SVECV-f12: benchmark of a composite scheme for accurate and cost effective evaluation of reaction barriers.

Oscar N. Ventura, Martina Kieninger, Aline Katz, Mauricio Vega-Teijido, Marc Segovia, Kenneth Irving*

Computational Chemistry and Biology Group, CCBG, DETEMA, Facultad de Química, Udelar, Montevideo, Uruguay

Abstract: A simple composite scheme is presented, based on a combination of density functional geometry and frequencies evaluation, valence energies obtained using the CCSD(T)-f12 method extrapolated to the complete basis set limit, and core-valence correlation corrections employing the MP2 method. The procedure was applied to the 38 reactions in Truhlar's HTBH38/08 and NHTBH38/08 databases and the errors in the barriers with respect to their best values are presented. Mean unsigned deviation (MUD) for the complete set of 68 independent barriers is 0.40 kcal mol^{-1}, compared to 1.31 kcal/mol for G4 and 1.62 kcal/mol for the dispersion-corrected M06-2X method. The accuracy of the procedure is also better that that of other calculations using composite methods of similar cost. The MUD of the new scheme on the barriers in the DBH24/08 subset (12 out of the 38 reactions in the other two sets) is 0.27 kcal mol^{-1}, better than that obtained at the expensive CCSD(T,full)/aug-cc-pCV(T+d)Z level (0.46 kcal mol^{-1}) and comparable to the most exact (and costly) Wn calculations (MUD=0.14 kcal mol^{-1}). The maximum unsigned deviation (MaxUD) for all the reactions studied is 0.99 kcal/mol. G4 and M06-2X, on the other side, exhibit MaxUDs of 6.7 and 8.0 kcal/mol respectively. The method was further tested against a subset of the reactions in the databases, for which the geometry and energies of all species were determined at the much more demanding CCSD(T)-F12//pVQZ-F12 level. These results showed that Truhlar's calculations in this subset are off the best values by a considerable amount, with an rmse of 0.56 kcal/mol. As a consequence, a new dataset of barrier heights, SV20, is presented. The SVECV-F12 procedure on this SV20 database results in rmse and MUD values of only 0.21 and 0.16 kcal/mol. The possible residual errors introduced by the approximations used for each component of the method are tested against more sophisticated calculations and shown to be accurate enough to obtain barriers well under the chemical precision limit at a reasonable cost for molecules of interest in atmospheric chemistry.

Keywords: Transition states; reaction barriers; chemical reactivity; CCSD(T)-F12; explicit correlation.

*Corresponding author Prof. Oscar N. Ventura

Computational Chemistry and Biology Group,

CCBG, DETEMA, Facultad de Química, Udelar,

Montevideo, Uruguay

E-mail: Oscar.N.Ventura@gmail.com

ORCID: *0000-0001-5474-0061*

Introduction

Accurate prediction of barrier heights for chemical reactions is a necessary, albeit not sufficient, prerequisite for the accurate calculation of reaction rates, as well as branching ratios for the products. While the prediction of the thermochemistry of chemical reactions in gas phase has reached the so-called chemical accuracy (± 1 kcal/mol) for molecules of up to a dozen non-hydrogen atoms, such precision is not yet possible concerning barrier heights in general (see, for instance, Vereecken *et al.* [1] for a comprehensive discussion). Since the barriers depend on the precise evaluation of both stable species (reactants and products) and transition states, normally very resource intensive methods are required to achieve high precision. CCSD(T) calculations [2-4] or beyond (including non-perturbative triple and quadruple excitations) are being presently employed in combination with very extensive basis sets (from quadruple up to sextuple zeta) extrapolated to the CBS limit, to obtain sub-kJ/mol precision [5]. But, of course, at an enormous computer cost.

Simplified procedures, usually known as composite methods [1], have been devised to reduce the computer cost to manageable levels, allowing the treatment of molecules with more than the barest minimum number of atoms. Well known procedures in this area are the complete basis set methods developed by Peterson and coworkers [6-10] (of which the most popular is CBS-QB3 [9,10]), Gn methods developed by Pople and collaborators [11-19] (which latest version are G4 and G4MP2 [14,18]), Weizmann-n (Wn) methods of Martin and coworkers [19-23] (the most accurate and expensive of which is W4 [23]) and the correlation consistent composite approach (ccCA) method of Wilson *et al.* [24-29]. Other less well-known or more purpose specific composite procedures, like the HL [30], HEAT [31, 32], and focal point analysis [33, 34] methods, have also being proposed and used less frequently in the literature. Wn, HEAT and focal point analysis are methods used normally for benchmarking, due to their high accuracy (maximum errors of less than 1 kJ/mol) but commensurate high demand of computer resources.

All composite methods have some characteristics in common. Normally, a density functional method (DFT) like B3LYP for instance, or MP2 perturbative theory are employed with relatively small basis sets to obtain geometries, frequencies and vibrational components of the thermodynamic functions. Single-point calculations are run then on the optimum geometries obtained, with progressively larger basis sets and post-Hartree-Fock methods to include the correlation energy of the valence shell (at the MP4, QCISD(T) or CCSD(T) levels). Finally, core-valence correlation energy may be included using especial basis sets with tight core functions (normally using some perturbative method of calculation) and, in some cases, empirical corrections are included as well.

As it is the case in other areas of computational chemistry, a compromise must be adopted between the accuracy desired and the resources available. Methods like CBS-QB3 are very fast and quite accurate on average for the study of thermochemistry, but they do not reach chemical accuracy (defined usually to mean ±1 kcal mol^{-1} or ~±4 kJ mol^{-1}) and in some cases the errors are quite large. G4 is much more accurate and costly, but still oftentimes fail (as an example, consider the barrier for the reaction $^{\bullet}CH_3 + FCl \rightarrow CH_3F + Cl^{\bullet}$ for which the G4 error, with respect to the best calculation, is more than 6 kcal/mol). Other methods, like Wn, are normally very expensive and only applicable to small molecules. Thus, the development of new schemes which show a good cost/accuracy ratio is still an active area of research.

Obviously, any approximate method may perform marvelously well for a given reaction and fail miserably for an unrelated one. Thus, it has been customary to assess the goodness of a new method according to the closeness to experiment or to very accurate calculation in some set(s) of chosen reactions. In the case of barrier heights, there exist two databases maintained by Don Truhlar and his group with this purpose. HTBH38 contains 19 hydrogen transfer reactions with values for the forward and reverse classical barrier heights (excluding ZPE) [35], and NHTBH38 containing also 19 reactions, this time including "heavy"-atom transfer, bimolecular nucleophilic substitution (SN2), association, and unimolecular reactions [36]. Two versions of these databases are available. In the 2004 version [35, 36], most of the barrier heights were derived from W1, some from the next computational level, W2h, or from accurate literature values. The 2008 version of the databases, which are those we used in this paper, contain also values calculated at the CCSDTQ5/CBS level via the W4 method by Karton *et al.* [37]. Zheng, Zhao and Truhlar [38] specifically developed a subset of those databases, consisting of a representative group of 12 reactions which they called DBH24/08, and used it to assess 348 model chemistries. They arrived to the conclusion that the best performance among those model chemistries was achieved by the CCSD(T)(full)/aug-cc-pCV(T+d)Z method with a mean unsigned error of 0.46 kcal/mol (~2 kJ mol^{-1}). The whole set of 38 reactions were used in this paper, to obtain a better sampling, but also the metrics for the reduced subset are given as a subproduct. The results obtained are compared to those published using other methods in the literature, to assess the behavior of this new simple scheme.

Computational details

The reactions in the HTBH38/04, NHTBH38/04 and modified data in the DBH24/08 databases were used as benchmark targets for the new composite method proposed. Individual barrier heights for the reactions were obtained from the repository at the University of Minnesota database web site [39]. Results at the G4 level for the 2004 version of the databases were taken from the paper by Curtiss, Redfern and Raghavachari [40]. These calculations were also checked against the 2008 databases, and the G4 values recalculated increasing the threshold for geometry optimizations and the density of the grid for calculation of the integrals in the B3LYP portion of the G4 scheme. Some small differences were noticed between the two sets of calculations. Results at the ccCA level of calculation were taken from the paper by Grimes *et al.* [29], while the most recent results by Karton at the W3lite-F12 level were also considered [5, 37].

The new composite method proposed, which will be called SVECV-f12 (a Simple Version of an Extrapolated, Core-Valence correlation corrected, CCSD(T)-f12 method) from now on, consists of three steps

1. Geometry optimization and frequency calculation were performed using a DFT method. M06-2X [41] was chosen because of its good accuracy, but with the caveat that it may give wrong results for transition states with large multireference character. Although the M06 methods and related include dispersion implicitly through the determination of the empirical parameters, in this paper it was explicitly included using Grimme's original D3 method [42]. Dunning's aug-cc-pVTZ basis set was adopted as a compromise between accuracy and speed of calculation [43]. Zero-point energy, as well as thermal and vibrational contributions to the final energies, were taken from the DFT calculations.
2. Single-point energy calculations using the DFT geometries were then performed at the CCSD(T)-f12 level [44,45], using both the pVDZ-f12 and pVTZ-f12 basis set [46]. Extrapolation to the CBS limit was performed using Martin's two-point extrapolation formula [47]:

$$E = E(CBS) + \frac{A}{(n+\frac{1}{2})^4} \text{ , n =2, 3}$$

3. Finally, the core-valence correlation energy was calculated at the MP2 level, as

$$\Delta E_{corr}^{CV} = E_{cc-pCVTZ}^{MP2}(full) - E_{cc-pCVTZ}^{MP2}(valence)$$

All geometry optimizations were performed using very tight thresholds and the finest available grids for evaluation of integrals. Transition state optimizations were started performing an analytical evaluation of the hessian at the DFT level employed. If necessary because of convergence problems, analytical evaluation of the hessian was repeated at each step in the geometry optimization procedure. DFT and MP2 calculations were performed with Gaussian 16, Revision C.01 [48], while CCSD(T)-f12 calculations were performed using Molpro version 2019.2 [49, 50]. Basis sets for the calculations were either the ones included in the codes employed or, when necessary, obtained from the Basis Set Exchange [43].

A few words are perhaps necessary as to the choice of the different components. It is known that the B3LYP DFT method used in G4 produces bad geometries in some cases. Curtiss *et al.* [40] attributed these failures in the group of reactions contained in the datasets to the bad description of transition states for fluorine-containing systems. A better method in this respect, M06-2X [41], which has a smaller error [38] for the determination of geometries was chosen in this work. Explicit empirical dispersion corrections [42] were included to describe better the non-covalent attractive forces present in some of the transition states in the databases. DFT methods exhibit a better convergence behavior than molecular orbital procedures with respect to basis sets. Therefore, no set larger than valence triple-zeta (enlarged by diffuse functions to treat better second-row atoms) is deemed necessary to achieve good convergence. Zheng, Zhao and Truhlar [38] even recommend a smaller basis set, cc-pVTZ+, as being more efficient with the same degree of accuracy. Aug-cc-pVTZ was adopted in this work but the point will be examined further with better methods for the geometry optimization.

