Analysis of Diamondoid Mechanosynthesis Tooltip Pathologies Generated via a Distributed Computing Approach

Damian G. Allis[1,*], Brian Helfrich[2], Robert A. Freitas Jr.[3], and Ralph C. Merkle[3]

[1]*Department of Chemistry, Syracuse University, Syracuse, NY 13244, USA*
[2]*Helcorp, Maplewood, NJ, 07040, USA*
[3]*Institute for Molecular Manufacturing, Palo Alto, CA 94301, USA*

The results of a combined molecular dynamics/quantum chemistry pathology study of previously reported organic (diamondoid) tooltips for diamondoid mechanosynthesis (DMS) are presented. This study, employing the NanoHive@Home (NH@H) distributed computing project, produced 80,000 tooltip geometries used in 200,000 calculations optimized at either the RHF/3-21G or RHF/STO-3G levels of theory based on geometries obtained from high-energy molecular dynamics simulations to produce highly deformed starting geometries. These 200,000 calculations have been catalogued, grouped according to energies and geometries, and analyzed to consider potentially accessible defect structures (pathologies) for tooltip geometries either binding a carbon dimer (C_2) feedstock or not containing the transported dimer feedstock. The transport and deposition of feedstock and the stability of the tooltip between dimer "loading" cycles are important geometries that must be considered as part of a tooltip stability analysis. The NH@H framework is found to be a useful method both for the study of highly deforming covalent geometries and, using lower-temperature MD simulations, for generating and optimizing molecular conformations (demonstrated using biotin, *n*-heptane, and *n*-octane in this study). The results of the pathology survey are discussed and general considerations for the exploration of DMS tooltip usability are explored.

Keywords: BOINC Platform, Conformational Searches, Diamondoid Mechanosynthesis, Distributed Computing, Molecular Manufacturing, NanoHive@Home, Tooltip Pathologies.

RESEARCH ARTICLE

1. INTRODUCTION

Molecular quantum mechanics (QM), classical molecular dynamics (MD) simulations, and their combination play unique roles in the field of molecular manufacturing by their use as the foundational tools directing nearly all research efforts. In the sub-field of diamondoid mechanosynthesis (DMS),[1,2] initial experimental feasibility studies have begun but are difficult due to the current inability to positionally control and deposit carbon (C) atoms or reactive organic fragments, although prior experimental studies demonstrate the feasibility of mechanical bond manipulation[3–5] and the first US patent on DMS was issued in 2010.[6] The positional-control feature of DMS defines its many differences from (and possible advantages over) traditional chemosynthetic approaches for the fabrication of complex nanostructures. In the field of

DMS, computational chemistry can be thought to serve the same role in experimental studies and engineering efforts that computer-aided design (CAD) performs in modern mechanical and electrical engineering, providing a framework for design and validation in efforts to greatly reduce the time and resource requirements of experimental studies performed in the absence of these predictive tools. The use of computational tools for design and initial testing influences all areas of DMS study, including the modeling of molecular tooltips for the delivery of feedstock to workpieces[1,2,7–11] and of workpiece reactivity during mechanosynthetic operations,[7–9,13] the design of useful DMS reaction sequences[2,9] and pathways for recharging spent tooltips,[2,12] and the modeling of feedstock deposition onto workpieces for the fabrication of structured surfaces and constrained molecular fragments.[2,9] Computational studies provide considerable information concerning the dynamics of structures, binding energies of feedstock to

---

*Author to whom correspondence should be addressed.

tooltips and workpieces, transition state barriers between possible defect structures (pathologies), and estimates of positional tolerances[8, 9, 12] based on the control of operational conditions (such as system temperatures and component composition). The utility of these studies is limited largely by the quality of the calculations that can be performed or, more specifically, the accessibility of computational resources capable of handling high-level calculations on large structures.

The area of study in DMS that has received the most attention to date has been the design and validation of tooltips used for the delivery of feedstock to workpieces.[1, 2, 7–12] A generic tooltip and division into its important components is shown in Figure 1. These studies consider the stability of feedstock bound to tooltips,[1, 2, 7, 8] the transfer energies of feedstock to workpieces,[2, 7–12] the optimization of tooltip designs to limit the potential for interactions between the tooltip framework and potentially reactive workpieces,[1, 2, 8] and the potential for defect structure formation.[1, 2, 8, 12] The formation of tooltip defect structures either with or without bound feedstock is, after the identification of candidate tooltip delivery geometries, the most important area of testing. The results of post-design tooltip studies may determine the acceptability of a tooltip as an experimental candidate, either removing a tooltip design from consideration based on the accessibility of geometries that prohibit recharging or deposition, or instigating the redesign of a tooltip to reduce the likelihood of a possible route to defect formation.

Two factors limit the identification of defects and their possible accessibility during mechanosynthetic operations. These include

(1) the number of potential defects accessible for a proposed tooltip and

(2) the computational resources required to model what may be multiple accessible defect structures from what may be multiple candidate tooltip structures.

In the interest of addressing both of these limiting factors, a distributed computing study was performed using
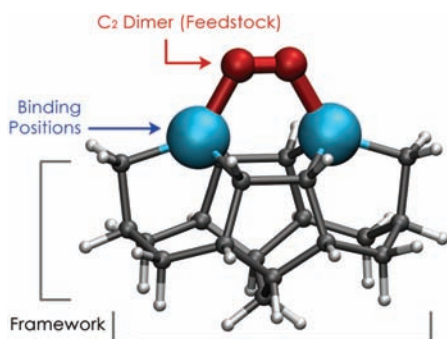

Fig. 1. The general design of a molecular tooltip for diamondoid mechanosynthesis. A tooltip is composed of a framework that holds, e.g., two binding atoms, a pair of binding atoms for the stable binding and controlled deposition of feedstock, and some feedstock (in this case, a carbon dimer) that is deposited onto a workpiece.

the NH@H project[16] and a series of tooltips that had been the subjects of two previous DMS studies. This NH@H study, known as the "NH@H tooltip pathology survey," employed a combined MD/QM methodology, the "Quantum Search for Minimum Alternatives in Kinetically-Accessible Space" (Q-SMAKAS), and combined the computational resources of 10,000 volunteered PCs to identify possible defect structures for 25 tooltips generated from high-temperature (highly deforming) MD simulations and optimized (via distributed computing) using quantum chemistry calculations. The analysis of the nearly 200,000 quantum chemistry calculations covering the generated structures and the presentation of possible defect structures for this set of 25 tooltips are described in the present work. The two previous DMS studies are summarized briefly below.

(1) *Horizontal Ge-Substituted Polymantane-Based C_2 Dimer Placement Tooltip Motifs for Diamond Mechanosynthesis:* This study[1] of carbon dimers (C_2) bound to two germanium (Ge) atoms within adamantane and related frameworks considered only pathologies that changed the binding modes of the dimer. The 24 tooltips for which high-level computational studies (B3LYP/6-311 + G(2d,p)//B3LYP/3-21G* relative energies and normal mode analyses) were performed were from an original set of 53 tooltips, of which 29 were determined by an AM1-driven MD simulation (combined semi-empirical quantum chemistry and classical dynamics) at 900 K to be too unstable for use in mechanosynthetic operations due to such identified problems as dimer orientation instability (a preference for the C_2 to orient perpendicular to the Ge...Ge axis to form a carbenic structure, see below), Ge–C bond strain due to the design of the C framework, large mobility of the tooltip framework after C_2 removal, and proximity of tooltip hydrogen (H) atoms to the Ge atoms (as H migration to the C_2 binding positions is one form of tooltip poisoning). This study was a follow-up to the first dimer tooltip proposal (2003) for the "DCB" and "MCB" motifs.[7]

(2) *Design and Analysis of a Molecular Tool for Carbon Transfer in Mechanosynthesis:* This study[8] considered a single tooltip (named "DC10c") in great detail, with the analysis of geometry, binding energies, defect structures, and transition state calculations performed in a manner that models how the subsequent studies of the 24 candidate tooltips above would be performed as either a means to acceptability or removal from possible experimental consideration. This study determined that the C_2-bound tooltip geometry was both a stable minimum energy structure and that several potential defect structures were, despite some that were determined to be more stable than the operational tooltip geometry, inaccessible based on their large calculated transition state barriers. The transition state search and defect structure analysis included both C_2 orientation and larger defects within the tooltip framework itself, with the generation of defects determined

2

J. Comput. Theor. Nanosci. 8, 1–23, 2011

from strain analyses at various positions, expectations based on organic mechanism considerations, and chemical intuition.

2. COMPUTATIONAL METHODS

NH@H[16] was a distributed computing system based on the Berkeley Open Infrastructure for Network Computing (BOINC) platform.[17] BOINC is a system that manages the distribution of computing jobs to participant computers on the Internet. Participants register and donate CPU time, after which data and applications are sent to the participant machine from a central server. After their computer has completed processing the data, it uploads the results to the central server and is then sent more data to process.

An initial geometry optimization of each tooltip, both with bound feedstock ("loaded") and without ("unloaded," see below), was performed to produce baseline structures. All energy minimizations were calculated using the open source Massively Parallel Quantum Chemistry (MPQC) program from Sandia National Laboratory.[18] The initial optimizations were performed both in-house and with the NH@H network. The majority of the tooltips were optimized at a Restricted Hartree-Fock level with the 3-21G Gaussian-type basis set.[19] Seven of the tooltips were determined to be too large for the 3-21G basis set, with the tradeoff of compute time versus chemical accuracy biased in favor of faster optimization of more structures and away from the more common strategy of slower but more accurate optimization of fewer structures. The need to accommodate the computing power of many participant machines (ca. February–May 2007) further argued for a

reduced basis set quality approach in these larger systems, resulting in the use of the STO-3G basis set.[20] STO-3G geometries are of sufficient quality to serve as the basis for later studies at higher levels of theory, given that the goal of the project was to generate a selection of potential failure modes for subsequent work. The tooltips processed at the RHF/STO-3G level were MCB5Ge, MCB57Ge, DCB75Ge, C100GeATD, C100GeATS, C100GeCTS, and MCB75Ge (see the Results section for the geometries).

The optimized tooltip geometries served as the starting structures for molecular dynamics (MD) simulations at high temperatures for 50,000 iterations, with snapshots of the tooltip saved to a file every 40 or 50 iterations. The thermal energy provided during the high temperature simulations perturbed each tooltip significantly. The snapshots provided kinetically accessible deformations of tooltip geometries, which were then subjected to quantum chemistry optimization at either the RHF/3-21G or RHF/STO-3G levels of theory. MD simulations were performed with the GROMACS software package[21] using the GROMOS87[22] force field with suitable modifications.[14, 15] The following parameters were employed for the simulations: timestep = 2.0 fs, neighbor list updated every ten steps, twin range electrostatics cut-offs with neighborlist cut-off = 0.8 nm and Coulomb cut-off = 1.4 nm, van der Waals cut-off = 0.8 nm, Berendsen-thermostat temperature coupling with time constant = 0.1 ps, and Berendsen pressure coupling.

Initially, a 50,000-iteration MD simulation at 1500 K was performed with each loaded and unloaded tooltip to establish a baseline level of tooltip deformation. The root mean square deviation (RMSD) of all tooltips at this

Table I. Unique geometries obtained from each molecular dynamics simulation (temperature in K as reported for each collection of structures) for the loaded (L) and unloaded (U) tooltips. For each tooltip, a representative structure code from each run is provided (Struct.) and is available as Supplemental Material. **L***n* and **U***n* labels correspond to structures provided in Figure 6. The numbers in bold within the **Struct.** columns are the total number of successful quantum chemistry optimizations for each tooltip snapshot, with the total number of snapshots submitted to quantum chemistry optimization provided below the tooltip label for each MD run (*n* **Runs**).

C100GeATD	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1210**	**1500 K**
L1	1t	0.0
L2	928t	159.2
1000 Runs	**263**	**4500 K**
U1	1s	0.0
U2	59s	−43.6
U3	861s	165.0
U4	627s	417.9
1000 Runs	**121**	**5000 K**

C100GeATSr5	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1094**	**1500 K**
L1	1v	0.0
L2	421v	138.6
L3	90v	279.1
1000 Runs	**636**	**3500 K**
U1	1w	0.0
U2	41w	81.2
1000 Runs	**711**	**3000 K**

C100GeATS	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1218**	**1500 K**
L1	1t	0.0
1000 Runs	**331**	**4500 K**
U1	5t	0.0
U2	412t	93.6
U3	626t	296.1
U4	995t	372.6
1000 Runs	**342**	**4500 K**

C100GeATSr6	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1050**	**1500 K**
L1	1x	0.0
1000 Runs	**711**	**2500 K**

C100GeCTS	Struct.	kJ/mol
L1	1v	0.0
L2	258v	261.3
L3	280v	326.1
1000 Runs	**661**	**3500 K**

C100GeCTS	Struct.	kJ/mol
U1	1w	0.0
U2	970w	102.8
1000 Runs	**829**	**3000 K**

temperature was around 0.02 Å. Quantum chemistry optimizations were then performed on the NH@H network using MPQC for each snapshot with the following convergence criteria and parameters: basis set = RHF/3-21G or RHF/STO-3G (as defined above), maximum energy gradient ≤0.001 Hartrees/Bohr, maximum atom displacement ≤0.001 Å, maximum optimization convergence attempts = 40, maximum energy convergence (SCF) cycles per optimization step attempt = 100. For this RMSD ≈ 0.02 Å pass of calculations, no failure modes were discovered. The representative structures for this temperature are tagged with "z" in the "Struct." column of Tables I–V. See below for more information.

