

Atomic pseudo-potentials for sp^2 carbon atoms

Alexander Punter, Paola Nava, and Yannick Carissan*

Aix Marseille Univ, CNRS, Centrale Marseille, iSm2, Marseille, France

E-mail: Yannick.Carissan@univ-amu.fr

Phone: +33 (0)491289168

Abstract

A pseudo-potential system for recreating an sp^2 carbon atom is built and tested as a building block for various pseudo-hydrocarbon chain and ring systems. This pseudo-system has a central charge of 1, thus it contains only one electron. It is employed in *ab-initio* calculations in which several physical characteristics including 1st ionisation and excitation energies, as well as the HOMO energy, are found to be well-reproduced by the pseudo-system.

1 Introduction

It is a common idea that a chemical system can be thought of as comprised of (at least) two parts: an active one, where most of the chemistry takes place, and an (apparently) inactive one, which must be taken into account in order to fulfill chemical requirements. Based on this general statement, many successful theoretical approaches have been developed. Among them can be cited QM/QM' or QM/MM methods (i), where the active region that usually contains several atoms is treated at a high level of calculation, while the inactive part is

treated at a lower level of theory.¹ Fragmentation methods split the complete problem into smaller parts and combine individual calculations on these fragments to recover the properties of the overall system.^{2,3} Frozen density embedding techniques (ii) replace part of a molecule, or its surroundings, by a frozen electronic density extracted on a reference system.⁴ They are an extension of the frozen core approximation, already used to reduce the number of parameters to be optimised in the self consistent field calculation. In the framework of this work, the pioneering work of Rivail *et al.* must be mentioned: the total wave function is optimized with some molecular orbitals kept frozen.⁵

In this article we shall focus on pseudo-potential techniques, which also divide the chemical problem into two parts. Effective core potentials (iii) are commonly employed for atoms: core electrons (and effects due to the corresponding nuclear charge) are replaced by an operator fast to evaluate and the active electrons are treated explicitly.⁶ In the same vein, model potentials (iv) replace a frozen electronic density (computed from the atom) by some operators, the number of which depends on the level of refinement required.⁷

The effective group potentials (v) bridge the frozen density and the core potential approaches: using core potential extraction techniques, effects due to the implicit electron density (and corresponding nuclear charge) are reproduced by a mono-electronic operator.^{8,9} However, while effective core potentials and model potentials intend to reproduce atomic properties, effective group potentials aim at mimicking the effect of atoms involved in one or more chemical bonds. The cyclopentadienyl group has, for instance, been successfully extracted.⁸ In this example, the active part consisted of six electrons (the π electrons) and six nuclear charges, with the pseudo-potential replacing all the rest, including the hydrogen atoms. Even if the σ/π separability suggested how to define the active and the inactive parts, carbon atoms are hybridized and the core/valence distinction is not strictly defined.

The theoretical background to extract effective group pseudo-potentials can be traced back to several contributions in the 80's and 90's.^{7,10-14} In 1992 Katsuki built molecular potentials based on Huzinaga's model potentials.^{15,16} In his work on this subject, Huzinaga

emphasises that pseudo-potentials should maintain three effects of the 'dormant' electrons on the active ones: the Coulomb, the exchange and the 'no-collapse' term.¹¹ Following his proposal, for an atom where active and dormant electrons have been separated, the Hamiltonian reads:

$$\hat{H} = \sum_{i=1}^n \hat{h}(i) + \sum_{i<j} \frac{1}{r_{ij}} \quad (1)$$

with $\frac{1}{r_{ij}}$ the bi-electronic interaction between explicitly treated active electrons and the mono-electronic operator:

$$\hat{h}(i) = -\frac{1}{2}\Delta_i - \frac{(Z - Z_c)}{r_i} + \hat{V}(i) + \hat{\sigma}(i) \quad (2)$$

where Δ_i is the Laplacian of the coordinates of electron i , and Z_c is the number of core electrons withdrawn from the reference system. The operator $\hat{\sigma}(i)$ is the 'no-collapse' term that prevents active electrons collapsing into the dormant region. The operator $\hat{V}(i)$ reproduces the Coulomb and exchange interactions, for which Hunzinger proposed a spectral formulation:

$$\hat{V} = r^{-n} \left[\sum_I A_I \exp(-\alpha_I r^2) + \sum_J B_J r \exp(-\beta_J r^2) \right] \text{ with } n = 1 \quad (3)$$

To take into account the fact that dormant electrons are removed from the system, the nuclear charge is modified by the Z_c value. Yet, as the effective charge felt by the active electrons is likely not to be an integer, the modification of the value of the nuclear charge is scaled in \hat{V} . In this operator we notice that the r^{-1} behavior is maintained for the first term (the Coulomb one).

In the present work, the aim is not to obtain pseudo-potentials for specific chemical groups, as in the case of the cyclopentadienyl anion cited hereabove. Instead, we want to propose a methodology for extracting a potential for a hybridised carbon atom (here an sp^2 carbon atom), which can then be used as a building block for constructing chemical systems. This carbon atom will contain explicitly one nuclear charge and one electron. Moreover, the

method should be usable out of the box in any standard quantum chemistry software. Thus, no modification of the source code should be done. This supplementary constraint will be fulfilled by strategically positioning the pseudo-potentials we intend to use.

