ELECTRONIC SUPPLEMENTARY INFORMATION

On the development of a gold-standard potential energy surface for the OH⁻ + CH₃I reaction

Domonkos A. Tasi, Tibor Győri and Gábor Czakó*

MTA-SZTE Lendület Computational Reaction Dynamics Research Group, Interdisciplinary Excellence Centre and Department of Physical Chemistry and Materials Science, Institute of Chemistry, University of Szeged, Rerrich Béla tér 1, Szeged H-6720, Hungary

*Email: gczako@chem.u-szeged.hu

Potential energy surfaces

For the OH^- + CH_3I reaction, utilizing the stationary points characterized with high-level *ab initio* methods,[1] an initial dataset (13 308 structures) is generated by two methods: modifying the Cartesian coordinates of the stationary points and changing the positions between the fragments of the reactants or products within 2.0 to 10.0 Å from each other. To obtain the energies of the structures of the initial dataset second-order Møller–Plesset perturbation theory (MP2)[2] with the correlation-consistent aug-cc-pVDZ(-PP) basis set[3,4] is used. For iodine relativistic small-core effective core potentials and the corresponding pseudo-potential (PP) basis sets are employed.[4] This initial dataset is used to start the ROBOSURFER program,[5] which adds the remaining points, using the same level of theory. The final version of the dataset contains 36 539 energy points. All fits presented in this work make use of the permutationally invariant polynomial approach.[6,7] The PES is fitted using fifth-order polynomial expansion of Morse-like variables, $\exp(-r_{ij}/a)$, where r_{ij} are the inter-atomic distances and $a = 3$ bohr. To determine the values of the 4693 polynomial coefficients we apply a weighted linear least-squares fit using the weight function of $E_0/(E + E_0) \times E_1/(E + E_1)$, where $E_0 = 94$ kcal/mol and $E_1 = 314$ kcal/mol, and E is the energy relative to the global minimum. We use the DGELSY and DGELS standard LAPACK subroutines to solve the linear least-squares problems.

The development of the PES by ROBOSURFER contains 86, 89, 65, 41, 50, and 49 iterations at 5, 10, 20, 30, 40, and 50 kcal/mol collision energies, respectively. In every iteration 24 trajectories are run with different impact parameters (b): 0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 4.0, and 9.0 bohr (3 trajectories at each b). The hard upper limit of acceptable potential energy is set to 200 kcal/mol relative to the global minimum optimized at MP2/aug-cc-pVDZ(-PP) and no lower limit is applied. The targeted PES accuracy is adjusted to 0.5 kcal/mol, and the full accuracy limit is set to 64 kcal/mol relative to free reactants. The latter value is set by considering the sum of the harmonic zero-point energy (ZPE) of the reactants obtained at MP2/aug-cc-pVDZ(-PP), and the maximum collision energy (50 kcal/mol), which is reduced by 14 kcal/mol as previous exprience[5] has indicated that the sum of reactant ZPEs and the collision energy may be an overly cautious overestimation. The HOLEBUSTER subprogram[5] is not used for the present PES development.

The final dataset (36 539 structures) is recomputed with the following levels of theory:
(1) DF-MP2-F12/aug-cc-pVTZ
(2) BCCD/aug-cc-pVDZ

(3)   CCSD-F12b/aug-cc-pVTZ

(4)   BCCD(T)/aug-cc-pVDZ

(5)   CCSD(T)-F12b/aug-cc-pVTZ

(6)   Composite

resulting about 36 530 *ab initio* energies for each PES (convergence problems are found only in a few cases). DF-MP2-F12, CCSD-F12b and CCSD(T)-F12b stand for the explicitly correlated methods of density-fitted second-order Møller–Plesset perturbation theory,[8] coupled cluster singles, doubles, and CCSD with perturbative triples,[9] respectively. BCCD and BCCD(T) represent Brueckner coupled cluster[10] doubles, and BCCD with perturbative triples. The composite *ab initio* energies are calculated as:

$$CCSD\text{-}F12b/aug\text{-}cc\text{-}pVTZ + BCCD(T)/aug\text{-}cc\text{-}pVDZ - BCCD/aug\text{-}cc\text{-}pVDZ. \qquad (1)$$

