Mindless Chemistry

Partha P. Bera, Kurt W. Sattelmeyer, Martin Saunders,[§] Henry F. Schaefer III, and Paul v. R. Schleyer*

Center for Computational Chemistry, University of Georgia, 1004 Cedar Street, Athens, Georgia 30605-2525

Received: December 6, 2005

Applications of an automated stochastic search procedure for locating all possible minima with a given composition are illustrated by the pentatomic molecules BCNOS, CAlSiPS, C_4B^-, C_4Al^-, and CBe_4^{2-}, as well as by C_6Be, the C_6Be^{2-} dianion, and C_6H_2. All previously identified minima were reproduced, and many new structures, often with nonintuitive geometries, were found.

After more than two centuries, chemists are still not able to predict the structures of all isomers with a given composition logically, thoroughly, and reliably. This is not surprising. As the number of atoms composing a molecule increases, the possible connectivities scale factorially! Hundreds, thousands, and then countless minima become conceivable. Experience, knowledge, and chemical intuition may offer suggestions, but the task of enumerating and establishing the relative stability of all possible structures rapidly becomes ill-suited and unfeasible for humans. Most chemists hesitate even to venture a guess as to what might be the most stable isomer of a compound with an unfamiliar composition.

Such problems can be daunting even for small molecules, especially when all atoms are different. Such species are seldom considered. Even small clusters with more than two elements are rare. Sun et al.[2−5] employed manual exploration techniques to find stable isomers of intriguing interstellar molecule candidates, such as SiC_2N, SiC_2P, SiC_3N, SiC_3P, and NC_3S. Each search began by choosing five or six possible connectivities for each molecule. Chemical intuition was used to guess all possible isomers with each of these connectivities. The task of evaluating all of them must have been very tedious and time-consuming.

Predicting the lowest-energy structure of a hypothetical pentatomic molecule, composed of five different, but familiar, atoms, e.g., B, C, N, O, S, would be an order of magnitude more complex! A thorough listing of all the possible connectivities must include those illustrated in Figure 1, along with all the associated permutations of the various atoms. Their geometric and conformational isomers may also be minima. The challenge of finding all the isomers of this and other examples is addressed in the present paper.

While conceivable, the manual exploration of each of the many hundred structural possibilities for BCNOS suggested by Figure 1 using quantum chemistry methods is an uninviting task, uncertain in its outcome. The rigorous manual solution of the next largest six unique atom puzzle, e.g., BCNOPS, surely would be precluded by the enormous labor required.
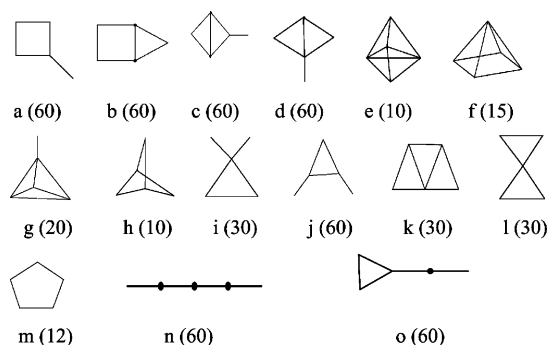
Figure 1. Connectivities of pentatomic molecules; the conformations also may vary. The number of possible permutations when all five atoms are different, as in BCONS and CAlSiPS, are given in parentheses.

There have been many efforts to deal with problems of similar complexity using automated procedures, but generally with the objective of finding the global minimum. In 2004, Zhao and Xie reviewed comprehensively the genetic algorithm-based methods employed to determine the most stable geometry of clusters.[6] R. Car and M. Parrinello's well-known "dynamic simulated annealing" combines molecular dynamics (MD) and density functional theory (DFT).[7] Searches employing "simulated annealing", "Monte Carlo metropolis search", and various "genetic algorithm techniques" are also well-documented.[7−10] Tomasulo and Ramakrishna gradually replaced the atoms of a starting structure one by one[11] and used the Car−Parrinello technique to relax the atomic positions and optimize electronic wave functions. However, most of the structures obtained this way could be anticipated easily, since only the relative positions of the atoms change. Babadova-Parvanova et al. employed the density functional tight binding level to search for low-energy isomers using the "single parent genetic algorithm" (DFTB/SPGA) method.[12] Bertolus et al. combined Car−Parrinello molecular dynamics with local energy minimization using DFT;[13] several molecular dynamics trajectories generated a large number of structures, e.g., several interesting Si_mC_n clusters. R. O. Jones has employed density functional methods for many years to study the structure and bonding of elemental as well as mixed element clusters.[14−16] Starting geometries were gener-

