

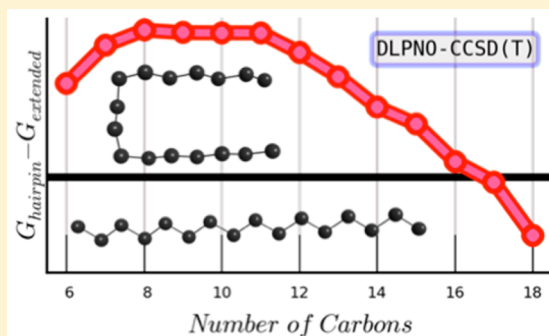
Domain Based Pair Natural Orbital Coupled Cluster Studies on Linear and Folded Alkane Chains

Dimitrios G. Liakos and Frank Neese*

Max Planck Institute for Chemical Energy Conversion, Stiftstr. 32-34, 45470 Mülheim an der Ruhr, Germany

S Supporting Information

ABSTRACT: In this study the question of what is the last unbranched alkane that prefers a linear conformation over a folded one is revisited from a theoretical point of view. Geometries have been optimized carefully using the most accurate theoretical approach available to date for such systems, namely, doubly hybrid density functional theory in conjunction with larger quadruple- ζ quality basis sets. The resulting geometries deviate significantly from previously reported ones and have a significant impact on the predicted energetics. Electronic energies were calculated using the efficient and accurate domain local pair natural orbital coupled cluster method with single-, double-, and triple substitutions (DLPNO-CCSD(T)) electronic structure method. Owing to the method's efficiency, we were able to employ up to quadruple- ζ quality basis sets for all hydrocarbons up to $C_{19}H_{40}$. In conjunction with carefully designed basis set extrapolation techniques, it is estimated that the electronic energies reported in this study deviate less than 1 kJ/mol from the canonical CCSD(T) basis set limit. Thermodynamic corrections were calculated with the PW6B95-D3 functional and the def2-QZVP basis set. Our prediction is that the last linear conformer is either $C_{16}H_{34}$ or $C_{17}H_{36}$ with the latter being more probable. $C_{18}H_{38}$ can be safely ruled out as the most stable isomer at 100 K. These findings are in agreement with the elegant experimental studies of Suhm and co-workers. Deviations between the current and previous theoretical results are analyzed in detail.



1. INTRODUCTION

Alkane chains, C_nH_{2n+2} , are the most fundamental organic molecules. Despite their simplicity their accurate description both theoretically and experimentally presents serious difficulties.^{1–8} The theoretical treatment of alkanes is an old problem in chemistry. Already in 1951 Pitzer tried to address the problem of calculating the height and shape of the potential barrier in ethane and other hydrocarbons.⁹ In 1975 P.J. Flory tried to “treat the structure and conformations accessible to the chain molecule in such a manner as will enable one to calculate configuration-dependent quantities and to average over all conformations, or spatial configurations, of the unperturbed chain.”¹⁰

Nevertheless two decades later the computational effort needed for accurate treatment was still considerable and only in 1991 higher level ab initio methods (MP4SDQ/6-31G*, a simplified version of fourth order Möller–Plesset perturbation theory where triple substitutions are omitted^{11,12}) were used for the “Investigation of intramolecular interactions in *n*-Alkanes”.¹³ It was clearly understood at the time that different conformations of a molecule may differ in stabilizing van der Waals interactions that should not be neglected.¹³ Studies then were made to show the effect of correlation in geometries but still one could go up to hexanes.¹⁴ Finally, in 1996 the “gold standard” of quantum chemistry, namely, the coupled cluster method with single-, double-, and perturbative triple excitations

(CCSD(T)^{15,16}), was used for the study of conformational energies and rotational barriers in *n*-alkanes,¹⁷ but the computational restrictions only allowed systems up to hexane to be studied. Similar works followed later.¹⁸

In 1997 Goodman systematically addressed the problem of “what is the longest unbranched alkane which has a linear global minimum?”¹⁹ At the time he used the “Hydrocarbon Force Field” MM2²⁰ and concluded that the first alkane for which the hairpin conformation is lower in energy than the linear one is $C_{18}H_{38}$. Using the more recent MM3* force field,²¹ the predicted crossover point raised to 24 ($C_{24}H_{50}$), while the protein and nucleic acid oriented AMBER force field^{22,23} predicted the change to happen at $C_{25}H_{52}$.

In his early study, Goodman also investigated semiempirical methods. Two significantly different results were obtained here: first, PM3²⁴ calculated the longest linear alkane to be $C_{11}H_{24}$, while the closely related AM1²⁵ method predicted $C_{59}H_{120}$; and second, MNDO²⁶ reversed the trend and predicted the hairpin structure to become increasingly unfavorable as the chain length increased. It should be mentioned that in 2006 Thomas et al., using OPLS-AA, predicted the smallest *n*-alkane with a hairpin geometry that is lower in energy than the all-trans conformer to be $C_{22}H_{46}$,²⁷ a value that is certainly too high, but

Received: March 20, 2015

nevertheless he noted that with a correction for GG sequences the true turning point is likely in the C_{16} – C_{18} range.²⁷

Since then a large number of calculations appeared in the literature approaching the problem from different angles.^{3,27–31} In 2009 Martin and his colleagues published a rigorous study on the conformational equilibria of alkanes up to octane.³² In their study they addressed all different aspects of the calculations, such as correlation method, effect of basis set, and performance of density functional methods.