A series of MD simulations were then performed for each tooltip in increasing temperature increments of 500 K until structures were produced with structural RMSDs of at least 0.04 Å. The final range of temperatures for the various loaded and unloaded tooltips were: 2000 K (tagged "y" in the "Struct." column of Tables I–V), 2500 K (tagged "x"), 3000 K (tagged "w"), 3500 K (tagged "v"), 4000 K (tagged "u"), 4500 K (tagged "t"), 5000 K (tagged "s"), 5500 K (tagged "r"), and 6000 K (tagged "q"). For this

Table II. Unique geometries obtained from each molecular dynamics simulation (temperature in K as reported for each collection of structures) for the loaded (L) and unloaded (U) tooltips. For each tooltip, a representative structure code from each run is provided (Struct.) and is available as Supplemental Material. **L***n* and **U***n* labels correspond to structures provided in Figure 7. The numbers in bold within the **Struct.** columns are the total number of successful quantum chemistry optimizations for each MD simulation, with the total number of structures submitted to quantum chemistry optimization provided below the tooltip label for each MD run (***n* Runs**).

DCB55AGe	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1088**	**1500 K**
L1	1t	0.0
1000 Runs	**363**	**4500 K**
U1	1u	0.0
U2	630u	−118.5
U3	898u	−83.8
U4	470u	−74.0
U5	909u	−1.4
U6	3u	16.3
U7	606u	94.0
U8	28u	234.7
1000 Runs	**549**	**4000 K**

MCB55Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**944**	**1500 K**
L1	32q	0.0
L2	682q	−49.8
L3	816q	−26.2
L4	129q	30.2
L5	161q	41.7
L6	958q	100.1
L7	839q	218.2
1000 Runs	**14**	**6000 K**
U1	1z	0.0
U2	746z	63.0
1000 Runs	**913**	**1500 K**

MCB5Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1196**	**1500 K**
L1	1y	0.0
1000 Runs	**899**	**2000 K**
U1	1z	0.0
1000 Runs	**916**	**2000 K**

DCB65Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**952**	**1500 K**
L1	1w	0.0
1000 Runs	**490**	**3000 K**
U1	1v	0.0
1000 Runs	**346**	**3500 K**

DCB55CGe	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1012**	**1500 K**
L1	1v	0.0
L2	225v	165.7
1000 Runs	**401**	**3500 K**
U1	1u	0.0
U2	701u	66.4
U3	81u	126.8
U4	424u	210.9
1000 Runs	**485**	**4000 K**

DCBIceane7Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1095**	**1500 K**
L1	1v	0.0
1000 Runs	**704**	**3500 K**
U1	1w	0.0
U2	10w	7.4
U3	431w	57.4
U4	503w	110.2
1000 Runs	**509**	**3500 K**

MCB75Ge	Struct.	kJ/mol
L1	1x	0.0
1000 Runs	**138**	**2500 K**

MCB57Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1148**	**1500 K**
L1	1y	0.0
1000 Runs	**827**	**2000 K**

DCB75Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1187**	**1500 K**
L1	1x	0.0
L2	79x	10.5
L3	274x	142.1
1000 Runs	**638**	**2500 K**
U1	1y	0.0
U2	381y	71.4
1000 Runs	**620**	**2000 K**

DCB57Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**987**	**1500 K**
L1	1u	0.0
1000 Runs	**443**	**4000 K**
U1	1u	0.0
U2	998u	−20.7
U3	969u	55.2
U4	843u	171.5
1000 Runs	**394**	**4000 K**

DCB6Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**977**	**1500 K**
L1	1v	0.0
1000 Runs	**490**	**3500 K**

RESEARCH ARTICLE

Table III. Unique geometries obtained from each molecular dynamics simulation (temperature in K as reported for each collection of structures) for the loaded (L) and unloaded (U) tooltips. For each tooltip, a representative structure code from each run is provided (Struct.) and is available as Supplemental Material. **L***n* and **U***n* labels correspond to structures provided in Figure 8. The numbers in bold within the **Struct.** columns are the total number of successful quantum chemistry optimizations for each MD simulation, with the total number of structures submitted to quantum chemistry optimization provided below the tooltip label for each MD run (***n* Runs**).

C110GeS	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1188**	**1500 K**
L1	1v	0.0
L2	613v	−11.6
L3	558v	21.8
L4	801v	97.5
L5	783v	124.2
1000 Runs	**698**	**3500 K**
U1	1v	0.0
U2	304v	27.6
U3	635v	190.5
1000 Runs	**322**	**3500 K**

TwistaneGe	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1223**	**1500 K**
L1	1u	0.0
L2	238u	−58.6
1000 Runs	**551**	**4000 K**

AdamGe22	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1195**	**1500 K**
L1	1w	0.0
L2	701w	−22.4
L3	865w	4.7
1000 Runs	**834**	**3000 K**
L1	1v	0.0
L4	787v	36.2
1000 Runs (D)	**746**	**3500 K**
U1	1w	0.0
U2	742w	45.2
U3	733w	204.9
U4	724w	276.5
1000 Runs	**354**	**3000 K**

C111Ge3	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1037**	**1500 K**
L1	1v	0.0
1000 Runs	**584**	**2500 K**

AdamGe33	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1200**	**1500 K**
L1	1w	0.0
L2	568w	85.2
L3	409w	97.9
1000 Runs	**742**	**3000 K**
U1	2z	0.0
U2	597z	6.0
U3	350z	238.1
U4	216z	277.6
1000 Runs	**422**	**1500 K**

Diad3Ge22	Struct.	kJ/mol
L1	2z	0.0
1250 Runs	**1003**	**1500 K**
L1	1x	0.0
L2	865x	111.7
1000 Runs	**670**	**3000 K**
U1	1w	0.0
U2	951w	−218.4
U3	37w	37.9
U4	872w	63.9
1000 Runs	**506**	**3000 K**

RMSD > 0.04 Å pass of calculations, failure modes were discovered. The final analysis of the generated structures was performed by energy comparisons, using a 1 kJ/mol cut-off for the grouping of distinct geometries. Structures of identical geometry that differed by more than 1 kJ/mol but less than 3 kJ/mol were sorted by inspection. For the performed tooltip calculations, defect geometries are the result of changes to covalent frameworks and not simply conformational changes. Despite this, alternate minima were determined in a few cases to lie close in energy to the starting geometries, thereby resulting in the small 1 kJ/mol starting structure delimiter.

An additional set of calculations were performed that only succeeded in generating conformations (and not structures with altered covalent frameworks) in three familiar molecules. A series of 1500 K MD simulations were

Table IV. Unique conformational geometries obtained from each molecular dynamics simulation (temperature in K as reported for each collection of structures) for the loaded (L) and unloaded (U) tooltips. For each tooltip, a representative structure code from each run is provided (Struct.) and is available as Supplemental Material. **L***n* and **U***n* labels correspond to structures provided in Figure 9. The numbers in bold within the **Struct.** columns are the total number of successful quantum chemistry optimizations for each MD simulation, with the total number of structures submitted to quantum chemistry optimization provided below the tooltip label for each MD run (***n* Runs**).

DCB55BGe	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**607**	**1500 K**
L1	1w	0.0
1000 Runs	**141**	**3000 K**

DCB55BGe	Struct.	kJ/mol
U1	1z	0.0
U2	681z	−71.4
U3	928z	−1.3
U4	860z	1.2
U5	301z	8.9
U6	441z	18.7
U7	978z	20.6
U8	524z	22.0
U9	655z	24.0
U10	710z	29.8
1000 Runs	**929**	**1500 K**

DCB5Ge	Struct.	kJ/mol
L1	1z	0.0
1250 Runs	**1006**	**1500 K**
L1	1w	0.0
L2	13w	126.2
L3	377w	150.0
L4	881w	159.8
1000 Runs	**561**	**4000 K**
U1	2x	0.0
U2	950x	−58.8
U3	237x	25.0
1000 Runs	**663**	**2500 K**

Table V. Unique geometries obtained from each molecular dynamics simulation (temperature in K as reported for each collection of structures) for the loaded (L) and unloaded (U) DC10c tooltips. For each tooltip, a representative structure code from each run is provided (Struct.) and is available as Supplemental Material. **L*n*** and **U*n*** labels correspond to structures provided in Figure 10. The numbers in bold within the **Struct.** columns are the total number of successful quantum chemistry optimizations for each MD simulation, with the total number of structures submitted to quantum chemistry optimization provided below the tooltip label for each MD run (*n* **Runs**). Simulations using deuterium masses for all H atoms are identified as (D).

DC10c	Struct.	kJ/mol		DC10c	Struct.	kJ/mol		DC10c	Struct.	kJ/mol
L1	1v	0.0		L1	1r	0.0		U1	1q	0.0
L2	719v	−3.1		L3	557r	19.3		U2	259q	−14.3
L6	563v	320.9		L5	798r	312.1		U3	889q	23.0
1000 Runs	**880**	**3500 K**		L7	841r	322.3		U4	159q	144.5
				L8	776r	337.1		U5	704q	168.7
L1	1w	0.0		**1000 Runs**	**490**	**5500 K**		U6	718q	201.9
L4	193w	307.7						U7	641q	288.5
1000 Runs (D)	**372**	**3000 K**						U8	901q	299.9
								1000 Runs	**952**	**6000 K**

performed for biotin, *n*-heptane, and *n*-octane in order to demonstrate the utility of the NH@H network as a tool for conformational sampling. For each of these three molecules, 1000 MD snapshots were generated and optimized with NH@H at the RHF/3-21G level of theory.

For the NH@H tooltip pathology survey, the data sent to each participant computer were tooltip geometries (primarily snapshots) and MPQC. Each tooltip geometry was sent to two different computers and both results were compared before acceptance as final data. If the two results were not in agreement, the geometry was sent to a third computer, and so on until a quorum of two matching results were received. Approximately 6,000 participants donated time on over 10,000 computers to perform nearly 200,000 energy minimizations over a period of three months, representing ∼2.5 million CPU hours of actual geometry optimization time in the present work.

3. RESULTS

3.1. Conformational Searches with NH@H: Biotin, *n*-Heptane, and *n*-Octane

The utility of the Q-SMAKAS methodology for the more general quantum chemistry study of molecular conformation searches was tested using the biomolecule biotin ($C_{10}H_{16}N_2O_3S$) and the linear hydrocarbons *n*-heptane (C_7H_{16}) and *n*-octane (C_8H_{18}). The results of these studies are presented below with the geometries for all calculated structures provided as Supplemental Material.

3.1.1. Biotin

As a test case for conformational searches, biotin is an interesting molecule on several levels beyond its biological significance. There are three (two O–H and one N–H) H-bond donors and five (two O, one S, and two lone pair N) H-bond acceptors responsible for the water-solubility of this B-complex vitamin. As most of these fragments are localized on the fused ureido/tetrahydrothiophene (U/T) rings at the "head" of the

biotin molecule, the presence of multiple local conformational minima along the covalent potential energy surface are made accessible by flexibility in the $-C_4H_8-$ "tail" linking the fused ring head to the carboxylic acid (COOH) pendant group. While multiple local minima have been predicted by previous studies,[23–25] the two most interesting conformations of note are a cyclic form predicted in the gas phase (where a pair of H bonds are formed between the U/T rings and the COOH) and the fully-extended form found in the solid-state (where the formation of multiple intermolecular H bonds is preferred over intramolecular H-bond formation).[25]

Figure 2 shows a sample of the identified local minima predicted to lie within 40 kJ/mol of the global minimum, as well as the least stable structure optimized to a minimum energy geometry. Among those conformations within 40 kJ/mol of the global minimum, the stabilities of the biotin molecule geometries are found to increase significantly with H-bond formation between the U/T rings and the COOH. Neglecting the small variations in $-C_4H_8-$ conformations among the most stable structures, the identified biotin minima can be grouped according to the number and arrangement of H bonds formed between the U/T head and COOH tail. These basic motifs are shown in Figure 2. The most stable structures are found to form two direct H bonds using the entire COOH ($C{=}O_{COOH}\ldots H{-}N_{U/T}$ and $O{-}H_{COOH}\ldots O{=}C_{U/T}$) tail. The second group of structures forms a smaller cyclic H-bonding system containing an oriented $O{-}H_{COOH}$ group that acts as both H-bond donor ($O{-}H_{COOH}\ldots O{=}C_{U/T}$) and acceptor ($H{-}O_{COOH}\ldots H{-}N_{U/T}$). The third group contains structures with only a single directed H bond ($C{=}O_{COOH}\ldots H{-}N_{U/T}$ or $O{-}H_{COOH}\ldots O{=}C_{U/T}$). The final group to lie within 40 kJ/mol of the lowest energy structure contains no H-bonding interactions between the polar fragments of the U/T ring and COOH but may include weak $C{-}H\ldots O/N$ H bonds to add slight stability and diversify the number of predicted minima in this category.

The least stable structures (of which several are found to lie 90 kJ/mol or more above the global minimum) contain $O{=}C{-}O{-}H$ carboxylic acid geometries where the
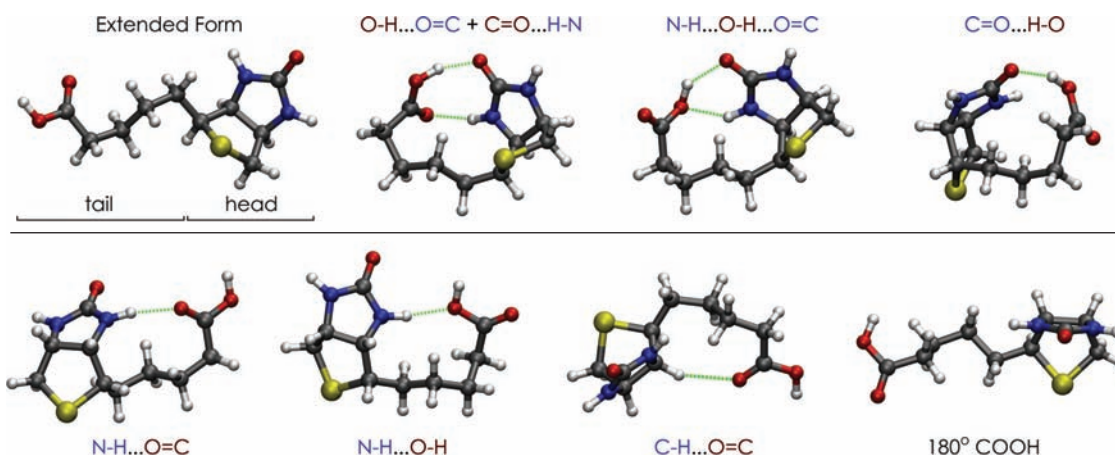
Fig. 2. Representative conformational minima within 40 kJ/mol of the predicted global minimum geometry ($O-H_{tail} \ldots O=C_{head} + C=O_{tail} \ldots H-N_{head}$) and the highest energy conformation identified from the HF/3-21G optimizations (180° COOH). Representative structure codes (top) and relative energies (ΔE_{Rel}, in kJ/mol) are also provided.

dihedral angle around the central C–O bond approaches 180 degrees. This local geometry is significant both to the analysis of the results and to the algorithm being used in the conformational survey. This geometry would be expected to be highly disfavored because it places the lone pair of the $O-H_{COOH}$ oxygen into its closest proximity with a lone pair of the $O=C_{COOH}$ oxygen, a geometry that combines the electrostatic repulsion of these two lone pairs with the breaking of a favorable interaction between the C=O lone pair and the H atom in the O–H. Nevertheless, this geometry is found to be an isolatable structure according to the SCF convergence algorithm in several of the calculated structures. *Even highly strained conformations can be trapped by a convergence algorithm in a quantum chemistry calculation* provided that the local potential well is deep enough to not allow the converger to sample alternate local minima that might lie much lower in energy.

