Topology of Cyclooctane Energy Landscape

Shawn Martin,[1, a)] Aidan Thompson,[2] Evangelos A. Coutsias,[3] and Jean-Paul Watson[4]

[1)] *Computer Science and Informatics, Sandia National Laboratories,Albuquerque, NM,87185-1316*

[2)] *Multiscale Dynamic Materials Modeling, Sandia National Laboratories, Albuquerque,NM,87185-1316*

[3)] *Department of Mathematics, University of New Mexico,Albuquerque,NM, 87131*

[4)] *Discrete Mathematics and Complex Systems, Sandia National Laboratories, Albuquerque,NM,87185-1318*

(Dated: 16 April 2010)

Understanding energy landscapes is a major challenge in chemistry and biology. Although a wide variety of methods have been invented and applied to this problem, very little is understood about the actual mathematical structures underlying such landscapes. Perhaps the most general assumption is the idea that energy landscapes are low-dimensional manifolds embedded in high-dimensional Euclidean space. While this is a very mild assumption, we have discovered an example of an energy landscape which is non-manifold, demonstrating previously unknown mathematical complexity. The example occurs in the energy landscape of cyclooctane, which was found to have the structure of a reducible algebraic variety, composed of the union of a sphere and a Klein bottle, intersecting in two rings.

[a)]Electronic mail: smartin@sandia.gov

I. INTRODUCTION

The energy landscape of a molecule can provide insight into that molecule's structure and function. Numerous approaches have been employed to understand the nature of these landscapes for different molecules[1,2], and some notable discoveries have been made, including the hypothesized existence of funnel-like structures that serve to guide complex conformational changes in proteins[3]. In general, however, energy landscapes remain poorly understood. Experimental methods such as nuclear magnetic resonance (NMR) and x-ray crystallography cannot easily detect atomic-scale conformational motion, and computational methods are quickly overwhelmed by the dimensionality of the problem ($3N$ for N atoms). As a result, very little is known about the basic mathematical structure underlying energy landscapes.

One of the approaches that has been applied in the study of energy landscapes is dimension reduction[4–11]. By representing a molecule with N atoms in Cartesian space, we can encode a given conformation with $3N$ variables. Since bond lengths, angles, steric constraints, *et cetera* influence the positions of the atoms, there are numerous relationships between the $3N$ variables. Dimension reduction can be used to gain insight into such relationships, ideally obtaining a lower dimensional ($\ll 3N$) but equivalent representation of the conformations. Of course dimension reduction is itself a difficult problem, requiring different assumptions depending on the approach taken. The standard approach is Principal Component Analysis (PCA)[12] and assumes a linear structure. More general approaches are non-linear and assume an orientable manifold structure[13,14].

While dimension reduction has provided insight into the study of molecular motion, there is no reason *a priori* to assume motion constrained to a manifold, either linear (as in the case of PCA) or non-linear. In fact, we have discovered an example of an energy landscape that is non-manifold. Our example uses cyclooctane, a cyclic alkane used in the manufacture of plastics, adhesives, fibers and coatings. This relatively simple molecule has been studied in chemistry for over 40 years[15,16], and has attracted particular attention in computational chemistry[11,16–20]. It has a surprisingly complex potential energy landscape, and a definitive characterization of the conformation space has remained elusive[11,16,20].

By applying techniques from the fields of computational algebraic geometry and topology, as well as non-linear dimension reduction, we have uncovered new insights into the energy landscape of cyclooctane. We find that the mathematical structure underlying the energy

landscape is a reducible algebraic variety, composed of the union of a sphere and a Klein bottle (a non-orientable manifold), intersecting in two rings. This result reveals previously unrealized complexity in molecular conformation and points towards the use of emerging tools that can provide additional understanding of energy landscapes.

II. BACKGROUND

A. Cyclooctane

Cyclooctane is a saturated eight-member cyclic compound with chemical formula C_8H_{16}. Cyclooctane has received attention in computational chemistry because it has multiple conformations of similar energy, a complex potential energy landscape, and significant steric influence from the hydrogen atoms on preferred conformations[15–17]. Cyclooctane is also simple enough that there are enumerative algorithms available which can provide a dense sampling of the conformation space[20,21]. These algorithms show from first principles that the conformation space has two degrees of freedom, suggesting that the space is a surface (but not necessarily a manifold).