In 2013 Suhm, Mata, and co-workers published a detailed combined experimental and high level ab initio study to address the problem of finding the last globally extended alkane.⁷ Experimentally they showed that at low temperatures $C_{18}H_{38}$ should be an upper bound with $n_c = 17$ being the most likely crossover point. These results were in agreement with compound ab initio quantum calculations. Subsequently, in 2014 Byrd, Bartlett, and Montgomery used different complete basis set (CBS) extrapolation schemes in conjunction with the CCSD(T) method to address the same problem. Through a slightly different route they concluded that “this finding is in complete agreement with the experimental results of Lüttswagger et al.”⁸

Beside the apparent final agreement between the contributions of Byrd et al.⁸ and Lüttswagger et al.,⁷ a more detailed study reveals some apparent disagreement. In addition, while Byrd et al. treated only even carbons from 8 to 18, Lüttswagger et al. treated even and odd carbons but only after $C_{14}H_{30}$.

In this study we revisit the problem using the more recently developed domain based pair natural orbital coupled cluster method DLPNO-CCSD(T). Owing to the high intrinsic efficiency and accuracy of this method, it can be applied in conjunction with large basis sets to large molecules and hence is able to avoid some of the approximations that are inherent in the earlier works.

2. THEORY AND TECHNICAL DETAIL

2.1. Technical Details. The ORCA quantum chemistry package was used for all calculations.³³ The basis set used for the geometry optimization was def2-QZVP.³⁴ For the single point calculations with wave function based correlation methods, (CCSD(T), MP2, and DLPNO-CCSD(T)), we used the cc-pVnZ and aug-cc-pVnZ^{35,36} basis sets with $n = D, T,$ and Q . For the MP2 and DLPNO-CCSD(T) cases where the resolution of the identity (RI or density fitting, DF) is needed, the corresponding auxiliary basis sets were used.³⁷

2.2. Geometries. One important aspect of the problem at hand is what method to use for the geometry optimization of the isomers and how important is the effect of the different choices. Lüttswagger et al. used the DF-LMP2 method^{38–41} in conjunction with the cc-pVTZ basis set. Byrd et al.⁸ used the canonical MP2⁴² method with the same basis set, and this choice was based on earlier studies on shorter alkanes^{30,32,43} and the finding by the authors that for longer alkanes the use of a small basis set leads to erroneous hairpin structures.

Recently, a detailed study by Grimme and co-workers addressed the problem of how to best optimize the structures of organic molecules in cases where weak interactions are important.⁴⁴ They concluded that the “Jacobs ladder” classification scheme also holds for molecular structures, and hence for higher accuracy they recommended the use of the double hybrid B2PLYP-D3 functional together with the QZVP basis set. Despite the fact that, owing to the MP2-like step in

this method, this is an expensive method, we nevertheless used it for our optimizations. In addition, Grimme and co-workers pointed out that the efficient PW6B95-D3 method provides rather accurate structures, and we used this method for a second set of geometry optimizations together with frequencies in order to get the thermochemical corrections. This was done since the calculation of frequencies using the B2PLYP-D3/QZVP combination would be computationally too expensive for the longer hydrocarbon chains.

2.3. Electronic Energies. In addition to the CCSD(T)/CBS calculation described in more detail below, DFT calculations were performed using the B2PLYP⁴⁵ and PW6B95⁴⁶ functionals both corrected for dispersion forces within the -D3 formalism.^{47,48} For the D3 term, we used the damping scheme of Becke and Johnson.^{49,50} In our DFT calculations we used the def2-QZVP³⁴ basis set both with B2PLYP-D3 and PW6B95-D3.

2.4. Thermochemistry. For the thermodynamic corrections we used the PW6B95-D3/def2-QZVP basis set for both optimization and frequencies. We calculated the zero-point energy using scaled frequencies. To this we added thermal internal energy corrections to get the enthalpy and then vibrational and rotational contributions due to entropy to get Gibbs free energy. For the vibrational entropy contributions we used the scheme proposed by Grimme.⁵¹ Our thermodynamic corrections, beside the ZPE, were calculated at 100 K in order to compare with the experiment.

3. RESULTS AND DISCUSSION

3.1. Effect of Geometry. Lüttswagger et al.⁷ studied the properties of C_nH_{2n+2} molecules with $n = 14, 15, 16, 17, 18, 19, 20, 21,$ and 22 . Byrd et al.⁸ chose only even numbered carbon chains and hence treated $n = 8, 10, 12, 14, 16,$ and 18 . With the idea that the molecules with smaller number of carbons are useful for internal method calibration and are computationally not very demanding, we treated all linear and folded alkanes from $n = 6$ to $n = 19$.

