4523

GLOBAL OPTIMIZATION OF PASSIVATED SI CLUSTERS AT THE AB INITIO LEVEL VIA SEMIEMPIRICAL METHODS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

CHEMISTRY

DECEMBER 2004

By

Yingbin Ge

Dissertation Committee:

John D. Head, Chairperson

Klaus D. Sattler

Karl Seff

Ralf I. Kaiser

Satoshi Takara

ACKNOWLEDGEMENTS

I wish to thank my research adviser Prof. John D. Head for his brilliant teaching, insightful research advice, endless patience, and trust throughout all the years of my graduate studies. I would also like to thank my dissertation committee members Prof. Klaus D. Sattler, Prof. Karl Seff, Prof. Satoshi Takara and Prof. Ralf I. Kaiser for their constructive and helpful comments and suggestions. I would also like to thank the Maui High Performance Computing Center (MHPCC) and the Univeristy of Hawaii's IBM SP system for providing ample computer time. I am also grateful to the Department of Chemistry at the University of Hawaii at Manoa for their financial support.

ABSTRACT

New strategies based on the evolutionary theory for globally optimizing binary molecular systems have been developed. We first determined the global minima of several Si_xH_y clusters using a genetic algorithm (GA) coupled with the fast AM1 semiempirical method to evaluate the cluster energies. However, we found that the AM1 and the ab initio energy rankings differ significantly. Therefore we proposed an improved iterative global optimization strategy which involved two separate genetic algorithms. The first algorithm is the cluster GA (CGA) that finds the Si_xH_y cluster global minimum. The second algorithm is the parametrization GA (PGA) that reparametrizes the AM1 method against ab initio data to produce the GA-optimized AM1 (GAM1) parameters. Convergence is considered achieved when the CGA stops producing new low energy clusters. However this method is very computationally demanding. Therefore we further examined whether the GAM1 parameters obtained for a small Si_xH_y stoichiometry could be successfully transferred to larger clusters. We have found that the GAM1 parameters derived from the Si_7H_{14} training set can be transferred to larger Si clusters such as $Si_{14}H_{20}$ with a similar level of H passivation. Unfortunately we still could not locate the diamond-lattice like $Si_{14}H_{20}$ global minimum since our original choice of genetic operators produce drastic structural changes which makes production of the lowest energy offspring difficult. Therefore we developed new genetic operators such as SiH_2 insertion/removal and H-shift operators to enhance the efficiency at finding low energy clusters. The CGA using the new genetic operators enabled us to locate the $Si_{14}H_{20}$ and other global minima that we had missed before. We found that fully passivated Si_xH_y clusters prefer diamond-lattice like structures while slightly under-

passivated Si_xH_{y-2} global minimum clusters adopt significantly different Si core. We also developed a GAM1* energy evaluation scheme to evaluate the relative energies for the Si_xF_y clusters, which enabled us to search the the Si_xF_y global minimum efficiently. Surprisingly the Si_xF_y clusters and Si_xH_y clusters prefer completely different Si_x core structures. In summary, we have successfully determined the global minima of various Si_xH_y and Si_xF_y clusters, which prepares the ground for further investigation of the optoelectronic properties of passivated Si clusters.

CONTENTS

ACKNOWLEDGEMENTS iii

ABSTRACT iv

LIST OF TABLES viii

LIST OF FIGURES ix

LIST OF ABBREVIATIONS xiv

1. INTRODUCTION 1

1.1 Unpassivated and Passivated Si clusters 1

1.2 Global Optimization Strategies 8

1.2.1 Global minimum 8

1.2.2 Simulated Annealing 10

1.2.3 Genetic Algorithm 11

1.3 Cluster Energy Calculation Methods 15

1.3.1 Hartree Fock Theory 16

1.3.2 Semiempirical Methods 21

1.4 Post-HF methods 25

1.4.1 Moller-Plesset Many-Body Perturbation Theory 25

1.4.2 Coupled Cluster Methods 28

1.4.3 Density Functional Theory Methods 30

1.5 Outline of the Dissertation 33

2. GLOBAL OPTIMIZATION OF H-PASSIVATED SI CLUSTERS USING A GENETIC ALGORITHM VIA THE AM1 METHOD 35

2.1 Introduction 35

2.2 Computational Method 37

2.3 Results and Discussion 41

2.4 Conclusion 61

3. GLOBAL OPTIMIZATION OF H-PASSIVATED SI CLUSTERS AT THE AB INITIO LEVEL VIA AN ITERATIVELY PARAMETRIZED SEMIEMPIRICAL METHOD 63
3.1 Introduction 63
3.2 Computational Method 67
3.3 Results and Discussion 72
3.4 Conclusion 82
4. INVESTIGATION OF THE TRANSFERABILITY OF THE GAM1 PARAMETERS 84
4.1 Introduction 84
4.2 Computation Method 87
4.3 Results and Discussion 93
4.4 Conclusion 115
5. FAST GLOBAL OPTIMIZATION OF Si_xH_y CLUSTERS: NEW MUTATION OPERATORS IN THE CLUSTER GENETIC ALGORITHM 117
5.1 Introduction 117
5.2 Computation Method 121
5.2.1 Cluster Genetic Algorithm (CGA) and the new mutation operators. 121
5.2.2 Parametrization genetic algorithm (PGA). 124
5.2.3 Semiempirical and ab initio methods. 125
5.3 Result and Discussion 126
5.4 Conclusion 133
6. FAST GLOBAL OPTIMIZATION OF F PASSIVATED SI CLUSTERS AT THE DFT LEVEL 136
6.1 Introduction 136
6.2 Computation Method 139
6.3 Results and Discussion 142
6.4 Conclusion 159
7. CONCLUSION 160
REFERENCES 167

LIST OF TABLES

1.1 Parameters used in Semiempirical Methods 24

3.1 Comparison of the AM1 and the final Si_6H_2 and Si_6H_6 GAM1 parameters. 73

4.1 Comparison of the AM1 parameters and the improved GAM1 parameters derived from a Si_7H_{14} training set. 96

4.2 Comparison of the Si_7H_{14} relative energies for MP2 optimized geometries versus the AM1 and GAM1 relative energies calculated at the MP2 optimized geometry (S. P.) and at the optimized semiempirical geometry (Opt.). The RMS deviation between the semiempirical and the MP2 energies is evaluated using equation (2). All energies are in kcal/mol. 97

5.1 Comparison of the AM1 parameters and the improved GAM1 parameters using the PGA. $GAM1^a$ are the present Si and H parameters and $GAM1^b$ are the previous Si only parameters with both sets being derived from the Si_7H_{14} training set [155]. $GAM1^c$ are the Si parameters obtained from the Si_6H_6 training set [154]. The local geometries and energies of the $GAM1^a$ training set was computed with the MP2/LanL2dz(d) method while the $GAM1^b$ and $GAM1^c$ training set used the MP2/6-31G* method. 127

5.2 Comparison of the Si_7H_{14} relative energies for MP2 optimized geometries versus the AM1 and GAM1 relative energies calculated at the MP2 optimized geometry (S. P.) and at the optimized semiempirical geometry (Opt.). The RMS deviation between the semiempirical and the MP2 energies is evaluated using equation (2). All energies are in kcal/mol. 128

6.1 Relative energies for the two Si_2F_4 and Si_2H_4 isomers with C_{2h} and C_s symmetry shown in figure 6.1. Listed are the MP2 and B3LYP optimized energies for various basis sets and the CCSD(T) single point energy computed with the B3LYP optimized geometry. All energies are in kcal/mol. 144

LIST OF FIGURES

1.1 The global minimum structures of Si_2 [22, 23], Si_3 [22, 23], Si_4 [22–26], Si_5 [23,24], Si_6 [22,24,25] and Si_7 [22–25] clusters may be considered established by calculation and experiment. 2

1.2 The global minima of Si_8, Si_9, and Si_{10} [22]. 3

1.3 The global minimum structures of the Si_{11} clusters predicted by different methods. (a) by MP2/6-31G(d), CCSD(T)/6-31G(d) and MP2/6-311G(2d), (b) by CCSD(T)/6-311G(2d) and (c) by MP2/6-311G(2d) [22]. 4

1.4 The prolate global minimum structures of the Si_{17-18} clusters predicted by a genetic algorithm using a DFT method with LDA to evaluate the cluster energies [28]. 5

1.5 The global minimum structures of the Si_2H_3 and Si_2H_4 clusters predicted by the CCSDT/cc-pVTZ method after the relativistic correction and the ZPE correction are included. 6

1.6 The global minimum structures of the Si_2F_3 and Si_2F_4 clusters predicted by the DFT method [93]. 8

1.7 Local Minimum (I and III) versus Global Minimum (II) 9

1.8 A typical genetic algorithm chart 12

2.1 Representative optimized cluster structures obtained with the AM1, B3LYP and MP2 methods. The structures are: (a) $Si_{10}H_4$ a-10, (b) $Si_{10}H_8$ f-1, and (c) $Si_{10}H_{20}$ d-10(5) dissociated into $Si_{10}H_{18}$ and a H_2 molecule. The three bond lengths are listed in the order AM1, (B3LYP) and [MP2] in Å. . . . 43

2.2 Ten lowest AM1 energy structures for $Si_{10}H_4$. The numbers label the B3LYP and MP2 energy rankings as explained in the text. 45

2.3 Ten lowest AM1 energy structures for $Si_{10}H_8$. 47

2.4 Ten lowest AM1 energy structures for $Si_{10}H_{12}$ 48

2.5 Twenty lowest AM1 energy structures for $Si_{10}H_{16}$. 50

2.6 Twenty lowest AM1 energy structures for $Si_{10}H_{20}$. 53

2.7 GA evolution for $Si_{10}H_{16}$. The graph shows how the cluster binding energy of the lowest (dark line) and highest (light line) energy individual in each population varies with generation number. 57

2.8 The $Si_{14}H_{20}$ cluster where the Si framework has a diamond lattice like structure. 58

2.9 Twenty lowest AM1 energy structures for $Si_{14}H_{20}$. 60

3.1 The AM1 and GAM1 versus MP2 relative energies for the Si_6H_2 clusters. All energies are in kcal/mol. 75

3.2 The Si_6H_2 MP2 optimized structures within 20 kcal/mol of the global minimum found by the global optimization procedure. The structures are labeled in order of increasing energy, where the numbers in the parentheses are the MP2 energies with ZPE corrections given in kcal/mol relative to structure (a). Structure (e) is the MINDO/3 global minimum previously found by Meleshko *et al.* [66]. 76

3.3 The AM1 and GAM1 versus MP2 relative energies for the Si_6H_6 clusters. All energies are in kcal/mol. 78

3.4 Si_6H_6 MP2 optimized structures. Structures (a-i) were obtained by our global optimization procedure and are labeled in order of increasing energy, where the numbers in the parentheses are the MP2 energies with ZPE corrections given in kcal/mol relative to structure (a). Structure (κ) is the lowest energy structure found by Miyazaki *et al.* related to structures (d), (g) and (h) by different H atom arrangements. Structure (x-z) are additional low energy structures previously found by Meleshko *et al.* and Miyazaki *et al.* [64, 66]. 79

4.1 The MP2/6-31G* locally optimized geometries for the Si_7H_{14} training set. In parentheses are the MP2 relative energies in kcal/mol. 94

4.2 Comparison of the Si_7H_{14} AM1 (left) and GAM1 (right) relative optimized energies against MP2/6-31G* relative optimized energies. Energies are in kcal/mol. 98

4.3 Comparison of AM1 and GAM1 relative optimized energies against MP2 relative optimized energies for representative stoichiometries $Si_{10}H_8$, $Si_{10}H_{16}$ and $Si_{14}H_{20}$. Energies are in kcal/mol. 101

4.4 MP2/6-31G* optimized $Si_{10}H_{16}$ structures. 103

4.5 MP2/6-31G* optimized $Si_{10}H_{14}$ structures and their corresponding relative energies. In parentheses are the MP2 relative energies in kcal/mol. 106

4.6 MP2/lanl2dz(d) optimized $Si_{14}H_{20}$ structures. In parentheses are the MP2 relative energies in kcal/mol. Structures (a-t) were obtained from the CGA while the structure (u) is the proposed global minimum at both GAM1 and MP2 level. Structures (v) and (w) were obtained from the seeded CGA. . 110

4.7 MP2/lanl2dz(d) optimized $Si_{14}H_{18}$ structures. In parentheses are the MP2 relative energies in kcal/mol. 114

5.1 Low energy structures for Si_6H_6 and $Si_{14}H_{20}$ where in parenthesis are listed their different relative MP2 energies in kcal/mol. For Si_6H_6: (a) MP2 global minimum, (b) GAM1 global minimum, and (c) one of the low energy structures from the CGA resembling (a) [154]. For $Si_{14}H_{20}$: (d) MP2 global minimum and (e) and (f) the two lowest energy structures found by the CGA [155]. 118

5.2 Examples showing the four new mutation operators acting on a parent cluster to produce an offspring which is then locally optimized. 122

5.3 The MP2/lanl2dz(d) locally optimized geometries for the $Si_{18}H_{24}$ low energy clusters obtained from the CGA. In parentheses are the MP2 relative energies in kcal/mol. 131

5.4 The MP2/lanl2dz(d) locally optimized geometries for the $Si_{18}H_{22}$ low energy clusters obtained from the CGA. In parentheses are the MP2 relative energies in kcal/mol. 132

5.5 The Si_xH_y MP2/LanL2dz(d) global minima: a. $Si_{10}H_{16}$; b. $Si_{10}H_{14}$; c. $Si_{14}H_{20}$; d. $Si_{14}H_{18}$; e. $Si_{18}H_{24}$; f. $Si_{18}H_{22}$. 134

6.1 The locally optimized Si_2H_4 and Si_2F_4 isomers. The Si_2H_4 geometry information are listed from top to bottom in the order: MP2/LanL2dz(d), B3LYP/LanL2dz(d) and CCSD(T)/cc-pVTZ [69] methods. The Si_2F_4 geometry information is in the order: MP2/LanL2dz(d), B3LYP/LanL2dz(d) and B3LYP/DZP++(d) [93] methods. 145

6.2 The Si_7F_{14} B3LYP locally optimized geometries for the original 14 training set compounds. In parentheses: (Si_7F_{14} B3LYP relative energy | corresponding Si_7H_{14} B3LYP and MP2 relative energies). All energies are in kcal/mol. 148

6.3 The left chart (a) shows the GAM1 relative energies of Si_7H_{14} clusters vs the B3LYP relative energies of the 14 Si_7F_{14} training compounds, while the right chart (b) shows the corrected GAM1* relative energies of Si_7H_{14} clusters vs the B3LYP relative energies of Si_7F_{14} clusters. The lines in the charts all have a slope of 1 but with various interceptions to best match with the data. 149

6.4 The B3LYP locally optimized geometries of the 3 low energy Si_7F_{14} clusters obtained by the CGA using the GAM1* parameters. (u) and (v) are Si_7F_{14} clusters containing five- and six-membered Si rings, respectively; (u^H) is the Si_7H_{14} global minimum. In parentheses: (Si_7F_{14} B3LYP and MP2 relative energies | Si_7H_{14} B3LYP and MP2 relative energies). All energies are in kcal/mol. 151

6.5 The left charts shows the GAM1 relative energies of Si_7H_{14} clusters versus the B3LYP relative energies of Si_7F_{14} clusters, while the right charts shows the GAM1* relative energies of Si_7H_{14} clusters versus the B3LYP relative energies of Si_7F_{14} clusters. The charts (a) and (b) are for the 14 Si_7F_{14} training compounds plus 12 more data for the low energy clusters predicted by the CGA coupled with the GAM1* energy function. The charts (c) and (d) are the amplified parts of the charts (a) and (b) after the 2 rightmost data with the highest B3LYP energies are excluded. The lines in the charts all have a slope of 1 but with various interceptions to best match with the data. 152

6.6 The B3LYP locally optimized geometries for the Si_8F_{14} low energy clusters located by the CGA. Structures (u) and (v) correspond to the lowest energy Si_8H_{14} clusters. In parentheses: (Si_8F_{14} B3LYP relative energy | Si_8H_{14} B3LYP and MP2 relative energies). All energies are in kcal/mol. 154

6.7 The $Si_{10}F_{16}$ low energy B3LYP locally optimized geometries located by the CGA on $Si_{10}F_{16}$. Structure (u) is the global minimum and (v) and (w) are low energy structures for $Si_{10}H_{16}$. In parentheses: ($Si_{10}F_{16}$ B3LYP relative energy | $Si_{10}H_{16}$ B3LYP and MP2 relative energies). All energies are in kcal/mol. 156

6.8 The $Si_{10}F_{14}$ low energy B3LYP locally optimized geometries located by the CGA. Structures (u-x) are the $Si_{10}H_{14}$ global minimum and low energy structures. In parentheses: ($Si_{10}F_{14}$ B3LYP relative energy | $Si_{10}H_{14}$ B3LYP and MP2 relative energies). All energies are in kcal/mol. 157

7.1 The Si_xH_y MP2/LanL2dz(d) global minima: a. $Si_{10}H_{16}$; b. $Si_{10}H_{14}$; c. $Si_{14}H_{20}$; d. $Si_{14}H_{18}$; e. $Si_{18}H_{24}$; f. $Si_{18}H_{22}$. 164

7.2 The Si_xL_y global minima at ab initio level: a. Si_7H_{14}; b. Si_7F_{14}; c. Si_8H_{14}; d. Si_8F_{14}; e. $Si_{10}H_{16}$; f. $Si_{10}F_{16}$; g. $Si_{10}H_{14}$; h. $Si_{10}F_{14}$. 165

LIST OF ABBREVIATIONS

AM1 Austin Model 1 (one of the semiempirical methods)

AO Atomic Orbital

B3LYP Becke 3-term functional; Lee, Yang, Parr exchange

BO Born-Oppenheimer

BS Basis Set

BSSE Basis Set Superposition Error

CCSD Coupled-Cluster Singles Doubles

CCSD(T) Coupled-Cluster Singles Doubles and non-iterative Triples correction

CCSDT Coupled-Cluster Singles Doubles and full Triples correction

CGA Cluster Genetic Algorithm

CNDO Complete Neglect of Differential Overlap

CI Configuration Interaction

CPU Central Processing Unit

CVD Chemical Vapor Deposition

DFT Density Functional Theory

DH Dunning-Hay Basis Set

EA Evolutionary Algorithm

ECP Effective Core Potential

GA Genetic Algorithm

GAM1 Genetic-algorithm parametrized AM1 semiempirical method

GAMESS the General Atomic and Molecular Electronic Structure System

GO Global Optimization

HF Hartree-Fock

HOMO Highest Occupied Molecular Orbital

HW Hay-Wadt effective core potential and basis set

IBM SP2 International Bussiness Machines Scalable Power parallel system 2

INDO Intermediate Neglect of Differential Overlap

LANL2DZ Los Alamos National Laboratory 2nd Double-Zeta basis set

LDA Local Density Approximation

LJ Lennard-Jones (clusters)

LSDA Local Spin Density Approximation

LUMO Lowest Unoccupied Molecular Orbital

MHPCC Maui High Performance Computing Center

MINDO Modified Intermediate Neglect of Differential Overlap

MINDO/3 Modified Intermediate Neglect of Differential Overlap version 3

MM Molecular Mechanics

MNDO Modified Neglect of Diatomic Overlap

MP Moller-Plesset (perturbation theory)

MP2 Moller-Plesset 2nd order perturbation theory

MP4 Moller-Plesset 4th order perturbation theory

MRCI Multi-Reference Configuration Interaction

NDDO Neglect of Diatomic Differential Overlap

NP Nondeterministic Polynomial

PGA Parametrization Genetic Algorithm

PL Photoluminescence

PM3 Parametric Method 3 (one of the semiempirical methods)

PSi Porous Silicon

QM Quantum Mechanics

QN Quasi-Newton method

RHF Restricted Hartree-Fock

RMS Root Mean Square

ROHF Restricted Open-shell Hartree-Fock

SA Simulated Annealling

SCF Self-Consistent Field

TB Tight Binding

UHF Unrestricted Hartree-Fock

ZINDO Zerner's INDO, a semiempirical quantum mechanical program developed by M. C. Zerner.

$$< \mu|\nu > = \int \phi_\mu^*(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)d\mathbf{r}_1$$

$$< \mu|\hat{O}|\nu > = \int \phi_\mu^*(\mathbf{r}_1)\hat{O}\phi_\nu(\mathbf{r}_1)d\mathbf{r}_1$$

$$(\mu\nu|\lambda\sigma) = \iint \phi_\mu^*(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)r_{12}^{-1}\phi_\lambda^*(\mathbf{r}_2)\phi_\sigma(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

$$< \mu\nu|\lambda\sigma > = \iint \chi_\mu^*(\mathbf{x}_1)\chi_\nu^*(\mathbf{x}_2)r_{12}^{-1}\chi_\lambda(\mathbf{x}_1)\chi_\sigma(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2$$

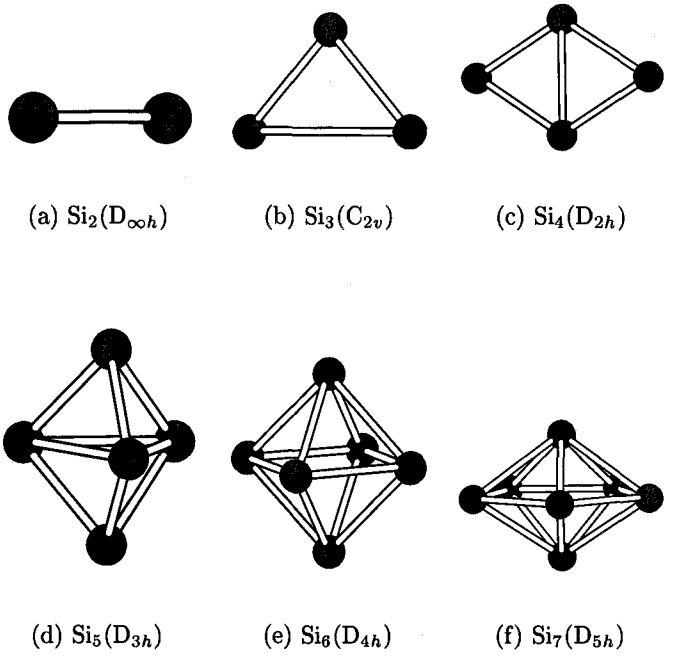
CHAPTER 1. INTRODUCTION

1.1 Unpassivated and Passivated Si clusters

Bulk Si has been widely utilized in the microelectronics industry for many years. However the nanometer sized Si clusters had not attracted too much interest until the discovery of the room-temperature photoluminescence (PL) from porous Si (PSi) was made by Canham in 1990 [1]. The nanometer sized porous Si emits bright light in the visible range of spectrum from red to blue, which suggests a novel and promising potential of fabricating various optoelectronic devices using the Si clusters. Quantum confinement [2–6] and surface effect [7–21] theories have been proposed to explain the mechanism of the PL from the porous Si. The structural features of the lowest energy isomers of the Si clusters are needed to correctly model their optoelectronic properties, however this information cannot readily be obtained from the experiments. The focus of this dissertation is to develop theoretical techniques for global optimizing, that is finding the lowest energy isomers, of Si clusters. Another reason why we are interested in the global optimization of the Si clusters is to uncover the mechanism of chemical vapor deposition (CVD) process. CVD is a chemical technique to form the Si thin films which are widely used for fabricating the Si based optoelectronic devices such as the solar cells. In the typical CVD process, small silicon hydride molecules, such as SiH_4 and Si_2H_6, serve as the precursors to form larger Si_xH_y intermediate molecules which are then deposited on the substrate. By investigating the most stable structures of the Si_xH_y molecules, we may better understand the mechanism of the CVD process.

The Si clusters can be categorized into 2 groups: the unpassivated Si only clusters

and the passivated Si clusters with various types of ligands on the surface Si atoms. The extensive study of the unpassivated Si clusters have been making great progress both theoretically and experimentally in the past 14 years [22–54]. One major achievement so far is that the same global minimum structures of the small sized Si clusters are consistently predicted by the experiments [24–26] and the theoretical calculations [22, 23]. The most stable isomers of the small Si clusters consisting 2-7 atoms are shown in figure 1.1, all of which adopt compact structures except that the Si_4 cluster prefers a rhombus planar structure over the more compact pyramid structure [22–26].



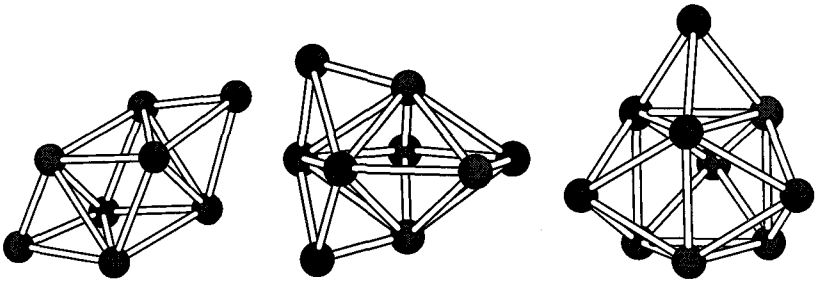
(a) $Si_2(D_{\infty h})$ (b) $Si_3(C_{2v})$ (c) $Si_4(D_{2h})$

(d) $Si_5(D_{3h})$ (e) $Si_6(D_{4h})$ (f) $Si_7(D_{5h})$

Figure 1.1: The global minimum structures of Si_2 [22, 23], Si_3 [22, 23], Si_4 [22–26], Si_5 [23,24], Si_6 [22,24,25] and Si_7 [22–25] clusters may be considered established by calculation and experiment.

The slightly larger Si_{8-11} global minimum clusters are obtained by Zhu and Zeng

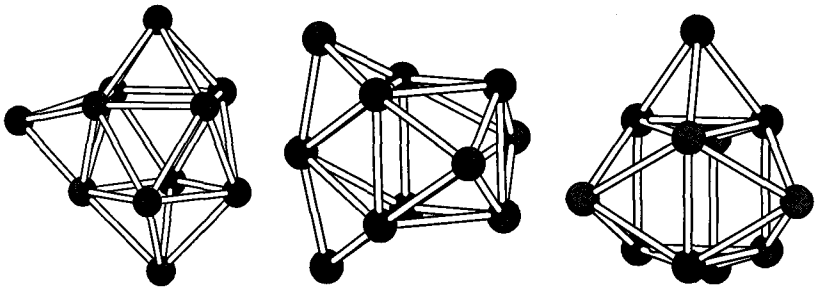
[22] using both the MP2 and CCSD(T) methods with a double-zeta 6-31G(d) and a larger triple-zeta 6-311G(2d) basis sets including the d-type polarization functions on the Si atoms. Their MP2 and CCSD(T) results find that the global minima of the Si_{8-10} clusters also adopt compact structures as shown in the figure 1.2. However their results obtained by the various methods using different basis sets predict the different global minimum structures shown in the figure 1.3 for the Si_{11} clusters. Unfortunately there is no experimental data available supporting the theoretical predictions for the most stable structures of Si clusters with more than 7 Si atoms.



(a) $Si_8(C_{2h})$ (b) $Si_9(C_{2v})$ (c) $Si_{10}(C_{3v})$

Figure 1.2: The global minima of Si_8, Si_9, and Si_{10} [22].

The number of stable isomers grows exponentially with the size of the Si clusters. Soon it becomes infeasible to search the global minimum for the slightly larger Si clusters by performing local optimizations for many structures which are fabricated simply by chemical intuition. Ho et al. investigated the possible global minimum structures of Si_{12-20} clusters by using a genetic algorithm (GA) coupled with a fast density functional theory (DFT) method using the local density approximation(LDA) [28]. They found that the Si clusters with 16-18 Si atoms preferred prolate structures where two examples

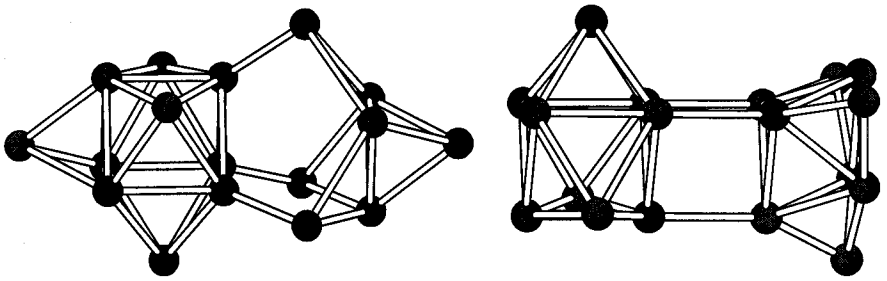


(a) C_s (b) C_{2v} (c) C_{3v}

Figure 1.3: The global minimum structures of the Si_{11} clusters predicted by different methods. (a) by MP2/6-31G(d), CCSD(T)/6-31G(d) and MP2/6-311G(2d), (b) by CCSD(T)/6-311G(2d) and (c) by MP2/6-311G(2d) [22].

for the Si_{17} and the Si_{18} global minima are shown in figure 1.4. The calculated ion mobilities of the Si_{17} and Si_{18} global minimum agree well with the experimental data. They also predicted that the Si_{19-20} global minimum start to adopt compact structures. However the experimental data suggests that a transition from the prolate to compact shape of the Si only clusters only occurs at around Si_{24}. Therefore they also constructed the prolate structures for the Si_{19-26} clusters which were also low energy isomers but more consistent with the results of the ion mobility experiment. It is worth noting that the Si_{20} cluster, unlike C_{20}, prefers a more compact structure instead of a hollow-center globe-shaped fullerene structure, which may be due to the rather weak Si-Si π bond and the fact that Si atoms favors tetrahedral coordination.

The unpassivated Si clusters are expected to have dangling bonds on the surface Si atoms and to be highly reactive with moisture or O_2 in the air. Defects due to the dangling bonds may also quench any photoluminescence which may be useful in a practical optoelectronic device. H atoms could serve as a ligand to reduce the dangling

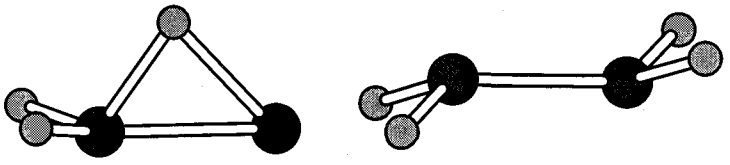


(a) $Si_{17}(C_{3v})$ (b) $Si_{18}(C_{3v})$

Figure 1.4: The prolate global minimum structures of the Si_{17-18} clusters predicted by a genetic algorithm using a DFT method with LDA to evaluate the cluster energies [28].

bonds on the surface of the Si clusters. For these reasons the research in this dissertation is more interested in the H-passivated Si clusters rather than the bare Si-only clusters. The structural, energetical and other electronic properties of the H-passivated Si clusters have been studied intensively both in theory and by experiments [55–90]. Recently small Si hydrides such as the Si_2H_3 and Si_2H_5 radicals and their corresponding deuterium substituted Si_2D_3 and Si_2D_5 radicals have been detected in low temperature silane matrices by Kaiser's group [87,88]. However it still remains extremely difficult to explore the stablest structures of slightly larger Si hydrides by experiments, therefore theoretical calculation can be very useful for predicting the geometric, energetic and spectroscopic properties of the low energy Si_xH_y clusters. Unfortunately many theoretical attempts to locate the lowest energy Si_xH_y clusters were not performed systematically. A very common way to search for the global minimum is to simply fabricate possible low energy clusters based on previous knowledge of the low energy structures of the Si only cluster [63–65], or to use chemical intuition to guess at likely Si_xH_y geometries [78]

and then perform a local geometry optimization. However this approach can easily miss the true global minimum for even very simple clusters. Even the search for the global minimum of the small Si_2H_y clusters turns out to be very challenging due to the many possible isomers which lie in a small energy range. Sax and Kalcher predicted the H_3Si-Si structure as the Si_2H_3 global minimum by using a MRCI technique with double and triple-zeta basis set to perform the local optimization [71]. The results of G2 calculations by Curtiss et al. agreed with Sax's result [72]. However Gong [73] predicted a different C_2 symmetry HSi-H-SiH structure to be the lowest energy isomer among the 6 Si_2H_3 isomers they studied in 1997. In 2003, Sari et al. investigated 10 isomers of Si_2H_3 and Si_2H_4 clusters and found the true global minimum structures shown in figure 1.5. Their results were obtained by using a CCSDT method with a quite substantive cc-pVTZ basis set and they also properly included the relativistic correction and the Zero Point Energy (ZPE) correction.



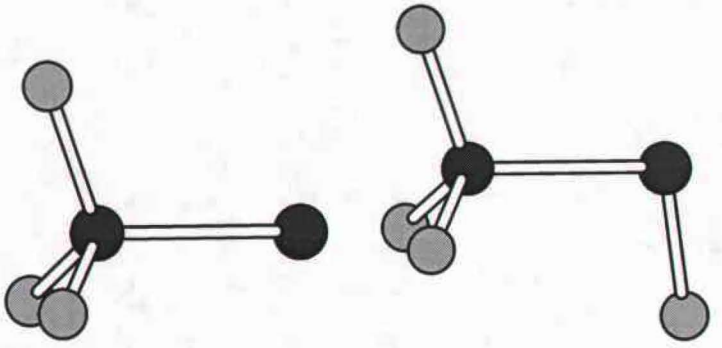
(a) $Si_2H_3(C_s)$ (b) $Si_2H_4(C_{2h})$

Figure 1.5: The global minimum structures of the Si_2H_3 and Si_2H_4 clusters predicted by the CCSDT/cc-pVTZ method after the relativistic correction and the ZPE correction are included.

Meleshko et al. have been investigating the global minimum structures of the Si_xH_y clusters using the Simulated Annealing (SA) algorithm, which is one of the many global optimization techniques, coupled with MINDO/3 semiempirical method to eval-

uate the cluster energy [66–68]. However, the unsatisfactory accuracy of the semiempirical MINDO/3 method makes their results fairly dubious and is discussed further in detail in the Chapter 3. All in all the many theoretical calculations have been done unsystematically and the results are sometimes contradictory. Furthermore, there are no or few experimental data clearly supporting the results of theoretical calculations. This dissertation describes several strategies which successfully and consistently find the global minimum structures of H passivated Si clusters with up to 18 Si atoms and 24 H atoms at the ab initio and DFT level systematically.

We are also interested in the ligand effect on the structure of Si cluster. F atoms may serve as another type of ligand with strong electronegativity other than H atoms. Feller and Dixon have investigated the heats of formation of very small SiF, SiF_2 and SiF_4 clusters using the CCSD(T) method with large basis set [91]. Schaefer's group has investigated the electron affinities of the silicon fluorides SiF_n (n=1-5) using the B3LYP method [92]. Schaefer and coauthors have also searched the global minimum isomers of $Si_2F_{(1-6)}$ neutral compounds and $Si_2F^-_{(1-6)}$ negative ions [93]. However, no research which searches for the global minimum structures of the Si_xF_y clusters with more than 2 Si atoms have been published so far to the best knowledge of the author. In chapter 6 we present the results of global optimization for the larger Si_xF_y clusters. It is worth noting that our calculations predict the Si_7F_{14}, Si_8F_{14}, $Si_{10}F_{14}$ and $Si_{10}F_{16}$ clusters to have very different global minimum structures compared with those of the corresponding Si_xH_y clusters, which suggests that various ligands can be used to fine-tune the structures and the optoelectronic properties of the Si cluster potentially. The global optimization of both the Si_xH_y and Si_xF_y clusters in our research are performed using genetic algorithms which are introduced in the next subsection.



(a) $Si_2F_3(C_s)$ (b) $Si_2F_4(C_{2h})$

Figure 1.6: The global minimum structures of the Si_2F_3 and Si_2F_4 clusters predicted by the DFT method [93].

1.2 Global Optimization Strategies

1.2.1 Global minimum

The goal of the global optimization is to find the global extreme of a N-dimensional function $f(x_1, x_2, ...x_N)$ and the corresponding $x_{1,2,...N}$ values. The global minimum is the local minimum with the lowest value while the global maximum is the local maximum with the highest value. The local minimum or maximum is defined as a point where its value is lower or higher than all its neighbor's values respectively. Mathematically speaking, the positive second derivatives indicate a local minimum, while negative second derivatives indicate a local maximum at the point where all of the first derivatives are zeroes.

The figure 1.6 gives a simple example of one dimensional function $f(x)$ with 2 local minima and a global minimum. Among the three valleys in the chart of $f(x)$ versus x, the I and III valleys are the local minima while only the the valley II is the global

minimum with the lowest value of function $f(x)$. A typical local optimization technique such as the Quasi-Newton (QN) method can only find a local minimum within the same catchment basin as the starting configuration, whereas the global optimization techniques aim to find the global minimum with the lowest possible function value [97–99].

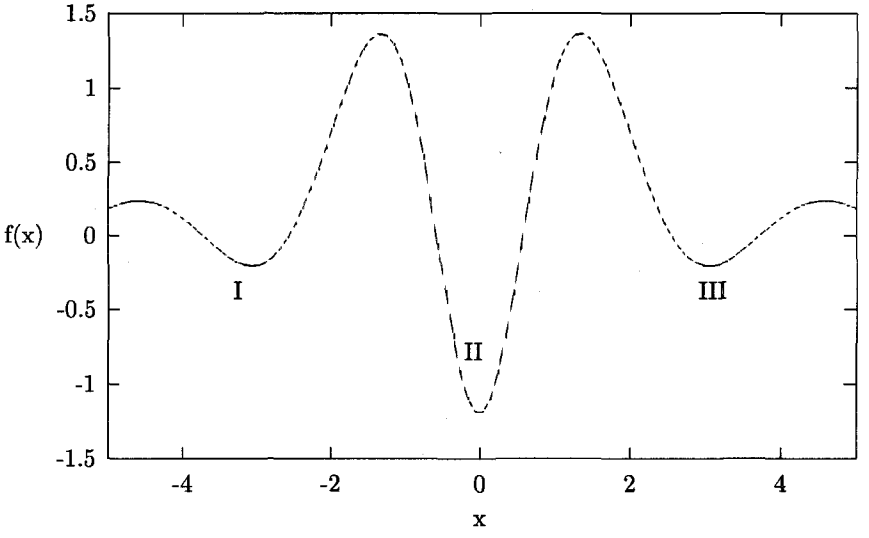


Figure 1.7: Local Minimum (I and III) versus Global Minimum (II)

The global optimization of the clusters and molecules are Nondeterministic Polynomial (NP) hard problems which can not be solved in polynomial time using a deterministic algorithm [94]. The number of the local minima increases at least exponentially with the increasing size of the system [94,95]. Tsai and Jordan have searched the local minimum of small sized Lennard-Jones clusters exhaustively and found more than 1000 local minima for a small 13-particle system and it was estimated that a 150-particle system has at least 10^{60} local minima by extrapolation [95]. To conquer the NP-hard global optimization problem, many global optimization techniques have been proposed. Due

to the space limitation, we will only briefly introduce the 2 most successful and widely used techniques: simulated annealing (SA) [97,98] and genetic algorithms (GA) [99].

1.2.2 Simulated Annealing

Simulated Annealing (SA) simulates the way in which a metal cools and freezes into a lowest energy crystalline structure [97]. SA's major advantage over local search methods is its ability to avoid becoming trapped at local minima. The SA algorithm based global optimization for the clusters starts from an initial geometry which may be randomly fabricated or intentionly designed. The corresponding energy of the initial geometry is evaluated. Then this geometry is randomly modified by moving one or more randomly selected atoms by random displacements. The energy of the new structure is also evaluated and compared with the energy of the previous geometry before the displacement. The SA algorithm accepts the new geometry if it has a lower energy as in a local search strategy. However, the SA also accepts some of the moves which leads to a higher energy cluster with a probability p:

$$p = min\{1, \quad e^{\frac{-\Delta E}{RT}}\} \tag{1.1}$$

where ΔE is the increase in cluster energy, R is 8.314 kJ/mol and T is the temperature parameter which keeps decreasing during the annealing process. The geometry keeps unchanged if the move is not accepted. At each temperature, many configurations are generated consecutively by performing the random movements which are accepted by the SA using the above probability equation. The temperature T is lowered gradually until absolute zero K is approached. Although the SA can in principle find the global

minimum if the number of geometries being evaluated at each temperature approaches infinity and the rate of lowering the temperature is slow enough, practically the SA-based global search often fails because the geometry gets trapped in a deep valley of a local minimum [96]. In contrast, genetic algorithms have the advantage of being able to jump between different catchment regions without needing the energy to pass over the barriers.

1.2.3 Genetic Algorithm

Genetic Algorithms (GA) are inspired by Darwin's evolutionary theory. The basic idea is that the strongest survives. GA has been widely used in various fields in chemistry such as the structure prediction of large biomolecules including protein, DNA and RNA, molecular design, protein docking, cluster or molecule structure global optimization, crystallography, parametrization of empirical and semiempirical method etc. [100–115]. Figure 1.7 shows a typical flow chart of a genetic algorithm which is used to solve global optimization problems. The first step is to generate the initial solutions to the problem. The solutions can be generated randomly or fabricated based on the previous knowledge or intuition. Although the latter approach needs to be applied with caution since it may introduce prejudice into the GA. The second step is to evaluate the fitness values of the initial solutions. The fitter the solution, the higher chance it has of passing its 'genes' (partial solution) to the offspring.

There are various parent selection schemes being used in the step 3. Good selection schemes keep a proper selection pressure. Too low selection pressure makes the bad parents have too much chance of being selected and generating bad offspring, while too high a selection pressure may cause the best solution already in the population to be selected exclusively which may easily lead to premature convergence. The simplest

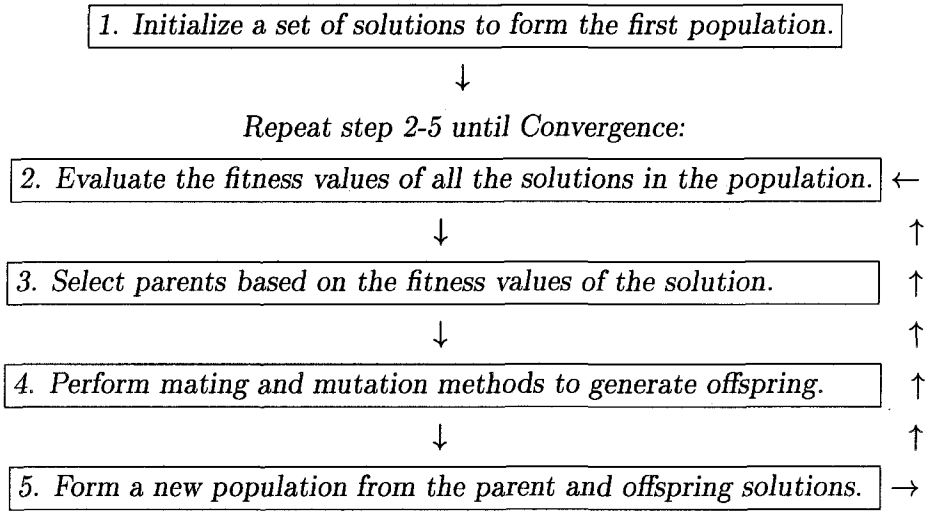


Figure 1.8: A typical genetic algorithm chart

selection scheme is the roulette selection, where the probability P_i of selecting the i-th individual can be expressed as:

$$P_i = \frac{f_i}{\sum_i f_i} \tag{1.2}$$

where f_i is the fitness value of the i-th individual in the population. We first generate a random number between 0 and 1, we then select the k-th individual if the random number is between $\sum_{i=1}^{k-1} f_i$ and $\sum_{i=1}^{k} f_i$. In Boltzman selection the probability becomes

$$P_i = \frac{e^{f_i/T}}{\sum_i e^{f_i/T}} \tag{1.3}$$

where T is a constant adjusting the distribution of the values of P_i. The above two

selection methods may cause premature convergence since the individual which is much fitter than the others may be selected exclusively. Therefore these two selection methods should be applied with enough caution to maintain a proper selection pressure during the evolution. The third type of selection methods is called the rank-based selection, where the probability of selecting one individual does not depend on its fitness value but on the ranking of the fitness of that individual. The rank-based selection overcomes the premature convergence problem by preventing an individual which is much fitter than all other individuals to be selected exclusively. The simplest rank-based method is to let the top-ranking individuals have equal chance of being selected while the bottom ranking individuals be discarded completely. Another rank-based method simply uses the rankings as the fitness values for every individuals and then the highest ranking individuals have the highest chance of being selected while other individuals also have fair chance of being selected with a probability which linearly depends on its ranking. Tournament selection belongs to another category of the selection methods which handles the premature convergence problem well. The tournament selection first selects a group of individuals randomly and then the best one in the group is selected as a parent to generate an offspring using the genetic operators including mating and mutation methods. The mating methods usually need 2 parents both of which pass the good 'genes' to the offspring while the mutation methods slightly change the 'genes' of the offspring which simulates the natural process. The genetic operators for generating clusters are explained in detail in the following chapters.

How does the population change generation by generation? There are various ways of updating a population. The simplest way and the most natural way is to let the offspring generation replace the whole parent population. However this is seldomly used in a GA for the cluster global optimization due to its low efficiency. A strong elitism

strategy is often used for cluster global optimization where the best of the parents and offspring are kept in the population while the others are discarded. The compromise between the above two extreme strategies is to keep some of the best individuals and some of the bad solutions to guarantee the diversity of the population.

One major problem of GA is to determine whether it is converged. Most GAs examine the fitness of the best individual in the population and claim that the GA is converged when the best solution in the population does not change or the best fitness value is not changed after a certain number of generations. Another possible criterion is to examine the worst solution in the population and claim that the GA is converged when the worst solution does not change after a few generations. The above two criteria can be coupled with the examination of the convergence of the average fitness value of the whole population. Some also consider a GA being converged after a well known global minimum is located for the purpose of evaluating the efficiency and effectiveness of the GA being tested [102]. It is also common that a GA is terminated for practical reasons after certain number of generations have been evolved or a certain amount of CPU time has been consumed due to the limitation of resources.

The genetic algorithm has become one of the most popular global optimization strategies. The solutions generated by genetic operators in the GA can easily switch between different local minima. Unlike simulation annealing, GA does not require a careful design of the temperature lowering scheme to avoid getting trapped in a local minimum. A GA-based global extreme search is more efficient than a random search or an exhaustive search for such types of problems where the offspring solutions being constructed from building blocks of good parent solutions have better fitness values than random solutions statistically. GA works better for the above problems by squeezing out the bad solutions by their better competitors providing a proper selection

pressure. Genetic algorithms usually can locate the global minimum in a reasonable number of fitness evaluations. However due to the stochastic characteristic of GAs, programs based on GAs do not always produce the same results even when providing the same input. In the situations where the global minimum is hard to find or needs to be confirmed, multiple GA runs with different starting populations are performed for a global optimization problem to ensure that a consistent global extreme is being located.