Using dimension reduction methods, we have analyzed the cyclooctane conformation space[11]. In our previous analysis, we used *trans*,*trans*-1,2,4-trifluorocyclooctane. In the present analysis, we use pure cyclooctane (wherever necessary, we have repeated our previous efforts using pure cyclooctane). We generated a dataset of 1,031,644 cyclooctane conformations, enumerated using a triaxial loop closure algorithm set to assume fixed bond length and angles[21]. Each conformation is placed in Cartesian space via the 3D position coordinates of the atoms in the molecule. The conformations are aligned to a reference conformation such that the Eckart conditions are satisfied[22], and the final positions of a given conformation are concatenated to obtain a vector in $\mathbb{R}^{72}$. The resulting collection is a dataset $\{\mathbf{x}_i\}_{i=1}^{1,031,644} \subset \mathbb{R}^{72}$ which is presumed to describe a surface. Repeating our previous work[11], we applied the Isomap dimension reduction algorithm[14]. Our analysis is summarized in Figure 1.

B. Terminology

At this point, we digress for a moment to clearly define our use of the terms "surface" and "landscape." In mathematics, the term *surface* is used to describe an object with two degrees of freedom, such as a 2D manifold, or a 2D algebraic variety. This is how we use the term surface. The term surface is used more loosely in the study of molecular science. In particular, the term Potential Energy Surface (PES) is often used in the study of energy landscapes. Here the term surface has no connection to 2D, but rather refers to the high dimensional function (typically $\gg 2$) underlying the free energy landscape. To confuse the matter further, the term energy landscape is also loosely defined, in that it can be used to refer to the free energy landscape (sometimes called the free energy surface) or the PES. We use the term energy landscape to refer to the PES.

C. Algebraic decomposition and triangulation

As in our previous analysis[11], we found that the Isomap representation of the cyclooctane conformation space is a 3D visualization of an object that requires 5D for a full representation. The conformation space appears to be the union of a sphere and an hourglass intersecting in two rings, but what are we missing by using 3D instead of the necessary 5D in Figure 1(c)?

To answer this question, we investigated the intersection rings by using local dimension estimation. We identified 2D and 3D neighborhoods within the dataset[23]. Perhaps unsurprisingly, the 3D neighborhoods were in the vicinity of the intersection rings in the Isomap reduction. However, these neighborhoods revealed an unexpected geometry. When projected using local PCA, the 3D neighborhoods of the intersection rings had the non-manifold geometry of two intersecting planes, as shown in Figure 2(a)-(c). We modeled the 3D neighborhoods by fitting two planes through the data points using quadratic polynomials. These polynomials were fit using least squares optimization, with constraints to guarantee that the solution factored as two intersecting planes (an algebraic decomposition). Although the minimization problem was nonlinear, we were able to solve it using the Singular Value Decomposition (SVD) followed by an eigenvalue decomposition[23]. An example of two fitted planes is shown in 2(c).

Our characterization of the intersection rings indicates that the cyclooctane conformation space has non-manifold self-intersections known as algebraic double curves. These singularities can be removed by decomposing the space into the sphere and hourglass seen in Figure 1(c). To perform this decomposition, we used point set triangulation. To reduce the computational effort for the triangulation, we first subsampled from the approximately 1M conformations to obtain a reduced dataset containing 6,040 conformations (used in Figures 1 and 2). The subsampling was performed in $\mathbb{R}^{24}$ using only ring atom coordinates such that no two points in the dataset were closer than $\epsilon = 0.12$ Euclidean distance in $\mathbb{R}^{24}$ (≈ 0.05 Å root mean square distance). The triangulation was obtained using a modified version of an incremental surface reconstruction algorithm for high dimensional data[23,24]. A triangulation of our 6,040 point dataset in shown in Figure 2(d). We performed the entire subsampling and triangulation process 5 times to verify our calculations. We produced 5 different triangulations with 6,040 ($\epsilon = 0.12$); 7,114 ($\epsilon = 0.11$); 8,577 ($\epsilon = 0.10$); 10,503 ($\epsilon = 0.09$); and 13,194 ($\epsilon = 0.08$) vertices.

III. RESULTS

A. Topological classification

We verified the accuracy of the 5 triangulations by computing Betti numbers. Betti numbers are topological invariants which quantify large scale features of a space[25]. The first Betti number counts the connected components of the space; the second Betti number counts the non-contractible loops; and the third Betti number counts enclosed volumes. We computed Betti numbers using the Plex toolbox (www.comptop.stanford.edu) and Linbox (www.linalg.org). We used Plex to compute boundary operator matrices and Linbox to compute matrix ranks. The Betti numbers are obtained from the matrix ranks.

For each of the 5 triangulations, we obtained Betti numbers (1,1,2). Betti numbers can be used to classify compact connected (manifold) surfaces without boundary, but since the cyclooctane surface is non-manifold, the Betti numbers (1,1,2) are uninformative. However, using our triangulation of the conformation space, we decomposed the object into the two components apparent in Figure 2(d): the outer sphere and the enclosed hourglass. Both of these objects were found to be compact connected (manifold) surfaces without boundary,

each sharing points on the two intersection rings.

