What Is the Longest Unbranched Alkane with a Linear Global Minimum Conformation?

Jonathan M. Goodman

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW

Short unbranched alkanes, such as heptane and hexane, prefer extended conformations. Polythene chains prefer folded conformations. How long a chain can be built before the global minimum conformation ceases to be linear? The question is a difficult one, because exhaustive conformation searches on molecules of the size where the transition is likely to occur are very time consuming. This study demonstrates that the answer is not more than 17 for the MM2* force field and that the question provides a sensitive measure of comparison between force fields, solvent models, and semiempirical methods.

INTRODUCTION

It is well-known that short, unbranched alkanes tend to have linear ground state conformations. However, longer alkanes do not. In the course of another study, a conformation of $C_{39}H_{80}$ was minimized and found to have a lower energy than the extended conformation of the same molecule (Figure 1). At what chain length does the change from a linear global minimum, such as that found for hexane, to a nonlinear one (which $C_{39}H_{80}$ must have, even though there is no reason to believe that the illustrated conformation is the global minimum) occur? What is the longest unbranched alkane which has a linear global minimum? We have found by chance that it is less than 39 carbon atoms long, but exactly how long is it?

It is important that the conformational properties of open-chain compounds are understood, in order to design compounds of with particular spatial properties[1] and to understand the factors controlling the folding of peptides and other biological polymers. Unbranched alkanes are the simplest examples of this class of molecules, so it might be hoped that it would be possible to calculate their properties accurately.

RESULTS

The first approach to the problem was to use exhaustive conformation searching. MacroModel[2] and the MM2* force field[3] were used to perform a conformation search (10 000 Monte Carlo steps) on nonane which took 9 h on an R4400 Silicon Graphics Indigo. Two hundred ninety-four structures were found within 50 kJ mol^{-1} of the global minimum energy structure, which turned out to be linear. A similar search on decane took 11 h and found 927 structures within 50 kJ mol^{-1} of the global minimum energy structure, which was also linear. Continuing this approach seemed unlikely to find the longest linear alkane, because the conformation search for each longer chain would be increasingly difficult. An alternative strategy is to start from the shortest chain for which the linear form is not the ground state ($C_{39}H_{80}$ at this

stage of the study) and reduce the size of this molecule by removing carbon atoms from the end of the chain. This strategy was unsuccessful. Removing a methyl group from the ends of the C39 chain generated structures which were higher in energy than the linear form of $C_{38}H_{78}$. Molecular dynamics was rather more effective, but a great deal of computer time was required to find low energy conformations. Neither of these approaches seemed promising for finding the longest linear alkane, which must be between 10 and 38 carbon atoms long.

The reason that the folded conformations are preferred for longer alkanes is that the attractive van der Waals forces outweigh the energetic cost of twisting the carbon chain away from the preferred extended conformation. Twisting a single carbon−carbon bond in the middle of a linear molecule from 180° to 60° has an energetic cost is about 3.7 kJ mol^{-1} (MM2*). The favorable van der Waals term for to alkane chains placed adjacent to each other is about 4.1 kJ mol^{-1} per methylene unit (from calculations on the pentane, decane, $C_{15}H_{32}$, $C_{20}H_{42}$, $C_{25}H_{52}$, $C_{30}H_{62}$, and $C_{35}H_{72}$ dimers).

The smallest number of twists to reverse the chain is two, which creates an unfavorable *syn*-pentane arrangement, at the cost of 13 kJ mol^{-1}. This arrangement does not produce two parallel chains. A combination of four twists, however, will reverse the chain without causing this steric clash and aligns the adjacent chains rather well. Six atoms are involved in this turn, which costs four times as much as a single twist (14.8 kJ mol^{-1}). Four optimally arranged methylene pairs should be required to balance the energy required for the twist, suggesting that the maximum length of a linear alkane should be something like $C_{14}H_{30}$. In fact, the first alkane for which the hairpin conformation is lower in energy than the linear one is $C_{18}H_{38}$.

A conformation search and a molecular dynamics simulation on $C_{17}H_{36}$ failed to find any conformations lower than the extended conformation. This cannot be regarded as a certain proof but is probably a strong indication that the preferred conformation of the global minimum structure ceases to be linear on changing from $C_{17}H_{36}$ to $C_{18}H_{38}$.

