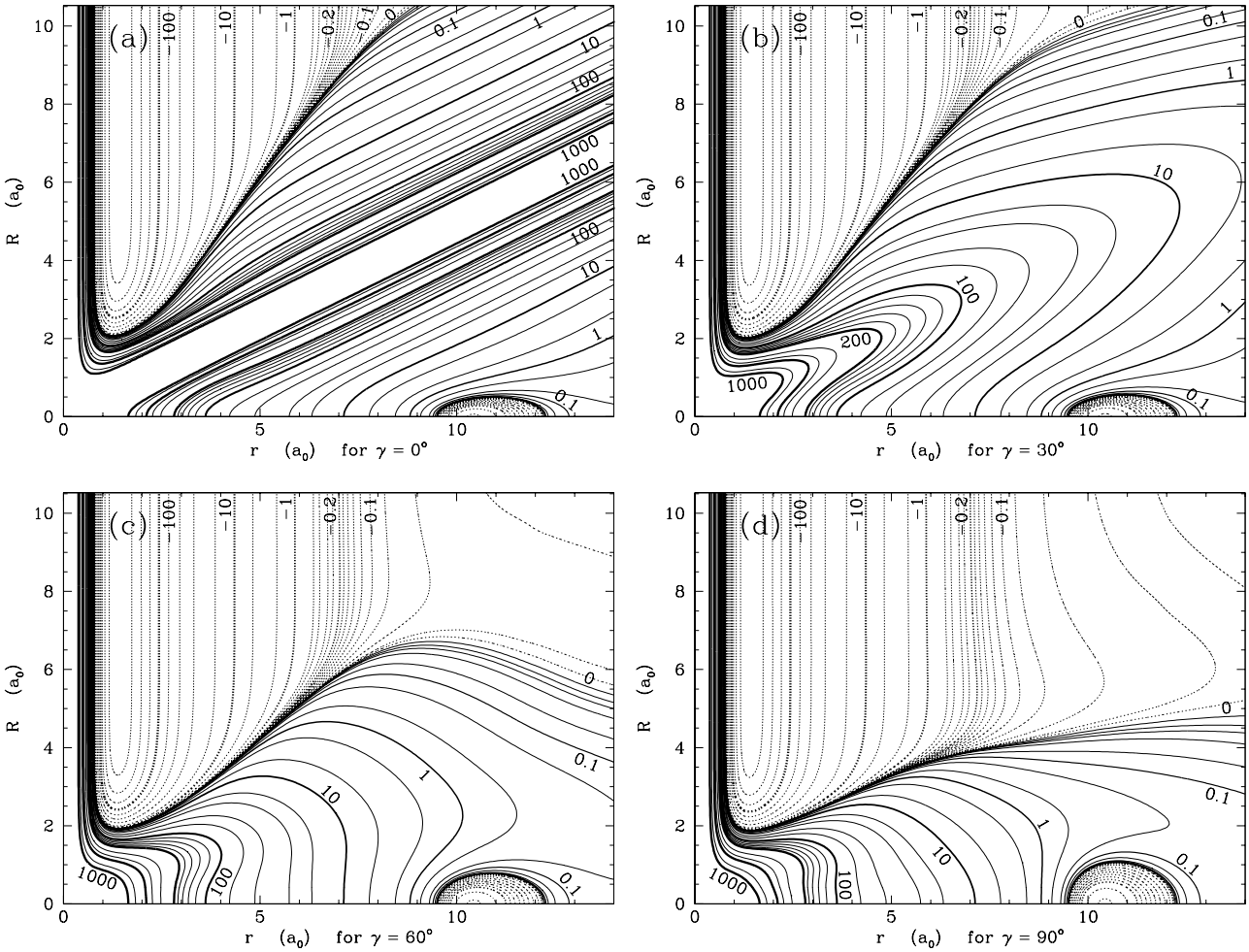

FIG. 6. Contour plots of the fitted "BMP" He-H_2 surface of this paper as a function of $H-H$ separation r and H_2-He separation R, at four angles γ (see Fig. 1). Positive contour levels (solid lines) lie at energies $E = \{$*0*, 0.005, 0.01, 0.02, 0.05, *0.1*, 0.2, 0.5, *1*, 2, 5, *10*, 20, 30, 50, 70, *100*, 120, 140, 160, 180, *200*, 300, 400, *500*, 750, and *1000* $mE_h\}$ relative to separated H+H+He (numbers in *italics* correspond to heavy contour lines — most of these are also labelled in the plots). Negative contour levels (dotted lines) lie at energies $E = \{-0.005, -0.01, -0.025, -0.05, -0.075, -0.1, -0.125, -0.15, -0.175, -0.2, -0.25, -0.3, -0.4, -0.5, -1, -2, -5, -10, -20, -40, -60, -80, -100, -120, -140,$ and $-160\ mE_h\}$. (a) $\gamma = 0°$, (b) $\gamma = 30°$, (c) $\gamma = 60°$, (d) $\gamma = 90°$. Note that the small ($E \sim -0.3\ mE_h$) "basin" near ($r = 10.5\ a_0, R = 0$) is spurious: see text.

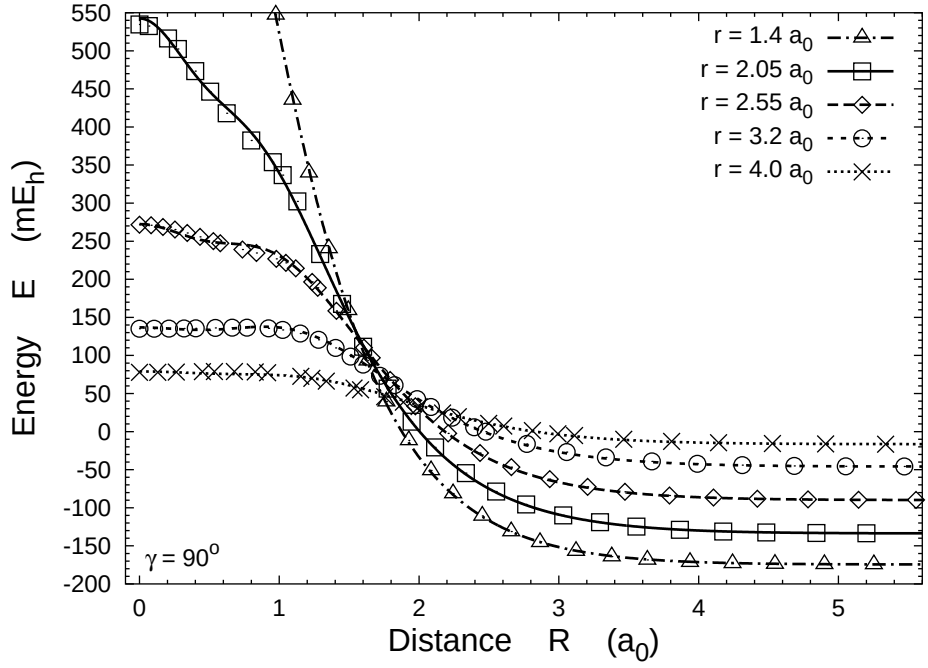

FIG. 7. Comparison of the BMP fit (lines) with *ab initio* energies (symbols) as a function of the $H_2 - He$ distance R at $\gamma = 90^\circ$ (T-shaped orientation), for five $H - H$ distances $r = 1.4$, 2.05, 2.55, 3.2, and 4.0 a_0.