Unsurprisingly, the Betti numbers of the spherical component were (1,0,1), which are the Betti numbers of a sphere. Surprisingly, the Betti numbers of the hourglass were (1,1,0). These are the Betti numbers of a Klein bottle, which is a compact connected non-orientable surface without boundary. There are many ways to describe a Klein bottle, but for our purposes the best way uses two Möbius strips. A Möbius strip is a non-orientable surface with boundary which can be obtained by taking a rectangular piece of paper, giving one of the ends a half-twist, then gluing the ends together. A Möbius strip has only one side and one edge. A Klein bottle can be obtained by gluing two Möbius strips together along their edges. This operation cannot be performed in 3D without tearing the resulting surface. Hence the Klein bottle cannot be embedded in less than 4D. This fact agrees with our previous observation that 5D is necessary to fully capture cyclooctane conformation space.

B. Dimension reduction

To further understand the cyclooctane conformation space, we examined projections from torsion space. Any molecule can be represented using internal dihedral angles known as torsion angles. In the case of cyclooctane, we have 8 carbon atoms in a ring so that we can describe the relative positions of the ring atoms using 8 torsion angles. Further, there is a homeomorphism (topological equivalence) of the conformation space using these 8 torsions to the conformation space we have previously employed (in $\mathbb{R}^{72}$). This homeomorphism arises from the fact that the 8 torsions can be used to place the 8 ring atoms in Cartesian space, and that the positions of the hydrogens can then be determined according to minimum energy[11].

We used distance geometry to derive 5 canonical conformations in terms of torsion angles[26]. Let u, v, w be defined such that

$$\begin{aligned}
\cos u &= (1 - \sqrt{2} - \cos\theta_b)/(1 + \cos\theta_b) \\
\cos v &= \cos^2\theta_b / \sin^2\theta_b \\
\cos w &= (3\cos^2\theta_b - 1)/\sin^2\theta_b,
\end{aligned}$$

where θ_b is the (fixed) bond angle, taken by us to be 115°.

Using u, v, w, we obtain conformations

$$\begin{aligned}
\mathbf{c}_r &= (u, -u, u, -u, u, -u, u, -u)^T \\
\mathbf{b}_1 &= (0, v, 0, -v, 0, v, 0, -v)^T \\
\mathbf{b}_2 &= (v, 0, -v, 0, v, 0, -v, 0)^T \\
\mathbf{c}_1 &= (v, 0, -v, w, -v, 0, v, -w)^T \\
\mathbf{c}_2 &= (v, -w, v, 0, -v, w, -v, 0)^T.
\end{aligned}$$

These conformations are known as the crown ($\mathbf{c}_r$), boat ($\mathbf{b}_1, \mathbf{b}_2$), and chair ($\mathbf{c}_1, \mathbf{c}_2$) conformations[15], shown in Figure 3(a). We have discovered that they give an orthogonal basis for the 5 most important dimensions in the 8D cyclooctane torsion space. We refer to this basis as the crown-boat-chair basis. The crown-boat-chair basis can be used to understand the geometry of the cyclooctane conformation space, shown in Figure 3(b-h): it can be used to reproduce analytically the previous result obtained using Isomap (b); demonstrate how the Klein bottle can be decomposed into two Möbius strips (e,f); and provide a fully 2D reduction of both the sphere (c,d) and Klein bottle (g,h) components.

C. Conformation space

Hendrickson identified 10 conformations in his seminal work on cyclooctane[15]: crown (Cr); chair-chair (CC); twist-chair-chair (TCC); boat (B); saddle, also known as twist-boat (TB); boat-boat (BB); boat-chair (BC); twist-boat-chair (TBC); chair (C); and twist-chair (TC). Each of Hendrickson's 10 conformations have certain symmetries when we consider ring atoms. These symmetries are obtained by rotation and reflection of the conformation in physical space. (When such a change occurs as a result of molecular motion, i.e. by changes in torsion angles, it is known as a pseudorotation.) One way to quantify these symmetries is by examining torsion angles. The crown conformation, for example, is highly symmetric and can exist in only two states: $\pm\mathbf{c}_r$. The boat conformation is also very symmetric but can exist in four states: $\pm\mathbf{b}_1$, $\pm\mathbf{b}_2$. In general, conformations on the spherical component of the cyclooctane space have the symmetry $(t_1, t_2, t_3, t_4, t_1, t_2, t_3, t_4)$ and absent additional symmetry will exist in eight states. Conformations on the Klein bottle component have no such constraint and absent additional symmetry will exist in sixteen states.