Two sets of comparison are relevant at this point. First, we make a comparison between the structures of the two previous studies and the current one. In Figure 1 and Figure 2 we present the backbone of hairpin structures for $C_{14}H_{30}$ and $C_{18}H_{38}$ from the two previous studies^{7,8} and the ones we used in our studies.

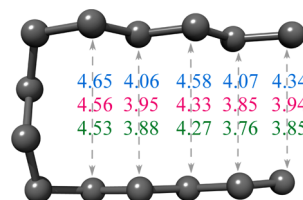


Figure 1. Backbone structure of $C_{14}H_{30}$ showing the distance between the two parallel chains. Blue numbers: Lüttswagger et al.⁷ Red numbers: B2PLYP-D3/def2-QZVP. Green numbers: Byrd et al.⁸

These two structures are representative of the general picture concerning the geometries. Two points deserve notice here. The first is that even though the two previous studies used the same basis set and essentially the same method there are fairly large differences in the C–C distances reaching up to 0.5 Å (for $C_{14}H_{30}$). Our results are overall in reasonably good agreement with Byrd et al.⁸ The significantly increased C–C distances

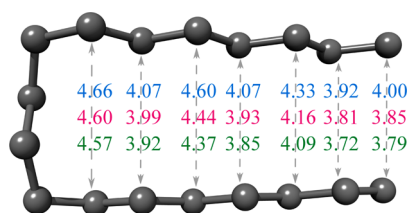


Figure 2. Backbone structure of $C_{18}H_{38}$ showing the distance between the two parallel chains. Blue numbers: Lüttchwager et al.⁷ Red numbers: B2PLYP-D3/def2-QZVP. Green numbers: Byrd et al.⁸

predicted by the LMP2 method will necessarily have an impact on the folding energetics given the steep R^{-6} decrease of the van der Waals interaction. That this is indeed the case will be borne out by the numbers presented below.

To check the effect of the different geometries we calculated the single point energies of the hairpin configurations for $C_{14}H_{30}$, $C_{16}H_{34}$, and $C_{18}H_{38}$, using the CCSD(T) method and cc-pVDZ basis set (Table 1). The combination of method and

Table 1. CCSD(T)/cc-pVDZ Calculated Energy Differences, in kJ/mol, for the Published Geometries with Respect to the Energy Calculated Using the B2PLYP-D3/def2-QZVP Optimized Structures

	current – Lüttchwager et al. ⁷	current – Byrd et al. ⁸
$C_{14}H_{30}$	6.11	2.75
$C_{16}H_{34}$	7.11	3.12
$C_{18}H_{38}$	8.04	3.54

basis set is of course not an accurate one. However, what we want to point out here is how significant the energy changes are that arise from the geometric differences.

The structures we used agree with the structures of Byrd et al. within 0.1 Å, with the structures of Byrd et al. being always slightly “tighter”. This will result, of course, in some limited extra stabilization of the hairpin structures. On the basis of ref 44 one would expect that the ideal geometries lay between the def2-B2PLYP-D3/def2-QZVP and the geometries from ref 8 with the geometries calculated with DFT being slightly more reliable. However, the deviations are so limited that no significant differences are expected. On the other hand the structures of ref 7 lead to significant deviations.

Since we will later calculate the thermodynamic corrections on a differently (PW6B95-D3/def2-QZVP) optimized structure, it is instructive to investigate the differences to our best (B2PLYP-D3/def2-QZVP) geometries (Figure 3).

From Figure 3 one observes that the two sets of geometries are nearly identical. This validates the use of these geometries

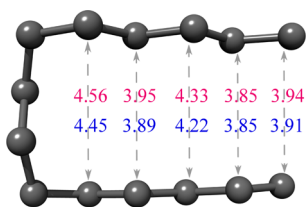


Figure 3. Backbone structure of $C_{14}H_{30}$ showing the distance between the two parallel chains. Red numbers: B2PLYP-D3/def2-QZVP. Blue numbers: PW6B95-D3/def2-QZVP/def2-QZVP.

in calculating the thermodynamic correction and adds further credibility to the employed structures.

3.2. Accuracy of Extrapolation Schemes. In the next step we will address the problem of calculating the electronic energy of the different molecules and postpone the addition of thermodynamic corrections. It has been shown⁵² that the CCSD(T) method is for all intent and purposes accurate enough for the problem at hand and that higher order corrections have minor effects. The problem though is that CCSD(T) needs large basis sets in order to fully exploit its inherent accuracy. This fact combined with the characteristic N^7 scaling of the method makes the required calculations quite challenging, even for present day computers.