1.3 Cluster Energy Calculation Methods

Even though the global optimization techniques such as the genetic algorithms dramatically reduce the number of local optimizations needing to be performed, a large number of local optimizations still have to been done to locate the global minimum. Non-empirical methods such as DFT or MP2 may give accurate cluster energies to within a few kcal/mol when compared with experimental data, however they both require a considerable amount of CPU time. The molecular mechanics (MM) methods are much faster and can deal with a large system of thousands atoms in a very short time, however they do not always provide reasonable energies. Another major drawback of the MM methods is that the bonding situation between atoms has to be specified in order to perform the MM calculation. The semiempirical methods can evaluate cluster energies and optimize geometries much faster than the non-empirical methods. An important feature of the semiempirical calculations is that they still have a solid quantum mechanics basis: no additional empirical information is needed to identify bonds between atoms in the semiempirical calculations. Since the semiempirical methods are based on the Hartree-Fock (HF) method, the fundamental concepts of the Hartree Fock theory and the various types of the semiempirical methods are briefly introduced in the

following subsections.

1.3.1 Hartree Fock Theory

The time-independent Schrödinger equation can be used to solve for the energy and the electronic state for the molecular systems:

$$\hat{H}\Psi = E\Psi \tag{1.4}$$

where the $\hat{H}$ is called a Hamiltonian operator which includes the kinetic energy and potential energy terms of all the electrons and nuclei of the studied molecular system:

$$\hat{H} = \frac{-1}{2M_A}\sum_A \nabla_A^2 - \frac{1}{2}\sum_i \nabla_i^2 - \sum_i \sum_A \frac{Z_A}{R_{Ai}} + \sum_{i<j} \frac{1}{r_{ij}} + \sum_{A<B} \frac{Z_A Z_B}{R_{AB}} \tag{1.5}$$

Since the nuclei are much heavier than the electrons, the Born-Oppenheimer (BO) approximation assumes that the kinetic energy contributed by the nuclei is omitable and the wave function of the whole molecular system can be approximated as the product of the wave function of electrons and the wave function associated with the nuclei. Within the BO approximation the electronic Hamiltonian for a specific nuclear configuration is:

$$\hat{H} = \frac{1}{2}\sum_i \nabla_i^2 - \sum_i \sum_A \frac{Z_A}{R_{Ai}} + \sum_{i<j} \frac{1}{r_{ij}} \tag{1.6}$$

The ground state Hartree-Fock wave function Ψ of a N-electron system is approximated by a single Slater determinant Ψ_0 that is constructed from the occupied

molecular orbitals χ_i:

$$\Psi_0 = |\chi_1\chi_2\cdots\chi_N> = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{1.7}$$

where χ_i is the spin orbital function which consists of the spatial function ψ_i and either an α or β spin function:

$$\chi_i(\mathbf{x}) = \begin{cases} \psi_i(\mathbf{r})\alpha(\omega) \\ \text{or} \\ \psi_i(\mathbf{r})\beta(\omega) \end{cases} \tag{1.8}$$

where $\mathbf{x} = \{\mathbf{r}, \omega\}$.

In Hartree-Fock Roothan method, the molecular orbital ψ_i are expressed as a linear combination of atomic-like orbitals or basis functions ϕ_μ:

$$\psi_i = \sum_\mu \phi_\mu C_{\mu i} \tag{1.9}$$

where $C_{\mu i}$ is a molecular orbital coefficient. The $C_{\mu i}$ are determined by iteratively solving the Fock matrix equation:

$$\mathbf{F\ C = S\ C\ E} \tag{1.10}$$

where for canonical MOs, $E_{ij} = \varepsilon_i\ \delta_{ij}$ and ε_i is the i-th molecular orbital energy. **S** is the overlap matrix consisting of integrals representing the overlap between the basis functions:

$$S_{\mu\nu} =< \mu|\nu >= \int \phi^*_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1)\ d\mathbf{r}_1 \tag{1.11}$$

The Fock **F** matrix elements are given by:

$$\begin{aligned} F_{\mu\nu} &= H_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma}[(\mu\nu|\sigma\lambda) - \frac{1}{2}(\mu\lambda|\sigma\nu)] \\ &= H_{\mu\nu} + G_{\mu\nu} \end{aligned} \tag{1.12}$$

where $G_{\mu\nu}$ is the two-electron term and $H_{\mu\nu}$ is the one-electron matrix element which consists of the kinetic energy and electron-nuclear attraction terms:

$$H_{\mu\nu} =< \mu|\frac{-\nabla^2}{2}|\nu > - \sum_B < \mu|\frac{Z_B}{R_B}|\nu > \tag{1.13}$$

and **P** is the density matrix or sometimes called charge-density bond-order matrix:

$$P_{\lambda\sigma} = \sum_i C_{\lambda i} C^*_{\sigma i} n_i \tag{1.14}$$

where n_i is the occupation number: 0 for unoccupied orbitals and 2 for the occupied orbitals in the closed-shell molecules. The two-electron integrals $(\mu\nu|\sigma\lambda)$ are given by:

$$(\mu\nu|\sigma\lambda) = \int \phi^*_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_1) r_{12}^{-1} \phi^*_\sigma(\mathbf{r}_2)\phi_\lambda(\mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.15}$$

The total Hartree-Fock electronic energy E_{el} is not simply the sum of the energies of all the occupied molecular orbitals but is obtained by using:

$$E_{el} = \frac{1}{2} \sum_\mu \sum_\nu P_{\mu\nu}(H_{\mu\nu} + F_{\mu\nu}) \tag{1.16}$$

The SCF procedure includes the following 5 steps:

1. Construct the **S** overlap matrix and the one-electron **H** Hartree matrix. Assume $\mathbf{F} = \mathbf{H}$ initially. The two-electron integrals $(\mu\nu|\sigma\lambda)$ are calculated at this stage for the non-direct HF methods, while the direct HF methods recompute the two-electron integrals during each construction of the Fock matrix in step 4.

2. Solve the Fock matrix equation (1.10) to get the molecular orbital coefficients **C**.

3. Construct the density matrix **P** by using the equation (1.14).

4. Construct the new **F** matrix with the above **P** matrix by using the equation (1.12).

5. Go to step 2 until the molecular coefficient vectors **C** becomes converged.

The total energy for a specific nuclear configuration of the whole molecular system can be obtained by summing up the electronic energy E_{el} and all the core repulsion energies:

$$E_{tot} = E_{el} + \sum_{A<B} E_{AB}^{core} \tag{1.17}$$

where in an ab initio calculation, $E_{AB}^{core} = Z_A Z_B / R_{AB}$.

The local minimum of a N-particle molecule has all the first derivatives of the total energy E_{tot} over the $3N$ Cartesian coordinates of the nuclei equal to zero:

$$\frac{\partial E_{tot}}{\partial Q_i} = 0 \quad for \;\; i = 1, 2, \ldots, 3N \tag{1.18}$$

where Q_i is the i-th Cartesian coordinate of the N-particle molecule.

The local minimum also has a positive definite Hessian matrix **H** where the elements H_{ij} of the symmetric Hessian matrix are given by:

$$H_{ij} = \frac{\partial^2 E_{tot}}{\partial Q_i \partial Q_j} \tag{1.19}$$

The above Hessian matrix is positive definite if and only if all its eigenvalues are

positive.

1.3.2 Semiempirical Methods

As we can see from the above subsection, the calculation for the two-electron integrals can be very time-consuming due to the large number of the possible combinations of $(\mu\nu|\lambda\sigma)$ terms which is proportional to the fourth power of the number of atomic orbital functions. For example, the number of the two-electron integrals increases 16-fold when the number of the atomic orbital functions is doubled.

The semiempirical methods use minimal basis sets and introduce parameters to make approximations to solve a Schrödinger equation more efficiently. To simplify the calculation for constructing the **H** and **F** matrices, Pople and co-workers proposed the zero differential overlap approximation (ZDO), which can be further sorted into three major categories: Complete Neglect of Differential Overlap (CNDO), Intermediate Neglect of Differential Overlap (INDO) and Neglect of Diatomic Differential Overlap (NDDO) [116–118]. The ZDO approximation simply says that the overlap between many atomic orbitals will be small and thus are negligible. The NDDO-based semiempirical methods usually are more accurate and thus currently more widely used than the CNDO or INDO-based methods due to having less approximations because more one-electron and two-electron integrals are retained. In the NDDO methods, the one-electron **H** matrix is constructed in the following 2 ways depending on whether the two atomic functions μ and ν are on the same center:

$$
\begin{aligned}
H_{\mu^A\nu^A} &= < \mu^A | \frac{-\nabla^2}{2} | \nu^A > - < \mu^A | \frac{Z_A}{R_A} | \nu^A > - \sum_{B \neq A} < \mu^A | \frac{Z_B}{R_B} | \nu^A > \\
&= U_{\mu^A\nu^A} - \sum_{B \neq A} < \mu^A | \frac{Z_B}{R_B} | \nu^A >
\end{aligned}
\tag{1.20}
$$

$$
\begin{aligned}
H_{\mu^A\nu^B} &= < \mu^A | \frac{-\nabla^2}{2} - \frac{Z_A}{R_A} - \frac{Z_B}{R_B} | \nu^B > \\
&= \left(\frac{\beta_{\mu^A} + \beta_{\nu^B}}{2} \right) S_{\mu^A\nu^B}
\end{aligned}
\qquad (A \neq B) \tag{1.21}
$$

where μ and ν are the basis functions. A and B label the different atomic centers. $U_{\mu^A\nu^A}$ consists of the one-center kinetic energy and electron-nuclear attraction terms for atom A, and is approximated by the semiempirical parameters. β_{μ^A} and β_{ν^B} are semiempirical parameters introduced to approximate the two-center one-electron resonance integrals $H_{\mu^A\nu^B}$.

The **F** matrix is constructed from the **H** matrix and the one- and two-center two-electron terms:

$$
F_{\mu^A\nu^A} = H_{\mu^A\nu^A} + \sum_{\lambda^A\sigma^A} P_{\lambda^A\sigma^A} [(\mu^A\nu^A|\sigma^A\lambda^A) - \frac{1}{2}(\mu^A\lambda^A|\sigma^A\nu^A)] + \sum_{A \neq B} \sum_{\lambda^B\sigma^B} P_{\lambda^B\sigma^B} (\mu^A\nu^A|\sigma^B\lambda^B) \tag{1.22}
$$

$$
F_{\mu^A\nu^B} = H_{\mu^A\nu^B} - \frac{1}{2} \sum_{\lambda^A\sigma^B} P_{\lambda^A\sigma^B} (\mu^A\lambda^A|\sigma^B\nu^B) \quad (A \neq B) \tag{1.23}
$$

In 1977, Dewar and Thiel developed the Modified Neglect of Differential Overlap

(MNDO) method [119] which is based on the NDDO approximation. The nuclear core repulsion term of the MNDO method is approximated through the following expression:

$$E_{AB}^{MNDO} = Z_A Z_B < s_A s_A | s_B s_B > [1 + e^{(-\alpha_A R_{AB})} + e^{(-\alpha_B R_{AB})}] \tag{1.24}$$

where $< s_A s_A | s_B s_B >$ is a two-center electron-electron repulsion integral.

Unfortunately the MNDO core repulsion term tends to overestimate the repulsion at the van der Waals's distances, and as a consequence poorly reproduces the heats of formation for the H bonds and bulky compounds where many atoms are within the van der Waal's distances.

To remedy deficiencies in the MNDO method, the Dewar group developed the Austin Model 1 (AM1) method in 1985 [120]. The AM1 method better models the van der Waal's forces after revising the core repulsion term E_{AB}^{core} by simply introducing three Gaussian-type functions with 9 extra K_{1-3}, M_{1-3}, L_{1-3} parameters:

$$E_{AB}^{AM1} = E_{AB}^{MNDO} + \frac{Z_A Z_B}{R_{AB}} \left(\sum_{i=1}^{3} K_{A_i} e^{(L_{A_i}(R_{AB} - M_{A_i})^2)} + \sum_{i} K_{B_i} e^{(L_{B_i}(R_{AB} - M_{B_i})^2)} \right) \tag{1.25}$$

The 3rd parametrization of the original MNDO models, the PM3 method, has the same scheme as the AM1 method to construct the **F** matrix [120–122]. However the PM3 method differs from the AM1 method in that the AM1 method derives the one-center, two-electron integrals from the atomic spectroscopy while the PM3 method treats those integrals as pure parameters which are chosen by fitting the calculated molecular properties to the corresponding experimental data. Table 1.1 lists

the different parameters used in these three popular MNDO, AM1 and PM3 semiempirical methods [119–122]. Other than the MNDO, AM1 and PM3 methods, some other semiempirical methods with the NDDO approximation have also been developed including MNDOC [123], MNDO/d [124–128], SAM1 [129,130] , and NDDO/G [131].

Table 1.1: Parameters used in Semiempirical Methods

Parameters	description	MNDO	AM1	PM3
U_{ss} and U_{pp}	atomic orbital 1-e$^-$ 1-center integrals	+	+	+
ζ_s and ζ_p	Slater atomic orbital exponents	+	+	+
β_s and β_p	1-e$^-$ 2-center resonance integral terms	+	+	+
α	core-core repulsion term	+	+	+
K_{1-n}	core-core repulsion parameters		+	+
L_{1-n}	core-core repulsion parameters		+	+
M_{1-n}	core-core repulsion parameters		+	+
G_{ss}, G_{pp}, G_{sp}, G_{p2}	1-center 2-e$^-$ repulsion integral	+	+	+
H_{sp}	1-center 2-e$^-$ exchange integral	+	+	+

In our initial calculations on the Si_xH_y clusters, we felt that the AM1 method gave the most useful results relative to the MNDO and PM3 methods. Later on we realized that we needed better results than provided with the original AM1 parameters. Therefore we developed techniques to reparameterize the AM1 method so that more accurate energy rankings for the Si_xH_y and Si_xF_y cluster were obtained. Chapter 3 through 6 describe in more detail this AM1 reparametrization. However we note that even the reparametrized AM1 method does not always give the same energy ranking as the ab initio or DFT calculations. Therefore we further determine the relative energies and energy rankings for these clusters more accurately using the ab initio and DFT methods after we use the CGA to prescreen the low energy clusters at the semiempirical level. A brief overview on post-HF and DFT methods is covered in the next section.

1.4 Post-HF methods

The Hartree-Fock methods ignore much of the electron correlation and give unreliable energy estimations, therefore some post-HF methods such as the MP2 [132] and B3LYP [133–135] DFT method are used in this dissertation to give better energy rankings for various isomers. We also apply the more expensive coupled cluster (CC) methods to evaluate the accuracy of the MP2 and DFT methods we have used. These above mentioned methods are briefly introduced in the following subsections.

1.4.1 Moller-Plesset Many-Body Perturbation Theory

In the Moller-Plesset Perturbation Theory (MPPT), the zero-order Hamiltonian $\hat{H}_0$ is the sum of the one-electron Fock operators:

$$\hat{H}_0 = \sum_{i=1}^{N} \hat{f}_i \tag{1.26}$$

where $\hat{f}_i$ is the Fock operator which is constructed from the one-electron kinetic energy and electron-nuclei repulsion terms and two-electron coulomb and exchange terms. If we arbitrarily assign an electron 1 to a spin orbital χ_a, we have the following equation:

$$\hat{f}_1 \chi_a(1) = \varepsilon_a \chi_a(1) \tag{1.27}$$

where ε_a is the orbital energy of the spin orbital χ_a.

The HF ground-state wave function Ψ_0 in equation (1.7) is an eigenfunction of the $\hat{H}_0$ operator with an eigenvalue $E^{(0)}$:

$$\hat{H}_0\Psi_0 = E^{(0)}\Psi_0 \tag{1.28}$$

where $E^{(0)}$ is the sum of the occupied molecular orbital energies:

$$E^{(0)} = \sum_{a=1}^{N} \varepsilon_a \tag{1.29}$$

The Hartree Fock energy E_{HF} is the sum of the zero-order energy $E^{(0)}$ and the the first-order energy correction $E^{(1)}$:

$$E_{HF} = E^{(0)} + E^{(1)} \tag{1.30}$$

where

$$\begin{aligned} E^{(1)} &= < \Psi_0|\hat{H} - \hat{H}_0|\Psi_0 > \\ &= -\frac{1}{2}\sum_{ab} < ab||ba > \end{aligned} \tag{1.31}$$

where $< ab||ba >$ are computed using the following formula:

$$< ij||kl > = < ij|kl > - < ij|lk > \tag{1.32}$$

where $< ij|kl >$ is often referred as a physicist's notation for the two-electron integrals:

$$
\begin{aligned}
< ij|kl > &= (ik|jl) \\
&= \int \chi_i^*(\mathbf{x}_1)\chi_j^*(\mathbf{x}_2) r_{12}^{-1} \chi_k(\mathbf{x}_1)\chi_l(\mathbf{x}_2)\ d\mathbf{x}_1 d\mathbf{x}_2
\end{aligned}
\tag{1.33}
$$

For a closed-shell system, the first-order energy $E^{(1)}$ is given by:

$$
E^{(1)} = -\sum_{ab}^{N/2} [2(aa|bb) - (ab|ba)] \tag{1.34}
$$

where a and b now signify occupied spatial orbital functions.

By using Rayleigh and Schrödinger perturbation theory the second-order energy correction $E^{(2)}$ is given by:

$$
E^{(2)} = \frac{1}{4} \sum_{ab}^{occ} \sum_{rs}^{vir} \frac{| < ab||rs > |^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \tag{1.35}
$$

The second-order energy correction $E^{(2)}$ for a closed-shell system is:

$$
E^{(2)} = 2 \sum_{abrs}^{N/2} \frac{(ar|bs)(ra|sb)}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} - \sum_{abrs}^{N/2} \frac{(ar|bs)(rb|sa)}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \tag{1.36}
$$

where ε_a and ε_b are energies of occupied molecular orbitals while ε_r and ε_s are energies of virtual molecular orbitals. Please note that a, b, r and s each now signify spatial orbitals in the above equation.

The MP2 energy of a molecular system E_{MP2} can be expressed as the sum of the

Hartree Fock energy and the second-order energy correction [132]:

$$\begin{aligned} E_{MP2} &= E_{HF} + E^{(2)} \\ &= E^{(0)} + E^{(1)} + E^{(2)} \end{aligned} \tag{1.37}$$

It is more expensive but from higher order of perturbation theory it is very common to also include the third- and the fourth-order energy correction which gives the E_{MP3} and E_{MP4} energies accordingly. In chapter 2, we show that the MP2 method provides very reasonable energies for the H-passivated Si clusters as well as the MP4 method.

1.4.2 Coupled Cluster Methods

The wave function in Coupled Cluster (CC) methods is taken to have the form:

$$\Psi_{cc} = e^{\hat{\mathbf{T}}}\Psi_0 \tag{1.38}$$

where Ψ_0 is the ground state HF wave function and $e^{\hat{\mathbf{T}}}$ has the following form:

$$e^{\hat{\mathbf{T}}} = \mathbf{1} + \hat{\mathbf{T}} + \frac{1}{2!}\hat{\mathbf{T}}^2 + \frac{1}{3!}\hat{\mathbf{T}}^3 + \frac{1}{4!}\hat{\mathbf{T}}^4 + ... = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{\mathbf{T}}^k \tag{1.39}$$

where $\hat{\mathbf{T}}$ is given by:

$$\hat{\mathbf{T}} = \hat{\mathbf{T}}_1 + \hat{\mathbf{T}}_2 + ... + \hat{\mathbf{T}}_N \tag{1.40}$$

where the $\hat{\mathbf{T}}_i$ operator operates on the ground state HF wave function to give all i-th excited Slater determinants. Because of the presence of $\hat{\mathbf{T}}_i^k$ excitations the $e^{\hat{\mathbf{T}}}\Psi_0$ includes all possible excited determinants like in a full CI. However using all $\hat{\mathbf{T}}_{1-N}$ operators is extremely computationally demanding, therefore truncated coupled cluster methods are frequently used such as the CCSD, CCSDT and CCSD(T) methods. In the CCSD method, $\hat{\mathbf{T}}$ is only the sum of $\hat{\mathbf{T}}_1$ and $\hat{\mathbf{T}}_2$, where $\hat{\mathbf{T}}_1$ and $\hat{\mathbf{T}}_2$ generates excited state functions by exciting 1 or 2 electrons from the occupied molecular orbitals to virtual orbitals:

$$\hat{\mathbf{T}}_1\Psi_0 = \sum_a^{occ}\sum_r^{vir} t_r^a \Psi_r^a \tag{1.41}$$

$$\hat{\mathbf{T}}_2\Psi_0 = \sum_{a,b}^{occ}\sum_{r,s}^{vir} t_{rs}^{ab} \Psi_{rs}^{ab} \tag{1.42}$$

The higher excitation in CCSD arises because $e^{\hat{\mathbf{T}}}$ has the form:

$$e^{\hat{\mathbf{T}}_1+\hat{\mathbf{T}}_2} = \mathbf{1} + \hat{\mathbf{T}}_1 + (\frac{1}{2!}\hat{\mathbf{T}}_1^2 + \hat{\mathbf{T}}_2) + (\frac{1}{3!}\hat{\mathbf{T}}_1^3 + \hat{\mathbf{T}}_2\hat{\mathbf{T}}_1) + ... \tag{1.43}$$

The CCSDT method is very computationally demanding even for a small molecular system due to its inclusion of the $\hat{\mathbf{T}}_3$ term, therefore the CCSD(T) method is often preferred over the CCSDT method since it is much faster because the triples constribution is estimated by perturbation theory.

1.4.3 Density Functional Theory Methods

DFT differs from the above mentioned post-HF methods by focusing on the electron probability density ρ rather than the wavefunction Ψ. Hohenberg and Kohn proved that the ground-state energy and all other ground-state electronic properties are uniquely determined by the electron density distribution in the space [136]. The exact ground-state electronic energy can be written as a function of the electron probability density ρ [137]:

$$E(\rho) = \sum_i \int \psi_i(1)^* \frac{-\nabla_1^2}{2} \psi_i(1)\, d\mathbf{r}_1 - \int \sum_A \frac{Z_A \rho(1)}{R_{A1}}\, d\mathbf{r}_1 + \int\int \frac{\rho(1)\rho(2)}{2r_{12}}\, d\mathbf{r}_1\, d\mathbf{r}_2 + E_{XC}(\rho) \tag{1.44}$$

where ψ_i is the Kohn-Sham orbital, and the electron density function ρ can be constructed from the Kohn-Sham orbitals:

$$\rho = \sum_i |\psi_i|^2 \tag{1.45}$$

The Kohn-Sham orbitals satisfy the self-consistent Kohn-Sham equations:

$$\left(-\frac{1}{2}\nabla^2 + V(\mathbf{r})\right)\psi_i = \varepsilon_i \psi_i \tag{1.46}$$

where ε_i is the energy of the i-th Kohn-Sham orbital. $V(\mathbf{r})$ is given by:

$$V(\mathbf{r}) = V_{ext}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\partial E_{XC}(\rho)}{\partial \rho} \tag{1.47}$$

where V_{ext} is the external potential. $E_{XC}(\rho)$ takes the electron exchange and correlation terms into account. The above Kohn-Sham equations are solved iteratively in the same way as the HF method. The total energy of the molecular system is not simply the sum of the eigenvalues ε_i. Instead some double counted terms should be subtracted:

$$E = \sum_i \varepsilon_i - \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \left(E_{XC}(\rho) - \int \rho(\mathbf{r}) \frac{\partial E_{XC}(\rho)}{\partial \rho} d\mathbf{r} \right) \tag{1.48}$$

Unfortunately, no one knows how to determine the E_{XC} functional exactly, therefore approximations have to be used. One common approximation is the so called local density approximation (LDA) which locally approximates the exchange correlation energy of an inhomogeneous system with that of a homogeneous electron gas [138]. In systems where open electronic shells are involved, local spin density approximation (LSDA) is used. LSDA is similar as LDA except the electron density ρ consists of two separate spin densities:

$$\rho(\mathbf{r}) = \rho_\uparrow(\mathbf{r}) + \rho_\downarrow(\mathbf{r}) \tag{1.49}$$

And the exchange-correlation energy E_{XC} is now a function of both $\rho_\uparrow(\mathbf{r})$ and $\rho_\downarrow(\mathbf{r})$ in the local spin density approximation. The SVWN method, one of the LSDA methods, adopts the Slater exchange [139] and the VWN correlation functional developed by

Vosko, Wilk, and Nusair [140]. The SVWN method predicts reasonable system energies and thus has been widely used. However the electron density does not always vary slowly in the three-dimensional space of the real chemical systems as in the nearly homogeneous electron gas. Therefore various methods with the gradient-corrected exchange and correlation functionals have been developed. They are also called generalized gradient approximation (GGA) methods. The most commonly used GGA exchange functionals include the Becke88 (B) [141], Perdew-Wang (PW91) [142], and Gill96 (G96) [143] functionals. The most commonly used GGA correlation functionals are LYP by Lee, Yang, and Parr [144], Perdew-Wang (PW91) [142], and Perdew 86 (P86) [145]. The third class of functionals which correct for deficiency in GGA functionals are hybrid functionals such as the B3LYP method [133–135]. The B3LYP method we use in this dissertation mixes the exchange energy calculated by the HF method with that obtained from DFT methods in order to increase the accuracy. The B3LYP method consists of Becke 3-parameter hybrid exchange functional combined with the Lee-Yang-Parr correlation functional [144]:

$$E_{XC}^{B3LYP} = \alpha_0 E_X^{HF} + (1-\alpha_0)E_X^{LDA} + \alpha_x E_X^B + (1-\alpha_c)E_C^{VWN} + \alpha_c E_C^{LYP} \qquad (1.50)$$

where the α_0, α_x, and α_c are semiempirical coefficients. The values of α_0, α_x, and α_c are 0.2, 0.72 and 0.81 respectively. E_X^{HF} is the exchange energy determined by the HF method. E_X^{LDA} is the local density approximation (LDA) exchange functional [138]. E_X^B is Becke's exchange correction [141]. E_C^{VWN} is the correlation energy functional developed by Vosko, Wilk, and Nusair [140]. E_C^{LYP} is the correlation energy functional developed by Lee, Yang and Parr [144].

1.5 Outline of the Dissertation

The rest of the dissertation is organized as follows. Chapter 2 describes the initial strategy of the global optimization for the Si_xH_y clusters using the fast AM1 method to prescreen the low energy Si_xH_y clusters. We then applied the more accurate MP2 [132] and B3LYP [133–135] methods to find the global minimum structure at the ab initio level. Our work demonstrated that the GA could be used to find the global minimum of Si_xH_y clusters. This work also showed that the AM1 method did not give the same energy ranking as the MP2 or B3LYP methods especially for the under passivated Si_xH_y clusters, which inspired us to consider reoptimizing the AM1 parameters for the Si atoms [120, 146].

In chapter 3 an iterative global optimizations strategy involving two genetic algorithms, a parametrization genetic algorithm (PGA) and a cluster genetic algorithm (CGA), is proposed and two successful examples of global optimization of underpassivated Si_6H_2 and Si_6H_6 clusters using this strategy are given in the chapter 3. This iterative strategy soon becomes very computationally demanding when it is applied to larger Si-H clusters with more than 10 Si atoms. Therefore we investigated the transferability of the GAM1 parameters obtained form a training set of smaller Si_7H_{14} compounds to evaluate the energies for the larger Si_xH_y clusters with comparable H passivation levels. The details of the transferability strategy and results of cluster global optimization of $Si_{10}H_{14}$, $Si_{10}H_{16}$, $Si_{14}H_{18}$ and $Si_{14}H_{20}$ clusters are included in the chapter 4.

Unfortunately we were still not able to locate the diamond-lattice like structure of $Si_{14}H_{20}$ in 3 CGA runs where 9000 local optimizations were performed, despite that the diamond-lattice like structure for $Si_{14}H_{20}$ giving the lowest energy by both GAM1 and MP2 calculations. The three genetic operators we used in the CGA [153–155] are

probably the main reason for not generating low energy clusters efficiently. Chapter 5 introduces a more efficient and effective cluster genetic algorithm with new mutation methods for the global optimization of Si_xH_y clusters. The CGA with these new mutational methods not only enables us to find the $Si_{14}H_{20}$ global minimum within 3000 local optimizations in 3 out of 5 CGA runs, but also allows us to locate the larger diamond-lattice like $Si_{18}H_{24}$ global minimum structure due to its high efficiency.

In chapter 6 we describe a compensation function coupled with the GAM1 method for the Si_xH_y clusters which allows us to quickly evaluate the relative energies of the Si_xF_y clusters. The fast energy evaluation method enables us to prescreen the low energy Si_xF_y clusters which are then evaluated at a more accurate DFT level. Different Si frameworks are found in the Si_xF_y global minimum clusters compared with those of the Si_xH_y global minimum.

Concluding remarks and directions for future research will be given in the last chapter 7.

CHAPTER 2. GLOBAL OPTIMIZATION OF H-PASSIVATED SI CLUSTERS USING A GENETIC ALGORITHM VIA THE AM1 METHOD

(This chapter is published in *J. Phys. Chem. B* **2002**, *106*, 6997.)

2.1 Introduction

Silicon nanoparticles have attracted a lot of interest because of their potential use as material for fabricating optoelectronic devices. A number of theoretical and experimental studies have investigated the stable geometries for small and medium sized Si only clusters [24,25,28,32–36,38–40,43]. Some type of surface passivated particle will be needed if Si nanoclusters are to be used in a robust commercial device. In this chapter we theoretically investigate the stable geometries formed by H-passivated Si clusters. While it seems most likely for a completely passivated nanocluster that the atoms in the underlying Si cluster should form a diamond lattice like arrangement, it is not obvious how many H atoms are required to completely passivate a Si cluster or indeed whether the diamond like structure for a Si framework actually does correspond to the most stable geometry. Estimation of the number of H atoms binding to a Si cluster is further complicated when the possibility of structures containing rings of Si atoms or the formation of relatively unstable Si-Si multiple bonds is considered. Even a medium-sized cluster can have a huge number of stable local geometries and an exhaustive search for the global structure minimum is extremely hard to perform. Several groups have theoretically explored some of the possible structures for small [55] to medium sized

[56,63–68,78] silicon hydride clusters. We report here a genetic algorithm (GA) we have developed as a strategy to find the global energy minimum corresponding to the most stable structure for various Si_xH_y clusters. The GA is a global optimization strategy inspired by the Darwinian evolution process. Since 1990, GAs have been widely used to search for the global minimum of atomic clusters and molecular agglomerates [28,46, 47,102,110,111]. Some of the previous workers have used simulated annealing methods coupled with either plane wave DFT [56,64] or semiempirical MINDO/3 [66] calculations as an alternative global optimization strategy for the silicon hydride clusters. We choose the GA approach because it can also identify many of the low energy cluster structures while seeking the global minimum. Furthermore, the stochastic nature of the GA is temperature independent and avoids the problem of a geometry getting trapped into a specific local minimum which can occur with simulated annealing as the molecular system is cooling down [96].

The GA explores a large number of different cluster conformations and consequentially requires a computational method which can evaluate the total energy of specific cluster geometry rapidly. Ho and coworkers obtained energy values via a TB-DFT method in their implementation of a GA to find the structures of medium-sized silicon clusters [28]. Presently these TB-DFT methods are only parametrized for use with a single element. Due to the binary nature of the Si_xH_y clusters we perform our energy evaluations with the semiempirical AM1 method [120]. Once the GA selects a specific cluster conformation the AM1 energy calculations are relatively fast enabling a rapid gradient based optimization to the nearest local minimum. The AM1 energy has the important feature of a solid quantum mechanical basis so that no additional empirical information is required to identify which atom pairs in the cluster combine to form bonds. Thus a H atom in a Si_xH_y cluster has the option of forming a bond with either

a Si atom or another H atom depending on which bond gives the lower cluster total energy. As a consequence, when a Si cluster is completely H-passivated, any remaining excess H atoms can pair up to form hydrogen molecules.

Simulated annealing optimization also involves a large number of energy function evaluations and Meleshko *et al.* selected the semiempirical MINDO/3 method for their procedure [66]. The computer times needed for either MINDO/3 or AM1 calculations are comparable, but experience by us and others [122, 126] suggest that AM1 is presently the semiempirical method which gives the most reliable optimum geometries and relative total energies for Si containing compounds. However, since we obtain several different geometries close in AM1 energy in a GA run on a specific Si_xH_y cluster, we also examine the relative energy ordering of a small subset of the low AM1 energy structures by performing ab initio Second Order Møller-Plesset (MP2) and Density-Functional Theory with the B3LYP functional (B3LYP) calculations using the optimized AM1 geometries as starting structures.

In the next section we describe in more detail the GA and the methods used for evaluating the cluster total energy. In the third section we first present results for application of the GA to a series of $Si_{10}H_m$ clusters and show how the cluster geometry evolves as the Si_{10} cluster becomes completely passivated with H. We then discuss the efficiency of the GA. In the final part of the results section we describe the global optimization of $Si_{14}H_{20}$ and the formation of a diamond-lattice structure for the Si atoms. Concluding remarks are given in the final section of the chapter.

2.2 Computational Method

To find the more stable isomers for a Si_xH_y cluster requires a large number of total energies at different cluster geometries to be evaluated. Standard AM1 semiem-

pirical energy calculations [120] are used in the GA to facilitate a computationally fast comparison of the relative stability of various optimized Si_xH_y geometries. We define the cluster binding energy BE_i for the i-th structure as

$$BE_i = E_{Si_xH_y} - xE_{Si} - yE_H \tag{2.1}$$

where $E_{Si_xH_y}$ is the total energy of the optimized cluster geometry and E_{Si} and E_H are the ^3P and ^2S ground state energies of the Si and H atoms, respectively.

Although AM1 is time-efficient for the GA, it could not be relied on to accurately order the stability of different cluster isomers close in energy. Hence, we pick up the best several structures found by our GA, and then further optimize them by ab initio Second Order Møller-Plesset (MP2) [132] and Density-Functional Theory using the B3LYP functional (B3LYP) [133–135] calculations. As suggested by Raghavachari and Rohfling [23] the MP2 and B3LYP geometry optimizations were performed using the Hay-Wadt effective core potential (ECP) and valence basis set for Si [147] augmented with a d function [23] and the Dunning-Hay basis for H [148]. The AM1 and MP2 calculations were performed using the GAMESS program [149] and the B3LYP calculations were performed with Gaussian 94 [150]. Raghavachari and Rohfling show the HW-ECP and basis set at the MP4 level does an excellent job of reproducing the results obtained with all electron calculations on Si only clusters [23]. In the present work we find calculations at the MP2 and MP4 levels to give essentially the same energy order for the $Si_{10}H_4$ and $Si_{10}H_{16}$ series of clusters; where some energy reordering does occur for some of the high energy $Si_{10}H_{16}$ structures which are within 2 kcal/mol of each of other. Krüger and Sax have recently found the 6-31G** B3LYP functional, as well as other functionals,

to give reasonable geometries and relative energies for different isomers of Si_2H_2 and Si_2H_4 [79]. We find the HW-ECP and basis with the B3LYP functional to reproduce the Krüger and Sax calculations. As a consistency check on the ab initio calculations, we report the results of both B3LYP and MP2 calculations instead of performing the much more computationally demanding MP4 calculations. A complete tabulation of the final cluster geometries and relative energies are available from the authors.

The GA works by randomly selecting and mating the more fit individuals in a generation to produce next generation of offspring, where the fitness is a measure of the energetic stability for an individual optimized cluster geometry. The GA selects new cluster conformations which are likely to have lower optimized energies than structures in previous generations. The population size of the GA is set to N, which is the total number of Si and H atoms in the cluster to be globally optimized. Our experience finds that a variable N is needed to prevent premature GA convergence for the larger clusters and to avoid low convergence efficiency for the smaller clusters. We generate N initial structures of the H-passivated Si clusters by first putting the Si atoms into a cubic box randomly, and then we randomly arrange H atoms outside the box. That is, we generate initial structures with the assumption that the stable structures of the H-passivated Si cluster will tend to have H atoms on the surface instead of inside. However to reduce this bias, we also allow a small number of H atoms inside the box nearby the borderline. Then we optimize the geometry of the above N structures by the AM1 method to create our ancestor population. The fitness f_i of the i-th initial individual cluster is based on a Boltzman distribution:

$$f_i = \exp\left(\frac{C - \mathrm{BE}_i}{NRT}\right) \tag{2.2}$$

where BE_i is the optimized AM1 cluster binding energy, and C is a constant, whose value is set to be the lowest energy obtained from the initial N ancestors. C is used to rescale the fitness f_i into a rather small range. The value of C will not affect the probability of one parent being selected. RT is set to 0.2 eV/atom, which corresponds to a temperature of 2300 K. Roulette selection is used to select N random parent couples for the next generation. The probability P_i of the i-th geometry being selected is determined by the normalized fitness of that cluster:

$$P_i = f_i / \sum_i f_i \tag{2.3}$$

The same individual cluster geometry can be selected more than once during the roulette selection. Each resulting parent is then randomly rotated about their center of mass. From previous work on atomic clusters we adopt three possible mating methods. One approach is to take the arithmetic mean of the cartesian coordinates from two parent geometries [47, 110]. The second method is to take a fragment of Si atoms from one parent and replace it with the fragment of the other parent with the same number of atoms, and also cut a fragment of H atoms from one parent and replace it with a fragment of H atoms in the other parent with the same number of atoms [47, 110]. The last method is to cut each parent into halves and then recombine one half from each parent to generate the offspring [102]. The first and second mating processes produce drastic mutations, while the third method retains much of the local structures of both parents [102]. The probabilities used for selecting the first, second and third methods are 25%, 25% and 50%, respectively. After the mating process, N offspring are generated and fully optimized again by the AM1 method. We will replace the least

stable structure of its parent population with a new offspring if it has relatively lower energy. A number of different convergence strategies have been proposed [46, 47]. To enable an equal treatment on the different sized clusters we found it best to take the GA as being converged when the lowest energy structure remains the same for the last $3N$ continuous generations. The final structures we tabulate correspond to the lowest optimized energies found during an entire GA run.

2.3 Results and Discussion

We start this section by examining the effectiveness of the GA at identifying chemically reasonable low energy structures of Si clusters. We applied the GA to a series of $Si_{10}H_m$ clusters with $m = 4, 8, 12, 16$ and 20. Clearly the Si framework will only be partially passivated by H when m is small, but it is less clear how many H atoms are needed to ensure all the Si atoms in a cluster are tetrahedrally coordinated and fully saturated with four bonds. The determination of the number of H atoms needed is further complicated by the Si atoms in the cluster being able to form ring structures or perhaps even Si=Si double bonds. One goal of the present work is not just to develop a method which finds the most stable structures of binary clusters with specific stoichiometries, but also to show that our GA approach can determine how many H atoms are preferred in a cluster. Following the discussion of the $Si_{10}H_m$ results and the effectiveness of the GA for investigating passivation of Si clusters, we extend our calculations to the $Si_{14}H_{20}$ cluster. The Si framework in $Si_{14}H_{20}$ can be arranged as a diamond-lattice like fragment, where this fragment has 20 Si dangling bonds which can be saturated with the 20 H atoms. We apply the GA to $Si_{14}H_{20}$ to determine whether the diamond like structure is the most stable structure for this cluster.

The key structural results for the series of $Si_{10}H_m$ clusters are presented in fig-

ures 2.1 through 2.6. Since we do obtain different relative stabilities by the semiempirical and non-empirical calculations each structure in the figures is labeled by c-X where c is an alphabetical character ordering the AM1 structures by decreasing binding energy and X is a number indicating the relative binding energy as calculated by the B3LYP functional. Generally we found the binding energy order calculated by MP2 follows the B3LYP ordering, but in the cases where they do differ we add the MP2 order in parenthesis to the c-X structure label. Thus structure d-9(7) of $Si_{10}H_4$ in figure 2.2 is the fourth AM1, ninth B3LYP and seventh MP2 lowest energy structure. The figures show the B3LYP optimized geometries obtained by starting with coordinates from the AM1 optimized structure produced by the GA. Despite the problems with the relative energy ordering, the geometries obtained with the AM1, MP2 and B3LYP for a particular structure all exhibit the same general features. To illustrate the degree of typical similarity we show in figure 2.1 AM1, MP2 and B3LYP structures obtained for the a-10 $Si_{10}H_4$, f-1 $Si_{10}H_8$ and d-11 $Si_{10}H_{20}$ clusters. It is worth noting that the lines connecting pairs of Si atoms are drawn by the graphics program only when the two atoms are within 2.45 Å of each other and a line only approximately indicates that a bond is actually formed between a pair of atoms. Below we present in more detail the different structural features found for the various $Si_{10}H_m$ clusters in turn.

$Si_{10}H_4$. Figure 2.2 shows the 10 structures with lowest AM1 energy obtained by the GA for $Si_{10}H_4$ and their rankings with the B3LYP and MP2 methods. The $Si_{10}H_4$ cluster attempts to form compact structures presumably so as to reduce the number of Si atoms not tetrahedrally coordinated. Each H atom binds to a different Si atom and each of these hydrogenated Si atoms are tetrahedrally coordinated with 3 additional neighboring Si atoms. The 3 most stable AM1 structures a-10, b-8(9) and c-7(8) differ

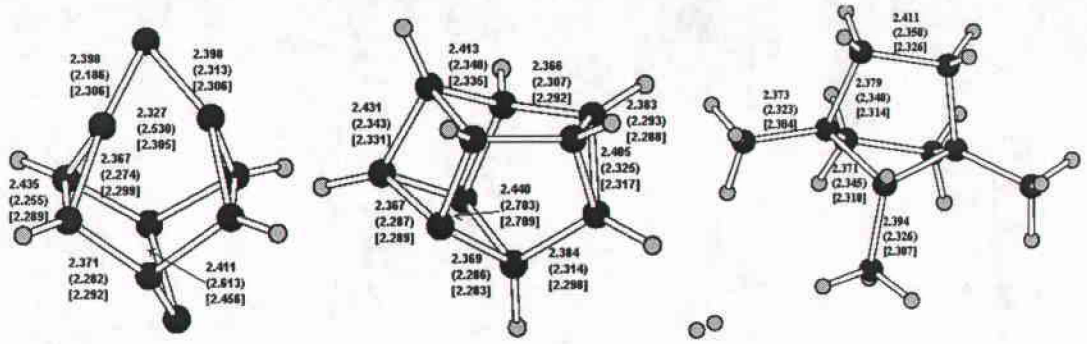


Figure 2.1: Representative optimized cluster structures obtained with the AM1, B3LYP and MP2 methods. The structures are: (a) $Si_{10}H_4$ a-10, (b) $Si_{10}H_8$ f-1, and (c) $Si_{10}H_{20}$ d-10(5) dissociated into $Si_{10}H_{18}$ and a H_2 molecule. The three bond lengths are listed in the order AM1, (B3LYP) and [MP2] in Å.

in energy by less than 0.5 kcal/mol. Structures a-10, b-8(9) and c-7(8) are the only structures we find that contain 2 Si atoms which are only doubly coordinated. The MP2 and B3LYP calculations show these structures to be the least stable of those shown in figure 2.2. The Si parameters used by the AM1 method were obtained from a set compounds where the Si atom is generally coordinated by 4 other atoms [146] and the AM1 parameters most likely are not optimally designed to reliably treat low coordinate Si atoms. The present calculations suggest that AM1 overestimates the stability of a biradical localized on on the 2 coordinate Si atoms. Figure 2.1 shows the AM1, B3LYP and MP2 methods produce the same basic optimized a-10 geometry, even though B3LYP and MP2 compute a-10 to be 56.2 and 82.3 kcal/mol, respectively, higher in energy than the e-1 structure. E-1 is the most stable structure calculated by B3LYP and MP2. Energetically structures e-1 through j-4 differ by 3.4 kcal/mol in the AM1 calculations, while the B3LYP and MP2 calculations produce differences of 24.6 and 28.6 kcal/mol, respectively. Geometrically structures e-1 and j-4 have a

similar Si_{10} skeleton made up of 5 Si_4 rings but with different H atom arrangements. A second type of Si_{10} framework containing one Si_3 ring is found for the i-2, g-3, f-5 and h-6 structures. Presumably the Si_3 ring is more strained than the larger rings causing the i-2 structure to be 4.0 (B3LYP) and 3.8 (MP2) kcal/mol less stable than e-1. The 3 membered rings in i-2 and g-3 have one of the Si atoms H passivated, whereas f-5 and h-6 have two of the Si atoms H passivated further enhancing the Si_3 ring strain. Even though j-4 does not contain any 3 membered Si rings, it maybe less stable than e-1 because two of the 3 coordinate Si atoms bind to 3 H passivated Si atoms thereby producing 2 isolated dangling bonds; in all the other structures every 3 coordinate Si atom is adjacent to at least another 3 coordinate Si atom.