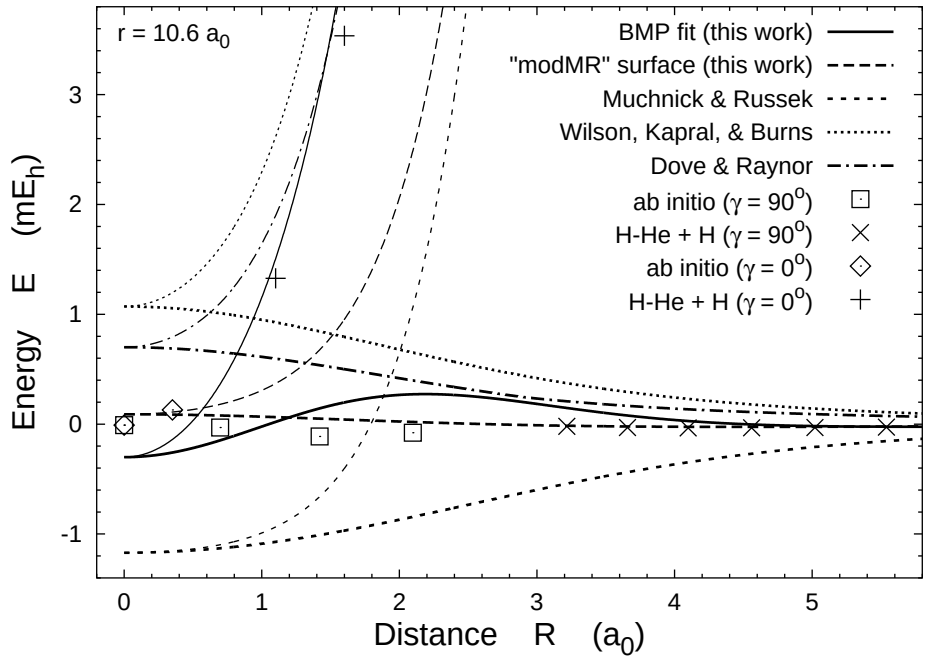

FIG. 8. Potential for the He atom "between" widely-separated H atoms ($r = 10.6\ a_0$), for displacements R of the He atom at two angles γ (from the center point). For $\gamma = 90°$ (T-shaped: heavy lines), as R increases, the He atom is moving out from between the H atoms. For $\gamma = 0°$ (linear: light lines), as R increases, the He atom approaches one of the H atoms and encounters the base of the repulsive "wall." Note that the "modMR" surface is identical to the Muchnick & Russek (Ref. 24) surface, except for the long-range (dispersion) terms.