Pseudo-potentials for hybridised carbon atoms have been already successfully extracted in a previous work of ours.¹⁷ There, our attention was focused on the 'no-collapse' term, and functions in the pseudo-potential shift up unwanted orbitals and correctly place the energy of the wanted orbitals. However, some pseudo-potentials had to be put precisely at the center of each bond. Thus, those pseudo-potentials could not be considered as atomic as they were not defined solely with respect to the position of the atom they applied to. The previous work should therefore be regarded as a proof-of-concept. In this new version, we include more physical meaning in our model, by placing pseudo-potentials that mimic Coulomb interactions among both the active and dormant electrons, and the shielded nuclear charge. The new pseudo-potentials are fully atomic: the position of the atom determines completely where the potentials are placed in space in the best possible manner when dealing with hybridised atoms (as hybridisation destroys isotropy, the preferred orientations must be given to define the pseudo-potential).

This article is structured as follows. In the first part, the method is defined in detail. The second part shows the results obtained, from the most crude to the most refined version of the method. The extracted pseudo-potentials are then employed for extended systems. Perspectives are presented in the conclusion.

2 Method

2.1 Pseudopotential definition

As in our previous work,¹⁷ we take CH_3^\bullet as our starting system to be reproduced. It is the smallest system containing one and only one sp^2 hybridised carbon atom. This choice allows us to isolate the different electrostatic interactions to build a physically meaningful model.

We make use of two kinds of Gaussian pseudo-potentials,¹⁸ of s and p shapes. As we want to avoid modifying the quantum chemistry software itself, the pseudo-potentials have a semi-local form and we choose a value of $n = 1$ to preserve the r^{-1} behaviour as in Huzinaga's work. The multi-centered pseudo-potentials that we build read as follows:

$$\hat{W} = r^{-n} \left[\underbrace{\sum_I A_I \exp(-\alpha_I r^2) |Y_{1,m}\rangle \langle Y_{1,m}|}_{p \text{ projectors}} + \underbrace{\sum_J C_J r \exp(-\gamma_J (r - r_J^0)^2) |Y_{0,0}\rangle \langle Y_{0,0}|}_{s \text{ projectors}} \right] \quad (4)$$

with $Y_{0,0}$ the s spherical harmonic, $Y_{1,m}$ the p spherical harmonics (the sum over m is not written explicitly to improve readability) and r_J^0 the relative fixed position of the J^{th} potential with respect to the origin of the pseudo-atom, which carries the pseudo-potentials. By analogy with Huzinaga model potentials defined in (3) we can say that the p projectors aim to mimic a coulombic interaction, while the s potentials have the task of recovering part of the bi-electronic interaction between the dormant and the active electrons. Thus, this second term is not stricly equivalent to Huzinaga's B_J terms as they aim at recovering the bi-eletronic interactions and the no-collapse term. This later role was devoted to the $\hat{\sigma}$ operator in Huzinaga's formulation.

As will be discussed in the following sections, the p pseudo-potential allows to reproduce atomic properties, when the s -potentials mimic electrostatic electron pair repulsion. The pseudo-potentials are mono-electronic operators, thus they shall not be able to reproduce fully bi-electronic interactions. Knowing this, we adjust the parameters of the s -potentials for the pseudo-system to have a wavefunction in which properties, such as spatial extent, energy and excitation energies, are as close as possible to those of the reference system.

Figure 1 displays the final pseudo-system: The CH_3^\bullet radical has been replaced by a hydrogen-like "pseudo-carbon", with a nuclear charge of $Z_{nucleus} = 1$, and one electron occupying the p_z orbital. There are now no H atoms, and the system is surrounded by three potential sets at a planar distance of d , each consisting of two s -shaped potentials with a

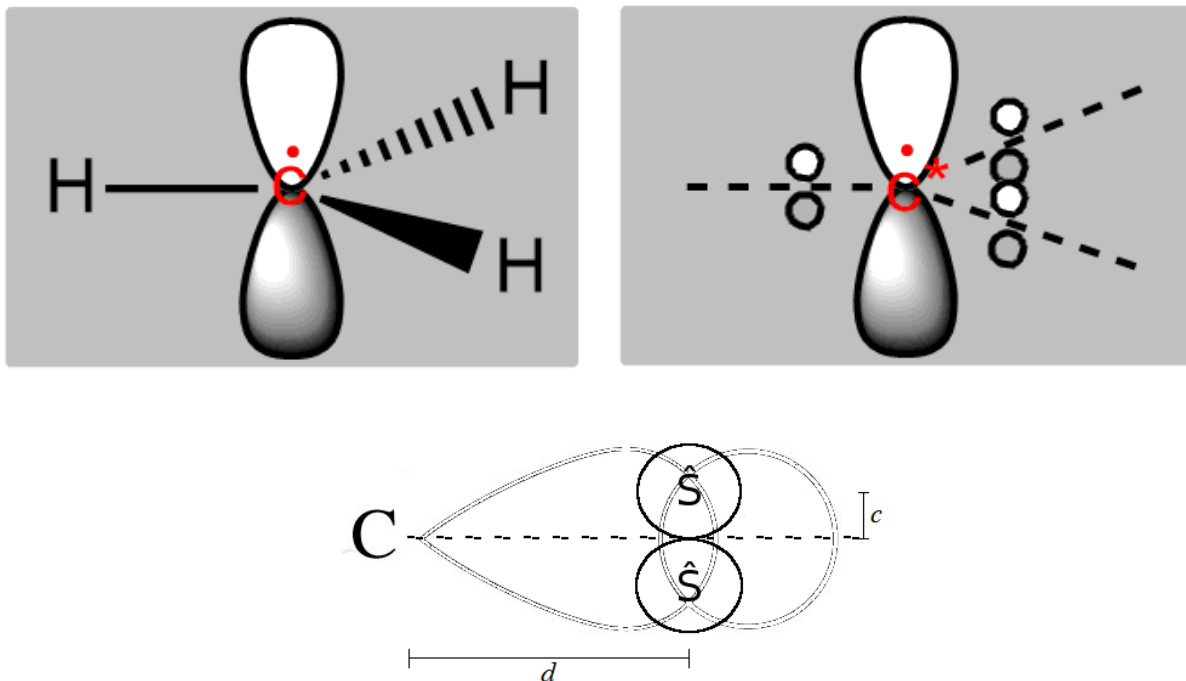


Figure 1: Diagrams of CH_3^\bullet (left) and pseudo- CH_3^\bullet (right, below) molecules. The pseudo- CH_3^\bullet diagrams display the s and p -potential positions, and the distances d and c .

distance above and below the xy -plane of c . A further p -shaped potential is applied directly to the pseudo-carbon.