$Si_{10}H_8$. Figure 2.3 shows the 10 most stable structures obtained for the $Si_{10}H_8$ cluster using the AM1 energy and the GA. The 4 extra H atoms continue to complete the tetrahedral coordination favored by Si and in a number of structures the Si skeleton from $Si_{10}H_4$ is retained. For example, structures b-6, c-2 and d-3(4) in figure 2.3 have the same Si_{10} framework as for e-1 of $Si_{10}H_4$; while the $Si_{10}H_8$ structures f-1, a-4(3), h-5, g-7 and e-8 have the same Si skeleton found for $Si_{10}H_4$ i-2 structure. In all of the $Si_{10}H_8$ structures, Si only binds to one H atom and figure 2.3 suggests the two remaining un-passivated Si atoms are always 3 coordinate. The AM1 energy variation for the different $Si_{10}H_8$ structures is much larger than found for $Si_{10}H_4$ and ranges over 8.6 kcal/mol for structures a-4(3) to f-1. At the B3LYP and MP2 levels we find f-1 to be the most stable structure. We think this structure is favored because the two unpassivated Si atoms can get within 2.7 Å of each other along the four membered ring diagonal and form a weak Si-Si bond. A comparison of the AM1, B3LYP and MP2 structures for f-1 are given in figure 2.1. The bond lengths predicted by the B3LYP

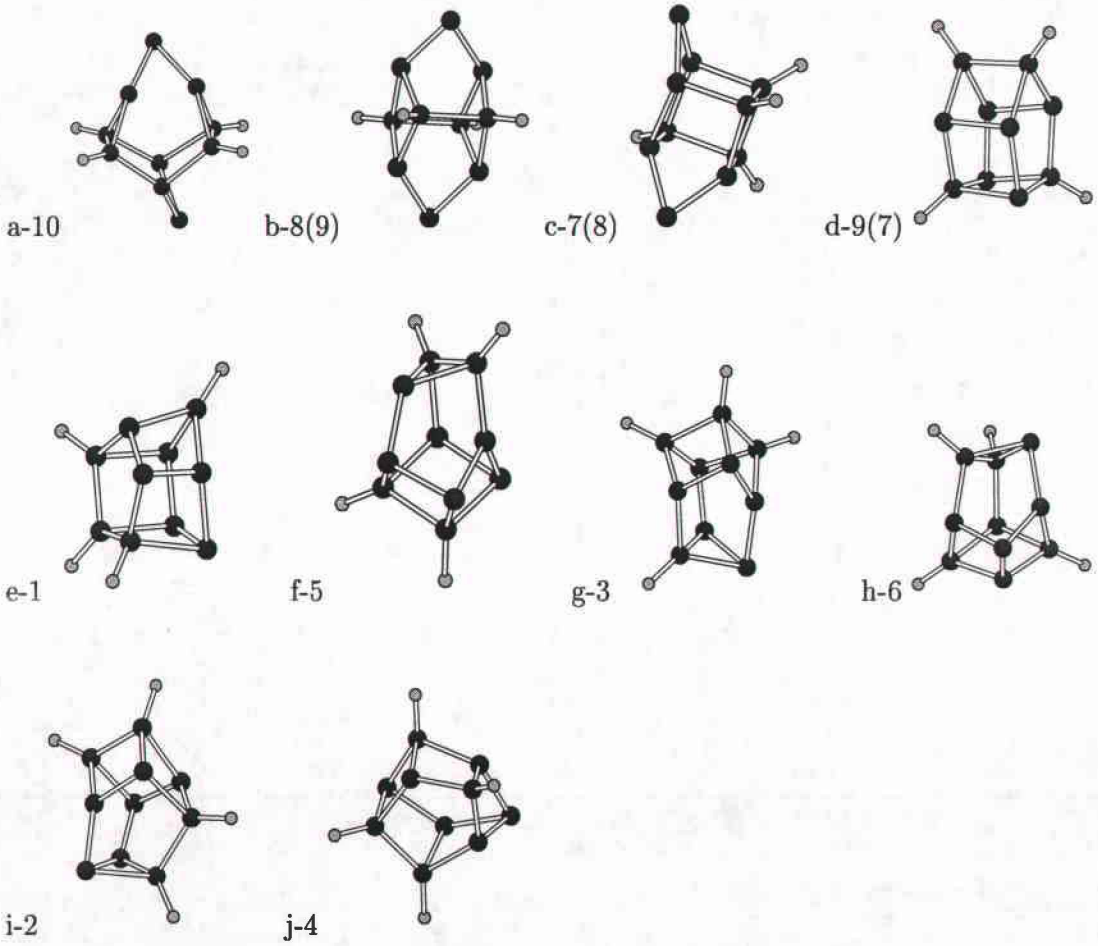


Figure 2.2: Ten lowest AM1 energy structures for $Si_{10}H_4$. The numbers label the B3LYP and MP2 energy rankings as explained in the text.

and MP2 methods are very similar with a deviation of less than 0.02 Å, while the AM1 distances are within 0.1 Å of the ab initio values. A comparison of the f-1 structure with the $Si_{10}H_4$ a-10 structure shows the Si-Si bond lengths of $Si_{10}H_8$ are longer due to more H passivation, and they are closer to the experimental 2.3517 Å Si-Si distance in a diamond lattice structure [151]. F-1 is 2.5 (B3LYP) and 5.3 (MP2) kcal/mol more stable than c-2, and both structures have the 2 unpassivated Si diagonally opposite each other in a four membered ring. However, in contrast to what was found for $Si_{10}H_4$, f-1 appears to be stabilized by a three membered ring of Si-H atoms relative to c-2. The next higher energy structures a-4(3) and d-3(4) resemble f-1 and c-2, respectively, but with the unpassivated Si atoms at different locations. The MP2 energy favors the 3 member ring and has a-4(3) 1.6 kcal/mol more stable than d-3(4) whereas B3LYP gives 2.6 kcal/mol for the reverse stability.

$Si_{10}H_{12}$. There are now enough H atoms in the $Si_{10}H_{12}$ clusters shown in figure 2.4 to ensure all of the Si atoms are four coordinate. Figure 2.4 also shows all of the structures to have 5 rings of Si atoms, this number of rings must be formed if each of the Si atoms is four coordinate and no multiple Si-Si bonds are formed. Typically 8 of the Si atoms bind with a single H atom and the two remaining Si atoms form SiH_2; although in the higher energy structures g-8(9) and e-9(8) a SiH_3 group is formed and one Si atom is tetrahedrally coordinated by 4 other Si atoms. Presumably because the Si coordination is now complete the low energy structures obtained by AM1 match better with the B3LYP and MP2 energy ordering. The two lowest energy structures b-1 and a-2 appear to have a Si framework quite different to those previously found for the $Si_{10}H_4$ and $Si_{10}H_8$ clusters and they differ in energy by 2.7, 1.3 and 0.7 kcal/mol at the AM1, MP2 and B3LYP levels, respectively. The next structure c-3(4) is 1.4

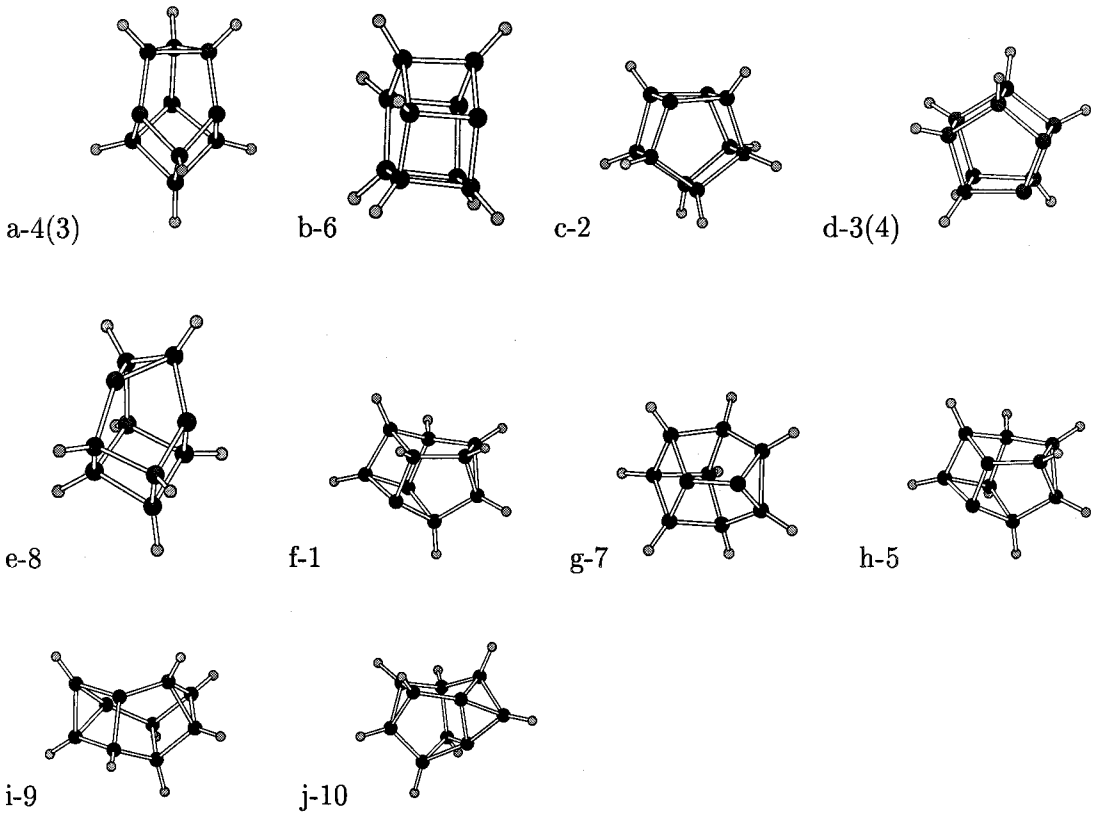


Figure 2.3: Ten lowest AM1 energy structures for $Si_{10}H_8$.

(AM1), 7.4 (B3LYP) and 10.2 (MP2) kcal/mol above the 2nd lowest energy structures and has a Si skeleton easily derived by adding H atoms to structures c-2 or d-3 found for $Si_{10}H_8$. Structure d-4(3) is only 2.5 (AM1), 0.8 (B3LYP) and -0.03 kcal/mol (MP2) above c-3(4) and corresponds to an opening of the Si_3 ring of the $Si_{10}H_8$ f-1 structure. New Si_{10} skeletons are also found for the f-5, h-6 and i-7 structures. Structures e-9(8) and g-8(9) have the same Si_9 framework but place the SiH_3 group at different positions and have their largest energy difference of 0.8 kcal/mol at the AM1 level. We also find the numbers of smaller three- and four-member Si rings in figure 2.4 are less than those of in the less H-passivated Si clusters suggesting there is now less strain on the Si_{10} framework.

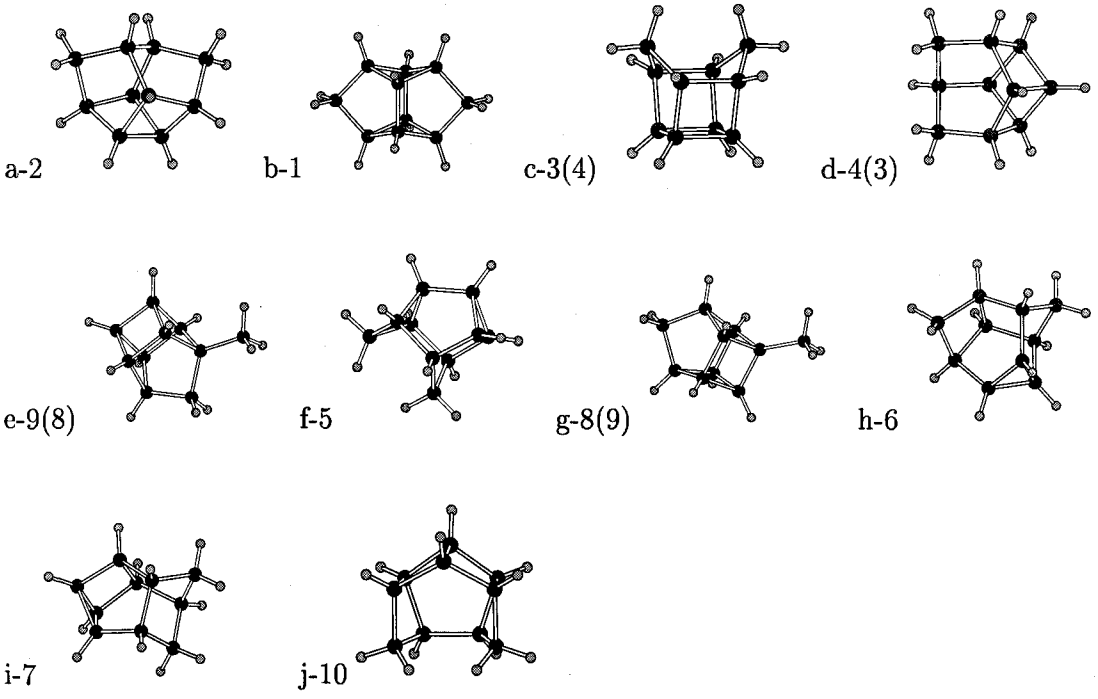


Figure 2.4: Ten lowest AM1 energy structures for $Si_{10}H_{12}$

$Si_{10}H_{16}$. In figure 2.5 we show the 20 $Si_{10}H_{16}$ structures corresponding to the 20 lowest AM1 optimized energies we obtain with the GA. Since the j-1 structure had the lowest B3LYP and MP2 energy among the first 10 AM1 structures we display 10 additional structures so as to ensure that other low energy structures, such as m-2(4), are not overlooked. The H atoms in the $Si_{10}H_{16}$ clusters continue to aid forming four coordinate Si atoms. Consistent with the formation of 4 new Si-H bond, each of the Si frameworks in figure 2.5, apart from c-19 and l-20, now contain only three rings of Si atoms and as a consequence the Si framework is substantially different to that found in $Si_{10}H_{12}$. Unfortunately, the AM1 energy values exhibit a much narrower range of energy variation, all of the structures in figure 2.5 are within 5.8 kcal/mol of each other. The structure with lowest B3LYP and MP2 energy, j-1, is particularly interesting because the Si framework adopts a diamond lattice like arrangement suggesting the H passivated Si clusters do prefer a diamond like structure for the Si atoms. Structures m-2(4), b-3(2) and a-4(3) are 5.1, 5.5 and 5.7 kcal/mol at the B3LYP level and 4.5, 2.8 and 3.2 kcal/mol at the MP2 level less stable than the j-1 structure. It is also interesting that the GA manages to find two enantiomers, clusters d and e, which are mirror images of each other.

Figure 2.5 also contains two structures with 4 rings of Si atoms, c-19 and l-20, which corresponds to a separate $Si_{10}H_{14}$ cluster and H_2 molecule. Thus, the GA is giving the first indication that the Si_{10} framework is starting to become saturated when there are 16 H atoms. Energetically c-19 is computed to be 24.6 (B3LYP) and 15.8 kcal/mol (MP2) less stable than the most stable B3LYP structure j-1. The binding energy expression given by equation (1) effectively becomes for the c-19 and l-20 structures

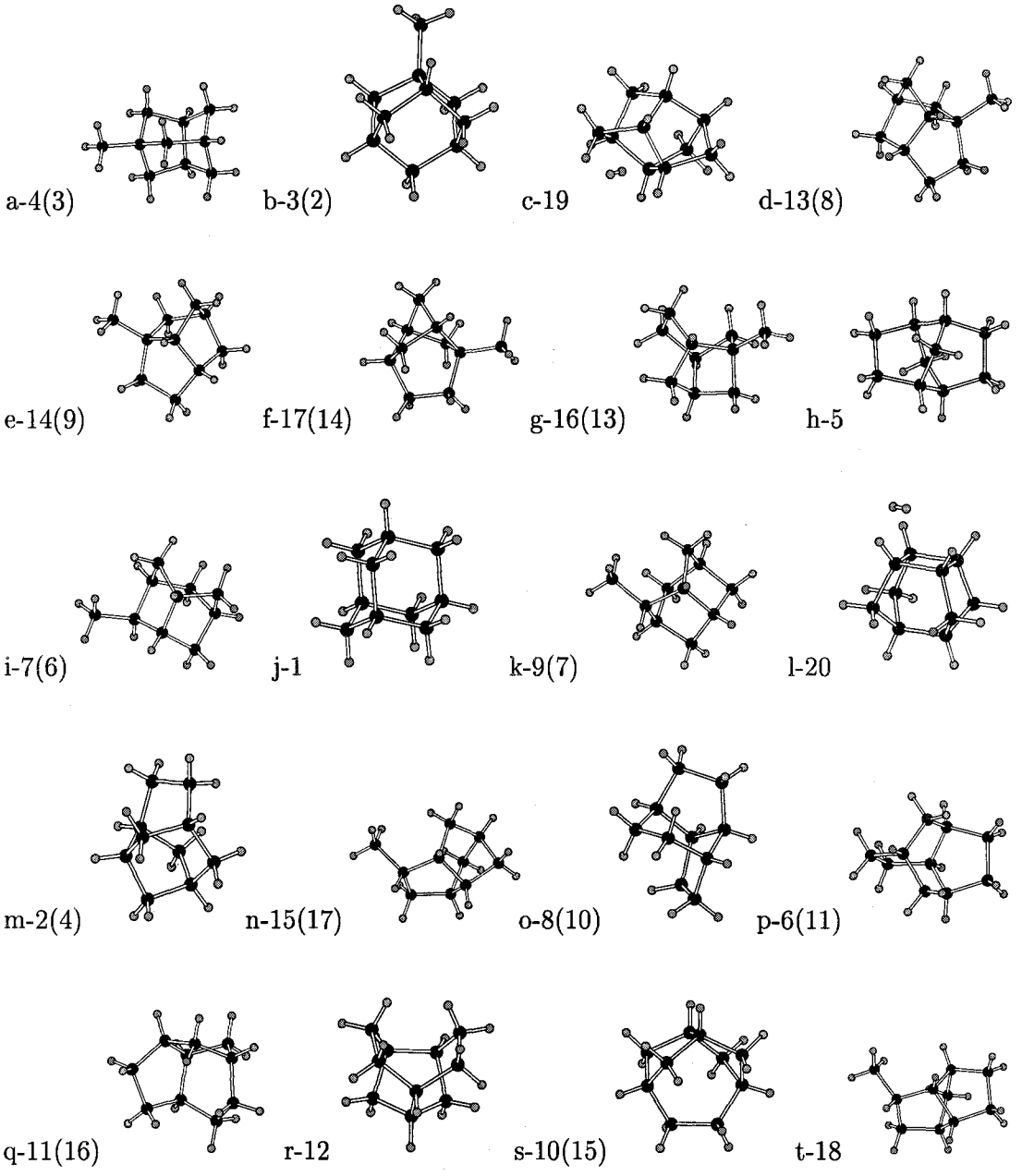


Figure 2.5: Twenty lowest AM1 energy structures for $Si_{10}H_{16}$.

$$\mathrm{BE}_i = E_{\mathrm{Si_{10}H_{14}}} + E_{\mathrm{H_2}} - 10E_{\mathrm{Si}} - 16E_{\mathrm{H}} \tag{2.4}$$

Equation (4) suggests that if $Si_{10}H_{14}$ was the preferred passivation product over $Si_{10}H_{16}$ at zero K, then the magnitude of the BE for c-19 should be larger than for the j-1 structure. That is, our GA strategy automatically identifies whether a lower $Si_{10}H_m$, $m < 16$, passivated cluster is a more stable structure. The $Si_{10}H_{14}$ cluster shown in figure 2.5 for c-19 should represent the lowest energy structure possible for $Si_{10}H_{14}$, although since c-19 and l-20 have the highest B3LYP and MP2 energies in the 20 structures shown we may not be sampling enough of the $Si_{10}H_{14}$ low energy structures to correctly identify the most stable $Si_{10}H_{14}$ cluster. At higher temperatures, the larger combined entropy of $Si_{10}H_{14}$ and H_2 should eventually favor their formation. An estimate for the entropic contribution can obtained from scaled AM1 frequency calculations [152]. In general, we do not find the preferred passivation product to be significantly altered when we include AM1 zero point energy corrections in equation (4). However, at 298 K we estimate from the vibrational calculations that $Si_{10}H_{14}$ and H_2 will only be preferred when their total BE is less than 6 kcal/mol above the $Si_{10}H_{16}$ BE. Since both the B3LYP and MP2 calculations give a BE for c-19 considerably larger that 6 kcal/mol than the j-1 structure we conclude $Si_{10}H_{16}$ has the preferred passivation.

$Si_{10}H_{20}$. Figure 2.6 shows the 20 lowest AM1 structures found for $Si_{10}H_{20}$. At the B3LYP and MP2 levels the 4 most stable $Si_{10}H_{20}$ structures given in figure 2.6 correspond to different isomers of a tetrasilyl cyclopentasilane. It is difficult to determine which isomer is the most stable since the 4 structures are within 0.8 (B3LYP) and 0.5 (MP2) kcal/mol of each other. The next 12 higher non-empirical energies correspond

to structures containing a $Si_{10}H_{18}$ cluster and one H_2 molecule. At the B3LYP level structures f-5(7) and t-16 range from 9.3 to 15.2 kcal/mol above the i-1(4) lowest energy structure, while at the MP2 level d-10(5) through t-16 are only 5.7 to 9.7 kcal/mol above the n-4(1) structure. For comparison, the AM1 energy varies only 5.3 kcal/mol for all 20 structures. Figure 2.1 shows the geometry calculated by AM1, B3LYP and MP2 for the d-10(5) structure in more detail and again confirms the B3LYP and MP2 methods predict very similar geometries, while AM1 exhibits larger geometric differences but preserves the main structural features. From the B3LYP and MP2 energies we can conclude that $Si_{10}H_{20}$ is the more stable passivated structure at 0 K. At 298 K we estimate via AM1 frequency calculations an entropic contribution favoring H_2 loss when the $Si_{10}H_{18}$ and H_2 energies are within 11 kcal/mol of the $Si_{10}H_{20}$ electronic energy. Thus we anticipate a relatively low temperature should cause $Si_{10}H_{20}$ to lose H_2 and produce $Si_{10}H_{18}$.

The four remaining structures g-17(18), a-18(17), e-19 and k-20 in figure 2.6 consist of a $Si_{10}H_{16}$ cluster and two H_2 molecules. The energetic cost at 0 K of forming these two hydrogen molecules ranges from 22.2 to 26.7 kcal/mol for B3LYP and 13.2 to 16.5 kcal/mol for MP2 and we can again conclude that the passivation in $Si_{10}H_{16}$ is less favored relative to $Si_{10}H_{20}$. Interestingly, following the same relative AM1 binding energy ordering, the $Si_{10}H_{16}$ structures in a-18(17) and g-17(18) in figure 2.6 match with the structures e-14(9) and q-11(16), respectively, shown in figure 2.5. However, the $Si_{10}H_{16}$ clusters shown in the e-19 and k-20 structures in figure do not match with any of the structures shown in figure 2.5. This result demonstrates that the GA in its search for a global minimum does not exhaustively find all of the other low energy structures possible for a particular cluster type. Entropic contributions at 298 K, estimated via AM1 frequency calculations stabilize the $Si_{10}H_{16}$ formation by about

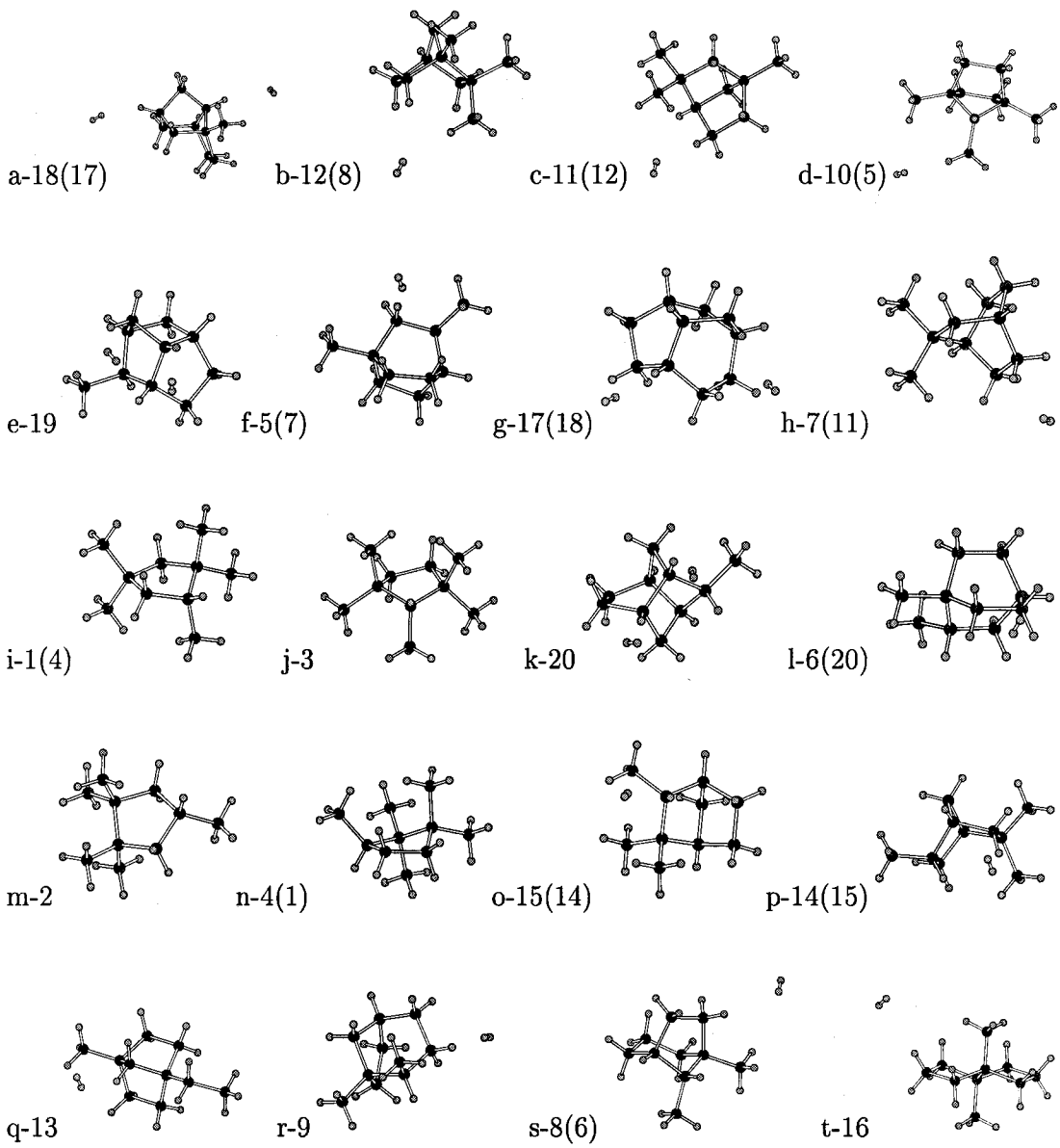


Figure 2.6: Twenty lowest AM1 energy structures for $Si_{10}H_{20}$.

20 kcal/mol relative to $Si_{10}H_{20}$ suggesting that $Si_{10}H_{16}$ is the preferred passivation of the Si_{10} cluster at higher temperatures.

$Si_{10}H_m$ structure evolution. To summarise: the Si_{10} framework changes from compact cages when only partially passivated by H atoms to more open structures when $m = 12, 16$ and 20 with all the Si atoms becoming four coordinate. Substantial changes in the Si_{10} framework occur between $Si_{10}H_{12}$ and $Si_{10}H_{16}$ because of the Si ring opening which must take place if the structures are to accept additional H atoms. Miyazaki *et al.* found a similar evolution of compact to open structures in their investigation of Si_6H_m clusters with $0 \leq m \leq 14$ [64]. Our GA results also suggest that at 0 K the Si_{10} framework to be fully passivated with 16 H atoms, although we do find that $Si_{10}H_{20}$ can form tetrasilyl cyclopentasilane isomers. The formation of these $Si_{10}H_{20}$ isomers is consistent with the Miyazaki *et al.* results which find the complete family of Si_6H_m clusters to be stable. Entropic contributions at higher temperatures appear to favor $Si_{10}H_{16}$ as the preferred passivation of the Si_{10} clusters. Unfortunately, the large number of energy evaluations needed by the GA means that it is only practical to estimate the entropic stabilization of differing amounts of H passivation only after identifying possible candidate structures at 0 K.

Effectiveness of the Genetic Algorithm. The above results for the $Si_{10}H_m$ clusters demonstrate the GA to be a useful tool for locating several of the low energy isomers possible for a particular cluster without biasing the calculations towards some preconceived notion of what the most stable structure might be. The GA serves to dramatically reduce the number of energy function evaluations while randomly evaluating many of the more stable isomers of a cluster. To test the consistency of our GA algorithm we performed two additional GA runs on each specific Si_xH_y cluster using

different randomly chosen starting structures. We found all three GA runs to give the same lowest energy structures for the $Si_{10}H_m$ clusters with $m = 4, 8, 12, 16$, while for the $Si_{10}H_{20}$ and $Si_{14}H_{20}$ slightly different lowest energy structures were produced but all with 3 kcal/mol of each other. In figure 2.7 we show the search evolution in one of our typical GA runs on the $Si_{10}H_{16}$ cluster, we find similar evolution patterns for all the other Si cluster described in this chapter. The dark line gives the binding energy for the $Si_{10}H_{16}$ structure with lowest energy, the lighter line gives the binding energy for the structure with highest energy in the population, the 26-th structure. In the first 25 generations, the binding energy of the most stable structure drops rapidly to -1846.2 kcal/mol and then is only slowly lowered in energy during the remaining generations. The energy of the 26th structure in the population decreases much more slowly, and this is a consequence of the mating and mutation process continuing to form unstable structures in each generation. After generation 70 only small energy changes occur for lowest and highest energy structures and we take the GA to be converged when the lowest energy structure does not change over the next $3N$ generations, where N is the total number of atoms in the cluster. The average number of generations needed to reach convergence for the $Si_{10}H_m$, $m = 4, 8, 12, 16, 20$ clusters was 83 ± 19, 93 ± 18, 93 ± 9, 136 ± 22 and 155 ± 26, respectively. Since we take the population of a generation as N, the average total number of AM1 geometry optimizations performed was 1162 ± 259, 1680 ± 316, 2039 ± 197, 3527 ± 578 and 4650 ± 769 for $m = 4, 8, 12, 16, 20$. Thus, with our implementation of the GA, the number of structures evaluated approximately grows only as $\sim O(N^2)$. Although this growth still requires the energy evaluation in the GA to be computed rapidly with some simple function such as provided by the AM1 semiempirical method. The present work shows that the semiempirical AM1 method to be useful for ranking energetically the different structures generated by the

GA. Unfortunately, the close spacing of the total AM1 energies for different structures requires that B3LYP or MP2 geometry optimizations are also performed to more accurately rank the final structures obtained by the GA. Indeed, one potential drawback to present approach is that the GA searches for the structure with lowest AM1 energy while the more physically correct structure might have significantly higher AM1 energy and not be detected when the higher level B3LYP and MP2 calculations are performed. Nonetheless, the GA appears to be doing a good job of providing a relatively small set of low energy candidate structures for further energy analysis by MP2 or B3LYP calculations. Our experience so far suggests this set from the GA contains structures which can give rise to the MP2 or the B3LYP global minimum. Obviously this claim can only be disproved if another optimization method produced even lower energy structures. The inability to directly perform B3LYP or MP2 calculations in the GA is most clearly demonstrated by noting that the actual B3LYP and MP2 calculations on the 10 or 20 lowest AM1 energy structures found by the GA utilized more computer time than all of the AM1 calculations performed in a GA run. Apart from when structures are close in energy, we find the B3LYP and MP2 methods both give very similar cluster geometries and rankings suggesting that we could just use one of the methods for an accurate structure analysis.

The $\sim O(N^2)$ growth of the number of functions evaluated coupled with the increased computational demands of AM1 calculations on larger clusters requires that other tricks be used in the GA search for the global structure of larger clusters. Below we describe the results of the GA applied to $Si_{14}H_{20}$. The best and safest scheme we have found so far for improved performance is to reduce both the SCF density convergence threshold to 10^{-3} a.u. and gradient convergence threshold to 10^{-3} a.u. in the geometry optimization when generating the initial population of $Si_{14}H_{20}$ structures.

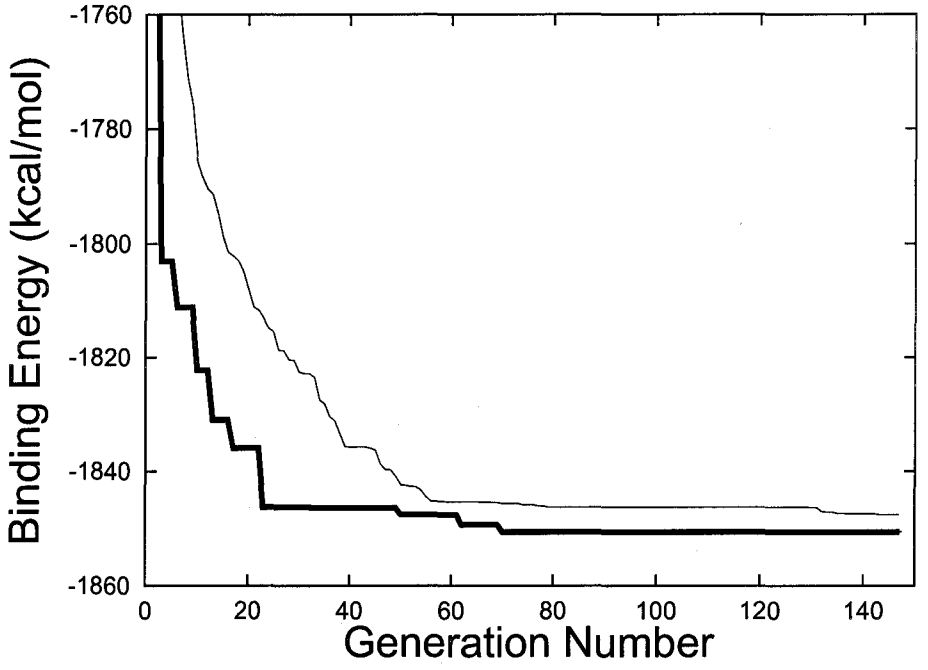


Figure 2.7: GA evolution for $Si_{10}H_{16}$. The graph shows how the cluster binding energy of the lowest (dark line) and highest (light line) energy individual in each population varies with generation number.

We find this significantly shortens the total CPU time without sacrificing the generality and reliability of the GA. The final $Si_{14}H_{20}$ structures described below were produced from 278 generations consisting of 34 structures. We are presently investigating other strategies for more efficient application of the GA to larger passivated Si clusters.

$Si_{14}H_{20}$. An important consideration in the study of nanomaterials is what size particle will cause the bulk like structure and electronic properties to be manifested. Silicon in passivated Si_xH_y clusters can in principle readily form a diamond like lattice. For example, figure 2.8 shows the optimized geometry obtained at the AM1, MP2 and B3LYP levels for the $Si_{14}H_{20}$ cluster, where the Si atoms were given an initial geometry

with a diamond-lattice like arrangement. The average nearest-neighbor Si-Si distance at 2.371 ± 0.009 (AM1), 2.324 ± 0.003 (B3LYP) and 2.312 ± 0.001 Å (MP2) are reasonably close to the experimental 2.3517 Å distance in bulk Si [151], and all of the Si atoms are tetrahedrally coordinated. However, the final structure obtained in an energy gradient based geometry optimization is dependent on the initial atom coordinates, and the final structure obtained is usually the local minimum closest to the starting geometry. If there was another $Si_{14}H_{20}$ structure at lower energy, a standard gradient based geometry optimization would not find it unless an appropriate starting geometry was given. The GA, on the other hand, does not have a single starting geometry nor is it bound to a particular valley on the potential energy hypersurface. The GA, in principle, should find the structure corresponding to a true global minimum.

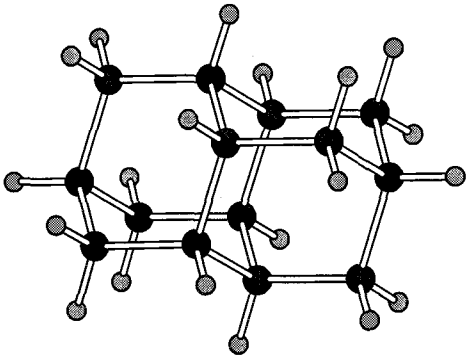


Figure 2.8: The $Si_{14}H_{20}$ cluster where the Si framework has a diamond lattice like structure.

Thus as a final test of the GA, using the AM1 energy function, we have attempted to determine whether the diamond like Si structure shown in figure 2.8 is the most stable structure for $Si_{14}H_{20}$. We use the refined method outlined in the previous subsection. The resulting 20 lowest AM1 energy structures we obtain for $Si_{14}H_{20}$ from the GA

are shown in figure 2.9. The GA is clearly successful as it finds 20 structures all with AM1 energies less than that obtained for the diamond structure in figure 2.8. Similar to our results for the $Si_{10}H_m$ clusters, we get different AM1, B3LYP and MP2 energy rankings for the structures. Unfortunately, the figure 2.8 diamond like structure is not amongst the structures obtained by the GA in figure 2.9. The GA structure with the lowest B3LYP and MP2 energies is s-1 and this structure is 5.4 (B3LYP) and 2.6 (MP2) kcal/mol above the diamond like structure in figure 2.9. Remarkably structure s-1 is very similar to the diamond structure apart from a 5, instead of a 6, member Si ring being formed and the addition of SiH_3 to the 5 membered ring. Structures j-2 and h-3 exhibit much larger structural differences from the diamond structure but they both differ by less than 0.7 kcal/mol from the s-1 MP2 and B3LYP energies. Thus our calculations show that the GA is not able to guarantee finding the global minimum structure evaluated with a B3LYP or MP2 energy. Nonetheless, it is reassuring that none of the structures in figure 2.9 have B3LYP or MP2 energies lower than the corresponding energy found for the diamond structure shown in figure 2.8. The GA may in fact be finding the global minimum relative to the AM1 energy, and the reason we are not finding the non-empirical global minimum structure is that we may not be including enough low energy individuals in the GA before satisfying the convergence criterion. Finally, since we are not finding any B3LYP or MP2 energies below the diamond like structure the present results are consistent with the diamond like structure shown in figure 2.8 as most likely being the global minimum at the B3LYP and MP2 levels. These relative B3LYP and MP2 energies support the idea that the most stable structure of small H passivated Si clusters is the one where the Si framework adopts a diamond lattice like structure.

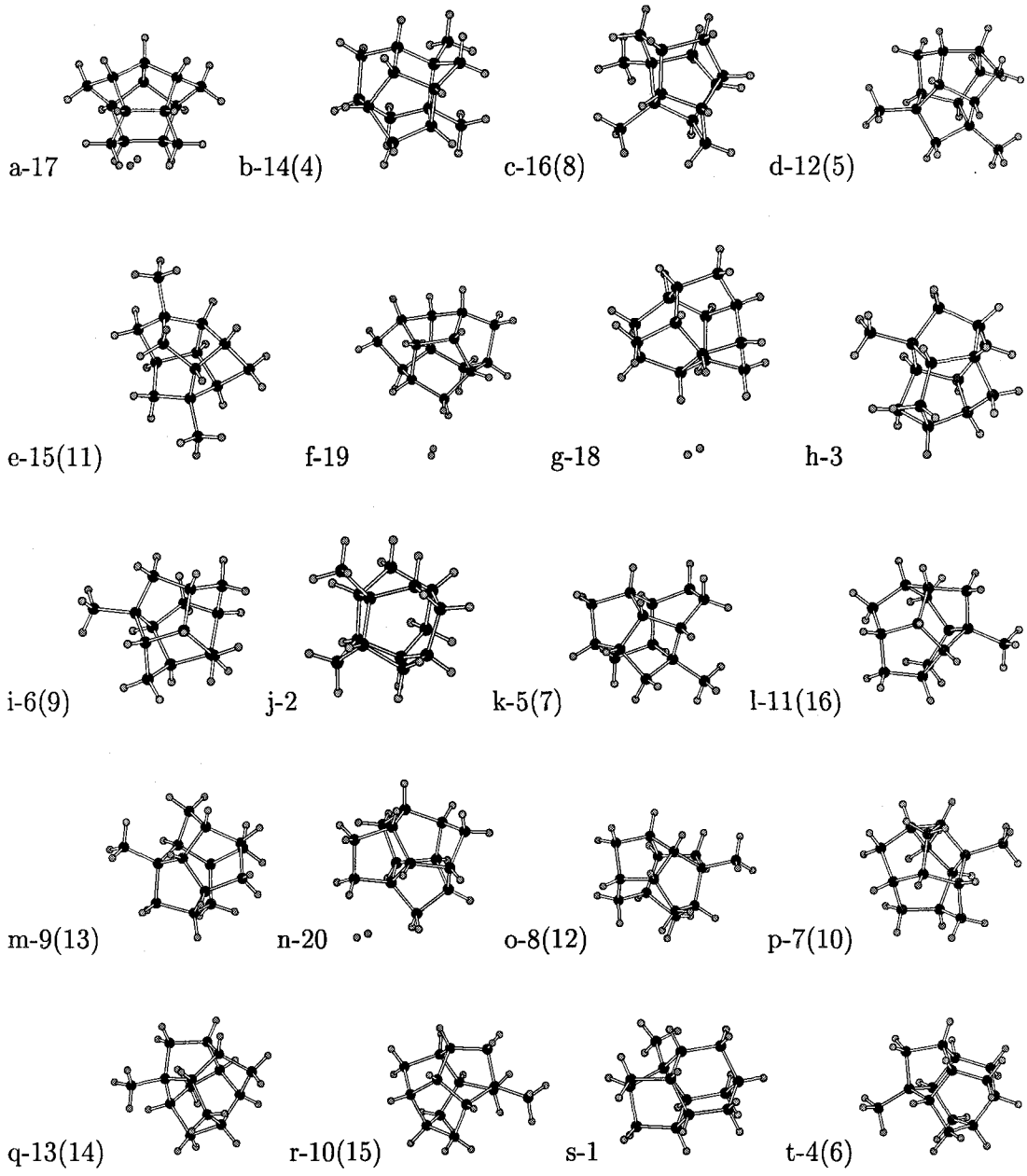


Figure 2.9: Twenty lowest AM1 energy structures for $Si_{14}H_{20}$.

2.4 Conclusion

We have developed a genetic algorithm (GA) using optimized AM1 semiempirical energies to find the lowest energy structures of Si_xH_y binary clusters. Coupling the GA with the AM1 energy makes the procedure flexible enough to treat different levels of H passivation for the Si framework in the cluster. The AM1 calculations are inexpensive to perform and, despite being parametrized largely using compounds consisting of four coordinate Si atoms, provide a reasonable starting guess for the partially and fully H-passivated Si clusters. However, quite different energy rankings are obtained with the presumably more physically realistic ab initio calculations. We have found that either B3LYP or MP2 calculations, starting from the optimized AM1 coordinates, can be used to produce very similar structure rankings.

We have applied the GA to the series of clusters $Si_{10}H_m$ with $m = 4, 8, 12, 16$ and 20. With increasing numbers of H atoms the low energy structures evolve from compact Si_{10} cages with some of the Si atoms not fully coordinated to more open structures with all of the Si atoms tetrahedrally coordinated. To accommodate more H atoms when all the Si atoms are four coordinate requires Si ring openings and substantial changes in the Si framework to take place. At the B3LYP and MP2 levels we find the global minimum structure for $Si_{10}H_{16}$ to have a Si framework with a diamond lattice like structure. We also find evidence that Si_{10} is essentially completely passivated with 16 H atoms, although the GA also indicates that different tetrasilyl cyclopentasilane isomers can be formed by $Si_{10}H_{20}$ at low temperature.

As a further test for how big a Si_xH_y cluster is needed to produce a bulk like Si lattice we also applied a refined GA, with reduced AM1 SCF and optimization tolerances for the initial generation, to the $Si_{14}H_{20}$ cluster. In separate AM1, B3LYP and MP2 calculations we determined the optimized geometry and energy for the $Si_{14}H_{20}$ cluster

arranged to have a diamond lattice like structure. The GA found 20 AM1 structures all with energies lower than and different from the possible $Si_{14}H_{20}$ diamond lattice like structure. However, the energies of the B3LYP and MP2 optimized geometries obtained by starting from the AM1 structures were all above the respective B3LYP or MP2 energies computed for the diamond lattice structure. Thus the present calculations are consistent with a relatively small Si_xH_y cluster forming a diamond lattice like structure as the global minimum at the B3LYP and MP2 energy levels.

CHAPTER 3. GLOBAL OPTIMIZATION OF H-PASSIVATED SI CLUSTERS AT THE AB INITIO LEVEL VIA AN ITERATIVELY PARAMETRIZED SEMIEMPIRICAL METHOD

(This chapter is published in *Inter. J. Quant. Chem.* **2003**, *95*, 617.)

3.1 Introduction

Nanometer sized Si particles have attracted much interest due to their potential optoelectronic properties which could be useful in various technological applications. We anticipate such a device would contain silicon clusters whose surfaces are passivated by some chemically inert ligands. In this chapter we theoretically examine using H atoms as passivation ligands and investigate the low energy structures and stabilities of Si_xH_y clusters. H atom serves as a prototype ligand since Si_xH_y clusters will generally be unstable in the air. There have been a number of other studies on Si_xH_y clusters. Chambreau *et al.* have performed flash pyrolysis of silane and disilane to produce Si_xH_y clusters in studies of chemical vapor deposition (CVD) [80]. They also used MP2 calculations to predict the geometries and relative stabilities of three Si_6H and five Si_6H_2 clusters by starting from geometries built by adding one or two H atoms onto the stable Si_6 cluster with D_{4h} symmetry. Unfortunately, they were not able to ascertain whether their fabricated Si_6H and Si_6H_2 clusters actually are the most stable structures. Meleshko *et al.* have also attempted to identify global minima for various silicon hydride clusters using simulated annealing and the MINDO/3 method to

evaluate the cluster energy [66]. Due to the inaccuracy of the MINDO/3 semiempirical method when compared against ab initio methods, it is questionable if they obtained the true Si_xH_y global minima. Miyazaki *et al.* have used density functional theory based methods to study the stable structures of Si_6H_{2n} (n=1-7) [64]. Their search for the local minima was started by using a combination of simulated annealing with pseudopotentials for Si and H atoms, along with optimizing structures initially built using chemical intuition, and by considering structures previously proposed by others.

Recently we have developed a global optimization strategy based on a genetic algorithm (GA) for theoretically investigating the structures of partially and completely H passivated Si clusters [153]. Since our global minimum search entails many cluster energy evaluations, we used the computationally fast AM1 semiempirical method [120] to identify the most stable Si_xH_y cluster, along with a number of low energy isomers. The AM1 energy calculation has the important feature of a solid quantum mechanical basis so that no additional empirical information is needed to identify which atom pairs in the binary Si_xH_y cluster combine to form bonds. Nonetheless we expect ab initio calculations at either the second order Møller-Plesset (MP2) [132] or the density-functional theory with B3LYP functional (B3LYP) levels [133–135] to better predict structures closer to experimental reality. In our previous work we hoped by evaluating the ab initio energy for the various low energy AM1 structures found by the GA we could identify the most stable Si_xH_y cluster at the ab initio level. The original AM1 parameters for Si were determined from a limited number of small molecules containing one or a few Si atoms [146], where the Si atoms were mostly four coordinate bonded. Thus it is not surprising that the AM1 parameters may fail to correctly rank the relative energies of Si clusters which have surface atoms not fully passivated or which have great ring strain. However, we also found the low energy AM1 structures of the fully H-passivated

Si clusters obtained by the global optimization were too close in relative energies when compared against the ab initio energies making it difficult to guarantee that the set of low energy AM1 structures always includes the ab initio global minimum. The AM1 method performs even worse when it is applied to evaluate the relative stabilities of barely H-passivated Si clusters. We find that the Si_6H_2 AM1 global minimum adopts a linear structure even though the ab initio calculations predict the linear structure to be highly unstable relative to the more compact structures such as D_{4h} Si_6 cluster with two H atoms each on one apex. Most of other low AM1 energy clusters adopt partially linear structures with many two coordinate Si atoms, which are also determined as unstable structures with high ab initio energies. Thus our strategy is not able to locate the ab initio low energy clusters if we only select from the top 10 or 20 AM1 low energy clusters to perform ab initio calculations. In this chapter we present an improved iterative global optimization strategy still based on using semiempirical calculations, but which should better converge to the correct ab initio energetic ranking. The method involves using two separate genetic algorithms which are invoked consecutively. The first GA, the cluster GA (CGA), corresponds to the previous GA we developed to find the Si_xH_y cluster global minimum [153]. A second and separate parametrization GA (PGA) is used to reparametrize the AM1 method against ab initio data, which is generated as we attempt to find the ab initio global minimum. The GA-reparametrized AM1 (GAM1) semiempirical method is intended to better reproduce the ab initio energetics of the clusters with a specific Si_xH_y stoichiometry contained in the training set.