In Figure 4 we have located each instance of Hendrickson's 10 conformations in our representation of the cyclooctane conformation space. Also shown in Figure 4 are two additional conformations, which we call peak (P) and saddle (S). P and S occur on the intersection rings and within the set of intersection ring conformations they are energy maxima (P) and minima (S). Figure 4 can be used to understand how particular cyclooctane conformers influence the topology and geometry of the full conformation space. As an example, consider the spherical component of the conformation space in Figure 4(a). There are two Cr conformers, related to each other through reflection. This reflection occurs through the center of the sphere, so that the two Crs are opposite each other on the north and south poles. At a high northern latitude, there are four TCCs, related to each other through rotation, thus forming a ring. In the southern hemisphere, there are four additional TCCs, forming their own ring, and related to the northern TCCs through reflection. Similarly, there are four northern and four southern CCs, Ss and Ps. On the equator, there are four Bs and BBs, related through rotation, again forming rings. Finally, there are eight TBs, related to each other by both rotation and reflection. Since reflection is through the center of the sphere, reflected TBs again lie on the equator. (In fact this also occurs with the Bs and BBs, except that the reflected Bs and BBs can also be obtained by rotation.) All of these symmetries are accommodated by the topology of the sphere. If we now consider distance as a constraint, we see how the particular conformers influence the geometry of the sphere. Both TCCs and CCs are very similar to Crs so that they are near the poles. They are also very similar to each other so form small diameter (high latitude) rings. By comparison, Bs and BBs are distinct from Crs and from each other so are far from the poles and form large diameter (equatorial) rings. These rings are then connected by meridians (e.g. Cr-TCC-S-B-S-TCC-Cr) to form the sphere.

Similar observations can be made to understand why the Möbius strip occurs in Figure 4(b). The S and P conformers are connected in a central ring corresponding to the northern S and P intersection ring in Figure 4(a). These conformations are related by rotation, i.e. the Ss are related by rotation, as are the Ps, while their reflections occur on the red Möbius strip (or southern intersection ring). The BCs also form a ring, again related by rotation, with the reflected ring occurring on the red Möbius strip. If we account for distance, the BC ring is approximately twice as long as the central S and P ring (note there are twice as many BCs as there are Ss or Ps). Further, each point on the central ring is close

in distance to two points occurring on opposite sides of the BC ring (e.g. a P is close to two opposite BCs). Thus we have a long ring with opposite points connected via a short ring: a Möbius strip. The TBC ring is similar to the BC ring. The Cs and TCs are analogous to the equatorial ring in Figure 4(b), in that they capture both rotation and reflection. The Cs and TCs also serve to connect the blue and red Möbius strips, forming the Klein bottle.

D. Energy landscape

In addition to conformation, we must of course consider energy to obtain a full understanding of cyclooctane. Using TINKER (dasher.wustl.edu), we computed the MM3 energy[27] for each conformation in our 6,040 point dataset. We also used TINKER to investigate four transition states previously discovered using quantum calculations[16,28]. These transition states are Cr (via TCC) $\rightarrow$ TS_1 $\rightarrow$ TBC, TBC $\rightarrow$ TS_2 $\rightarrow$ BC, TBC $\rightarrow$ TS_3 $\rightarrow$ TBC, and BC $\rightarrow$ TS_4 $\rightarrow$ BB. The third state TS_3 is a pseudorotation of TBC. (In addition, B and C were found to be transition states[28], although these are not included in our analysis.) We verified that these four conformations were transition states using vibrational analysis with MM3 energy in TINKER. Finally, we used structure minimization and TINKER's implementation of the Elber-Karplus Lagrangian multiplier-based reaction path following algorithm[29] to locate paths through conformation space connecting the various conformers via the transition states. The results of these investigations are shown in Figure 5.

Figure 5 summarizes most of what is known about cyclooctane's geometry. There are three conformation families: crown (Cr, CC, and TCC), boat (B, BB, and TB), and boat-chair (BC, TBC, C, and TC). There are two major transition states TS_1 and TS_4 which provide interconversion between the three families, and there are additional minor transition states TS_2, TS_3 (which appears to be the TC conformation), B, and C which provide interconversions within the families. Finally, the topology of the landscape can be described as a sphere intersecting a Klein bottle. In fact, it was previously thought that the conformation space of cyclooctane could be qualitatively modeled as a sphere intersecting a torus[30]. This observation was made using ring puckering coordinates and a limited sample of conformations[26].

Perhaps more interesting is how the topology of the cyclooctane energy landscape supports the fact that the BC is the dominant conformation. We first observe the relatively

high energy of the boat family, making B, TB, and BB unlikely conformations. Next we note that while BC is the global energy minimum, it is only slightly lower than Cr, with an approximate difference of 1 kcal/mol. Why then is BC the dominant conformation? We have observed that only two distinct Cr conformers exist, located on opposite poles of the spherical component of the conformation space. These conformers are accessible from the rest of the conformational space only via the relatively high energy TS_1 transition state. In contrast, there are sixteen distinct BC conformers. Each of these is connected via low energy TS_2 barriers to two nearby TBC conformers, forming two rings of connected states on the two Mobius strips. Furthermore, the TBC conformers in the two rings are connected to each other via TS_3 (TC) transition states. Hence the BC and TBC conformers form a large highly connected cluster of low energy states. In summary, the topology of the Klein bottle, combined with the energy landscape, causes cyclooctane to heavily favor BC conformers over Cr conformers.