After computing the distributions of the energy points for the different PESs, in the case of CCSD(T)-F12b/aug-cc-pVTZ the dataset is reduced by 17 structures from 36 531 to 36 514. This truncated set is also employed for fitting. All PESs are fitted with the same conditions as described above. The root-mean-square (RMS) fitting errors and the size of the dataset for each PES can be found in Table S1. The RMS deviations of the relative energies with respect to the MP2/aug-cc-pVDZ reference are 2.87 [DF-MP2-F12/aug-cc-pVTZ], 3.88 [BCCD/aug-cc-pVDZ], 4.50 [CCSD-F12b/aug-cc-pVTZ], 4.08 [BCCD(T)/aug-cc-pVDZ], 4.49(4.18) [CCSD(T)-F12b/aug-cc-pVTZ(after removing 17 outliers)], and 4.02 [Composite] kcal/mol (see Figure 1).

Ab initio **computations for selected structures**

Several *ab initio* methods are used to obtain the relative energies of selected representative structures from the final dataset:

(1)   HF/aug-cc-pVDZ

(2)   MP2/aug-cc-pVDZ

(3)   MP2-F12/aug-cc-pVDZ

(4)   MP2/aug-cc-pVTZ

(5)   MP2-F12/aug-cc-pVTZ

(6)   DF-MP2-F12/aug-cc-pVTZ

(7)   BCCD/aug-cc-pVDZ

(8)   BCCD(T)/aug-cc-pVDZ

(9)   BCCD(T)/aug-cc-pVTZ

(10) OQVCCD/aug-cc-pVDZ

(11) OQVCCD(T)/aug-cc-pVDZ

(12) CCSD/aug-cc-pVDZ

(13) CCSD-F12b/aug-cc-pVDZ

(14) CCSD-F12b/aug-cc-pVTZ

(15) CCSD(T)/aug-cc-pVDZ

(16) CCSD(T)-F12b/aug-cc-pVTZ

(17) Composite [Eq. (1)]

(18) CCSDT/aug-cc-pVDZ

The HF, OQVCCD, OQVCCD(T), and CCSDT denote Hartree–Fock,[11] optimized-orbital quasi-variational coupled cluster doubles, OQVCCD with perturbative triples,[12] and coupled-cluster singles, doubles, and triples,[13] respectively. All electronic structure computations, except CCSDT, are carried out with the MOLPRO program package,[14] for CCSDT the MRCC program[15] is used, interfaced to MOLPRO.

As seen above we use the F12b variant of the explicitly-correlated CC methods, which is known to have smoother convergence behaviour than F12a toward the complete-basis-set limit.[16] However, due to favourable error cancellation, the CCSD(T)-F12a method may provide more accurate results with small basis sets up to aug-cc-pVTZ.[16] Therefore, F12b was recommended for computations with aug-cc-pVQZ or larger basis sets,[16] whereas F12a is preferred for small basis sets such as cc-pVDZ as used in ref. 17. To test the performance of the F12a and F12b CC methods for the $OH^- + CH_3I$ system, we compute the relative energies of several different structures using both CCSD(T)-F12a and CCSD(T)-F12b with the aug-cc-pVTZ basis. As Table S2 shows, the F12a and F12b results agree with an RMS deviation of 0.30 kcal/mol, which is less than the fitting errors of our PESs. Furthermore, we think that the above RMS value overestimates the difference between the two methods, because our test configurations include many structures with strong multi-reference character usually providing larger differences ($\Delta = 0.2$–0.6 kcal/mol) between F12a and F12b than the single-reference structures ($\Delta = \sim 0.05$ kcal/mol) with T_1 diagnostics less than 0.02 (Table S2). On the basis of these test computations we conclude that the choice of the F12 variant does not have significant effects on the quality of the present PESs, in agreement with our previous study.[18]