One way to avoid this N^7 scaling, used by Byrd, is the frozen natural orbital (FNO) method.^{53–55} In this method one uses the MP2 density matrix to create new virtual orbitals and then replaces the Hartree–Fock orbitals by MP2 virtual natural orbitals. Finally the virtual space is truncated by examining the virtual occupation numbers. Byrd et al. used a threshold of 1×10^{-4} , resulting in 40% of the virtual orbitals being dropped. Then in the CCSD FNO(T) version of the method the converged T_1 and T_2 amplitudes were used to compute the perturbative triples contribution in conjunction with cc-pVTZ basis set.⁸

Another way to address this problem is to exploit the systematic convergence of the energy, both Hartree–Fock and correlation with the basis set. Then using some smaller basis sets to extrapolate to the complete basis set (CBS) limit. Here we will follow the terminology used before.⁵⁶ Two kinds of extrapolation are of main use when one tries to reach the CBS limit. In the first one called EP1^{56–59} the CBS total energy is approximated as

$$E_{\text{CBS}}^{\text{EP1}}(M; X, Y) = \underbrace{\frac{E_{\text{HF}}(Y)e^{-a\sqrt{X}} - E_{\text{HF}}(X)e^{-a\sqrt{Y}}}{e^{-a\sqrt{X}} - e^{-a\sqrt{Y}}}}_{\text{Hartree–Fock}} + \underbrace{\frac{X^\beta E_{\text{corr}}^M(X) - Y^\beta E_{\text{corr}}^M(Y)}{X^\beta - Y^\beta}}_{\text{Correlation}} \quad (1)$$

where M is the method used (e.g., MP2, DLPNO-CCSD(T)) and X and Y are the cardinal numbers of the two basis sets used (e.g., 2 for cc-pVDZ and 3 for cc-pVTZ). $E_{\text{corr}}^M(X)$ is the correlation energy calculated with method M and basis set corresponding to cardinal number X. The second extrapolation scheme (EP2) is the one in eq 2.

$$E_{\text{CBS}}^{\text{EP2*}}(M; X, Y, X') = \underbrace{E_{\text{CBS}}^{\text{EP1}}(M; X, Y)}_{\text{method's M CBS}} + \underbrace{(E_{\text{CCSD(T)}}(X') - E_M(X'))}_{\text{CCSD(T) adjustment}} \quad (2)$$

Here we add a “*” in the definition of the EP2 scheme to indicate out that we can customize the method such that X and X' are not the same. In the traditional EP2 method these two are the same but in principle there is no theoretical reason for this, and as we will show below it can be beneficiary to use a different X, as we will do in our calculations. In what follows we will use the EP2* extrapolation scheme using as method M the DLPNO-CCSD(T) method.

Next we need to analyze the error one expects from the EP2* method in order to decide the final technical details

concerning the basis sets that we are going to use. We define as $E_{\text{CCSD(T)}}^{\infty}$ the true complete basis set limit for CCSD(T). Then the error of the EP2* method would be $\text{Error} = E_{\text{CCSD(T)}}^{\infty} - E_{\text{CBS}}^{\text{EP2*}}(\text{M}; X, Y, X')$. We can analyze this error:

$$\text{Error} = E_{\text{CCSD(T)}}^{\infty} - E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y) - [E_{\text{CCSD(T)}}(X') - E_{\text{M}}(X')] \quad (3)$$

The term $E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y)$ represents the CBS energy calculated with the EP1 formalism for method M using basis sets X and Y. Nevertheless this term should also contain an error with regard to the real CBS of method M at infinity, and then we can rewrite this term in the following way:

$$\begin{aligned} \Delta E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y) &= E_{\text{M}}^{\infty} - E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y) \\ \Rightarrow E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y) &= E_{\text{M}}^{\infty} - \Delta E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y) \end{aligned} \quad (4)$$

Substituting eq 4 into eq 3 we get

$$\begin{aligned} \text{Error} &= E_{\text{CCSD(T)}}^{\infty} - E_{\text{M}}^{\infty} + \Delta E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y) \\ &\quad - [E_{\text{CCSD(T)}}(X') - E_{\text{M}}(X')] \end{aligned} \quad (5)$$

and after some reorganization we end up with

$$\begin{aligned} \text{Error} &= \underbrace{(E_{\text{CCSD(T)}}^{\infty} - E_{\text{M}}^{\infty}) - (E_{\text{CCSD(T)}}(X') - E_{\text{M}}(X'))}_{\text{A}} \\ &\quad + \underbrace{\Delta E_{\text{CBS}}^{\text{EP1}}(\text{M}; X, Y)}_{\text{B}} \end{aligned} \quad (6)$$

In what follows we will try to estimate the errors of term A and term B for the method that we are going to use. It should be emphasized again that X' of term A and X of term B need not be the same. Term A shows how constant the energy differences remain between CCSD(T) and the method M as the size of the basis set increases.