This gives multiple variables we can use to manipulate the properties of the system. We can alter the strength and diffuseness of the p and s potentials themselves, as well as vary the distances d and c by moving the s -potentials. In this article, we fix $c = 0.25 \text{ a.u.}$.

2.2 Making Potentials Physically Meaningful

It is possible to obtain the potential parameters entirely by empirical means, in principle. However, we can make a more informed guess at some starting parameters from which to optimise. Clearly the assumption that $Z = 1$ is unrealistic. Slater's Rules¹⁹ suggest that, with the screening effect, the p_z electron of a carbon atom should experience a charge of $Z = 2.4$. To mimic the effect of an electron-screened nucleus, we can use a p_z pseudo-potential. In order to make an educated guess of the parameters of this new potential,

we need to find an expression for $Z_{eff}(\langle r \rangle)$; Z_{eff} being the total nuclear attraction the electron would feel in the real CH_3^\bullet system, taking into account the screening effect of the core electrons that would be present in an all-electron model; and $\langle r \rangle$ being the expected distance of the electron from the nucleus.

The analytical form of the p_z orbital for a hydrogen-like atom is²⁰

$$\phi_{210} = \frac{1}{\sqrt{\pi}} \frac{Z_{eff}^{\frac{5}{2}}}{2a_0} r e^{-\frac{Z_{eff}r}{2a_0}} \cos \theta \quad (5)$$

and from this we obtain

$$\langle \phi_{210} | r | \phi_{210} \rangle = \frac{5a_0}{Z_{eff}} \quad (6)$$

Next, we need to find a value for $\langle r \rangle$. We can do this by calculating

$$\langle r \rangle = \langle \psi_{p_z} | r | \psi_{p_z} \rangle \quad (7)$$

where ψ_{p_z} are the molecular orbitals extracted from a reference calculation of CH_3^\bullet . From Equation 7, we have $\langle r \rangle \approx 1.8 \text{ a.u.}$ and therefore we can see from Equation 6 that $Z_{eff} \approx 3.6$. Even if we shall use the semi-local pseudo-potential definition for the calculations, for a quick evaluation of overlap effects between the basis set and the pseudo-potential, we assume a non-local pseudo-potential definition. We define:

$$S = \langle \psi_{p_z} | \chi \rangle \quad (8)$$

the overlap between a molecular orbital ψ_{p_z} , and χ , taken from our non-local pseudo-potential definition¹¹

$$\chi = e^{-\alpha r^2}, \quad \hat{p} = Z_{pseudo} |\chi\rangle \langle \chi| \quad (9)$$

leading us to

$$\langle \hat{Z} \rangle = \langle \psi_{p_z} | \hat{Z} | \psi_{p_z} \rangle = \langle \psi_{p_z} | Z_{pseudo} | \chi \rangle \langle \chi | \psi_{p_z} \rangle = Z_{pseudo} S^2 \quad (10)$$

Finally, knowing that our hydrogen-like pseudo-system already contains a charge, $Z_{nucleus} = 1$, which we subtract from the desired Z_{eff} with which we want to influence our p_z electron, leaving us with

$$Z_{pseudo} = (Z_{eff} - Z_{nucleus})S^{-2} \quad (11)$$

We now have the power to choose a p_z pseudo-potential solely based on the Gaussian exponent, and the Z_{eff} then follows from the above. Clearly, the exponent should be chosen to give us a strong overlap with the molecular orbital we wish to influence. Hence we arrive at a p_z -potential that should be physically meaningful. We chose the exponent in Table 3, which gave the maximum overlap possible (≈ 0.79).

3 Results, Development and Discussion

3.1 CH₃, with s -potentials

Table 1: Coefficients and exponents for s -only pseudo-potentials for CH₃[•], optimised to give the all-electron HOMO reference energy of -10.537 eV (HF) and -6.726 eV (PBE0). $d = 0.5$ a.u., as defined in Figure 1.

	Coefficient	Exponent
HF	-2.594	1.0
	-4.788	5.0
	-7.524	10.0
PBE0	-2.605	1.0
	-4.873	5.0
	-7.678	10.0

We aim first at reproducing the values for the CH₃[•] radical as given in Table 1. This reference CH₃[•] is created and has its geometry optimised under Hartree-Fock (HF). The reference geometry gives a C - H distance of 2.0466 a.u., and so we pick $d = 2.0$ a.u. as a starting guess for the planar distance from the pseudo-carbon, d , of our pseudo-potentials. The pseudo-system is then set up, erasing the hydrogen atoms, setting the carbon charge $Z_{nucleus} = 1$ and applying s pseudo-potentials, as well as selecting the correct orbital for