For the calculation of the main component of the energies, the explicitly correlated CCSD(T)-f12 method was chosen, which of course is more expensive than CCSD(T). However, it has been noticed in previous calculations that the convergence properties of CCSD(T)-f12 with the increase of the basis set are much better than for CCSD(T). Thus, only DZ and TZ basis sets are needed for obtaining a good extrapolated energy and the less demanding basis sets compensate for the more expensive calculation method. Varandas has recently performed a very complete comparison of the CCSD(T) and CCSD(T)-F12 convergence toward the CBS limit which is worth mentioning [51]. Nevertheless, this point will be further explored in this paper. An experiment was conducted also using CCSD(T) for the calculation of the core-valence correlation correction of a subset of the reactions, comparing it with the MP2 results to judge their precision and accuracy. A triple-zeta basis set augmented with tight core functions was chosen for these calculations.

In the course of this research some hints pointed toward the possibility that Truhlar's values for the barriers of some reactions were suboptimal. A more complete set of calculations was then performed. Geometry optimizations, frequencies and energy determinations were performed at the CCSD(T)-F12/pVnZ-F12 level (n=D, T and Q). Needless to say, even for molecules as small as those in this set, the CCSD(T)-F12/pVQZ-F12 calculations (for which numerical evaluation of the Hessian was necessary) consume a large amount of CPU-time and memory. However, these calculations probably afford the best geometrical structures known in general for the transition states involved. Presumably, the barriers obtained in that way are also of high quality.

Results

Geometries obtained at the M06-2X/D3/aug-cc-pVTZ level are given in the Supporting Information Section. Table SM1 in there contains the energies (total energy, ZPE, enthalpy and free energies) of all the species considered. Only the values of direct and reverse barriers as well as the errors with respect to Truhlar's values (called "best" for simplicity) will be considered in this section.

Table 1 shows the reactions in the HTBH38/08 database, listing the "best" calculated values for the direct and reverse reactions, as well as the errors of the M06-2X/D3/aug-cc-pVTZ, G4 and SVECV-f12 calculations performed in this paper with respect to those values. Table 2 shows the same for the NHTBH38/08 database. Two statistical metrics have been included in both tables, mean unsigned deviation (MUD) and root mean square deviation (RMSD). Repeated values, corresponding to equal forward and reverse barriers for symmetric reactions, have been excluded from the evaluation of MUD and RMSD, resulting in 68 different barriers in total, divided in 36 for HTBH38/08, 32 for NHTBH38/04 and 22 for DBH24/08 (included already in the former). Anyway, very small difference in MUD and RMSD are noticed if repeated values are not excluded. The selected barrier heights from the previous two sets that are included in the DBH24/08 subset, are specified in the footnotes to the tables. Collective metrics for the three data sets, along with values taken from the literature, are shown in Table 3. The correlation between individual values obtained with the new method and "best" ones is shown in Fig. 1.

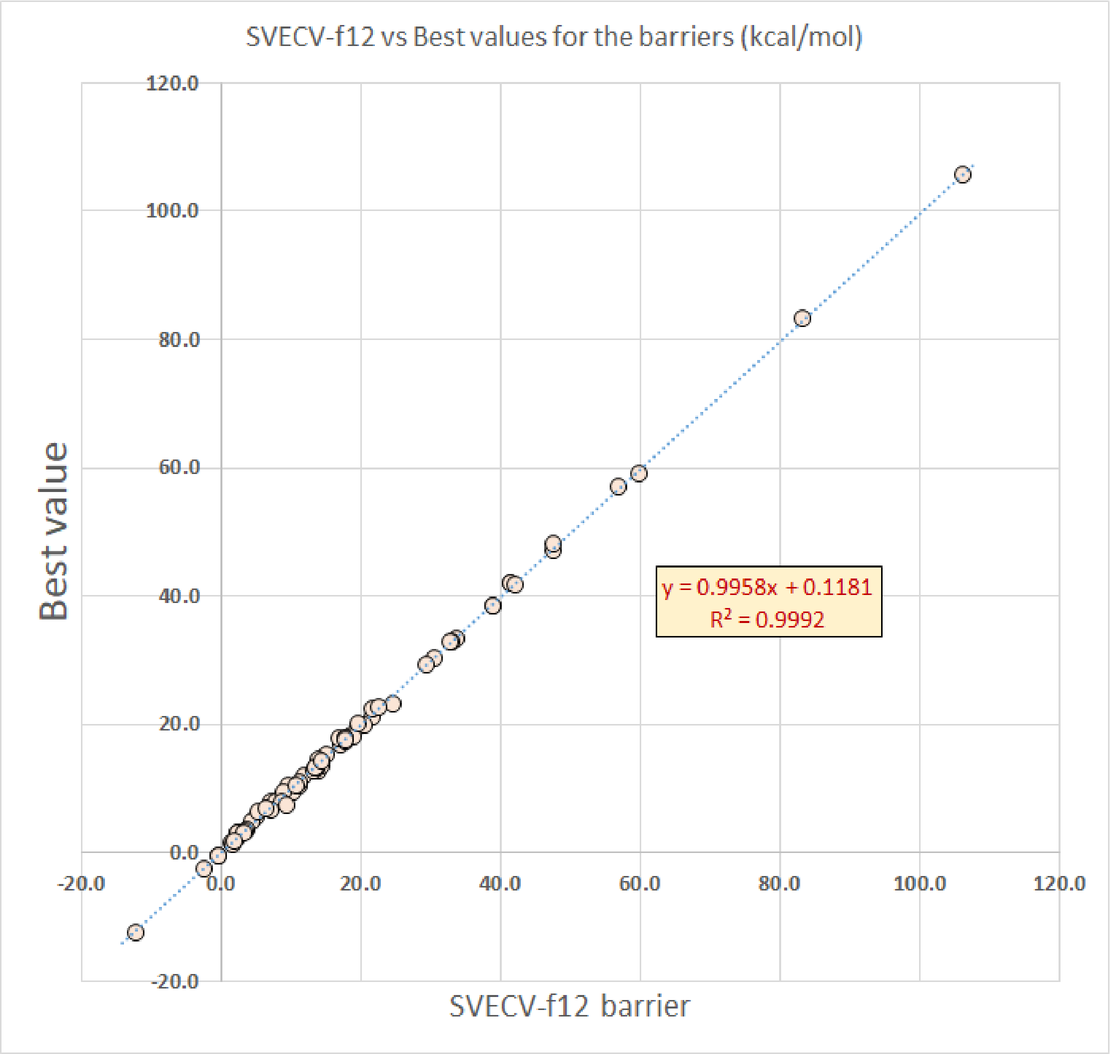


Figure 1. Linear correlation between SVECV-f12 calculated barriers and "best" known values (in kcal/mol). The linear equation and correlation coefficient are depicted on the graph.

Table 1. Theoretical barrier heights (in kcal/mol) obtained in this paper at the G4, M06-2X/D3/aug-cc-pVTZ and SVECV-f12 levels, for the reactions in the HTBH38/08 dataset, together with "best" values taken from [38, 39].

		Reaction[f,g]	Barrier[a]	G4[b]	M062X[b,c]	SVECV-f12[b]	Best[d]
R01	1	$H^\bullet + HCl \rightarrow H_2 + Cl^\bullet$	$V_f^{\neq}$	3.13	4.80	4.90	5.70
	2		$V_r^{\neq}$	7.10	7.54	6.95	7.86
R02	3	$^\bullet OH + H_2 \rightarrow H_2O + H^\bullet$	$V_f^{\neq}$	6.91	4.97	5.35	4.90
	4		$V_r^{\neq}$	21.92	21.05	21.60	21.20
R03	5	$^\bullet CH_3 + H_2 \rightarrow CH_4 + H^\bullet$	$V_f^{\neq}$	13.02	5.60	11.83	12.10
	6		$V_r^{\neq}$	15.28	15.54	14.91	15.30
R04	7	$^\bullet OH + CH_4 \rightarrow {}^\bullet CH_3 + H_2O$	$V_f^{\neq}$	7.09	5.42	6.31	6.50
	8		$V_r^{\neq}$	19.84	11.56	19.49	19.60
R05	9	$H^\bullet + H_2 \rightarrow H_2 + H^\bullet$	$V_f^{\neq}$	10.47	11.56	9.67	9.60
	10		$V_r^{\neq}$	10.47	11.56	9.67	9.60
R06	11	$^\bullet OH + NH_3 \rightarrow H_2O + {}^\bullet NH_2$	$V_f^{\neq}$	4.60	2.63	3.56	3.00
	12		$V_r^{\neq}$	14.77	12.21	13.84	12.70
R07	13	$HCl + {}^\bullet CH_3 \rightarrow Cl^\bullet + CH_4$	$V_f^{\neq}$	1.08	-6.22	1.43	1.70
	14		$V_r^{\neq}$	7.32	6.46	6.55	7.06
R08	15	$^\bullet OH + C_2H_6 \rightarrow H_2O + {}^\bullet C_2H_5$	$V_f^{\neq}$	4.20	3.04	3.63	3.20
	16		$V_r^{\neq}$	20.73	19.62	20.36	19.90
R09	17	$F^\bullet + H_2 \rightarrow HF + H^\bullet$	$V_f^{\neq}$	1.72	0.48	1.54	1.42
	18		$V_r^{\neq}$	33.09	31.87	33.65	33.40
R10	19[e]	$O + CH_4 \rightarrow {}^\bullet OH + {}^\bullet CH_3$	$V_f^{\neq}$	9.09	0.04	8.77	7.90
	20		$V_r^{\neq}$	14.09	11.74	14.21	13.47

R11	21	$H^{\bullet} + PH_3 \rightarrow {}^{\bullet}PH_2 + H_2$	$V_f^{\neq}$	1.54	3.56	2.81	3.10
	22		$V_r^{\neq}$	24.33	25.99	24.55	23.20
R12	23[e]	$H^{\bullet} + HO^{\bullet} \rightarrow H_2 + O$	$V_f^{\neq}$	10.75	9.69	10.75	10.50
	24		$V_r^{\neq}$	13.48	11.45	13.12	12.87
R13	25	$H^{\bullet} + H_2S \rightarrow H_2 + HS^{\bullet}$	$V_f^{\neq}$	2.50	4.33	3.67	3.50
	26		$V_r^{\neq}$	17.26	18.47	16.98	16.76
R14	27[e]	$O + HCl \rightarrow {}^{\bullet}OH + Cl^{\bullet}$	$V_f^{\neq}$	8.75	7.08	10.32	9.57
	28		$V_r^{\neq}$	9.99	8.07	10.00	9.36
R15	29[e]	${}^{\bullet}NH_2 + {}^{\bullet}CH_3 \rightarrow CH_4 + NH$	$V_f^{\neq}$	9.43	0.37	8.65	8.0
	30		$V_r^{\neq}$	23.00	20.85	21.69	22.40
R16	31[e]	${}^{\bullet}NH_2 + C_2H_5 \rightarrow NH + C_2H_5$	$V_f^{\neq}$	9.86	7.59	9.21	7.50
	32		$V_r^{\neq}$	19.64	17.63	18.70	18.30
R17	33	${}^{\bullet}NH_2 + C_2H_6 \rightarrow NH_3 + {}^{\bullet}C_2H_5$	$V_f^{\neq}$	11.55	10.23	11.20	10.40
	34		$V_r^{\neq}$	17.90	17.23	17.65	17.40
R18	35	${}^{\bullet}NH_2 + CH_4 \rightarrow NH_3 + {}^{\bullet}CH_3$	$V_f^{\neq}$	14.56	12.80	13.88	14.50
	36		$V_r^{\neq}$	17.13	9.36	16.78	17.80
R19	37	s-trans cis-C_5H_8 → s-trans cis-C_5H_8	$V_f^{\neq}$	40.00	38.42	38.78	38.40
	38		$V_r^{\neq}$	40.00	38.42	38.78	38.40
		MUD [g]		0.94	2.04	0.54	
		RMSD [g]		1.13	3.34	0.65	

[a] $V_f^{\neq}$ and $V_r^{\neq}$ are the barrier heights for the forward and reverse reactions respectively, in kcal/mol, excluding zero point energy and spin-orbit corrections.