Table S2 shows the T_1-diagnostic[19] values for the nine different structures shown in Fig. 2 and Fig. S1. T_1 diagnostics can be used to assess the multi-reference character of the

system; T_1-diagnostic values above 0.02-0.03 predict strong non-dynamic correlation effects. For the structures where the T_1 value is below 0.03 the traditional (T), the Brueckner-type (T), and the full-T correlation energy contributions are the same within 0.4 kcal/mol as seen in Table S2. For the geometries where single bonds, C–I bond in the present case, break homolyticaly the T_1 diagnostic values can be larger than 0.1 indicating strong multi-reference behaviour. In this case the traditional perturbative (T) approximation fails, resulting in too negative (T) energies as shown in Table S2. The reason of this failure is twofold: (1) the perturbation theory may approach a singular region, because some HF orbitals may become quasi-degenerate resulting in small numbers in the denominator of the perturbative (T) correlation energy expression and/or (2) the single HF reference is not good causing substantial errors in the CCSD relative energies, which are amplified by the perturbative approximation.[20] If we do not use the perturbative approach and compute the full T correction interatively, this problem does not occur and the contribution of the triples becomes smaller. The Brueckner coupled-cluster theory uses a non-HF reference (the Brueckner orbitals are optimized to incorporate the effects of single excitations), which may behave better in strongly correlated cases, resulting in a more reasonable (T) correction and better agreement with the full T values as shown in Table S2. The effect of the reference function is seen in our numerical results, as the CCSD and BCCD relative energies differ by about 5 kcal/mol for the structures with $T_1 > 0.1$, whereas the BCCD and OQVCCD results agree with each other within 0.2 kcal/mol. Thus, we conclude that the Brueckner−orbitals-based BCCD(T) method can be useful for PES developments in the regions where the configurations have significant multi-reference characters caused by, for example, a single bond cleavage. Here it is important to note that even if in the present study our numerical test computations have found that the BCCD(T) method agrees slightly better with the CCSDT results than OQVCCD(T), the latter, which was developed to describe multi-reference systems with a single-reference function, may also become useful in future applications, especially in regions where multiple bonds break.[12,20]

To confirm the correct numerical computation of the unphysically large traditional (T) corrections, we have performed the CCSD(T) computations both with MOLPRO[14] and MRCC[15] program packages, which have different implementations of the same CCSD(T) method. As seen in Table S2, the two different programs give the same CCSD(T) relative energies with an RMS deviation of only 0.0004 kcal/mol, thus, the failure of the traditional (T) method is not caused by implementation issues in the *ab initio* codes.

Quasiclassical trajectory computations

QCT computations are performed for the $OH^-(v = 0) + CH_3I(v = 0)$ reaction on the analytical *ab initio* PESs described above. Standard normal-mode sampling is used to prepare the vibrational ground states ($v = 0$) of OH^- and CH_3I and the rotational temperatures are set to 0 K. The initial orientations of the reactants are randomly sampled and the initial distance of the reactants is 40 bohr with a given b. The trajectories are run at collision energies of 5, 20, and 50 kcal/mol using a time step of 0.0726 fs. At each collision energy, b is scanned with the step size of 1 bohr from 0 to b_{max}, where the probability of the reaction becomes 0. At each b, 5000 trajectories are run, and each trajectory is propagated until the largest interatomic separation becomes 1 bohr larger than the largest initial one. Therefore, about 210 000 trajectories are run for each PES, thus more than 1 million trajectories are studied in present work. Cross sections are obtained by a b-weighted numerical integration of the opacity functions (reaction probabilities as a function of b).

Table S1. The root-mean-square fitting errors (in kcal/mol) and the size of the datasets of various PESs for the OH⁻ + CH₃I reaction

Level of theory[a]	Number of structures	Energy ranges (kcal/mol)		
		0 – 94	94 – 188	188 – 471
MP2/aug-cc-pVDZ	36539	1.02	1.76	1.68
DF-MP2-F12/aug-cc-pVTZ	36538	0.99	1.73	1.73
BCCD/aug-cc-pVDZ	36532	1.02	1.76	1.56
CCSD-F12b/aug-cc-pVTZ	36532	0.99	1.72	1.79
BCCD(T)/aug-cc-pVDZ	36532	1.03	1.74	1.51
CCSD(T)-F12b/aug-cc-pVTZ	36531	1.09	1.83	2.06
CCSD(T)-F12b/aug-cc-pVTZ'	36514	1.02	1.72	1.76
Composite [Eq. (1)]	36529	1.00	1.69	1.81

[a] The prime symbol (') denotes the removal of 17 configurations with relative energies below the MP2 data by more than 50 kcal/mol.