GAs have been used previously to find parameters for semiempirical methods. Rossi and Truhlar used a GA to obtain specific reaction parameters (SRP) for use in dynamics calculations on the $Cl + CH_4 \rightarrow HCl + CH_3$ reaction [112]. The NDDO-SRP parameters were obtained by fitting to reference ab initio data calculated at relevant

points along the specific reaction pathway. Cundari *et al.* have optimized the PM3 [122] parameters of Tc using a GA to match the PM3 optimized geometries of the Tc compounds with their crystal structures by comparing their atom distances [113]. Their GA-revised PM3 parameters significantly improved the performance compared with the original PM3 parameters which was obtained by the interpolation of the parameters for Mo and Ru. Our approach resembles the strategy proposed by Hartke of iteratively modifying an empirical potential for Si clusters against the true cluster energy calculated by some ab initio method [49]. As a preliminary test for this strategy, Hartke used the Stillinger-Weber potential [48] with modifications by Gong [37] for the empirical potential and another empirical potential by Bolding and Anderson [45] to simulate the ab initio energy [49]. More recently, Hartke has applied this approach to find the global minima of small Si clusters using density functional theory as the true cluster energy [50]. In the present work we use the MP2 ab initio method to define the true cluster energy. Rather than have a fixed training set for the specific Si_xH_y cluster type, we adopt the idea by Hartke of expanding the training set used in the GAM1 parametrization iteratively. We first use the CGA to generate around fifty to sixty structures to be included into the training set used by the PGA. Once a set of GAM1 parameters is obtained by the PGA we repeat the CGA to find the set of lowest energy clusters derived from the new semiempirical parameters. Any new cluster structures generated are iteratively included into the training set and the parametrization is repeated with the PGA to obtain the next GAM1 parameters. The cluster global optimization at the ab initio level is repeated for at least five GAM1 parameter optimization cycles and is continued until a consistent set of low energy clusters is obtained.

In the next section we describe in more detail how the ab initio global optimization strategy is implemented along with a description of the two coupled GAs. In third

section we illustrate examples of the global optimization strategy for the Si_6H_2 cluster with low H passivation and Si_6H_6 cluster with medium H passivation. We include a comparison to the previously found low energy SI_6H_2 and Si_6H_6 structures. Concluding remarks are given in the final section.

3.2 Computational Method

The determination of the global optimization of Si_xH_y at the ab initio level consists of five main steps:

1. Initial training set design. We perform an initial crude cluster global optimization of the Si_xH_y cluster and terminate it prematurely once fifty unique AM1 locally optimized geometries are generated. We visually examine any two AM1 optimized geometries which have energy differences of less than 0.1 kcal/mol for distinct structures. This step is intended to generate a number of different Si_xH_y clusters with different structural features spread over an approximately 100 kcal/mol energy range. In contrast later cluster global optimization steps are intended to generate structures with low energies. We do not adopt the idea by Hartke of simply generating random training compounds since we find Si hydrides have a more complicated potential energy surface than bare Si clusters. Many of our randomly generated structures have high MP2 energies. We find both low and high energy structures are needed to ensure that the PGA finds new GAM1 parameters which reasonably match with the MP2 energies for the more stable clusters. We then compute the single point MP2 energies for the initial 50 training compounds. We also use the MP2 method to locally optimize the 5 geometries with the lowest single point MP2 energies. To speed up these MP2 optimizations we use reduced convergence conditions with the maximum gradient $< 10^{-3}$ a. u. and root mean square (RMS) gradient $< 0.33 \times 10^{-3}$ a. u.. We find the reduced

convergence criterion usually gives MP2 energies within 0.1 kcal/mol of those obtained by the default GAMESS convergence condition (maximum gradient $< 10^{-4}$ a. u. and RMS gradient $< 0.33 \times 10^{-4}$ a. u.). The 50 AM1 optimized geometries and the 5 derived MP2 optimized geometries are all included into the training set for the first cycle of the GAM1 parametrization. We do selectively exclude some unphysical clusters which have excessively high MP2 energies. Perhaps surprisingly, this criterion caused us not to include the linear Si_6H_2 AM1 global minimum structure into the training set.

2. GAM1 parameter optimization using the parametrization GA (PGA). The semiempirical calculations are all performed using the original AM1 equations. We only redetermine the 16 parameters of the Si atom: the atomic core integrals U_{ss} and U_{pp}, the Slater type orbital exponents ζ_s and ζ_p, the resonance β_s and β_p, the core-core repulsion term α, and the 9 core repulsion function parameters (K_{1-3}, L_{1-3} and M_{1-3}) [146].

Our experience suggests that a PGA with the population size 100 balances the efficiency and effectiveness of the optimization well. The initial 100 sets of different GAM1 parameters are randomly assigned values with a range of $\pm 30\%$ from the original AM1 parameters. For each distinct parameter set a GAM1 single point energy is evaluated for every geometry in the training set. We define the fitness function f to be the RMS deviation between the GAM1 and MP2 single point energies times a penalty coefficient P:

$$f = P[\sum_{i=1}^{N_c} (E_i^{\mathrm{GAM1}} - E_i^{\mathrm{MP2}} + D)^2 / N_c]^{\frac{1}{2}} \tag{3.1}$$

where D is introduced as a displacement to minimize the RMS deviation between the

GAM1 and MP2 relative energies:

$$D = \sum_{i=1}^{N_c} (E_i^{\mathrm{MP2}} - E_i^{\mathrm{GAM1}})/N_c \tag{3.2}$$

Since it is possible that some of the trial GAM1 parameters do not enable the SCF step to converge for all the training compounds, we evaluate the fitness function using the N_c training compounds reaching SCF convergence. We also introduce P as the penalty coefficient which takes into account when there are fewer SCF calculations reaching convergence:

$$P = e^{2.0(N_T - N_c)/N_T} \tag{3.3}$$

where N_T is the total number of the training compounds. The smaller the fitness value, the better our GAM1 energies match with the ab initio energies.

After the initial 100 fitness values are obtained, the top 10% GAM1 parameter sets with the smallest fitness values are copied directly into the offspring generation. The top 30% (including the top 10% mentioned above) are put into the breeding pool with equal weight, while the least fit 70% are eliminated from the generation. Pairs of parents are randomly chosen from the breeding pool to perform the uniform crossover to generate the remaining 90% offsprings. Each gene of an offspring has a 50%-50% chance of inheriting the gene of one of the two parents at the same position. The population size is kept constant during the evolution. A 6.25% (1 out of 16 GAM1 parameters) rate of mutation is performed on every offspring. The mutation works by randomly picking up one parameter and changing its value randomly within ±20%. The best

fitness value in one generation drops rapidly during the first few generations and starts to converge after 10 to 20 generations. We perform the PGA for 50 generations to guarantee well converged final GAM1 parameters with a small fitness value. We then replace the previous AM1 or GAM1 parameters with the new GAM1 parameters and repeat the global optimization of the Si_xH_y cluster.

3. Cluster global optimization using the cluster GA (CGA). This strategy is only briefly described here since it is essentially same as our previous genetic algorithm used for globally optimizing the Si_xH_y clusters [153]. The global optimization starts from a few initial random structures of the Si clusters as our ancestor population. The fitness f_i of the i-th initial individual cluster is based on a Boltzmann distribution:

$$f_i = \exp\left(\frac{C - \mathrm{BE}_i}{NRT}\right) \tag{3.4}$$

where BE_i is the optimized GAM1 cluster binding energy, and C is a constant, whose value is set to be the lowest energy obtained from the initial N ancestors, where the N is the total number of atoms in that molecule. C is used to rescale the fitness f_i into a rather small range. RT is set to 0.2 eV/atom. Roulette selection is used to select N random parent couples for the next generation. The probability P_i of the i-th geometry being selected is determined by the normalized fitness of that cluster:

$$P_i = f_i / \sum_i f_i \tag{3.5}$$

One mating method is to take the arithmetic mean of the cartesian coordinates from two parent geometries. The second method is to take a fragment of Si atoms from one parent and replace it with the fragment of the other parent with the same number

of atoms, and also cut a fragment of H atoms from one parent and replace it with a fragment of H atoms in the other parent with the same number of atoms. The last method is to cut each parent into halves and then recombine one half from each parent to generate the offspring . The first and second mating processes produce drastic mutations, while the third method retains much of the local structures of both parents . The probabilities used for selecting the first, second and third methods are 25%, 25% and 50%, respectively. After the mating process, N offspring are generated and fully optimized again by the GAM1 method. We replace the least stable structure of its parent population with a new offspring if this offspring has relatively lower energy and it differs with all other parents to guarantee the variety of the population. The above procedure is repeated until the lowest energy structure does not change for $3N$ generations.

4. Augmentation of the training set. After the CGA, we select ten GAM1 local minimum with the lowest GAM1 energies and evaluate their single point MP2 energies. One GAM1 structure with the lowest MP2 single point energy is optimized using the MP2 method. The above ten new GAM1 structures and the derived MP2 optimized structure are all added to the training set and used in the next cycle of the GAM1 parametrization starting at step 2. At least five GAM1 parameterization cycles are performed using steps 2, 3 and 4 and the whole global optimization procedure is repeated until either the average single point MP2 energies of the ten lowest GAM1 energy clusters are converged to 10 kcal/mol or when at least half of the different GAM1 parameter sets predict the same global minimum.

5. Determination of relative stabilities of the lowest MP2 energy clusters. After the final GAM1 parameters and the resulting low GAM1 energy clusters are found, we fully optimize the ten clusters with lowest GAM1 energies at the MP2 level.

From these ten optimized structures along with consideration of the other previously optimized MP2 geometries, we select the ones with the lowest MP2 energies within a range of 20 kcal/mol. We then further compute the zero point energy (ZPE) and other thermodynamic properties at the MP2/6-31G* level using Scott and Radom vibrational scaling factors [152]. This completes the determination of the global minimum at the MP2 level. As further fine tuning we have also performed single point MP4 and CCSD(T) energy calculations on some of the low energy MP2 optimized geometries.

All of the MP2 and AM1 calculations were performed using the GAMESS program [149]. The MP2 calculations were performed using a 6-31G* basis set [156,157]. Honea *et al.* have used the same MP2/6-31G* method to calculate the Raman frequencies of Si_y ($y = 4, 6, 7$) which are consistent with the experimental frequencies [25]. Li *et al.* also use the same method to compute IR frequencies for the Si_6 clusters and their theoretical results agree well with the experimental results [24]. The GAM1 semiempirical calculations were also performed using a slightly modified version of the GAMESS program which enabled the inclusion of the GAM1 parameters for the Si atom. The MP4 and CCSD(T) calculations were performed using the Gaussian 94 package [150]. The PGA and CGA optimizations were performed with computer codes developed by us.

3.3 Results and Discussion

We illustrate our strategy for two different silicon hydride clusters with low and medium H passivation: Si_6H_2 and Si_6H_6 clusters. The original AM1 and the two sets of final GAM1 Si atom parameters are tabulated in the table 3.1. The GAM1 Si atom parameters optimized for the Si_6H_6 cluster stoichiometry are generally closer to the AM1 parameters than the GAM1 parameters obtained from the Si_6H_2 clusters. This

is not too surprising since the original AM1 Si parameters were obtained using mostly compounds where the Si atoms were four coordinate bonded [146]. The Si GAM1 parameters determined for the Si_6H_2 clusters show large increases in the core integrals U_{ss} and U_{pp} and the resonance parameters β_s and β_p and is probably a reflection of the greater degree of multiple Si–Si bonding taking place in these clusters. For both cluster types we find the core repulsion parameters L_1, L_2 and L_3 to adopt a wide range of values during the GAM1 parametrization runs, to confirm the original expectation that these parameters play a less important role in getting the correct energy in the original AM1 method [120].

Table 3.1: Comparison of the AM1 and the final Si_6H_2 and Si_6H_6 GAM1 parameters.

Parameters	AM1	GAM1(Si_6H_2)	GAM1(Si_6H_6)
U_{ss}/eV	-33.953622	-42.809285	-34.499993
U_{pp}/eV	-28.934749	-31.966910	-26.766126
ζ_s/au	1.830697	1.456994	1.450233
ζ_p/au	1.284953	1.315649	1.265556
β_s/eV	-3.784952	-5.248804	-4.359965
β_p/eV	-1.968123	-4.419519	-2.560926
α/Å $^{-1}$	2.257816	2.352404	2.183698
K_1	0.250000	0.305731	0.281696
K_2	0.061513	0.072892	0.098506
K_3	0.020789	0.026816	0.017541
L_1	9.000000	10.012125	19.046254
L_2	5.000000	5.522481	4.288868
L_3	5.000000	2.014358	7.737797
M_1	0.911453	1.351186	0.995406
M_2	1.995569	1.807782	1.510432
M_3	2.990610	2.875304	1.655458

1. **Si_6H_2.** Figure 3.1 compares the AM1 and GAM1 relative energies against the MP2 relative energies for all the Si_6H_2 clusters included in final training set. The

reparametrized GAM1 method now does a good job of approximating the energy ranking given by the MP2 relative energies while the AM1 relative energies hardly shows any linear relationship when compared against the MP2 relative energies. The AM1 Si_6H_2 global minimum has a linear structure and is 211 kcal/mol less stable than the lowest MP2 energy cluster we locate at the MP2 level.

Figure 3.2 shows the seven most stable Si_6H_2 structures, within the 20 kcal/mol range, we have obtained at the MP2 level. It is worth noting that in figure 3.2 we draw a line between two Si atoms to suggest there is a bond when the distance between the two Si atoms is less than 2.8 Å. The relative MP2 energies (including ZPE) of these Si_6H_2 structures are included in the parentheses after each label in the increasing order. Structure (a) consisting of an elongated Si_6 octahedron, with elongation occurring for the Si atom bound to the two H atoms, is what we determine to be the global minimum at the MP2 level. The C_s structure (e) was previously determined as the MINDO/3 global minimum by Meleshko *et al.* [66]. Thus our method finds four ab initio structures (a-d) which are more stable than the MINDO/3 global minimum by 10.7, 7.6, 3.4 and 2.4 kcal/mol, respectively. It is worth noting that structure (d) has the H_2 molecule separated from the Si_6 framework and correspond to Si_6H_2 losing H_2. Others have identified the D_{4h} Si_6 compressed octahedron structure in (d) the Si_6 global minimum and it has been characterized experimentally by Raman and IR spectroscopy [24, 25]. However our Si_6 in (d) structure is slightly distorted from the D_{4h} symmetry owing to the presence of the H_2 molecule about 3.7 Å away from the nearest Si atom. The fact that we find structure (d) suggests that even though the GAM1 parameter set is optimized for the Si_6H_2 stoichiometry, it can still provide a reasonable estimates of the relative stabilities of the clusters with different stoichiometries and slightly lower H passivation. Structures (a-c) and (f) are the four most stable clusters investigated

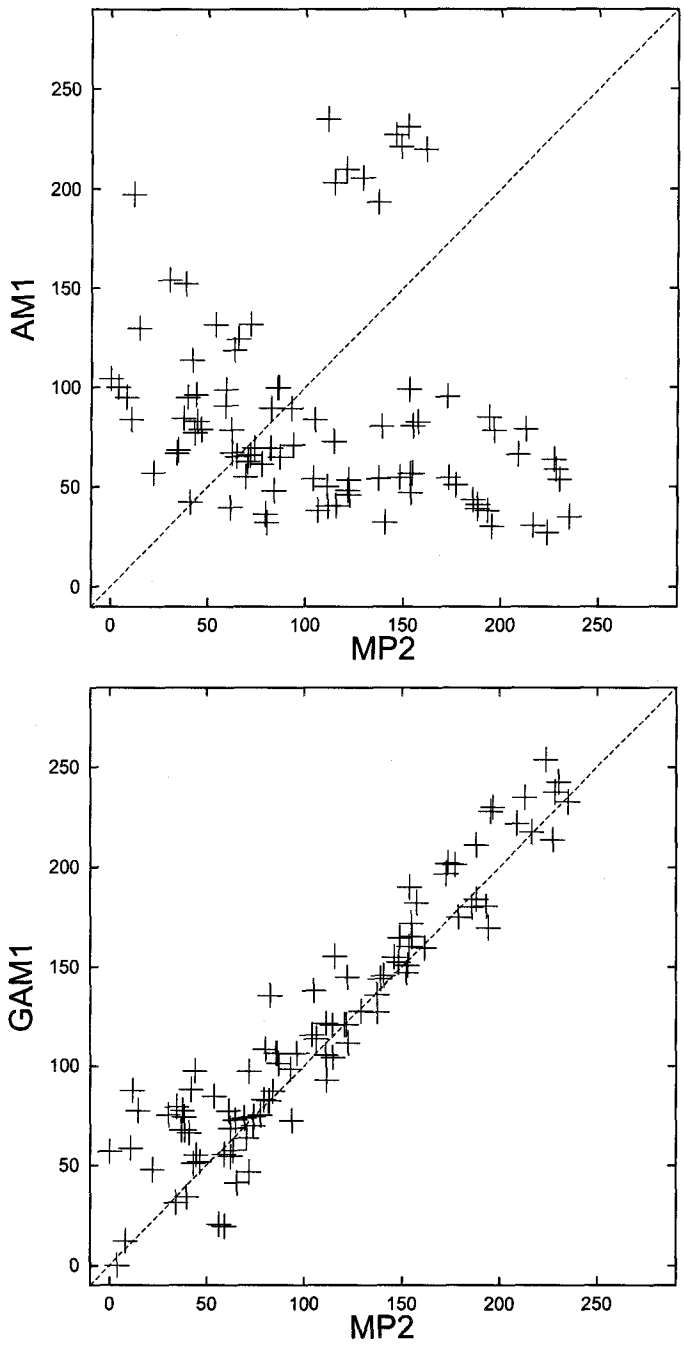


Figure 3.1: The AM1 and GAM1 versus MP2 relative energies for the Si_6H_2 clusters. All energies are in kcal/mol.

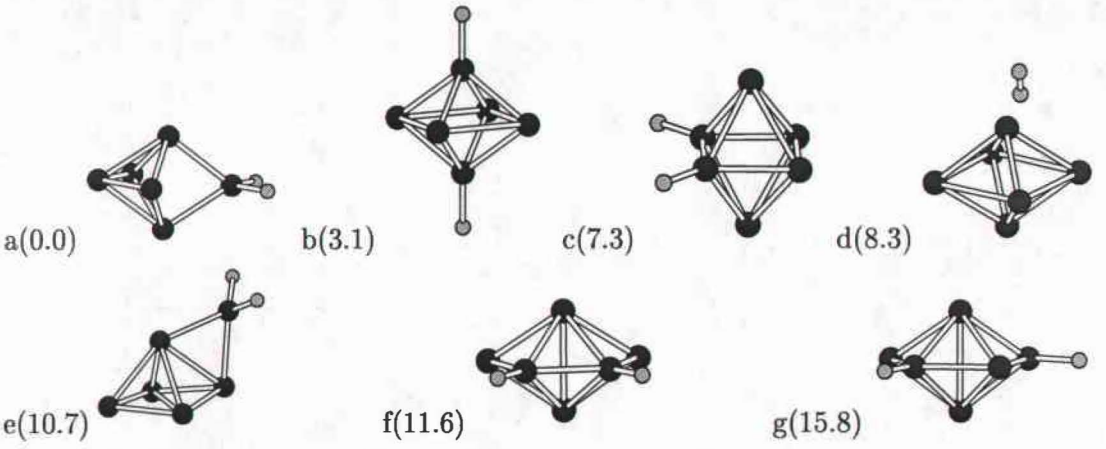


Figure 3.2: The Si_6H_2 MP2 optimized structures within 20 kcal/mol of the global minimum found by the global optimization procedure. The structures are labeled in order of increasing energy, where the numbers in the parentheses are the MP2 energies with ZPE corrections given in kcal/mol relative to structure (a). Structure (e) is the MINDO/3 global minimum previously found by Meleshko *et al.* [66].

by Chambreau *et al.* which they obtained by initially adding H atoms onto the D_{4h} Si_6 framework [80]. They used G3 theory [158,159] to evaluate the relative energies of the above clusters and ranked the structures (a-c) in the reverse energy order to our ranking. Their relative energies (including ZPE) of the structures (a-c) are 0.0, -1.2 and -2.2 kcal/mol respectively. However our MP2 calculations still agree with their G3 calculations by ranking the structures (a-c) as the most stable structures within 8 kcal/mol of each other. When we further evaluate the MP4 and CCSD(T) single point energies at the MP2 optimized geometries we obtain (a), (b) and (c) relative energies (including ZPE) of 0.0, 1.9 and 5.7 kcal/mol with MP4 and 0.0, -0.8 and 2.9 with CCSD(T) showing that there are small energy differences between all three structures. The calculations by Chambreau *et al.* suggest the strategy of fabricating low energy

Si_6H_2 clusters from the stable Si_6 framework is reasonable one, but in addition to the vertical compression and horizontal expansion of the Si_6 framework found in (f) and (g) our global optimization strategy demonstrates the ability to find another low energy Si_6H_2 structure (e) which has a very different Si_6 framework. All seven of the low energy structures shown in figure 3.2 are about 200 kcal/mol at the MP2 level more stable than the global minimum structure found using AM1.

2. Si_6H_6. The two charts in figure 3.3 show the deviation of the (G)AM1 relative energies from the MP2 relative energies for all the Si_6H_6 clusters in the final training set. The AM1 method does a better job at getting relative energies for the Si_6H_6 clusters than it did for the Si_6H_2 clusters with lower H passivation. However, the new GAM1 parameters dramatically improves the linear relationship between the semiempirical and the MP2 energies.

Figure 3.4 shows the nine lowest MP2 energy clusters (a-i) within the 20 kcal/mol range we have located by our global optimization method. We also include in figure 3.4 three low energy Si_6H_6 structures (x-z) described previously by Meleshko *et al.* [66] and the four structures (κ, x-z) discussed by Miyazaki *et al.* [64]. Overall we find the Miyazaki *et al.* (κ) structure to be the cluster with lowest energy in figure 3.4 but as we discuss below (κ) is closely related to structures (d), (g) and (h) found by us. MP4 and CCSD(T) single point energy calculations on the MP2 optimized geometries shown in figure 3.4 provide the same energy ranking as the MP2 calculations, apart for structure (g) which becomes 1.1 (MP4) and 3.4 (CCSD(T)) kcal/mol more stable than (f). As a result of the GAM1 reparametrization all our identified low energy structures (a-i) are more stable than structure (y) which was found to be the AM1 global minimum.

The nine low energy structures (a-i) exhibit three different types of Si_6 frameworks, where each of these frameworks are surrounded by various arrangements of the 6 H

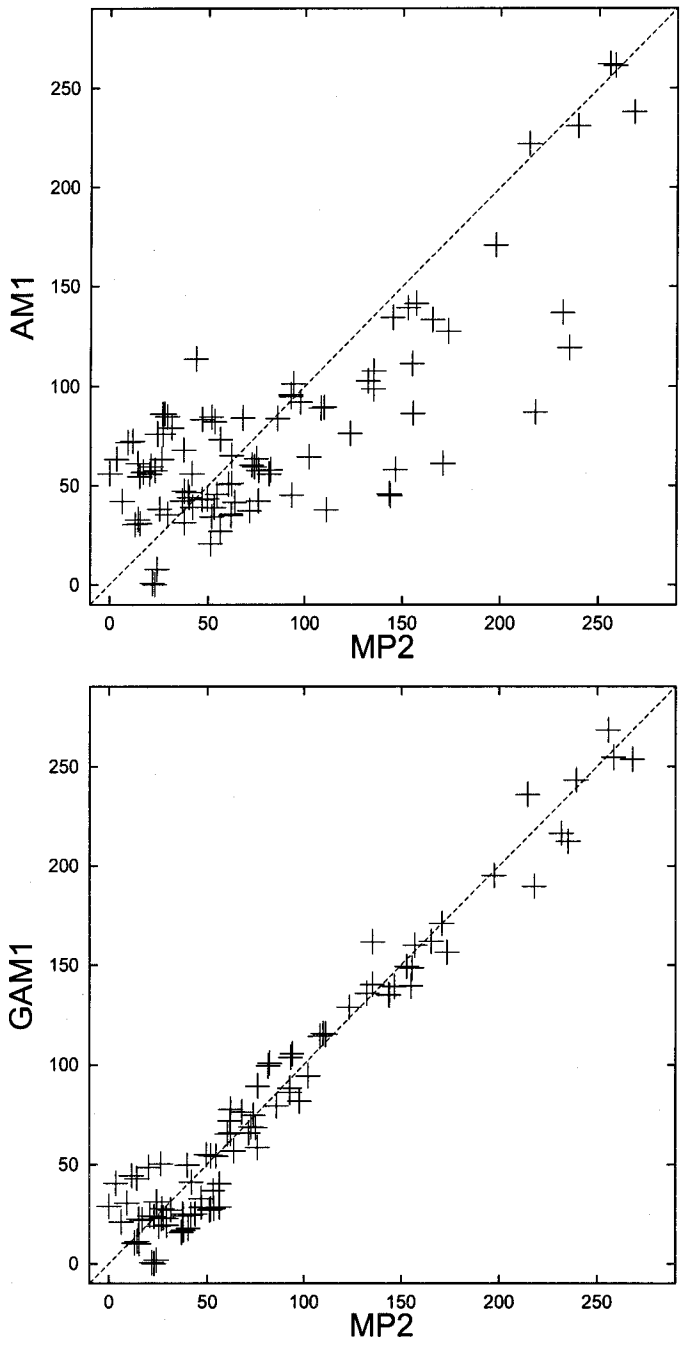


Figure 3.3: The AM1 and GAM1 versus MP2 relative energies for the Si_6H_6 clusters. All energies are in kcal/mol.

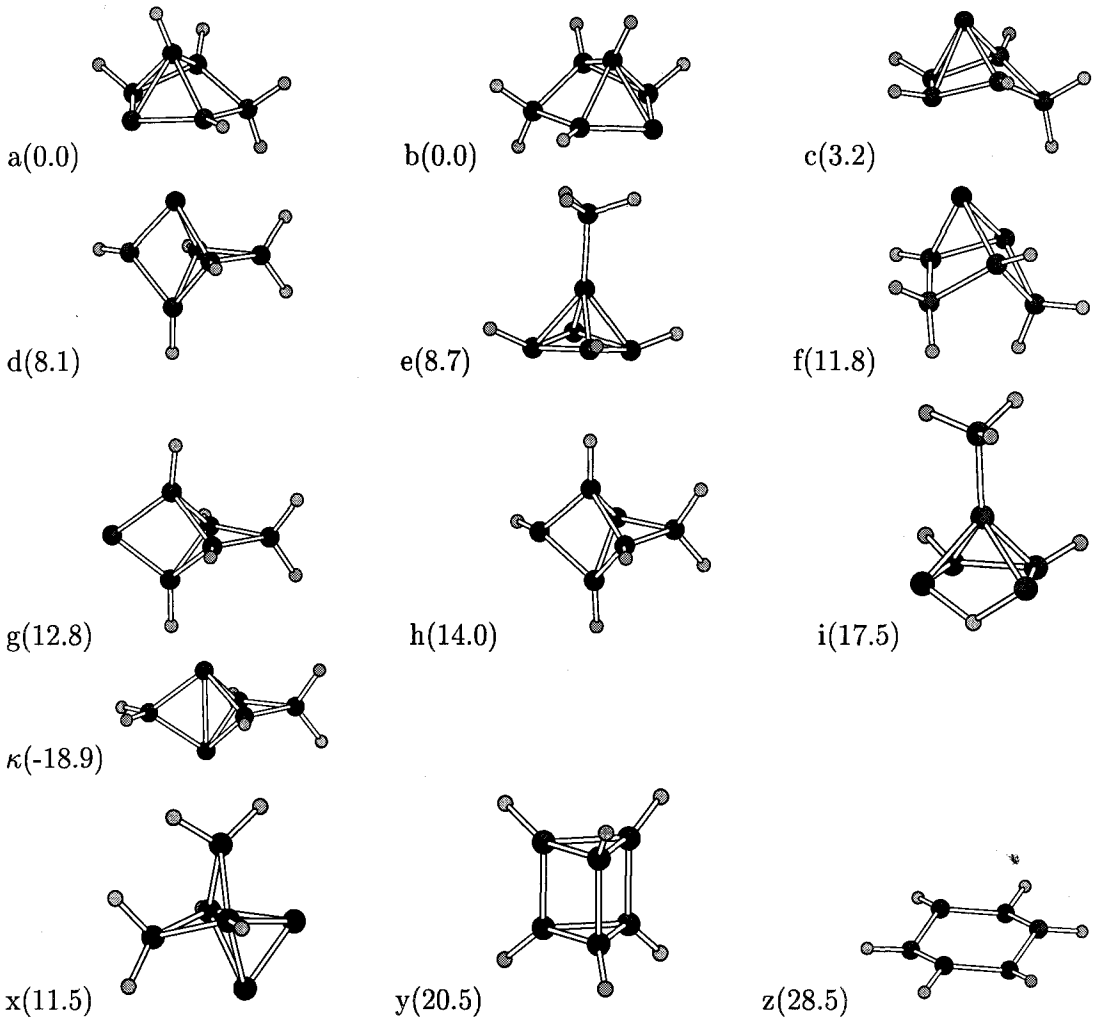


Figure 3.4: Si_6H_6 MP2 optimized structures. Structures (a-i) were obtained by our global optimization procedure and are labeled in order of increasing energy, where the numbers in the parentheses are the MP2 energies with ZPE corrections given in kcal/mol relative to structure (a). Structure (κ) is the lowest energy structure found by Miyazaki *et al.* related to structures (d), (g) and (h) by different H atom arrangements. Structure (x-z) are additional low energy structures previously found by Meleshko *et al.* and Miyazaki *et al.* [64, 66].

atoms. The Si framework in structures (a), (b), (c) and (f) consists of a Si_5 ring with one Si atom above. Structures (d), (g) and (h) contain a bipyramidal Si_5 cluster with the sixth Si atom bridging two of the equatorial Si atoms. The (κ) structure also has this Si_6 framework. Finally, structures (e) and (i) are built on a Si_5 tetragonal pyramid with a silyl, SiH_3, group attached to the top the pyramid. Clearly these Si frameworks in Si_6H_6 are very different to those found for low energy Si_6H_2 clusters and this must obviously be a consequence of the higher opportunity for Si to achieve tetravalency provided by the 4 extra H atoms. It is interesting that none of the (a-i) structures found by our cluster optimization strategy contain any of the Si_6 frameworks which occur in structures (x), (y) and (z). We actually do find the (y) Si_6 framework type but because its MP2 relative energy exceed our 20 kcal/mol energy retention threshold we do not include it with the other low energy structures found during the global optimization. The lower ab initio energy of structure (x) suggests it might be expected to be amongst our final set of low energy structures. Unfortunately (x) is eliminated from ab initio consideration since its optimized GAM1 energy is 36.9 kcal/mol higher than for the (a) and (b) structures, where even this GAM1 energy represents a big improvement from the 64.7 kcal/mol calculated for the optimized AM1 structure.

The Si_6 framework plays the primary role in determining the relative energies of the different the Si_6H_6 clusters, with the energy of a specific cluster appearing to be finer tuned by the H atom arrangement around the Si_6 framework. For instance, our global minimization strategy locates the degenerate structures (a) and (b) which are enantiomers of each other. Structure (c) is 3.2 kcal/mol higher in energy as the H atom on the top Si atom shifts to a Si atom in the Si_5 ring. There is a slightly larger energy change as two SiH_2 groups are formed in (f) while decreasing the number of Si-Si bonds between the top Si atom and the Si_5 ring atoms. In (c) this Si-Si bond length was

2.5 Å and in (f) the top Si atom is 3.4 Å from the Si atom in the SiH_2. Likewise, for structures (e) and (i), we find a 8.8 kcal/mol MP2 energy increase as one of the basal H atom shifts from a Si-H bond to bridging a pair of Si atoms. Structure (i) is a true local minimum representing an intermediate structure connecting (e) to another (e)-like structure but with the H atom shifted to the neighboring unpassivated basal Si atom suggesting that there are relatively low barriers to H atom rearrangement on the Si_6H_6 clusters.

We could make the same kind of discussion for the different H atom arrangements on clusters (d), (g), (h) and (κ) except that (κ) is significantly lower in energy than our global minima (a) and (b) as well as (d). One explanation for the lower stability of (κ) becomes apparent when one notes that (g) and (h) have two H atoms each bound to both apical Si atoms in the Si_5 bipyramid, structure (d) has lower energy since there is only one H atom bound one of the available apical Si atoms. In (κ) there are no H atoms binding with either apical Si atoms and the distance between the two apical Si atoms is reduced to 2.71 Å suggesting they are starting to form a Si-Si bond. Miyazaki *et al.* [64] initially choose to investigate the (κ) structure because the Si_6 framework corresponds to the "tetrahedral-bond-network" (TBN) structure designation by Patterson and Messmer [32].

Our global optimization strategy is having difficulty finding the (κ) structure for two main reasons. The AM1 optimized (κ) structure is 12.2, 14.4 and 20.2 kcal/mol less stable than the (d), (g) and (h) structures and although we improve the energy ranking with our final GAM1 parameters (κ) is still 7.8, 19.7 and 7.5 kcal/mol higher in energy than the (d), (g) and (h) structures, respectively. As a consequence (κ) is never selected to join the parametrization training set so that it can more easily be identified as one of the lower energy structures. Secondly, although the three mat-

ing operations we use to make new generations of structures in our CGA appear to be good at finding the different low energy Si_x frameworks, these mating operations do this while retaining many of the Si-H bonds. As a consequence the present CGA explores many different Si_x frameworks but only a subset of the possible H atom arrangements on these frameworks. Our work demonstrates the well known feature that genetic algorithms are good tools for searching for the lower energy local minimum on a potential energy hypersurface, but the rate of convergence of the genetic algorithm towards the global minimum is not necessarily enhanced by information from any of the low energy structures closely related to the global minimum [100]. However, a good feature of the present iterative global optimization strategy is that it does enable us to identify the energetically important Si_x frameworks in the Si_xH_y clusters. We are currently working on developing an improved strategy for examining different H atom arrangements on these Si frameworks. Even with the present procedure, after visually inspecting the different H atom arrangements on a specific Si_x framework, it would not be too a laborious task to manually check a few closely related structures for the possibility that they have lower energies.

3.4 Conclusion

We have attempted to develop a strategy for the global optimization of binary Si_xH_y clusters at the ab initio level using two coupled genetic algorithms. One GA (CGA) searches amongst the many structure candidates for potential low energy clusters via fast semiempirical calculations so that the energy of only a few structures, including hopefully the global minimum, need to be evaluated using ab initio MP2/6-31G* calculations. Since semiempirical calculations can often give a different energy ranking to the ab initio ranking we use a second GA (PGA) to reparametrize the semiempirical

method as different new ab initio structures with low energy appear from the CGA.

To illustrate our global optimization strategy we have attempted to find the ab initio global minima for the Si_6H_2 and Si_6H_6 clusters. Reparametrizing AM1 for specific Si_xH_y stoichiometries certainly improves the ability of the semiempirical method to rank the energy of different structure types. For Si_6H_2 we found seven low MP2 energy structures within 20 kcal/mol of each other. The global minimum has the two H atoms attached to a single Si atom on a distorted Si_6 octahedron. Our Si_6H_2 global minimum is over 10 kcal/mol more stable than the MP2 energy of the global structure previously found by Meleshko *et al.* [66] using MINDO/3 calculations. Our results for Si_6H_6 are more complicated and we find nine lowest MP2 energy structures within 20 kcal/mol of each other. Unfortunately, none of structures correspond to the structure which is 18.9 kcal/mol lower in energy, and presumably the global minimum, found previously by Miyazaki *et al.* [64]. Nonetheless, our global optimization strategy does find several Si_6H_6 clusters with the same Si_6 framework as the global minimum but with different H atom arrangements. Suggestions for manual checks of possible lower energy structures can be made by visually inspecting the H atom arrangements on the Si framework.

Our experience with the Si_6H_2 and Si_6H_6 clusters suggests we have developed a useful and efficient ab initio global optimization strategy. For the binary Si_xH_y clusters the method is especially useful for finding the important low energy Si_x frameworks and are now trying to develop a better procedure for investigating the best arrangement the H atoms can adopt around the Si framework.

CHAPTER 4. INVESTIGATION OF THE TRANSFERABILITY OF THE GAM1 PARAMETERS

(This chapter is published in *J. Phys. Chem. B* **2004**, *108*, 6025.)

4.1 Introduction

Nanometer sized Si particles have potentially interesting optoelectronic properties which could be useful in various technological applications. Since Si only clusters are highly unstable due to dangling bonds on the surface Si atoms a practical device will most likely be fabricated from silicon clusters whose surfaces are passivated by some chemically inert ligand. For these reasons we have been investigating the structures of Si_xH_y clusters where the H atoms serve as a prototypical passivating ligand. Several other groups have also been theoretically and/or experimentally exploring the possible structures for small to medium sized silicon hydride clusters [55, 56, 63–69, 78, 80]. Not only is finding the Si_xH_y global minimum structure important, but we also determine how many H atoms are actually needed to completely passivate the Si_x core. We applied a genetic algorithm (GA) [100–103] to find the cluster with the global minimum energy at the ab initio level for a particular Si_xH_y stoichiometry [153]. We used the AM1 semiempirical method [120] as a fast computational procedure to screen the energies for the many different Si_xH_y clusters generated during the GA and then searched among the low AM1 energy structures for the one with lowest ab initio energy. Our search strategy starts from a set of random Si_xH_y structures and avoids the explicit introduction of structures which might be expected from chemical intuition to have low energy. When the Si_x core is over passivated the method automatically produces Si_xH_{y-2} and H_2

structures. Our calculations confirmed the chemically expected result that the Si cores of the small to medium sized Si_xH_y clusters favor forming diamond-lattice like structures provided they are surrounded with enough passivating ligands [153]. We have proposed that the Si_{10} and Si_{14} clusters need 16 and 20 H atoms to be fully passivated so as to eliminate any dangling bonds on the Si surface and avoid forming structures containing small and strained three or four member Si rings. Unfortunately the energy ranking determined by the original AM1 method differed appreciably from the ab initio energy ranking. Even for well passivated Si_xH_y clusters, the ab initio global minimum is often not the AM1 global minimum. Where we designate the ab initio global minimum to mean the lowest energy isomer which has been found so far by any method at the ab initio level. For instance, previously we found nine $Si_{10}H_{16}$ structures with AM1 energies lower than the cluster with a diamond lattice like Si_{10} core corresponding to the ab initio global minimum [153]. To guarantee finding the true ab initio global minimum for the $Si_{10}H_y$ clusters, we had to examine up to 20 of the lowest AM1 energy clusters at the ab initio level. However, this strategy still failed when we attempted to find the global minimum of $Si_{14}H_{20}$ at the ab initio level. Since the number of local minima increases exponentially [95], while the total energy scales only linearly, with the size of the cluster, we estimate that the larger Si_xH_y clusters will have an exponentially increased number of local minima residing in a very narrow energy range. A few kcal/mol relative energy deviation between the AM1 method and the ab initio method could easily lead to totally different energy rankings and explains why the true ab initio global minimum of the larger $Si_{14}H_{20}$ clusters was not found among the 20 clusters with lowest AM1 energies [153]. We identified the true ab initio $Si_{14}H_{20}$ global minimum by considering Si core fragments with a diamond lattice like structures.

To improve our reliability at finding the ab initio global minimum for a Si_xH_y

cluster we developed a second global optimization strategy which utilizes two coupled genetic algorithms [154]. One of the GAs (the PGA) is used to reoptimize the AM1 parameters for Si [146] based on an iteratively augmented training set of Si_xH_y clusters treated at the ab initio level, and the other GA (the CGA) is essentially our original GA [153] used to search for the cluster ab initio global minimum but now using the reoptimized AM1 parameters for Si. The usual philosophy with a semiempirical method is to find a set of parameters which are applicable to a wide range of chemical environments. The goal of the GA optimized AM1 (GAM1) parameters is to have parameters which more accurately predict the energy but for a narrower range of chemical systems. In our studies on Si_6H_2 and Si_6H_6 clusters the GAM1 parameters significantly improved the semiempirical energy ranking of the different structures for a specific stoichiometry [154]. For instance, the final Si_6H_2 GAM1 parameters enabled the CGA to find the same global minimum structure at both the ab initio and semiempirical levels. Whereas the Si_6H_6 GAM1 parameters resulted in the CGA finding three nearby low energy local minima having the same Si framework as the Si_6H_6 ab initio global minimum. Since the GAM1 parameters for Si automatically adjust to different extents of H-passivation we feel this approach improves our ability at ultimately finding the ab initio global minimum for Si_xH_y clusters with various stoichiometries. Hartke has used a similar approach in the global optimization of small Si only clusters where a Si empirical potential is parametrized against the true cluster energy calculated by some ab initio method [49,50]. Unfortunately, the GAM1 reparametrization work means that the coupled GA method is not very efficient when we try to globally optimize larger clusters such as the well-passivated $Si_{10}H_{16}$ cluster. In this chapter we explore how transferable the GAM1 parameters obtained from a Si_7H_{14} training set are to larger clusters with comparable levels of passivation. We choose Si_7H_{14} since it can form a

series of structures close in energy which are representative of Si core being completely H-passivated. Clearly, if the Si_7H_{14} GAM1 parameters produce a ranking matching the ab initio calculations on larger clusters we can then reliably use the much faster single GA approach to identify the global minimum for these larger clusters.

The remainder of the chapter is organized as follows. Details on the computational methods are presented in the next section. We describe the two genetic algorithms used for finding: (1) the optimum semiempirical parameters (PGA), and (2) the cluster global minimum (CGA). We also describe how the semiempirical and ab initio calculations are performed. The results and discussion are presented in the third section. This section starts by describing the results from the GAM1 parametrization using the Si_7H_{14} training set and the PGA. We then explore the transferability Si_7H_{14} GAM1 parameters to larger Si_xH_y clusters having different levels of H-passivation. The final part of the results and discussion section demonstrates the range of application of the GAM1 parameters in the global optimization of the $Si_{10}H_{16}$, $Si_{10}H_{14}$, $Si_{14}H_{20}$ and $Si_{14}H_{20}$ clusters. Our results show for the clusters with high levels of H-passivation that the Si_7H_{14} GAM1 parameters outperform the AM1 method by better matching with the MP2 relative energies. Both the GAM1 and MP2 calculations consistently rank the diamond-lattice like structures of $Si_{10}H_{16}$ and $Si_{14}H_{20}$ as the most stable isomers. Concluding remarks are given in the final section of this chapter.

4.2 Computation Method

1. Parametrization genetic algorithm (PGA).

The main goal of this chapter is to find the correct global minimum of various well-passivated Si_xH_y clusters at the MP2 [132] level. We start by selecting a training set consisting of relatively small Si_7H_{14} clusters as being suitable for reparametrizing

the AM1 method. Si_7H_{14} is large enough that it can adopt various configurations, ranging from linear structures through three to seven-member ring structures, which represent common partial structures found in a well-passivated Si_xH_y cluster. The 14 different MP2 optimized Si_7H_{14} structures selected to be in the training set are shown in Figure 4.1 along with their corresponding MP2 energies computed using a 6-31G* basis set [156, 157] listed in parentheses. We include low and some high energy structures in the training set to ensure that both low and high energy structural features are reasonably modeled by the GAM1 parameters. In our previous work on the Si_6H_2 and Si_6H_6 stoichiometries we found a cluster global optimization (CGA) run with the new GAM1 a useful check that we have sufficient low energy structures in the parameter training set [154].

The semiempirical calculations are all performed using the original AM1 equations. We only reoptimize the 16 Si atom parameters: the atomic core integrals U_{ss} and U_{pp}, the Slater type orbital exponents ζ_s and ζ_p, the resonance β_s and β_p, the core-core repulsion term α, and the 9 core repulsion function parameters (K_{1-3}, L_{1-3} and M_{1-3}) [146]. In our IJQC paper, we simply changed the values of the above 16 AM1 parameters in the data block of mpcdat.src in the GAMESS package, and did not correctly take into account changes needed for the one center two electron integral terms [154]. In the present work we now correctly compute the one center two electron integral terms: fortunately we find our previous GAM1 parameters [154] are not significantly affected by the small errors present in the one center terms.

The PGA is only slightly modified from our earlier work [154]. We still use a population size of 100 since we find this balances the efficiency and effectiveness of the parameter global optimization. The initial 100 sets of different GAM1 parameters are randomly assigned real number values with a smaller range of $\pm 3\%$ from the original

AM1 parameters. After the initial 100 fitness values are obtained, the top 10% GAM1 parameter sets with the smallest fitness values are copied directly into the offspring generation. The top 30% (including the top 10% mentioned above) are put into the breeding pool with equal weight, while the least fit 70% are eliminated from the generation. Pairs of parents are randomly chosen from the breeding pool to perform the uniform crossover to generate the remaining 90% offspring. Each gene of an offspring has a 50%-50% chance of inheriting the gene of one of the two parents at the same position. The mutation amplitude is also reduced to 2% since we find small modification of the original AM1 parameters is good enough for matching relative energies of well-passivated Si_xH_y clusters calculated by the semiempirical and the ab initio methods. We also want to avoid dramatic change of the AM1 parameters when using this relatively smaller Si_7H_{14} training set.

The GAM1 parameters are found by minimizing a fitness function f consisting of an energy and a gradient part:

$$f = f_E + f_G \tag{4.1}$$

For each distinct parameter set a GAM1 single point energy is evaluated for every geometry in the training set. We define the energy related fitness function f_E to be the root mean square (RMS) deviation between the GAM1 and MP2 single point energies times a penalty coefficient P:

$$f_E = P[\sum_{i=1}^{N_c}(E_i^{\mathrm{GAM1}} - E_i^{\mathrm{MP2}} + D)^2/N_c]^{\frac{1}{2}} \tag{4.2}$$

where D is the difference between the MP2 and GAM1 average energies for the training compounds and is introduced to maximize the coincidence of the GAM1 and MP2 relative energies:

$$D = \sum_{i=1}^{N_c} (E_i^{\mathrm{MP2}} - E_i^{\mathrm{GAM1}})/N_c \tag{4.3}$$

Since it is possible that some of the trial GAM1 parameters cause poor SCF convergence, we evaluate the fitness function using only the N_c training compounds which are SCF converged. The penalty coefficient P is designed to increase when one or more SCF calculation fails to converge:

$$P = e^{2.0(N_T - N_c)/N_T} \tag{4.4}$$

where $N_T(= 14)$ is the total number of compounds in the training set. The gradient related fitness value is determined from RMS of the GAM1 energy gradient, G_i^{GAM1}, evaluated at the MP2 optimized geometry for structure i in the training set:

$$f_G = P[\sum_{i=1}^{N_c} (G_i^{\mathrm{GAM1}})^4/N_c]^{\frac{1}{2}} \tag{4.5}$$

Rossi and Truhlar used a GA to obtain specific reaction parameters for use in dynamics calculations on the $Cl + CH_4 \rightarrow HCl + CH_3$ reaction [112]. They also used a fitness function which included both an energy and a gradient term. However, our experience suggests that the decrease in the deviation of energy after the GAM1 optimization is

roughly proportional to the square of GAM1 RMS gradient when starting from a MP2 optimized geometry. We find including the higher fourth power on the gradient term is more effective at reducing differences between the GAM1 and the MP2 optimized geometries. We did not include the gradient term in the fitness function in our IJQC paper [154] since we found that the geometry and energy of the low H-passivated Si_xH_y clusters could change dramatically during local optimization even when the starting geometry had a small gradient.