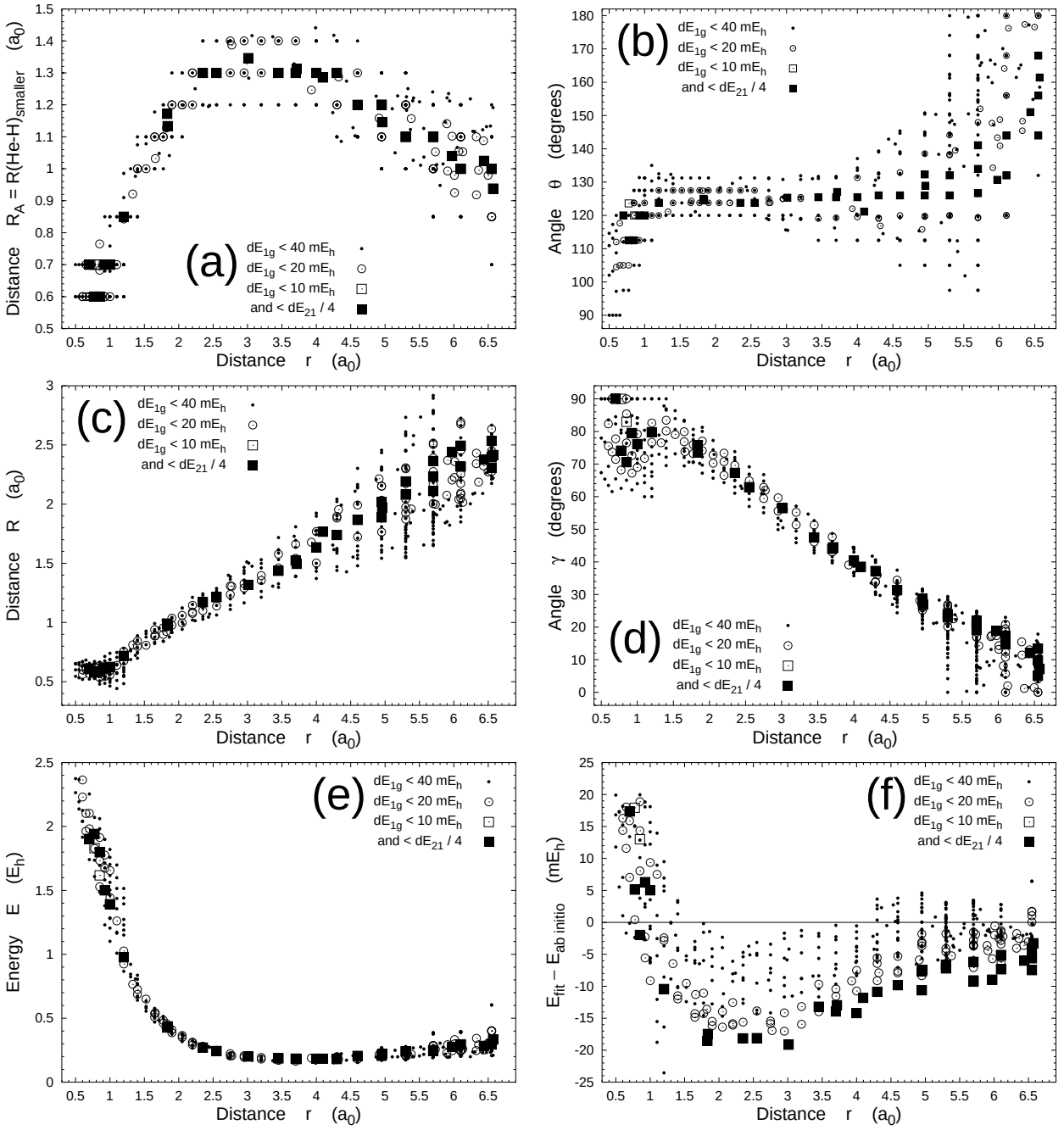

FIG. 9. The approximate position of the conical intersection between the ground state and the first excited state, as a function of the H − H separation r. Plotted points lie close to the conical intersection, as they have small differences dE_{1g} between ground state and first excited state energies ($20 - 40\ mE_h$ for dots, $10 - 20\ mE_h$ for open circles, $0 - 10\ mE_h$ for open and filled squares — filled squares also have $dE_{1g} < 0.25\, dE_{21}$, where dE_{21} is the energy difference between the first and second excited states). (a) The smaller He − H separation R_A, (b) the angle θ of the He atom relative to the closer H atom, (c) the distance R to the center of the H_2 molecule, (d) the angle γ relative to the H_2 molecule, (e) the energy E relative to separated H + H + He, and (f) the error in the fitted BMP surface near the conical intersection.