3.1.2. n-Heptane and n-Octane

The linear alkanes *n*-heptane and *n*-octane are nonpolar molecules whose conformations are generated by rotation about C–C σ-bonds and are the result of a minimization of the total number of eclipsing H...H geometries for which maximum electrostatic repulsion occurs between neighboring $-CH_2-CH_2-$ fragments. In the case of *n*-heptane, the range of identified minima extends from the fully extended global minimum to an "S"-shaped maximum 31.5 kJ/mol higher in energy. For *n*-octane, this range extends to 41.1 kJ/mol in another highly twisted "S"-conformation that was only found in a single structure optimization. A selection of conformations for both molecules is shown in Figure 3, with the final structure groupings of the 1000 starting geometries for each molecule determined first by relative energies and secondly by RMSD analysis. The approximate 3.2 kJ/mol increase

in energy with each C–C rotation away from the most stable geometry is reproduced in the conformational groupings of both the *n*-heptanes and *n*-octanes.

3.2. Carbon (C_2) Dimer Tooltips and Pathologies

The 25 tooltip structures considered in the NH@H study are taken from two previous computational studies. All but one of the tooltips used were part of a larger survey by Freitas et al.,[1] This first survey identified structures that were predicted at the B3LYP/6-311 + G(2d,p)//B3LYP/3-21G* level of theory to bind C_2 either by the binding mode that uses the C_2 to bridge two Ge atoms separated by the organic tooltip framework or by defect modes that involve rearrangement of the C_2 dimer at its Ge...Ge binding position. The search for additional defect structures was not performed in the first survey due to the large number of tooltips and the absence of a thorough computational procedure for automating the defect structure process. One additional tooltip, DC10c, was the subject of a comprehensive optimization, defect geometry, and transition state analysis by Allis and Drexler.[8] This tooltip is distinct from the Ge-based tooltips in both the manner of C_2 binding in the completed (loaded) tooltip and the use of π-conjugation as the means for stabilization of the tooltip in the C_2 dimer-free (unloaded) tooltip — the Ge tooltips use only the relative binding energies of the C_2 to the framework Ge atoms and the stability of the framework C–C bonds relative to potential C–Ge structure rearrangement. The tooltip structures corresponding to the "operational" geometries (bound C_2 dimer on a stable optimized framework) for all 25 structures considered in the NH@H study are shown in Figures 4 and 5.

As a means to clarify terms used generally in the previous mechanosynthetic tooltip analyses but not otherwise common in a strictly chemical context, several terms used in this analysis are defined below.
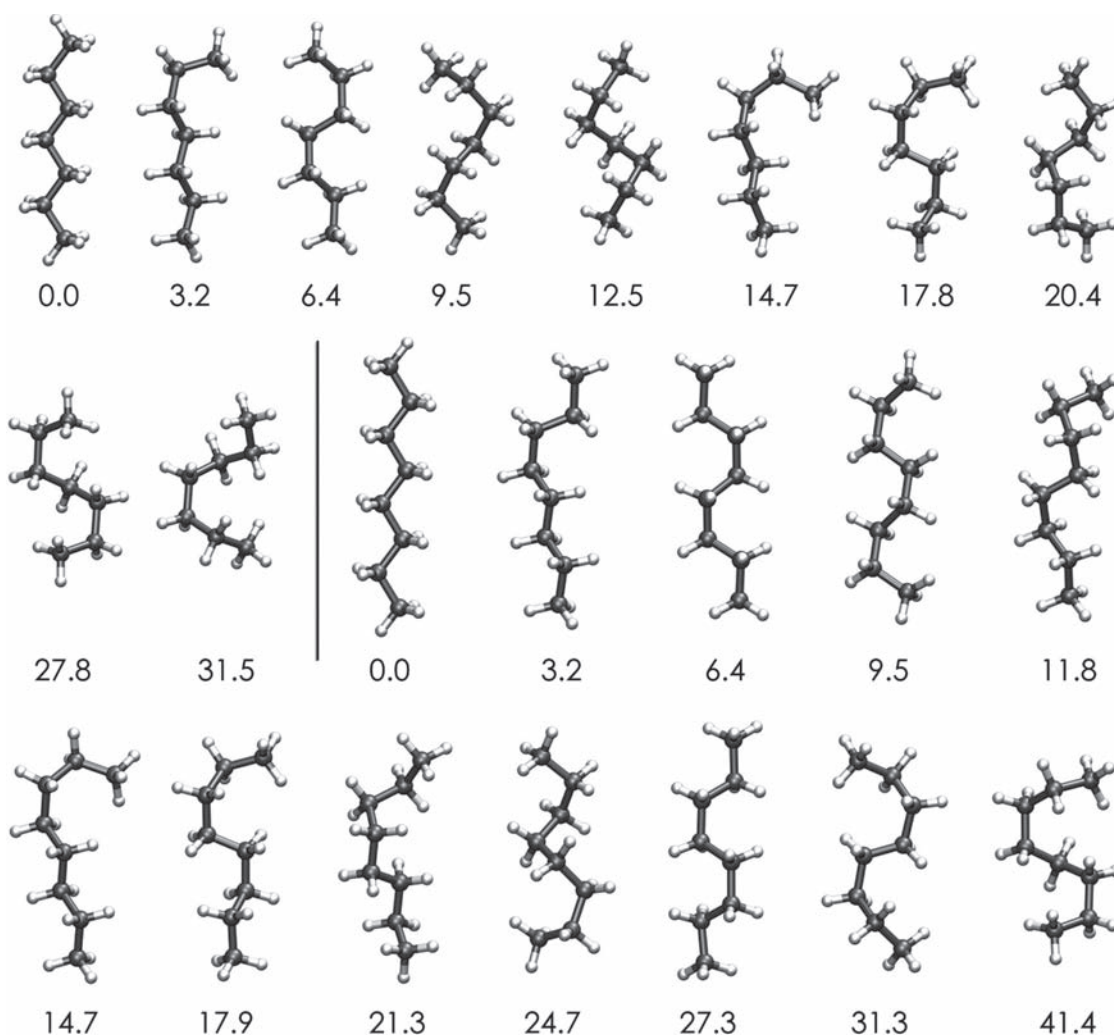
Fig. 3. Relative energy groupings (energies in kJ/mol) and representative structures from the *n*-heptane and *n*-octane structure optimizations.

3.2.1. Definitions

Loaded Tooltip—A loaded tooltip (LT) is one binding the C_2 dimer at both binding positions in the tooltip. The LT is distinct from C_2 dimer binding modes that do not bind the C_2 dimer at either end in the tooltip binding position (a few examples of which are provided as pathologies in the analysis).

Unloaded Tooltip—An unloaded tooltip (UT) is one in which the C_2 dimer is not present. For use in mechanosynthetic operations, rechargeable tooltips (one for which a new C_2 dimer has been added) require that both the LT be stable for positional control and the UT be stable after C_2 dimer deposition. A tooltip that undergoes any structural rearrangement upon C_2 dimer removal or that has any nearby accessible potential minima that involve minor structural changes that modify the C_2 dimer binding area is not a reliable tooltip for multiple operations.

Pathology—A *pathology*, as defined in previous mechanosynthetic studies,[1,2,12] is a defect structure in an operational context in the same way that a structural

isomer in a chemical context has the same molecular formula but a different atomic framework arrangement. Pathologies, like structural isomers, can appear very similar in geometry to the desired molecule, or can be completely rearranged in geometry, or may possess an intermediate geometry.

Ln and Un—The naming convention in this study for each tooltip and tooltip pathology is to label the base geometry (the expected LT or UT) as L1 (loaded base geometry) and U1 (unloaded base geometry). Subsequent structure labels are ordered from most stable to least stable pathology. In several cases, pathologies occur that are more stable than the expected LT and UT structures but are still labeled from $L(n+1)$ and $U(n+1)$. In Tables I–V, these Ln and Un labels are provided with a structure labeled in a "$N\mathbf{n}$" format, where N is the snapshot number and $\mathbf{n}$ ($= y, x, w, v, u, t, s, r, q$) is a letter corresponding to the bath temperature of the MD run (see the Methods section above). These codes correspond to representative geometries whose coordinates are provided as Supplemental
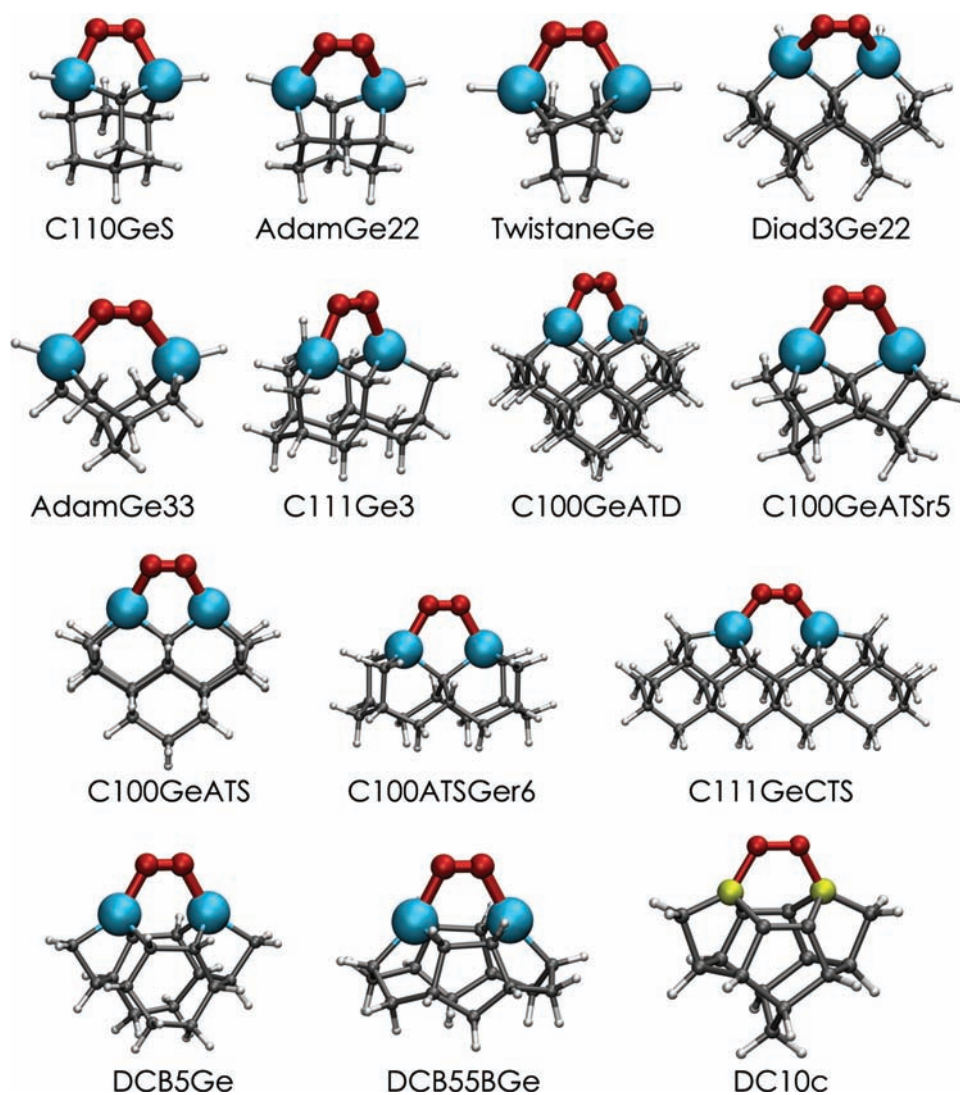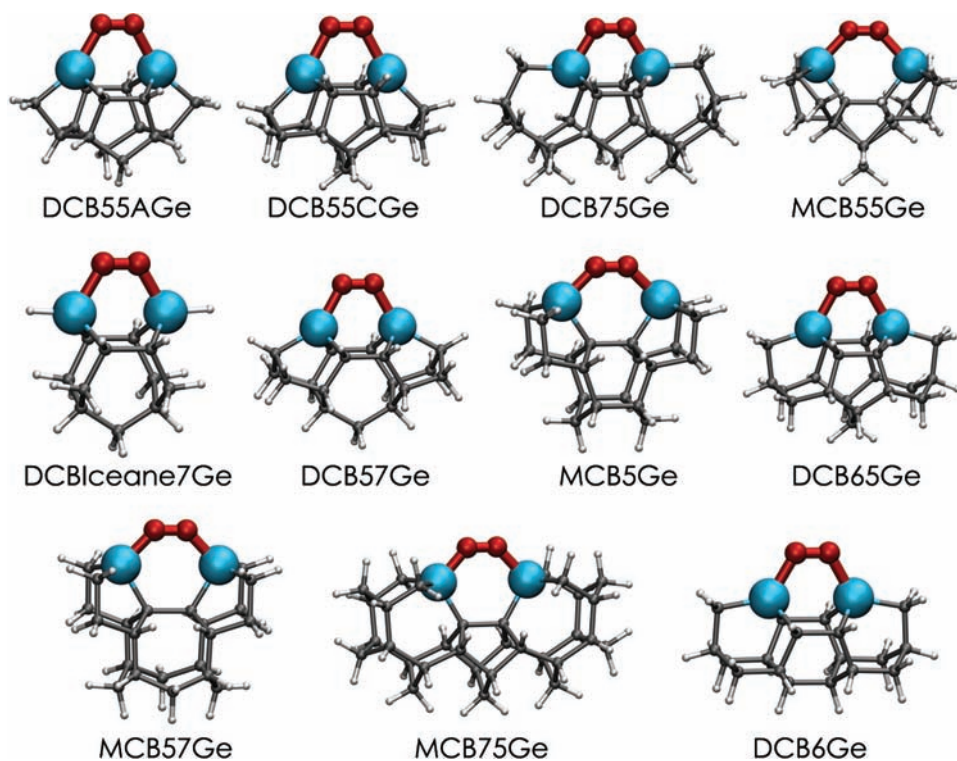
Fig. 4. Loaded tooltip (LT) geometries used in the NH@H study. Carbon dimers (C_2) are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. Germanium atoms (in all but the DC10c tooltip) are shown in blue. DC10c C atoms used for C_2 dimer binding are shown in yellow.