2. Cluster Genetic Algorithm (CGA).

We have slightly modified the previous genetic algorithm used for globally optimizing the Si_xH_y clusters [153] to facilitate more efficient calculations in a multiprocessor computing environment. The global optimization starts from N initial random structures of the Si_xH_y clusters as our ancestor population, where N is the population size which depends on the size of the molecule. In this chapter we use a population size of 60 for $Si_{10}H_y$ stoichiometries and 120 for the larger $Si_{14}H_y$ ones. The fitness value f_i of the i-th initial individual cluster is taken as the absolute value of the GAM1 energy of the locally optimized cluster. The lower the energy of an individual, the higher is its fitness. Tournament selection instead of Boltzmann selection is used to select random parent couples which mate to produce offspring with guaranteed variety for the next generation. Each parent used for the mating process is chosen by first randomly selecting eight individual clusters from the current population and the one with lowest energy is then taken as one parent. In order to run our cluster global optimization on a multiprocessor environment more efficiently, we reorganized our CGA source code so that it no longer explicitly treats separate generations. Previously we only started work on a new generation after all individuals in the previous generation have been locally optimized: the new strategy now performs parent selection, mating process and local

optimization for a new individual as soon as any cluster local optimization job is finished on a processor. This new strategy reduces the wall clock time since it avoids processors being idle while waiting for other local optimization jobs in the same generation to be completed.

Generally we want the mating methods to keep the good features of the better (lower energy) parents while some mutations are performed to help maintain the variety of the population. Three mating methods are adopted. One is to take the arithmetic mean of the Cartesian coordinates from two parent geometries. The second method is to take a fragment of Si atoms from one parent and replace it with the fragment of the other parent with the same number of atoms, and also cut a fragment of H atoms from one parent and replace it with a fragment of H atoms in the other parent with the same number of atoms. The last method is to cut each parent into halves and then recombine one half from each parent to generate the offspring. The first and second mating processes produce drastic mutations, while the third method retains much of the local structures of both parents. The probabilities used for selecting the first, second and third methods are 25%, 25% and 50%, respectively. After the mating process, an offspring is generated and fully optimized again by the GAM1 method. To avoid same structures occupying the population repeatedly, we replace the least stable structure of the current population with a new offspring only if this offspring has lower energy and not identical with any individual in the current population. We have developed an efficient and effective strategy to determine whether the offspring generated is identical with any cluster already present in the current population. We simply compare the energy and principal moments of inertia of new offspring with the values for every individual in the current population. We assume that the new offspring is identical with a previous individual when their energies differ by less than 10^{-3} a.u.

and their three principal moments of inertia all agree within 1%. Our experience shows this threshold is tight enough to keep distinct offspring clusters while not too tight to have many essentially identical individuals occupying the population.

3. Semiempirical and ab initio methods.

All of the semiempirical and ab initio calculations were performed using the GAMESS program [149]. The GAM1 parameters were introduced into the GAMESS program by modifying the appropriate Si atom AM1 parameters. Initially the CGA finds the many low energy Si_xH_y clusters using the GAM1 method, then we determine the ab initio global minimum by locally optimizing the 20 lowest GAM1 energy structures with the MP2 method [132]. The ab initio calculations use a 6-31G* basis for the clusters containing 10 or less Si atoms [156, 157]; whereas the Hay-Wadt effective core potential (ECP) and valence basis set for Si [147] augmented with a *d* function [23] and the Dunning-Hay basis for H [148] was used to reduce the computational effort for the larger $Si_{14}H_{20}$ and $Si_{14}H_{18}$ clusters. Consistent with the work of others [23,79], our tests show that the relative energies of the well-passivated Si clusters calculated by the MP2 method match well with the much more expensive methods which include higher levels of electron correlation such as MP4 and CCSD(T). We also performed some vibrational frequency calculations to estimate zero-point energy (ZPE) corrections and Gibbs free energies using HF/6-31G* calculations for $Si_{10}H_{16}$ and $Si_{10}H_{14}$ and HF/3-21G calculations for the larger $Si_{14}H_{20}$ and $Si_{14}H_{18}$ clusters. All of the frequencies were scaled using Scott and Radom correction factors [152].

4.3 Results and Discussion

1. GAM1 Parameter Global Optimization Using the Si_7H_{14} Training Set

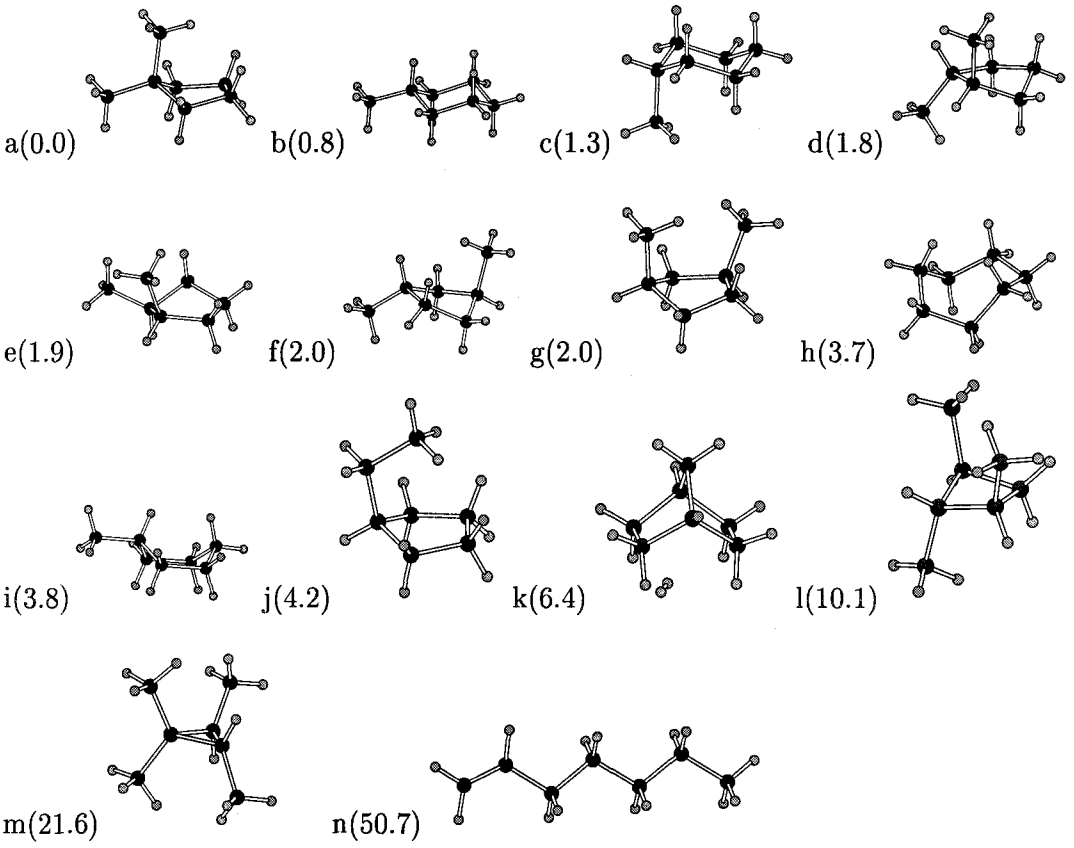


Figure 4.1: The MP2/6-31G* locally optimized geometries for the Si_7H_{14} training set. In parentheses are the MP2 relative energies in kcal/mol.

The Si_7H_{14} training set and the relative MP2 energies for the locally optimized structures are shown in figure 4.1. Both the AM1 and the new GAM1 parameters obtained from the Si_7H_{14} training set are listed in table 4.1. The original AM1 parameters for the Si atom were obtained from a wide variety of compounds containing one or more Si atoms [146] and consequently it should not be surprising that different GAM1 parameters are obtained when treating a narrower range of compound types. On the other hand, we should not anticipate major changes in the parameters since the AM1 method has generally been found to give reasonable relative energies for Si containing compounds [146]. The most significant changes occur for the ζ_s and β_s parameters which are reduced by 18% and 12%, respectively, while the rest of the GAM1 parameters from the Si_7H_{14} training set differ by less than 10% from the corresponding AM1 parameter. Previously the GAM1 parameters obtained using the Si_6H_2 and Si_6H_6 training sets showed greater differences from the original AM1 parameters presumably because of the presence of a greater fraction of low coordinate Si atoms in these training sets [154].

Table 4.2 compares the relative MP2 optimized energies against the AM1 and GAM1 relative energies for all of the Si_7H_{14} training compounds. Table 4.2 includes the single point AM1 and GAM1 relative energies computed at the MP2 optimized geometries and the AM1 and GAM1 energies obtained by reoptimizing the MP2 structures with the semiempirical calculations. Structure 1(a) containing a five membered Si ring is found to be the global minimum for all three methods and we use the 1(a) MP2 optimized geometry as the zero reference for each of the MP2, AM1 and GAM1 relative energies. Table 4.2 shows the Si_7H_{14} training set to contain six MP2 optimized structures within 2 kcal/mol of the global minimum. The GAM1 single point calculations have the smallest RMS deviation and do the best job at reproducing the small

Table 4.1: Comparison of the AM1 parameters and the improved GAM1 parameters derived from a Si_7H_{14} training set.

Parameters	AM1	GAM1
U_{ss}/eV	-33.953622	-34.269578
U_{pp}/eV	-28.934749	-30.694633
ζ_s/au	1.830697	1.500623
ζ_p/au	1.284953	1.356890
β_s/eV	-3.784952	-3.321166
β_p/eV	-1.968123	-1.862159
α/Å $^{-1}$	2.257816	2.249007
K_1	0.250000	0.249818
K_2	0.061513	0.065260
K_3	0.020789	0.020458
L_1	9.000000	10.141849
L_2	5.000000	5.309296
L_3	5.000000	4.221081
M_1	0.911453	0.865394
M_2	1.995569	1.955345
M_3	2.990610	2.812023

MP2 energy differences in table 4.2. The GAM1 single point calculations give a definite improvement over the AM1 single point relative energies. Since in a CGA we would locally optimize various cluster structures with the semiempirical method first before performing an ab initio optimization, it is also important to notice in table 4.2 that the GAM1 RMS deviation increases slightly due to the small energy relaxation which occurs when each MP2 optimized structure is reoptimized using the GAM1 parameters. Nonetheless, the RMS deviation for GAM1 optimized structures is still considerably better than the AM1 RMS deviations. The small GAM1 energy relaxation after optimization is consistent with an improvement in the geometry capabilities, as well as providing more reliable energy rankings, for the GAM1 parameters obtained from using Si_7H_{14} training compounds. The average RMS cartesian energy gradients for station-

ary point structures (excluding 1(k)) are 6.4×10^{-3} and 0.9×10^{-3} a.u. for AM1 and GAM1, respectively. The 1(k) stationary point structure has much larger RMS AM1 and GAM1 gradients of 9.4×10^{-3} and 6.6×10^{-3} a.u. due to large gradients on the isolated H_2 molecule which are independent of the Si atom parameters values. In figure 4.2 the relative AM1 and GAM1 optimized structure energies are plotted against the corresponding MP2 energy and this provides another illustration that the resulting Si_7H_{14} GAM1 parameters do a better job at simulating the expensive MP2/6-31G* calculations than the AM1 method.

Table 4.2: Comparison of the Si_7H_{14} relative energies for MP2 optimized geometries versus the AM1 and GAM1 relative energies calculated at the MP2 optimized geometry (S. P.) and at the optimized semiempirical geometry (Opt.). The RMS deviation between the semiempirical and the MP2 energies is evaluated using equation (2). All energies are in kcal/mol.

Structure	Si ring size	MP2	AM1		GAM1	
			S. P.	Opt.	S. P.	Opt.
1(a)	5	0.0	0.0	-2.8	0.0	-0.3
1(b)	6	0.8	6.5	3.5	1.9	1.8
1(c)	6	1.3	6.1	3.4	1.8	1.7
1(d)	5	1.8	3.7	0.6	1.9	1.5
1(e)	5	1.9	3.7	0.5	1.9	1.4
1(f)	5	2.0	3.4	0.5	1.7	1.4
1(g)	5	2.0	3.2	0.4	1.5	1.1
1(h)	7	3.7	10.2	7.1	3.8	3.5
1(i)	6	3.8	6.3	3.5	2.5	2.2
1(j)	5	4.2	5.8	3.4	2.6	2.4
1(k)	5, 5	6.4	2.3	-2.0	5.9	4.4
1(l)	4	10.1	10.7	6.5	9.2	8.0
1(m)	3	21.6	18.9	15.1	20.8	19.5
1(n)		50.7	49.3	42.6	50.5	47.3
RMS Dev.			2.9	3.5	0.7	1.1

Except for structure 1(n), all the Si_7H_{14} training compounds shown in figure 4.1

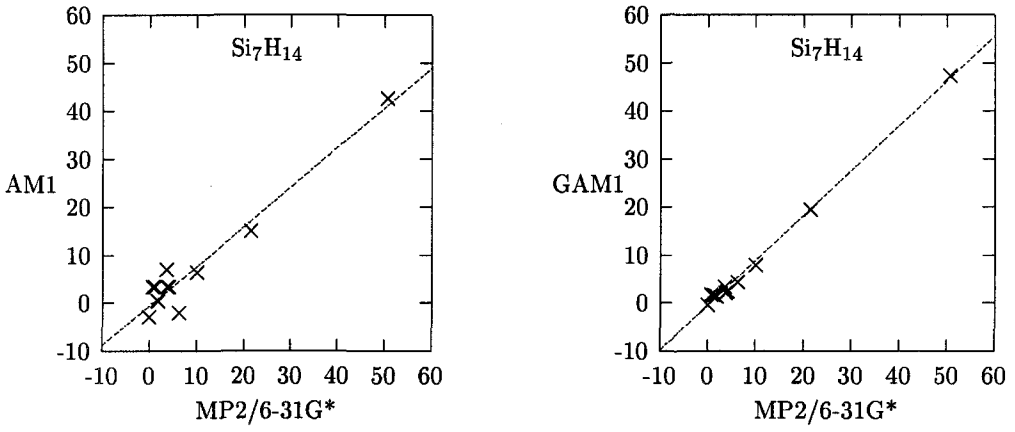


Figure 4.2: Comparison of the Si_7H_{14} AM1 (left) and GAM1 (right) relative optimized energies against MP2/6-31G* relative optimized energies. Energies are in kcal/mol.

contain a Si atom ring structure with sizes between three and seven Si atoms. Table 4.2 lists the Si ring size for each training compound and we see at the MP2 level that the seven lowest energy structures which are within 2 kcal/mol of each other contain either a five or a six membered Si ring. The AM1 calculations in table 4.2 consistently underestimate the stability of both six and seven-member ring structures relative to the five-member ring structures. This AM1 trend is found for the two low energy six membered ring structures 1(b) and 1(c) and for the slightly higher energy structures 1(h) and 1(i). The new Si_7H_{14} GAM1 parameters are able to correct for this deficiency and better reproduce the ab initio relative energies for the six and seven membered ring structures.

The bicyclic structure 1(k) is an example of Si_7H_{14} automatically losing a H_2 molecule because of over passivation. Despite containing 2 five membered Si rings, AM1 predicts a low relative energy for structure 1(k) which suggests the AM1 parameters

favors forming more compact Si_xH_y structures. This result may explain why in our previous research work [153] we obtained many low relative energy structures with a H_2 molecule detached at the AM1 level which had appreciably higher relative energies at the ab initio level. The GAM1 method increases the single point relative energy of 1(k) to 5.9 kcal/mol although the energy of the optimized GAM1 structure still relaxes to 4.4 kcal/mol. 1.4 kcal/mol of the energy change is due to the H_2 relaxing from the optimized MP2 geometry to the equilibrium AM1 geometry as a result of the relatively large H_2 energy gradients described above for the 1(k) single point structure. Consequently, we expect the present GAM1 parameters should also give a better relative energy estimate when a H_2 molecule is lost from a Si_xH_y cluster.

2. Transferability of the Si_7H_{14} GAM1 Parameters

To test the transferability of the Si_7H_{14} GAM1 parameters to larger Si_xH_y clusters we have repeated the semiempirical local optimizations on all of the $Si_{10}H_y$, $y = 4, 8, 12, 16, 20$ and the $Si_{14}H_{20}$ cluster structures we obtained from our earlier cluster global optimization effort [153]. The $Si_{10}H_y$ and $Si_{14}H_{20}$ clusters contain Si atoms in a range of environments. Low passivation is exhibited for the clusters with the $Si_{10}H_4$ and $Si_{10}H_8$ stoichiometries since they all contain at least two Si radicals. Three of the $Si_{10}H_4$ structures we have obtained even contained two biradical Si atoms. The ten low AM1 energy $Si_{10}H_{12}$ structures we found previously all have the Si atoms four-fold coordinated, but these structures still have significant ring strain due to the presence of small three or four-member Si rings indicative of intermediate passivation. Each of the Si atoms in the different $Si_{10}H_{16}$, $Si_{10}H_{20}$ and $Si_{14}H_{20}$ clusters are essentially tetrahedrally coordinated with either other Si or H atoms and the Si_x core is well or even over-passivated by H atoms. This last group of clusters does not contain any three or four-member Si atom rings, further suggesting that the Si cores of these clusters have

the similar level of passivation as the low energy Si_7H_{14} clusters. As expected we find the GAM1 method obtained using the Si_7H_{14} training compounds dramatically outperforms the original AM1 method for $Si_{10}H_{16}$, $Si_{10}H_{20}$ and $Si_{14}H_{20}$ clusters. Figure 4.3 displays the typical energy-comparison charts for the under-passivated $Si_{10}H_8$ structures and for the two well-passivated stoichiometries $Si_{10}H_{16}$ and $Si_{14}H_{20}$. Figure 4.3 shows the general result that for well-passivated stoichiometries, such as $Si_{10}H_{16}$ and $Si_{14}H_{20}$, the GAM1 method produces local optimized relative energies which match well with the MP2 relative energies; whereas for the stoichiometries representative of a lower level of H passivation, such as $Si_{10}H_8$, the GAM1 parameters gives results which are not any better than those obtained from the original AM1 parameters. Our earlier work [154] suggests that a different set of GAM1 parameters obtained from a training set built from Si clusters corresponding to a lower level of H-passivation is needed to better reproduce the MP2 energies for this stoichiometry. Finally, it is worth noting that the diamond-lattice like $Si_{10}H_{16}$ and $Si_{14}H_{20}$ cluster, described in more detail below, both produce the lowest energies with the GAM1 and MP2 methods.

3. CGA only Global Optimization

In our IJQC paper we advocated the reoptimization of the AM1 parameters for each specific Si_xH_y stoichiometry [154]. Unfortunately, performing a global optimization using a coupled PGA and CGA soon becomes an expensive undertaking for large clusters such as $Si_{10}H_{16}$ and $Si_{14}H_{20}$. The results in the previous subsection suggest that the GAM1 parameters obtained for a certain level of H-passivation may be transferred to larger cluster with comparable levels of H-passivation. In this subsection we analyze whether the GAM1 parameters obtained from the Si_7H_{14} training set do enable a more reliable CGA only search for the ab initio global minimum of $Si_{10}H_{16}$ and $Si_{14}H_{20}$ than when using the AM1 parameters. In addition, to get some feel for the range of applica-

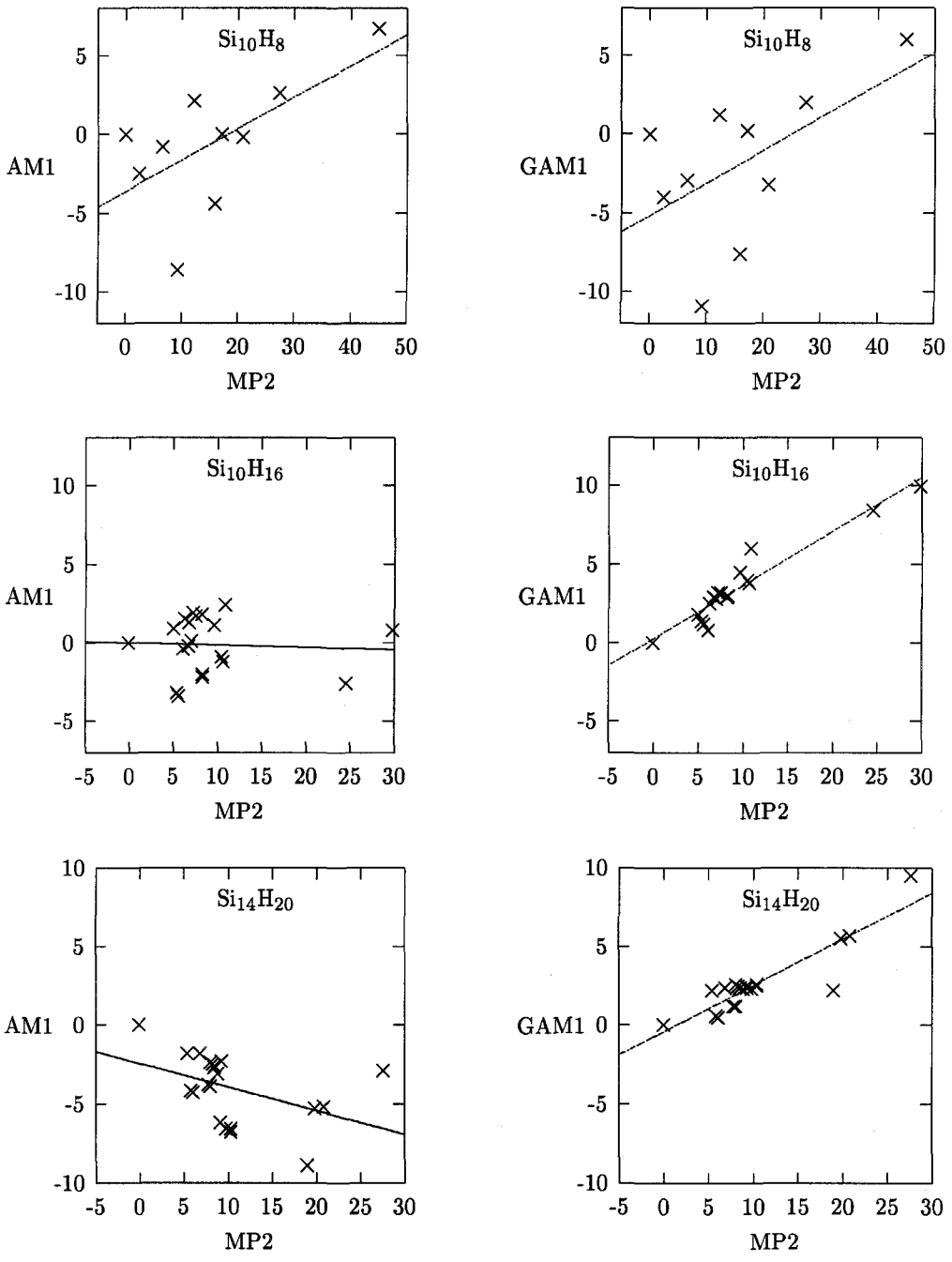


Figure 4.3: Comparison of AM1 and GAM1 relative optimized energies against MP2 relative optimized energies for representative stoichiometries $Si_{10}H_8$, $Si_{10}H_{16}$ and $Si_{14}H_{20}$. Energies are in kcal/mol.

bility of the Si_7H_{14} GAM1 parameters we also use the CGA to find the global minimum for the $Si_{10}H_{14}$ and $Si_{14}H_{18}$ stoichiometries which correspond to slightly lower extents of H-passivation.

$Si_{10}H_{16}$: We have repeated the global optimization of $Si_{10}H_{16}$ using the CGA and the GAM1 parameters obtained from the Si_7H_{14} training set to see if we still arrive at the same results as before [153]. Figure 4.4 shows the twenty lowest energy $Si_{10}H_{16}$ cluster structures listed in order of increasing MP2 energy found by the CGA. The global minimum 4(a) has a diamond-lattice like Si_{10} core structure. While the diamond-lattice structure is not unexpected, it is worth emphasizing that the CGA obtains this structure without any prior Si chemical bonding or structural bias being explicitly included in the algorithm. In our first CGA run, the global minimum structure 4(a) was found after only performing 630 local GAM1 geometry optimizations. Nonetheless, to ensure that the 4(a) structure is the true GAM1 global minimum, we kept the CGA running until 9000 local GAM1 geometry optimizations had been completed. We checked this result by running two other shorter CGA runs where only 3000 local geometry optimizations were performed but still found structure 4(a) as having the lowest total energy. The structures shown in figure 4.4 are the twenty distinct lowest GAM1 energy clusters found by the CGA which were subsequentially optimized at the MP2/6-31G* level. In the case of two enantiomers being located, we simply pick one of the structures to run the MP2 local optimization.

The Si_7H_{14} GAM1 parameters gives the diamond lattice fragment $Si_{10}H_{16}$ as the lowest energy cluster, whereas in our earlier work the AM1 calculations ranked this structure as having the tenth lowest energy [153]. In addition, the GAM1 calculations now produces quite a number of other low energy structures which are different from what were previously obtained [153]. We find the twenty lowest energy GAM1 structures

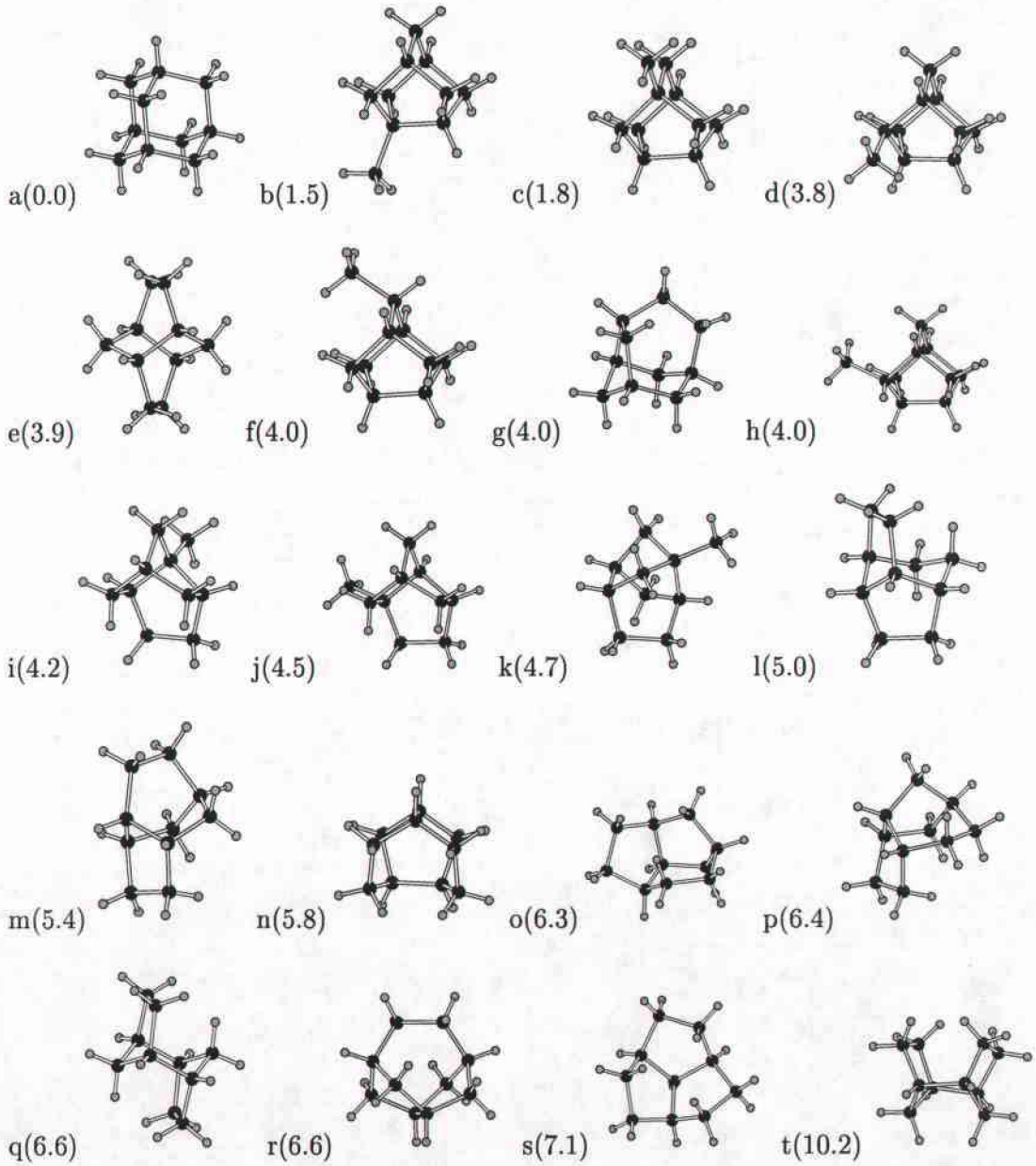


Figure 4.4: MP2/6-31G* optimized $Si_{10}H_{16}$ structures.

to produce a 10.2 kcal/mol energy range at the MP2/6-31G* level compared with the previous twenty lowest AM1 structures which spanned a 23.5 kcal/mol MP2/lanl2dz(d) energy range. The AM1 calculations also produced two structures (3rd and 12th lowest energy) consisting of $Si_{10}H_{14}$ and H_2 which we used earlier to suggest that $Si_{10}H_{16}$ may be over H-passivated [153]. Now, none of the present twenty low energy GAM1 structures shown in figure 4.4 have a H_2 dissociating from the $Si_{10}H_{16}$ nor do we find any structures to contain high energy three or four-member Si rings. The GAM1 and MP2 methods both rank the diamond-lattice like structure 4(a) as the global minimum while the AM1 method ranked structures 4(b) and 4(c) 3-4 kcal/mol below the diamond-lattice like structure 4(a). The five low energy structures 4(b-d), 4(f) and 4(h) have basically the same Si core framework but with a silyl group at different positions. Their Si frameworks differ from the diamond-lattice like structure 4(a) by forming a five-member ring after removing a SiH_2 from one of the six-member Si rings. Structure 4(e), like 4(a), also consists of 4 six-member Si rings but with an arrangement unrelated to the diamond lattice. Whereas 4(g) can be derived from 4(a) by shifting a SiH_2 from one of the six-member Si rings to another ring to produce a five- and a seven-member Si rings. The narrow MP2 relative energy range for structures 4(a) to 4(h) illustrates that the energy cost of interconverting the diamond lattice like structure of 4(a) to a related distorted structure can be very small. Some insight into why AM1 does not give 4(a) as its global minimum is indicated by the results in table 4.2 which shows the AM1 energy for six member Si rings may be as much as 3 kcal/mol too high. In general we find the Si_7H_{14} GAM1 parameters to yield locally optimized $Si_{10}H_{16}$ structures which are very close to the final MP2 optimized structures, whereas the AM1 optimized structures tend to exhibit much larger geometry differences from the ab initio structures. We have also checked that the $Si_{10}H_{16}$ cluster energy rankings are not affected by the inclusion of

zero point energy (ZPE) corrections. The vibrational frequency calculations show the ZPE to vary by less than 0.4 kcal/mol between the twenty different $Si_{10}H_{16}$ structures.

To summarize, by performing a CGA using the GAM1 parameters obtained from the Si_7H_{14} training set and then computing the locally optimized geometries of the resulting low energy $Si_{10}H_{16}$ structures we obtain stronger evidence supporting our previous proposal that the small to medium sized Si_xH_y clusters prefer to adopt diamond-lattice like structures [153]. In addition we find no low energy structures where the $Si_{10}H_{16}$ loses a H_2 molecule in agreement with our earlier work that this stoichiometry is the preferred level of complete H-passivation for the Si_{10} core [153].

$Si_{10}H_{14}$: To test the range of applicability for the GAM1 parameters from the Si_7H_{14} training set we have searched for the global minimum of the $Si_{10}H_{14}$ stoichiometry. Three independent CGA runs were performed. In one run 9000 local geometry optimizations were performed while the two other shorter runs used only 3000 local geometry optimizations simply to check whether the same global minimum was located by the multiple runs. They all find the global minimum structure 5(a) within 780, 600 and 300 local geometry optimizations, respectively. The geometries of the twenty lowest GAM1 energy structures were further optimized using the MP2/6-31G* method and the resulting structures are shown in figure 4.5 in the order of their MP2 relative energies. Our calculations find the GAM1 energy ranking for the different $Si_{10}H_{14}$ clusters still agrees well with the MP2 relative energies. Both the GAM1 and MP2 methods give structure 5(a) as the global minimum. Thus, even for Si_xH_y stoichiometries which have slightly lower levels of H-passivation, we feel the GAM1 parameters still provide a reasonable energy ranking.

The relative energy range for the lowest twenty $Si_{10}H_{14}$ structures shown in figure 4.5 is about 20 kcal/mol at the MP2 level. Thus, $Si_{10}H_{14}$ has a smaller number

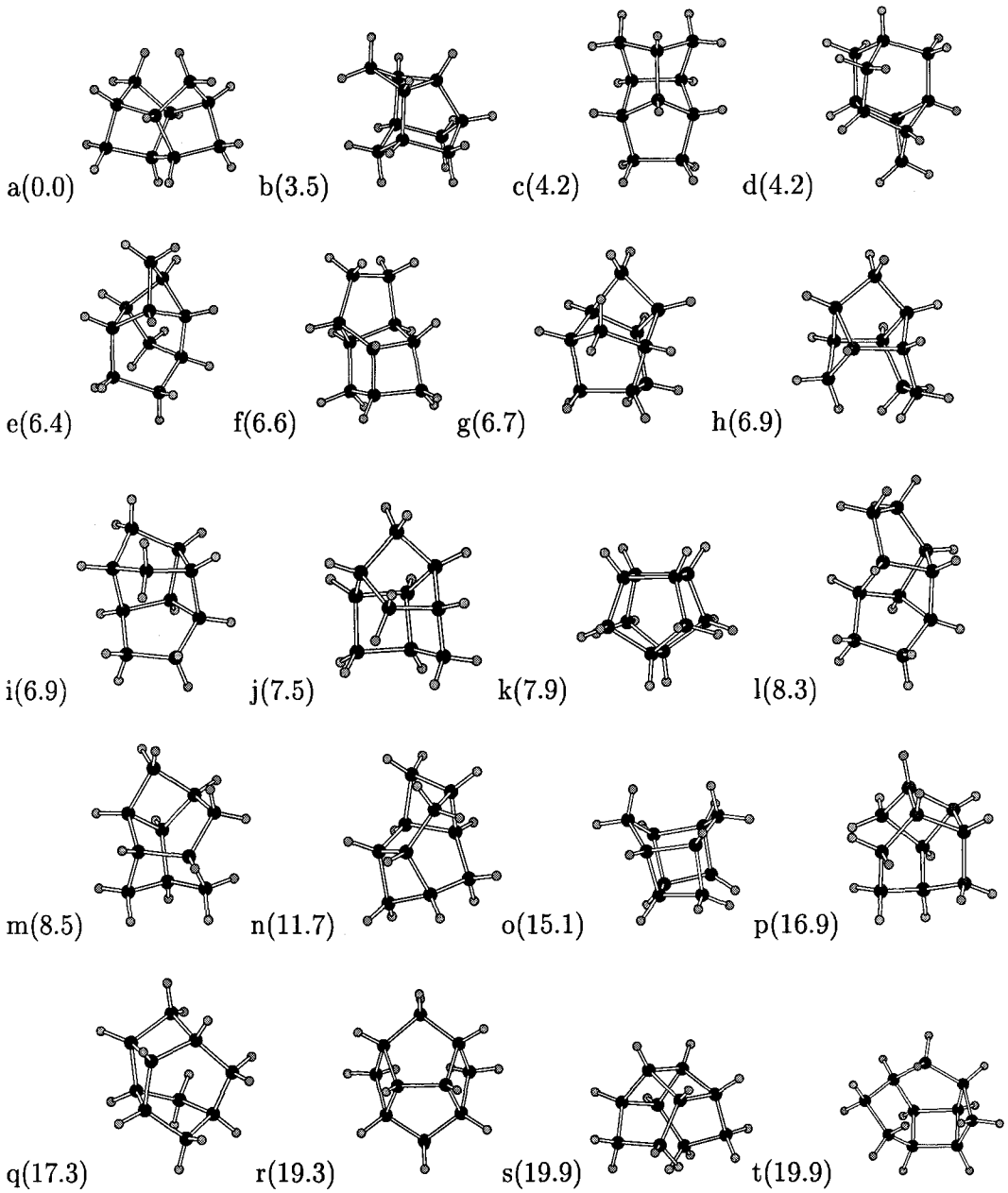


Figure 4.5: MP2/6-31G* optimized $Si_{10}H_{14}$ structures and their corresponding relative energies. In parentheses are the MP2 relative energies in kcal/mol.

of low energy isomers than we found for the $Si_{10}H_{16}$ stoichiometry. We think this is because the $Si_{10}H_{14}$ structures are more compact and they do not form several structures close in relative energy where a silyl group is just bound at different locations on the same Si_9 core. In contrast to $Si_{10}H_{16}$, the low energy $Si_{10}H_{14}$ structures only contain at most two six-member Si rings. Interestingly when we used the AM1 method in the CGA we obtained the molecule complex consisting of the structure 5(a) and a H_2 molecule as the third lowest AM1 energy $Si_{10}H_{16}$ structure [153]. As noted previously, the results in table 4.2 indicates that AM1 assigns too high an energy to six-member Si rings and this may explain why we found the $Si_{10}H_{14}$ global minimum so readily before. It is also worth noting that all structures except 5(a) and 5(c) in figure 4.5 contain at least one four-membered Si ring and the data in table 4.2 suggests that Si_xH_y clusters containing a four-membered Si ring are likely to have higher relative energy than clusters with five or six membered Si rings. This appears to be generally true for the different $Si_{10}H_{14}$ structures although 5(b) and 5(d) do contain four membered Si rings while still being close in energy to 5(c). This explains why clusters with surface silyl groups are absent because the remaining Si_9 core would contain more high energy three- or four-membered Si rings. Another consequence is that in figure 4.5 we do not find any $Si_{10}H_{14}$ structures corresponding to an even lower H-passivation, where H_2 is lost, since this would also result in a structure which contains more small Si rings or even a Si atom with dangling bonds. This is also supported by our previous work where we found that the lowest MP2 energy $Si_{10}H_{12}$ clusters have at least two four-membered Si rings [153].

To determine what level of H-passivation Si_{10} prefers we computed the Gibbs free energy obtained by augmenting the MP2/6-31G* electronic energy with data from a HF/6-31G* vibrational frequency calculation on the optimized structures. The com-

puted Gibbs free energy indicates that the $Si_{10}H_{16}$ 4(a) structure is around 7.8 kcal/mol more stable than the $Si_{10}H_{14}$ 5(a) structure and a H_2 molecule at 0 K. Not surprisingly, at room temperature this Gibbs free energy difference is reduced appreciably to 2.5 kcal/mol by a translational entropy contribution arising mainly from the H_2 molecule, suggesting that the H_2 and the $Si_{10}H_{14}$ 5(a) complex are nearly as stable as the $Si_{10}H_{16}$ global minimum structure.

The $Si_{10}H_{14}$ global minimum 5(a) demonstrates why cluster global optimization techniques, such as the CGA, are useful computational tools. Clearly, structure 5(a) cannot be readily generated by simply removing two H atoms from the $Si_{10}H_{16}$ global minimum structure 4(a) and then performing a local geometry optimization. The Si atoms in 4(a) are arranged as a fragment of the bulk Si lattice with several Si_6 rings, whereas Si framework in 5(a) is quite different with mostly Si_5 rings. In our previous CGA on $Si_{10}H_{16}$ we may have correctly found the $Si_{10}H_{14}$ global minimum using the AM1 parameters because as figure 4.5 shows the low energy $Si_{10}H_{14}$ structures generally only contain one six-member Si ring [153]. The GAM1 parameters from the Si_7H_{14} training set now enable us to confirm more carefully that the $Si_{10}H_{14}$ and $Si_{10}H_{16}$ global minima do have quite different structures.

$Si_{14}H_{20}$: We performed three independent CGA runs each with 9000 local geometry optimizations to search for the global minimum for the $Si_{14}H_{20}$ stoichiometry using the GAM1 parameters obtained from the Si_7H_{14} training set. The twenty lowest energy GAM1 structures 6(a-t) after locally optimizing at the MP2/lanl2dz level and are shown in figure 4.6. Two of the CGA runs locate the structure 6(c) as the lowest GAM1 energy cluster, while the other run locates the structure 6(a) as the lowest GAM1 energy cluster. Unfortunately the CGA does not succeed in finding the ab initio $Si_{14}H_{20}$ global minimum which all of our calculations so far suggest to be structure 6(u) corresponding

to a Si_{14} core having a diamond lattice like structure fully passivated by 20 H atoms. In our previous CGA using AM1, we found at least twenty $Si_{14}H_{20}$ structures with AM1 energies less than 6(u) [153] but now both the MP2 and GAM1 calculations produce the lowest energy for 6(u). We still do not find the diamond-lattice structure 6(u) among any of the structures obtained by the three CGA runs even though it is another 0.4 kcal/mol lower in GAM1 energy than the twenty 6(a-t) clusters found by the CGA. One reason why finding 6(u) is so difficult to find is because the GAM1 energies for the twenty 6(a-t) clusters only have a 3 kcal/mol spread. The 19 lowest $Si_{14}H_{20}$ structures found in the CGA are also all within 5 kcal/mol of each other at the ab initio level and generally correspond to a fairly similar Si_{13} core structure with a silyl group at different locations. We did obtain 6(h) which does not contain any silyl groups and appears to have a Si_{14} core resembling closest the global minimum structure 6(u).

To test further whether the diamond-lattice like structure is the global minimum at the GAM1 level, we ran a seeded CGA where we included structure 6(u) as one of the starting structures in the initial generation instead of using all randomly fabricated ancestors. The seeded CGA has both advantages and disadvantages when compared against the regular CGA without any seeds in the first population. The seeded GA obviously introduces prejudice into the CGA since it generates low energy clusters faster than the regular GA simply because from the beginning the seeded CGA runs contain low energy structures which can be utilized to provide good partial structures to form low energy offspring. However, we do not find the seeded GA to produce any new $Si_{14}H_{20}$ structures with lower GAM1 energy than the diamond-lattice structure 6(u), supporting that the 6(u) structure is the GAM1 global minimum. The seeded GA also locates the structure 6(c) as the next lowest GAM1 energy cluster. In addition two new low GAM1 energy structures 6(v) and 6(w) which have very similar Si framework as the MP2 global

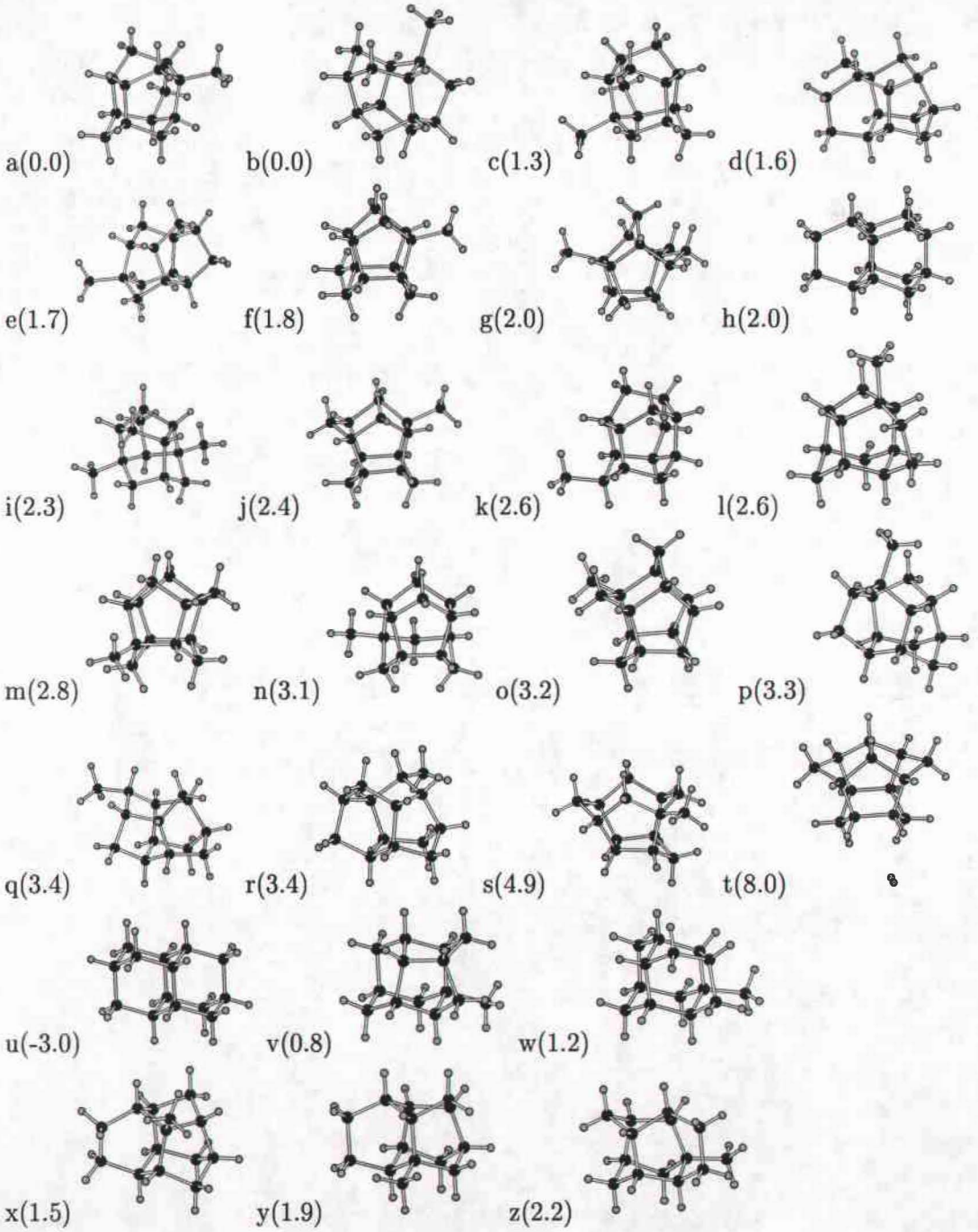


Figure 4.6: MP2/lanl2dz(d) optimized $Si_{14}H_{20}$ structures. In parentheses are the MP2 relative energies in kcal/mol. Structures (a-t) were obtained from the CGA while the structure (u) is the proposed global minimum at both GAM1 and MP2 level. Structures (v) and (w) were obtained from the seeded CGA.

minimum 6(u) were located. To explore more possible low energy cluster structure at the ab initio level, we performed another seeded GA starting with the 6(u) structure and 20 $Si_{14}H_{20}$ cluster structures we obtained in our previous research [153]. This second seeded CGA located three more low GAM1 energy clusters 6(x-z) which appear to be higher energy modifications of 6(u) but are different from the seed structures and the 6(a-w) structures. The present results further support that the structure 6(u) is the global minimum at both the GAM1 and MP2 level.