In Figure 4 we plot the difference between CCSD(T) and DLPNO-CCSD(T) calculated with cc-pVnZ and aug-cc-pVnZ

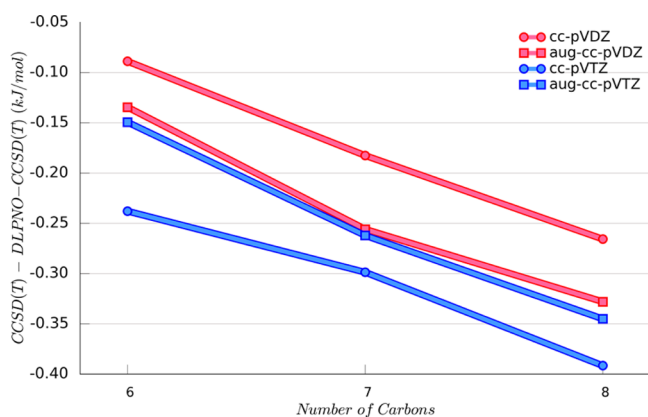


Figure 4. CCSD(T)-DLPNO-CCSD(T) “isomerization” energy calculated with various basis sets.

with $n = \text{D, T}$. Since we could not do calculations with more than 8 carbons for the aug-cc-pVTZ basis set with CCSD(T), we show only these results here. In Figure S1 of Supporting Information we present the calculations with up to 19 carbons for the remaining basis sets.

From Figure 4 it is evident that DLPNO-CCSD(T) is an excellent approximation to CCSD(T) with deviations of less than 0.4 kJ/mol. As we show in Supporting Information Figure S1 the deviations for longer chains become larger, as they should be due to the additivity of error for longer chains. The purpose of the extrapolation is to even eliminate these small deviations. At the aug-cc-pVDZ level the energy difference between CCSD(T) and DLPNO-CCSD(T) is essentially the same with the one at aug-cc-pVTZ, and since aug-cc-pVTZ is a sufficiently large basis set, we do not expect any significant change as the basis sets further approach saturation. This means that if one uses aug-cc-pVDZ for term A, the expected error is negligible. In order to be on the safe side we will estimate a maximum error of -0.2 kJ/mol coming from term A if we use the aug-cc-pVDZ basis set to calculate it. We note that a negative deviation here means that the method M overstabilizes the extended conformer with regard to the hairpin one with respect to the reference values.

Next we turn our attention to term B. This term does not depend on CCSD(T) but only on the method M, in our case DLPNO-CCSD(T), and describes how converged the EP1 extrapolated result is with respect to the M's basis set limit. In Supporting Information Figure S2 we show the convergence for the $E_{\text{hairpin}} - E_{\text{extended}}$ conformations using the MP2 method. In Supporting Information Figure S3 we show the error convergence with respect to EP1 extrapolated MP2 energies using cc-pVTZ and cc-pVQZ energies. There we see that the maximum deviation between aug-cc-pV(T/Q)Z and cc-pV(T/Q)Z extrapolated values is below 1 kJ/mol and only happens at $\text{C}_{19}\text{H}_{40}$, which is already outside of the candidate range for representing the critical chain length. For $\text{C}_{17}\text{H}_{36}$ the error is less than 0.75 kJ/mol. We should again note here that a positive deviation means that the hairpin conformer is overstabilized compared to the hairpin one with respect to the reference values. It is also noticed that terms A and B seem to have opposite sign pointing to an additional cancellation of errors that would decrease the final deviations from the canonical CCSD(T)/CBS results even more.

Using DLPNO-CCSD(T), the same quantities were calculated up to cc-pVQZ. We show the convergence in the isomerization energy in Supporting Information Figure S4 and the error convergence in Supporting Information Figure S5. It is noticed that the pattern of convergence is exactly the same between MP2 and DLPNO-CCSD(T). Hence this provides additional evidence for assuming the same error for EP1-(DLPNO-CCSD(T); cc-pVTZ, cc-pVQZ).

Thus, from term A we expect a negative error of up to 0.2 kJ/mol while for term B term we expect a positive error of up to 1 kJ/mol. Overall this means that if one uses aug-cc-pVDZ for term A and cc-pVTZ/cc-pVQZ for term B, the final electronic energy estimate should at most lead to an overstabilization of the extended conformer of less than 1 kJ/mol. For the region of 17 carbons this maximum error estimate should drop to around 0.5 kJ/mol with respect to real CCSD(T)/CBS result and on the side of overstabilizing the hairpin conformer.

3.3. Best Electronic Structure Results. In Figure S6 of the Supporting Information we present the electronic part of the $E_{\text{hairpin}} - E_{\text{extended}}$ energy difference, using various methods and also including the two most recent ab initio theoretical studies. In order to make it easier to read in Figure 5 we cut out the interesting region ranging from 14 to 18 carbons. Our best result is calculated with the EP2* scheme as described before with the DLPNO-CCSD(T) method. As pointed out above, we