Material. They do not represent the numbers of times a tooltip was identified as a minimum, only one member of the set of tooltips with identical energies and geometries after the quantum chemistry optimization.

C_2 dimer—The C_2 carbon dimer is a feedstock fragment employed for mechanosynthetic depositions. C_2 has been the feedstock of choice for several previous computational studies as a predictably manageable building unit for diamondoid, nanotube, and graphene fabrication.[2, 7–9] The C_2 dimers in all images are shown in red.

Framework—The framework in these tooltips are the C atoms that hold the Ge atoms (or any C_2-dimer binding atom positions) in place. A change to the geometry of the framework (as occurs in some pathologies) must involve changes to the organic framework and may or may not involve changes to the C_2 dimer binding position (specifically, to the positions of the Ge atoms). A framework

defect may still serve as a proper tool for C_2 dimer recharging and mechanosynthetic operations.

Hydrogen inversion—The majority of all identified pathologies are the result of one or more H atoms from the organic framework inserting into the three-dimensional cage geometry of the tooltip. These structures result from the very high temperatures used in the MD simulations to generate the deformed geometries and the large H-atom displacements that occur as a result of the combination of H-atom momentum and long time steps. Generally, the H-inverted structures generated from the MD simulations are chemically inaccessible because of the significant energies involved both in the insertion and in the final H-trapped minima. As a simple procedure for removing these geometries from the structure analysis, simply placing a 400 kJ/mol energy cut-off on acceptable tooltips for included structures removes all but a few H-inverted

RESEARCH ARTICLE

Fig. 5. Loaded tooltip (LT) geometries used in the NH@H study. Carbon dimers (C_2) are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. Germanium atoms are shown in blue.

cases. Some of the tooltip structures are flexible enough at their framework bases that H inversion, while energetically inaccessible under normal conditions, leads to pathologies within the range of other pathologies, and so these structures are included in the analysis for comparative purposes. This simulation issue is addressed in greater detail in the Discussion section.

3.2.2. Ge-Tooltip Structures Optimized at the RHF/STO-3G Level of Theory

3.2.2.1. *C100GeATD (Table I, Fig. 6)*. A single LT pathology (L2, +159.2 kJ/mol) is identified from the 4500 K MD simulation generated by the migration of a framework C–H H atom onto the C_2 dimer and formation of a C_2 dimer-$C_{framework}$ single bond to produce a C–C–Ge three-member ring. Three UT pathologies are identified from a 5000 K MD simulation, including a more stable structure generated by the migration of a nearby C–H H atom to a tooltip Ge (U2), another structure with a broken Ge–C bond (U3), and another with considerable σ_{C-C} framework rearrangement and the formation of one C=C double bond (U4).

3.2.2.2. *C100GeATS (Table I, Fig. 6)*. The 1500 K and 4500 K MD simulations of the LT produced only the base geometry (L1). Three pathologies of the UT are identified in the 4500 K MD simulation that are less stable than the UT base structure (U1), with the U1 optimization itself resulting in the formation of a Ge–Ge bond (or

close contact; see below). The three pathologies all optimize with one or more broken Ge–C bonds. Pathology U2 includes Ge–C bond breaking and the formation of one C=C double bond. Pathology U3 has an H atom inserted into the tooltip framework and significant displacement of one Ge from its U1 position. Pathology U4 has one broken Ge–C bond and H migration. Both U3 and U4 represent defects not expected to occur in a larger tooltip framework, as the surface C–H H atoms involved in migration would be replaced with additional C–C framework bonds. In the isolated tooltips, the resulting H rearrangements produce structures of significantly higher energy than the base structure (U1).

3.2.2.3. *C100GeCTS (Table I, Fig. 6)*. Two LT pathologies are identified in the 3500 K simulation. The first (L2) involves the breaking of one C–C framework bond and a combination of H migration and C=C double bond formation which occurs away from the C_2 dimer binding position. Pathology L3 similarly involves framework carbon bond rearrangement with the preservation of the C_2 dimer binding geometry at the two Ge atoms. In L3, a single Ge–C bond is broken and, via several small rearrangements, a framework cyclopropyl ring is formed. The UT base structure (U1) optimizes with the formation of a Ge–Ge single bond. The only identified pathology from the 3000 K simulation (U2) combines a broken Ge–C bond with the formation of one C=C double bond and the migration of an H atom to one Ge atom that lies
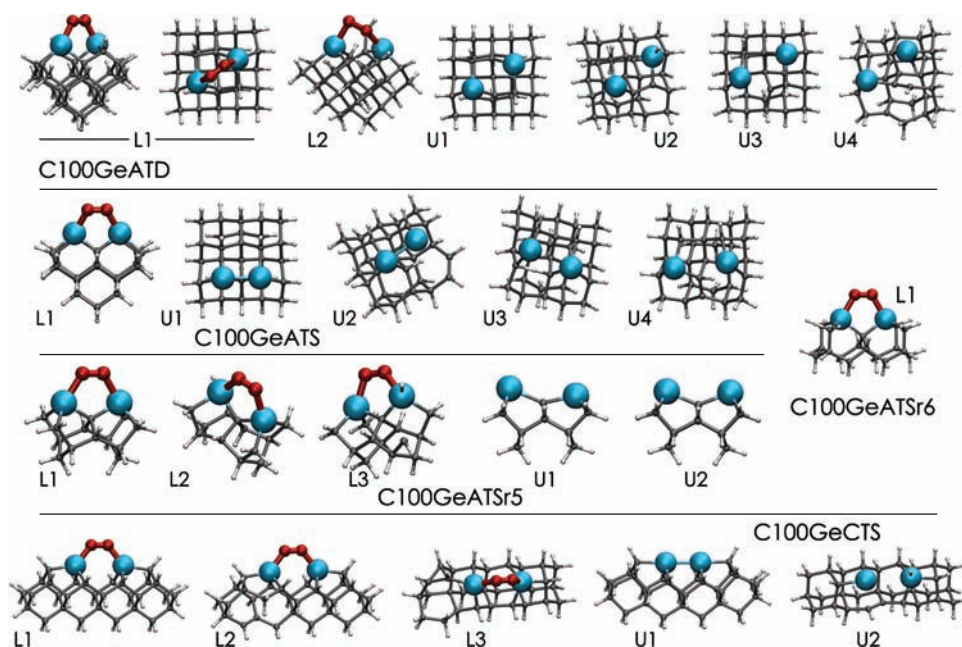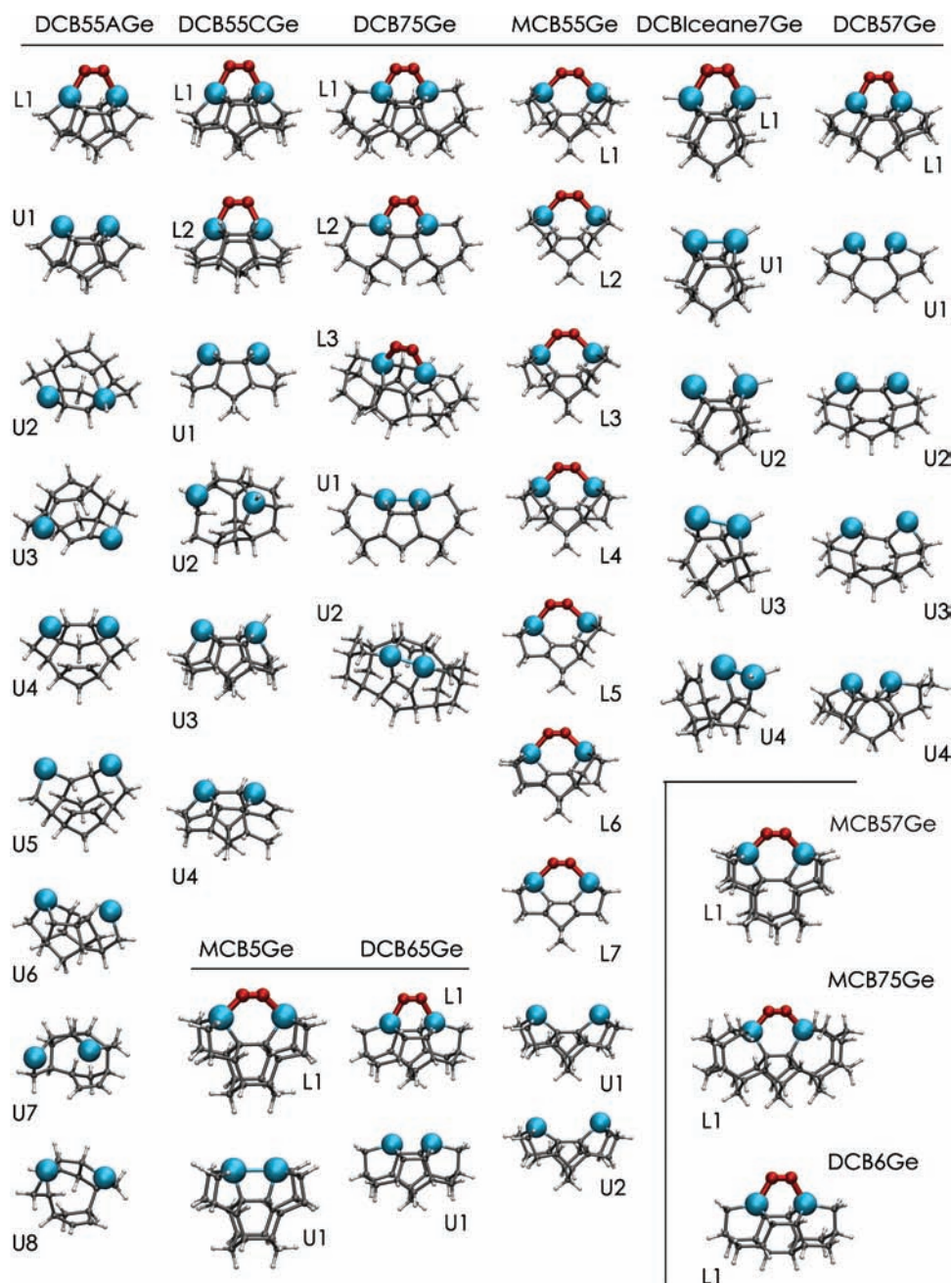
Fig. 6. LT and UT base geometries (L1 and U1, respectively) and identified pathologies. Carbon dimers (C_2) in the LT structures are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. Germanium atoms are shown in blue. The relative energies of these structures are provided in Table I.

103 kJ/mol higher in energy than U1; the source of this H atom is the nearest C atom in the C=C bond.

3.2.2.4. DCB75Ge (Table II, Fig. 7). Two LT pathologies were identified from the 2500 K MD simulation. The first of these (L2) is a conformational change of the two methylene (CH_2) fragments at the base of the tooltip and represents only a 10 kJ/mol change in energy. The second and more significant defect (L3) is the product of an H inversion into the tooltip framework. As the DCB75Ge framework is larger at the base, this H inversion occurs with a relatively small change in energy compared to other H-inverted structures (142 kJ/mol). In both pathologies, the addition of a larger support structure would be expected to remove these accessible defects.

A single UT pathology is identified in the 2000 K MD simulations. The base structure (U1) optimizes with the formation of a Ge–Ge single bond. In pathology U2, the Ge–Ge single bond is retained, but a single C–H H atom adjacent to one Ge atom is inverted into the flexible framework of the tooltip, producing a geometry only 71 kJ/mol higher than the U1 base geometry. The addition of a larger support structure might be expected to remove this accessible defect, although the close proximity of the inverted H atom to the two Ge atoms (and, therefore, to the C_2 dimer binding position) makes framework addition a potentially problematic solution to the potential H inversion problem.

3.2.2.5. MCB5Ge (Table II, Fig. 7). The LT 1500 K and 2000 K MD simulations produced no pathologies for this structure. The UT 2000 K MD simulations also produced no pathologies.

3.2.2.6. MCB57Ge (Table II, Fig. 7). The LT 1500 K and 2000 K MD simulations produced no pathologies for this structure. Energy convergence for the UT structures was not achieved in the allotted time and within the 40-step SCF limit.

3.2.2.7. MCB75Ge (Table II, Fig. 7). The LT 2500 K MD simulations produced no pathologies for this structure. Energy convergence for the UT structures was not achieved in the allotted time and within the 40-step SCF limit.