---

* Corresponding author. E-mail: schleyer@chem.uga.edu.
§ Department of Chemistry, Yale University, New Haven, CT 06520 (U.S.A.).

ated by either MD simulated annealing or density functional tight binding or selected from DFT-based computations or experimental results. Choosing starting geometries becomes more and more complex when the number of atoms increases. Even though their energies were lowest or competitive, hollow cage isomers of gold clusters were not located in a recent genetic algorithm search.[17] Alexandrova et al. used genetic algorithm-based techniques on randomly generated structures to search for global minima.[18] Godecker in 2004 proposed a "minima hopping" method to scan the entire potential energy surface.[19] Other methods such as "conformational flooding" address the same issue.[20] A major problem which bedevils any search method is that the potential energy surface (PES) at lower theoretical levels, employed initially for practical considerations, may differ from the PES at higher, more accurate levels. We have encountered cases where structures, important on higher-level PESs, do not exist as minima at lower levels!

The purpose of this paper is to present illustrative examples demonstrating how structures generated through stochastic methodology[21] can be used to explore PESs having many minima automatically with the least possible human involvement. Truly rigorous searches require knowledge of all possible isomers to be certain that the global minimum has been located.

Fortunately, Saunders's fast, comprehensive, and automated "kick" method produces isomers and conformers simply and effectively with little thought or effort.[21] A similar random search and minimization technique was proposed by Lloyd and Johnston while exploring aluminum clusters.[22] The automated random structure generation procedure implemented here enables searches for unknown isomers much more easily than manual explorations. Neither preconceived structures nor bonding principles are involved! All the atoms are placed at the same point initially and then are "kicked" randomly within a box of chosen dimensions, e.g., a 2 Å cube. The method developed further in this work generates up to 1000 unbiased starting geometries and submits them automatically for optimization with, e.g., *Gaussian 94*,[23] as employed here, or other electronic structure programs, to a bank of rapid, coarse-grained parallelized PCs. While many jobs die quickly (the atoms are too close, or the self-consistent field (SCF) convergence is not achieved) and other jobs do not achieve geometrical convergence, the "yield" of completed optimizations typically ranges between 10% and 50%. Most of these completed jobs from initial batches of, e.g., 100 "kick runs" are unique minima in the examples we have explored! (If desired, promising partially optimized jobs may be completed in the usual manner, and other search procedures, e.g., utilizing the gradient and Hessian information, may be employed. However, this requires human effort. It is easier just to run more "batches".)

Subsequently, an auxiliary program searches each batch of outputs for successful completion (convergence to a minimum), eliminates redundancies, orders the energies, and lists the archive summaries of the unique isomers. Selected structures (e.g., the most stable or otherwise attractive isomers) obtained in this manner at lower levels of theory may then be refined manually by employing more sophisticated methods. In essence, lower-level searches generate "candidates" for refinement at higher levels rapidly. All "kick" optimizations proceed in point group C_1, but often lead very nearly to higher symmetries, which may then be imposed and the harmonic frequencies computed.

To ensure that all the minima have been located, additional batches of kick jobs are run until no new structures are obtained. While it is impossible to be absolutely certain that all minima have been found, this problem exists to a greater extent with
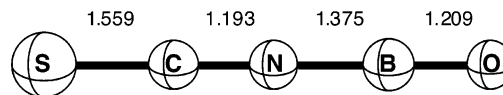
**Figure 2.** The linear BCONS global minimum geometry (in Å) at the cc-pVTZ/CCSD ab initio level.

other methods. Minima may lie beyond the resolution capabilities of the method employed, or the "box" dimensions may be too restrictive. However, the latter may be varied to increase the percentage of successful optimizations or the probability of finding structures with planar or elongated shapes. The purpose of this paper is to illustrate the method we have developed by applying it to several diverse problems. Singlet PESs of the following molecules were explored using the method explained above.