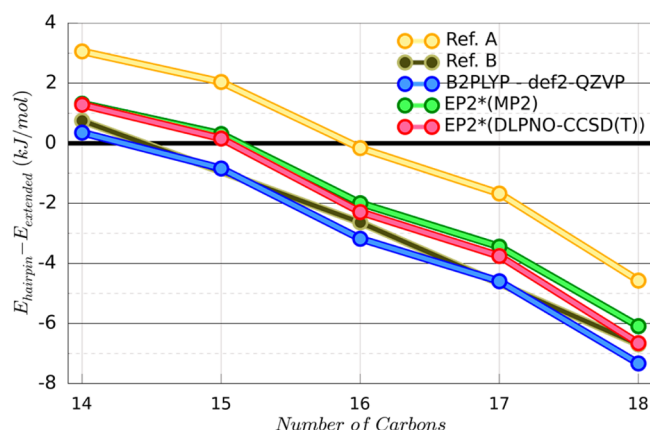


Figure 5. $E_{\text{hairpin}} - E_{\text{extended}}$ with various methods. Reference A: Lüttswager et al.⁷ Reference B: Byrd et al.⁸ EP2*(MP2): EP2(MP2; aug-cc-pVTZ, aug-cc-pVQZ, aug-cc-pVDZ). EP2*(DLPNO-CCSD(T)): EP2(DLPNO-CCSD(T); cc-pVTZ, cc-pVQZ, aug-cc-pVDZ).

do expected these electronic energies to be accurate within ~ 0.5 kJ/mol of the canonical CCSD(T)/CBS limit. This error bar is important to keep in mind, given that the electronic energy drop per carbon appears to be only 2–3 kJ/mol.

It is instructive to compare the DLPNO-CCSD(T) results with other electronic structure methods. First, using the EP2*(MP2) scheme, the MP2 energy was extrapolated using the aug-cc-pVTZ/aug-cc-pVQZ basis sets. Then we calculated the CCSD(T) – MP2 energy difference in conjunction with the aug-cc-pVDZ basis set. Finally we added these two terms. It appears that its performance is very good, giving results that are essentially identical with the DLPNO-CCSD(T) ones, showing that MP2 itself should not have any significant problems in this case. Second, the B2PLYP-D3 method has been suggested to give results close to CCSD(T) when the def2-QZVP basis set is used.⁵¹ Indeed we see that its performance is excellent with deviations from our best results of less than 1 kJ/mol and in very close agreement with the results of Byrd et al. There is a minor tendency to overstabilize the hairpin conformation.

The results from the work of Byrd et al.⁸ in the interesting region are within 1 kJ/mol of our reference calculations. Of course they do not have the results for $\text{C}_{15}\text{H}_{32}$ and $\text{C}_{17}\text{H}_{36}$, but given the agreement with $\text{C}_{14}\text{H}_{30}$, $\text{C}_{16}\text{H}_{34}$, and $\text{C}_{18}\text{H}_{38}$ there is no reason to expect that something drastically different should happen for the missing members of the series.

Concerning the work of Lüttswager et al.⁷ a more significant deviation of about 2 kJ/mol exists. While this is a rather limited deviation in absolute terms, it does matter for the case of weak interactions. The most obvious cause of the deviations is the rather significant deviations in the optimized geometries. The authors used a local MP2 method with the cc-pVTZ basis set to optimize the structures. As shown before, Byrd et al. also used the MP2 method with the cc-pVTZ basis set, but the geometries turned out to be different. The LMP2 method that Lüttswager et al. used appears to underestimate the weak interactions, thus producing more “open” structures for the hairpin conformer and thus destabilizing the hairpin with respect to the extended conformers.

3.4. Thermodynamical Corrections. In Figure 6 we show the thermodynamic corrections obtained in the two previous studies and the current one for the interesting region C14 to C18. In Figure S7 of the Supporting Information the same results for the whole range of molecules is documented. For the

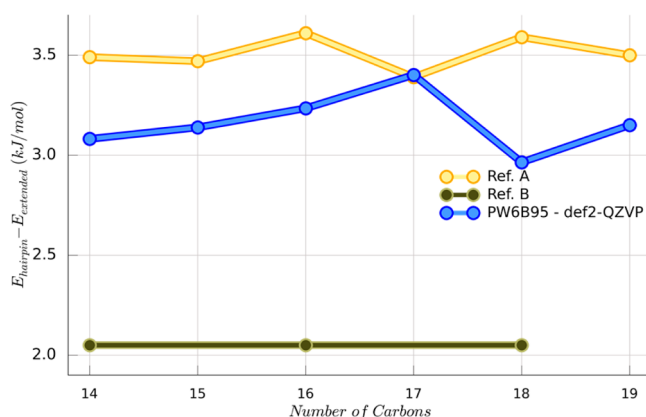


Figure 6. Thermodynamical correction with various methods. Reference A: Lüttswager et al.⁷ Reference B: Byrd et al.⁸ PW6B95-def2-QZVP: ZPE + U (100 K) – 100 × S(100 K).

calculation of the PW6B95-D3/def2-QZVP zero-point energy (ZPE) correction we used the scaling factor 0.970 proposed by Truhlar and co-workers.⁶⁰ This scaling factor was proposed for a different basis set (6-31+G(d,p)), but given that the scaling factors of DFT methods do not vary significantly with the basis set, we felt that this value would be closer to the experimental one compared to the unscaled one. In any case, the maximum difference between the scaled and the unscaled ZPE is 0.25 kJ/mol in our calculation, and thus not significant.