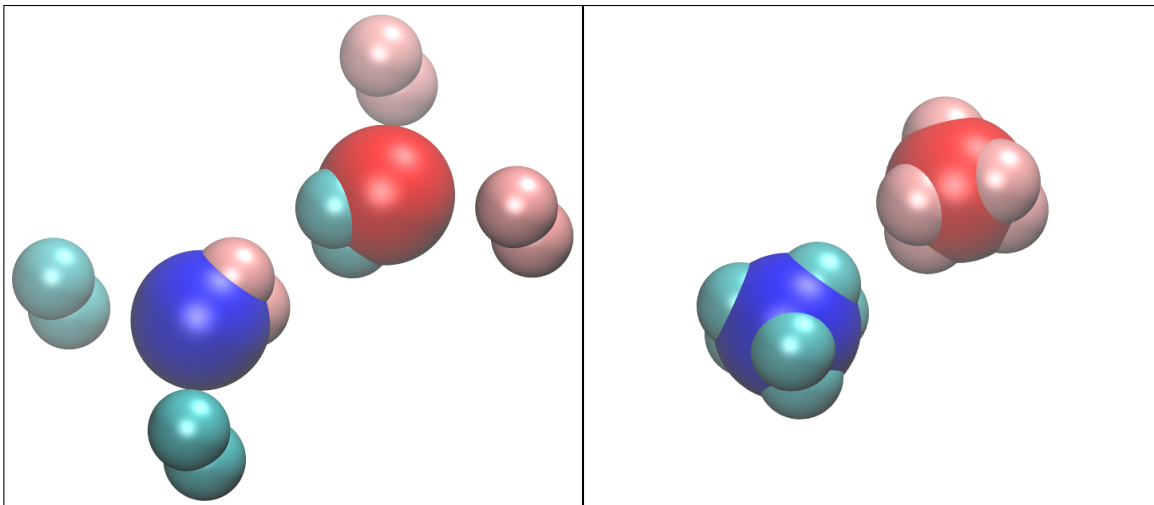


Figure 2: Diagrams of pseudo-ethene with $d = 2.0$ a.u. (left), and $d = 0.5$ a.u. (right). The first pseudo-carbon is displayed in blue, with its s pseudo-potentials in cyan, and the second pseudo-carbon is in red, with its potentials in pink.

the remaining electron. Table 1 displays some of our results. Promisingly, we are able to produce many sets of potentials that give the correct energy.

3.2 Ethene, with s and p -potentials

Next, we take some of these potentials to create a pseudo-ethene system, with the results shown in Table 2. All potentials tested with $d = 2.0$ a.u. gave results several orders of magnitude away from the reference value. From Figure 2 we may see the reason. One of the potential sets from each carbon is closer to the neighbouring carbon than the neighbour's own potential sets, thus both pseudo-carbons are affected by potentials which do not belong to them. At the shorter range $d = 0.5$ a.u., the HOMO energy is of the right magnitude, though with errors of 30%. Attempts to eliminate this error lead us to the additional use of the p_z potential *alongside* the s potentials.

The next step adds a p -potential centred on the pseudo-carbon, with Table 3 displaying the results. As before, $d = 0.5$ a.u.. The p_z potential is selected using the procedure described in Section 2.2, with the exponent chosen to give the maximum possible overlap with the p_z orbital, and the matching Z_{eff} coefficient calculated from the exponent and overlap. The

Table 2: HOMO energies (π , eV) for ethene and pseudo-ethene. The pseudo-potentials used in the pseudo-ethene are taken from optimised values for pseudo-CH₃[•] with only s -potentials and an exponent of 10.0.

	d	s coefficient	π
HF ^a			-10.363
PBE0 ^a			-6.632
HF	2.0	-7.521	-9597.0
HF	0.5	-7.521	-7.905
PBE0	0.5	-7.678	-8.447

^a All-electron reference values.

s -potentials are then optimised once more to give the correct HOMO energy for CH₃[•]. We again take these potentials to create a pseudo-ethene molecule, with the results shown in Table 3. We can see that these potentials seem to transfer more effectively from the CH₃[•] system to the ethene, suggesting therefore that whilst the s -potentials can affect both the p_z and π orbitals, they cannot alone describe the relationship between them.

Table 3: π orbital energy values (eV) for pseudo-ethene, using s and p pseudo-potentials. The pseudo-potentials are taken from a pseudo-CH₃[•] system optimised to give the correct all-electron HOMO energy.

	p coefficient	p exponent	
p_z potential	-3.267	0.295	
Calculation Type	s coefficient	s exponent	π
HF	2.772	1.0	-13.654
	6.173	5.0	-14.011
	10.381	10.0	-14.061
PBE0	3.483	1.0	-10.325
	9.801	5.0	-10.409
	18.351	10.0	-12.543

Having successfully created a pseudo-ethene with the correct HOMO, we attempt to have the pseudo-system replicate other properties of the real system: the singlet-triplet excitation energy (Δ_{ST}) in the SCF framework (energy difference between the triplet and the singlet mono-reference calculations), the ionisation energy (IE) and the energy of the HOMO orbital (ε_{HOMO}). Reference values for the singlet-triplet $\pi-\pi^*$ excitation and first ionisation energies of ethene are given in Table 4. Testing the relevant energies for the optimised pseudo-systems

above, we can see from Table 4 that the early results are not promising. However, after we abandon the notion of sticking strictly to a p_z -potential exponent that gives the maximum overlap with the real orbital, we discover there is a "sweet spot" of potential coefficients and exponents around which the correct values begin to emerge. Table 4 shows our optimal result, chosen to give HOMO, triplet - singlet and ionisation energies closest to the reference values.