[b] This work.

[c] M06-2X/aug-cc-pVTZ calculations including Grimme's D3 empirical dispersion energy.

[d] CCSD(T,Full)/aug-cc-pCV(T+d)Z//QCISD/MG3 from ref. [38] and references therein, as collected in https://comp.chem.umn.edu/db/dbs/htbh38.html [39].

[e] The O atom and NH species were calculated in the triplet ground state.

[f] Barrier heights 7, 8, 23, 24, 25 and 26 are also considered in the DBH24/08 dataset.

[g] Duplicated values 10 and 38 (symmetrical reactions) were taken into account only once for the calculation of MUD (mean unsigned deviation) and RMSD (root mean square deviation).

Table 2. Theoretical barrier heights obtained in this paper at the G4, M06-2X/D3/aug-cc-pVTZ and SVECV-f12 levels for the reactions in the NHTBH38/08 dataset, together with "best" values taken from ref. [39].

		Reaction[f,g]	Barrier[a]	G4[b]	M062X[b,c]	SVECV-f12[b]	Best[d]
R20	1	$H^{\bullet} + N_2O \rightarrow {}^{\bullet}OH + N_2$	$V_f^{\neq}$	17.19	17.22	18.86	18.14
	2		$V_r^{\neq}$	82.04	82.28	84.21	83.22
R21	3	$H^{\bullet} + FH \rightarrow HF + H^{\bullet}$	$V_f^{\neq}$	41.47	37.78	41.47	42.18
	4		$V_r^{\neq}$	41.47	37.78	41.47	42.18
R22	5	$H^{\bullet} + ClH \rightarrow HCl + H^{\bullet}$	$V_f^{\neq}$	18.41	18.55	17.73	18.00
	6		$V_r^{\neq}$	18.41	18.55	17.73	18.00
R23	7	$H^{\bullet} + FCH_3 \rightarrow HF + {}^{\bullet}CH_3$	$V_f^{\neq}$	29.19	29.76	30.50	30.40
	8		$V_r^{\neq}$	56.36	53.41	56.92	57.00
R24	9	$H^{\bullet} + F_2 \rightarrow HF + F^{\bullet}$	$V_f^{\neq}$	1.46	3.27	2.16	2.27
	10		$V_r^{\neq}$	102.3	109.0	106.1	105.8
R25	11	${}^{\bullet}CH_3 + FCl \rightarrow CH_3F + Cl^{\bullet}$	$V_f^{\neq}$	0.06	4.53	6.99	6.75
	12		$V_r^{\neq}$	52.63	59.24	59.78	59.16

R26	13	$F^- + CH_3F \rightarrow FCH_3 + F^-$	$V_f^{\neq}$	-0.03	0.01	-0.40	-0.34
	14		$V_r^{\neq}$	-0.03	0.01	-0.40	-0.34
R27	15	$F^-\cdots CH_3F \rightarrow FCH_3\cdots F^-$	$V_f^{\neq}$	11.44	14.99	13.53	13.38
	16		$V_r^{\neq}$	11.44	14.99	13.53	13.38
R28	17	$Cl^- + CH_3Cl \rightarrow ClCH_3 + Cl^-$	$V_f^{\neq}$	3.47	1.86	2.28	3.10
	18		$V_r^{\neq}$	3.47	1.86	2.28	3.10
R29	19	$Cl^-\cdots CH_3Cl \rightarrow ClCH_3\cdots Cl^-$	$V_f^{\neq}$	13.59	13.28	13.29	13.41
	20		$V_r^{\neq}$	13.59	13.28	13.29	13.41
R30	21	$F^- + CH_3Cl \rightarrow FCH_3 + Cl^-$	$V_f^{\neq}$	-8.74	-13.57	-12.31	-12.54
	22		$V_r^{\neq}$	20.86	22.22	19.52	20.11
R31	23	$F^-\cdots CH_3Cl \rightarrow FCH_3\cdots Cl^-$	$V_f^{\neq}$	-2.76	3.55	3.46	3.44
	24		$V_r^{\neq}$	30.12	32.30	29.34	29.42
R32	25	$OH^- + CH_3F \rightarrow HOCH_3 + F^-$	$V_f^{\neq}$	-2.94	-2.51	-2.59	-2.44
	26		$V_r^{\neq}$	18.14	18.18	17.67	17.66
R33	27	$OH^-\cdots CH_3F \rightarrow HOCH_3\cdots F^-$	$V_f^{\neq}$	43.41	50.73	47.59	47.20
	28		$V_r^{\neq}$	7.11	12.23	11.10	10.96
R34	29	$H^\bullet + N_2 \rightarrow HN_2^\bullet$	$V_f^{\neq}$	13.85	13.68	14.29	14.36
	30		$V_r^{\neq}$	9.50	11.22	10.78	10.61
R35	31	$H^\bullet + CO \rightarrow HCO^\bullet$	$V_f^{\neq}$	1.52	3.67	3.10	3.17
	32		$V_r^{\neq}$	20.73	23.10	22.60	22.68

R36	33	$H^\bullet + C_2H_4 \rightarrow {}^\bullet C_2H_5$	$V_f^{\neq}$	-0.41	2.71	1.90	1.72
	34		$V_r^{\neq}$	39.64	43.63	42.08	41.75
R37	35	$CH_3^\bullet + C_2H_4 \rightarrow CH_3CH_2CH_2^\bullet$	$V_f^{\neq}$	6.13	5.94	6.25	6.85
	36		$V_r^{\neq}$	32.27	34.05	32.99	32.97
R38	37	HCN → HNC	$V_f^{\neq}$	48.01	46.01	47.51	48.07
	38		$V_r^{\neq}$	32.83	33.10	32.68	32.82
		MUD		1.85	1.31	0.29	
		RMSD		2.64	1.73	0.39	

[a] $V_f^{\neq}$ and $V_r^{\neq}$ are the barrier heights for the forward and reverse reactions respectively, in kcal/mol, excluding zero point energy.

[b] This work.

[c] M06-2X/aug-cc-pVTZ calculations including Grimme's D3 empirical dispersion energy.

[d] CCSD(T,Full)/aug-cc-pCV(T+d)Z//QCISD/MG3 from ref. [38] and references therein, as collected in https://comp.chem.umn.edu/db/dbs/htbh38.html [39].

[e] Barrier heights 1, 2, 5, 6, 11, 12, 19, 20, 23, 24, 25, 26, 29, 30, 33, 34, 37, and 38 are also considered in the DBH24/08 dataset.

[f] Duplicated values 4, 6, 14, 16, 18 and 20 (symmetrical reactions) were considered only once for the calculation of MUD (mean unsigned deviation) and RMSD (root mean square deviation).

Table 3. Comparison of quality statistics for different methods applied to the reaction height datasets employed in this paper. Values in kcal/mol

Dataset[a]	Metric[b]	This work			Wilson[c]			Truhlar[d]			Curtiss[g]		Karton[h]
		G4	M06-2X	SVECV-f12	ccCA-S4	ccCA-P	G3B	CCSD(T)	M06-2X	G4	G4(B3LYP)	G4(QCISD)	W3lite-F12
DBH24/08	MUD	1.50	1.31	**0.27**				0.46[e]	0.93	0.58			0.14
(22 barriers)	RMSD	2.54	2.10	**0.36**									0.21
	MaxUD	6.69	8.04	**0.99**									
HTBH38/08	MUD	0.94	2.04	**0.54**	0.91	0.92	1.71				0.91	0.95	
(36 barriers)	RMSD	1.13	3.34	**0.65**									
	MaxUD	2.57	8.44	**1.71**	2.88	2.95	5.04				2.39	2.79	
NHTBH38/08	MUD	1.85	1.31	**0.29**	0.98	0.99	2.28	0.96[f]			1.81	0.47	
(32 barriers)	RMSD	2.64	1.73	**0.39**									
	MaxUD	6.69	4.40	**0.99**	3.67	3.75	6.70				8.48	1.93	
All reactions	MUD	1.37	1.69	**0.42**	0.94	0.95	1.98	1.10[f]			1.36	0.71	
(68 barriers)	RMSD	1.99	2.70	**0.54**									
	MaxUD	6.69	8.44	**1.71**									

[a] HTBH38 and NHTBH38 data sets consist each of 19 reactions with forward and reverse barriers, but since some of them are symmetrical, barriers in this case were counted only once; same with the DBH24 data set which is a subset of the combination of the other two; the number of barrier heights employed in each case is written under the name of the set.

[b] MUD = mean unsigned deviation, RMSD = root mean square deviation, MaxUD = maximum unsigned deviation.

[c] Ref. [29], geometries optimized at the B3LYP/6-31G(d) level for the G3B calculations and at the B3LYP/6-31G(2df,p) level for ccCA; ccCA-S4 and ccCA-P differ only in the form of the extrapolation formula used to obtain the CBS MP2 energy.

[d] Refs. [38] and [36], geometries calculated at the QCISD/MG3 level.

[e] CCSD(T)(full)/aug-cc-pCV(T+d)Z calculations.

[f] QCISD(T)/MG3 calculations.

[g] Ref. [40], G4(B3LYP) is the conventional procedure, G4(QCISD) a modified G4 method where geometries are calculated at the QCISD/MG3 level.

[h] Ref. [5], MUD and RMSD quoted as 0.9 kJ mol^{-1} and 0.6 kJ mol^{-1} respectively.

The main results of this paper are summarized in Fig. 1 and Table 3. The correlation between the SVECV-f12 values and the "best" ones (mostly at the W1 and W2h level [35, 36 39], but those in the DBH24 data base updated using the W4 method [37-39]) is very good on a range of energy barriers larger than 100 kcal/mol.

Discussion

General

The first observation in Table 3 is that the MUD of the present method is smaller than the best ones available in the literature using other composite schemes of similar cost. Even in the worst case, the HTBH38/08 dataset, the MUD is almost half of that obtained with the ccCA methods or

Figure 2. MUD, RMSD, maximum positive and negative deviations plotted for the methods used in this wok and for the original G4 data in Curtiss *et al.* [40] paper. All values in kcal/mol.