3.2.3. Ge-Tooltip Structures Optimized at the RHF/3-21G Level of Theory

3.2.3.1. C100GeATSr5 (Table I, Fig. 6). Two LT pathologies were produced from the 3500 K MD simulations. The first (L2) involves H inversion into the bottom of the tooltip framework which, being a more open structure, produces a pathology below the 400 kJ/mol H inversion cut-off. This pathology would not form in a larger framework, where this part of the tooltip would be embedded in a larger C framework. The second pathology (L3) involves the breaking of a Ge–C bond and the migration of one H atom from the broken Ge–C C atom onto the broken Ge–C Ge atom. The two identified UT pathologies are unique in that the geometries are similar in appearance but differ by 81.2 kJ/mol. The more stable structure (U1) is less symmetric than U2 with a 0.1 Å increase in the Ge–C distances from its bound carbons and a 0.2 Å reduction in the H_2C–Ge–CH_2 C...C separation.

3.2.3.2. C100GeATSr6 (Table I, Fig. 6). The LT 1500 K and 2500 K MD simulations produced no pathologies for this structure. Energy convergence for the UT structures

*J. Comput. Theor. Nanosci. 8, 1–23, **2011***
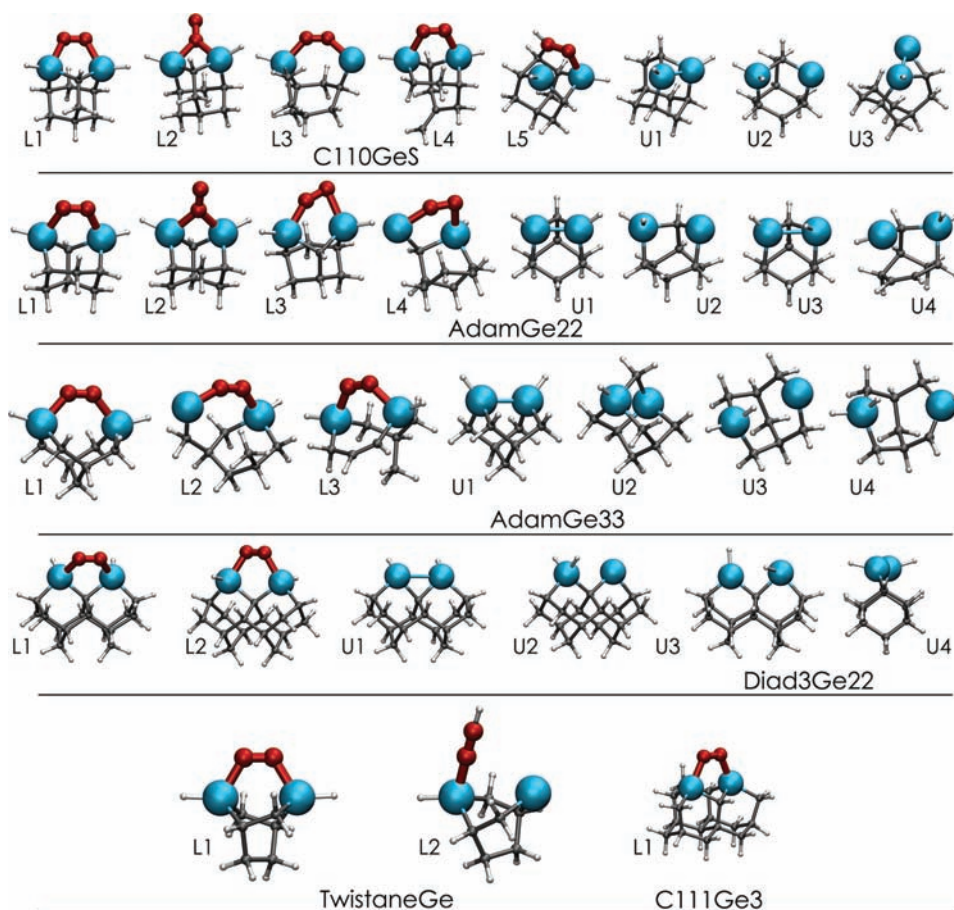
RESEARCH ARTICLE

11

Fig. 7. LT and UT base geometries (L1 and U1, respectively) and identified pathologies. Carbon dimers (C_2) in the LT structures are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. Germanium atoms are shown in blue. The relative energies of these structures are provided in Table II.

was not achieved in the allotted time and within the 40-step SCF limit.

3.2.3.3. *C110GeS (Table III, Fig. 8).* Four LT pathologies are identified from the 3500 K MD simulations, including one of slightly lower energy. Pathology L2 differs from the base LT geometry (L1) with the rotation of the C_2 dimer to form a structure with two Ge–C bonds to the same C_2 dimer atom. While the calculations here predict this geometry to lie 11.6 kJ/mol lower in energy, the original study on this system (at the

B3LYP/6-311 + G(2d,p)//B3LYP/3-21G* level of theory) predicted the base L1 geometry to be 54 kJ/mol more stable. Tooltip pathology L3 involves the breaking of one Ge–C bond and the migration of the Ge–H H atom to the broken Ge–C C atom. L4 is formed by the breakage of a framework C–C bond, the formation of one C=C double bond at the base, and the migration of one C–H H atom from a C=C bond C atom to the framework C atom formerly bound to the C=C double bond C atom. The final pathology (L5) features a broken Ge–C bond with

Fig. 8. LT and UT base geometries (L1 and U1, respectively) and identified pathologies. Carbon dimers (C_2) in the LT structures are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. Germanium atoms are shown in blue. The relative energies of these structures are provided in Table III.

migration of the C_2 dimer, breaking a Ge–C_2 dimer bond to form a $C_{framework}$–C_2 dimer bond. The new $C_{framework}$–C_2 dimer bond in L5 also produces H migration from the framework C ($C_{framework}$) to the C_2 dimer.

The UT 3500 K MD simulations produced two pathologies, with the geometry optimization of the base structure (U1) resulting in the formation of a Ge–Ge bond. The first pathology (U2) shifts one Ge–H H atom to the other Ge–H Ge atom, producing a completed GeH_2 fragment, one H-vacant Ge atom, and a breaking of the Ge–Ge bond in a structure only 27 kJ/mol higher in energy than U1. This defect can be prevented from forming in a larger tooltip framework where the Ge–H H atoms are replaced with additional framework C atoms. The higher-energy pathology (U3) combines the breaking of a Ge–C bond with Ge–H → C–H H migration to form a stable C framework and a Ge–Ge–C three-member ring.

3.2.3.4. C111Ge3 (Table III, Fig. 8). The LT 1500 K and 2500 K MD simulations produced no pathologies for this structure. Energy convergence for the UT structures was not achieved in the allotted time and within the 40-step SCF limit.

3.2.3.5. AdamGe22 (Table III, Fig. 8). The LT MD simulations at 3000 K and 3500 K produced three pathologies either more stable or only slightly less stable than the LT base structure (L1). Pathology L2 is 22.4 kJ/mol more stable than L1, the result of a rotation of the C_2 dimer such that both Ge atoms are bound to the same C atom in the C_2 dimer. This geometry is similar to that predicted for C110GeS-L2 but is more stable relative to its base geometry by 11 kJ/mol. Pathology L3 shifts the C_2 dimer such that one dimer π-bond is now directly bound to one Ge atom and the second C_2 dimer C atom is shifted towards this Ge atom to form a coordinating interaction. This structure is a small feedstock rearrangement with a small relative energy change but is of great significance because of the amount of structural flexibility identified and the potential problems associated with positional control raised by the small barrier to dimer motion. Pathology L4 is a derivative of the L3 geometry, with a single Ge–C bond breaking and the Ge–H H atom from the Ge–C Ge atom shifting to the Ge–C C atom. Despite the large change in geometry, this pathology lies only 36.2 kJ/mol higher in energy than base geometry U1.

Three UT pathologies are identified from the 1500 K MD simulations, with the base structure (U1) optimization producing a Ge–Ge bond. Pathology U2 optimizes with both Ge–H H atoms shifted to the same Ge atom (retaining the Ge–Ge bond but leaving one Ge with a diminished valence). Pathology U3 optimizes with one Ge–H H atom shifted to a position between the two Ge atoms, retaining the original Ge–H bond and producing a longer-range Ge–H...Ge interaction. Pathology U4 combines the breaking of a Ge–C bond with migration of a C–H H atom to produce one GeH_2 fragment, one C=C π-bond, and one Ge–H pendant group coordinated to the new C=C double bond.

3.2.3.6. AdamGe33 (Table III, Fig. 8). The two LT pathologies identified for AdamGe33 involve Ge–C bond breaking, the results of which are significantly deformed structures. Pathology L2 breaks one Ge–C bond and migrates the Ge–H H atom to the Ge–C C atom to produce a free methyl group. Pathology L3 breaks a framework C–C bond and migrates one C–H H atom to complete the formation of both a free methyl group and a framework C=C double bond. Three UT pathologies are identified from the 1500 K MD simulations, with the base structure (U1) optimization producing a Ge–Ge bond. Pathology U2 is only slightly higher in energy than U1 but involves a significant rearrangement that retains the original connectivity but shifts a complete methylene (CH_2) group from a framework position to a bridging position between the two Ge atoms. Pathology U3 migrates one Ge–H H atom to the other Ge atom to form a complete GeH_2 fragment, one H-free Ge atom, and a broken Ge–Ge bond. Pathology U4 is similar in geometry to U3 but shifts one local CH_2 H atom to the H-free Ge atom to form a Ge=C double bond and complete valences at all atoms.

3.2.3.7. Diad3Ge22 (Table III, Fig. 8). A single LT pathology is identified for Diad3Ge22 from the 3000 K MD simulations. The Ge and C atom positions in pathology L2 are identical to those of the base tooltip (L1), with the structure difference occurring in the orientation of the two Ge–H H atoms which now point in opposite directions instead of residing on the same side of the tooltip as in L1. The three UT pathologies, all involving only the two Ge–H bonds, feature the breakage of the Ge–Ge bond formed from the optimization of U1. Pathology U2 has both Ge–H H atoms on the same Ge atom, while U3 and U4 contain either one bridging Ge–H...Ge H atom (U3) or both Ge–H atoms pointing in opposite directions on the tooltip surface (U4). Defect U4 is similar in geometry to the LT pathology L2. These accessible pathologies are H-dependent and may not occur in systems where the Ge–H bonds are replaced with Ge–$C_{framework}$ bonds in a larger tooltip handle structure.

3.2.3.8. DCB55AGe (Table II, Fig. 7). No LT pathologies are identified from either the 1500 K or 4500 K MD simulations. Several UT pathologies are identified that

display increasing degrees of tooltip deformation away from the base geometry (U1). The lowest-energy structure (U2) combines Ge–C bond breaking with C–H → Ge–H H migration to form an open structure with a Ge–H bond and a C=C double bond. Pathology U3 finds a single Ge–C bond broken at both Ge atoms which produces two incomplete Ge valences and a single C=C double bond formation, with all other C valences satisfied with H atoms. Pathology U4 is a complete cleavage of the two Ge–C bonds on one side of the tooltip with formation of a C=C double bond to produce an open tooltip structure. These first three pathologies are all significantly more stable than the base UT geometry. Pathology U5 represents a considerable change to the base framework connectivity, combining both the Ge–C bond cleavage and C=C double bond formation in U4 with a single H inversion into the tooltip framework, producing a structure only 1.4 kJ/mol more stable than the base U1 geometry. In contrast, pathology U6 retains much of the base tooltip geometry. Here, two Ge–C bonds are broken on a single Ge, forming a C–C bond at this position. The result is a tooltip with a Ge bound only to one C atom but coordinated to two nearest-neighbor H atoms. Pathology U7 is a more disrupted arrangement of U5, combining Ge–C bond cleavage and C=C double bond formation with H inversion at a different C–H position in the framework. Finally, pathology U8 involves the structural rearrangement of U7 with a second H inversion to produce a close H...H interaction within the tooltip framework.

3.2.3.9. DCB55BGe (Table IV, Fig. 9). No LT tooltip pathologies are identified for L1 at either 1500 K or 3000 K in the MD simulations. The UT geometries are varied in nature but are not properly described as pathologies. Instead, the structures from U1–U10 are largely conformational variations at the base of this tooltip, with the most stable geometry also involving the formation of a Ge–Ge bond (U2). Tooltip U2, as the only structure with a close Ge–Ge interaction, appears to be balancing the formation of this Ge–Ge bond with a constrained framework conformation such that the Ge–Ge bond formation produces a structure only 70 kJ/mol more stable than several of the more open tooltip geometries. The overlay of these structures (Fig. 9) shows that the U2 geometry having the single Ge–Ge bond is unique in its Ge geometry from all other tooltips, which group into more open Ge...Ge faces. The energy range for all 10 of these structures (100 kJ/mol) is low in comparison to the other conformation-specific tooltip series (DCB5Ge) wherein framework constraints place the relative energies over a 160 kJ/mol range.