BCONS. In 1976, Baudisch synthesized a chemical novelty, a substituted five-membered BCNOS ring composed of 5 different atoms.[24] This feat inspired us to search for all the possible singlet pentatomic minima involving only these 5 familiar main group elements (without any substituents). Independent batches of 100 or 200 jobs employed the kick method at the respectable B3LYP/6-31G* level of theory. The kick size in the cubic box was varied from 2.0 to 2.5 Å. Since some atoms must move far to reach an equilibrium position, the number of optimization cycles permitted was increased (from the default). Ten PCs running simultaneously processed about 20 jobs per hour; roughly 30% of the jobs completed satisfactorily. Redundancies in each batch (energies within 0.000 01 au) were eliminated; duplicates appearing in more than 1 batch were noted; and the kick runs were continued until no new structures were generated. A total of 1000 kicks gave 103 unique geometries. The energies of these minima ranged over 314 kcal/mol!

The best isomers had type **n** (Figure 1) topologies. Surprisingly, linear SCNBO was the global minimum, despite its seemingly unfavorable S=C double bond. Linear OCNBS followed next, then NCSBO (bent at sulfur), NCOBS (bent at oxygen), linear SNCBO, and other permutations of these elements. Many isomers of type **o** (Figure 1) with a three-membered ring and a two-atom side chain constituted the second most stable structural group. Of these, planar NCB(OS) had the lowest energy, followed by CNB(OS) where the first three atoms form the linear tail (see type o, Figure 1).

As a check, we explored all 60 possible type **n** permutations manually; 28 of these survived as minima, but some were bent rather than linear. Others rearranged, and a few formed three-membered rings. All of these minima had been found previously by the kick procedure. Remarkable variety illustrating the connectivities of Figure 1 characterized the other, higher-energy structures and included a five-membered ring, cages, and bridged isomers, as well as examples of partially dissociated CO, NO, and CS complexes (with the corresponding three-atom fragments). The refinement and characterization of these minima using a higher level of theory (cc-pVTZ/CCSD) resulted in the same ordering of energies for the ten lowest-energy isomers (see Supporting Information Table 1). Figure 2 shows the geometry of the global minimum reoptimized at the cc-pVTZ/CCSD level of theory.

CAlSiPS. Our second example retained carbon but included 4 second-row elements known to eschew multiple bonding. Indeed, the results were quite different from BCNOS: There were no linear minima. The fourth 100-run batch did not result in any significant new isomers. Fifty minima in an 86 kcal/mol energy range were found. Oddly, **1** (type **c**) was the lowest-energy form (global minimum) at B3LYP/6-31G* (the CCSD/

Letters

J. Phys. Chem. A, Vol. 110, No. 13, 2006 **4289**
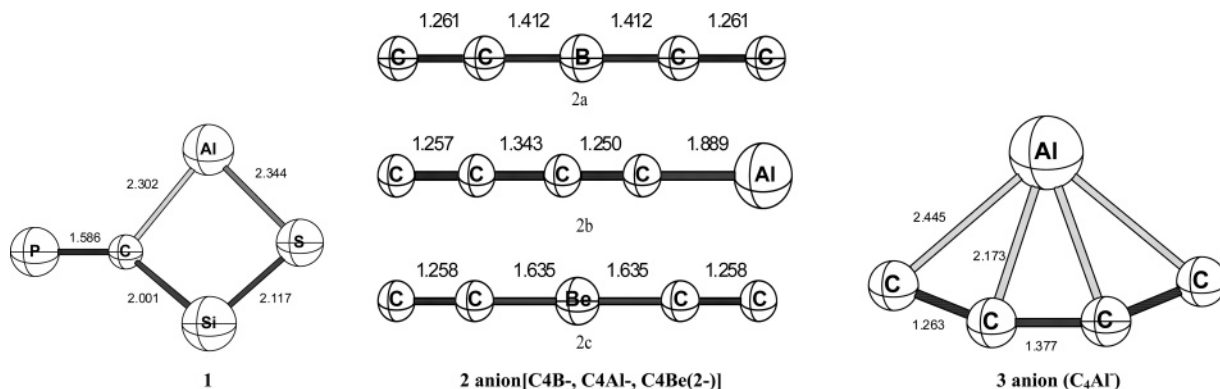
cc-pVTZ geometry is shown below). Type **o** structures were favored by the next several lower-energy isomers at B3LYP/6-31G*.