IV. DISCUSSION

We have applied tools from the emerging fields of computational algebraic geometry and topology, as well as non-linear dimension reduction, to better understand the energy landscape of cyclooctane. Because cyclooctane is a relatively simple molecule, it has been well studied for the past 40 years. Even so, the tools we used have yielded new insights. First, it is surprising that the underlying conformation space is non-manifold. Although the space was known to be an algebraic variety, it was not known that the variety had singularities. Second, the existence of a non-orientable component (the Klein bottle) was not known and not expected. To our knowledge this discovery is unique in the area of molecular motion. In addition, the existence of the Klein bottle is not just a mathematical artifact. Instead, its presence is intimately tied to the geometry of cyclooctane conformation, and together with the energy landscape can be used to summarize and explain the molecular motion of cyclooctane.

Although our findings have a certain intrinsic interest, we should ask how they influence the standard view of energy landscapes and how they might affect typical algorithms used in computational chemistry. To address these questions, we first note that the cyclooctane conformations studied here were obtained using a rigid geometry (fixed bond lengths and an-

gles). We might speculate that the structure we observed was an artifact of this assumption, and that it might dissolve upon relaxing this constraint. The assertion that such structures are not artificial is known as the rigid geometry hypothesis[20,31], and is based on the observation that deformations of bond lengths and angles are up to two orders of magnitude stiffer than deformations of torsion angles. As an example, our experience using Tinker did not assume a rigid geometry, but our results supported the rigid geometry hypothesis, in that the transition paths we found were near the constrained surface.

Next consider the existence of the intersection rings. We might imagine that the energy gradient at a point on an intersection ring would preferentially push a cyclooctane conformation onto one or the other of the sphere or Klein bottle. If this were the case we would again conclude that the rings were an artifact. However, this is not the case. The TINKER calculations demonstrate that canonical conformations on the sphere (e.g. the crown conformation) and the Klein bottle (e.g. the boat-chair conformation) are linked via low energy transition pathways that pass through the vicinity of an intersection ring. We conjecture that the intersection rings may contain higher-order saddle points, such that a given conformation on a ring could fall towards a pole of the sphere, the equator of the sphere, or onto the Klein bottle.

Given that such structures exist in cyclooctane conformation space, and probably exist for other molecules, we might ask what effect these structures would have on typical algorithms from computational chemistry. On one hand, the existence of intersection singularities and non-orientable surfaces imply some amount of unrecognized complexity for modeling molecular behavior. Do multiple sheets in conformation space imply special requirements for sampling in Monte Carlo or Molecular Dynamics methods? Does it matter that we are sampling from a non-orientable structure? On the other hand, the ability to reduce dimension and provide a minimal representation is likely to improve results for many algorithms, especially algorithms that are not focused specifically on energy minima. We could, for example, use conformations obtained geometrically to seed certain methods, such as elastic band and string type path searching algorithms.

Finally, there is nothing particularly unique about cyclooctane or the methods we used in our analysis. Although our tools were specialized to 2D and algebraic double curves, the general concepts apply to any number of singularities in arbitrary dimensions (albeit with significantly increased computational complexity). In the future, these or similar tools might

well be applied to increasingly complex molecules (such as proteins), yielding additional interesting and unexpected results which would be difficult to otherwise obtain.

ACKNOWLEDGMENTS

This work has been supported by the Computer Science Research Fund (CSRF) under the Advanced Computing and Simulation (ASC) program at Sandia National Laboratories. E. A. Coutsias acknowledges partial support from NIH-NIGMS Grants No. R01-GM081710 and R01-GM090205. We thank Eric Cyr and Scott Mitchell for discussions of this work. Sandia is a multipurpose laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

REFERENCES

[1]D. Wales, *Energy Landscapes: with Applications to Clusters, Biomolecules, and Glasses* (Cambridge University Press, 2003).

[2]A. Pohorille and C. Chipot, *Free Energy Calculations: Theory and Applications in Chemistry and Biology* (Springer-Verlag, 2007).

[3]K. A. Dill and H. S. Chan, Nat. Struct. Bio. **4**, 10 (1997).

[4]A. E. Garcia, Phys. Rev. Lett. **68**, 2696 (1992).

[5]S. Hayward and N. Go, Ann. Rev. Phys. Chem. **46**, 223 (1995).

[6]S. B. Nolde, A. S. Arseniev, V. Y. Orekhov, and M. Billeter, Proteins: Struct. Func. Genet. **46**, 250 (2002).

[7]Y. Levy and A. Caflisch, J. Phys. Chem. B **107**, 3068 (2003).