The correction includes contributions from the ZPE and thermal contributions, as well as contributions due to vibrational and rotational entropy. For the vibrational entropy we used the model proposed by Grimme that treats low-lying vibrational modes by a free-rotor approximation.⁵¹ The suggested values of 100 cm^{-1} for ω_0 and 4.0 for α were used. Again the difference between using the model and not using it in our case has no significant effect, the only exception being C16 where there is a difference of 0.5 kJ/mol. In all other cases the difference is less than 0.2 kJ/mol.

From Figure 6 we see that the thermodynamic corrections in this study lie in between those predicted by the two previous studies. Compared to the work of Lüttswager et al. the values obtain in this work differ by only 0.5 kJ/mol in favor of the linear conformer. These very limited deviations probably also arise from the differences in the geometries together with the use of unscaled ZPE values in their study. The difference with the thermodynamic correction of Byrd et al. is larger; the origin of the deviations is not clear to us. In any case we do believe that our thermodynamic corrections should not have an error of larger than 1–1.5 kJ/mol.

3.5. Total Energies. Finally in Figure 7 we plot the result with the total energies in the region C14 to C18 and in Supporting Information Figure S8 for the whole range of molecules. For $\text{C}_{15}\text{H}_{32}$ the linear conformer is more stable by around 4 kJ/mol. Given that the expected error bar for our calculations is below 2 kJ/mol, we believe we can safely predict that at 100 K the linear conformer should be the most stable conformation. $\text{C}_{18}\text{H}_{38}$ clearly presents the opposite behavior, in this case the hairpin conformer is more stable by about 4 kJ/mol and thus we conclude that this is the more stable conformer. In between these two cases we have $\text{C}_{16}\text{H}_{34}$ and $\text{C}_{17}\text{H}_{36}$. Both of them are inside our conservative error range, but $\text{C}_{16}\text{H}_{34}$ seems to be marginally safer as a linear conformer. In any case, the last linear conformer at 100 K should be either

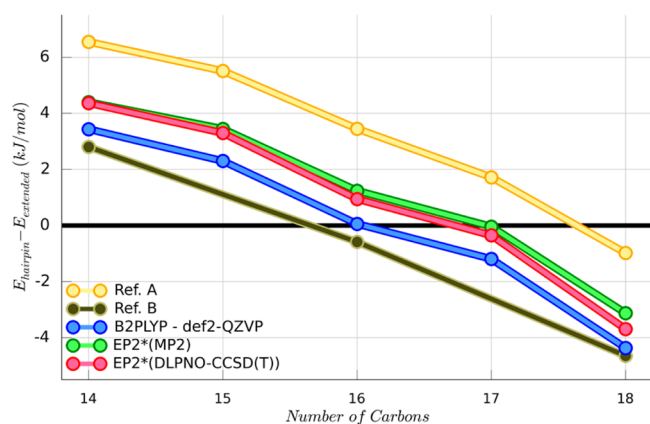


Figure 7. $E_{\text{hairpin}} - E_{\text{extended}}$ with various methods Reference A: Lüttswager et al.⁷ Reference B: Byrd et al.⁸ EP2*(MP2): EP2(MP2; aug-cc-pVTZ, aug-cc-pVQZ, aug-cc-pVDZ). EP2*(DLPNO-CCSD(T)): EP2(DLPNO-CCSD(T); cc-pVTZ, cc-pVQZ, aug-cc-pVDZ).

$\text{C}_{16}\text{H}_{34}$ or $\text{C}_{17}\text{H}_{36}$ with $\text{C}_{17}\text{H}_{36}$ being most probable. These results agree well with the experimental result that “ $n_c = 18$ assignment may just be an upper bound”.⁷ The results also agree with the two previous theoretical results to “a prediction of $n_c = 17 \pm 1$ ”⁷ and “... hairpin preference takes place at $n \geq 16$ ”.⁸

For completeness we also added the results calculated with EP2*(MP2) and B2PLYP-3D/def2-QZVP where we added our thermodynamic correction.

4. CONCLUSIONS

In this study we have revisited the question of what is the last globally stable extended alkane at 100 K from a theoretical perspective. Using a combination of the recently published DLPNO-CCSD(T) method and a tailored basis set extrapolation scheme, we were able to treat the electronic problem to an accuracy of better than 1 kJ/mol relative to the CCSD(T) basis set limit. In addition the most comprehensive thermodynamic correction treatment to date was obtained using large basis set DFT results. We did not treat the configurational entropy problem, but since our results refer to 100 K, we believe that its contribution should be negligible.

Overall we predict that the last nonbranched alkane that prefers to be in the linear conformation is either $\text{C}_{16}\text{H}_{34}$ or $\text{C}_{17}\text{H}_{36}$ with the latter being the most probable. These results agree well with the experimental findings and are in good agreement with previous theoretical works on the subject. The limited deviations between our and previous theoretical results have been carefully analyzed.

■ ASSOCIATED CONTENT

Supporting Information

Figures S1–S8 as well as the Cartesian coordinates for all structures used in this article. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00265.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: frank.neese@cec.mpg.de.