G4. A very enlightening result is that obtained for the NHTBH38/08 dataset (which then carries over into the result for the combined sets). In this case, the MUD for the new method is about three times smaller than that of ccCA and five times smaller than that of conventional G3B or G4 (using B3LYP for geometry optimization). Substitution of the B3LYP geometries by the QCISD ones, as in the work by Curtiss *et al*. [40] reduces the factor of 5-6 to about 1.5, still with SVECV-f12 as the front runner. This demonstrates –as pointed out by those authors—that most of the error in the G4 results stems from the use of B3LYP to obtain geometries. The better M06-2X DFT method has been adopted in this work for the SVECV-f12 calculations. Concerning the energies obtained with the DFT method, it is noticeable that the MUD is always larger than both the G4 and SVECV-f12 values. However, the geometries are sound, since the calculation of the SVECV-f12 energies on top of those geometries give excellent results.

As was to be expected, since Karton's W3lite-f12 method is very similar to the other Wn methods used to obtain the target results, the error in the barrier obtained with this method is about half that obtained with our SVECT-f12 procedure. Even if that method is dubbed as "lite", it still requires the use, for instance, of quadruple basis sets for the evaluation of the energy at the CCSD-f12 level and the use of the very costly CCSDT(Q) method to include the effect of higher excitations. The reduction in the MUD with respect to the novel procedure is still only 0.17 kcal/mol (less than 1 kJ mol^{-1}), which is probably unimportant for ordinary calculations.

A second observation concerns the point already mentioned in the Introduction about the spread of the calculated barriers. The results are shown graphically in Fig. 2, where data of M06-2X, G4 and SVECV-f12 (present work) as well as the original G4 and modified G4 calculations of Curtiss *et al.* are plotted. It is immediately obvious that the spread of the results at the M06-2X and G4 levels are too big. The effect of tightening the thresholds does produce an observable but small effect in the G4 results, which are more affected nonetheless by the change in the method for the geometry optimization [40]. The spread is then reduced to about one third of the original one. Nonetheless, neither the MUD nor the spread are as good as for the SVECV-f12 procedure. An interesting observation is that all methods cluster their outliers on the negative region, i.e. where the calculated height of the barrier is lower than the best theoretical value. This asymmetry is less marked in the case of the SVECV-f12 method, and even in this case the maximum positive deviation (1.71 kcal/mol) is larger than the maximum negative deviation (-1.34 kcal/mol). Notice that in this case the behavior is opposite to that of the G4 or M062X methods, the largest error is found as an overestimation of some barriers, not underestimation as in the other methods.

A concern of these calculations is the presence of "difficult" cases, i.e. those where the methods misbehave to a larger extent than on average. For G4, it was already mentioned that this is caused mainly for bad geometries at the underlying B3LYP geometry optimization level. Curtiss *et al.* [40] describe this problem with respect to two reactions, (9) and (11) in Table 2. For the first of this reactions, $H^{\bullet} + F_2 \rightarrow HF + F^{\bullet}$, the B3LYP structure exhibits an unrealistically long H···F forming bond in the transition state, 4.412 Å, while the QCISD/MG3 optimization affords a much more reasonable distance of 1.615 Å in fair agreement with the M06-2X value we obtained of 1.721 Å (the F-F bond is 1.480 Å at the QCISD/MG3 level and 1.420 Å at the M06-2X level). The situation is even worse for reaction (11), ${}^{\bullet}CH_3 + FCl \rightarrow CH_3F + Cl^{\bullet}$, where not only the F···C distance in the transition state is much longer in B3LYP than QCISD (3.101 Å vs. 2.064 Å) but also the geometry and bonding patterns are distorted. B3LYP TS is mostly described by a FH interaction, while the QCISD calculation shows a more reasonable structure with a three-fold symmetry axis and the F-C interaction. This is exactly the same structure predicted by M06-2X and even the distance of the bond been formed between F and C, 2.078 Å, is very well in agreement with the QCISD value. Further discussion of geometries will be done later.

On the other side, Grimes *et al.* [29] described also those same structures as problematic, as well as a third one, that corresponding to the ionic pair in reaction (23) which exhibits the same structural problem than the others, due to the drawbacks of B3LYP. However, they also describe four other reactions which are pairwise troublesome, two of them for the G3B method, the other two for the ccCA methods (see their Table IV, problematic reactions). Their values, as well as those obtained in this work, are displayed in Table 4.

Table 4. Problematic reactions for ccCA and G3B cited in Grime's *et al.* work [29]. Values in Kcal/mol.

	Reaction	Best[a]	ccCA-S4[b]	ccCA-P[b]	G3B[b]	G4[c]	M06-2X-G3[c]	SVECV-f12[c]
2	$H_2 + Cl \longrightarrow H + HCl$	7.86	5.82	5.75	4.59	7.1	7.54	6.95
19	$O + CH_4 \longrightarrow OH + CH_3$	13.7	14.37	14.36	8.66	14.09	11.74	14.21
27	$F^{-}\cdots CH_3OH \longrightarrow FCH_3\cdots OH^{-}$	47.2	43.53	43.45	43.67	43.41	50.73	47.59
31	$H + CO \longrightarrow HCO$	3.17	5.82	5.92	9.87	1.52	3.67	3.10
	MUD		2.26	2.32	4.64	1.65	1.58	0.47
	RMSE		2.51	2.57	4.83	2.11	2.04	0.56

[a] Values from ref. [39].

[b] Ref. [29].

[c] This work.

MUD and RMSE values for the four reactions were also added in this table, in order to make a comparison across methods. They show that the problematic cases are largely corrected when larger levels of theory are used. All methods employed in this paper show a MUD much smaller than those from ref. [29]. In particular, the new procedure gives a MUD which is about one third as large as the second best and well in agreement with the overall MUD for the whole set of reactions.

Are "best" values really the best ones?

In the previous analysis it was assumed that the results from ref. [38] are the most accurate ones and used them to assess the errors of the SVECV-f12 values. Nonetheless, since SVECV-f12 includes explicitly correlated calculations, it may be interesting to obtain highly accurate values at that level. For that reason CCSD(T)-F12/pVnZ-F12 calculations (n=D,T,Q) were performed on a subset of reactions from Table 1 and 2. Energy calculations were performed on top of geometries optimized at the same level, so that the whole procedure is consistent. Since this is a costly procedure, especially for geometry optimization of the transition states with the largest cc-pVQZ-F12 basis set, only some of the reactions included in Tables 1 and 2 were processed. The results are collected in Table 5, in comparison with those of ref. [38] and our own SVECV-12 ones.

The conclusion of this study is that if one accepts, as we think, that the CCSD(T)-F12/cc-pVQZ//CCSD(T)-F12/cc-pVQZ calculations are the most accurate feasible ones, then those in ref. [38] exhibit a non-insignificant error. Comparing the three statistics reported in the table, Truhlar's CCSD(T,Full)/aug-cc-pCV(T+d)Z//QCISD/MG3 values exhibit an RMS and MUD more than two times those of SVECV-f12, and a too large MaxUD, outside our reference limit of 1 kcal/mol. As can be seen comparing the values in the table, this maximum error can be tracked down to the reverse barrier in reaction R06 (barrier 12 in Table 1). Other smaller errors (but still larger than 1 kcal/mol) occur for reactions R11 (barrier 22 in Table 1), R14 (barrier 27 in Table 1) and R20 (barrier 2 in Table 2). The bottom line is that if the errors of SVECV-F12 in this subset are expressed with respect to our best calculations instead than to those of ref. [38], the rms is reduced from 0.51 to 0.21 kcal/mol, the mean unsigned deviation from 0.40 to 0.16 kcal/mol and the maximum deviation from 1.35 to 0.46.

Looking to the errors, one can conclude that a SVECV-F12 calculation tends to give similar results to a CCSD(T)-F12/cc-pVTZ-F12//CCSD(T)-F12/cc-pVTZ-F12 one, obviously with much less effort, since the optimization in SVECV-F12 is done at the DFT level and the procedure requires only two single-point calculations at the CCSD(T)-F12 level. These calculations were performed only for a subset of 20 reactions in the original HTBH38 and NHTBH38 databases. This dataset will be called SV20 reaction barriers database (SV20-RBD) and it is available in the Supporting Information section.

Since on average, the SVEVC-F12 seems to perform similarly to CCSD(T)-F12/cc-pVTZ-F12//CCSD(T)-F12/cc-pVTZ-F12 (even if an extrapolation to the CBS limit is included in the former) it may be convenient to analyze the different approximations used in the procedure, to identify systematic errors they might introduce. The SV20-RBD will be utilized to that end and the results are discussed in the next section.

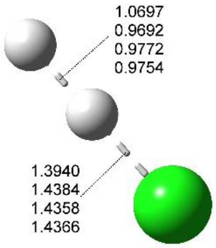
Possible weak points of the method

According to the definition of the procedure given in section **2**, the possible weak point of the procedure are: (i) the level of geometry optimization, (ii) the level used for obtaining the core-valence correlation energy and (iii) the formula employed for the CCSD(T)-f12 extrapolations. Another possible weak point is the method employed to obtain the ZPE, but it is not relevant for this study because Truhlar's reference energies [38, 39] do not include ZPE.

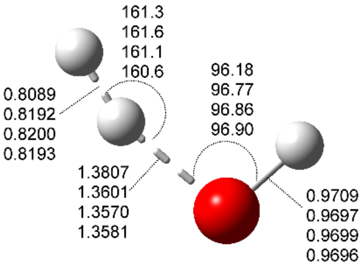
Geometries. The effect of the geometries was explored in two ways. On one hand, the difference between optimum geometries determined at the CCSD(T)-f12/cc-pVnZ-F12 (n=D, T, Q) and those determined at the MO6-2X/3/aug-cc-pVTZ level was analyzed. On the other hand, the CCSD(T)-F12/pVnZ energies obtained at the CCSD(T)-F12/pVnZ optimized geometries were compared to those obtained at the same level of calculation, but with the DFT geometries.

The first set of comparisons is summarized in Fig. 3 where the internal coordinates calculated at the four levels of geometry optimization used in this paper for the reactions in the SV20-RBD are collected. One can compare there the geometries of the transition states and verify that the average difference between geometries is not larger than about 3% in in general, except for R01and R09. A more precise observation can be constructed considering all symmetry-unique bond distances and angles for all the species and performing a percentage root mean square deviation of the optimum CCSD(T)-F121/pVQZ-F12 values. This calculation is included as Table SM2 in the Supporting Information section. Suffice it to say here that % RMSD at the M06-2X/D3/aug-cc-pVTZ, CCSD(T)-F12/cc-pVDZ-F12 and CCSD(T)-F12/cc-pVTZ-F12 levels are 2.3%, 1.7% and 2.1%, with maximum errors of 9.7%, 8.8% and 10.7%.