Table S2. T_1 diagnostics, energy corrections (kcal/mol), and energies (kcal/mol) of representative structures obtained by different *ab initio* levels of theory (energy corrections and energies are relative to those of the reactants)

Structure[a]	T_1[b]	(T)[c]	B-(T)[d]	full-T[e]	F12a[f]	F12b[g]	Δ[h]	MOLPRO[i]	MRCC[j]	Δ[k]
Fig. 2 (left)	0.014	−0.86	−0.73	−0.86	18.64	18.59	0.05	19.26	19.26	0.00
Fig. 2 (right)	0.157	−120.50	−32.04	−26.73	−57.68	−57.35	0.33	−61.02	−61.02	0.00
Fig. S1 (A)	0.085	−11.30	−8.43	−9.39	40.96	41.27	0.31	36.87	36.87	0.00
Fig. S1 (B)	0.109	−74.52	−43.73	−28.68	−2.25	−2.87	0.63	−2.67	−2.67	0.00
Fig. S1 (C)	0.027	−1.99	−2.27	−1.91	51.10	51.08	0.02	47.82	47.82	0.00
Fig. S1 (D)	0.133	−98.02	−34.96	−27.27	−29.98	−29.94	0.05	−34.20	−34.20	0.00
Fig. S1 (E)	0.019	−1.58	−1.63	−1.48	32.70	32.75	0.05	36.79	36.79	0.00
Fig. S1 (F)	0.155	−134.31	−33.23	−27.13	−74.24	−74.03	0.21	−75.63	−75.63	0.00
Fig. S1 (G)	0.094	−32.34	−12.06	−13.92	12.85	13.27	0.42	10.29	10.29	0.00
RMS							**0.30**			**0.00**

[a] Structures shown in Figure 2 and Figure S1.

[b] T_1 diagnostic values obtained at the CCSD-F12b/aug-cc-pVTZ level of theory.

[c] (T) corrections obtained as CCSD(T)-F12b – CCSD-F12b using the aug-cc-pVTZ basis.

[d] Brueckner (T) corrections obtained as BCCD(T) – BCCD using the aug-cc-pVDZ basis.

[e] Full-T corrections obtained as CCSDT – CCSD using the aug-cc-pVDZ basis.

[f] CCSD(T)-F12a/aug-cc-pVTZ relative energies.

[g] CCSD(T)-F12b/aug-cc-pVTZ relative energies.

[h] Absolute differences (kcal/mol) between the F12a and F12b relative energies.

[i] CCSD(T)/aug-cc-pVDZ relative energies obtained by MOLPRO.

[j] CCSD(T)/aug-cc-pVDZ relative energies obtained by MRCC.