3.2.3.10. DCB55CGe (Table II, Fig. 7). The only identified LT pathology (L2) is the product of a CH_2 torsion at the base of the tooltip on the opposite side of the C_2 dimer binding position. As a result, this defect would not
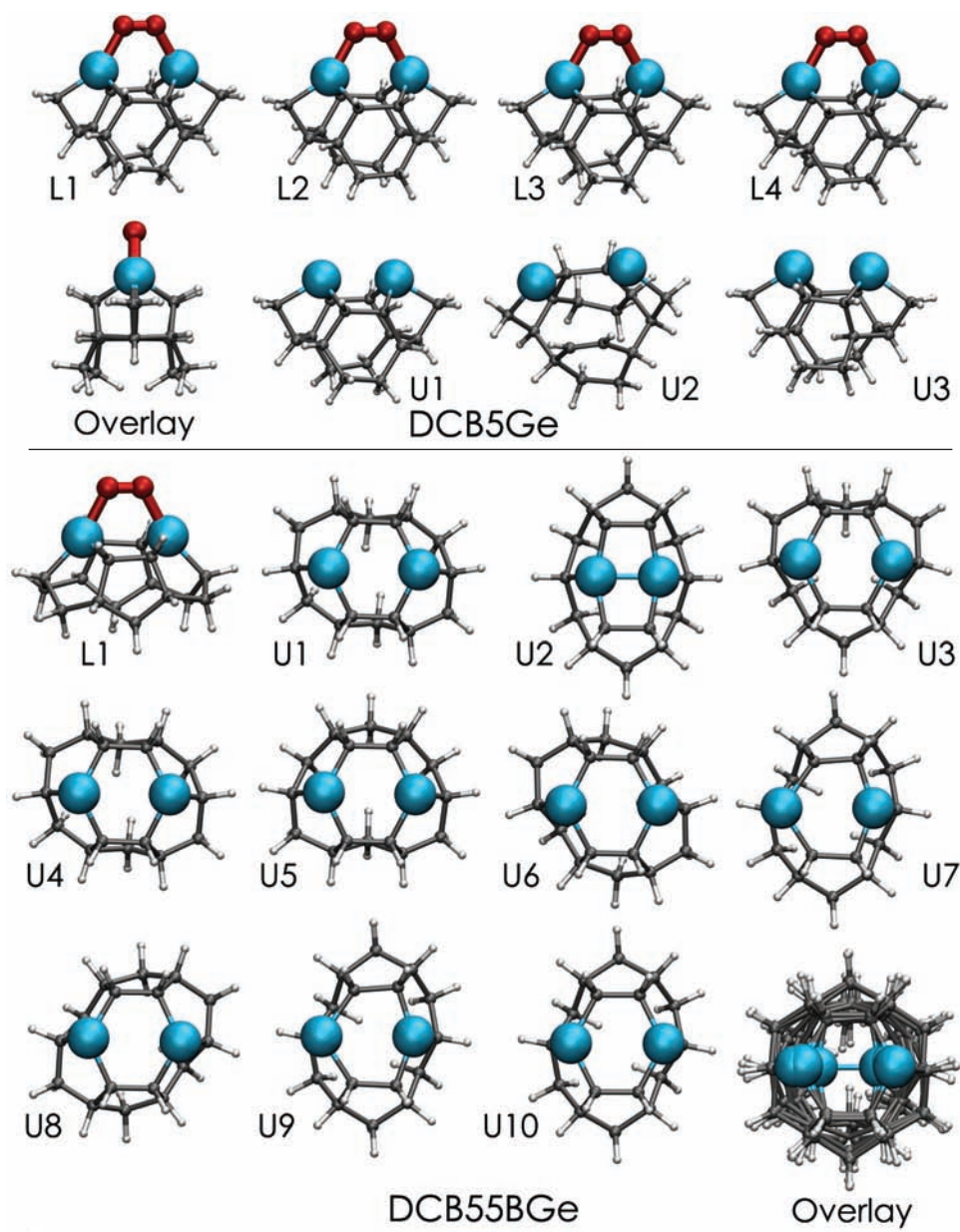
Fig. 9. LT and UT base geometries (L1 and U1, respectively) and identified pathologies for DCB5Ge and DCB55BGe. Carbon dimers (C_2) in the LT structures are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. Germanium atoms are shown in blue. The relative energies of these structures are provided in Table IV. Overlay images are generated from RMSD alignment.

necessarily affect the C_2 dimer binding in base structure L1 if the tooltip was embedded within a larger framework. Three UT pathologies are predicted at higher energies than the base structure (U1). The first of these, by inspection, is the most disruptive, involving H migration to the Ge atoms and C=C double bond formation to produce a greatly deformed structure at moderately higher energy (66 kJ/mol). Pathology U3 is the result of a single C–H → Ge–H H migration. Pathology U4 combines C–C framework bond breaking with H migration to produce an open structure containing a framework C=C double bond and a nearby methyl group.

3.2.3.11. DCB57Ge (Table II, Fig. 7). No LT pathologies are identified for the base tooltip (L1) at either 1500 K or 4000 K in the MD simulations. Three UT pathologies are identified, including one of lower energy. Pathology U2 combines two broken Ge–C bonds on the same side of the tooltip framework with C=C double bond formation. Pathology U3 is similar to U2 but includes a single inverted H atom into the tooltip framework. The opening of the frame from Ge–C bond breaking provides enough space within the tooltip to make this inverted geometry only 70 kJ/mol higher than U2. Pathology U4 combines C–C framework bond breaking with H migration to

produce a free methyl group and a C atom with only two C–C σ-bonds to it.

3.2.3.12. *DCB5Ge (Table IV, Fig. 9)*. The three predicted LT pathologies for DCB5Ge are the result of conformational differences in the free –CH$_2$–CH$_2$- fragments at the tooltip base. With the tooltip framework constraining these –CH$_2$–CH$_2$- groups to a small separation, the conformational energy differences for L2, L3, and L4 are quite large despite the only minor apparent structural changes. The sampling of all conformational differences and the localization of these differences to the base of the tooltip is best represented by the overlay of all four Ln tooltips (Fig. 9).

The UT pathologies represent the extremes in observed types of deformations for all tooltips. Pathology U2 is the result of breaking both Ge–C framework bonds and of the formation of a C=C double bond—the Ge atoms do not gain H atoms from other positions to complete the valence. Pathology U3 is a H-inverted geometry that is included in the tooltip summary because the DCB5Ge framework is flexible enough that this inverted geometry is only 25 kJ/mol higher in energy than the base structure (U1).

3.2.3.13. *DCB65Ge (Table II, Fig. 7)*. The LT 1500 K and 3000 K MD simulations produced no pathologies for this structure. The UT 3500 K simulations produced no pathologies for this structure.

3.2.3.14. *DCB6Ge (Table II, Fig. 7)*. The LT 1500 K and 3500 K MD simulations produced no pathologies for this structure. Energy convergence for the UT structures was not achieved in the allotted time and within the 40-step SCF limit.

3.2.3.15. *DCBIceane7Ge (Table II, Fig. 7)*. The LT 1500 K and 3500 K MD simulations produced no pathologies for the base structure (L1). Three low-lying UT pathologies are identified for the base structure (U1), which itself optimizes with the formation of a Ge–Ge single bond (or close contact). Pathology U2 involves the migration of a Ge–H H atom to the other Ge atom, producing one Ge with a complete valence and one constrained in the RHF calculation to a ground state singlet that lies only 7 kJ/mol higher than the base structure. Pathology U3 combined Ge–C bond breaking with Ge–H → C–H H migration between the same Ge–C atoms. Pathology U4 is a highly deformed geometry that involves Ge–C bond cleavage at both Ge atoms and Ge–H → Ge–H$_2$ H migration to form one C=C double bond, one GeH$_2$ fragment, and one Ge atom bound by only one Ge–Ge bond and one Ge–C bond.

3.2.3.16. *MCB55Ge (Table II, Fig. 7)*. Six alternative LT geometries are identified for MCB55Ge but none are best described as pathologies, as all differ only by the orientations of the four –CH$_2$–CH$_2$- fragments bridging the two Ge atoms to the more constrained framework base. Tooltips L2, L3, L4, L5, and L6 all lie within a narrow energy range. Tooltip L7 is the least stable version of MCB55Ge with all of the –CH$_2$–CH$_2$-fragments oriented

to maximize H…H repulsion on the tooltip framework sides. It is noteworthy that L2 and L7 are both of C_{2v} symmetry and represent the minimum and maximum in the range of energies for the –CH$_2$–CH$_2$-conformations.

A single UT pathology is identified (U2) to produce a pair of structures (U1 and U2) very similar to tooltip C100GeATSr5 (Fig. 6). In MCB55Ge, the energy difference for the symmetric base structure (U1) and more stable asymmetric structure (U2) is only 18 kJ/mol smaller than the same difference for C100GeATSr5. In both, the geometry of one Ge atom is more sp^3-like in nature while the other Ge is more sp^2-like, indicating that this difference in geometry is the result of intramolecular charge-transfer in the singlet-constrained RHF wavefunction to form an anionic Ge (sp^3) and cationic Ge (sp^2) pair within the same molecule. This interpretation is consistent with the Mulliken charge analyses, since the Ge atom charges are identical for the symmetric tooltips but different (approximately 0.8 e$^-$ in both cases) for the asymmetric structures.

3.2.3.17. *TwistaneGe (Table III, Fig. 8)*. A single LT pathology is identified in the 4000 K MD simulation. Pathology L2 breaks one Ge–C$_2$ dimer bond and migrates the associated Ge–H H atom to the C$_2$ dimer to produce a valence-filled Ge-[C$_2$ dimer]-H acetylene structure and one Ge atom with an incomplete valence. This structure is predicted to be 58 kJ/mol more stable than L1. As the defect appears to be largely a result of H migration upon C–Ge bond breaking, the removal of the Ge–H H atoms in a larger framework might prevent this defect from forming. Energy convergence for the UT structures was not achieved in the allotted time and within the 40-step SCF limit.

3.2.4. DC10c (Table 5, Fig. 10)

Seven LT pathologies are identified for the DC10c tooltip, of which two (L2 and L3) are potentially accessible defects. Pathology L2 is the result of a symmetry-breaking of two separate $C_{framework}$–$C_{dimer-binding}$ bonds and the bending of the C$_2$ dimer towards the two broken $C_{framework}$–$C_{dimer-binding}$ C framework atoms to form a highly strained four-member ring in a slightly more open tooltip structure. This defect lies only 3.1 kJ/mol lower in energy than the base tooltip geometry (L1). The previous B3LYP/6-31G(d,p) study on the DC10c tooltip also identified this geometry as a potential defect of concern, but found the transition state barrier for this rearrangement to lie a considerable 243 kJ/mol higher in energy than the L1 base structure, making it inaccessible under proposed operational conditions.[8] A second pathology (L3) involves the breaking of one $C_{framework}$–$C_{dimer-binding}$ bond and $C_{framework}$–H → $C_{dimer-binding}$–H H migration to produce a tooltip with one $C_{framework}$=C double bond and one $C_{dimer-binding}$=C double bond, thereby completing the valences of all atoms in a highly strained geometry. The high-energy pathology
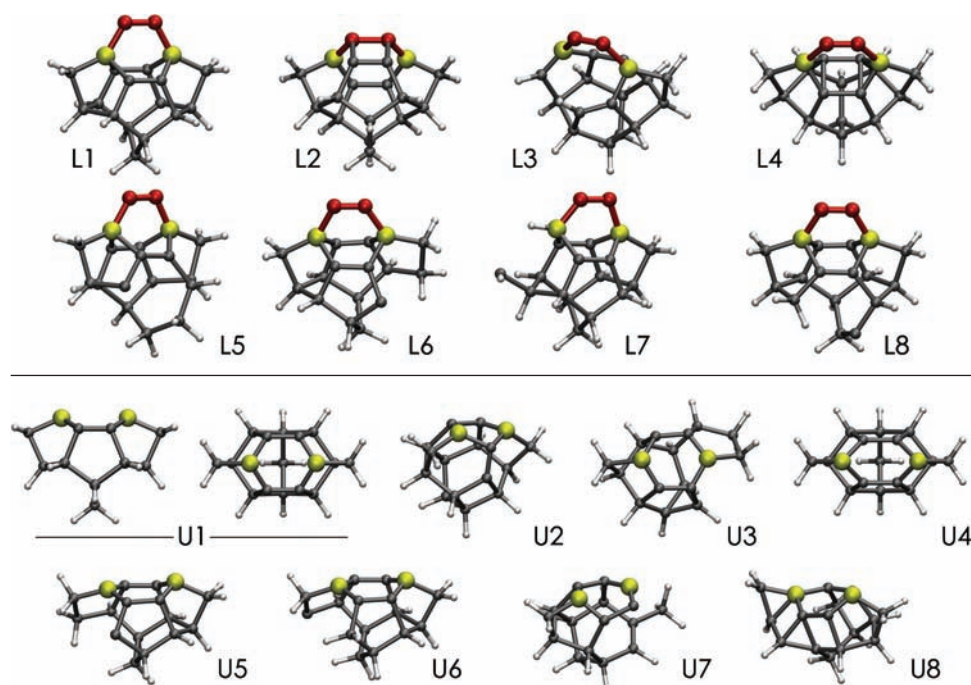
Fig. 10. LT and UT base geometries (L1 and U1, respectively) and identified pathologies for the DC10c tooltip. Carbon dimers (C_2) in the LT structures are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. The C atoms used to bind the C_2 dimer are shown in yellow. The relative energies of these structures are provided in Table V.

L4 is noteworthy for its formation of a four-member ring from the two C=C double bonds on either side of the C_2 dimer binding positions. All other pathologies (L5, L6, L7, L8) involve considerable framework deformations due to framework C–C bond breaking.

Seven UT pathologies are identified for DC10c, of which two are close in energy to the base geometry (U1). Pathology U2 involves the breaking of one [$C_{\text{dimer-binding}}$–C=C] $C_{\text{dimer-binding}}$–C bond and rotation of the associated C=C double bond such that both C_2 dimer binding positions (yellow) are now connected to the same C atom. The binding face of the DC10c tooltip is, despite this significant geometry change, partially retained. Pathology U3 is the result of breaking framework C–C bonds below the C_2 dimer binding surface followed by formation of a three-member ring, nevertheless also preserving the C_2 dimer binding position. Pathology U4 is an interesting defect structure as it preserves the overall UT geometry but breaks the strained conjugation of the six-member ring in favor of two shorter C=C double bonds on either side of the dimer binding position. This defect might be considered the upper-limit geometry on a ring-breathing vibrational mode. The high-energy pathology U7 incorporates a C_2 dimer within the framework after breaking a C–C bond between the framework and one of the dimer binding positions. Pathologies U5, U6, and U8 all, remarkably, completely preserve the geometry of the strained π-system at the C_2 dimer binding position but involve various framework C–C σ-bond rearrangements and H migrations.