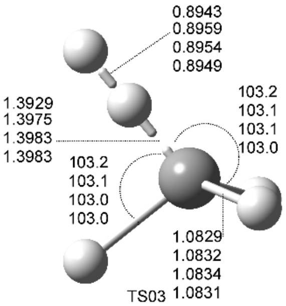
The conclusion is that any of the CCSD(T)-F12 optimized geometries are reasonably near to the best calculated ones, but even that the DFT geometries are reasonably good, with errors of about 0.01 to 0.03 Å in bond lengths and 1-2 degrees in bond angles on average, but with larger errors in some cases. The most important point, however, is not whether geometries in general are precise enough to compare with microwave spectroscopy data, for instance, but how do these approximate geometries influence the total energy of the transition state and the height of the barriers. We have then collected in Table 5 the total energies obtained at the CCSD(T)-F12/cc-pVnZ-F12 levels (n=D, T, Q) using full optimization with each basis set, and the energies obtained when the M06-2X/D3/aug-cc-pVTZ geometries were used instead, calculated the errors in the barriers and reported the errors at each level.



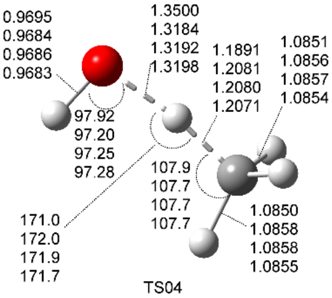
TS01



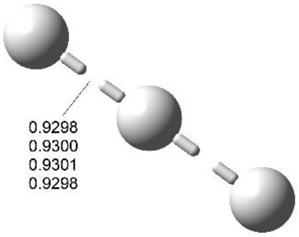
TS02



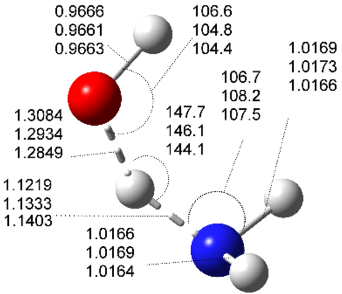
TS03



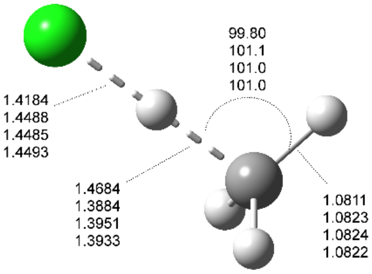
TS04



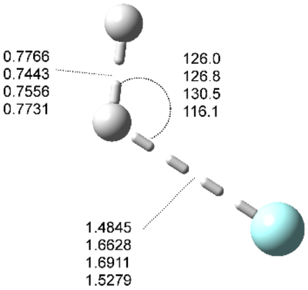
TS05



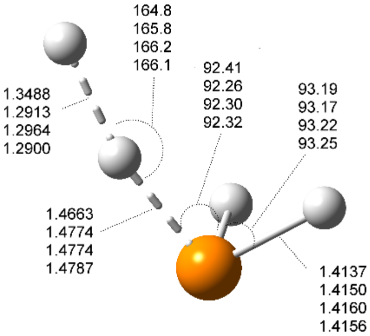
TS06



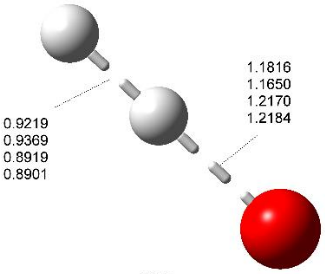
TS07



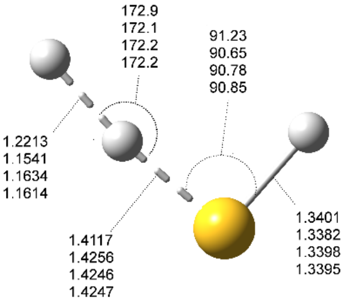
TS09



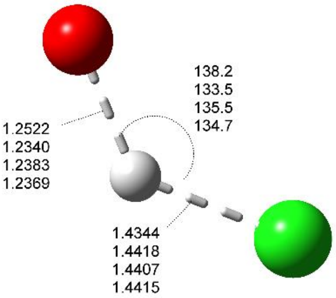
TS11



TS12



TS13



TS14

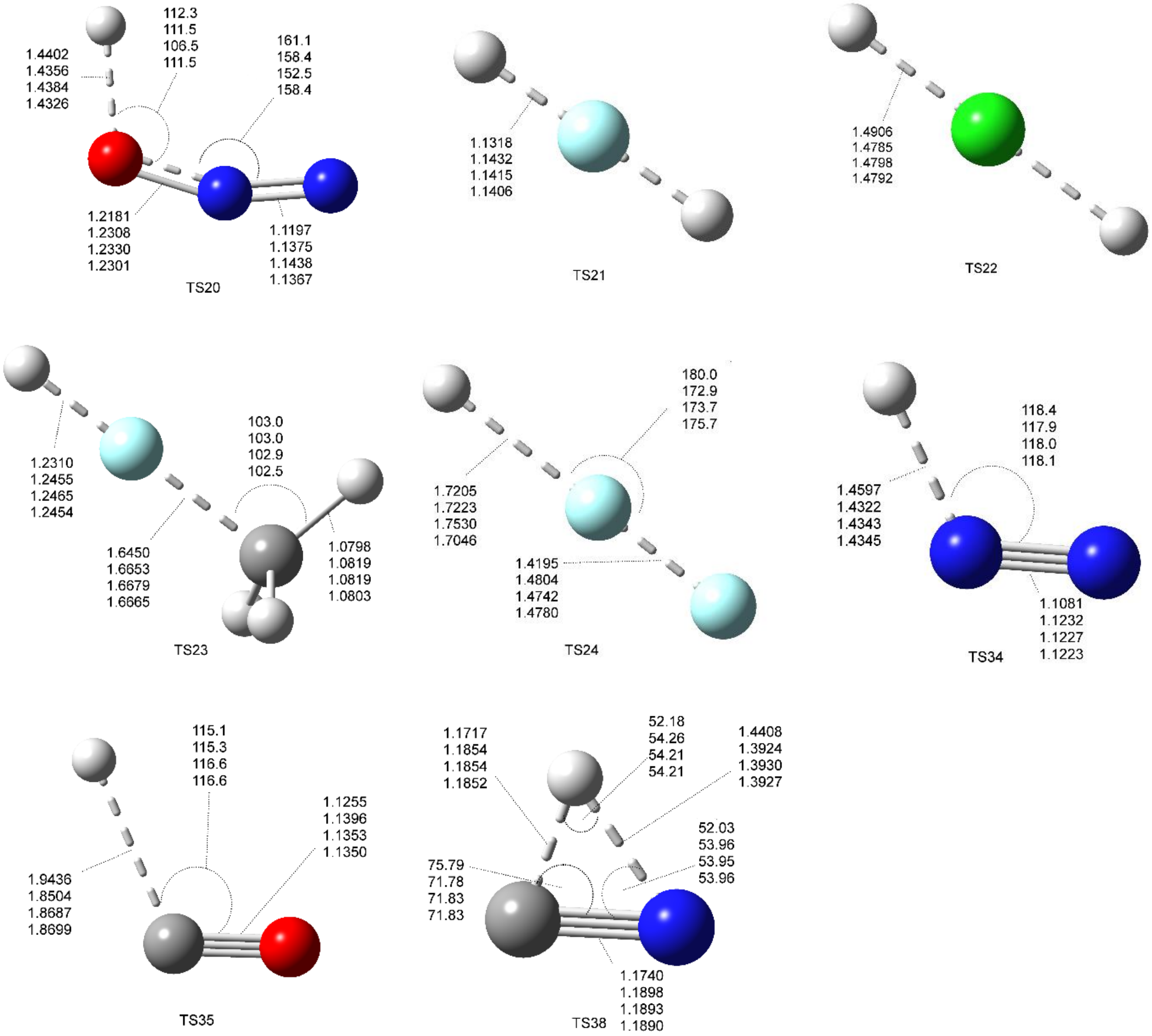


Fig. 3. Structure of the transition states in the SV20-RBD. From top to bottom, distances and angles optimized at the M062X/D3/aug-cc-pVTZ and CCSD(T)-F12 with the thre basis setspVDZ-F12, pVTZ-F12 and pVQZ-F12. Distances in Å and angles in º.

Table 5. Forward and reverse barriers (in kcal/mol) for some of the reactions considered in Tables 1 and 2, calculated at different levels and compared with the best results in reference [38]

				CCSD(T)-F12/cc-pVnZ-F12//CCSD(T)-F12/cc-pVnZ-F12				
		Reaction	**Barrier**	**n=Q**	**n=T**	**n=D**	**Ref.[38]**	**SVECV-F12**
R01	1	$H + HCl \longrightarrow TS1 \longrightarrow H_2 + Cl$	$V_f^\#$	5.36	5.32	5.90	5.70	4.90
	2		$V_r^\#$	7.13	7.19	7.25	7.86	6.95
R02	3	$OH + H_2 \longrightarrow TS2 \longrightarrow H + H_2O$	$V_f^\#$	5.33	5.41	5.53	4.90	5.35
	4		$V_r^\#$	21.70	21.59	21.47	21.20	21.60
R03	5	$CH_3 + H_2 \longrightarrow TS3 \longrightarrow H + CH_4$	$V_f^\#$	11.88	11.87	11.99	12.10	11.83
	6		$V_r^\#$	14.97	14.96	15.11	15.30	14.91
R04	7	$OH + CH_4 \longrightarrow TS4 \longrightarrow CH_3 + H_2O$	$V_f^\#$	6.36	6.42	6.54	6.50	6.31
	8		$V_r^\#$	19.64	19.52	19.35	19.60	19.49
R05	9	$H + H_2 \longrightarrow TS5 \longrightarrow H_2 + H$	$V_f^\#$	9.70	9.72	9.85	9.60	9.67
	10		$V_r^\#$	9.70	9.82	9.85	9.60	9.67
R06	11	$OH + NH_3 \longrightarrow TS6 \longrightarrow H_2O + NH_2$	$V_f^\#$	3.97	4.09	4.13	3.00	3.56
	12		$V_r^\#$	14.30	14.34	14.29	12.70	13.84
R07	13	$HCl + CH_3 \longrightarrow TS7 \longrightarrow Cl + CH_4$	$V_f^\#$	1.80	1.78	2.29	1.70	1.43
	14		$V_r^\#$	6.66	6.73	6.77	7.06	6.55
R09	17	$F + H_2 \longrightarrow TS9 \longrightarrow HF + H$	$V_f^\#$	1.51	1.34	1.58	1.42	1.54
	18		$V_r^\#$	33.71	33.31	33.15	33.40	33.65
R11	21	$H + PH_3 \longrightarrow TS11 \longrightarrow PH_2 + H_2$	$V_f^\#$	2.93	2.92	3.11	3.10	2.81
	22		$V_r^\#$	24.61	24.64	24.74	23.20	24.55
R12	23	$H + HO \longrightarrow TS12 \longrightarrow H_2 + O$	$V_f^\#$	10.94	10.81	10.37	10.50	10.75
	24		$V_r^\#$	13.30	13.42	13.68	12.87	13.12
R13	25	$H + H_2S \longrightarrow TS13 \longrightarrow H_2 + HS$	$V_f^\#$	3.89	3.86	4.18	3.50	3.67
	26		$V_r^\#$	17.03	17.03	16.95	16.76	16.98
R14	27	$O + HCl \longrightarrow TS14 \longrightarrow OH + Cl$	$V_f^\#$	10.61	10.54	11.09	9.57	10.32
	28		$V_r^\#$	10.02	9.79	9.13	9.36	10.00
R20	1	$H + N_2O \longrightarrow TS20 \longrightarrow OH + N_2$	$V_f^\#$	19.09	19.92	19.30	18.14	18.86
	2		$V_r^\#$	84.24	85.24	85.05	83.22	84.21
R21	3	$H + FH \longrightarrow TS21 \longrightarrow HF + H$	$V_f^\#$	41.54	41.61	42.09	42.18	41.47
	4		$V_r^\#$	41.54	41.61	42.09	42.18	41.47
R22	5	$H + ClH \longrightarrow TS22 \longrightarrow HCl + H$	$V_f^\#$	17.77	17.49	18.04	18.00	17.73
	6		$V_r^\#$	17.77	17.49	18.04	18.00	17.73
R23	7	$H + FCH_3 \longrightarrow TS23 \longrightarrow HF + CH_3$	$V_f^\#$	30.57	30.63	30.95	30.38	30.50
	8		$V_r^\#$	56.94	57.02	57.26	57.02	56.92
R24	9	$H + F2 \longrightarrow TS24 \longrightarrow HF + F$	$V_f^\#$	2.37	2.52	2.49	2.27	2.16
	10		$V_r^\#$	105.63	105.73	106.03	105.80	106.08
R34	29	$H + N_2 \longrightarrow TS34 \longrightarrow HN_2$	$V_f^\#$	14.52	14.59	15.00	14.36	14.29
	30		$V_r^\#$	10.83	10.79	10.56	10.61	10.78
R35	31	$H + CO \longrightarrow TS35 \longrightarrow HCO$	$V_f^\#$	3.27	3.32	3.60	3.10	3.17
	32		$V_r^\#$	22.24	22.58	22.66	22.68	22.60
R38	37	$HCN \longrightarrow TS38 \longrightarrow HNC$	$V_f^\#$	47.85	47.77	47.55	48.07	47.51
	38		$V_r^\#$	32.92	32.79	32.47	32.82	32.68
RMS					**0.25**	**0.37**	**0.56**	**0.21***(0.51)*
MUD					**0.14**	**0.32**	**0.42**	**0.16***(0.40)*
MaxUD					**1.00**	**0.89**	**1.60**	**0.46***(1.35)*