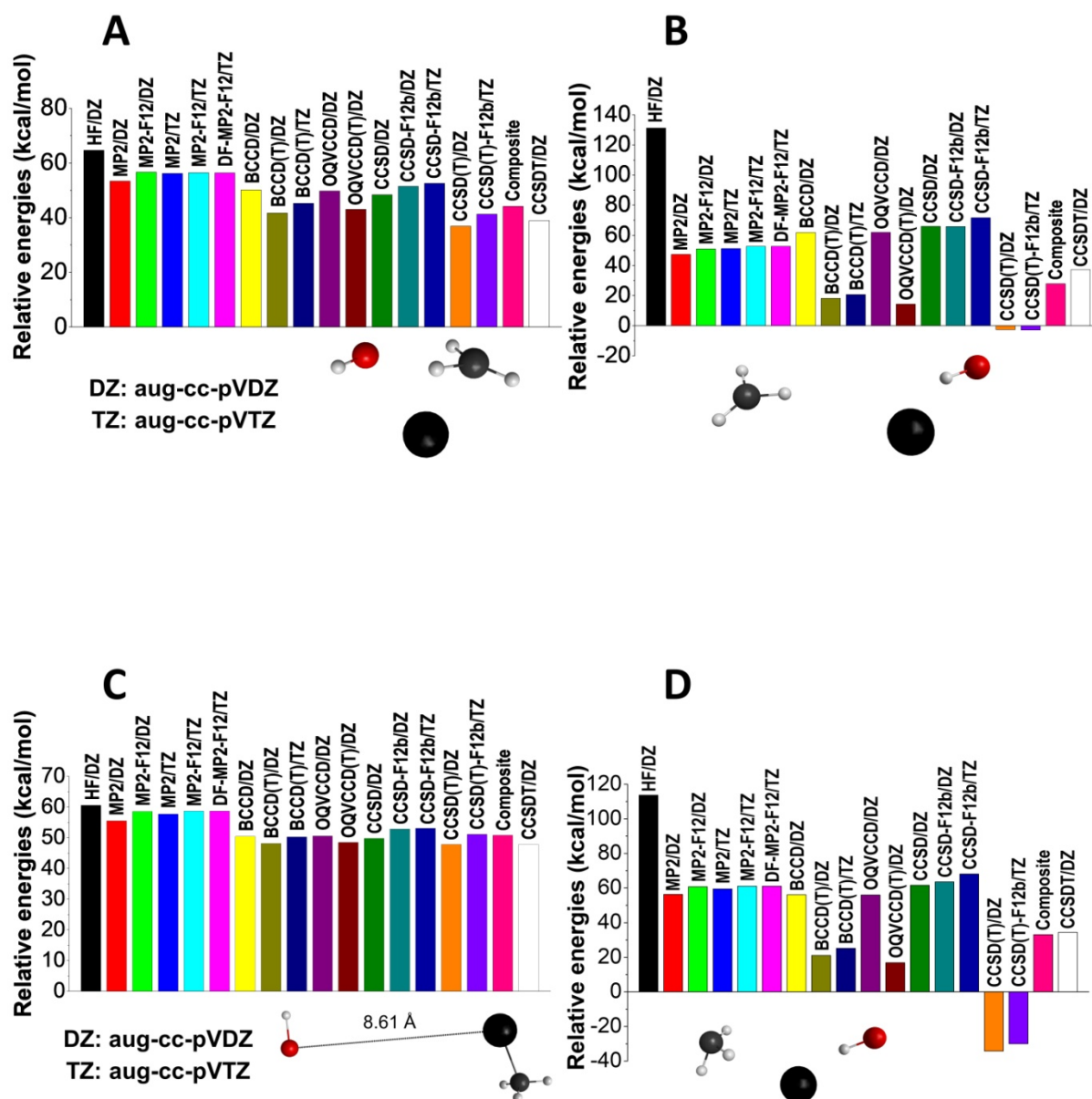[k] Absolute differences (kcal/mol) between the MOLPRO and MRCC results. RMS deviation is 0.0004 kcal/mol.