Table 4: s -potential fits to ethene properties as defined in the text (eV). All values are from HF calculations, and $d = 0.5$ a.u.

s coefficient	s exponent	Δ_{ST}	Ionisation	ε_{HOMO}
<i>Reference Values</i>		-3.533	-9.091	-10.363
p coefficient -3.267; p exponent 0.295				
0.552	0.1	-3.533	-27.158	-27.31
0.608	1.0	-3.533	-28.247	-28.395
0.936	5.0	-3.533	-29.583	-29.722
<i>Optimal Values</i>		p coefficient -3.91; p exponent 0.624		
1.5	0.5	-3.533	-9.806	-10.062

3.3 $C_{2n}H_{2n+2}$, $n = 2 - 12$

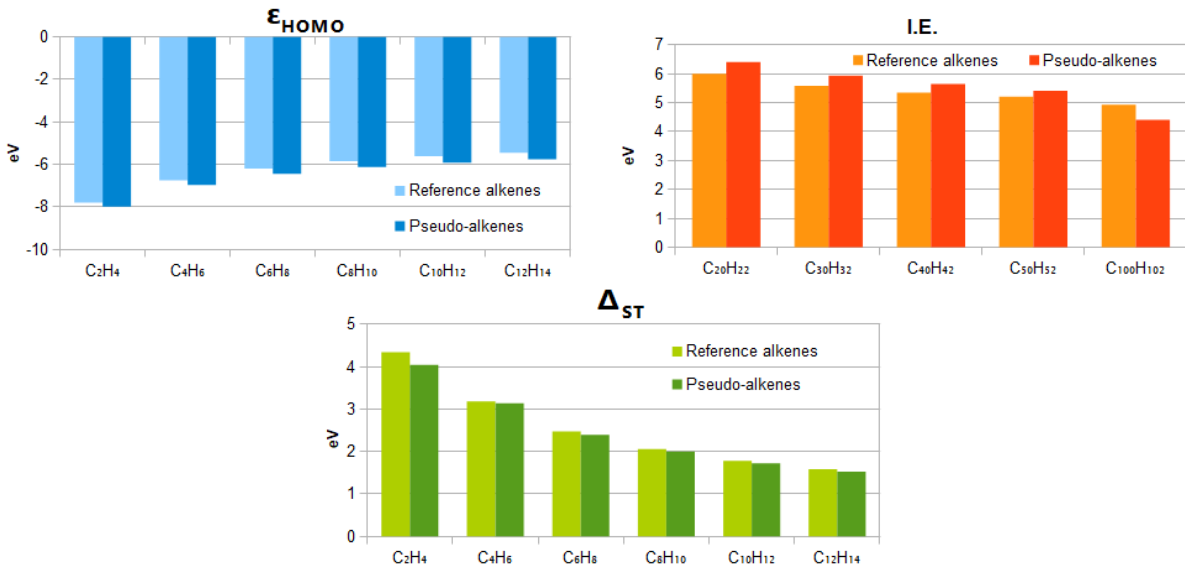


Figure 3: DFT and TD-DFT (PBE0) comparison of reference and pseudo-system energies across a range of chain alkenes and computational methods.

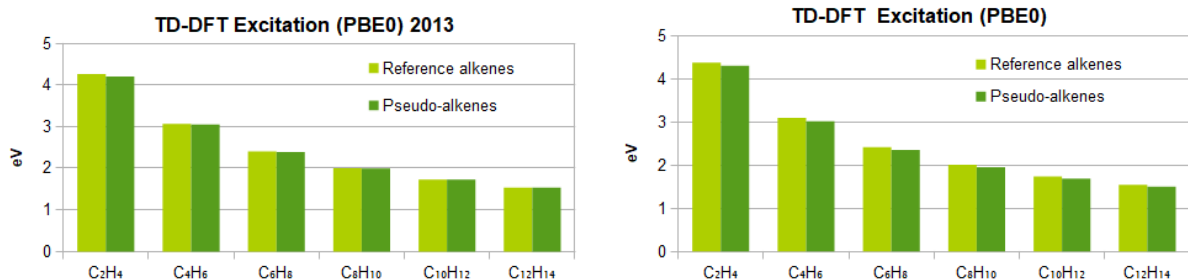


Figure 4: Comparison of pseudo-alkenes with previous¹⁷ and current potentials using TD-DFT excitation energies.

Table 5: %-errors across calculation types for short chain alkenes (C₂-C₁₂)

Calculation Type	HF	PBE0	PBE	TPSS	TPSSH
Mean $\pi - \pi^*$ Triplet-Singlet error (%)	4.1	3.6	7.5	13.3	11.4
Mean Ionisation error (%)	5.3	6.0	8.1	10.1	9.0
Mean Singlet HOMO error (%)	1.3	4.0	8.5	12.0	9.7
Mean TD-DFT Excitation error (%)	-	2.6	-	-	-

Taking these new potentials we test them against a series of chain alkenes up to length C₁₂H₁₄, using a variety of functionals. In each case, the geometry of the reference system is optimised according to the method used, before taking the reference geometry and applying the pseudo-potentials from Table 4. Figures 3 and 4 show the results. Table 5 gives a breakdown of the percentage errors for each method across all molecules tested. The pattern of increasing HOMO energy and decreasing cation-singlet and triplet-singlet energies seen in the reference systems is well replicated by the pseudo-alkenes, with the energies following the same gradient.

Also compared are TD-DFT results for this system and results of a previous work of some of the authors, which match to within 3%.¹⁷ Let us recall here that in our previous work potentials were placed at the center of bonds.

3.4 Large Systems

The potentials derived above are also tested on larger systems. Figure 5 shows the Δ_{ST} , the IE and HOMO energy values for several ring systems. As with the short chain alkenes,

the general trend of the results is well-replicated by the pseudo-systems, and the percentage errors, displayed in Table 6, are similar.