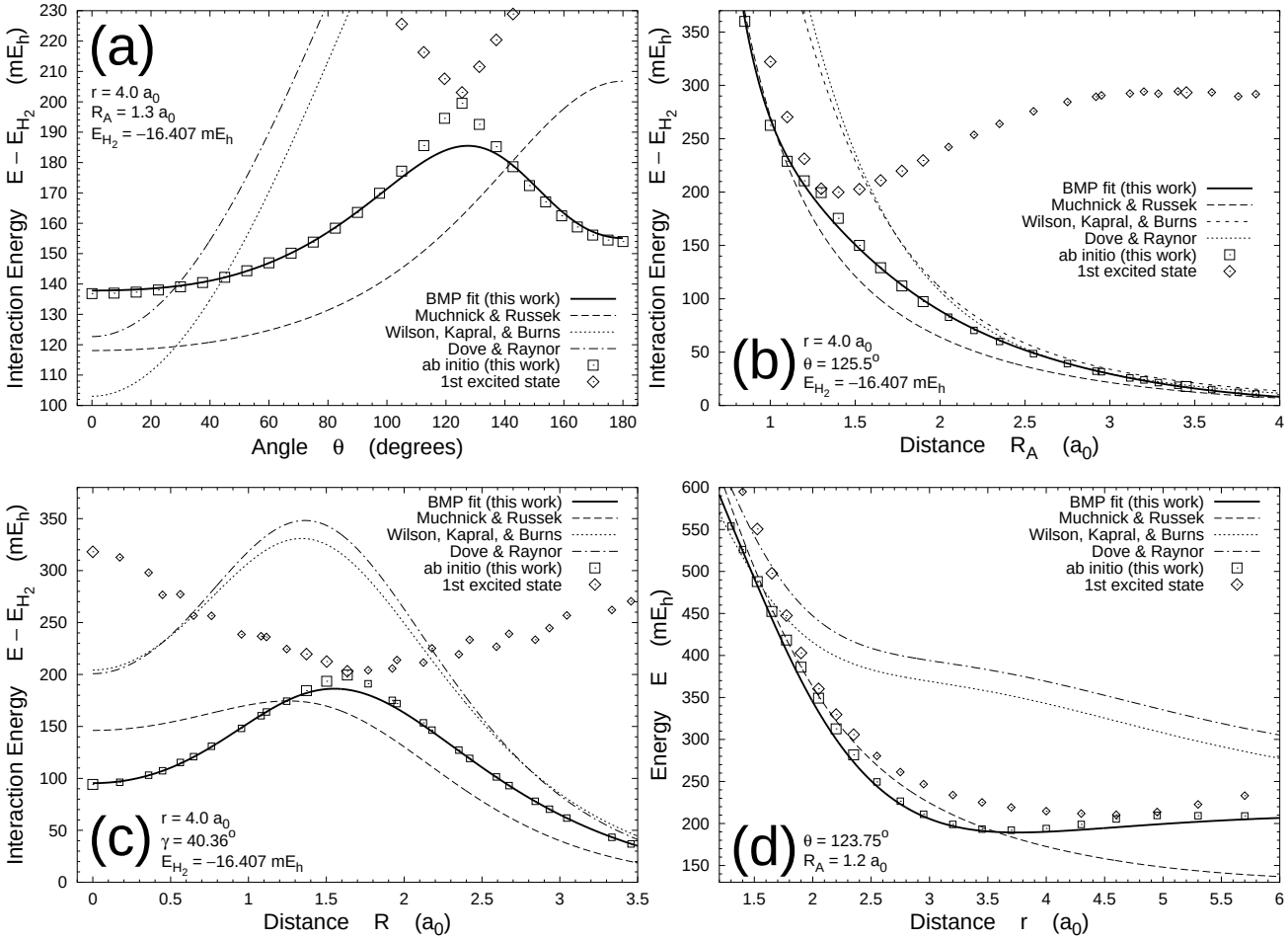

FIG. 10. Four different cuts through (or very near) conical intersection points of the He-H_2 surface, showing *ab initio* ground state energies (squares) and first excited state energies (diamonds). Smaller symbols indicate points near but not on the cut; their energies have been shifted for plotting purposes by the difference in the BMP surface energy corresponding to moving these points onto the cut (as expected, this works well for the ground state, but less well for the first excited state). Also shown are the fitted BMP surface of this work (solid line) and the previous surfaces of Muchnick & Russek (Ref. 24) (dashed line), Wilson, Kapral, & Burns (Ref. 17) (dotted line), and Dove & Raynor (Ref. 20) (dot-dashed line). The Schaefer & Köhler surface (Ref. 22) is not plotted, as this would require extrapolating it far beyond its range of validity.