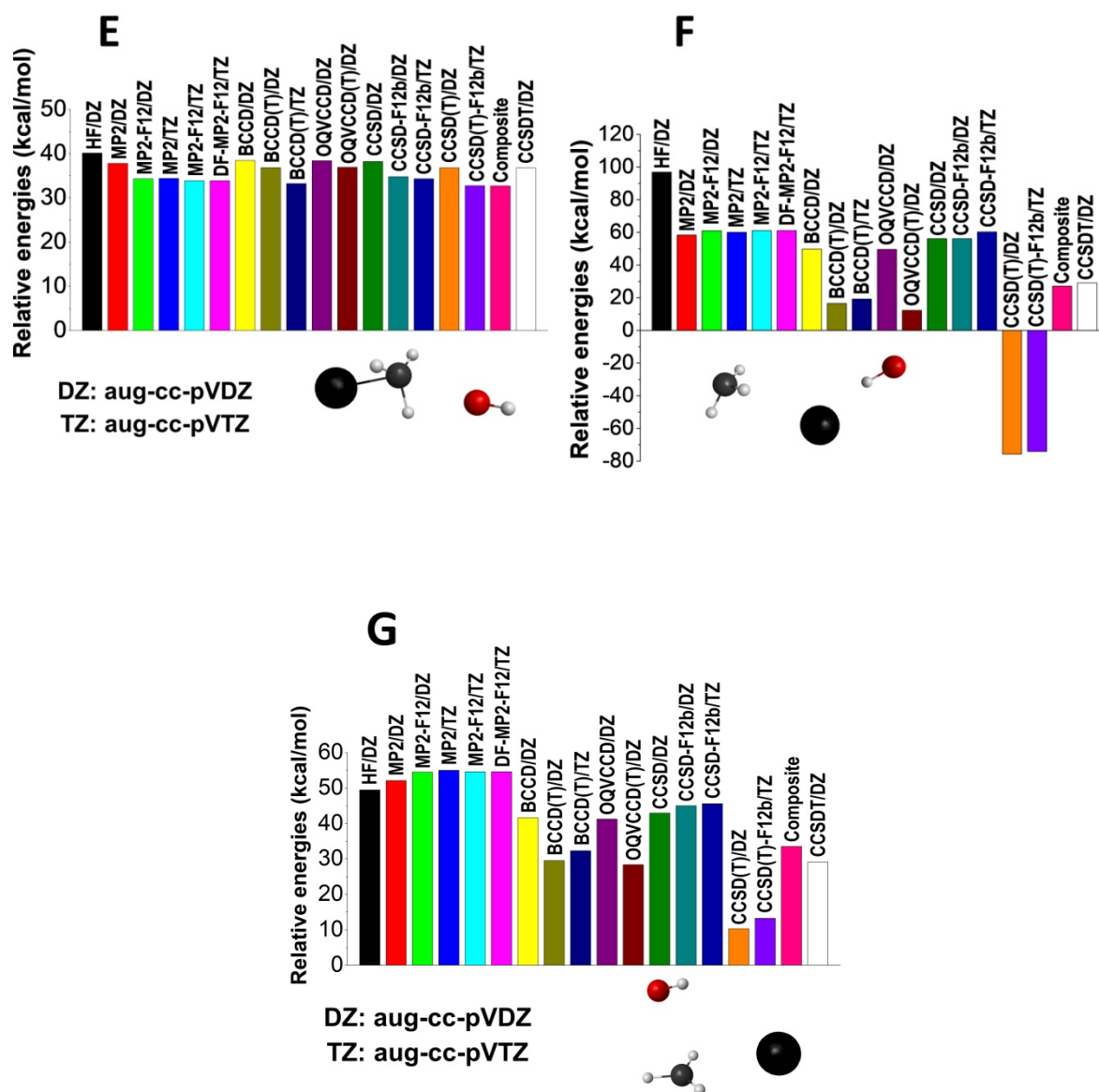Figure S1 continues on the next page

Figure S1 continues

Figure S1. Energies of representative structures relative to the reactants obtained by different *ab initio* levels of theory.
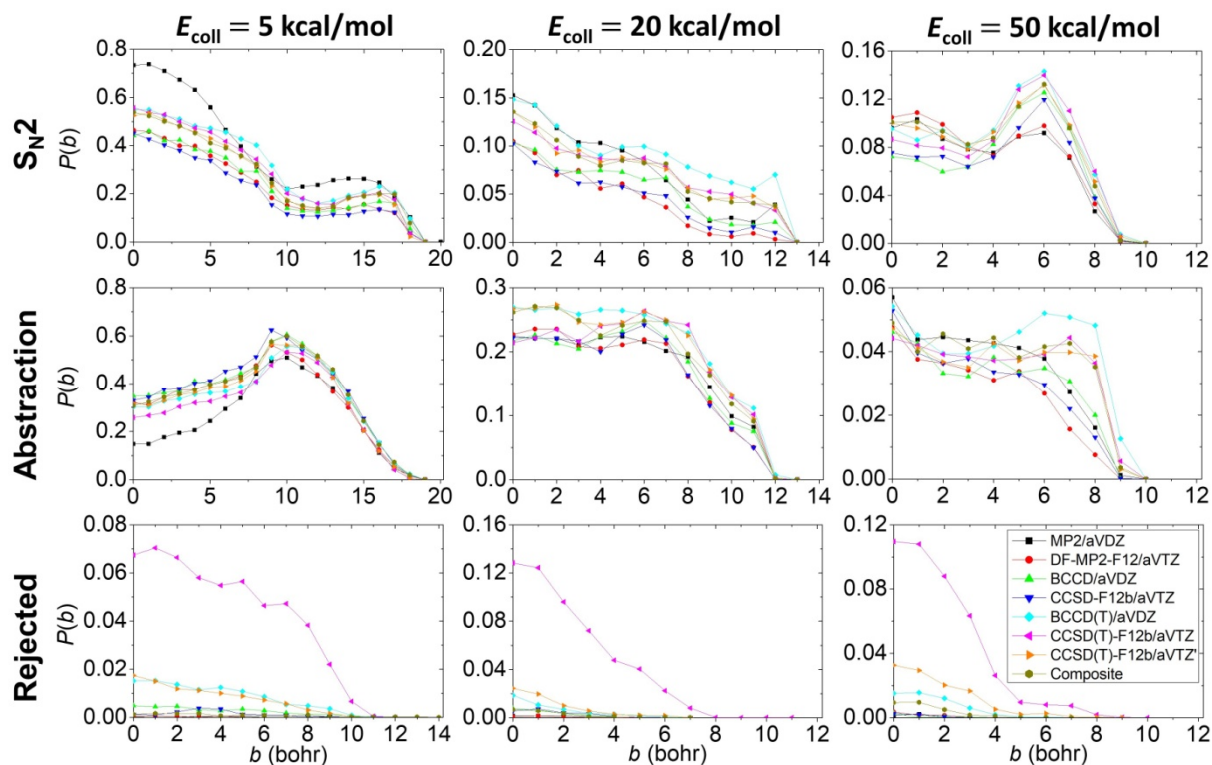
Figure S2. Opacity functions of the S_N2 ($I^- + CH_3OH$), proton-abstraction ($H_2O + CH_2I^-$), and rejected channels of the $OH^- + CH_3I$ reaction obtained on the different analytical *ab initio* PESs at collision energies of 5, 20, and 50 kcal/mol. Rejected trajectories result in unphysical, energetically non-available products (*e.g.*, dissociation to many fragments). The prime symbol (') denotes the removal of 17 configurations with relative energies below the MP2 data by more than 50 kcal/mol.