[8]M. L. Teodoro, G. N. Phillips, and L. E. Kavraki, J. Comp. Bio. **10**, 617 (2003).

[9]P. Das, M. Moll, H. Stamati, L. E. Kavraki, and C. Clementi, Proc. Nat. Acad. Sci. **103**, 9885 (2006).

[10]A. Altis, M. Otten, P. H. Nguyen, R. Hegger, and G. Stock, J. Chem. Phys. **128**, 245102 (2008).

[11]W. M. Brown, S. Martin, S. N. Pollock, E. A. Coutsias, and J.-P. Watson, J. Chem. Phys. **129**, 064118 (2008).

[12]M. Kirby, *Geometric Data Analysis: an Empirical Approach to Dimensionality Reduction and the Study of Patterns* (Wiley, 2001).

[13]S. Roweis and L. K. Saul, Science **290**, 2323 (2000).

[14]J. Tenenbaum, V. de Silva, and J. Langford, Science **290**, 2319 (2000).

[15]J. B. Hendrickson, J. Am. Chem. Soc. **89**, 7036 (2002).

[16]W. R. Rocha, J. R. Pilego, S. M. Resende, H. F. Dos Santos, M. A. De Oliveira, and W. B. De Almeida, J. Comput. Chem. **19**, 524 (1998).

[17]R. K. Bharadwaj, Mol. Phys. **98**, 211 (2000).

[18]Z. Chen and F. A. Escobedo, J. Chem. Phys. **113**, 11382 (2000).

[19]J. Pérez, K. Nolsøe, M. Kessler, L. Garcia, E. Pérez, and J. L. Serrano, Acta. Cryst., 585(2005).

[20]J. M. Porta, L. Ros, F. Thomas, F. Corcho, J. Cantó, and J. J. Pérez, J. Comput. Chem. **28**, 2170 (2007).

[21]E. A. Coutsias, C. Seok, M. J. Wester, and K. A. Dill, Int. J. Quant. Chem. **106**, 176 (2006).

[22]A. Y. Dymarksky and K. N. Kudin, J. Chem. Phys. **122**, 124103 (2005).

[23]S. Martin and J.-P. Watson, submitted to Comput. Geom. Theory Appl.(2010).

[24]D. Freedman, Comput. Geom. Theory Appl. **36**, 106 (2007).

[25]G. Rote and G. Vegter, in *Effective Computational Geometry for Curves and Surfaces*, edited by J.-D. Boissonnat and M. Teillaud (Springer-Verlag, 2006) pp. 277–312.

[26]See Supplementary Material Document No. _____ for derivation of canonical conformations and topological analysis using ring puckering coordinates. For information on Supplementary Material, see http://www.aip.org/puservs/epaps.html.

[27]Y. H. Y. N. L. Allinger and J. H. Lii, J. Am. Chem. Soc. **111**, 8551 (1989).

[28]K. B. Wiberg, J. Org. Chem. **68**, 9322 (2003).

[29]R. Elber and M. Karplus, Chem. Phys. Lett., 375(1987).

[30]D. G. Evans and J. C. A. Boeyens, Acta Cryst., 663(1988).

[31]K. D. Gibson and H. A. Scheraga, Journal of Computational Chemistry **18**, 403 (1997).