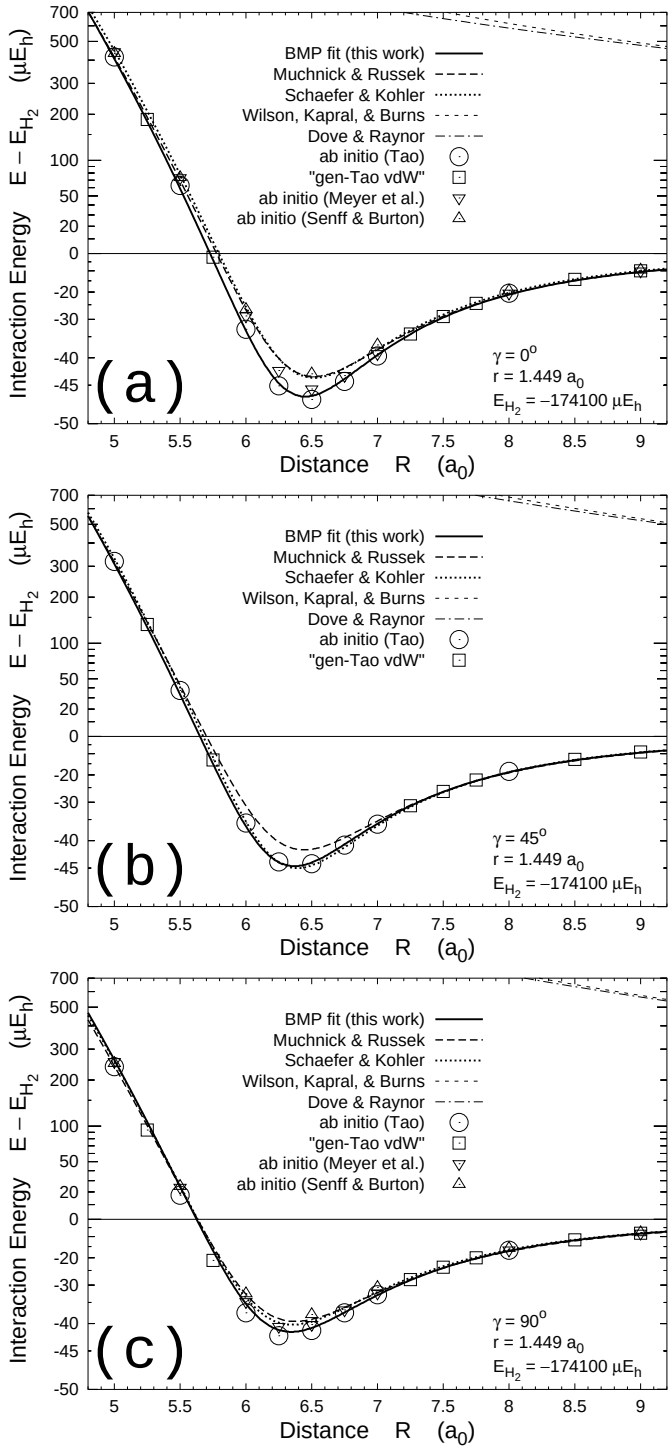

FIG. 11. Comparison of the fitted BMP surface (solid lines) in the van der Waals well at three different orientations with the accurate *ab initio* energies of Tao (Ref. 31) (circles) and the accurate "gen-Tao vdW" energies generated from them (squares); some less-accurate earlier *ab initio* energies (triangles) are also shown, as well as some earlier analytic surfaces (dotted/dashed lines: as in Fig. 10). Note the "shifted-logarithmic" energy scale, used to emphasize effects at the bottom of the van der Waals well. (a) $\gamma = 0°$, (b) $\gamma = 45°$, (c) $\gamma = 90°$.