We notice that many of the low energy $Si_{14}H_{20}$ clusters again have very similar Si_{13} frameworks but with a silyl group at different surface positions. In contrast, structure 6(u) contains a distinct Si_{14} framework which cannot as readily be obtained from the mating procedures when either of the two parents have a silyl group. Structure 6(h) is one of the only structures without a silyl group but it would take a very fortunate cluster cut and mating operation before structure 6(u) could be generated. A part of the reason that structure 6(u) may be so difficult to obtain is due to the covalent nature of the Si and H bonds, the global optimization of clusters composed of only one element will always easier because there will always be some atoms with incomplete bonding. We had a similar problem in the global optimization of the incompletely passivated Si_6H_6 stoichiometry, where we could find fairly easily the optimal Si_6 framework but our mating processes did not readily allow the identification of the best H atom positions [154]. Presently we are investigating whether a mutation operation could be developed which would enable a SiH_2 group to be inserted between or removed from two other neighboring Si atoms.

Interestingly, the highest energy structure 6(t) shown in figure 4.6 corresponds to $Si_{14}H_{20}$ losing a H_2 molecule and is suggestive of possible over H passivation. Previously, and obviously unrealistically, we obtained this same 6(t) structure as the AM1 global

minimum of the $Si_{14}H_{20}$ stoichiometry [153]. Thus, although we still do not obtain diamond-lattice like structure 6(u) as the lowest energy structure in the CGA, using the GAM1 parameters from the Si_7H_{14} training set does produce a semiempirical method which excels the AM1 method by correctly ranking the 6(u) as the global minimum in perfect agreement with the ab initio method.

$Si_{14}H_{18}$: To further investigate whether the $Si_{14}H_{18}$ cluster in the above 6(t) structure is also the GAM1 and/or MP2 global minimum cluster with $Si_{14}H_{18}$ stoichiometry, we have performed three independent CGA runs each with 9000 local geometry optimizations to search for the global minimum with the $Si_{14}H_{18}$ stoichiometry. All three CGA runs locate the same $Si_{14}H_{18}$ global minimum after 1290, 6420 and 1800 local geometry optimizations. The resulting twenty lowest GAM1 energy $Si_{14}H_{18}$ clusters after locally optimizing at the MP2 level are shown in figure 4.7. As expected the $Si_{14}H_{18}$ GAM1 global minimum, structure 7(a), is the same $Si_{14}H_{18}$ cluster shown in the 6(t) structure. Since 7(a) contains only three six-member Si rings our previous CGA on $Si_{14}H_{20}$ with AM1 [153] and the present CGA with the GAM1 parameters appear to have little difficulty in finding this global minimum. Indeed, we can now understand our previous result since the present work shows that the AM1 calculations are assigning too high a relative energy to six-member Si rings and the $Si_{14}H_{18}$ in 6(t) (or 7(a)) has fewer six-member rings than the other lower energy $Si_{14}H_{20}$ structures. The fact that 7(a) is also the lowest MP2 energy cluster further justifies the transferability of the GAM1 parameters obtained using a Si_7H_{14} training set. Unlike the 6(u) diamond-lattice structure where all the six-member Si rings adopt the chair conformation, the three six-member Si atom rings in structure 7(a) all have the boat conformation. We expect the H atoms at the "flagpole" positions on the Si_6 rings to repel each other since they are separated by only $\approx$ 2.9 Å. This may explain why the structures 7(b) and

7(c) are less than 1 kcal/mol higher in energy than 7(a) since 7(b) and 7(c) lose two flagpole H atoms and its underlying Si to give a silyl group which binds at different positions on the remaining Si_{13} framework. Besides structures 7(b) and 7(c), many other low MP2 energy clusters such as structures 7(d-h) have Si frameworks similar to the 7(a) structure and as a consequence are within a small 2 kcal/mol energy range. The energy range for the twenty lowest energy $Si_{14}H_{18}$ clusters is only 5.3 kcal/mol and is significantly narrower than the 19.9 kcal/mol determined for $Si_{10}H_{14}$. We think this is because all Si atoms in $Si_{14}H_{18}$ are able to remain tetrahedrally coordinated without requiring any three- or four-member Si rings to be formed.

The inclusion of the HF/3-21G ZPE calculation suggests that the Gibbs energy of $Si_{14}H_{20}$ 6(u) diamond-lattice like structure is 7.6 kcal/mol lower than the total Gibbs energy of $Si_{14}H_{18}$ 7(a) structure and a H_2 molecule at 0 K. However, the entropy contribution from the translational mode of the H_2 molecule causes the total Gibbs free energy of a H_2 molecule and a $Si_{14}H_{18}$ 7(a) cluster to be only 2.7 kcal/mol higher than that of the diamond-lattice like $Si_{14}H_{20}$ global minimum cluster at the room temperature.

The CGA with the Si_7H_{14} GAM1 parameters find a global minimum 7(a) for $Si_{14}H_{18}$ which is consistent with our earlier CGA results on $Si_{14}H_{20}$ using AM1. The $Si_{14}H_{18}$ global minimum cannot be simply obtained by removing two H atoms from the $Si_{14}H_{20}$ global minimum 6(u) but it does contain only three six-member Si rings enabling both AM1 and GAM1 to locate 7(a). However, since the Si_7H_{14} GAM1 parameters produce $Si_{14}H_{20}$ and $Si_{14}H_{18}$ global minima and low energy structures better matching the MP2 rankings we have greater confidence in the present results.

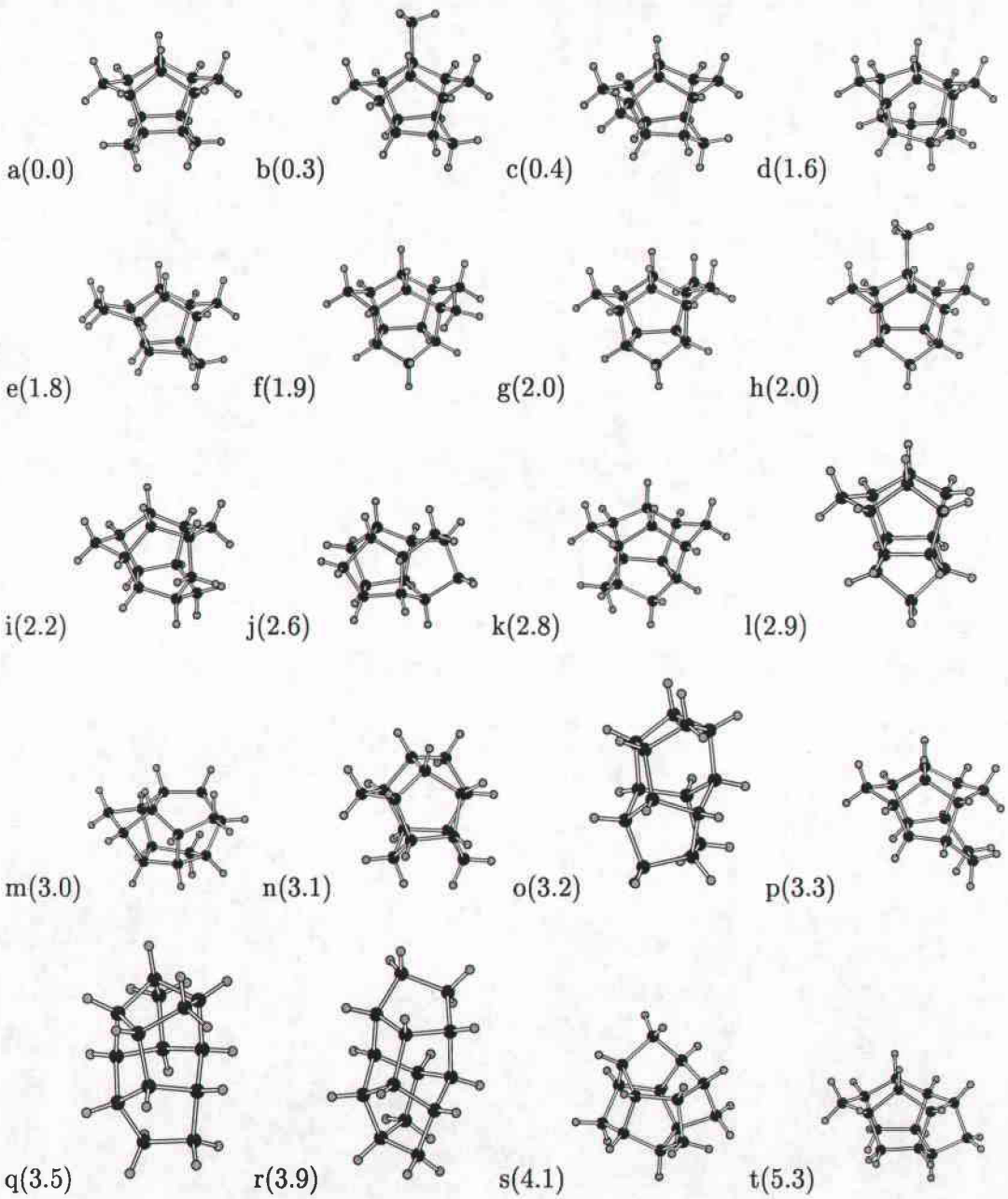


Figure 4.7: MP2/lanl2dz(d) optimized $Si_{14}H_{18}$ structures. In parentheses are the MP2 relative energies in kcal/mol.

4.4 Conclusion

We have applied a parametrization genetic algorithm (PGA) to determine the GAM1 parameters using a Si_7H_{14} training set consisting of 14 different structures locally optimized at the MP2/6-31G* level. The resulting GAM1 method reproduces the Si_7H_{14} MP2/6-31G* relative energies and geometries in better agreement than the original AM1 method. The GAM1 parameters appear to correct for the too high relative energy calculated by AM1 for structures containing six member Si rings.

Rather than perform coupled PGA and cluster genetic algorithm (CGA) to find the global minimum structure for large clusters we have investigated the transferability of the GAM1 parameters derived from a Si_7H_{14} training set to larger clusters. We first evaluated the GAM1 optimized energies of the clusters with various stoichiometries ($Si_{10}H_y$, $y = 4, 8, 12, 16, 20$ and $Si_{14}H_{20}$) we located in previous research work [153]. Our results show that the GAM1 and MP2 relative energies and energy rankings match well for the $Si_{10}H_{16}$, $Si_{10}H_{20}$ and $Si_{14}H_{20}$ stoichiometries which have the similar level of passivation as the Si_7H_{14} stoichiometry.

To further illustrate the transferability of this GAM1 parameters for larger clusters, we have globally optimized the $Si_{10}H_{14}$, $Si_{10}H_{16}$, $Si_{14}H_{18}$ and $Si_{14}H_{20}$ clusters using the Si_7H_{14} GAM1 parameters to evaluate the cluster energies. We find the GAM1 and MP2 energy rankings match well for each of the 4 above stoichiometries and the GAM1 global minimum clusters also have the lowest MP2 energies. Unlike for the AM1 calculations, we find the diamond-lattice like $Si_{10}H_{16}$ and $Si_{14}H_{20}$ structures to have lowest GAM1 and MP2 energies. These results further confirm that the $Si_{10}H_{16}$ and $Si_{14}H_{20}$ clusters prefer diamond-lattice structures at 0 K. However, vibrational frequency calculations show that the total Gibbs free energy of the $Si_{10}H_{14}$ or $Si_{14}H_{18}$ global minimum and a H_2 molecule is close to that of the diamond-lattice $Si_{10}H_{16}$ or $Si_{14}H_{20}$ cluster

respectively at the room temperature.

Finally, the ability to transfer GAM1 parameters obtained from a Si_7H_{14} training set to larger well-passivated Si_xH_y clusters should enable us to develop a fast global optimization strategy for larger Si_xH_y clusters at the ab initio level providing the GAM1 parameters are obtained from a training set built from a specific stoichiometry with a comparable level of passivation.

CHAPTER 5. FAST GLOBAL OPTIMIZATION OF Si_xH_y CLUSTERS: NEW MUTATION OPERATORS IN THE CLUSTER GENETIC ALGORITHM

(This chapter is published in *Chem. Phys. Lett.* **2004**, *398*, 107.)

5.1 Introduction

Nanometer sized Si particles have promising optoelectronic properties which could be useful in various technological applications. Since Si only clusters are highly unstable due to dangling bonds on the surface Si atoms a practical device will most likely be fabricated from silicon clusters whose surfaces are passivated by some chemically inert ligand. For these reasons we have been theoretically determining the globally optimized structures of Si_xH_y clusters where the H atoms serve as a prototypical passivating ligand. This structural information is difficult to obtain experimentally and is needed to correctly model the optoelectronic properties of Si nanoparticles. To find the global minimum structure of Si_xH_y clusters at the ab initio level we have recently developed a cluster genetic algorithm (CGA) strategy [153–155]. In our first procedure we used the AM1 semiempirical method [120] as a fast computational tool for prescreening the energy of the different Si_xH_y structures before performing the more demanding ab initio calculations [153]. We soon realized that for an effective CGA we needed better energy prescreening than provided by AM1. This led us to develop a second approach where we used an iterative strategy involving both a parametrization genetic algorithm (PGA) and a cluster genetic algorithm (CGA) [154]. The PGA reparametrized the AM1

method against ab initio data to produce the GA-optimized AM1 (GAM1) parameters for a specific Si_xH_y stoichiometry [154]. We obtained two different GAM1 parameter sets for the Si_6H_2 and Si_6H_6 clusters which significantly improved the energy ranking ability of the semiempirical method. This iterative strategy enabled us to find the Si_6H_2 global minimum structure at the ab initio level and three nearby low energy local minima which closely resemble the Si_6H_6 ab initio global minimum. In figure 5.1 we show the Si_6H_6 ab initio 1(a) global minimum and two GAM1 structures obtained by the CGA: structure 1(b) is the GAM1 global minimum and 1(c) is one of the low energy structures with the same Si_6 core as the ab initio 1(a) global minimum but with a different H atom arrangement.

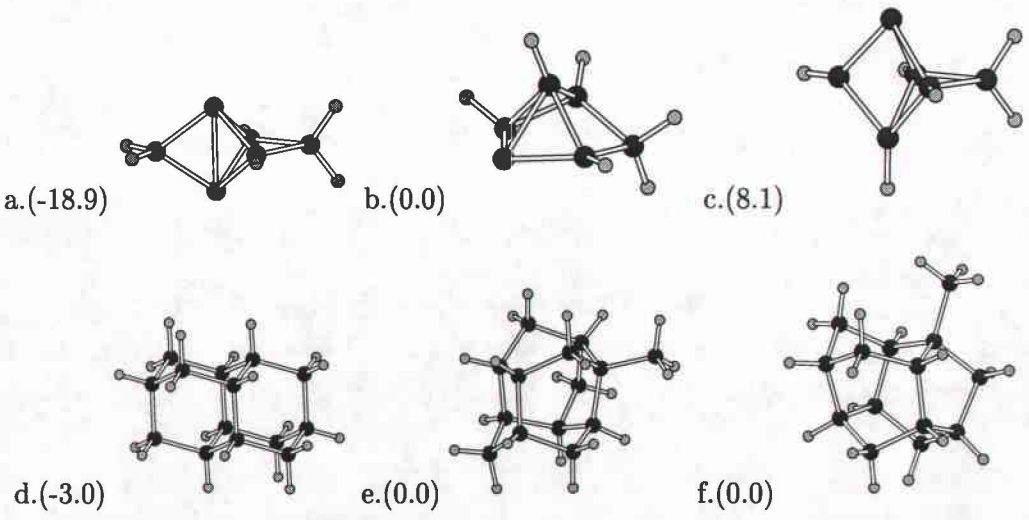


Figure 5.1: Low energy structures for Si_6H_6 and $Si_{14}H_{20}$ where in parenthesis are listed their different relative MP2 energies in kcal/mol. For Si_6H_6: (a) MP2 global minimum, (b) GAM1 global minimum, and (c) one of the low energy structures from the CGA resembling (a) [154]. For $Si_{14}H_{20}$: (d) MP2 global minimum and (e) and (f) the two lowest energy structures found by the CGA [155].

Obtaining specific GAM1 parameters for each stoichiometry soon becomes very computationally demanding and the coupled PGA/CGA procedure is impractical for Si_xH_y clusters with more than 10 Si atoms. Therefore in our most recent work we investigated under what conditions the GAM1 parameters from a small Si_xH_y cluster can be applied to a larger cluster so that only the CGA global search need be performed [155]. We choose Si_7H_{14} as the starting stoichiometry since it can adopt various possible structures which model a Si_7 core completely passivated by H atoms. We found that the GAM1 parameters obtained by fitting to the MP2 optimized structures and energies for a Si_7H_{14} training set were transferable to larger clusters such as the $Si_{14}H_{20}$ clusters providing there was a comparable level of H passivation. Semiempirical calculations using the Si_7H_{14} GAM1 parameters and the MP2 method both rank the diamond-lattice like $Si_{10}H_{16}$ and $Si_{14}H_{20}$ structures as the global minimum. We also found the Si_7H_{14} GAM1 parameters reproduced the MP2 energy ranking for stoichiometries corresponding to slightly lower levels of H passivation, such as in the $Si_{10}H_{14}$ and $Si_{14}H_{18}$ clusters, but the quality of the GAM1 energy ranking deteriorated for stoichiometries where the Si_x core is only partially H passivated [155].

In contrast to our initial work [153], the Si_7H_{14} GAM1 parameters now enabled the CGA to correctly identify the global minimum of $Si_{10}H_{16}$, $Si_{10}H_{14}$ and $Si_{14}H_{18}$ clusters [155]. Unfortunately, the CGA was still not able to find the 1(d) $Si_{14}H_{20}$ global minimum, and instead the lowest energy structures such as 1(e) and 1(f) shown in figure 5.1 were obtained. We take the diamond-lattice like structure to be the global minimum since the CGA never generated another $Si_{14}H_{20}$ structure which has a lower ab initio or GAM1 semiempirical energy. Many of the low energy $Si_{14}H_{20}$ structures found by the CGA, such as 1(e) and 1(f) closely resemble the 1(d) global minimum but have one of the SiH_2 units present in one of the six-membered Si rings moved to form a

SiH_3 group on the cluster surface. The CGA is unable to rapidly find the $Si_{14}H_{20}$ global minimum because none of the three mating methods used in the CGA readily create offspring where the surface SiH_3 is recombined back into Si_{x-1} core. We have a similar problem due to the mating operations not generating different H atom arrangements while maintaining the same Si_6 core structure when we globally optimize the Si_6H_6 cluster.

Three different mating operators were used in our previous CGA [153–155]. The most important mating operator is the cut-and-paste method first proposed by Deaven and Ho in a GA for global optimization of atomic clusters [102]. The method cuts two parent clusters into halves along randomly orientated planes and recombines two halves from different parents into a offspring structure: a more fit offspring inherits desirable partial structures from the parents. The other two mating operators were really designed to cause mutations and introduce additional structural diversity into the cluster population. The coordinate averaging method takes the mathematical mean of the cartesian coordinates for two parent cluster geometries to produce cartesian coordinates of the offspring cluster. The fragment mixing operator replaces a fragment of the Si core from one parent structure by a Si fragment from the other parent and also replaces a fragment of the H shell of one parent structure by a fragment of the H shell from the other parent. Although both the coordinate averaging and fragment mixing operators introduce structural diversity to the CGA they often result in unphysical starting cluster geometries which causes many of the local geometry optimizations to fail. The goal of this chapter is to describe 4 new mutation operators which replace the coordinate averaging and fragment mixing operators in our previous CGA. We show this new CGA, which produces new offspring using the original cut-and-paste operator and the 4 new mutation operators, is now able to find the global minima for Si_6H_6 and

$Si_{14}H_{20}$. In addition, there is an overall reduction in the number of energy evaluations needed before finding the Si_xH_y global minimum with the new CGA.

The remainder of the chapter is organized as follows. In next section we first describe in detail the new mating operators now utilized by the CGA. We also give a brief description of the PGA used to obtain new Si and H GAM1 parameters for the Si_7H_{14} training set along with other general computational details. We then have the results and discussion section where we present the new GAM1 parameters and describe the improved performance of the new CGA. In the final part of this section we describe the testing the new CGA to globally optimize the larger $Si_{18}H_{24}$ and $Si_{18}H_{22}$ stoichiometries. Concluding remarks are given in the final section.

5.2 Computation Method

5.2.1 Cluster Genetic Algorithm (CGA) and the new mutation operators.

We use the same basic CGA as used in our earlier work [155] but have modified the genetic operators which produce new offspring from the parent clusters. We have kept our original cut-and-paste mating method and discard the other two genetic operators which often resulted in too drastic structural modifications.

We now add the four new mutation/mating operations which are shown in figure 5.2. The first three of these operations are mutation operations and can be thought of as analogues to the add/etch operations which Wolf and Landman used in their global optimization studies of large Lennard-Jones clusters [160]. The covalent bonding network ubiquitous to Si_xH_y clusters makes implementing the add/etch operations more difficult and requires that each Si atom be assigned a functional group identification which we simply determine through connectivity information obtained from the internuclear separation matrix. For each Si atom in a cluster we determine it to be

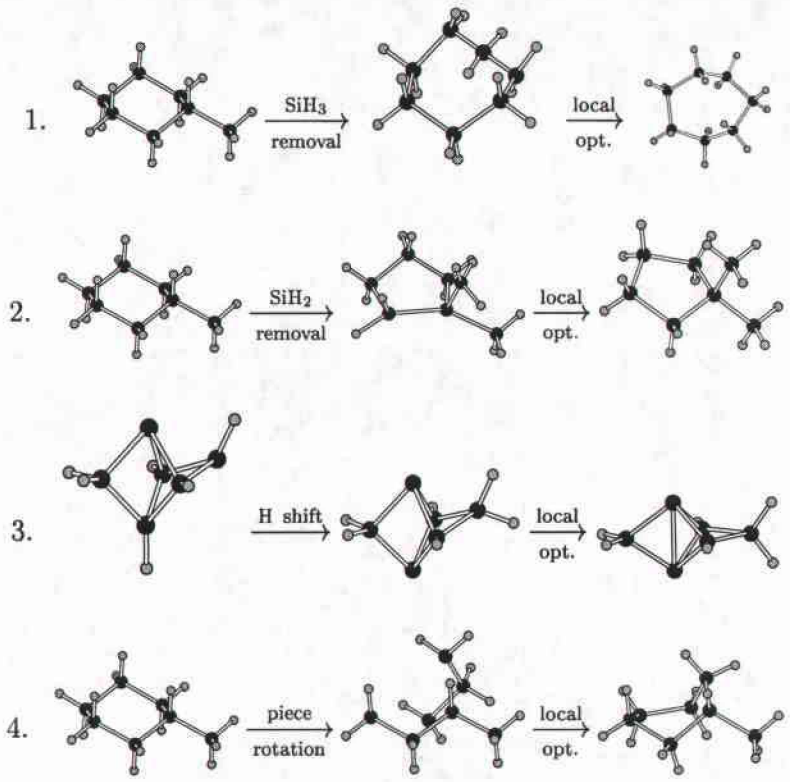


Figure 5.2: Examples showing the four new mutation operators acting on a parent cluster to produce an offspring which is then locally optimized.

connected to another either H or Si atom when their internuclear separations are less than 120% of the usual Si-H and Si-Si bond length (1.48 and 2.35 Å) respectively. Thus a SiH_3 group consists of a Si atom connected to 3 H atoms and a SiH_2 group has a Si atom connected to 2 H atoms.

The first new mutation operation, shown in figure 5.2, removes a peripheral SiH_3 group and replaces it with a single H atom and inserts a SiH_2 group between two randomly selected neighboring Si atoms to produce either a longer Si chain or an expanded Si ring. Figure 5.2 illustrates how the SiH_3 removal operator expands the 6-membered

Si-ring in silyl-hexasilane to a 7-membered ring. The second mutation operation is just the reverse procedure of the first one: it removes a SiH_2 group from the cluster and then a randomly chosen H atom is replaced by a SiH_3 group. To reduce making a drastic geometry change the two Si atoms which were neighboring the SiH_2 group are pulled slightly towards each other to induce the formation of a new Si-Si bond. The second mutation operation either reduces the size of a ring or shortens a chain by one Si atom. The SiH_2 removal operation in figure 5.2 transforms silyl-hexasilane to the 5-membered ring structure disilyl-pentasilane. The third mutation operation is a H shift operator which simply shifts a H atom from one Si atom to another. It is only applied to clusters which contain incompletely passivated Si atoms as illustrated for the Si_6H_6 cluster in figure 5.2. The last operator, a piece-rotation mutation method, is a single-parent cut-and-paste operator and is similar to the operator used by Rata *et. al.* [46] for the global optimization of Si-only clusters. The Si_xH_y clusters is randomly cut into two parts and then one half of the cluster is randomly rotated relative to the other half. Figure 5.2 shows the piece-rotation operator acting on silyl-hexasilane causing the chair conformation to be mutated to a distorted boat structure. When applying the CGA to well H passivated Si_xH_y clusters we use the cut-and-paste, SiH_3 removal, SiH_2 removal and piece-rotation operators with equal probability; whereas in the global optimization of low-passivated Si clusters the cut-and-paste, piece-rotation and H-shift genetic operators are selected with 25%, 25% and 50% probabilities, respectively.

The global optimization starts from 120 randomly generated locally optimized structures as the ancestor population. The fitness value of an individual is taken as the absolute value of the GAM1 energy of the locally optimized cluster. Tournament selection is used to select the parents needed to produce offspring. Since the calculations are performed on either a 16 or a 32 processor computer we do not explicitly treat

separate generations. The CGA performs parent selection, mating/mutation process and generates a new trial offspring structure as soon as the local geometry optimization of a cluster included in the parent population is finished on a processor. To avoid identical structures dominating the population new offspring which already exist in the current population are discarded.

5.2.2 Parametrization genetic algorithm (PGA).

To ensure that the semiempirical parameters give energy rankings and geometries which are consistent with the results from the ab initio calculations on the large well passivated Si_xH_y clusters reported in this chapter we used the PGA to obtain new GAM1 parameters for H and Si. We used the same Si_7H_{14} training set as before [155] but, as described below, the ab initio calculations now utilize the LanL2dz(d) effective core potential (ECP) and basis set [23, 147, 148]. The 16 AM1 parameters: U_{ss}, U_{pp}, ζ_s, ζ_p, β_s, β_p, α, K_{1-3}, L_{1-3} and M_{1-3} were reoptimized for Si, whereas only 13 AM1 parameters were reoptimized for the H atom since there are no p orbital parameters [120, 146]. The new H and Si GAM1 parameters were obtained with the same basic PGA as described previously [155]. The new GAM1 parameters are obtained by globally minimizing a fitness function f consisting of an energy and a gradient part:

$$f = P\left(\left[\sum_{i=1}^{N_c}(E_i^{\mathrm{GAM1}} - E_i^{\mathrm{MP2}} + D)^2/N_c\right]^{\frac{1}{2}} + \left[\sum_{i=1}^{N_c}(G_i^{\mathrm{GAM1}})^4/N_c\right]^{\frac{1}{2}}\right) \qquad (5.1)$$

where E_i^{GAM1} and G_i^{GAM1} are the GAM1 energies and RMS gradients and E_i^{MP2} is the corresponding MP2 geometry optimized energy for the i-th member of the training set. Some of the trial GAM1 parameter sets make the SCF convergence difficult and

f is only evaluated with the N_c converged semiempirical SCF calculations. D is the difference between the MP2 and GAM1 average energies for the training compounds and is introduced to maximize the coincidence of the GAM1 and MP2 relative energies:

$$D = \sum_{i=1}^{N_c} (E_i^{\mathrm{MP2}} - E_i^{\mathrm{GAM1}})/N_c \tag{5.2}$$

P is a penalty coefficient which is unity when all the semiempirical calculations achieve SCF convergence:

$$P = e^{2.0(N_T - N_c)/N_T} \tag{5.3}$$

where $N_T(= 14)$ is the total number of compounds in the training set.

5.2.3 Semiempirical and ab initio methods.

All of the semiempirical and ab initio calculations were performed using the GAMESS program [149]. The 14 Si_7H_{14} training set compounds and the 20 lowest GAM1 energy Si_xH_y clusters obtained in the CGA were locally optimized with the MP2 method [132]. Since we are finding the global minimum of fairly large Si_xH_y clusters the LanL2dz(d) scheme consisting of the Hay-Wadt effective core potential (ECP) and valence basis set for Si [147] augmented with a d function ($\zeta = 0.45$) [23] and the Dunning-Hay basis for H [148] was used. Consistent with the work of others [23, 79], our tests show that the relative energies of the well-passivated Si clusters calculated by the MP2 method match well with the much more expensive methods which include higher levels of electron correlation such as MP4 and CCSD(T).

5.3 Result and Discussion

Table 5.1 lists the best new H and Si GAM1 parameters obtained from 3 separate PGA runs using the Si_7H_{14} training set with optimized geometries and energies computed using the MP2/LanL2dz(d) method. Also included in table 5.1 are the original H and Si AM1 parameters [120, 146], and the Si GAM1 parameters from the previous Si_7H_{14} [155] and Si_6H_6 [154] training sets. The Si_6H_6 GAM1 parameters listed here are obtained with the one-center two-electron integrals being computed correctly [120, 154, 155]. For Si_7H_{14} the new Si GAM1 parameters are now closer to the original AM1 values than the previous GAM1 Si only parameters. We think this is because reoptimizing the H and Si parameters reduces the likelihood of the Si parameters compensating for deficiencies due to fixed AM1 H parameters. The GAM1 ζ_s for the Si and H atoms are decreased by 12% and 14%, respectively; while the Si ζ_p is slightly increased by 3%. The reduction in the Si ζ_s/ζ_p ratio is probably the main origin for the improved performance of the new Si and H GAM1 semiempirical method since the the other parameters are not appreciably modified during the reparametrization. After completing the PGA step, the new Si and H GAM1 parameters were used to locally optimize the Si_7H_{14} training compounds. We find the resulting new GAM1 geometries are in much better agreement with the MP2 geometries than with the AM1 parameters. Previously the main geometry discrepancy was for training compound (k) which is not strictly a Si_7H_{14} monomer but really consists of separate Si_7H_{12} and H_2 molecules. With the previous Si only GAM1 parameters, the geometry for H_2 is determined by the AM1 parameters and results in a much too short 0.677 Å H-H bond length. Now, with the new Si and H GAM1 parameters, we obtain a 0.734 Å H-H bond length which is very close to the 0.738 Å calculated by the MP2/LanL2dz(d) method. Consequently, using the new Si and H GAM1 parameters to locally optimize structure

(k) from it's MP2/LanL2dz(d) training set geometry results in a small energy relaxation of 0.2 kcal/mol versus the 1.5 kcal/mol relaxation obtained previously. Table 5.2 shows that the MP2/LanL2dz(d) and MP2/6-31G(d) methods agree well in predicting the relative energies for the 14 Si_7H_{14} training compounds. Also the GAM1 method provides much more reliable relative energies than the AM1 method using the ab initio data as a reference.

Table 5.1: Comparison of the AM1 parameters and the improved GAM1 parameters using the PGA. GAM1[a] are the present Si and H parameters and GAM1[b] are the previous Si only parameters with both sets being derived from the Si_7H_{14} training set [155]. GAM1[c] are the Si parameters obtained from the Si_6H_6 training set [154]. The local geometries and energies of the GAM1[a] training set was computed with the MP2/LanL2dz(d) method while the GAM1[b] and GAM1[c] training set used the MP2/6-31G* method.

	H		Si			
Param.	AM1	GAM1[a]	AM1	GAM1[a]	GAM1[b]	GAM1[c]
U_{ss}/eV	-11.396427	-11.356885	-33.953622	-33.723108	-34.269578	-32.918202
U_{pp}/eV			-28.934749	-29.628321	-30.694633	-26.183913
ζ_s/au	1.188078	1.049572	1.830697	1.578831	1.500623	1.685995
ζ_p/au			1.284953	1.327590	1.356890	1.300521
β_s/eV	-6.173787	-5.900161	-3.784952	-3.716276	-3.321166	-3.282843
β_p/eV			-1.968123	-1.893157	-1.862159	-3.040449
α/Å $^{-1}$	2.882324	2.751108	2.257816	2.336323	2.249007	2.141541
K_1	0.122796	0.119580	0.250000	0.231462	0.249818	0.183072
K_2	0.005090	0.005199	0.061513	0.062800	0.065260	0.114281
K_3	-0.018336	-0.017641	0.020789	0.019162	0.020458	0.021690
L_1	5.000000	4.932948	9.000000	9.739164	10.141849	5.598440
L_2	5.000000	5.089244	5.000000	5.347795	5.309296	6.848934
L_3	2.000000	1.929803	5.000000	4.490209	4.221081	3.091319
M_1	1.200000	1.028925	0.911453	0.858012	0.865394	0.919845
M_2	1.800000	1.816682	1.995569	1.885485	1.955345	1.608310
M_3	2.100000	2.452170	2.990610	2.946521	2.812023	8.058081

We have repeated the CGA runs using both the new Si_7H_{14} Si and H GAM1

Table 5.2: Comparison of the Si_7H_{14} relative energies for MP2 optimized geometries versus the AM1 and GAM1 relative energies calculated at the MP2 optimized geometry (S. P.) and at the optimized semiempirical geometry (Opt.). The RMS deviation between the semiempirical and the MP2 energies is evaluated using equation (2). All energies are in kcal/mol.

	Si ring	MP2		AM1		GAM1	
Structure	size	lanl2dz(d)	6-31G*	S. P.	Opt.	S. P.	Opt.
1(a)	5	0.0	0.0	0.0	-5.0	0.0	-0.4
1(b)	6	0.2	0.8	6.5	1.3	1.4	1.0
1(c)	6	0.8	1.3	6.2	1.2	0.7	0.1
1(d)	5	1.6	1.8	3.9	-1.6	1.5	0.6
1(e)	5	1.8	1.9	3.8	-1.6	1.3	0.2
1(f)	5	1.8	2.0	3.6	-1.7	1.3	0.9
1(g)	5	1.8	2.0	3.4	-1.8	1.0	0.0
1(h)	7	3.1	3.7	10.4	4.9	2.9	2.1
1(i)	6	3.2	3.8	6.1	1.3	1.8	0.8
1(j)	5	3.8	4.2	5.9	1.2	1.3	0.6
1(k)	5, 5	8.7	6.4	3.9	-4.2	8.0	7.8
1(l)	4	11.2	10.1	11.0	4.3	10.0	9.2
1(m)	3	22.7	21.6	19.2	12.9	20.6	19.2
1(n)		51.8	50.7	45.2	40.4	51.5	51.0
RMS Deviation			0.9	3.9	4.3	0.9	1.1

parameters and the new mutation operators to globally optimize the $Si_{10}H_{16}$, $Si_{10}H_{14}$, $Si_{14}H_{20}$ and $Si_{14}H_{18}$. We performed 5 CGA runs on each stoichiometry and obtained not only the same global minimum structures for the $Si_{10}H_{16}$, $Si_{10}H_{14}$ and $Si_{14}H_{18}$ stoichiometries as in our previous research, but the CGA now produces the diamond-like lattice structure as the GAM1 global minimum of the $Si_{14}H_{20}$ cluster. The final global minimum structures for each stoichiometry are illustrated in figure 5.5. To check that the semiempirical method correctly ranks the low energy structures we locally optimize the clusters with the 20 lowest GAM1 energies at the MP2 /LanL2dz(d) level. So far none of these MP2/LanL2dz(d) local geometry optimizations have resulted in a cluster with an energy ranking significantly different from the GAM1 ranking. The new mutation methods dramatically improve the efficiency of the CGA. It took 560 $\pm$ 242, 1170 $\pm$ 1041 and 3170 $\pm$ 2826 local optimizations to locate the global minimum of the $Si_{10}H_{14}$, $Si_{10}H_{16}$ and $Si_{14}H_{18}$ clusters in our previous research [155]. These numbers drop to 438 $\pm$ 300, 841 $\pm$ 241 and 1448 $\pm$ 518 respectively. Most importantly, we located the diamond-lattice like $Si_{14}H_{20}$ structure as the GAM1 global minimum in 3 out of 5 CGA runs using only 3000 local optimizations whereas before the global minimum was not located in any of the 3 CGA runs where 9000 local optimizations were performed per CGA [155].

To explore the effectiveness of the H-shift mutation method, we performed 5 CGA runs to globally optimize the Si_6H_6 clusters using the Si only Si_6H_6 GAM1 parameters list in table 5.1. Each CGA was allowed to perform a maximum of 1000 local optimizations. The previous CGA found several low energy structures with the same Si_6 core as in the global minimum structure, but the new H-shift mutation method now enables us to also find the preferred H atom locations. The Si_6H_6 global minimum structure 1(e) was found in 4 of the CGA runs after only 142, 459, 776 and 950 local optimization.

Finding the global minimum with fewer energy evaluations demonstrate that the four new mutation operators cause different possible Si_xH_y cluster structures to be explored than with the previous CGA. The new mutation operators only modify the structure of a few atoms in a cluster and this must better enable the good structural features already present in a cluster to be passed onto the next generation. The fact we are now able to find the global minima for the $Si_{14}H_{20}$ and Si_6H_6 clusters, along with being consistent with our previous searches for the $Si_{10}H_{16}$, $Si_{10}H_{14}$ and $Si_{14}H_{18}$, improves our overall confidence with the new CGA at being able to find the global minimum for larger Si_xH_y stoichiometries. To further test the new CGA we have applied it to the $Si_{18}H_{24}$ and $Si_{18}H_{22}$ clusters. We chose the $Si_{18}H_{24}$ cluster because it can adopt a global minimum corresponding to another diamond-lattice like structure shown in the figure 5.5. We performed five CGA runs utilizing 15000 local optimizations for both the $Si_{18}H_{24}$ and $Si_{18}H_{22}$ stoichiometries. We further locally optimized the low energy $Si_{18}H_{24}$ and $Si_{18}H_{22}$ clusters at the MP2 level. The 10 lowest MP2 energy $Si_{18}H_{24}$ and $Si_{18}H_{22}$ clusters are shown in the figure 5.3 and figure 5.4 respectively. It is worth noting that only one of the CGA runs, after 11280 local geometry optimizations, finds the $Si_{18}H_{24}$ diamond-lattice like structure 3(a) as the GAM1 and MP2 global minimum while the four other CGA runs find structures which are 0.01, 1.0, 1.5 and 3.2 kcal/mol higher in GAM1 energy. The structure 3(e) resembles the structure 3(a) however the repulsion between the two flagpole H atoms in structure 3(e) raises the cluster energy by around 4 kcal/mol. Figure 5.4 shows the MP2 global minimum 4(a) for $Si_{18}H_{22}$ which was obtained after 1599 local optimizations in one of the CGA runs. The $Si_{18}H_{22}$ MP2 global minimum is the second lowest GAM1 energy cluster. The GAM1 global minimum 4(g) located in the other four CGA runs has a 0.4 kcal/mol lower GAM1 energy but a 2.0 kcal/mol higher MP2 energy. The disagreement of the MP2 and

GAM1 predicted global minimum structures may be starting to indicate the limit of applicability range of the Si_7H_{14} GAM1 parameters in the global minimum search of larger Si_xH_y clusters. Again the removal of 2 H atoms from the completely passivated Si_{18} core produces a global minimum structure which is not simply related to the fully H passivated structure.

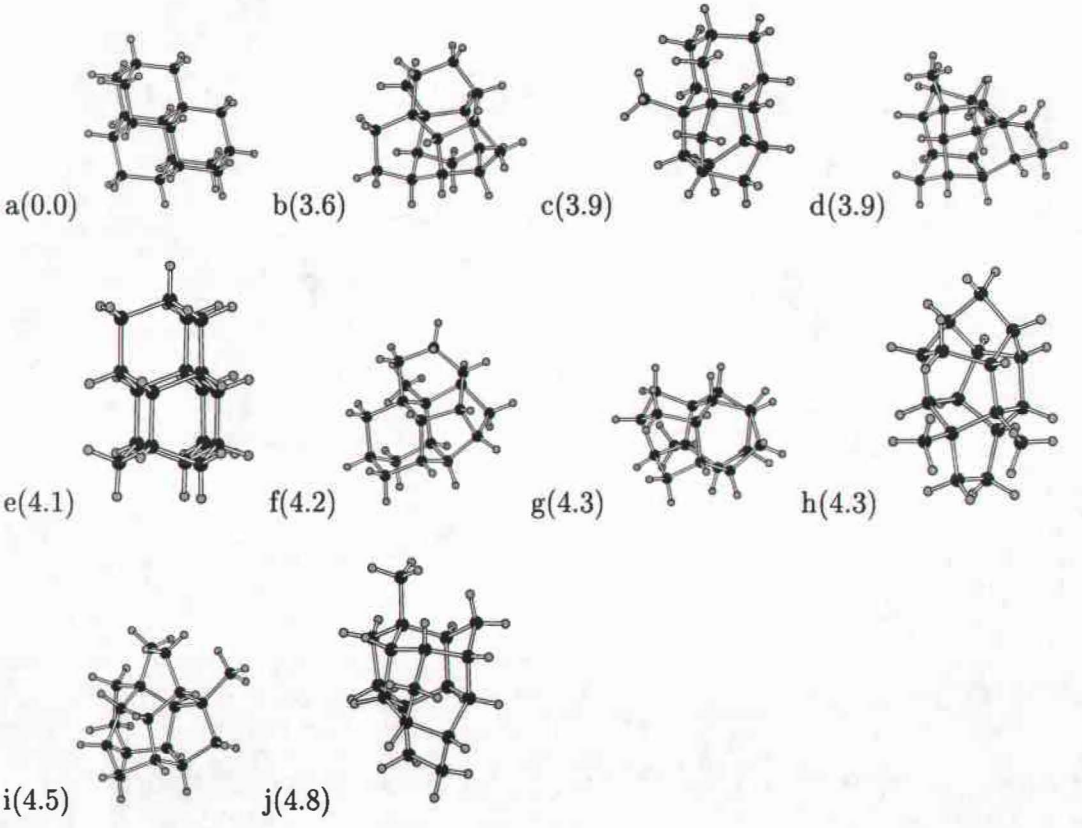


Figure 5.3: The MP2/lanl2dz(d) locally optimized geometries for the $Si_{18}H_{24}$ low energy clusters obtained from the CGA. In parentheses are the MP2 relative energies in kcal/mol.

In summary figure 5.5 shows that the slightly under H passivated $Si_{10}H_{14}$, $Si_{14}H_{18}$ and $Si_{18}H_{22}$ stoichiometries have global minima which do not show any obvious struc-

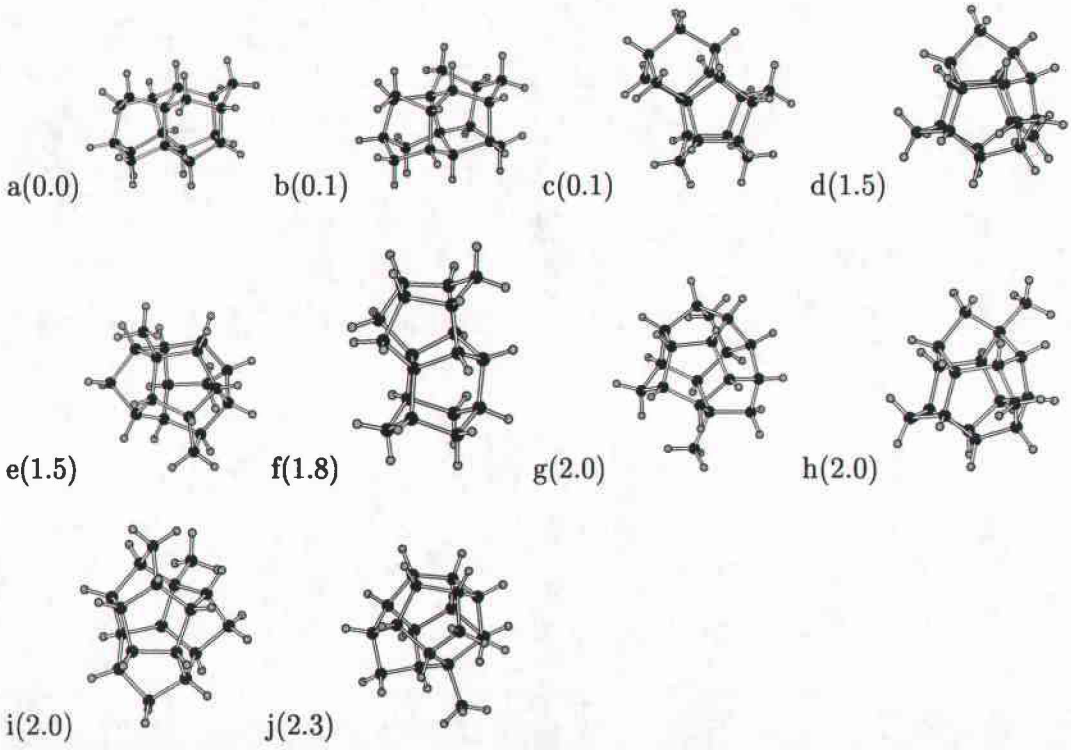


Figure 5.4: The MP2/lanl2dz(d) locally optimized geometries for the $Si_{18}H_{22}$ low energy clusters obtained from the CGA. In parentheses are the MP2 relative energies in kcal/mol.

tural similarities. The $Si_{18}H_{22}$ global minimum does appear to retain more of the diamond-lattice like structure than in the smaller $Si_{14}H_{18}$ and $Si_{10}H_{14}$ clusters. This suggests that in larger under H passivated Si clusters more of the bulk Si structure will be retained, but in the region of the incomplete H passivation there will be a structure corresponding essentially to a Si lattice defect. However, the lattice defect structure in Si_xH_{y-2} cannot be simply derived by removing two H atoms from the diamond-lattice like Si_xH_y global minimum and then performing a local geometry optimization.

5.4 Conclusion

We have developed a faster cluster genetic algorithm (CGA) for finding the global minimum of Si_xH_y clusters. The key components to the CGA are the genetic operators. We continue to use the cut-and-paste mating operator to produce an offspring from two parent clusters. We also introduce 4 new mutation operators: SiH_3 removal, SiH_2 removal, H shift and piece rotation. These new operators better retain good structural information between different generations in the CGA thereby enabling a global minimum structure to be found with fewer cluster energy evaluations.

In our previous CGA we had difficulty finding the global minima for $Si_{14}H_{20}$ and Si_6H_6. The SiH_3 and SiH_2 removal operators enable the new CGA to locate a diamond-lattice like structure which is the $Si_{14}H_{20}$ global minimum. For the low H passivated clusters, the H shift operator assists with generating the H atom arrangement present such as in the Si_6H_6 global minimum.