References

[1] D. A. Tasi, Z. Fábián and G. Czakó, *J. Phys. Chem. A*, 2018, **122**, 5773.

[2] C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.

[3] T. H. Dunning, Jr., *J. Chem. Phys.*, 1989, **90**, 1007.

[4] K. A. Peterson, D. Figgen, E. Goll, H. Stoll and M. Dolg, *J. Chem. Phys.*, 2003, **119**, 11113.

[5] T. Győri and G. Czakó, *J. Chem. Theory Comput.*, 2020, **16**, 51.

[6] B. J. Braams and J. M. Bowman, *Int. Rev. Phys. Chem.*, 2009, **28**, 577.

[7] J. M. Bowman, G. Czakó and B. Fu, *Phys. Chem. Chem. Phys.*, 2011, **13**, 8094.

[8] H.-J. Werner, T. B. Adler and F. R. Manby, *J. Chem. Phys.*, 2007, **126**, 164102.

[9] T. B. Adler, G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2007, **127**, 221106.

[10] K. A. Brueckner, *Phys. Rev.*, 1954, **96**, 508.

[11] W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, *Molecular Orbital Theory*, Wiley: New York, 1986.

[12] J. B. Robinson and P. J. Knowles, *Phys. Chem. Chem. Phys.*, 2012, **14**, 6729.

[13] J. Noga and R. J. Bartlett, *J. Chem. Phys.*, 1987, **86**, 7041.

[14] MOLPRO, version 2015.1, a package of *ab initio* programs, H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz and others, see http://www.molpro.net.

[15] MRCC, a quantum chemical program suite written by M. Kállay, P. R. Nagy, Z. Rolik, D. Mester, G. Samu, J. Csontos, J. Csóka, B. P. Szabó, L. Gyevi-Nagy, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei and B. Hégely, See also Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki and M. Kállay, *J. Chem. Phys.*, 2013, **139**, 094105, as well as: www.mrcc.hu.

[16] G. Knizia, T. B. Adler and H.-J. Werner, *J. Chem. Phys.*, 2009, **130**, 054104.

[17] M. B. Oviedo, N. V. Ilawe and B. M. Wong, *J. Chem. Theory Comput.*, 2016, **12**, 3593.

[18] G. Czakó, I. Szabó and H. Telekes, *J. Phys. Chem. A*, 2014, **118**, 646.

[19] T. J. Lee and P. R. Taylor, *Int. J. Quant. Chem.*, 1989, **S23**, 199.

[20] J. B. Robinson and P. J. Knowles, *J. Chem. Phys.*, 2013, **138**, 074104.