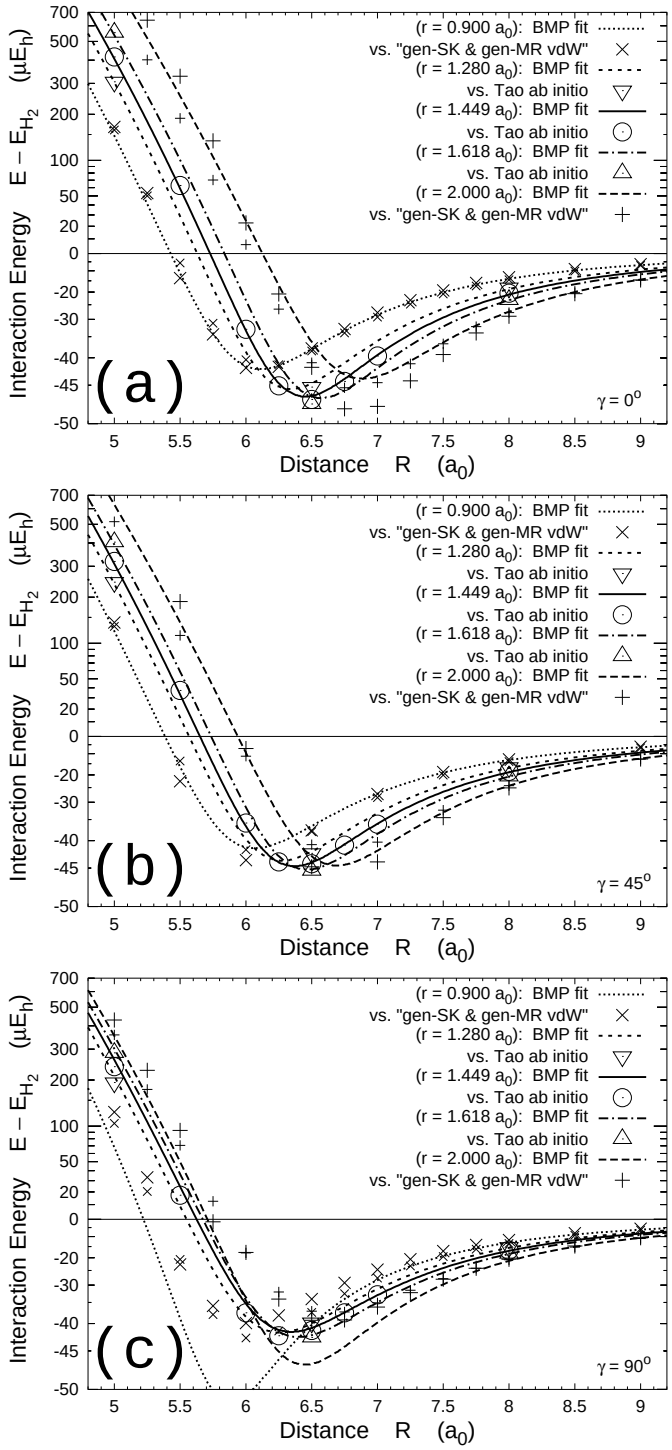

FIG. 12. Effect of H_2-molecule size r on the van der Waals well, at three different orientations (note that $E_{H_2} = \{-83571.4, -171346.4, -174100.0, -167643.3, \text{and } -138203.3\ \mu E_h\}$ for $r = \{0.9, 1.28, 1.449, 1.618, \text{and } 2.0\ a_0\}$, respectively); the energy scale is the same as in Fig. 11. The fitted BMP surface (lines) is compared to the near-H_2-equilibrium *ab initio* energies of Tao (Ref. 31) (circles and triangles) and, at more extreme r values, the much-less-accurate generated energies described in § II A 4 ("gen-SK vdW": larger crosses and plusses, "gen-MR vdW": smaller ones). For the sake of clarity, the "gen-Tao vdW" points are omitted. (a) $\gamma = 0°$, (b) $\gamma = 45°$, (c) $\gamma = 90°$.