A similar analysis can be performed with the MM3* force field.[4] These suggest that the longest linear alkane should

---

LONGEST UNBRANCHED ALKANE

J. Chem. Inf. Comput. Sci., Vol. 37, No. 5, 1997 **877**
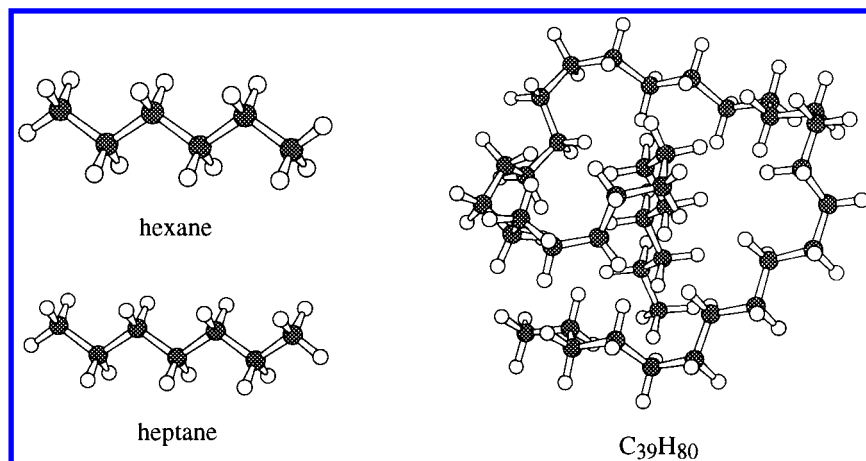

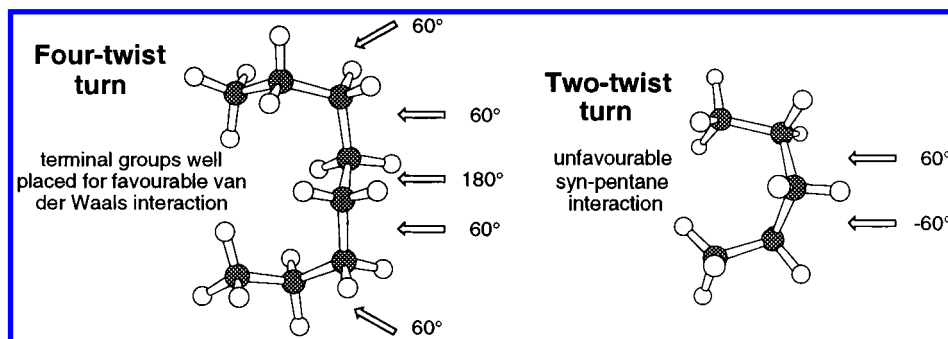**Figure 1.** Conformaitons of some unbranched alkanes.

Figure 2. A turn with four twists creates a better alignment of the terminal groups than a two twist turn.

be $C_{24}H_{50}$. The AMBER* force field[5] favors $C_{25}H_{52}$ as the longest.

These analyses ignore the effects of solvent. These were investigated using the continuum water and chloroform models[6] of MacroModel and the AMBER* force field. Water may be expected to compress the system, favoring the smallest possible surface area for an alkane, and so reduce the length of the longest chain with a linear ground state. Consistent with this expectation, the longest linear chain is $C_{21}H_{44}$, four carbon atoms shorter than the crossover point for the AMBER* calculation without the continuum solvent correction.

Chloroform, a much less polar solvent than water, may have the opposite effect. The choice between a favorable van der Waals interaction with the chain and itself or with the chain and water may be evenly balanced. It turns out that the greatest chain length for which the linear structure is lower in energy than the hairpin is $C_{142}H_{286}$. If a model for hexane, or some other hydrocarbon solvent, were available, it is likely that the hairpin form would never be favored, because the solvent would provide a favorable van der Waals attraction along the whole length of the alkane chain.

Semiempirical methods were also investigated, using VAMP[7] and MOPAC[8] which gave consistent results. PM3[9] calculated the longest linear alkane to be $C_{11}H_{24}$ which is rather shorter than is expected from the force-field results. AM1,[10] which is a rather similar semiempirical method, gives dramatically different results. The longest linear alkane is now $C_{59}H_{120}$, which is considerably longer than expected from the force-field results and remarkably different to the

Table 1. Shortest Chains Known To Have Nonlinear Global Minima

class	model	shortest chain
isolated	MM2	$C_{18}H_{38}$
molecules	MM3	$C_{25}H_{52}$
	AMBER	$C_{26}H_{54}$
solvent effects	AMBER/water	$C_{22}H_{46}$
	AMBER/chloroform	$C_{143}H_{288}$
semiempirical	PM3	$C_{12}H_{26}$
molecular orbital	AM1	$C_{60}H_{122}$
theory	MNDO	no limit

PM3 result. The relative merits of these two semiempirical molecular orbital methods have been debated in the literature, but this debate has not centered on the key parameterizations of carbon and hydrogen, for which the properties of the two methods are rather similar.[11] The earlier model MNDO[12] was now tried, and the results suggested the opposite trend to all others: the hairpin structure becomes increasingly unfavorable as the chain length increases. This demonstrates how MNDO tends to overestimate the repulsion between nonbonded atoms at around their van der Waals separation.[13] This deficiency was corrected in AM1 and PM3 by the addition of an extra term to the expression for core−core repulsion. The different parameterizations have a dramatically different effect on the global minima of hydrocarbons.