The table has three sections. In the first and second ones, the comparison is made between the total energies of the transition states and the barriers obtained with the cc-pVDZ-F12 and cc-pVTZ-12 using either the optimized geometry at each CCSD(T)-F12 level or the M06-2X/D3/aug-cc-pVTZ optimized geometry. The MUD of the energies are just 0.12 and 0.13 kcal/mol respectively, meaning that the choice of the DFT method to perform the geometry optimization is a sensible one. One could think that even a small error in the energy of the TS may lead to a large error in the barriers, but this is not the case. Using the barriers calculated directly at the CCSD(T)-F12/cc-pVnZ level (n = D or T), optimizing also the reactants and products, of course, and comparing them with those obtained using single-point calculations at the M06-2X/D3/aug-cc-pVTZ optimized geometries, the rmsd is 0.43 kcal/mol (n=D) and 0.19 kcal/mol (n=T) respectively. Then, finally, in the third section of the table is a comparison between the barriers obtained at the extrapolated CCSD(T)-F12/CBS level to those obtained at the fully optimized CCSD(T)-F12/cc-pVQZ-F12 level. The energies at the extrapolated level are generally about 0.8 kcal/mol more negative than those at the optimized level and the rms and MaxUD statistics are just 0.20 kcal/mol and 0.47 kcal/mol respectively.

The general conclusion is that there is a small effect in the barriers provoked by the use of a DFT optimized geometry in the single-point calculations instead of the fully optimized geometry at the CCSD(T)-F12 level, but that this error is very small. Since full geometry optimization at a level like CCSD(T)-F12/cc-pVQZ-F12 is possible only for very small molecules, the choice taken of using these DFT optimized values seems reasonable.

Table 6. Comparison between total energies of transition states and the barriers for the reactions in the SV20 dataset, using geometries optimized at every CCSD(T)-F12 level and those at the M06-2X/D3 level

Reaction	Transition State	cc-pVDZ				
		Optimized	M06-2X/D3	Difference		
				Total Energy	Barrier	
					Forward	Reverse
R01	TS01	-460.842896	-460.843404	0.32	0.33	0.20
R02	TS02	-76.822686	-76.822695	0.01	-0.19	0.26
R03	TS03	-40.923939	-40.923935	0.00	-0.10	-0.14
R04	TS04	-116.095678	-116.095793	0.07	-0.10	0.38
R05	TS05	-1.657574	-1.657573	0.00	-0.15	-0.15
R06	TS06	-132.142188	-132.141647	-0.34	-0.49	-0.31
R07	TS07	-500.118391	-500.118753	0.23	-0.25	0.12
R09	TS09	-100.816796	-100.816645	-0.09	-0.16	0.47
R11	TS11	-343.197912	-343.198014	0.06	-0.12	-0.06
R12	TS12	-76.141546	-76.141590	0.03	0.60	-0.35
R13	TS13	-399.441300	-399.441473	0.11	-0.18	0.19
R14	TS14	-535.324719	-535.324711	-0.01	-0.48	0.89
R20	TS20	-184.927217	-184.926755	-0.29	-0.01	-0.81
R21	TS21	-100.801605	-100.801524	-0.05	-0.59	-0.59
R22	TS22	-460.822914	-460.822846	-0.04	-0.70	-0.70
R23	TS23	-140.048222	-140.048241	0.01	-0.36	-0.31
R24	TS24	-199.846636	-199.845763	-0.55	0.07	-0.94
R34	TS34	-109.878459	-109.878286	-0.11	-0.29	0.39

R35	TS35	-113.678130	-113.678239	0.07	-0.13	0.18
R38	TS38	-93.222702	-93.222856	0.10	0.66	0.67
			MUD	0.12	RMSE	0.43
					MaxUD	0.49

Reaction	Transition State	cc-pVTZ				
		Optimized	M06-2X/D3	Difference		
				Total Energy	Barrier	
					Forward	Reverse
R01	TS01	-460.861604	-460.861995	0.25	0.29	0.19
R02	TS02	-76.835705	-76.835728	0.01	-0.06	0.14
R03	TS03	-40.930671	-40.930668	0.00	0.01	0.01
R04	TS04	-116.114461	-116.114577	0.07	0.01	0.21
R05	TS05	-1.658505	-1.658505	0.00	-0.02	-0.02
R06	TS06	-132.163649	-132.164101	0.28	0.18	0.26
R07	TS07	-500.142826	-500.143114	0.18	0.21	0.11
R09	TS09	-100.834155	-100.833784	-0.23	-0.06	0.17
R11	TS11	-343.208467	-343.208550	0.05	0.06	0.03
R12	TS12	-76.153232	-76.153484	0.16	0.29	0.04
R13	TS13	-399.456156	-399.456290	0.08	0.11	0.09
R14	TS14	-535.353942	-535.353965	0.01	0.09	0.24
R20	TS20	-184.952912	-184.953761	0.53	0.16	-0.19
R21	TS21	-100.819985	-100.819927	-0.04	-0.10	-0.10
R22	TS22	-460.841575	-460.841521	-0.03	-0.15	-0.15
R23	TS23	-140.072056	-140.072074	0.01	-0.04	-0.07
R24	TS24	-199.880988	-199.880223	-0.48	0.18	-0.58

R34	TS34	-109.892987	-109.892804	-0.11	0.09	0.14
R35	TS35	-113.694427	-113.694457	0.02	0.09	-0.22
R38	TS38	-93.234682	-93.234853	0.11	0.43	0.35
			MUD	0.13	RMSE	0.19
					MaxUD	0.58

Reaction	Transition State	cc-pVQZ Optimized	extrap. CBS	Energy Difference		
				Total Energy	Barrier	
					Forward	Reverse
R01	TS01	-460.866644	-460.868538	-1.19	0.47	0.19
R02	TS02	-76.839680	-76.840315	-0.40	-0.01	0.10
R03	TS03	-40.932670	-40.933037	-0.23	0.05	0.07
R04	TS04	-116.120093	-116.121187	-0.69	0.05	-0.23
R05	TS05	-1.658871	-1.658833	0.02	0.03	0.03
R06	TS06	-132.170156	-132.172003	-1.16	0.41	0.46
R07	TS07	-500.149570	-500.151687	-1.33	0.35	0.18
R09	TS09	-100.838630	-100.839815	-0.74	-0.03	0.06
R11	TS11	-343.210774	-343.212258	-0.93	0.13	0.07
R12	TS12	-76.156619	-76.157670	-0.66	0.19	0.19
R13	TS13	-399.459524	-399.461504	-1.24	0.21	0.05
R14	TS14	-535.361792	-535.364261	-1.55	0.30	0.02
R20	TS20	-184.961970	-184.963265	-0.81	0.23	0.03
R21	TS21	-100.825192	-100.826403	-0.76	0.07	0.07
R22	TS22	-460.846880	-460.848092	-0.76	0.04	0.04
R23	TS23	-140.078951	-140.080461	-0.95	0.08	0.02

R24	TS24	-199.890642	-199.8923504	-1.07	0.22	-0.45
R34	TS34	-109.897057	-109.8979131	-0.54	0.23	0.05
R35	TS35	-113.699196	-113.7001639	-0.61	0.17	0.07
R38	TS38	-93.238013	-93.23907487	-0.67	0.34	0.24
			MUD	0.82	RMSE	0.20
					MaxUD	0.47

Core-valence correlation energy.

In this new procedure, likewise what it is done in other composite methods in the literature, the core-valence correlation energy contribution was calculated at the MP2 level using the cc-pCVTZ basis sets. Two questions can be posed. On one hand, is MP2 reasonable enough or CCSD(T) would be more accurate? And on the other hand, is the basis set accurate enough or a larger one should be used?

A limited attempt at calculating the influence of this factors was done estimating the degree of accuracy of MP2/cc-pCVTZ calculations with respect to CCSD(T)/cc-pCVTZ and CCSD(T)/aug-cc-pCV5Z for the reaction energies, and forward and reverse barriers on the six reactions R01, R02, R03, R04, R06 and R07 (R05 was excluded because all species comprise only hydrogens). The results are shown in Table 7.