The improved rate of finding the global minimum was further demonstrated by using the new CGA to find the global minima for the $Si_{18}H_{24}$ and $Si_{18}H_{22}$ clusters. The global minima structures for $Si_{18}H_{24}$ and $Si_{18}H_{22}$ are consistent with our previous results that if the Si_xH_y cluster is fully H passivated then the Si_x core forms a diamond-

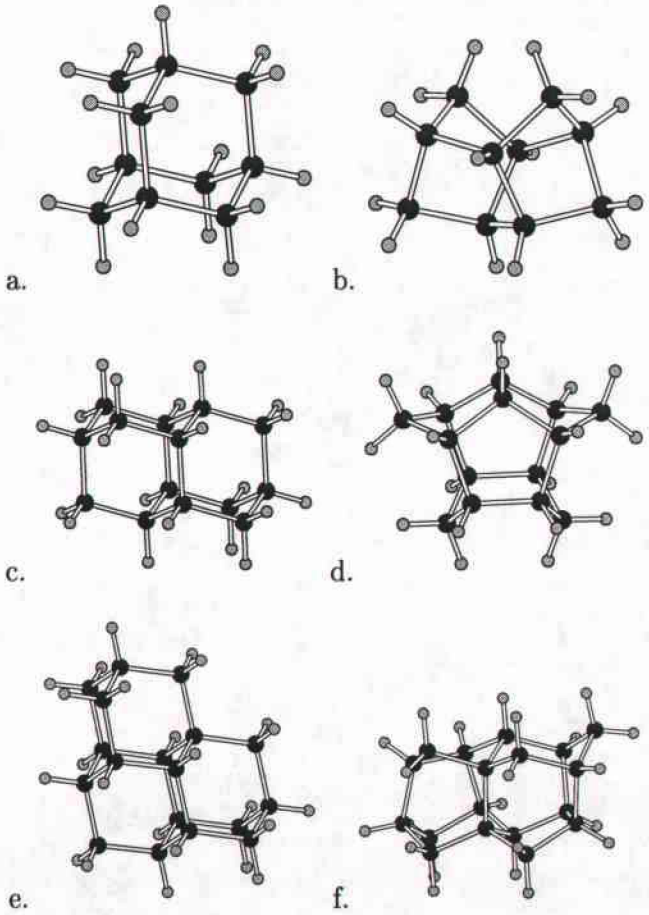


Figure 5.5: The Si_xH_y MP2/LanL2dz(d) global minima: a. $Si_{10}H_{16}$; b. $Si_{10}H_{14}$; c. $Si_{14}H_{20}$; d. $Si_{14}H_{18}$; e. $Si_{18}H_{24}$; f. $Si_{18}H_{22}$.

lattice like structure, while the Si_x core in Si_xH_{y-2} is significantly different from the diamond-lattice structure. Consequently, an important utility of the CGA is to show that there is no simple structural relationship between the global minima of clusters like $Si_{18}H_{24}$ and $Si_{18}H_{22}$.

CHAPTER 6. FAST GLOBAL OPTIMIZATION OF F PASSIVATED SI CLUSTERS AT THE DFT LEVEL

(A paper based on this chapter is submitted to *Mol. Phys.*)

6.1 Introduction

Nanometer sized silicon clusters have recently attracted considerable interest because they exhibit intense photoluminescence (PL) and have the potential of being developed into a practical optoelectronic device. The nanoparticles contrast with bulk silicon which only shows a low intensity PL owing to it being an indirect gap semiconductor. Quantum confinement [1, 6] and surface effect [10, 20, 21, 83] theories have been proposed to explain the mechanism of the PL in silicon clusters. The progress in methods development, through the work of the Handy group and others, means that molecular quantum mechanics can now be used to calculate the optoelectronic properties for relatively large atomic and molecular clusters. For instance, Zhou, Friesner and Brus (ZFB) recently used density functional theory (DFT) to calculate the electronic structure for 1 to 2 nanometer diameter silicon clusters as large as $Si_{87}H_{76}$ [83]. ZFB treated several different sized Si clusters passivated by either H, oxide, OH, hydrocarbon and F ligands. The calculations enable the characterization of single Si nanoparticles rather than the ensemble of particles with uncertain size ranges present in most real samples. In order to do these calculations, ZFB appear to make the chemically reasonable assumption that the most stable nanoclusters consist of a Si core with the same diamond-lattice like structure as found in bulk Si. ZFB then passivate dangling surface Si bonds of the diamond-lattice like core with a ligand of interest and perform a

full local optimization to get the geometry of specific clusters. Degoli *et al.* make a similar diamond-lattice like assumption in their ground and excited state calculations on clusters ranging from Si_5H_{12} to $Si_{35}H_{36}$ [163]. One of the reasons why quantum mechanical calculations are so useful in this research is because it is experimentally very difficult to determine the geometry of an individual cluster or even if a cluster has the anticipated stoichiometry. Unfortunately, a full local geometry optimization does not generally guarantee that the lowest possible energy structure and corresponding electronic properties for the cluster has been computed.

We have been developing strategies using a genetic algorithm (GA) for globally optimizing at the *ab initio* level Si_xH_y clusters [153–155, 164]. The search procedure entails performing many local geometry optimizations of different cluster structures and we use a reparametrized semiempirical AM1 [120] method as a fast energy prescreening method before using a relatively small number *ab initio* calculations to identify the global minimum [154,155]. Using a semiempirical method, rather than some of the other available empirical potentials, avoids having to explicitly specify bond types present in the clusters and automatically allows a Si_xH_y clusters to lose H_2 if it is overpassivated. By using this procedure we have demonstrated that the lowest energy structures for the clusters $Si_{10}H_{16}$, $Si_{14}H_{20}$ and $Si_{18}H_{24}$ do consist of a diamond-lattice like Si_x core fully passivated by H atoms as assumed by ZFB. However, our calculations also show that it should not be all that difficult to generate relatively stable under H passivated clusters $Si_{10}H_{14}$, $Si_{14}H_{18}$ and $Si_{18}H_{22}$ which, as figure 5.5 indicates, have very different Si_x core structures from the fully H passivated clusters [155, 164]. The Si_x cores in the under H passivated clusters could be regarded as a defect in the bulk Si lattice and these defects could actually be the origin of the bright PL in Si nanoparticles. The under H passivated Si clusters demonstrate why global optimization methods are so important:

their structures cannot be simply obtained by local optimization calculations initiated by simply removing two H atoms from the $Si_{10}H_{16}$, $Si_{14}H_{20}$ and $Si_{18}H_{24}$ global minima.

We are in the process of extending our global optimization strategy to Si_xL_y clusters with other ligands L. One of difficulties with going beyond H as the ligand L is that the semiempirical energy evaluation becomes much slower. For instance an AM1 calculation on Si_7F_{14} involves 84 valence basis functions which is double the number of basis functions required by a Si_7H_{14} calculation. To speed up the semiempirical calculations we have been investigating replacing the F ligand with a pseudo H like atom and use a Si_xF_y training set to get the F parameters. This idea can be very useful if the GAM1 energy rankings or even the relative GAM1 energies of the Si_xF_y clusters simulates those of the Si_xH_y clusters. Unfortunately we have found that the relative energies of the Si_7H_{14} and the corresponding Si_7F_{14} isomers differ dramatically. However we have also found that a very simple empirical energy correction function can be developed to correct the GAM1 relative energies of the Si_7H_{14} clusters to match the B3LYP relative energies of the Si_7F_{14} clusters. We applied this corrected GAM1 (GAM1*) energy scheme to perform the cluster global optimization for the Si_7F_{14} clusters and found 3 new low B3LYP energy Si_7F_{14} isomers other than the training compounds. We also have attempted to search the low energy Si_8F_{14} clusters to investigate the transferability of this GAM1* energy evaluation scheme. We then use the same GAM1* energy function to find the low B3LYP energy isomers for larger compounds such as $Si_{10}F_{16}$ and $Si_{10}F_{14}$. Our results indicates that the energy rankings of the low energy Si_xF_y structures are very different to those found for Si_xH_y.

The rest of the chapter is organized as follows. In the next section we outline the computational methods used. In the results and discussion section we compare the *ab initio* energies and structures for Si_2L_4, Si_7L_{14}, Si_8L_{14}, $Si_{10}L_{14}$ and $Si_{10}L_{16}$ clusters with

L = H and F. The final section of the chapter contains concluding remarks.

6.2 Computation Method

We used 14 Si_7H_{14} training compounds for the reparameterization of the GAM1 method to evaluate the energies of the well H-passivated Si clusters in our previous research [155, 164]. In this chapter we simply replace the H atoms with F atoms to form the starting geometries of the corresponding Si_7F_{14} isomers. We then locally optimized the above 14 Si_xF_y compounds using the B3LYP method with a LanL2dz basis set [147] with a d polarization function on each Si or F [23,148] atom and a p-type diffuse function on the more electronegative F atoms [148]. The B3LYP calculations produce reasonable relative energies for the different Si_xH_y and Si_xF_y isomers compared with the much more expensive CCSD(T) calculation as illustrated by the Si_2H_4 and Si_2F_4 examples in the next section.

Since we have developed a GAM1 method that evaluates the energies reasonably for the Si hydride and successfully obtained the global minimum and other low energy Si_xH_y clusters [164], one intuitive thought would be to simply replace the H ligands in low energy Si_xH_y clusters by F atoms to generate the corresponding low energy Si_xF_y clusters. However, this thought can be useful only if the Si_xF_y isomers have very similar relative energies and energy rankings as the corresponding Si_xH_y isomers with similar Si cores. Unfortunately we have found that the relative GAM1 energies of the Si_7H_{14} training compounds differ dramatically with the relative B3LYP energies of the 14 corresponding Si_7F_{14} isomers. However, after carefully examining the B3LYP energies of the 14 Si_7F_{14} isomers and the GAM1 energies of the corresponding Si_7H_{14} isomers we realized that a simple empirical function can be designed to correct the GAM1 energies of the Si_7H_{14} training compounds to match well with the relative B3LYP energies of

the Si_7F_{14} clusters:

$$E^{\mathrm{GAM1}^*} = C_0 \times E^{\mathrm{GAM1}} + C_1 \times N_{\mathrm{H_3Si-}} + C_2 \times N_{\mathrm{(SiH_3)_2Si=}} + C_3 \times N_{\mathrm{H_2}} \qquad (6.1)$$

where E^{GAM1^*} is the corrected GAM1 energies to match with the B3LYP energies of the Si_7F_{14} clusters. E^{GAM1} is the GAM1 energies of a Si_7H_{14} cluster with the same Si framework as the corresponding Si_7F_{14} training compound. $N_{\mathrm{H_3Si-}}$ is the number of SiH_3 groups of the Si_xH_y cluster. $N_{\mathrm{(SiH_3)_2Si=}}$ is the number of Si atoms which attach with at least 2 SiH_3 groups. $N_{\mathrm{H_2}}$ is the number of H_2 molecules coming off from the Si_xH_y clusters. We consider a Si-H or a H-H single bond being formed if their distance is within 120% of the usual Si-H or H-H distance. A H_3Si- is identified where a Si atom has bonds with three H atoms. A $(SiH_3)_2Si =$ is identified where a Si atom has bonds with at least two Si atoms that are part of a H_3Si- group. C_{0-3} are the coefficients which are to be optimized to match the GAM1* relative energies with the B3LYP relative energies for the Si_xF_y clusters.

By using the above empirical equation, we simplify the Si_xF_y cluster global optimization problem into a much easier Si_xH_y cluster global optimization problem. Unlike our previous CGA for the global optimization of the Si_xH_y clusters, we now use the GAM1* energy evaluation scheme to evaluate the Si_xH_y cluster energy in the cluster GA. The GAM1* energy evaluation scheme is actually based on the previous GAM1 method we developed in the chapter 4 [164] but with some extra empirical correction terms. After we obtain the Si_xH_y clusters with low GAM1* energies by the CGA, we then generate Si_xF_y cluster by repacing H atoms in the above Si_xH_y clusters with F atoms. We further locally optimize these Si_xF_y clusters using the B3LYP method to

provide more accurate energy rankings for the Si_xF_y clusters.

The Si_xL_y cluster global optimizations are performed with our recently improved cluster genetic algorithm (CGA) [164]. The Deaven and Ho cut-and-paste mating method [102] along with three other mutation operators are applied to generate offspring from parents. The cut-and-paste method cuts two parent clusters into halves along randomly orientated planes and recombines two halves from different parents into a offspring structure. The first mutation operation removes a SiL_3 group and replaces it with a single L atom and inserts a SiL_2 group between two randomly selected neighboring Si atoms to produce either a longer Si chain or an expanded Si ring. The second mutation operation is just the reverse procedure of the first one: it removes a SiL_2 group from the cluster and then a randomly chosen L atom is replaced by a SiL_3 group. To reduce making a drastic geometry change the two Si atoms which were neighboring the SiL_2 group are pulled slightly towards each other to induce the formation of a new Si-Si bond. The last operator, a piece-rotation mutation method, is a single-parent cut-and-paste operator and is similar to the operator used by Rata *et. al.* [46] for the global optimization of Si-only clusters. The Si_xL_y clusters is randomly cut into two parts and then one half of the cluster is randomly rotated relative to the other half. The above 4 genetic operators are applied with equal probability.

The global optimization starts from 120 randomly generated locally optimized structures as the ancestor population. The fitness value of an individual cluster is taken as the GAM1 semiempirical energy of the locally optimized cluster. Tournament selection is used to select the parents needed to produce offspring. Since the calculations are performed on either a 16 or a 32 processor computer we do not explicitly treat separate generations. The CGA performs parent selection, mating/mutation process and generates a new trial offspring structure as soon as the local geometry optimization

of a cluster included in the parent population is finished on a processor. To avoid identical structures dominating the population new offspring which already exist in the current population are discarded. The final set of clusters with lowest GAM1 energies are locally optimized with *ab initio* calculations.

All of the prescreening GAM1 semiempirical calculations were performed using the GAMESS program [149]. Since we only examine well passivated Si_xL_y clusters we use the GAM1 semiempirical parameters obtained previously from the Si_7H_{14} training set [155] for the Si_xH_y clusters. We have also derived the GAM1* energy evaluation scheme for the Si_xF_y clusters using a pseudo H-like atom and a Si_7F_{14} training set which is described above. Since we are searching for accurate low energy clusters of fairly large Si_xL_y clusters *ab initio* calculations were performed with the DFT method using the B3LYP functional [133–135] and the Gaussian package [150]. We used the LanL2dz(d) scheme consisting of the Hay-Wadt effective core potential (ECP) and valence basis set for Si [147] augmented with a d function ($\zeta_d = 0.45$) [23] and the Dunning-Hay basis for H and F [148]. Because of the strong F electronegativity the F basis set was augmented with a p diffuse function ($\zeta_p = 0.074$) and a d polarization function ($\zeta_d = 0.90$) [148]. The accuracy of the B3LYP energy ranking for some of the smaller clusters was checked against MP2 [132] calculations performed using the GAMESS program.

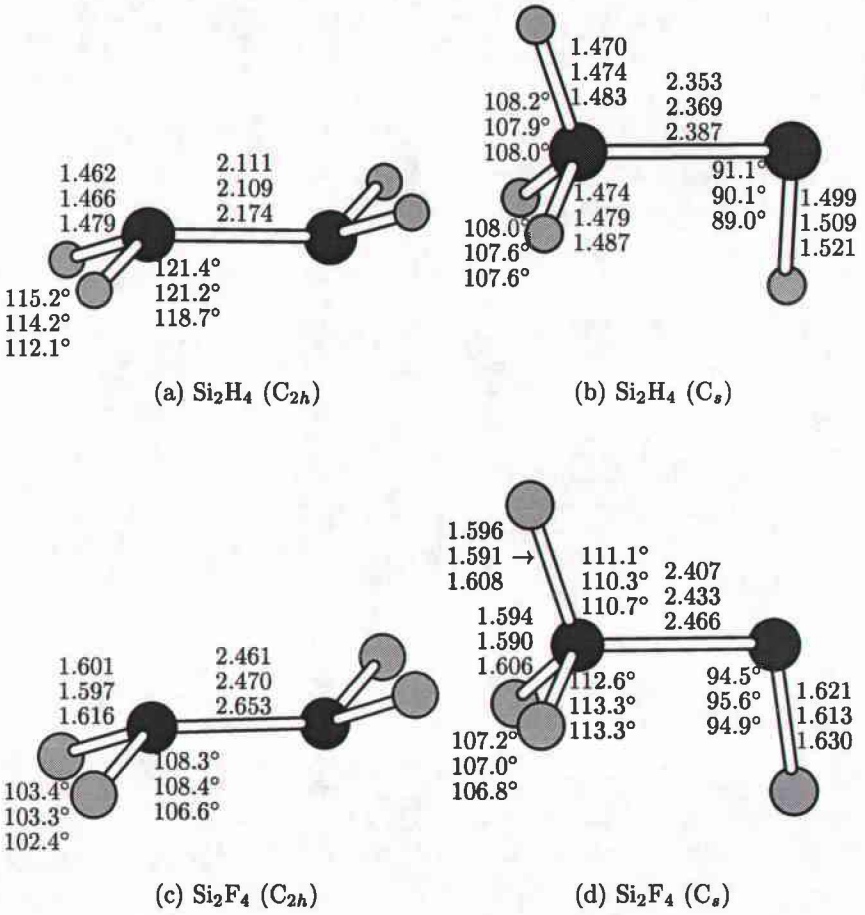
6.3 Results and Discussion

Si_2L_4. Only two Si_2L_4 local minimum types are possible when the ligand L binds to just one of the Si atoms as shown in figure 6.1. The disilene $L_2Si{=}SiL_2$ isomer is not a simple analogue of ethylene and instead favors forming a trans-bent structure with C_{2h} symmetry. However, even with this relatively small molecule, the preferred lowest energy isomer varies with the different H or F ligand. Recent work by Sari *et al.* provides

an extensive review of past work on the different Si_2H_4 isomers and they use highly correlated *ab initio* coupled cluster theories [69] to predict that the silene $H_2Si{=}SiH_2$ is 6 kcal/mol more stable than $HSi\text{-}SiH_3$. Table 6.1 shows that the our calculations on Si_2H_4 using LanL2dz(d) MP2, B3LYP and CCSD(T) give results which are reasonably consistent with the Sari *et al.* results. Fewer calculations appear to have previously been performed on the Si_2F_4 species. Li *et al.* used DFT methods with various functionals to investigate structures, thermochemistry and electron affinities for the Si_2F_n $(n = 1-6)$ series [93]. Consistent with the Li *et al.* work, our calculations which are summarized in table 6.1 find the fluorinated disilene (C_{2h}) is now less stable than $FSi\text{-}SiF_3$ (C_s) by 6-12 kcal/mol depending on the method used. Si_2F_4 represents the upper limit to the system size we can treat using highly correlated *ab initio* methods with our available computing resources and table 1 demonstrates that the LanL2dz(d)/B3LYP calculations should give adequate results for the larger clusters described in the rest of the chapter. The important observation from these Si_2L_4 calculations is that the F ligand appears to favor forming a trifluorosilyl, SiF_3, group, whereas the H ligand prefers to be evenly distributed on the two Si atoms. This F preference for forming SiF_3 groups while H favors being more evenly dispersed on all of the available Si atoms turns out to be a common theme in the global minima structures described below for the larger Si_xL_y clusters.

Table 6.1: Relative energies for the two Si_2F_4 and Si_2H_4 isomers with C_{2h} and C_s symmetry shown in figure 6.1. Listed are the MP2 and B3LYP optimized energies for various basis sets and the CCSD(T) single point energy computed with the B3LYP optimized geometry. All energies are in kcal/mol.

	LanL2dz(d)			6-311+G(d)			6-311+G(3df)		
Isomers	MP2	B3LYP	CCSD(T)	MP2	B3LYP	CCSD(T)	MP2	B3LYP	CCSD(T)
F_2SiSiF_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$FSiSiF_3$	-12.7	-8.1	-10.2	-11.0	-6.4	-9.3	-11.5	-6.7	N.A.
H_2SiSiH_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$HSiSiH_3$	12.1	10.5	11.9	8.5	5.8	8.8	9.6	6.2	8.5



(a) Si_2H_4 (C_{2h})

(b) Si_2H_4 (C_s)

(c) Si_2F_4 (C_{2h})

(d) Si_2F_4 (C_s)

Figure 6.1: The locally optimized Si_2H_4 and Si_2F_4 isomers. The Si_2H_4 geometry information are listed from top to bottom in the order: MP2/LanL2dz(d), B3LYP/LanL2dz(d) and CCSD(T)/cc-pVTZ [69] methods. The Si_2F_4 geometry information is in the order: MP2/LanL2dz(d), B3LYP/LanL2dz(d) and B3LYP/DZP++(d) [93] methods.

Si_7L_{14}. Our original motive for examining the relative energies of different Si_7H_{14} clusters was to develop a training set which could be used in a parametrization genetic algorithm to obtain improved AM1 parameters, which we designated as GAM1 parameters, for the Si atoms in fully passivated Si_xH_y clusters [155]. Si_7H_{14} is large enough that it can adopt various configurations, ranging from linear through three- to seven-membered Si ring structures, which represent common partial structures found in a well passivated Si_xH_y cluster. In figure 6.2 we show the locally optimized structures and relative energies for the Si_7F_{14} cluster obtained after replacing all the H atoms with F in our previous set of 14 Si_7H_{14} training compounds. To illustrate the change in relative energies of the silicon fluorides we still order the Si_7F_{14} structures in figure 6.2 in the original energy ranking order we obtained for Si_7H_{14} with MP2/6-31G* calculations. Structures 2(a) through 2(g) contain either five- or six-membered Si rings which for Si_7H_{14} have an energy spread of 2 kcal/mol. Previously we found with MP2/6-31G* calculations the five-membered Si ring structure 2(a) to be the global minimum and 0.8 and 1.3 kcal/mol more stable than structures 2(b) and 2(c) [155]. There is clearly still a small energy difference between the five- and six-membered Si ring Si_7H_{14} structures, but it is reversed, when B3LYP/LanL2dz(d) and MP2/LanL2dz(d) calculations are performed. Consistent with an increase in strain with reduced ring size, the Si_7H_{14} relative *ab initio* energies for the 2(l) and 2(m) structures with four- and three-membered Si rings present are around 10 and 23 kcal/mol, respectively, higher in energy. In strong contrast for Si_7F_{14} the five-membered Si ring structure 2(a) is approximately 20 kcal/mol lower in energy than the two structures 2(b) and 2(c) containing a six-membered Si ring. Figure 6.2 also shows significant changes in the energy rankings for many of the other Si_7F_{14} isomers. Due to the much larger Si-F bond energy relative to the F-F and Si-Si bond energies, structure 2(k) shows that the Si clusters should

favor high passivation by F atoms. We have used the Si_7F_{14} B3LYP energies to obtain the C_{0-3} GAM1* parameters for Si_xF_y clusters. The values of C_{0-3} are determined to be 0.842511, -0.009020 a.u., -0.016175 a.u., and 0.304159 a.u. respectively. We then compare the linear relationships between the B3LYP relative energies of the 14 Si_7F_{14} training compounds and the GAM1(*) relative energies of the corresponding Si_7H_{14} clusters. Chart 3(a) in the figure 6.3 shows that the relative GAM1 energies of the Si_7H_{14} does not match with the B3LYP relative energies of the Si_7F_{14} clusters at all, while the chart 3(b) shows that the corrected GAM1* relative energies match with the B3LYP relative energies of the Si_7F_{14} clusters nearly perfectly.

The GAM1* energy scheme is then used in the CGA to search for the Si_7F_{14} global minimum and other low energy structures which ultimately should be included in the Si_7F_{14} training set. In figure 6.4 we show three new Si_7F_{14} structures found by the CGA with even lower B3LYP energies than those obtained for the original training set clusters. Structure 4(a) is the Si_7F_{14} B3LYP/LanL2dz(d) global minimum and in figure 6.4 we now take structure 4(a) energy as the zero reference. Structures 4(a) and 4(b) consist of a four-membered Si ring with three trifluorosilyl groups attached, while 4(c) at only slightly higher B3LYP energy but a lower MP2 energy contains a three-membered Si ring with four SiF_3 groups attached. We include in figure 6.4 the five-membered Si ring global minimum 4(u^H) and the lowest energy six-membered ring 4(v) structures for Si_7H_{14}. Remarkably, in the local geometry optimizations we have found F to only slightly perturb the Si_7 frameworks from their original Si_7H_{14} positions. The only major exception to this is found for the low energy non-planar five-membered ring Si_7H_{14}, structure 4(u^H) in figure 6.4, which becomes planar in Si_7F_{14}. The energies of the Si_7H_{14} structures 2(l) and 2(m) suggest there must be a large ring strain present in the four- and three-membered Si_7F_{14} rings but this must now be compensated by

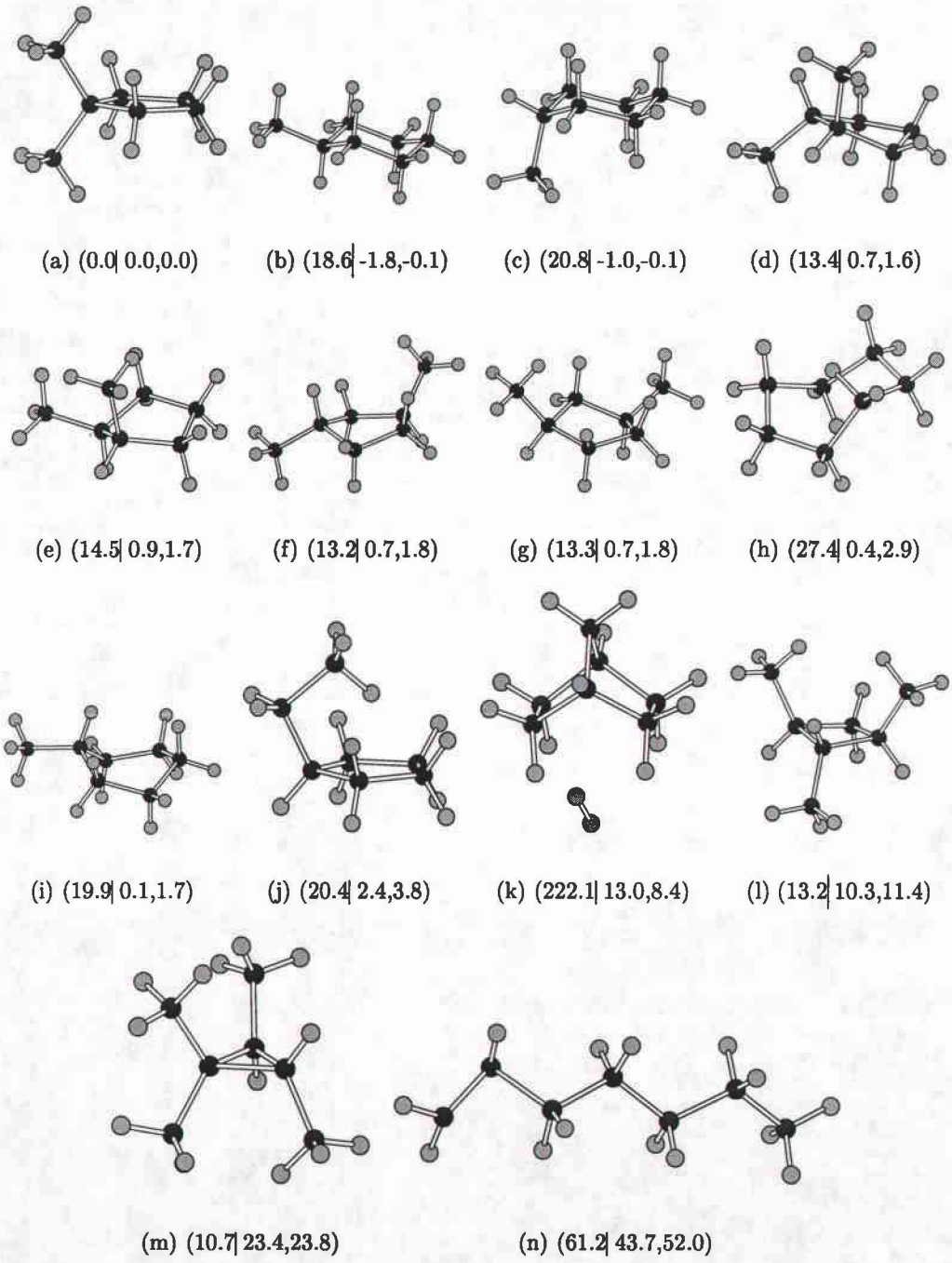
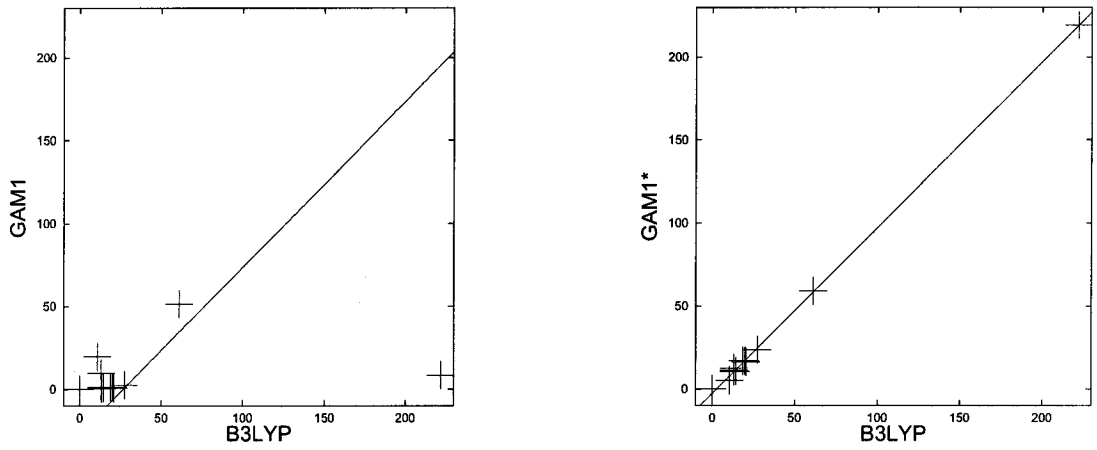


Figure 6.2: The Si_7F_{14} B3LYP locally optimized geometries for the original 14 training set compounds. In parentheses: (Si_7F_{14} B3LYP relative energy | corresponding Si_7H_{14} B3LYP and MP2 relative energies). All energies are in kcal/mol.



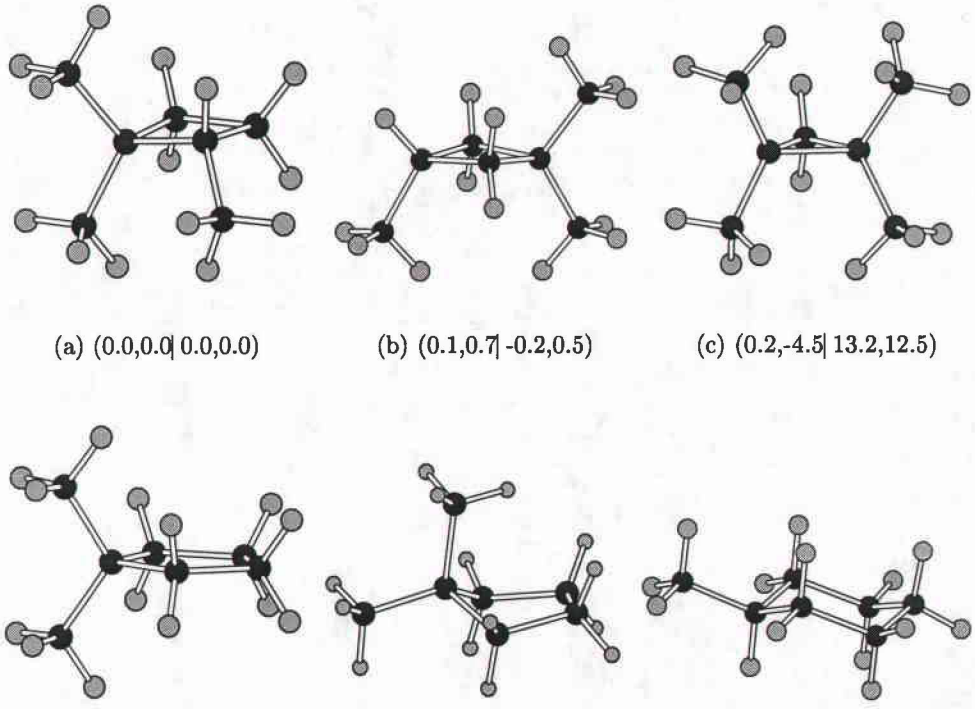
(a) (b)

Figure 6.3: The left chart (a) shows the GAM1 relative energies of Si_7H_{14} clusters vs the B3LYP relative energies of the 14 Si_7F_{14} training compounds, while the right chart (b) shows the corrected GAM1* relative energies of Si_7H_{14} clusters vs the B3LYP relative energies of Si_7F_{14} clusters. The lines in the charts all have a slope of 1 but with various interceptions to best match with the data.

forming the three or four trifluorosilyl groups. In general figures 6.2 and 6.4 show that the Si_7F_{14} structures with the greater number of SiF_3 groups have the lower energies. For the series of Si_7F_{14} structures containing five-membered rings we find the relative energy ordering: two SiF_3 groups attached to same Si ring atom 2(a) < two SiF_3 groups attached to different Si ring atoms 2(d-g) < Si_2F_5 group attached to single Si ring atom 2(j). This type of analysis can be used to rationalize the relative energies of the other structures. To examine the reliability and predictability of the new GAM1* energy evaluation scheme, we have investigated the linear relationships between the B3LYP relative energies and the corresponding GAM1(*) relative energies for the 14 training compounds plus 12 more isomers we obtained from the CGA. The left charts in the

figure 6.5 are the GAM1 relative energies of Si_7H_{14} clusters versus the B3LYP relative energies of Si_7F_{14} clusters, while the right charts are the GAM1* relative energies of Si_7H_{14} clusters versus the B3LYP relative energies of Si_7F_{14} clusters. The charts 5(c) and 5(d). are the amplified parts of the charts 5(a) and 5(b) respectively after the 2 rightmost data with the highest B3LYP energies are excluded. These four charts clearly show that the GAM1* relative energies match with the B3LYP relative energies much better than the GAM1 relative energies.

Si_8L_{14}. As an initial test of the transferability of the GAM1* parameters from the Si_7F_{14} training set to larger clusters with the same degree of F passivation we applied the CGA to find the global minimum for Si_8F_{14}. We also determined the Si_8H_{14} global minimum using separate CGA runs and the GAM1 Si parameters from the Si_7H_{14} training set. In figure 6.6 we show the 8 Si_8F_{14} structures with the lowest B3LYP energies which are within 10 kcal/mol of the global minimum structure 6(a) which we take as the zero energy reference. Included in figure 6.6 are two higher energy Si_8F_{14} structures 6(u) and 6(v) which we find as the Si_8H_{14} B3LYP and MP2 global minimum and the next lowest energy structures, respectively. Structure 6(b) is the second lowest Si_8F_{14} and the third lowest Si_8H_{14} B3LYP energy structure. Again we observe that the Si_8H_{14} and Si_8F_{14} global minimum structures are very different to each other. However, the H and F ligands cause structural changes similar to those we have already found in Si_2L_4 and Si_7L_{14}. The H atoms in the low energy Si_8H_{14} are evenly spread over all the Si atoms and five- and six-membered Si rings with low strain are formed. Structure 6(b) can be derived from 6(v) by moving a SiL_2 group from the six-membered Si ring to produce the SiL_3 group and the five-membered ring in 6(b). The low energy Si_8F_{14} structures have up to three SiF_3 side groups and as a consequence smaller three- or four-membered Si rings are present in these structures. The Si_8F_{14} global minimum



(a) (0.0,0.0| 0.0,0.0) (b) (0.1,0.7| -0.2,0.5) (c) (0.2,-4.5| 13.2,12.5)

(u) (0.5,1.7| -10.0,-10.0) $(u)^H$ Si_7H_{14} isomer (v) (19.1,25.2| -11.8,-10.0)

Figure 6.4: The B3LYP locally optimized geometries of the 3 low energy Si_7F_{14} clusters obtained by the CGA using the GAM1* parameters. (u) and (v) are Si_7F_{14} clusters containing five- and six-membered Si rings, respectively; (u^H) is the Si_7H_{14} global minimum. In parentheses: (Si_7F_{14} B3LYP and MP2 relative energies | Si_7H_{14} B3LYP and MP2 relative energies). All energies are in kcal/mol.

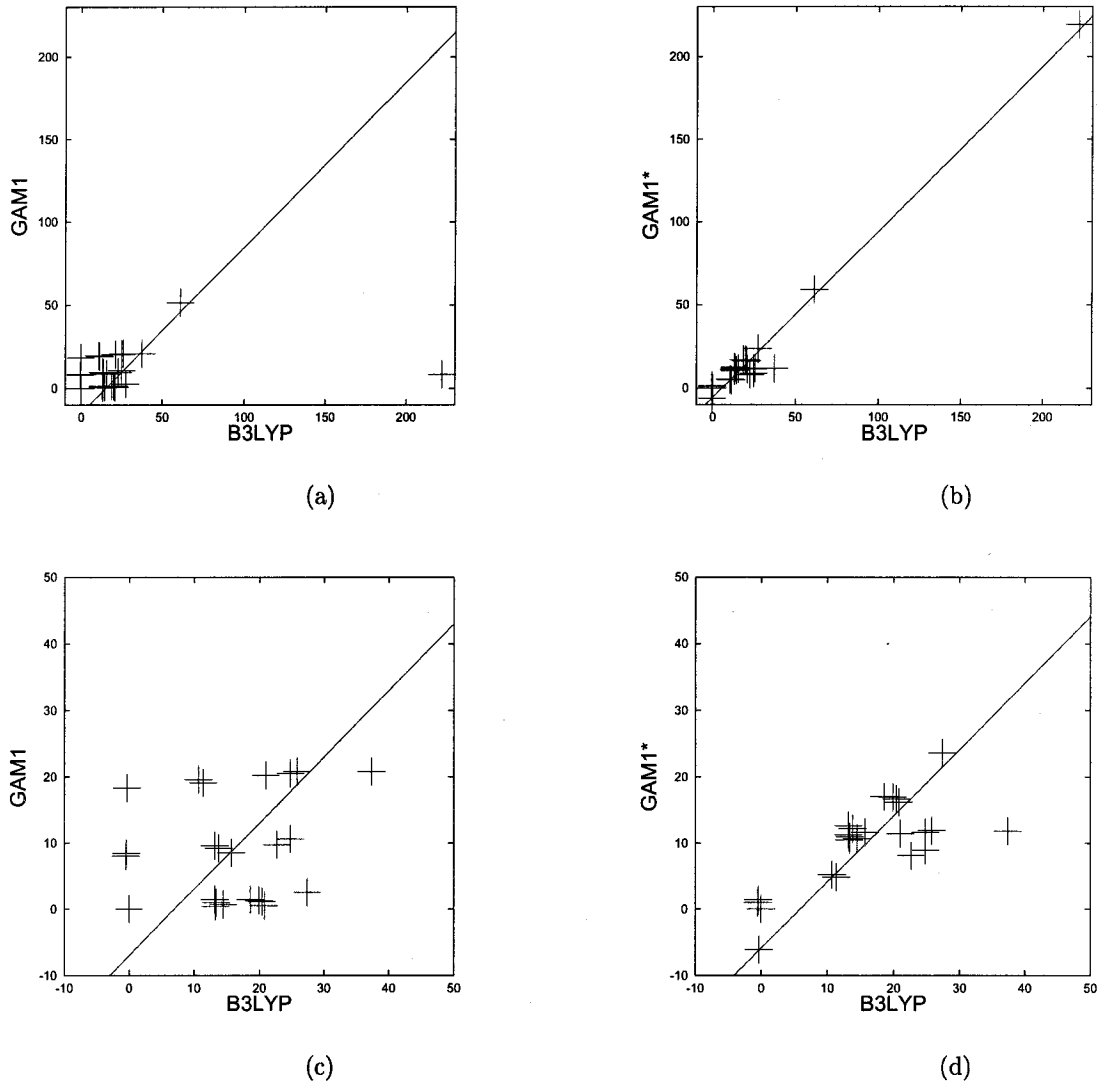


Figure 6.5: The left charts shows the GAM1 relative energies of Si_7H_{14} clusters versus the B3LYP relative energies of Si_7F_{14} clusters, while the right charts shows the GAM1* relative energies of Si_7H_{14} clusters versus the B3LYP relative energies of Si_7F_{14} clusters. The charts (a) and (b) are for the 14 Si_7F_{14} training compounds plus 12 more data for the low energy clusters predicted by the CGA coupled with the GAM1* energy function. The charts (c) and (d) are the amplified parts of the charts (a) and (b) after the 2 rightmost data with the highest B3LYP energies are excluded. The lines in the charts all have a slope of 1 but with various interceptions to best match with the data.

6(a) shows again that trifluorosilyl groups attached to the same Si atom is energetically favored despite the strain caused by the presence of four-membered Si rings. In contrast, structure 6(b) is formed from five- and six-membered Si rings which should have only a small amount of ring strain and as a consequence the single SiF_3 group is able to lower its energy relative to the other structures containing more SiF_3 groups. A similar analysis can be performed for structures 6(d) and 6(e). Structure 6(d) contains the strained three- and four-membered Si rings but this is compensated by having two SiF_3 attached to one Si atom and another SiF_3 group bonded to a different Si atom. Whereas structure 6(e) should have less ring strain, but it's relative energy is only stabilized the presence of two SiF_3 groups.

$Si_{10}L_{16}$. This is the smallest cluster where the Si_{10} core can adopt the bulk Si diamond-lattice like structure. Each Si atom in the $Si_{10}L_{16}$ diamond-lattice like structure is tetrahedrally coordinated by other Si or L atoms and the six-membered Si rings should have little, if any, Si ring strain. As assumed by ZFB and Degoli *et al.* and others, it is very chemically reasonable to assume that the global minimum for the $Si_{10}L_{16}$ is a diamond-lattice like structure. We have now performed an extensive series of global optimization studies on the $Si_{10}H_{16}$ stoichiometry, and computed GAM1, MP2 and B3LYP locally optimized geometries for a wide variety of different $Si_{10}H_{16}$ structures and conclude that the diamond-lattice like structure 7(u), in figure 6.7, is the *ab initio* global minimum [153,155,164]. The fact that our current CGA strategy [155,164] is able to find diamond-lattice like structures for $Si_{14}H_{20}$ and $Si_{18}H_{24}$ as well as $Si_{10}H_{16}$ enhances our confidence with the CGA method at being able to find the correct global minimum for other stoichiometries which do not have a diamond-lattice like structure as the global minimum. In figure 6.7 we show the B3LYP global minimum, structure 7(a), for $Si_{10}F_{16}$ using the CGA and the GAM1* parameters for energy prescreening.

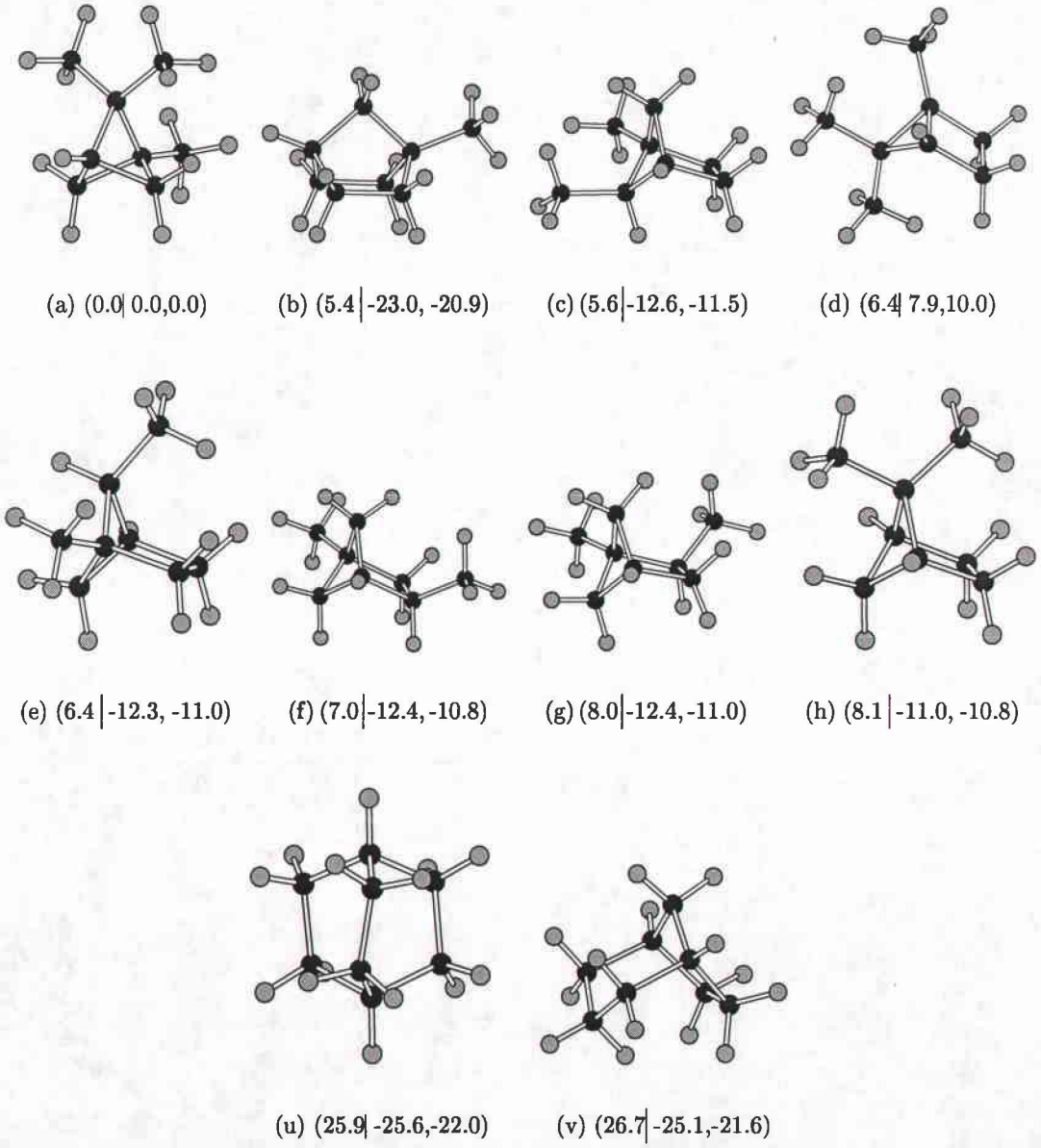
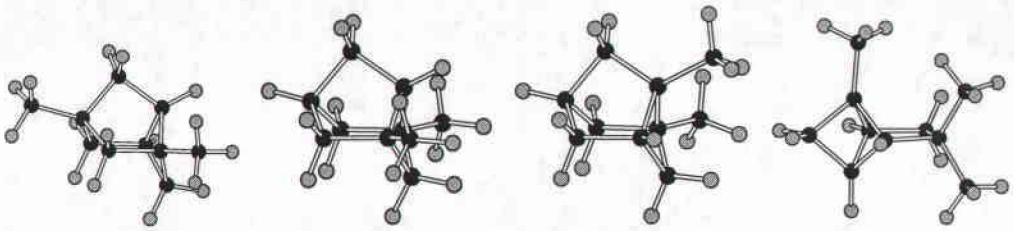


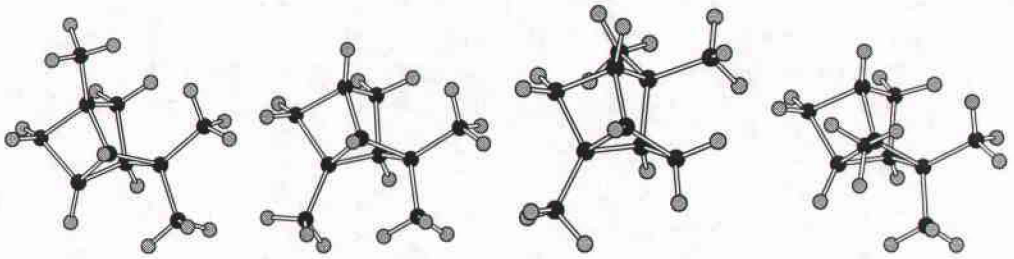
Figure 6.6: The B3LYP locally optimized geometries for the Si_8F_{14} low energy clusters located by the CGA. Structures (u) and (v) correspond to the lowest energy Si_8H_{14} clusters. In parentheses: (Si_8F_{14} B3LYP relative energy | Si_8H_{14} B3LYP and MP2 relative energies). All energies are in kcal/mol.