FIGURE CAPTIONS

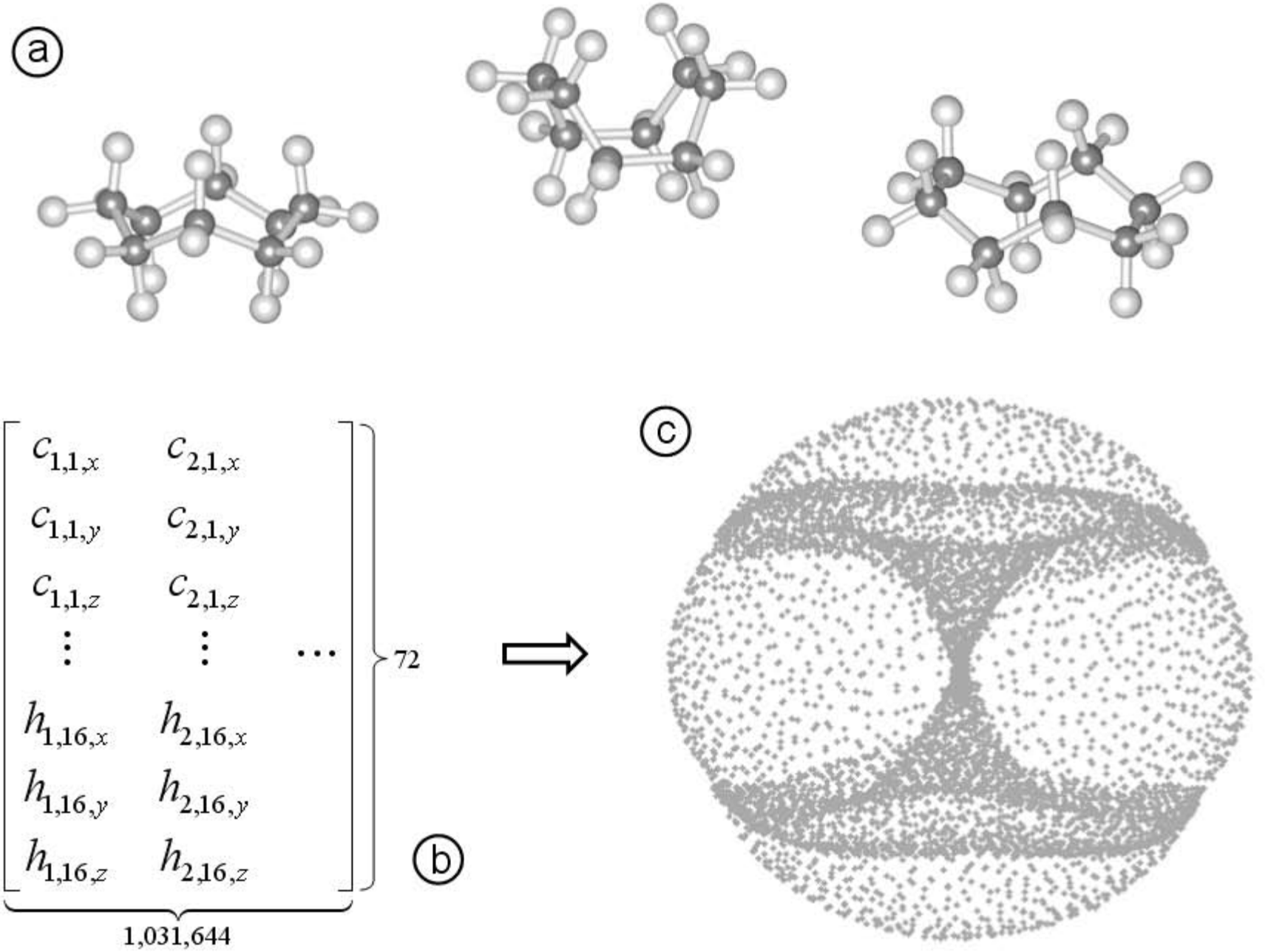
FIG. 1. Conformation space of cyclooctane. Here we show how the set of conformations of cyclooctane can be represented as a surface in a high dimensional space. On the top row (a), we show various conformations of cyclooctane. On the lower left (b), these conformations are represented by the 3D coordinates of their atoms. The coordinates are concatenated into vectors and shown as columns of a data matrix. As an example, the entry $c_{1,1,x}$ of the matrix denotes the x-coordinate of the first carbon atom in the first molecule. On the lower right (c), the Isomap method is used to obtain a lower dimensional visualization of the data.

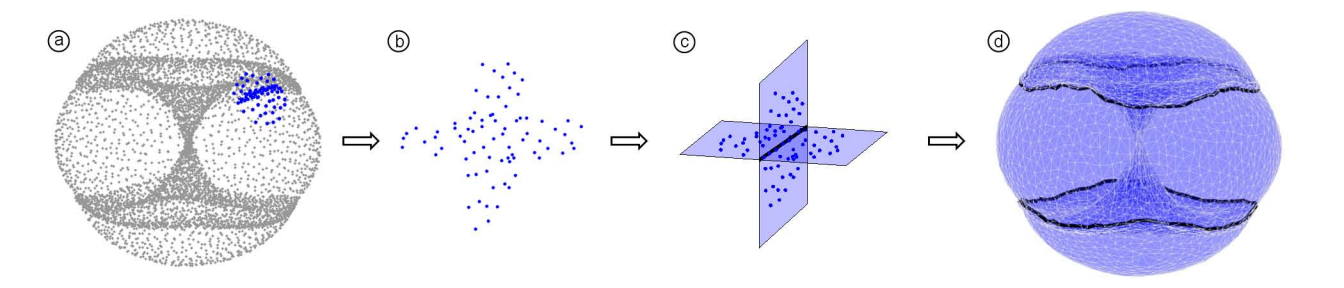
FIG. 2. Local geometry and triangulation of cyclooctane data. In (a)-(c) we show the local non-manifold geometry of the intersection rings in the cyclooctane conformation space. In (d) we show our triangulation of the dataset. A neighborhood centered on a point in the upper ring is shown in (a) using the Isomap coordinates; a PCA projection of this neighborhood from the original space ($\mathbb{R}^{72}$) is shown in (b); the local characterization of this neighborhood as two intersecting planes is shown in (c); and the triangulation obtained using an incremental surface reconstruction algorithm modified to use our local algebraic decomposition is shown in (d).