C_4B^-, C_4Al^-, and C_4Be^{2-}. The search for all the isomers of these isoelectronic species is inherently much less complex, since four atoms are the same. While the global minima of the C_4Be^{2-} and the "magic number" C_4B^- species have been predicted by analogy with the isoelectronic linear neutral C_5,[25−28] many of their isomers have been overlooked.

The C_4B^- anion signal was, by far, the most intense in a 1988 experiment on the cluster ions formed in the laser-ablation plasma of boron carbide.[29] The kick method located 12 singlet minima, ranging in energy by more than 127 kcal/mol at B3LYP/6-31G*. The best of these, CCBCC$^-$ (**2a**), is 30 kcal/mol more stable than its linear BCCCC$^-$ counterpart. The third and fifth best species (relative energies 60 and 78 kcal/mol, respectively) are of type **k** (Figure 1), with planar tetracoordinate boron and carbon. There are no analogous isoelectronic neutral C_5 or even C_4Be^{2-} minima, but a similar C_{2v} planar tetracoordinate aluminum structure (**3 anion**) is the second most stable C_4Al^- isomer. The C_4Al^- global minimum, linear CCCCAl$^-$, is 17 kcal/mol more stable; a D_{2d} type **l** structure is the only other low-energy isomer (**2b**).

Although small dianions in isolation are generally metastable toward Coulomb explosion, C_4Be^{2-} has been observed experimentally in the gas phase.[30] Dreuw and Cederbaum computed ten structures, but claimed that only the three linear forms, the lowest-energy CCBeCC^{2-} (**2c**) as well as BeCCCC^{2-} and CBeCCC^{2-}, were minima.[31] Shi and Kais reported that two other isomers (planar types **o** and **l**, with Be in the center of each) also were minima.[32]

The kick method, run at the B3LYP/6-31G* level some 500 times until no new minima appeared, revealed 12 isomers with energies ranging over 163 kcal/mol. (In contrast, only 5 minima were obtained at HF/6-31G*.) Besides the 3 linear forms, 3 were of type **o** and 2 of type **j**. Types **a** (external Be), **c**, **e** (trigonal Be), and **l** (D_{2d} rather than the reported planar form,[29] which has 3 imaginary vibrational frequencies) also were represented. Other singlet minima at B3LYP/6-31G* are unlikely.

C_6Be and C_6Be^{2-}. About 500 kick computations resulted in 33 C_6Be minima at the B3LYP/STO-3G level; only 15 of these survived when refined at B3LYP/6-31G*. A cyclic C_{2v} structure (**4**) was the most stable isomer. The structures of **4** and the **6** dianion, reoptimized at the cc-pVTZ/CCSD level of theory, are shown below.

Dreuw and Cederbaum[31] examined the experimentally known C_6Be^{2-}[30] and claimed that only "linear isomer[s]....correspond to minima on the potential energy surface". The kick method disagrees. Both the HF/STO-3G and B3LYP/6-31G* PESs have more than ten minima ranging over 120 kcal/mol. At the B3LYP