Also included in figure 6.7 are the different locally optimized B3LYP $Si_{10}F_{16}$ structures within 10 kcal/mol of the global minimum found by the CGA procedure. Once again $Si_{10}H_{16}$ and $Si_{10}F_{16}$ have vastly different low energy structures. The H atoms in $Si_{10}H_{16}$ prefer to be spread evenly over the Si_{10} core, while the F atoms favor forming trifluorosilyl groups in the low energy $Si_{10}F_{16}$ structures. As a consequence the fluorinated Si_{10} core contains some four-membered Si rings which are stabilized by the one or two SiF_3 groups attached to the ring Si atoms. For both the H and F ligands there are two series of structures 7(a-c) and 7(e-i,k) which have a narrow energy range presumably because the members of each series contain a similar Si_8 core and two SiL_3 groups attached at different Si core positions. It is interesting that in all of the low energy $Si_{10}F_{16}$ structures shown in figure 6.7 each of the SiF_3 groups bond with a core Si atom which is in turn bonded to three other Si core atoms. Structures 7(a) to 7(l) were all taken from the $Si_{10}F_{16}$ CGA and the $Si_{10}H_{16}$ optimized energies were obtained by replacing all F atoms by H atoms. When we examine the positions of the SiH_3 groups found in low energy structures obtained by the $Si_{10}H_{16}$ CGA runs we do find examples of the silyl group bonding to Si core atoms which are also bonded to other H atoms and not just three other Si core atoms [155]. Obviously, a Si core atom already bonded to a strongly electronegative F atom is going to have less electrons available than a Si core atom only bonding to other Si core atoms. When one takes into consideration the low energy structures found for Si_2F_4, Si_7F_{14} and Si_8F_{14} and the energy stability gained by forming a SiF_3 group, the 7(a) global minimum structure for $Si_{10}F_{16}$ does become chemically reasonable even though it is very different from the diamond-lattice like $Si_{10}H_{16}$ global minimum.

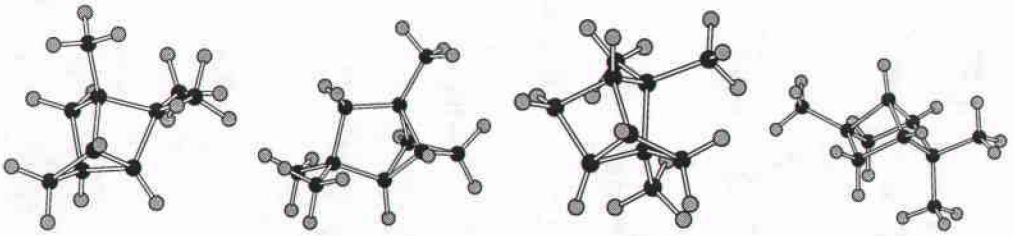
$Si_{10}L_{14}$. One important decision when performing calculations on this passivated Si clusters is deciding how many ligands are needed to adequately passivate the Si



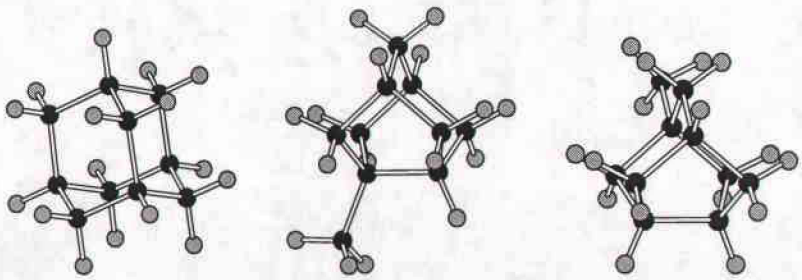
(a) (0.0| 0.0,0.0) (b) (1.6| -0.1,0.1) (c) (1.7| -0.4,-1.0) (d)(4.6|14.4,13.3)



(e)(5.4 | 15.5, 14.9) (f) (5.5 | 15.7, 15.1) (g) (5.9 | 15.7, 14.4) (h) (7.1 | 15.1, 14.0)

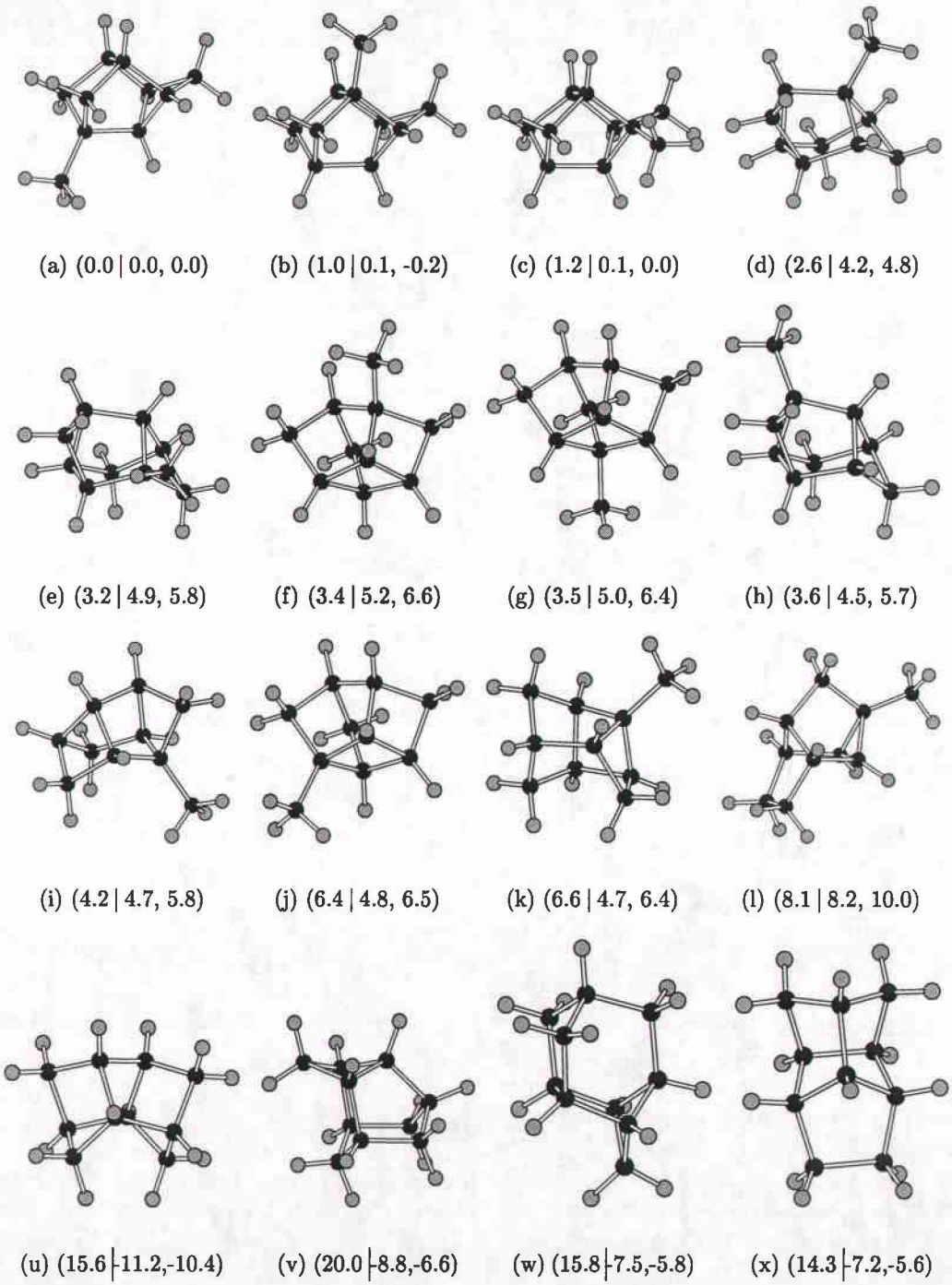


(i) (7.4| 15.6,14.7) (j) (8.1| 14.7,13.5) (k) (8.5 | 16.3, 14.7) (l) (9.6| 20.6,22.5)



(u) (35.4| -16.5,-13.8) (v) (15.7| -11.1,-10.6) (w) (15.4| -10.9,-10.2)

Figure 6.7: The $Si_{10}F_{16}$ low energy B3LYP locally optimized geometries located by the CGA on $Si_{10}F_{16}$. Structure (u) is the global minimum and (v) and (w) are low energy structures for $Si_{10}H_{16}$. In parentheses: ($Si_{10}F_{16}$ B3LYP relative energy | $Si_{10}H_{16}$ B3LYP and MP2 relative energies). All energies are in kcal/mol.



(a) (0.0 | 0.0, 0.0) (b) (1.0 | 0.1, -0.2) (c) (1.2 | 0.1, 0.0) (d) (2.6 | 4.2, 4.8)

(e) (3.2 | 4.9, 5.8) (f) (3.4 | 5.2, 6.6) (g) (3.5 | 5.0, 6.4) (h) (3.6 | 4.5, 5.7)

(i) (4.2 | 4.7, 5.8) (j) (6.4 | 4.8, 6.5) (k) (6.6 | 4.7, 6.4) (l) (8.1 | 8.2, 10.0)

(u) (15.6 | -11.2,-10.4) (v) (20.0 | -8.8,-6.6) (w) (15.8 | -7.5,-5.8) (x) (14.3 | -7.2,-5.6)

Figure 6.8: The $Si_{10}F_{14}$ low energy B3LYP locally optimized geometries located by the CGA. Structures (u-x) are the $Si_{10}H_{14}$ global minimum and low energy structures. In parentheses: ($Si_{10}F_{14}$ B3LYP relative energy | $Si_{10}H_{14}$ B3LYP and MP2 relative energies). All energies are in kcal/mol.

core. In the introduction we noted that one of the reasons for using the AM1 based semiempirical method is that it automatically enables L_2 to be lost from the cluster if there is an excess of the passivating ligands. In our previous work we have examined the energetics of $Si_{10}H_{16}$ losing H_2 and have concluded this may be only favored at high temperatures [153, 155]. To determine the energy of the H_2 loss reaction we have determined the $Si_{10}H_{14}$ global minimum. In figure 6.8 we show the low B3LYP energy structures obtained in the global optimization of $Si_{10}H_{14}$ and $Si_{10}F_{14}$ using the CGA. We have previously noted that the $Si_{10}H_{14}$ global minimum 8(u) cannot be obtained by simply removing two H atoms from the $Si_{10}H_{16}$ global minimum and then performing a local geometry optimization. To maintain an even distribution of H atoms the low energy $Si_{10}H_{14}$ clusters compensates for the reduced H passivation by forming structures containing some five- and even a four-membered Si rings. We find structure 8(a) as the $Si_{10}F_{14}$ global minimum. It appears to adjust to the lower F passivation by only slightly reducing the Si ring sizes relative to $Si_{10}F_{16}$ and instead 8(a) now contains a single trifluorosilyl group. In figure 6.8 we show the $Si_{10}F_{14}$ structures which are within 10 kcal/mol of the global minimum and do not find any structures containing two trifluorosilyl groups presumably because these structures would have appreciably higher Si ring strain. Analogous to what we found for $Si_{10}L_{16}$, we find three series of structures 8(a-c), 8(d,e,h) and 8(f,g,j) with narrow energy ranges, where each series consist of the same basic Si_9 core framework but with the SiL_3 group bound at different Si core atom positions. We find the Si core atoms which bond with SiF_3 to be coordinated by three other Si atoms. Overall, because of the much weaker F-F and Si-Si bond energies relative to the Si-F bond we do not anticipate that $Si_{10}F_{16}$ will readily lose F_2.

6.4 Conclusion

This study demonstrates that the global minima structures for Si_xH_y and Si_xF_y clusters are very different to each other. We have determined the global minima for Si_2L_4, Si_7L_{14}, Si_8L_{14}, $Si_{10}L_{16}$ and $Si_{10}L_{14}$ with L = H and F by using the cluster genetic algorithm developed in our laboratory. Our results show that the low energy Si_xF_y clusters cannot be guaranteed to be identified by simply replacing the H atoms in low energy Si_xH_y clusters by F and then performing a local geometry optimization.

In the low energy well H passivated Si clusters, the H atoms prefer to be fairly evenly distributed over all the Si core atoms enabling five- and six-membered Si rings with little ring strain to be formed. As a consequence, clusters with the appropriate stoichiometry, such as $Si_{10}H_{16}$, have global minima with the Si cores resembling the diamond-lattice structure in bulk Si as might be anticipated by chemical intuition. In contrast the low energy well F passivated Si clusters like to form structures containing one or more trifluorosilyl groups at the energy expense of forming highly strained four- and even three-membered Si rings. The trend of the different Si_xF_y clusters favoring the formation of SiF_3 groups means that the $Si_{10}F_{16}$ global minimum with strained five- and four-membered Si rings in the Si core we obtained is also chemically reasonable.

CHAPTER 7. CONCLUSION

In this dissertation we developed a new strategy based on the evolutionary theory for globally optimizing the binary molecular system. Our results show that this new strategy dramatically excels random search and exhaustive search in the perspective of efficiency. Also it outperforms the global minimum searching simply by chemical intuition which often misses the true global minimum and many other low energy clusters. We applied this new strategy for globally optimizing the various Si_xH_y and Si_xF_y clusters coupled with the AM1 and GAM1 semiempirical methods for energy evaluation. Further MP2 and B3LYP calculations are performed to provide more reliable energy rankings for the Si_xH_y and Si_xF_y low energy isomers we obtained from the CGA.

As an initial project we adopted the fast AM1 semiempirical method for the local optimization and energy evaluation of clusters in the GA evolution process. We used the CGA to prescreen the low energy structures for the clusters $Si_{10}H_y$ with $y = 4, 8, 12, 16, 20$ and $Si_{14}H_{20}$ at the AM1 level. We then performed a more accurate energy ranking of the low energy structures generated by the GA with B3LYP and MP2 calculations. We have found that the B3LYP and MP2 energy rankings agree very well. However the AM1 relative energies of the low energy isomers we obtained from the CGA do not match with the MP2 or B3LYP relative energies especially for the low H passivated Si clusters. The Si_{10} framework evolves from compact structures to more open structures with increasing numbers of H atoms and we have found that Si_{10} was essentially fully passivated with 16 H atoms. Adding H atoms causes substantial changes in the Si_{10} framework once all the Si atoms are four coordinate. We have also found that the Si framework in the global geometry of the fully H passivated

clusters such as $Si_{10}H_{16}$ and $Si_{14}H_{20}$ to adopt the structure of a bulk Si lattice fragment providing B3LYP and MP2 energy calculations were performed.

To better guarantee the ab initio accuracy and yet retain the efficiency of the AM1 method, we proposed an improved iterative global optimization strategy involving two separate genetic algorithms which are invoked consecutively. One is the cluster GA (CGA) to find the Si_xH_y cluster global minimum. The second parametrization GA (PGA) is used to reparametrize the AM1 method against ab initio data from a training set of Si_xH_y clusters to produce the GA-optimized AM1 (GAM1) parameters. The overall procedure is considered to be converged when the low energy clusters located by the CGA reach energetic or structural convergence. To illustrate this iterative strategy we globally optimized the Si_6H_2 and Si_6H_6 clusters and successfully located the global minimum of the Si_6H_2 clusters and low energy Si_6H_6 clusters. Even though we did not locate the Si_6H_6 global minimum, several low energy clusters we obtained have same Si framework as the Si_6H_6 global minimum structure but with one or two H atoms on different positions. Generally speaking, this iterative strategy is very effective in finding the cluster global minimum at the ab initio level. However, the drawback of this strategy is that we have to iteratively develop a different GAM1 parameter set for every different cluster stoichiometry.

The above iterative strategy involving two GAs for each specific stoichiometry soon becomes very computationally demanding for larger clusters such as $Si_{10}H_{16}$. In our third project we have applied a parametrization genetic algorithm (PGA) to determine the GAM1 parameters using a Si_7H_{14} training set consisting of 14 different structures locally optimized at the MP2 level. The resulting GAM1 method reproduces the Si_7H_{14} MP2 relative energies and geometries in much better agreement than the original AM1 method. We then have investigated the transferability of the GAM1 parameters

derived from a Si_7H_{14} training set to larger clusters. We first evaluated the GAM1 optimized energies of the clusters with various stoichiometries ($Si_{10}H_y$, $y = 4, 8, 12, 16, 20$ and $Si_{14}H_{20}$) we located in previous research work [153]. Our results show that the GAM1 and MP2 relative energies and energy rankings match well for the $Si_{10}H_{16}$, $Si_{10}H_{20}$ and $Si_{14}H_{20}$ stoichiometries which have the similar level of passivation as the Si_7H_{14} stoichiometry but not for low H passivated Si clusters. To further illustrate the transferability of this GAM1 parameters for larger clusters, we have globally optimized the $Si_{10}H_{14}$, $Si_{10}H_{16}$, $Si_{14}H_{18}$ and $Si_{14}H_{20}$ clusters using the Si_7H_{14} GAM1 parameters to evaluate the cluster energies. We found that the GAM1 and MP2 energy rankings match well for each of the 4 above stoichiometries and the GAM1 global minimum clusters also have the lowest MP2 energies. Unlike for the AM1 calculations, we found the diamond-lattice like $Si_{10}H_{16}$ and $Si_{14}H_{20}$ structures to have both lowest GAM1 and MP2 energies. These results further confirm that the $Si_{10}H_{16}$ and $Si_{14}H_{20}$ clusters prefer diamond-lattice structures at 0 K. However, our calculations also show that the total Gibbs free energy of the $Si_{10}H_{14}$ or $Si_{14}H_{18}$ global minimum and a H_2 molecule is close to that of the diamond-lattice $Si_{10}H_{16}$ or $Si_{14}H_{20}$ cluster respectively at the room temperature.

Even with the GAM1 method parametrized to reproduce the ab initio result, we still had difficulty finding the global minima for $Si_{14}H_{20}$ and Si_6H_6. Our fourth project is to develop a faster cluster genetic algorithm (CGA) for finding the global minimum of Si_xH_y clusters. The key components to the CGA are the genetic operators. However we realized that our previous genetic operators are not very efficient in removing/producing the silyl groups for the well H passivated Si clusters or moving H atoms around for the low H passivated Si cluster. Therefore we decided to replace our previous coordinate averaging method and fragment mixing method with more effective genetic operators.

We still keep the cut-and-paste mating operator to produce an offspring from two parent clusters. Four new mutation operators are introduced: SiH_3 removal, SiH_2 removal, H shift and piece rotation. These new operators better retain good structural information between different generations in the CGA thereby enabling a global minimum structure to be found with fewer cluster energy evaluations. The SiH_3 and SiH_2 removal operators enable the new CGA to locate a diamond-lattice like structure which is the $Si_{14}H_{20}$ global minimum. For the low H passivated clusters, the H shift operator assists with generating the H atom arrangement present such as in the Si_6H_6 global minimum. The improved rate of finding the global minimum was further demonstrated by using the new CGA to find the global minima for the larger $Si_{18}H_{24}$ and $Si_{18}H_{22}$ clusters.

In our fifth project we demonstrated that the global minima structures for Si_xH_y and Si_xF_y clusters are very different. We have determined the global minima for Si_2L_4, Si_7L_{14}, Si_8L_{14}, $Si_{10}L_{16}$ and $Si_{10}L_{14}$ with L = H and F by using the cluster genetic algorithm (CGA) coupled with semiempirical methods to evaluate the cluster energies. Our results show that the low energy Si_xF_y clusters cannot be obtained by simply replacing the H atoms in low energy Si_xH_y clusters by F atoms and then performing a local geometry optimization. The low energy H passivated Si clusters prefer five- and six-membered Si rings given enough H ligands. As a consequence, clusters with the appropriate stoichiometry, such as $Si_{10}H_{16}$ and $Si_{14}H_{20}$, have global minima with a diamond-lattice like Si core. In contrast the low energy F passivated Si clusters like to form structures containing one or more trifluorosilyl groups at the expense of forming highly strained four- and even three-membered Si rings.

In summary, our first major discovery in this dissertation is that if the Si_xH_y cluster is fully H passivated then the Si_x core forms a diamond-lattice like structure, while the Si_x core in Si_xH_{y-2} is significantly different from the diamond-lattice structure.

This finding is well illustrated in figure 7.1 which displays the global minimum structures of $Si_{10}H_{16}$, $Si_{10}H_{14}$, $Si_{14}H_{20}$, $Si_{14}H_{18}$, $Si_{18}H_{24}$ and $Si_{18}H_{22}$ stoichiometries.

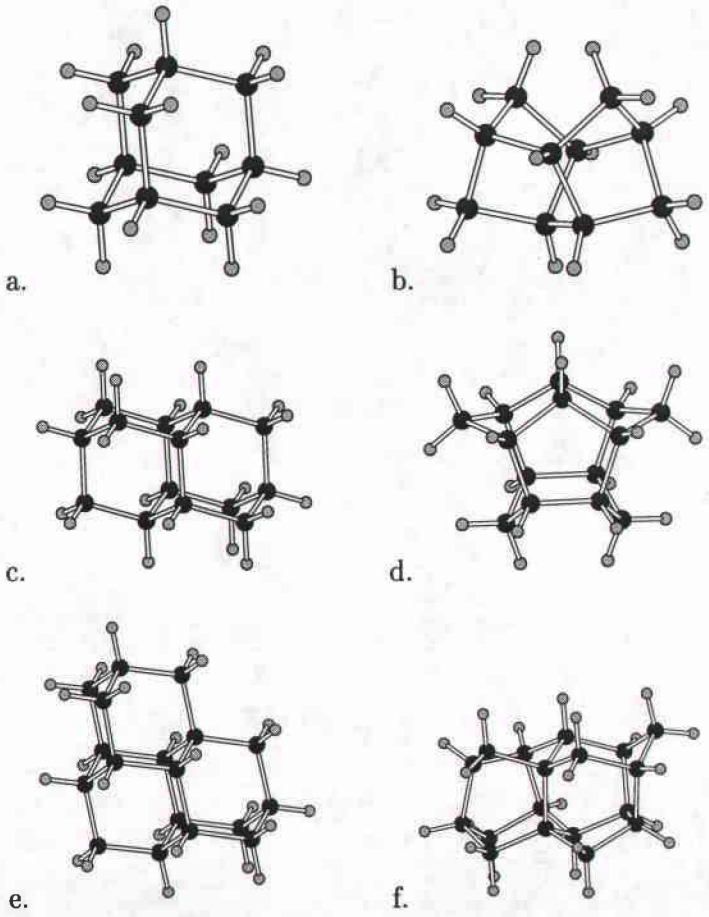


Figure 7.1: The Si_xH_y MP2/LanL2dz(d) global minima: a. $Si_{10}H_{16}$; b. $Si_{10}H_{14}$; c. $Si_{14}H_{20}$; d. $Si_{14}H_{18}$; e. $Si_{18}H_{24}$; f. $Si_{18}H_{22}$.

The second major discovery in this work is that the well-passivated Si_xF_y and Si_xH_y clusters prefer very different structures. The global minimum structures of Si_7L_{14}, Si_8L_{14}, $Si_{10}L_{16}$ and $Si_{10}L_{14}$ clusters (with L = H or F atom) displayed in figure 7.2 show that the well-passivated Si_xH_y clusters prefer five or six-membered Si rings, while the

corresponding Si_xF_y clusters prefer SiF_3 groups being formed at the expense of ring strain from the small four- and even three-membered rings.

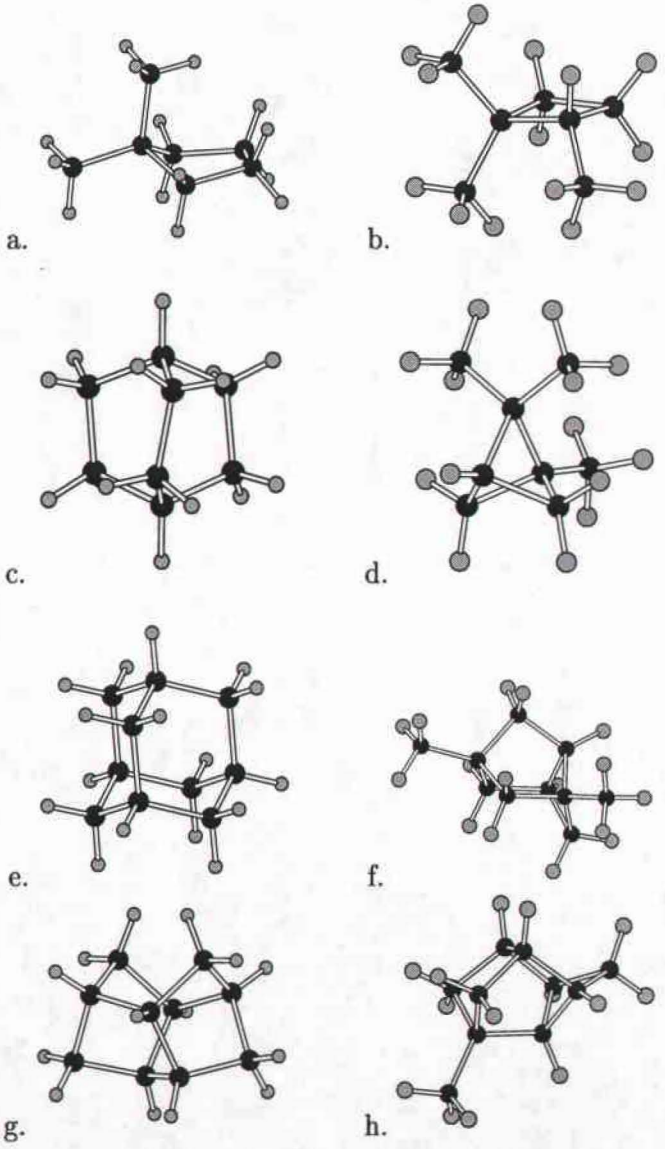


Figure 7.2: The Si_xL_y global minima at ab initio level: a. Si_7H_{14}; b. Si_7F_{14}; c. Si_8H_{14}; d. Si_8F_{14}; e. $Si_{10}H_{16}$; f. $Si_{10}F_{16}$; g. $Si_{10}H_{14}$; h. $Si_{10}F_{14}$.

So far we have developed various techniques to efficiently locate the global min-

imum and other low energy clusters for both H and F passivated Si clusters. These techniques should after proper modifications be able to perform cluster global optimizations of other binary molecular systems such as $Si_x(OH)_y$ and $Si_x(OCH_3)_y$ where the -OH and -OCH_3 groups can be treated as mono-ligands. Another interesting follow-up project will be to study the excited states of the various Si_xL_y low energy clusters we obtained where L represents a H or F ligand. The excited state investigation may lead us to uncover the mechanism of the photoluminescence of porous Si and the nanometer sized passivated Si clusters.

REFERENCES

[1] Canham, L. T. *Appl. Phys. Lett.* **1990**, *57*, 1046.

[2] Calcott, P. D. J.; Nash, K. J.; Canham, L. T.; Kane, M. J.; Brumhead, D. J *J. Phys.: Condens. Matter* **1993**, *5*, L91.

[3] Calcott, P. D. J.; Nash, K. J.; Canham, L. T.; Kane, M. J.; Brumhead, D. J *J. Lumin.* **1993**, *57*, 257.

[4] Nash, K. J.; Calcott, P. D. J.; Canham, L. T.; Needs, R. J. *Phys. Rev. B* **1995**, *51*, 17698.

[5] Cullis, A. G.; Canham, L. T.; Calcott, P. D. J. *J. Appl. Phys.* **1997**, *82*, 909.

[6] Wilson, W. L.; Szajowski, P. F.; Brus, L. E. *Science* **1993**, *262*, 1242.

[7] Raghavachari, K.; Pacchioni, G. *J. Chem. Phys.* **2001**, *114*, 4657.

[8] Chabal, Y. J.; Raghavachari, K.; Zhang, X.; Garfunkel, E. *Phys. Rev. B* **2002**, *66*, 161315.

[9] Zyubin, A. S.; Mebel, A. M.; Lin, S. H.; Glinka, Yu. D. *J. Chem. Phys.* **2002**, *116*, 9889.

[10] Zhou, F.; Head, J. D. *J. Phys. Chem. B* **2000**, *104*, 9981.

[11] Gole, J. L.; Dedul, F. P.; Grantier, D.; Dixon, D. A. *Phys. Rev. B.* **1997**, *56*, 2137.

[12] Gole, J. L.; Dixon, D. A. *J. Phys. Chem. B* **1997**, *101*, 8101.

[13] Gole, J. L.; Dixon, D. A. *Phys. Rev. B* **1998**, *58*, 4761.

[14] Gole, J. L.; Prokes, S. M. *J. Phys. Chem. B* **1998**, *102*, 33.

[15] Gole, J. L.; Dixon, D. A. *Phys. Rev. B* **1998**, *57*, 12002.

[16] Gole, J. L.; Dixon, D. A. *J. Appl. Phys.* **1998**, *83*, 5985.

[17] Gole, J. L.; Dixon, D. A. *J. Phys. Chem. B* **2000**, *104*, 1777.

[18] Stutzmann, M; Brandt, M. S.; Rosenbauer, M.; Fuches, H. D.; Finkbeiner, S.; Weber, J.; Deak, P. *J. Lumin.* **1993**, *57*, 321.

[19] Brandt, M. S.; Stutzmann, M *Solid State Commun.* **1995**, *93*, 473.

[20] Wolkin, M. V.; Jorne, J.; Fauchet, P. M.; Allan, G.; Delerue, C. *Phys. Rev. Lett.* **1999**, *82*, 197.

[21] Filonov, A. B.; Ossicini, S.; Bassani, F.; d'Avitaya, F. A. *Phys. Rev. B* **2002**, *65*, 195317.

[22] Zhu, X; Zeng, X. C. *J. Chem. Phys.* **2003**, *118*, 3558.

[23] Rohlfing, C. M.; Raghavachari, K. *Chem. Phys. Lett.* **1990**, *167*, 559.

[24] Li, S.; Van Zee, R. J.; Weltner, J. W.; Raghavachari, K. *Chem. Phys. Lett.* **1995**, *243*, 275.

[25] Honea, E. C.; Ogura, A.; Murray, C. A.; Raghavachari, K.; Sprenger, W. O.; Jarrold, M. F.; Brown, W. L. *Nature* **1993**, *366*, 42.

[26] Arnold, C. C.; Neumark, D. M. *J. Chem. Phys.* **1993**, *99*, 3353.

[27] Voss, D. *Science* **1996**, *273*, 751.

[28] Ho, K. M.; Shvartsburg, A. A.; Pan, B.; Lu, Z. Y.; Wang, C. Z.; Wacker, J. G.; Fye, J. L.; Jarrold, M. F. *Nature* **1998**, *392*, 582.

[29] Liu, B.; Lu, Z. Y.; Pan B. C.; Wang C. Z.; Ho, K. M.; Shvartsburg, A. A.; Jarrold, M. F. *J. Chem. Phys.* **1998**, *109*, 9401.

[30] Shvartsburg, A. A.; Jarrold, M. F.; Liu, B.; Lu, Z. Y.; Wang C. Z.; Ho, K. M. *Phys. Rev. Lett.* **1998**, *81*, 4616.

[31] Wang, C. Z.; Pan, B. C.; Ho, K. M. *J. Phys.: Condens. Matter* **1999**, *11*, 2043.

[32] Patterson, C. H.; Messmer,R. P. *Phys. Rev. B* **1990**, *42*, 7530.

[33] Rothlisberger, U.; Andreoni, W.; Parrinello, M. *Phys. Rev. Lett.* **1994**, *72*, 665.

[34] Kaxiras, E. *Phys. Rev. B* **1997**, *56*, 13455.

[35] Gong, X. G.; Zheng, Q. Q.; He, Y. *J. Phys.: Condens. Matter* **1995**, *7*, 577.

[36] Gong, X. G. *Phys. Rev. B* **1987**, *52*, 14677.

[37] Gong, X. G. *Phys. Rev. B* **1993**, *47*, 2329.

[38] Gingerich, K. A.; Ran, Q.; Schmude Jr., R. W. *Chem. Phys. Lett.* **1996**, *256*, 274.

[39] Jackson, K.; Pederson, M.; Wang, C.; Ho, K. *Phys. Rev. A* **1999**, *59*, 3685.

[40] Sieck, A; Sieck, A.; Porezag, Y.; Frauenheim, T.; Pederson, M. R.; Jackson, K. A. *Phys. Rev. A* **1997**, *56*, 4890.

[41] Pederson, M. R.; Jackson, K.; Porezag, Y.; Hajnal, Z.; Frauenheim, T. *Phys. Rev. B* **1996**, *54*, 2863.

[42] Jackson, K.; Pederson, M. R.; Porezag, Y.; Hajnal, Z.; Frauenheim, T.*Phys. Rev. B* **1997**, *55*, 2549.

[43] Mitas, L.; Grossman, J. C.; Stich, I.; Tobik, J. *Phys. Rev. Lett.* **2000**, *84*, 1479.

[44] Bachels, T; Schafer, R; *Chem. Phys. Lett.* **2000**, *324*, 365.

[45] Bolding, B. C.; Anderson, H. C. *Phys. Rev. B* **1990**, *41*, 10568.

[46] Rata, I.; Shvartsburg, A. A.; Horoi, M.; Frauenheim, T.; Siu, K. W. M.; Jackson, K. A. *Phys. Rev. Lett.* **2000**, *85*, 546.

[47] Iwamatsu, M. *J. Chem. Phys.* **2000**, *112*, 10976.

[48] Stillinger, F. H.; Weber, T. A. *Phys. Rev. B* **1985**, *31*, 5262.

[49] Hartke, B. *Chem. Phys. Lett.* **1996**, *258*, 144.

[50] Hartke, B. *Theoret. Chem. Acc.* **1998**, *99*, 241.

[51] Jarrold, M. F. *Science* **1991**, *252*, 1085.

[52] Hudgins, R. R.; Imai, M.; Jarrold, M. F. Dugourd, P. *J. Chem. Phys.* **1999**, *111*, 7865.

[53] Marsen, B.; Lonfat M.; Scheier, P.; Sattler K. *J. Electron Spectroscopy and Related Phenomena* **2000**, *109*, 157.

[54] Marsen, B.; Lonfat M.; Scheier, P.; Sattler K. *Phys. Rev. B* **2000**, *62*, 6892.

[55] Katzer, G.; Ernst, M. C.; Sax, A. F.; Kalcher, J. *J. Phys. Chem. A* **1997**, *101*, 3942.

[56] Onida, G.; Andreoni, W. *Chem. Phys. Lett.* **1995**, *243*, 183.

[57] Chakraborti, N.; Kumar R. *J. Phase Equalib.* **2003**, *24*, 132.

[58] Chakraborti, N.; Prasad, R. *B. Mater. Sci.* **2003**, *26*, 127.

[59] Chakraborti, N.; De, P. S.; Prasad R. *Mater. Lett.* **2002**, **55**, 20.

[60] Chakraborti, N.; Misra, K.; Bhatt, P.; Barman, N.; Prasad, R. *J. Phase Equalib.* **2001**, *22*, 525.

[61] Chakraborti, N.; De, P. S.; Prasad R. *Z Metallkunde* **1999**, *90*, 508.

[62] Prasad, R.; Chakraborti, N. *Met. Mater. Process* **1998**, *10*, 203.

[63] Miyazaki, T.; Uda, T.; Stich, I.; Terakura, K. *Surf. Sci.* **1997**, *377*, 1046.

[64] Miyazaki, T.; Uda, T.; Stich, I.; Terakura, K. *Chem. Phys. Lett.* **1996**, *261*, 346.

[65] Miyazaki, T.; Uda, T.; Stich ,I.; Terakura, K. *Chem Phys. Lett.* **1998**, *284*, 12.

[66] Meleshko, V.; Morokov, Y.; Schweigert, V. *Chem. Phys. Lett.* **1999**, *300*, 118.

[67] Meleshko, V. P.; Morokov, Y. N.; Schweigert, V. A. *J. Struc. Chem.* **1999**, *40*, 10.

[68] Meleshko, V. P.; Morokov, Y. N.; Schweigert, V. A. *J. Struc. Chem.* **1999**, *40*, 503.

[69] Sari, L.; McCarthy, M. C.; Schaefer, H. F.; Thaddeus, P. *J. Am. Chem. Soc.* **2003**, *125*, 11409.

[70] Pak, C.; Rienstra-Kiracode, J. C.; Schaefer, H. F. III *J. Phys. Chem. A* **2000**, *104*, 11232.

[71] Sax, A. F.; Kalcher, J. *J. Phys. Chem.* **1991**, *95*, 1768.

[72] Curtiss, L. A.; Raghavachari, K.; Deausch P. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *95*, 2433.

[73] Gong, X. G.; Guenzburger, D.; Saitovitch, E. B. *Chem. Phys. Lett* **1997**, *275*, 392.

[74] Hirao, M.; Uda, T. *Surface Sci.* **1994**, *306*, 87.

[75] Greeff, C. W.; Lester, W. A. Jr. *J. Chem. Phys.* **1997**, *106*, 6412.

[76] Klein, P.; Urbassek, H. M. *Phys. Rev. B* **1999**, *60*, 5478.

[77] Jursic, B. S. *J. Molecular Struct.* **2000**, *497*, 65.

[78] Swihart, M. T.; Girshick, S. L. *Chem. Phys. Lett.* **1999**, *307*, 527.

[79] Krüger, T.; Sax, A. F. *J. Comput. Chem.* **2001**. *22*, 151.

[80] Chambreau, S. D.; Wang, L. M.; Zhang, J. S. *J. Phys. Chem. A* **2002**, *106*, 5081.

[81] Zhang, S. B.; Branz, H. M. *Phys. Rev. Lett.* **2001**, *87*, 105503.

[82] Zhang, S. B.; Branz, H. M. *Phys. Rev. Lett.* **2000**, *84*, 967.

[83] Zhou, Z.; Friesner R. A.; Brus L. *J. Am. Chem. Soc.* **2003**, *125*, 15599.

[84] Zhou, Z.; Brus L.; Friesner R. A. *Nano Lett.* **2003**, *3*, 163.

[85] Benedict, L. X.; Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli G.; Klepeis, J. E. *Phys. Rev. B* **2003**, *68*, 085310.

[86] Graeger, E. W.; Grossman, J. C.; Williamson, A. J.; Galli, G. *Phys. Rev. Lett.* **2003**, *90*, 167402.

[87] Sillars, D.; Bennett, C. J.; Osamura, Y.; Kaiser, R. I. *Chem. Phys. Lett.* **2004**, *392*, 541.

[88] Sillars, D.; Bennett, C. J.; Osamura, Y.; Kaiser, R. I. *Chem. Phys.* **2004**, *305*, 141.

[89] Takahashi, M.; Sakamoto, K. *J. Phys. Chem. A* **2004**, *108*, 5710.

[90] Izumi, S.; Sato, Y.; Hara, S.; Sakai, S. *Surf. Sci.* **2004**, *560*, 1.

[91] Feller, D.; Dixon, D. A. *J. Phys. Chem. A* **1999**, *103*, 6413.

[92] King, R. A.; Mastryukov, V. S.; Schaefer, H. F. *J. Chem. Phys.* **1996**, *105*, 6880.

[93] Li, G.; Li, Q.; Xu, W.; Xie, Y; Schaefer, H. F. *Mol. Phys.* **2001**, *99*, 1053.

[94] Garey, M. R.; Johnson, D. S. *Computers and intractability: a guide to the theory of NP-completeness*, San Francisco: W. H. Freeman, 1980.

[95] Tsai, C. J.; Jordan, K. D. *J. Phys. Chem.* **1993**, *97*, 11227.

[96] Judson, R. S.; Colvin, M. E.; Meza, J. C.; Huffer, A.; Gutierrez, D. *Inter. J. Quantum Chem.* **1992**, *44*, 277.

[97] Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller E. *J. Chem. Phys.* **1953**, *21*, 1087.

[98] Kirkpatrick, S.; Gelatt, C. D. Jr.; Vecchi, M. P. *Science* **1983**, *220*, 671.

[99] Goldberg, D. *Genetic Algorithm in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.

[100] R. Judson, in Reviews in Computational Chemistry, Vol. 10, Eds. K. B. Lipkowitz, D. B. Boyd, VCH, New York, 1999, p. 1.

[101] Johnston, R. L. *Dalton Trans.* **2003**, 4193.

[102] Deaven, D. M.; Ho, K. M. *Phys. Rev. Lett.* **1995**, *75*, 288.

[103] Hartke, B. *J. Comput. Chem.* **1999**, *20*, 1752.

[104] Cartwright, H. M., in Structure and Bonding, Vol. 110, Series Ed. Mingos D. M. P., Vol. Ed. R. L. Johnston, Springer-Verlag Berlin Heidelberg, 2004, p. 1.

[105] Hartke, B., in Structure and Bonding, Vol. 110, Series Ed. Mingos D. M. P., Vol. Ed. R. L. Johnston, Springer-Verlag Berlin Heidelberg, 2004, p. 33.

[106] Harris, K. D. M.; Johnston, R. L.; Habershon, S., in Structure and Bonding, Vol. 110, Series Ed. Mingos D. M. P., Vol. Ed. R. L. Johnston, Springer-Verlag Berlin Heidelberg, 2004, p. 55.

[107] Woodley S. M., in Structure and Bonding, Vol. 110, Series Ed. Mingos D. M. P., Vol. Ed. R. L. Johnston, Springer-Verlag Berlin Heidelberg, 2004, p. 95.

[108] Gillet, V. J., in Structure and Bonding, Vol. 110, Series Ed. Mingos D. M. P., Vol. Ed. R. L. Johnston, Springer-Verlag Berlin Heidelberg, 2004, p. 133.

[109] Unger, in Structure and Bonding, Vol. 110, Series Ed. Mingos D. M. P., Vol. Ed. R. L. Johnston, Springer-Verlag Berlin Heidelberg, 2004, p. 153.

[110] Niesse, J. A.; Mayne, H. R. *J. Chem. Phys.* **1996**, *105*, 4700.

[111] Luo, Y. H.; Zhao, J.; Qiu, S.; Wang, G. *Phys. Rev. B* **1999**, *59*, 14903.

[112] Ross I.; Truhlar, D. G.; *Chem. Phys. Lett.* **1995**, *233*, 231.

[113] Cundari, T. R.; Deng, J.; Fu, W. *Inter. J. Quant. Chem.* **2000**, *77*, 421.

[114] Bazterra, V. E.; Ferraro, M. B.; Facelli J. C. *J. Chem. Phys.* **2002**, *116*, 5984.

[115] Bazterra, V. E.; Ferraro, M. B.; Facelli J. C. *J. Chem. Phys.* **2002**, *116*, 5992.

[116] Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S129.

[117] Pople, J. A.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, 9136.

[118] Pople, J. A.; Beveridge D. L.; Dobosh, P. A. *J. Chem. Phys.* **1967**, *47*, 2026.

[119] Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *69*, 3858.

[120] Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

[121] Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.

[122] Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221.

[123] Thiel, W. *J. Am. Chem. Soc.* **1981**, *103*, 1413.

[124] Thiel, W.; Voityuk, A. A. *Theor. Chim. Acta.* **1992**, *81*, 391.

[125] Thiel, W.; Voityuk, A. A. *Inter. J. Quantum Chem.* **1992**, *44*, 807.

[126] Thiel, W.; Voityuk, A. A. *J. Mol. Struct.* **1994**, *313*, 141.

[127] Thiel, W.; Voityuk, A. A. *J. Phys. Chem.* **1996**, *100*, 616.

[128] Thiel, W.; Voityuk, A. A. *Theor. Chim. Acta.* **1996**, *93*, 315.

[129] Dewar, M. J. S.; Jie C. X.; Yu, J. G. *Tetrahedron* **1993**, *49*, 5003.

[130] Holder A. J.; Dennington R. D.; Jie C. X. *Tetrahedron* **1994**, *50*, 627.

[131] Voityuk, A. A.; Zerner, M. C.; Rosch, N. *J. Phys. Chem. A* **1999**, *103*, 4553.

[132] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

[133] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[134] Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

[135] Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554.

[136] Hohenberg P.; Kohn W. *Phys. Rev.* **1964**, *136*, B864.

[137] Kohn W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

[138] Dirac, P. A. M. *Proc. Cam. Phil. Soc.* **1930**, *26*, 376.

[139] Slater, J. C. *Phys. Rev.* **1951**, *81*, 385.

[140] Vosko, S. J.; Wilk, L.; Nusair, M.; *Can. J. Phys.* **1980**, *58*, 1200.

[141] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

[142] Perdew, J. P.; Wang Y.; *Phys. Rev. B*, **1991**, *45*, 13244.

[143] Gill, P. M. W. *Mol. Phys.* **1996**, *89*, 433.

[144] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

[145] Perdew, J. P.; *Phys. Rew. B* **1986**, *33*, 8822.

[146] Dewar, M. J. S.; Jie, C. *Organometallics* **1987**, *6*, 1486.

[147] Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 284.

[148] Dunning, Jr., T. H.; Hay, P. J. *Modern Theoretical Chemistry*, Ed. Schaefer, III, H. F.; Plenum: New York, 1977; Vol. 3, p. 1.

[149] GAMESS, The General Atomic and Molecular Electronic Structure System, Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, Jr., J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

[150] 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*, Revision E.2, Gaussian, Inc., Pittsburgh PA, 1995.

[151] *CRC Handbook of Chemistry and Physics*, 1st ed.; CRC Press: Boca Raton, Fl., 1990.

[152] Scott, A. P; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.

[153] Ge, Y.; Head, J. D. *J. Phys. Chem. B* **2002**, *106*, 6997.

[154] Ge, Y.; Head, J. D. *Inter. J. Quant. Chem.* **2003**, *95*, 617.

[155] Ge, Y.; Head, J. D. *J. Phys. Chem. B* **2004**, *108*, 6025.

[156] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. *J. Chem. Phys. 1982*, **77**, 3654.

[157] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, **56**, 2257.

[158] Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

[159] Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110*, 7650.

[160] Wolf, M. D.; Landman, U. *J. Phys. Chem. A* **1998**, *102*, 6129.

[161] Dewar, M. J. S.; Zoebisch, E. G. *Theochem.* **1988**, *180*, 1.

[162] Pople, J. A.; Head-Gordon, M.; Raghavavhari, K. *J. Chem. Phys.* **1987**, *87*, 5968.

[163] Degoli, E.; Cantele, G.; Luppi, E.; Magri, R.; Ninno, D.; Bisi, O.; Ossicini, S. *Phys. Rev. B* **2004**, *69*, 155411.

[164] Ge, Y.; Head, J. D. submitted to *Chem. Phys. Lett.*