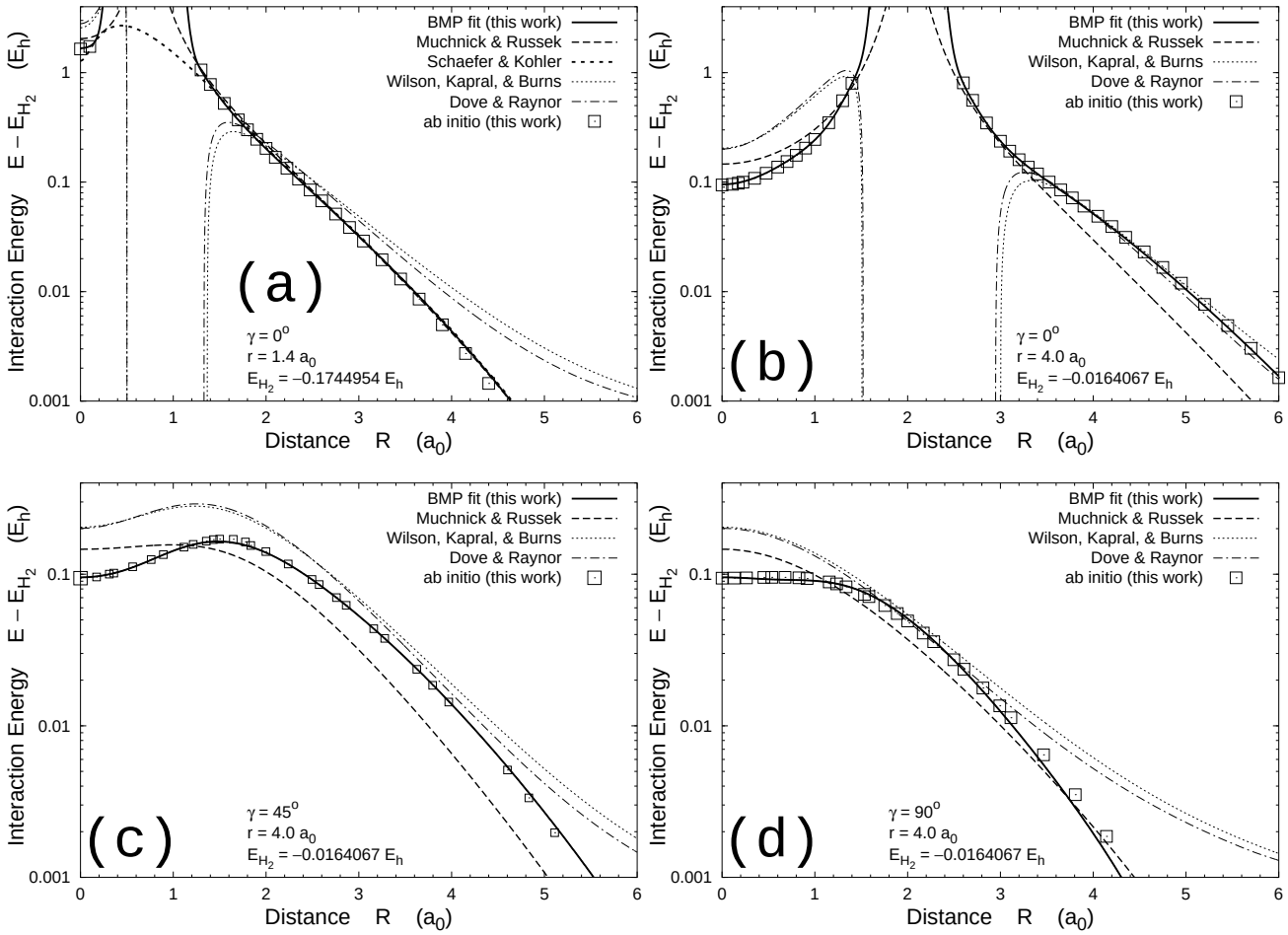

FIG. 13. Comparison of the fitted BMP surface of this work (and also previous analytic surfaces) to the *ab initio* energies, for two H_2-molecule sizes r. Only for the lowest-energy points does the (largely systematic) uncertainty of $\sim 0.6\ mE_h$ in the *ab initio* energies approach or exceed the size of the symbols. (a) $r = 1.4\ a_0$, $\gamma = 0°$: near equilibrium H_2; the "hole" is visible in the Wilson, Kapral, & Burns and Dove & Raynor surfaces (Refs. 17, 20), but the other surfaces are quite accurate. (b) $r = 4.0\ a_0$, $\gamma = 0°$: important for dissociation and high-excitation H_2; the "hole" in the Wilson, Kapral, & Burns and Dove & Raynor surfaces is at even lower energy, and even the Muchnick & Russek surface (Ref. 24) has relatively large fractional errors. The Schaefer & Köhler surface (Ref. 22) is not plotted, as this is too far to extrapolate beyond its range of validity. (c) $r = 4.0\ a_0$, $\gamma = 45°$: similar to (b), but note expanded vertical scale. Smaller symbols correspond to points near but not quite on the $\gamma = 45°$ cut, shifted as in Fig. 10. (d) $r = 4.0\ a_0$, $\gamma = 90°$: similar to (c).