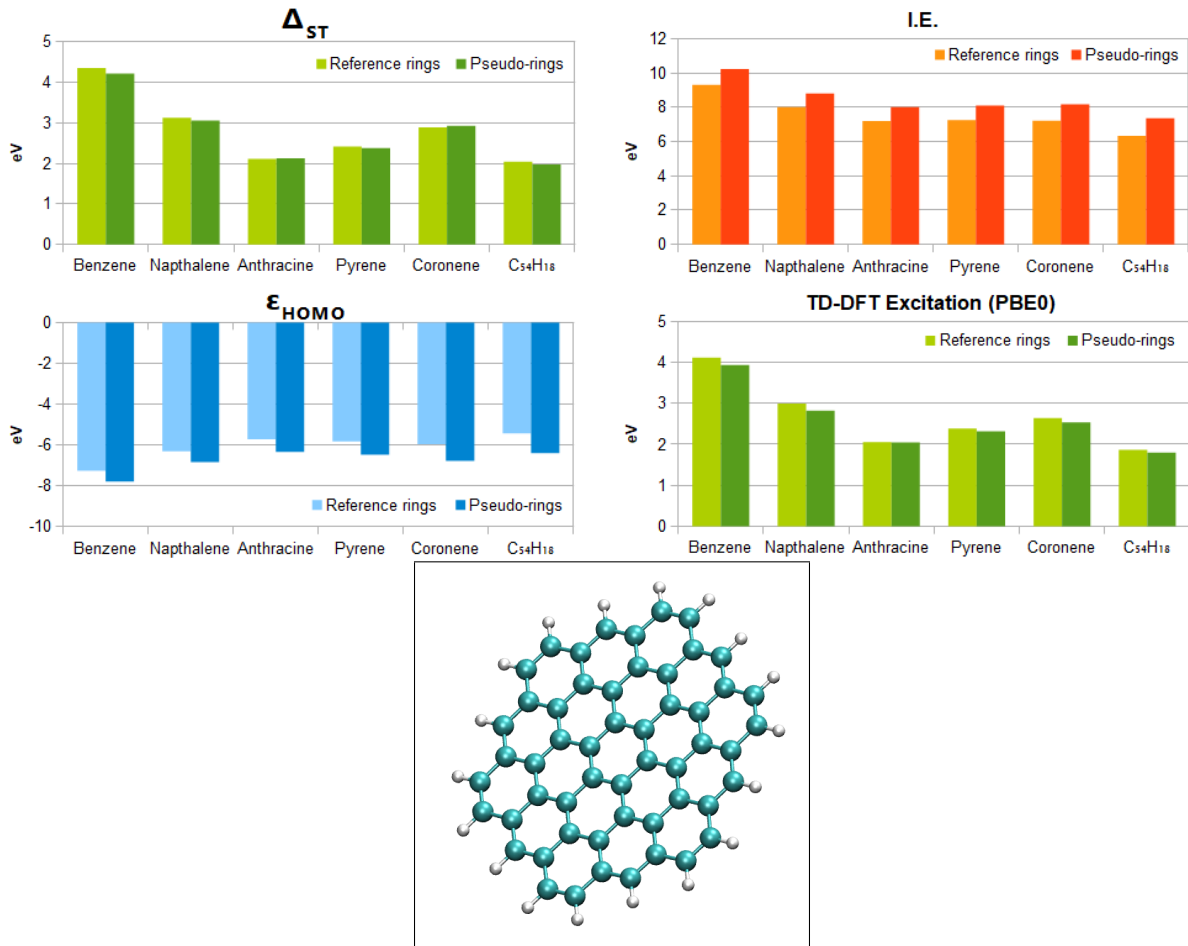


Figure 5: DFT and TD-DFT (PBE0) comparison of reference and pseudo-system energies across a range of ring molecules. C₅₄H₁₈ is shown in the bottom.

Figure 6 and Table 7 refer to longer alkene chains (n up to 100). The pattern of decreasing ionisation potentials and Δ_{ST} with increasing HOMO energy is still followed, with the absolute error remaining consistent. However, differences in the triplet-singlet energies between the reference and pseudo-systems become significant, notably for the largest case.

Unlike for the previous systems, there is a large discrepancy between Δ_{ST} and TD-DFT results. This apparent failure of the pseudo-potentials is to be found in the representation of the triplet state. The expectation values of the S^2 operator for the triplet calculations are plotted in Figure 7, which shows that the spin contamination of the triplet state computed