level, the C_{2v} $C_2BeCCCC^{2-}$ structure (**5 dianion**) is only 4 kcal/mol less stable the linear CCBeCCCC^{2-} global minimum (**6 dianion**).
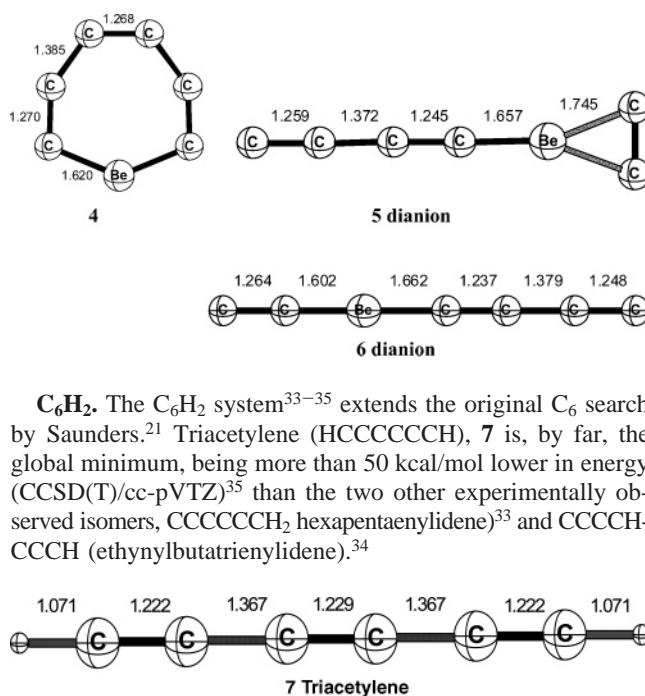
C_6H_2. The C_6H_2 system[33−35] extends the original C_6 search by Saunders.[21] Triacetylene (HCCCCCCH), **7**, is, by far, the global minimum, being more than 50 kcal/mol lower in energy (CCSD(T)/cc-pVTZ)[35] than the two other experimentally observed isomers, CCCCCCH$_2$ hexapentaenylidene)[33] and CCCCH-CCCH (ethynylbutatrienylidene).[34]

7 Triacetylene
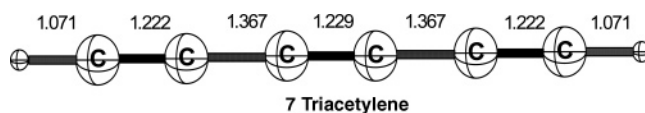
The two additional hydrogen atoms in C_6H_2 increase the number of minima from those reported for C_6[21] dramatically. Indeed, an earlier attempt "to investigate 'all' reasonable structures" was defeated by the "depressingly large"[35] number of structures which would have to be processed manually. Several thousand "kicks" located 80 C_6H_2 isomers, spanning a range of 175 kcal/mol at B3LYP/DZP. These included all 9 reported earlier.[33,34] Moreover, of the 32 isomers within 100 kcal/mol of triacetylene at the B3LYP/DZP level, 30 survived further refinement; all are minima at CCSD(T)/cc-PVTZ. As expected, structure **7** also has the lowest energy at this level.

In conclusion, the kick procedure facilitates unbiased and nearly automatic searches of potential energy surfaces. It is an important addition to the existing methods. Little thought and human effort are required, and numerous isomers not anticipated by chemical intuition are located. While no method can guarantee 100% recovery of all possible minima, this goal can be approached by repeating the "kick" procedure again and again. While low starting theoretical levels employ fewer computational resources and result in a large number of isomers, it may be better strategy to begin with higher levels, as these are less likely to be misleading.

4290 *J. Phys. Chem. A, Vol. 110, No. 13, 2006*

Letters

The often-employed genetic algorithm (GA) method[6,8,9] only attempts to locate the global energy minimum structure. In practice, this not always is achieved.[6,8,9,17] The concept of GA is inferior in comparison to methods that search for all the minima. However, the latter methods are not applicable currently to larger systems, which often are the focus of GA procedures. When two or more isomers have nearly the same energy, there is the further general problem that entropy differences may determine the lowest free energy. Many approaches, for example, systematic search, thermal annealing, distance geometry, and genetic algorithm,[6-16,18,22] have been promoted, but for systems of the moderate size we have discussed, none of these alternatives appear to be superior to the stochastic search procedure employed here.

Supporting Information Available: Cartesian coordinates and total energies (and selected list of structures for BCNOS) of the molecules depicted. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Bera, P. P.; Sattelmeyer, K. W.; Saunders: M.; Schaefer, H. F.; Schleyer, P. v. R. S. *Automated Stochastic Method to search for Isomers and Conformers*; Molecular Quantum Mechanics Conference: The No Nonsense Path to Progress, Cambridge University, England, 2004.

(2) Ding, Y.; Li, Z.; Huang, X.; Sun, C. *J. Phys. Chem. A* **2001**, *105*, 5896.

(3) Chen, G.; Ding, Y.; Huang, X.; Zhanng, H.; Li, Z.; Sun, C. *J. Phys. Chem. A* **2002**, *106*, 10408.

(4) Liu, H.; Huang, X.; Chen, G.; Ding, Y.; Sun, C. *J. Phys. Chem. A* **2004**, *108*, 6919.