Table 7. Core-valence correlation energy contribution to reaction energies and barriers at the MP2 and CCSD(T) levels wit the cc-pCVTZ and aug-cc-pV5Z basis sets. Values in kcal/mol

	Core-Valence Contribution								
	Reaction energy			Forward Barrier			Reverse Barrier		
	cc-pCVTZ		aug-cc-pCV5Z	cc-pCVTZ		aug-cc-pCV5Z	cc-pCVTZ		aug-cc-pCV5Z
Reactions	MP2	CCSD(T)	CCSD(T)	MP2	CCSD(T)	CCSD(T)	MP2	CCSD(T)	CCSD(T)
Reaction 1: H + HCl --> H2 + Cl	-0.43	-0.29	-0.27	0.01	0.00	0.05	-0.42	-0.29	-0.22
Reaction 2: OH + H2 --> H + H2O	-0.26	-0.25	-0.27	-0.05	-0.04	-0.03	0.21	0.21	0.24
Reaction 3: CH3 + H2 --> H + CH4	-0.18	-0.20	-0.21	-0.06	-0.06	-0.04	0.13	0.14	0.17
Reaction 4: OH + CH4 --> CH3 + H2O	-0.07	-0.05	-0.06	-0.01	0.00	-0.01	0.06	0.05	0.07
Reaction 6: OH + NH3 --> H2O + NH2	0.06	0.06	0.08	0.00	0.01	0.04	-0.05	-0.05	-0.03
Reaction 7: HCl + CH3 --> Cl + CH4	0.24	0.09	0.06	-0.06	-0.06	-0.05	-0.31	-0.14	-0.12

Contributions to reaction energies are small in most cases (always below 0.5 kcal/mol). In general, the calculation at the MP2 and CCSD(T) levels using the cc-pVCTZ basis set are comparable. In some cases, like reactions R01 and R07, the MP2 correction is about double the size of the CCSD(T) contribution. In this case, the difference seems to stem from only one of the hemireactions. In all cases, the contributions at the CCSD(T) level using the smaller or the larger basis sets are the same. In conclusion, the use of a more sophisticated and expensive method to calculate the core-valence correlation seems unnecessary.

Extrapolation. Since the purpose is to have a very simple procedure which can be routinely applied, a fast extrapolation scheme which does not extrapolate separately the Hartree-Fock and correlation energy components was used. Admittedly, this can introduce an error which has not been investigated further in this work. The papers of Helgaker *et al.* [52], Hill *et al.* [53] and Varandas [51, 54], among many others cited there, can be explored for other extrapolation options. It is however worthwhile to notice (see Table 6) that the MUD of the energy difference between the CCSD(T)-F12/CBS//M06-2X/D3/aug-cc-pVTZ and CCSD(T)-F12/cc-pVQZ-F12//CCSD(T)-F12/cc-pVQZ-F12 calculations is only 0.82 kcal/mol, with the former always lower than the latter except in the case of the all-hydrogens reaction R05. It seems then that the suggested extrapolation scheme is a reliable simple procedure to obtain the energies to be used for the calculation of the barriers (which rmse of only 0.20 kcal/mol is well within the stated goal).

Timing

A final word must be said about the timing of the methods. Truhlar estimated the relative cost of each method as the single-processor CPU time for calculating an energy gradient of the molecule phosphinomethanol divided by the time for an MP2/6-31+G(d,p) energy gradient calculation, using the same computer program on the same computer. In this paper, a relative estimate of the cost of the procedure was obtained measuring the CPU time necessary for the SVECV-f12 complete calculation of all the species included in a given reaction (reactants, products and transition states), calculating the ratio for the same process at the G4 level, and then averaging these ratios for the 38 reactions. Results are shown in Table 8.

Table 2. Total CPU time per processor at the M06-2X/D3/aug-cc-pVTZ, G4 and SVECV-f12 levels and difference between the M006 and SV methods (in seconds). The ratio between the SV and G4 times for each reaction is given in the last column. Maximum, minimum and rms of the ratio are given in the last columns, as well as the CPU-time (in hours) for the whole set

Reaction		M06-2X/D3	G4	SVECV-f12	SV-M06	Ratio SV/G4
R01	$H + HCl \longrightarrow H_2 + Cl$	39	48	60	21	1.3
R02	$OH + H_2 \longrightarrow H + H2O$	695	165	824	129	5.0
R03	$CH_3 + H_2 \longrightarrow H + CH_4$	1876	1369	2321	444	1.7
R04	$OH + CH_4 \longrightarrow CH_3 + H_2O$	5909	1523	6828	919	4.5
R05	$H + H_2 \longrightarrow H_2 + H$	27	44	45	18	1.0
R06	$OH + NH_3 \longrightarrow H_2O + NH_2$	680	693	1086	406	1.6
R07	$HCl + CH_3 \longrightarrow Cl + CH_4$	9863	1013	10545	682	10.4
R08	$OH + C_2H_6 \longrightarrow H_2O + C_2H_5$	6044	4710	22089	16045	4.7
R09	$F + H_2 \longrightarrow HF + H$	221	105	316	94	3.0
R10	$O + CH_4 \longrightarrow OH + CH_3$	689	723	1119	431	1.5
R11	$H + PH_3 \longrightarrow PH_2 + H_2$	3807	666	4107	299	6.2
R12	$H + HO \longrightarrow H_2 + O$	454	195	533	79	2.7
R13	$H + H_2S \longrightarrow H2 + HS$	418	272	557	139	2.0
R14	$O + HCl \longrightarrow OH + Cl$	399	281	547	149	2.0
R15	$NH_2 + CH_3 \longrightarrow CH_4 + NH$	7533	4504	7803	269	1.7
R16	$NH_2 + C_2H_5 \longrightarrow C_2H_6 + NH$	40503	6776	47685	7182	7.0
R17	$C_2H_6 + NH_2 \longrightarrow NH_3 + C_2H_5$	15467	4883	42109	26642	8.6
R18	$NH_2 + CH_4 \longrightarrow CH_3 + NH_3$	22708	832	23703	995	28.5

R19	s-trans cis-C_5H_8 ⟶ same	78873	8061	113512	34639	14.1
R20	H + N_2O ⟶ OH + N_2	2600	1570	3069	470	2.0
R21	H + FH ⟶ HF + H	138	107	229	91	2.1
R22	H + ClH ⟶ HCl + H	116	98	170	53	1.7
R23	H + FCH_3 ⟶ HF + CH_3	660	858	1190	530	1.4
R24	F + H2 ⟶ HF + H	551	580	742	191	1.3
R25	CH_3 + FCl ⟶ CH_3F + Cl	2777	3241	4404	1628	1.4
R26	F(-) + CH3F ⟶ FCH3 + F(-)	2591	496	5855	3264	11.8
R27	F(-)···CH_3F ⟶ FCH_3···F(-)	3594	1482	11086	7491	7.5
R28	Cl(-) + CH_3Cl ⟶ $ClCH_3$ + Cl(-)	1085	1066	2780	1695	2.6
R29	Cl(-)···CH_3Cl ⟶ CH_3Cl···Cl(-)	2556	496	5855	3299	11.8
R30	F- + CH_3Cl ⟶ FCH_3 + Cl-	4518	3226	6142	1624	1.9
R31	F(-)···CH_3Cl ⟶ FCH_3···Cl(-)	5906	3529	10695	4789	3.0
R32	OH- + CH_3F ⟶ $HOCH_3$ + F-	5919	1314	9145	3226	7.0
R33	OH(-)···CH_3F ⟶ $HOCH_3$…F(-)	9187	2822	18567	9380	6.6
R34	H + N_2 ⟶ HN_2	629	353	833	204	2.4
R35	H + CO ⟶ HCO	207	227	401	194	1.8
R36	H + C_2H_4 ⟶ C_2H_5	3754	3887	7874	4119	2.0
R37	CH_3 + C_2H_4 ⟶ $CH_3CH_2CH_2$	6924	4678	19328	12404	4.1
R38	HCN ⟶ HNC	320	279	866	546	3.1
	TOTAL TIME (hours)	70	19	110	40	
	Maximum					28.5
	Minimum					1.0
	RMSE					7.0

The process is of course not exact, because in a heterogeneous cluster it will depend on the nodes being active for a given calculation, the load of the clusters, the size of the molecules (especially the transition states), the difficulty in the calculation of the TSs, etc. At any rate, a rule of thumb says that the SVECV-f12 procedure will be on average about 7 times more costly than a G4 calculation (according to the rmse) with some of the calculations as fast as G4 (ratio SV/G4 = 1) and others much slower (ratio SV/G4 = 28.5). For the whole set of reactions, the theoretical CPU time (discounting specific problems of some of the calculations that led to reruns, etc) for the SVEVC-f12 procedure was 110 hours, about 6 times larger than the G4 time (19 hours). Most of that time (appx. 70 hours) was spent in the optimization procedure. It must be noted also that to increase the precision of G4 calculations, the QCISD method was suggested instead of B3LYP, meaning that the time for running the set would concomitantly increase.

Notwithstanding what was said in the previous paragraph, the calculations are in general much faster than CCSD(T)/CBS, especially if geometries are calculated at some CCSD(T) level, because CBS extrapolation for CCSD(T) requires at least a quadruple-zeta basis set. It is also noted that the present SVECV-f12 procedure is faster than Wn procedures. Therefore, both in terms of accuracy (much better than all 348 chemical models in Zheng's paper [38]) and speed (faster than the best CCSD(T)/CBS and Wn methods) this method seems to be a reasonable compromise for the accurate study of barrier heights of chemical reactions involving middle size species.

Conclusions

A new simple composite scheme, SVRCV-f12, based on a combination of density functional geometry and frequencies evaluation, inclusion of explicit correlation using the CCSD(T)-f12 method extrapolated to the complete basis set limit, and core-valence correlation corrections employing the MP2 method has been developed and tested on Truhlar's HTBH38/08, NHTBH38/08, and DBH24/08 databases of 68 forward and reverse reaction barriers corresponding to 38 different reactions. The results have been compared with G4 and M06-2X results and with data present in the literature employing other procedures.

The SVRCV-f12 calculated values correlate very well with the accurate and expensive W1 and W4 barriers for the studied reactions. A correlation coefficient larger than 0.999 reveals an almost perfect straight line for barriers spanning a range larger than 100 kcal/mol. MUD values obtained for the datasets were in all cases smaller than those obtained with other composite methods of comparable cost (G3B, G4, ccCA) and better than those obtained with very expensive CCSD(T)/CBS calculations. In the case of the landmark DBH24/08 smaller database, a MUD of 0.31 kcal/mol (i.e. about 1.3kJ/mol) was obtained, to be compared to a 0.46 kcal/mol MUD value obtained at the CCSD(T,Full)/aug-cc-pCV(T+d)Z level, the best one obtained by Zheng *et al.* [38] after testing 348 different chemical models on the set. The MUD for the whole set of 68 barriers was 0.43 Kcal/mol (about 2kJ/mol) and all the values lied in the interval [-1.34, 1.71] Kcal/mol.

A comparison of the values obtained by Zheng *et al.* [38] with our SV20 dataset of very accurate CCSD(T)-F12/cc-pVQZ-F12//CCSD(T)/cc-pVQZ-F12 reaction barriers showed them to exhibit a non-insignificant deviation, with an rmse of 0.56 kcal/mol and several errors larger than 1 kcal/mol. Thus, when the SVECV-F12 results were compared to those more accurate values, much smaller rmse (0.21 kcal/mol instead of 0.51 kcal/mol) and MUD (0.16 kcal/mol instead of 0.40 kcal/mol) were found. Most importantly, none of the values was farther away than 0.5 kcal/mol from the best counterpart.

The proposed composite scheme can therefore be a useful tool for the calculation of rate coefficients, since the barrier heights are very accurate. Further application of this procedure to larger and more complex transition states and reaction barriers is processed and will be reported elsewhere.