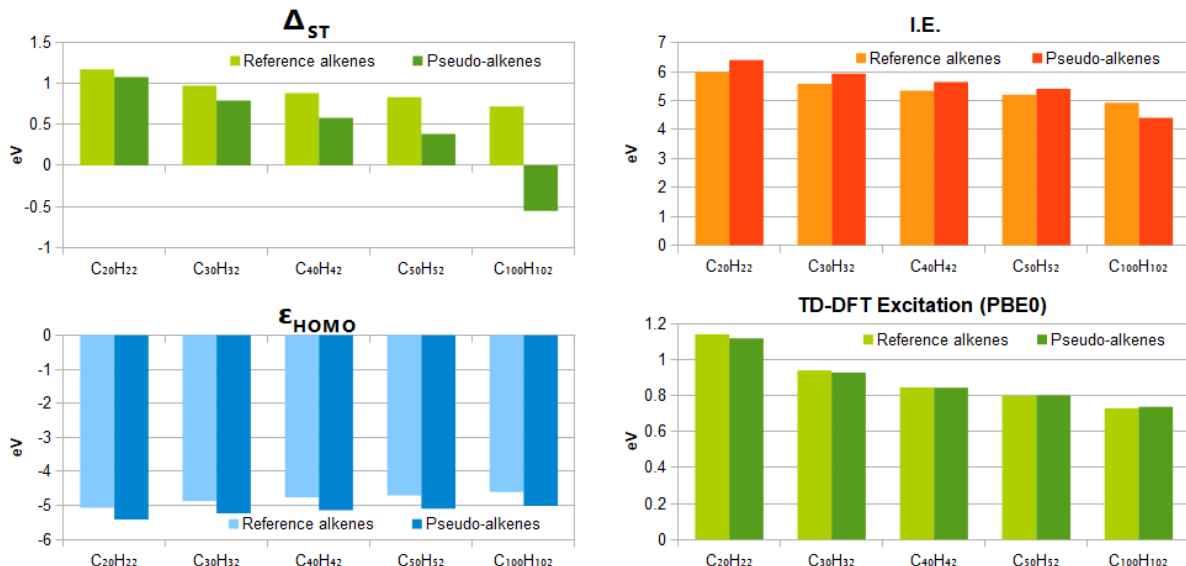


Figure 6: DFT and TD-DFT (PBE0) comparison of reference and pseudo-system energies across a range of long chain alkenes (C₂₀-C₁₀₀).

Table 6: %-errors across calculation types for ring-systems

Calculation Type	HF	PBE0	PBE	TPSS	TPSSH
Mean $\pi - \pi^*$ Triplet-Singlet error (%)	3.3	1.9	1.7	6.8	6.7
Mean Ionisation error (%)	7.4	10.0	12.5	14.4	13.0
Mean Singlet HOMO error (%)	3.4	9.3	13.4	17.1	14.7
Mean TD-DFT Excitation error (%)	-	0.099	-	-	-

as a single configuration (*i.e.* in a SCF framework) increases in both reference and pseudo-potential cases. Yet, this effect is strengthened in the pseudo-potential calculations. These results suggest that the triplet state cannot be represented as one configuration but as a linear combination of mono-electronic excitations as done through TD-DFT.

In order to show that the recovering of the agreement between the pseudo-potential and the reference calculations is not an artefact, we give in Table 8 the weight and nature of each excitation (weight larger than 3%) in the description of the triplet excited state for C₅₀H₅₂ (other values can be found in the SI, which exhibit the same trends). As can be seen, the agreement is very good.

These results show that the pseudo-potentials that we have extracted are able to reproduce the π systems in a variety of situations which are not part of their extraction set. The

Table 7: %-errors across calculation types for long chain alkenes (C₂₀-C₁₀₀)

Calculation Type	HF	PBE0	PBE	TPSS	TPSSH
Mean $\pi - \pi^*$ Triplet-Singlet error (%)	55.3	85.8	83.4	239.8	320.2
Mean Ionisation error (%)	25.1	6.7	9.4	10.3	11.6
Mean Singlet HOMO error (%)	1.8	7.3	11.3	16.7	13.6
Mean TD-DFT Excitation error (%)	-	0.005	-	-	-

Table 8: Comparison of the weights (all electron *vs.* pseudo-potentials) of the excitations obtained with TD-DFT to represent the triplet excited state from the closed shell singlet state. Example case of C₅₀H₅₂.

Excitation MO	Weight(%)	
	Ref.	Pseudo.
25 a" \rightarrow 26 a"	77.0	67.1
24 a" \rightarrow 27 a"	10.5	13.1
23 a" \rightarrow 28 a"	3.6	5.2

molecular orbital virtual space is also well described (cf. Table 8), which demonstrates that the good agreement with reference calculations is physically grounded.

3.5 Computational Details

All Hartree-Fock, DFT and time-dependent DFT (TD-DFT) energy calculations are performed with TURBOMOLE 7.1.²¹ The basis set used throughout is def-SV(P).²² Wherever possible, planar (C_S) symmetry is used. The convergence energy is 10^{-7} H (\$`scfconv` = 7) for SCF and 10^{-6} H for DFT. When running these calculations the occupation of orbitals is specified manually in the TURBOMOLE control file.

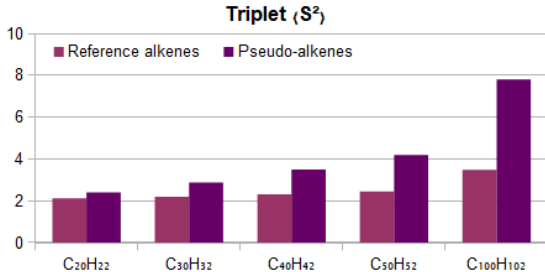


Figure 7: Comparison of S^2 expectation values obtained for the calculation of the first triplet configuration in a SCF formalism, for reference and pseudo-systems.

Chain Alkenes, Ring Molecules. In addition to Hartree-Fock calculations, DFT is used with PBE0, PBE, TPSS and TPSSh functionals.²³⁻²⁶ The integration grid size is set at *m*4. Also used are TD-DFT calculations, where the Tamm-Dancoff approximation (CIS)²⁷ is switched on to avoid triplet instability.

3.6 Optimisation

In earlier calculations with only *s*-potentials, optimisation was performed by choosing a range of exponent values and attempting to optimise the coefficient at each to produce the HOMO reference energy. Once the *p_z* potential was added, the *s*-potentials were optimised afterward.