Notes

The authors declare no competing financial interest.

■ REFERENCES

- (1) Schleyer, P. v. R.; Williams, J. E.; Blanchard, K. R. *J. Am. Chem. Soc.* **1970**, *92*, 2377–2386.
- (2) Herrebout, W. A.; van der Veken, B. J.; Wang, A.; Durig, J. R. *J. Phys. Chem.* **1995**, *99*, 578–585.
- (3) Grimme, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460–4464.
- (4) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M. *J. Chem. Phys.* **2006**, *124*, -.
- (5) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. *Org. Lett.* **2006**, *8*, 3631–3634.
- (6) Balabin, R. M. *J. Phys. Chem. A* **2009**, *113*, 1012–1019.
- (7) Lüttswager, N. O. B.; Wassermann, T. N.; Mata, R. A.; Suhm, M. A. *Angew. Chem., Int. Ed.* **2013**, *52*, 463–466.
- (8) Byrd, J. N.; Bartlett, R. J.; Montgomery, J. A. *J. Phys. Chem. A* **2014**, *118*, 1706–1712.
- (9) Pitzer, K. S. *Discuss. Faraday Soc.* **1951**, *10*, 66–73.
- (10) Flory, P. J. *Science* **1975**, *188*, 1268–1276.
- (11) Krishnan, R.; Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4244–4245.
- (12) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *Int. J. Quantum Chem.* **1988**, *34*, 377–382.
- (13) Tsuzuki, S.; Schafer, L.; Goto, H.; Jemmis, E. D.; Hosoya, H.; Siam, K.; Tanabe, K.; Osawa, E. *J. Am. Chem. Soc.* **1991**, *113*, 4665–4671.
- (14) Frey, R. F.; Cao, M.; Newton, S. Q.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1993**, *285*, 99–113.
- (15) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968–5975.
- (16) Lee, T. J.; Rendell, A. P.; Taylor, P. R. *J. Phys. Chem.* **1990**, *94*, 5463–5468.
- (17) Smith, G. D.; Jaffe, R. L. *J. Phys. Chem.* **1996**, *100*, 18718–18724.
- (18) Allinger, N. L.; Fermann, J. T.; Allen, W. D.; Schaefer, H. F., III. *J. Chem. Phys.* **1997**, *106*, 5143–5150.
- (19) Goodman, J. M. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 876–878.
- (20) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.
- (21) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- (22) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
- (23) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.
- (24) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (25) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (26) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
- (27) Thomas, L. L.; Christakis, T. J.; Jorgensen, W. L. *J. Phys. Chem. B* **2006**, *110*, 21198–21204.
- (28) Hoffmann, R. W. *Angew. Chem., Int. Ed.* **2000**, *39*, 2054–2070.
- (29) Salam, A.; Deleuze, M. S. *J. Chem. Phys.* **2002**, *116*, 1296–1302.
- (30) Klauda, J. B.; Brooks, B. R.; MacKerell, A. D.; Venable, R. M.; Pastor, R. W. *J. Phys. Chem. B* **2005**, *109*, 5300–5311.
- (31) Klauda, J. B.; Pastor, R. W.; Brooks, B. R. *J. Phys. Chem. B* **2005**, *109*, 15684–15686.
- (32) Karton, A.; Gruzman, D.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 8434–8447.
- (33) Neese, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (34) Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753–12762.
- (35) Dunning, J. T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (36) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (37) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.
- (38) Hampel, C.; Werner, H. J. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

- (39) Schutz, M.; Manby, F. R. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3349–3358.
- (40) Werner, H.-J.; Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2003**, *118*, 8149–8160.
- (41) Werner, H.-J.; Schütz, M. *J. Chem. Phys.* **2011**, *135*, 144116 .
- (42) Möller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.
- (43) Martin, J. M. L. *J. Phys. Chem. A* **2013**, *117*, 3118–3132.
- (44) Grimme, S.; Steinmetz, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042.
- (45) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108–034116.
- (46) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
- (47) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104–154119.
- (48) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (49) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *122*, 154104.
- (50) Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2005**, *123*, 024101 .
- (51) Grimme, S. *Chem.–Eur. J.* **2012**, *18*, 9955–9964.
- (52) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023–5063.
- (53) Sosa, C.; Geertsens, J.; Trucks, G. W.; Bartlett, R. J.; Franz, J. A. *Chem. Phys. Lett.* **1989**, *159*, 148–154.
- (54) Klopper, W.; Noga, J.; Koch, H.; Helgaker, T. *Theor. Chem. Acc.* **1997**, *97*, 164–176.
- (55) Taube, A. G.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *128*, 164101.
- (56) Liakos, D. G.; Neese, F. *J. Phys. Chem. A* **2012**, *116*, 4801–4816.
- (57) Klopper, W.; Kutzelnigg, W. *J. Mol. Struct. (THEOCHEM)* **1986**, *135*, 339–356.
- (58) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- (59) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.
- (60) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 2872–2887.