A molecular dynamics simulation (MM2*, 1 ns simulation at 300 K; 1000 structures saved and reminimized; one week's computer time on an R3000 Silicon Graphics Indigo) on $C_{39}H_{80}$ found a lower energy structure than the hairpin. This shows some resemblance to a paperclip and is illustrated in Figure 4. The complexity of these systems show that the conformation analysis of linear alkanes is still mysterious.
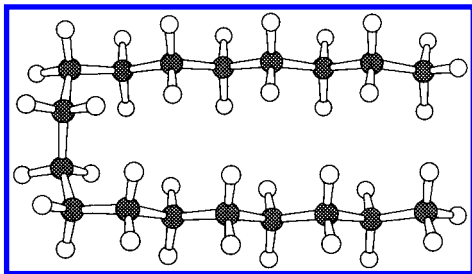
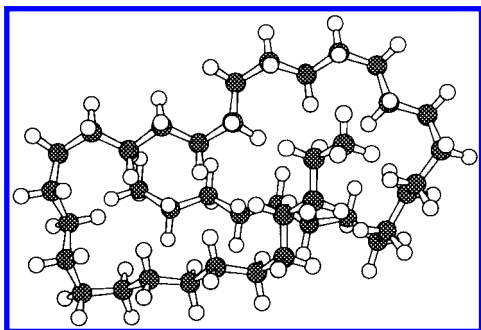**Figure 3.** C$_{18}$H$_{38}$, the global minimum? (MM2*).

Figure 4. Lowest known minimum for C$_{39}$H$_{80}$ (MM2*).

CONCLUSION

How long is the longest linear unbranched alkane? Conformation searching is not a useful approach to solving this problem, but the introduction of favorable motifs, such as a hairpin turn created by twisting four torsion angles, can provide an answer rapidly. These calculations demonstrate that C$_{18}$H$_{38}$ does not have a linear ground state according to the MM2* force field, and it is probably the shortest chain with this property. Other force fields produce similar results. A polar solvent will favor the folded structures, but a less polar solvent will increase the preference for extended forms. Comparison of extended and folded conformations highlights the differences in parameterization of AM1, PM3, and MNDO. This study has implications for the folding of other long chains of atoms, such as proteins.

REFERENCES AND NOTES

(1) Göttlich, R.; Kahrs, C. B.; Krüger, J.; Hoffmann, R. W. Open Chain Compounds with Preferred Conformations. *Chem. Commun.* **1997**, 247−251.
(2) Mohamedi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. Macro-Model- an Integrated Software System for Modeling Organic and Bioorganic Molecules using Molecular Mechanics. *J. Comput. Chem.* **1990**, *11*, 440−467.
(3) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127−8134.
(4) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular Mechanics−The MM3 Force-Field for Hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566.
(5) (a) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. A New Force-Field for Molecular Mechanical Simulation of Nucleic-Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765−784. (b) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force-Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230−252.
(6) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semi-analytical Treatment of Solvation for Molecular Mechanics and Dynamics *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(7) Vamp V5.5; Oxford Molecular Ltd.: Magdalen Centre, Oxford Science Park, Sandford-on-Thames, Oxford, OX4 4GA.
(8) MOPAC version 6.0. Stewart, J. J. P. MOPAC- a SemiEmpirical Molecular Orbital Program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−45.
(9) Stewart, J. J. P. Optimization of Parameters for SemiEmpirical Methods. 1. Method. *J. Comput. Chem.* **1989**, *10*, 209−220.
(10) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. The Development and Use of Quantum-Mechanics Molecular-Models. 76. AM1-A New General-Purpose Quantum-Mechanical Molecular-Model *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(11) Stewart, J. J. P. Optimization of Parameters for SemiEmpirical Methods. 2. Applications. *J. Comput. Chem.* **1989**, *10*, 221−264.
(12) Dewar, M. J. S.; Thiel, W. Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.
(13) Dewar, M. J. S.; Jie, C.; Yu, J. SAM1. The First of a New Series of General Purpose Quantum Mechanical Molecular Models. *Tetrahedron* **1993**, *23*, 5003−5038.

CI9704219