Once we started to look at excitation and ionisation energies however, optimisation became more complicated. Optimisations were at first performed of the *s* and *p*-potentials to reach the HOMO energy of ethene as before. With the different potential variables available, we produced a range of optimised potential sets. The best set of these potentials was then chosen and the values altered by hand in order to match as closely as possible three separate reference values: the singlet HOMO energy, the singlet-triplet $\pi - \pi^*$ excitation energy, and the cation-singlet energy. All optimisations used the Brent method in SciPy’s optimisation library, with a tolerance of 1.48×10^{-08} , and used standard Hartree-Fock calculations.²⁸

4 Conclusion

In this work, we tackled the two main problems of the initial version of our molecular potentials. Firstly, the new pseudo-potentials for sp^2 hybridised atoms are completely atomic and, even if the directionality of the bonding pattern has to be fulfilled by correctly positioning the *s* potentials, no potentials need to be added relative to the position of two atoms. Secondly, we gave a physical meaning to all the pseudo-potential terms. Contrary to our previous attempt, we do not rely on the "no collapse" term, which is modeled here by forc-

ing the molecular occupation (*vide infra*). We could show that not only were the occupied orbitals well-reproduced by the use of these new potentials, but also that the virtual space is of good quality for excited states calculation.

In the framework of this study, there are three logical next steps: the addition of an explicit "no collapse" term, the reproduction of the gradient through the parameterisation of additional terms, and finally the use of such potentials in the framework of QM/MM calculations as a replacement of hydrogen based link atoms. The first step is needed in order to avoid forcing the molecular orbital occupation and to avoid spurious virtual orbitals in the active space. We are confident that the first step could be easily done, as we have already extracted such terms in the previous version of our potentials. The reproduction of the gradient is more of a challenge as many electrons were removed from the system. Finally, QM/MM tests can be started quickly, as we could fix the local geometry of the potentials in the QM part. The MM part would be taken care of without modification of the standard parameters.

Supporting Information Available

A listing of the contents of each file supplied as Supporting Information should be included. For instructions on what should be included in the Supporting Information as well as how to prepare this material for publications, refer to the journal's Instructions for Authors.

The following files are available free of charge.

- ResultsandGraphs.ods: Complete results of calculations and associated graphs.
- coordinates folder: Contains all coordinate files used to generate the results in "Results & Graphs.ods", ordered by molecule shape, all-electron/pseudo-molecular calculation and functional/method.

This material is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. *Chemical Reviews* **2015**, *115*, 5678–5796.
- (2) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *The Journal of Physical Chemistry A* **2001**, *105*, 293–307.
- (3) Steinmann, C.; Fedorov, D. G.; Jensen, J. H. *PLoS ONE* **2012**, *7*.
- (4) Wesolowski, T. A.; Shedge, S.; Zhou, X. *Chemical Reviews* **2015**, *115*, 5891–5928.
- (5) Assfeld, X.; Rivail, J.-L. *Chemical Physics Letters* **1996**, *263*, 100 – 106.
- (6) Dolg, M.; Cao, X. *Chemical Reviews* **2012**, *112*, 403–480.
- (7) Huzinaga, S. *Canadian Journal of Chemistry* **1995**, *73*, 619–628.
- (8) Carissan, Y.; Bessac, F.; Alary, F.; Heully, J.-L.; Poteau, R. *International Journal of Quantum Chemistry* **2006**, *106*, 727–733.
- (9) Raynaud, C.; del Rosal, I.; Jolibois, F.; Maron, L.; Poteau, R. *Theoretical Chemistry Accounts* **2010**, *126*, 151–163.
- (10) Nicolas, G.; Durand, P. *The Journal of Chemical Physics* **1980**, *72*, 453–463.
- (11) Huzinaga, S. *Journal of Molecular Structure: THEOCHEM* **1991**, *234*, 51–73.
- (12) Hernandez, N. C.; Sanz, J. F. *Journal of chemical physics* **2000**, *113*, 6082–6087.
- (13) Romero, J. A. M.; Sanz, J. F. *Journal of chemical physics* **1993**, *99*, 1255–1261.
- (14) Mejias, J. A.; Sanz, J. F. *Journal of chemical physics* **1994**, *102*, 327–336.
- (15) Katsuki, S. *Canadian Journal of Chemistry* **1992**, *70*, 285–289.

- (16) Katsuki, S. *The Journal of Chemical Physics* **1993**, *98*, 496–501.
- (17) Drujon, J.; Carissan, Y. *Journal of Computational Chemistry* **2013**, *34*, 49–59.
- (18) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley and Sons Ltd, 2002.
- (19) Slater, J. C. *Physical Review* **1930**, *1*.
- (20) Tuckerman, M. E. CHEM-UA 127: Advanced General Chemistry I. 2011; http://www.nyu.edu/classes/tuckerman/adv.chem/lectures/lecture_8/.
- (21) *TURBOMOLE V7.1 2016*; a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- (22) Schäfer, A.; Horn, H.; Ahlrichs, R. *Journal of Chemical Physics* **1992**, *97*.
- (23) Perdew, J.; Ernzerhof, M.; Burke, K. *Journal of Chemical Physics* **1996**, *105*.
- (24) Perdew, J.; Burke, K.; Ernzerhof, M. *Physical Review letters* **1996**, *77*.
- (25) Tao, J.; Perdew, J.; Staroverov, V.; Scuseria, G. *Physical Review letters* **2003**, *91*.
- (26) Staroverov, V.; Scuseria, G.; Tao, J.; Perdew, J. *Journal of Chemical Physics* **2003**, *119*.
- (27) Peach, M. J. G.; Williamson, M. J.; Tozer, D. J. *Journal of Chemical Theory and Computation* **2011**, *7*.
- (28) Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open source scientific tools for Python. 2001; <http://www.scipy.org/>.

Graphical TOC Entry