Acknowledgments

Sustained financial support to our research by ANII, CSIC and Pedeciba in Uruguay are gratefully acknowledged. Part of the ideas developed in this document were conceived during stays at the Forschungszentrum Jülich GmbH (Germany) while holding an Alexander von Humboldt-Stiftung fellowship, and at the Scuola Normale Superiore in Pisa. Suggestions by Dr. Luck Verecken (Jülich) and Dr. Zoi Salta (Pisa) are gratefully acknowledged. The current state of the procedure was developed during a stay at Tropos.de, Leipzig (Germany) hosted by Prof. Dr. Hartmut Herrmann under the sponsorship of the Alexander von Humboldt-Stiftung. Part of the calculations presented in this paper were carried on at the Cluster.uy supercomputer center of Uruguay.

References

[1] L. Vereecken, D. R. Glowaki, M. J. Pilling, Theoretical chemical kinetics in tropospheric chemistry: methodologies and applications, Chem. Rev. 115 (2015) 4063.

[2] J. M. L. Martin, Computational thermochemistry: A brief overview of quantum mechanical approaches, Annu. Rep. Comput. Chem. 1 (2005) 31.

[3] R. J. Bartlett, M. Musial, Coupled-cluster theory in quantum chemistry, Rev. Mod. Phys. 79 (2007) 291.

[4] W. Klopper, R. A. Bachorz, C. Hättig, D. P. Tew, Accurate computational thermochemistry from explicitly correlated coupled-cluster theory, Theor. Chem. Acc. 126 (2010) 289.

[5] A. Karton, Highly Accurate CCSDT(Q)/CBS Reaction Barrier Heights for a Diverse Set of Transition Structures: Basis Set Convergence and Cost-Effective Approaches for Estimating Post-CCSD(T) Contributions, J. Phys. Chem. A 123 (2019) 6720.

[6] G. A. Petersson, T. G. Tensfeldt, J. A. Montgomery, A complete basis set model chemistry. III. The complete basis set-quadratic configuration interaction family of methods, J. Chem. Phys. 94 (1991) 6091.

[7] J. A. Montgomery, J. W. Ochterski, G. A. Petersson, A complete basis set model chemistry. IV. An improved atomic pair natural orbital method, J. Chem. Phys. 101 (1994) 101.

[8] J. W. Ochterski, G. A. Petersson, J. A. Montgomery, A complete basis set model chemistry. V. Extensions to six or more heavy atoms, J. Chem. Phys. 104 (1996) 2598.

[9] J. A. Montgomery, M. J. Frisch, J. W. Ochterski, G. A. Petersson, A complete basis set model chemistry. VI. Use of density functional geometries and frequencies, J. Chem. Phys. 110 (1999) 2822.

[10] J. A. Montgomery, M. J. Frisch, J. W. Ochterski, G. A. Petersson, A complete basis set model chemistry. VII. Use of the minimum population localization method, J. Chem. Phys., 112 (2000) 6532.

[11] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, L. A. Curtiss, Gaussian-1 theory: A general procedure for prediction of molecular energies, J. Chem. Phys. 90 (1989) 5622.

[12] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, J. A. Pople, Gaussian-3 (G3) theory for molecules containing first and second-row atoms, J. Chem. Phys. 109 (1998) 7764.

[13] L. A. Curtiss, P. C. Redfern, K. Raghavachari, J. A. Pople, Gaussian-3X (G3X) theory: Use of improved geometries, zero-point energies, and Hartree–Fock basis setsJ. Chem. Phys. 114 (2001) 108.

[14] L. A. Curtiss, P. C. Redfern, K. Raghavachari, J. Chem. Phys. 126 (2007) 084108.

[15] L. A. Curtiss, K. Raghavachari, P. C. Redfern, A. G. Baboul, J. A. Pople, Gaussian-4 theory, Chem. Phys. Lett. 314 (1999) 101.

[16] L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, Gaussian-3 theory using scaled energies, J. Chem. Phys. 112 (2000) 1125.

[17] L. A. Curtiss, P. C. Redfern, K. Raghavachari, J. A. Pople, Gaussian-3 theory: a variation based on third-order perturbation theory and an assessment of the contribution of core-related correlation, Chem. Phys. Lett. 313 (1999) 600.

[18] L. A. Curtiss, P. C. Redfern, K. Raghavachari, Gaussian-4 theory using reduced order perturbation theory, J. Chem. Phys. 127 (2007) 124105.

[19] L. A. Curtiss, P. C. Redfern, K. Raghavachari, V. Rassolov, J. A. Pople, Gaussian-3 theory using reduced Mo/ller-Plesset order, J. Chem. Phys. 110 (1999) 4703.

[20] J. M. L. Martin, G. de Oliveira, Towards standard methods for benchmark quality ab initio thermochemistry—W1 and W2 theory, J. Chem. Phys. 111 (1999) 1843.

[21] A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, J. Gauss, W3 theory: Robust computational thermochemistry in the kJ/mol accuracy range, J. Chem. Phys. 120 (2004) 4129.

[22] A. Karton, E. Rabinovich, J. M. L. Martin, B. Ruscic, W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions, J. Chem. Phys. 125 (2006) 144108.

[23] A. Karton, J. M. L. Martin, Explicitly correlated Wn theory: W1-F12 and W2-F12, J. Chem. Phys. 136 (2012) 124114.

[24] N. J. DeYonker, T. R. Cundari, A. K. Wilson, The correlation consistent composite approach (ccCA): An alternative to the Gaussian- methods, J. Chem. Phys. 124 (2006) 114104.

[25] N. J. DeYonker, T. V. Grimes, S. Yockel, A. Dinescu, B. Mintz, T. R. Cundari, A. K. Wilson, J. Chem. Phys. 125 (2006) 104111.

[26] D. S. Ho, N. J. DeYonker, A. K. Wilson, T. R. Cundari, The correlation-consistent composite approach: Application to the G3/99 test set, J. Phys. Chem. A 110 (2006) 31.

[27] N. J. DeYonker, D. S. Ho, A. K. Wilson, and T. R. Cundari, Computational s-Block Thermochemistry with the Correlation Consistent Composite Approach, J. Phys. Chem. A 111 (2007) 10776.

[28] N. J. DeYonker, K. A. Peterson, G. Steyl, A. K. Wilson, T. R. Cundari, Quantitative computational thermochemistry of transition metal species, J. Phys. Chem. A 111 (2007) 11269.

[29] T. V. Grimes, A. K. Wilson, N. J. DeYonker, T. R. Cundari, Performance of the correlation consistent composite approach for transition states: A comparison to G3B theory, J. Chem. Phys. 127 (2007) 154117.

[30] J. A. Miller, S. J. Klippenstein, From the Multiple-Well Master Equation to Phenomenological Rate Coefficients: Reactions on a C3H4 Potential Energy Surface, J. Phys. Chem. A 107 (2003) 2680.

[31] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, J. F. Stanton, HEAT: High accuracy extrapolated ab initio thermochemistry, J. Chem. Phys. 121 (2004) 11599.

[32] Y. J. Bomble, J. Vázquez, M. Kállay, C. Michauk, P. G. Szalay, A. G. Császár, J. Gauss, J. F. Stanton, High-accuracy extrapolated ab initio thermochemistry. II. Minor improvements to the protocol and a vital simplification, J. Chem. Phys. 125 (2006) 064108.

[33] A. East, C. Johnson, W. Allen, Characterization of the X̃ 1A' state of isocyanic acid, J. Chem. Phys. 98 (1993) 1299.

[34] M. S. Schuurman, S. R. Muir, W. D. Allen, H. F. Schaefer, Toward subchemical accuracy in computational thermochemistry: Focal point analysis of the heat of formation of NCO and [H, N, C, O] isomers, J. Chem. Phys. 120 (2004) 11586.

[35] Y. Zhao, B. J. Lynch, D. G. Truhlar, Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics, Phys. Chem. Chem. Phys. 7 (2005) 43.

[36] Y. Zhao, N. González-García, D. G.Truhlar, Benchmark Database of Barrier Heights for Heavy Atom Transfer, Nucleophilic Substitution, Association, and Unimolecular Reactions and Its Use to Test Theoretical Methods, J. Phys. Chem. A 109 (2005) 2012.

[37] A. Karton, A. Tarnopolsky, J.-F. Lamère, G. C. Schatz, J. M. L. Martin, Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional

and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics, J. Phys. Chem. A 112 (2008) 12868.

[38] J. Zheng, Y. Zhao, D. G. Truhlar, The DBH24/08 database and its use to assess electronic structure model chemistries for chemical reaction barrier heights, J. Chem. Theory Comput. 5 (2009) 808.

[39] The Minnesota databases 2.0, University of Minnesota, https://comp.chem.umn.edu/db/index.html. Last accessed 1/30/2021.

[40] L. A. Curtiss, P. C. Redfern, K. Raghavachari, Assessment of Gaussian-4 theory for energy barriers, Chem. Phys. Lett. 499 (2010) 168.

[41] Y. Zhao, D. G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, Theor. Chem. Acc. 120 (2008) 215.

[42] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.

[43] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, T. L. Windus, New basis set exchange: An open, up-to-date resource for the molecular sciences community, J. Chem. Inf. Model. 59 (2019) 4814. Basis Set Exchange ver2, https://www.basissetexchange.org/. Last accessed 1/27/2021.

[44] T. B. Adler, G. Knizia, H.-J. Werner, A simple and efficient CCSD (T)-F12 approximation, J. Chem. Phys. 127 (2007) 221106.

[45] G. Knizia, T. B. Adler, H.-J. Werner, Simplified CCSD (T)-F12 methods: Theory and benchmarks, J. Chem. Phys. 130 (2009) 054104.

[46] K. A. Peterson, T. B. Adler, H.-J. Werner, Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B–Ne, and Al–Ar, J. Chem. Phys. 128 (2008) 084102.

[47] J. M. L. Martin, Ab initio total atomization energies of small molecules—towards the basis set limit, Chem. Phys. Lett. 259 (1996) 669.

[48] Gaussian 16, Revision C.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2019.

[49] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, Molpro: a general-purpose quantum chemistry program package, WIREs Comput. Mol. Sci. 2 (2012) 242.

[50] MOLPRO, version 2019.2, a package of ab initio programs, H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, S. J. Bennie, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, S. J. R. Lee, Y. Liu, A. W. Lloyd, Q. Ma, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, T. F. Miller III, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, M. Welborn, , see https://www.molpro.net.

[51] A. J. C. Varandas, Canonical versus explicitly correlated coupled cluster: Post-complete-basis-set extrapolation and the quest of the complete-basis-set limit, Int. J. Quantum Chem., 2020, https://doi.org/10.1002/qua.26598.

[52] T. Helgaker, W. Klopper, and D. P. Tew, Quantitative quantum chemistry, Mol. Phys. 106, 2107, 2008.

[53] J. G. Hill, K. A. Peterson, G. Knizia, H.-J. Werner, Extrapolating MP2 and CCSD explicitly correlated correlation energies to the complete basis set limit with first and second row correlation consistent basis sets, J. Chem. Phys. 131, 194105, 2009.

[54] A. J. C. Varandas, Extrapolating to the one-electron basis-set limit in electronic structure calculations, J. Chem. Phys. 126, 244105, 2007.