(5) Liu, H.; Huang, X.; Chen, G.; Ding, Y.; Sun, C. *J. Phys. Chem. A* **2004**, *108*, 11828.

(6) Zhao, J.; Xie, R. *J. Comput. Theor. Nanosci.* **2004**, *1*, 117.

(7) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(8) Back, T. *Evolutionary Algorithm in Theory and Practice: Evolutionary Programming, Genetic Algorithms*; Oxford University Press: Oxford, 1990.

(9) Chaudhuri, P.; Bhattacharyya, S. P. *Chem. Phys.* **1999**, *241*, 313.

(10) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.

(11) Tomasulo, A.; Ramakrishna, M. V. *J. Chem. Phys.* **1996**, *105*, 10449.

(12) Babadova-Parvanova, P.; Jackson, K. A.; Srinivas, S.; Horoi, M.; Kohler, C.; Seifert, G. *J. Chem. Phys.* **2002**, *116*, 3576.

(13) Bartolus, M.; Finocchi, F.; Millie, P. *J. Chem. Phys.* **2004**, *120*, 4333.

(14) Jones, R. O.; Gantefor, G.; Hunsicker, S.; Pieperhoff, P. *J. Chem. Phys.* **1995**, *103*, 9549.

(15) Jones, R. O.; Lichtenstein, A. I.; Hutler, J. *J. Chem. Phys.* **1997**, *106*, 4556.

(16) Jones, R. O. *J. Chem. Phys.* **1999**, *110*, 5189.

(17) Wang, J.; Jellinek, J.; Zhao, J.; Chen, Z.; King, R. B.; Schleyer, P. v. R. *J. Phys. Chem. A* **2005**, *109*, 9265.

(18) Alexandrova, A.; Boldyrev, A.; Fu, Y.; Yang, X.; Wang, X.; Wang, L. *J. Chem. Phys.* **2004**, *120*, 5709.

(19) Goedecker, S. *J. Chem. Phys.* **2004**, *120*, 9911.

(20) Grubmuller, H. *Phys. Rev. Lett.* **1995**, *E52*, 1893.

(21) Saunders, M. *J. Comput. Chem.* **2004**, *25*, 621.

(22) Lloyd, L. D.; Johnston, L. *Chem. Phys.* **1998**, *236*, 107.

(23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*; Gaussian, Inc.: Pittsburgh, PA, 1995.

(24) Baudisch, H. University of Erlangan-Nuremberg, 1976.

(25) Galazutdinov, G.; Musaev, F.; Nirski, J.; Krelowski, J. *Mon. Not. R. Astron. Soc.* **2001**, *325*, 1332.

(26) Bernath, P. F.; Hinkle, K. H.; Keady, J. J. *Science* **1989**, *244*, 562.

(27) Motylewski, T.; Vaizert, O.; Giesen, T. F.; Linnartz, H.; Maier, J. P. *J. Chem. Phys.* **1999**, *111*, 6161.

(28) Galazutdinov, G.; Petlewski, A.; Musaev, F.; Moutou, C.; Lo Curto, G.; Krelowski, J. *Astron. Astrophys.* **2002**, *395*, 969.

(29) Becker, S.; Dietze, H. J. *Int. J. Mass Spectrom. Ion Processes* **1988**, *82*, 287.

(30) Klein, J.; Middleton, R. *Nucl. Instrum. Methods. Phys. Res., Sect. B* **1999**, *159*, 8.

(31) Dreuw, A.; Cederbaum, L. S. *J. Chem. Phys.* **2000**, *112*, 7400.

(32) Shi, Q.; Kais, S. *J. Am. Chem. Soc.* **2002**, *124*, 11723.

(33) Langer, W. D.; Velusamy, T.; Kuiper, T. B. H.; Peng, R.; McCarthy, M. C.; Travers, M. J.; Kovacs, A.; Gottlieb, C. A.; Thaddeus, P. *Astrophys. J.* **1997**, *480*, L63.

(34) McCarthy, M. C.; Thaddeus, P. *Astrophys. J.* **2002**, *569*, L55.

(35) Sattelmeyer, K. W.; Stanton, J. F. *J. Am. Chem. Soc.* **2000**, *122*, 8220.