4. DISCUSSION

4.1. Unloaded Tooltips and Ge–Ge Bond Formation

Several base geometry UT optimizations (C100GeATS, C100GeCTS, C110GeS, AdamGe22, AdamGe33, Diad3Ge22, DCB55BGe, DCB75Ge, MCB5Ge, DCBIceane7Ge) produced structures with covalent Ge–Ge bonds. These structures are significant for two reasons. First, the formation of a Ge–Ge bond requires either that the tooltips be flexible enough to allow for the two Ge atoms to move within bonding distance of one another or that the Ge atoms already be in close enough proximity in the LT that the removal of the C_2 dimer enables Ge–Ge bond formation without significant motion in the tooltip framework. Second, a stabilizing Ge–Ge bond (and the potential motions within the tooltip to accommodate this bond formation) increases the transition state barrier corresponding to C_2 dimer addition back onto the tooltip. While the Ge–Ge bond may be weak due to the orientation of the orbitals used to form this bond in highly constrained tooltips, the barrier to Ge–Ge bond insertion for these systems is still expected to be higher by some amount than tooltips for which the framework completely prohibits Ge–Ge bond formation (systems which account for the majority of the analyzed tooltips).

The formation of a Ge–Ge bond and the added stability it provides to a UT geometry may, for some systems, be an advantage. Any stabilizing interactions within the open tooltip may serve to increase the barriers to structural

rearrangement, H migration, etc. between the time a C_2 dimer is deposited on a workpiece and the time the tooltip is recharged with a new C_2 dimer. In effect, such bond formations in the Ge-tooltip series are similar in potential applicability to the strained π-framework on the C_2 dimer binding face of the DC10c tooltip.

4.2. More Stable Pathologies and Transition State Barriers

The identification of tooltip pathologies in this MD-driven quantum chemistry study indicate that alternative framework geometries for some cases are more stable than the LT or UT structures designed for use in mechanosynthetic operations. But these tooltip pathologies are similar to structural isomers in any other organic molecule. The identification of a more stable structure does not provide any insight into the energy barrier over which an operational mechanosynthetic geometry must pass in order to convert into a non-operational geometry. Transition state calculations must be performed to provide any insights into the barriers for these conversions. The original transition state barrier calculations for the DC10c tooltip indicated that, although identified pathology L2 is 1.3 kJ/mol more stable than "operational" tooltip geometry L1, the transition state for the L1 → L2 rearrangement is 51 kJ/mol at the B3LYP/6-31G(d,p) level of theory. While some of the failure modes predicted to lie within 100 kJ/mol of the operational geometry for a tooltip appear to be related by a single transition state coordinate, it is clear that many other predicted failure modes involve considerable structure rearrangement of the framework atoms and thus likely are the result of multi-step transition state conversions from one structure to the next.

The NH@H failure mode survey for the tooltips simply provides ground state geometries on the covalent potential energy surface. Identifying from among the failure modes within some energy range which of the operational tooltip structures that are accessible within a thermal regime in a working system is a time-consuming but very important subsequent step in any continued developmental survey of these tooltips.

4.3. Hydrogen-Inverted Tooltip Geometries and Larger Tooltip Frameworks

The MD simulations for each tooltip generated multiple structures at higher energies that have been removed from all but a few of the tooltip summaries. These largely-ignored tooltip pathologies are the results of *hydrogen inversion*, a type of defect that finds one or more H atoms inserting into the cage framework as a result of a methodological mismatch of atom momentum and classical time step in the MD simulations. The problems associated with H motions in classical MD simulations are familiar at all levels of simulation, from protein and DNA dynamics studies in aqueous media at room temperature to these very high temperature/high energy deformation simulations at up-to 4500 K. Three common workarounds for large H atom displacements per time step are

(1) the use of smaller time steps to recalculate the forces on the H atoms,

(2) the artificial increase of the mass of the H atoms to reduce their net displacement over some set time step (such as re-massing H atoms to deuterium or higher), and

(3) the subsuming of the H atoms into the associated "heavy" atom to remove the H motion entirely from the simulation.

A representative set of pathologies from a single tooltip is shown in Figure 11 for DCB65Ge. This tooltip is useful as an example because the tooltip cage is reasonably wide for the small number of atoms, producing an enclosed region within the tooltip that can expand and adjust as one or more H atoms minimize within the framework. The seven identified H-insertion pathologies are shown in Figure 11 with their representative codes (these structures are available as Supplemental Material) and the energies of these structures relative to the optimized loaded tooltip (in kJ/mol). Despite large expected repulsive energies upon insertion, three of these defects lie within 200 kJ/mol of the base structure, a range comparable to the energies of defect geometries in several other cases among the tooltips. Furthermore, the only pathologies identified for the DCB65Ge tooltip are those involving H insertion, indicating that the tooltip framework itself is otherwise quite resistant to structural deformation.

There are two reasons why such structures are not likely to be of benefit in a pathology analysis and thus are excluded in all but a few cases in the tooltip survey. First, the process of H inversion is a known issue of the MD methodology and is, therefore, something better treated by employing methods that remove the large H atom motions across all temperature ranges. Second, framework H atoms in these tooltip structures are, in effect, included in the quantum chemistry analysis to complete the valences of sp^3 carbons that would, in a complete tooltip framework, be replaced by bonds to other atoms as a way to anchor the tooltip onto some workpiece. An example structure is provided in Figure 12. While an ideal tooltip might contain within its base some fragment of a diamondoid lattice that easily provides the means for its inclusion into a larger mechanosynthetic assembly, many of the studied tooltips are not so ideally designed, thereby requiring a connective "bridge" that links the tooltip to some larger structure. In the hypothetical DCB65Ge bridge framework shown in Figure 12, it is seen that all but the C–H H atoms closest to the Ge atoms have been replaced by C–C bonds, removing all but one (559 w) of the predicted H-inversion pathologies from any further defect analysis.

There are a few ways to script the identification and removal of inverted species from the tooltip survey if
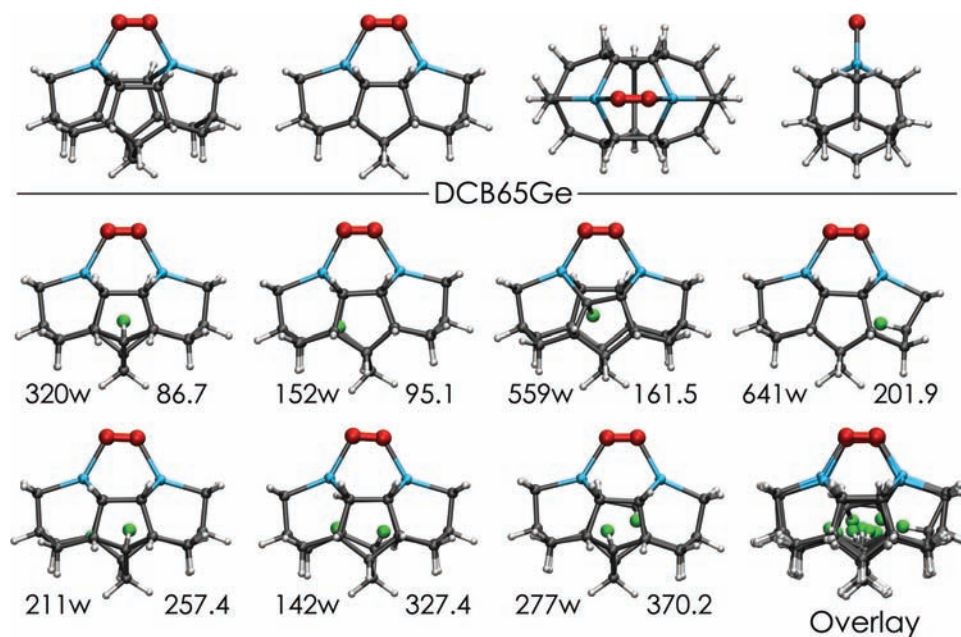
Fig. 11. Predicted H inversion structures for the DCB65Ge (four orientations shown at top) tooltip. Carbon dimers (C_2) in the LT structures are shown in red. Framework C and H atoms are shown in dark grey and white, respectively. Germanium atoms are shown in blue. Inverted H atoms are shown in green. Structure numbers (*NNN*w) and relative energies (in kJ/mol) are provided for each structure.

efforts are not otherwise taken to remove large (inverting) H motions, although each has its potential drawbacks. One approach is to perform H-only molecular mechanics (MM) minimizations on the tooltips before submitting them to the quantum chemistry queue. Complete MM minimizations (all atom MM) in such cases would not be expected to be beneficial because the connectivity of the atoms in the tooltips is predefined in the MM parameter assignment process. In this case, structures will optimize around their covalent connectivity and not, when the atomic positions are sufficiently different, the proximity of atoms to one another—which is the type of optimization obtained by a quantum chemistry analysis when no particular connectivity is assumed. In the absence of a procedure to adjust the H atom positions for MD time step samples where the H atoms are either within a cage framework or well within the sum of the van der Waals radii of a nearby atom, the quantum chemistry analysis must be performed to identify a structure with a more proper geometry than the MD structure. After that, the separations between atoms

are expected to be far more reasonable and the removal of H-inversion geometries can be performed algorithmically by identifying H atoms residing within the tooltip frameworks.

The use of increased mass (deuteration) to reduce the prevalence of H-inversion geometries in the MD simulations was tested on the DC10c LT structure at 3000 K. (A 3500 K AdamGe22 MD simulation was also performed, but this structure does not contain a framework cage that can accommodate an inverted H atom without significant structural rearrangement). While the overall number of H-inverted defects in LT DC10c dropped slightly, a 3000 K simulation is, unsurprisingly, disruptive enough to an organic framework to still produce many H inversion pathologies. If such structures are to be avoided but not removed from the quantum chemistry analysis, alternative approaches at the force field or MD level should be considered.

4.4. The NH@H Network and "Best-Practices" Considerations

The NH@H tooltip pathology survey generated a considerable amount of data and, after the analysis of the resulting data, yielded a selection of tooltip pathologies to serve as the basis for subsequent studies of tooltip defect pathways. The most important benefit gained from the NH@H approach was the massive distribution of the time-consuming quantum chemistry optimizations to external resources (i.e., home computers). As only the structure distribution, and not the quantum chemistry calculation, is
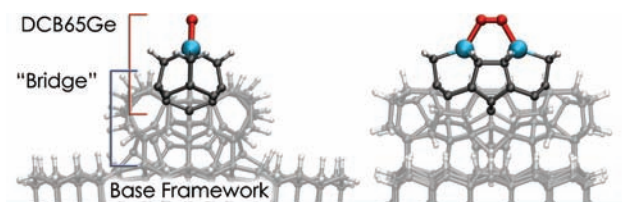

Fig. 12. The DCB65Ge tooltip within a larger diamondoid framework, showing how additional framework C–C connections remove H atoms that underwent H inversion in the MD simulations (for the base DCB65Ge structure, see Fig. 11).

distributed across the network, the success of an examination is resource-dependent in a manner different from a dedicated cluster, time-share, or national laboratory queue-based facility. The speed of a calculation and, ultimately, the quality of the calculation for a particular analysis are dictated by the quality of the computers owned by the participants. The RHF/3-21G and RHF/STO-3G calculations were not selected for their accuracy, but instead were selected because they provide *ab initio* answers that virtually any home computer (ca. 2006) would have the RAM, CPU performance, and hard drive capacity to run successfully. While the dedicated participants in this first NH@H study may not have had any reservations about committing full resources to these calculations, it is clear that submitting a tooltip the size of the C100GeATD, C100GeCTS, or any of the larger structures to a home computer for geometry optimization at an MP2/6-311+G(2d,p) level of theory would very likely risk depleting those machines of all storage space, since the estimated storage requirements for such calculation range as high as 500 GB based on previous trials. Such large runs also risk depleting participants of all patience as single geometry optimization cycles would execute over the course of days instead of minutes. The use of a survey of representative available computers at the start of a series of quantum chemistry studies can be of great benefit in identifying the constraints a researcher must place on their investigation.

4.5. Identified Minima versus Accessible Minima

In the absence of transition state calculations or MD simulations on larger tooltip frameworks to remove degrees of freedom in some atoms, it is not known how many of these tooltips can be ignored simply for energetic reasons—either due to large rearrangement barrier energies or due to one-step inaccessibility because of the presence of multiple barriers to complete a rearrangement to a predicted minimum. Further, as demonstrated in the calculated energy differences for the C110GeS-L2 tooltip, the relative energies of these structures may be very dependent on the level of theory used for their calculation. While classes of deformation can be catalogued, these geometries only serve as a structural basis. Higher levels of theory are required for the prediction of relative energies and, subsequently, transition state barriers to rearrangements involving only a single coordinate or the modeling of multiple steps along a sequence to more significant rearrangements.

4.6. Hydrogen Migration

In most stable tooltips, H migration is assumed to be the most accessible route to the formation of inoperative structures (what previous studies[1] have called "hydrogen poisoning"). Among the pathologies identified in the NH@H study, only two of the LT defects (C100GeATD-L2 and C100GeATSr5-L3) contain H-migrated Ge–H

bonds (the addition of valence electrons to the Ge atoms by H migration then affects the recharging step in any mechanosynthetic procedure). Only one of these finds the H atom bound to the C_2 dimer itself (C100GeATD-L2). With the deposition of a C_2 dimer, both Ge atoms in each tooltip contain unfilled valences, providing sites for potential H migration. A number of UT pathologies are the result of Ge–H H migration, including DCB55AGe-[U2], DCB55CGe-[U2,U3], DCBIceane7Ge-[U2,U3,U4], C100GeATD-[U2], C100GeS-[U2], AdamGe22-[U2,U4], AdamGe33-[U3], Diad3Ge22-[U2], and C100GeCTS-[U2]. While the Ge–H bonds are predicted to be weaker than the C–H bonds even in strained systems, the energy differences need only be enough to enable fractional Ge–H H migration either to influence the recharging of a tooltip with a C_2 dimer or, possibly, to enable some other structural transformation in the tooltip to occur. One possible method for reducing the potential for H migration to the Ge atoms is to employ tooltips that, upon C_2 dimer deposition, develop strained Ge–Ge bonds as observed in several of the tooltip structures and, therefore, increase the barrier to migration by requiring intramolecular Ge–Ge bond scission before Ge–H H-bond formation can occur.

4.7. Ge–C Framework Bond Breaking

All of the tooltips in this study are designed to facilitate dimer deposition and tooltip retraction with the only modification to the covalent framework of the tooltip being the loss of the C_2 dimer. Among the Ge-containing tooltips, four LT pathologies (C100GeATSr5-[L3], C110GeS-[L3], AdamGe22-[L4], AdamGe33-[L2]) are identified with only Ge–C$_{framework}$ bond breaking (as opposed to more significant deformations involving additional framework modification), while numerous UT structures contain this defect mode. The few LT forms are noteworthy for their small relative energy differences from ground state geometries. Only C100GeATSr5-[L3] lies above 100 kJ/mol from its ground state geometry. The UT structures run the gamut from significantly more stable (e.g., DCB55AGe-[U2] is 119 kJ/mol more stable than its base geometry) to significantly less stable (e.g., C100GeATD-[U4] is 418 kJ/mol less stable than its base geometry). These calculated energies are less reliable than those computed at higher levels of theory and those combined with proper transition state barriers, but the fact that a large range is observed makes such defect modes of considerable importance as tooltip analyses are performed either to redesign or to remove candidate tooltips from consideration for performing mechanosynthetic operations.

As part of the operation of a tooltip deposition process, the possibility of Ge–C$_{framework}$ bond breaking[12] is another important issue at the point where the presumably single covalent bonds between the C_2 dimer and the workpiece are formed and the tooltip is ready to be retracted. The

20

J. Comput. Theor. Nanosci. 8, 1–23, 2011

RESEARCH ARTICLE

formation of single C–C bonds between the C_2 dimer and workpiece does not result in the breaking of the Ge–[C_2 dimer]–Ge Ge–C bonds, as the C_2 dimer is itself bound to each tooltip as a strained triple-bond system and, upon workpiece bond formation, is a less strained double-bond system. (One might consider it an ethene functionalized with tooltip and workpiece.) It is at this point in the deposition process that the relative strength of the Ge–$C_{framework}$ tooltip bonds compared to the Ge–C_{dimer} bonds becomes the determining factor for whether or not Ge–$C_{framework}$ bond breaking is an important defect mode, as has previously been observed in simulations during step 3 of a C–Ge tooltip retraction event during a hydrogen abstraction tooltip recharge sequence.[12] The defects identified in the NH@H study with broken Ge–$C_{framework}$ bonds may not themselves be accessible in isolation, but the additional bonding modes and resulting strain in the entire system as part of a mechanosynthetic operation makes such defect modes important in the overall design analysis of a diamondoid structure that is to be fabricated. The study of such defect modes requires the modeling of the deposition process, which can be performed at a considerable cost in computational resources.[12]

4.8. Combined Ge–$C_{framework}$ Bond Breaking and C=C π-Formation

The formation of broken Ge–C/C=C bond pathologies are noteworthy, both in the context of the operational issue described above and in the manner by which a symmetric bond breaking in these tooltips can produce very stable geometries that then require large transition state barriers to be present for the mechanosynthetic operability of the tooltip. This class of defect mode occurs in several of the UT structures (AdamGe22-[U4], DCB55AGe-[U4,U5,U7,U8], DCBIceane7Ge-[U4], DCB57Ge-[U2,U3], DCB5Ge-[U2]) but, remarkably, did not occur in any of the LT structures. The presence of a conjugated unit on the Ge atoms (C_2 dimer) might be expected to stabilize Ge radical formation in the symmetric bond scissions by pairing these electrons in an extended π-system, but no such structures were identified in the quantum chemistry optimizations though they may have been present in some form during the MD simulations.

As with the single Ge–$C_{framework}$ bond breaking defects described above, the paired Ge–C/C=C pathologies may be of greater concern during a mechanosynthetic deposition than among structures *en route* to deposition. In a deposited [$C_{workpiece}$]–C=C–[$Ge_{tooltip}$] structure, the formation of C=C π-system pathologies becomes possible if the [$C_{workpiece}$–C_{dimer}]–$Ge_{tooltip}$ C_{dimer}-Ge bond is of comparable or greater strength than the $Ge_{tooltip}$–$C_{framework}$ Ge–C bonds such that the retraction of the tooltip results in the symmetric Ge–C bond breaking and C=C π-bond formation with retention of the C_{dimer}–$Ge_{tooltip}$ bonds.

4.9. Ge–Ge Bond Formation

This class of structures is described in section *Discussion I* above. As noted in the Hydrogen Migration section, the formation of these strained bonds in the UT structures may not be detrimental to tooltip operation but may, by increasing the barrier to other defect modes, serve as a form of stabilization during the time between deposition and C_2 dimer recharging.

4.10. C_2 Positional Variation

Remarkably, despite the high temperatures and otherwise large deformations identified in many of the tooltip structures, specific structural deformations at the Ge–[C_2 dimer]–Ge position were identified in only a few cases from the MD-based quantum chemistry optimizations. These structures include C100GeATD-[L2], C100GeS-[L2,L5], AdamGe22-[L2,L3,L4] and TwistaneGe-[L2], among which only those for C110GeS and AdamGe22 were considered as part of tooltip pathologies in the original analysis.[1] The results of the RHF/3-21G optimizations for the carbene pathologies in C100GeS and AdamGe22 are noteworthy for the reversal in their relative energies to the base structures compared to the original B3LYP/6-311 + G(2d,p)//B3LYP/3-21G* survey, which placed C100GeS-[L2] 54 kJ/mol higher in energy than its base structure and AdamGe22-[L2] 44 kJ/mol higher than its base structure. Here and generally, both the quality of the basis set and the inclusion of electron correlation (by way of the B3LYP density functional) are expected to make a considerable difference in the accuracy of the estimated relative energies.

4.11. Conformational Differences

In most rigid tooltip designs, the identification of conformational minima is interesting but otherwise not of significance, as conformational flexibility at the tooltip base is, like H inversion, removed from all structures as a result of embedding the structure within a larger and more constraining tooltip framework. Nevertheless, the results of the DCB55BGe UT simulations are noteworthy for the predicted changes to Ge...Ge separation in the many conformations and the impact that the control over conformation might have under operational conditions. Several of the DCB55BGe UT structures are constrained to conformations that separate the Ge atoms such that Ge–Ge bond formation, as predicted for DCB55BGe-[U2], does not occur. If it is found in subsequent studies that the barrier height for Ge–Ge bond breaking and C_2 dimer reloading is too high for proposed recharging pathways using one conformation, then the design of a larger tooltip framework that confines the end of the DCB55BGe (or similar) tooltip to a conformation that maximizes Ge...Ge separation could provide a way to improve the recharging process. This

approach to tooltip design optimization via conformational control of the base has not been considered in previous studies and is one of the more interesting results to emerge from this initial NH@H study.

5. CONCLUSION

The study of defect structures is a key step in the development cycle for DMS tooltips because the potential energy surface for rearrangement will determine the utility of a tooltip over either single or multiple positioning and deposition cycles. The automation of the identification of defects and possible operational-to-defect rearrangements is difficult to perform due to
(1) the computational cost of such surveys at appropriately high levels of theory,
(2) the difficulty in confirming that every possible failure mode that is thermally or operationally accessible has been considered, and, importantly,
(3) the absence of experimental data to identify which defects and, subsequently, which proposed pathways, are most likely.

While these same issues are often handled statistically (partial yields) or by modification of experimental conditions in chemical synthesis, a failure in a single mechanosynthetic operation may corrupt every subsequent step in a fabrication process. This situation is of additional concern as the possibility exists that a fabrication process may not be directly monitorable at every step, instead relying on characterization after some number of operations have been performed to determine the success of a partial fabrication cycle.

The procedure employed in the NH@H study, which combines high-energy (deforming) MD simulations with quantum chemistry optimizations over a distributed computing platform, provides a mechanism for sampling a wide range of both accessible and inaccessible pathologies. Multiple approaches to initial structure sampling can be performed through the use of an empirical potential in the MD simulation to produce highly deformed starting geometries. An alternative to high temperatures and short time steps is lower temperatures and longer time steps, resulting in runs that produce structures deformed enough so that quantum chemistry optimization cannot restore the atomic arrangement to its starting ("ready") geometry. The use of very high temperatures in the MD simulations produces deformations which are based on initial connectivity and associated force constants, thus the high temperatures enable the sampling of geometries that are far from the geometry of the ready tooltip. Despite the considerable deformations resulting from the MD simulations, a number of tooltips retained much of their starting structure or optimized in the MPQC calculations back to their ready tooltip geometries. Just as a number of approaches in the MD simulations can be employed to generate structures

for quantum chemistry analysis, the use of the MPQC program provides researchers with a wealth of high-level quantum chemistry methods. The quality of the calculations performed is a function of the available computational resources. In dedicated facilities (where the NH@H platform would also work as a simulation-and-queue system), the limits of available resources are known. In distributed computing projects relying on the donation of cycles by participants and their computers, the balance of theory and resources (which may vary considerably among participants) must be an important factor during planning.

Identification of stable tooltip pathologies should be the first validation process after the initial identification of candidate tooltips, but still represents just one step in a more complete protocol for the characterization of potentially viable tooltip designs intended to be proposed for experimental consideration. With one or more pathologies identified, the next step is the identification of transition state geometries connecting ready and defect geometries, where single transition states may connect two structures or multi-step transition pathways may be required to connect intermediate geometries to "ready" and defect geometries in both the loaded and unloaded cases.

References

1. R. A. Freitas, Jr., D. G. Allis, and R. C. Merkle, *J. Comput. Theor. Nanosci.* 4, 433 (**2007**).
2. R. A. Freitas, Jr. and R. C. Merkle, *J. Comput. Theor. Nanosci.* 5, 760 (**2008**).
3. N. Oyabu, O. Custance, I. Yi, Y. Sugawara, and S. Morita, *Phys. Rev. Lett.* 90, 176102 (**2003**).
4. Y. Sugimoto, P. Pou, O. Custance, P. Jelinek, M. Abe, R. Perez, and S. Morita, *Science* 322, 413 (**2008**).
5. A.-S. Duwez, S. Cuenot, C. Jérôme, S. Gabriel, R. Jérôme, S. Rapino, and F. Zerbetto, *Nature Nanotechnol.* 1, 122 (**2006**).
6. R. A. Freitas Jr., US Patent No. 7,687,146 March (**2010**).
7. R. C. Merkle and R. A. Freitas, Jr., *J. Nanosci. Nanotechnol.* 3, 319 (**2003**).
8. D. G. Allis and K. E. Drexler, *J. Comput. Theor. Nanosci.* 2, 45 (**2005**).
9. J. Peng, R. A. Freitas, Jr., R. C. Merkle, J. R. Von Ehr, J. N. Randall, and G. D. Skidmore, *J. Comput. Theor. Nanosci.* 3, 28 (**2006**).
10. B. Temelso, C. D. Sherrill, R. C. Merkle, and R. A. Freitas, Jr., *J. Phys. Chem. A* 110, 11160 (**2006**).
11. B. Temelso, C. D. Sherrill, R. C. Merkle, and R. A. Freitas, Jr., *J. Phys. Chem. A* 111, 8677 (**2007**).
12. D. Tarasov, N. Akberova, E. Izotova, D. Alisheva, M. Astafiev, and R. A. Freitas, Jr., *J. Comput. Theor. Nanosci.* 7, 325 (**2010**).
13. J. Peng, R. A. Freitas, Jr., and R. C. Merkle, *J. Comput. Theor. Nanosci.* 1, 62 (**2004**).
14. A. R. Van Buuren, S.-J. Marrink, and H. J. C. Berendsen, *J. Phys. Chem.* 97, 9206 (**1993**).
15. D. Spoel, A. R. Buuren, D. P. Tieleman, and H. J. C. Berendsen, *J. Biomol. NMR* 8, 229 (**1996**).

16. B. Helfrich, Distributed and Grid Computing—Science Made Transparent for Everyone, Principles, Applications, and Supporting Communities, edited by M. Weber, Rechenkraft.net, Marburg (**2008**), Chap. 2, pp. 2–7.

17. Berkeley Open Infrastructure for Network Computing, http://boinc.berkeley.edu/

18. C. L. Janssen, I. B. Nielsen, M. L. Leininger, E. F. Valeev, and E. T. Seidl, The Massively Parallel Quantum Chemistry Program (MPQC), Version 2.3.1, Sandia National Laboratories, Livermore, CA, USA (**2004**).

19. J. S. Binkley, J. A. Pople, and W. J. Hehre, *J. Am. Chem. Soc.* 102, 939 (**1980**); M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, and W. J. Hehre, *J. Am. Chem. Soc.* 104, 2797 (**1983**); K. D. Dobbs and W. J. Hehre, *J. Comput. Chem.* 7, 359 (**1986**).

20. W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* 2657 (**1969**); W. J. Hehre, R. Ditchfield, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* 2769 (**1970**); W. J. Pietro, B. A. Levy, W. J. Hehre, and R. F. Stewart, *J. Am. Chem. Soc.* 19, 2225 (**1980**).

21. E. Lindahl, B. Hess, and D. van der Spoel, *J. Mol. Model.* 7, 306 (**2001**).

22. W. F. van Gunsteren and H. J. C. Berendsen, Groningen Molecular Simulation (GROMOS) Library Manual, BIOMOS B. V.: Nijenborgh 16, Groningen, The Netherlands (**1987**).

23. Y. Lei, H. Li, R. Zhang, and S. Han, *J. Phys. Chem. B* 108, 10131 (**2004**).

24. A. A. Strzelczyk, J. Cz. Dobrowolski, and A. P. Mazurek, *J. Mol. Struct. (Theochem)* 541, 283 (**2001**).

25. D. G. Allis, A. M. Fedor, T. M. Korter, J. E. Bjarnason, and E. R. Brown, *Chem. Phys. Lett.* 440, 203 (**2007**).

RESEARCH ARTICLE