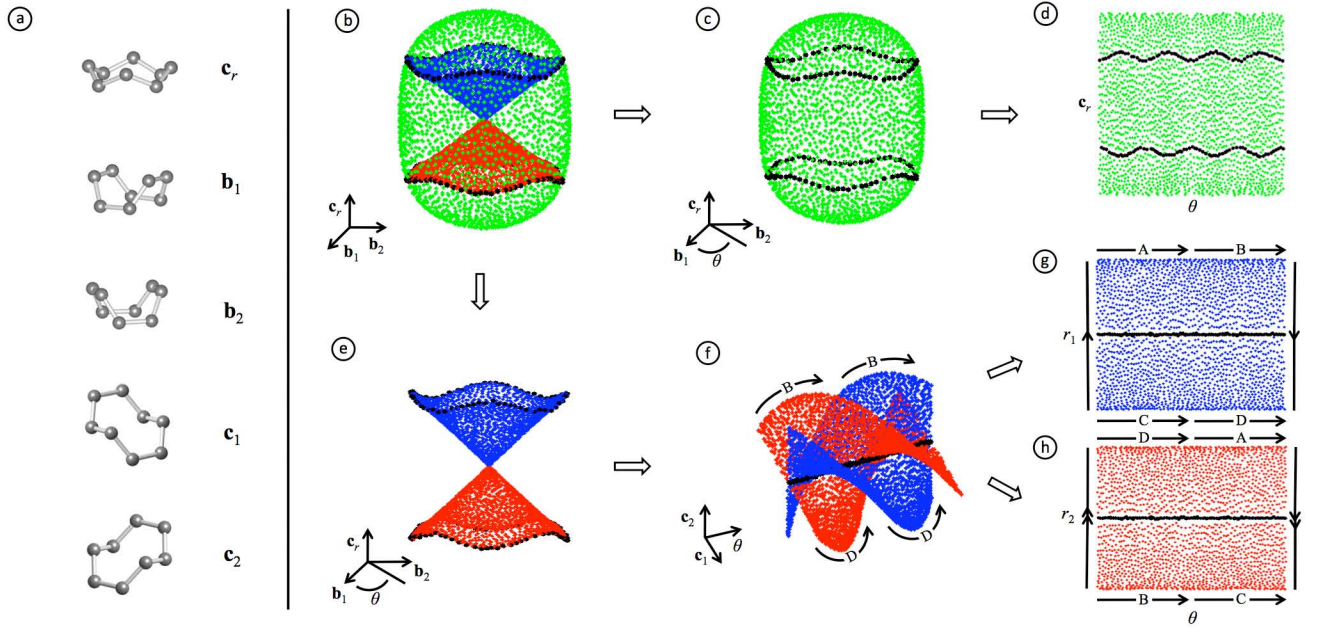
FIG. 3. Projections from torsion space. Here we show how the canonical crown-boat-chair basis can be used to produce a fully reduced 2D representation of the cyclooctane conformation space. The canonical conformations corresponding to the crown ($\mathbf{c}_r$), boat ($\mathbf{b}_1, \mathbf{b}_2$), and chair ($\mathbf{c}_1, \mathbf{c}_2$) are shown in (a). The projection of the space onto ($\mathbf{b}_1, \mathbf{b}_2, \mathbf{c}_r$) is shown in (b), analytically reproducing the results previously obtained using Isomap. We use green to represent the spherical component (c) of the conformation space and blue/red to represent the Klein bottle component (e). The intersection rings are shown using black. The spherical component (c) of the conformation space can be reduced to 2D by varying the azimuthal angle (θ) between 0 and 2π in the ($\mathbf{b}_1, \mathbf{b}_2$) plane, as shown in (d). The Klein bottle component can be reduced to 2D by decomposition into two Möbius strips, apparent using ($\mathbf{c}_1, \theta, \mathbf{c}_2$) coordinates (f), where θ again varies between 0 and 2π. The Möbius strips in (f) are in fact helicoids, which can be parameterized using signed distances r_1 and r_2 from the line $\mathbf{c}_1 = \mathbf{c}_2 = 0$. The signed distances can be used to reduce the two Möbius strips to 2D (g,h). In (g) and (h), we show how the two Möbius strips can be glued together to form the Klein bottle using labeled arrows to show equivalences (e.g. — $A \rightarrow$). It is interesting to note that representation of the Klein bottle in (b,e) has two deficiencies. First, the Möbius strips have been folded over the intersection rings so that the hourglass shape actually consists of two sheets, although it appears as one. Second, the apparent singularity at the origin is in fact an artifact of the projection. In actuality, the Möbius strips are joined according to the equivalences in (g,h), albeit using dimensions which are not visible in the ($\mathbf{b}_1, \mathbf{b}_2, \mathbf{c}_r$) coordinates.

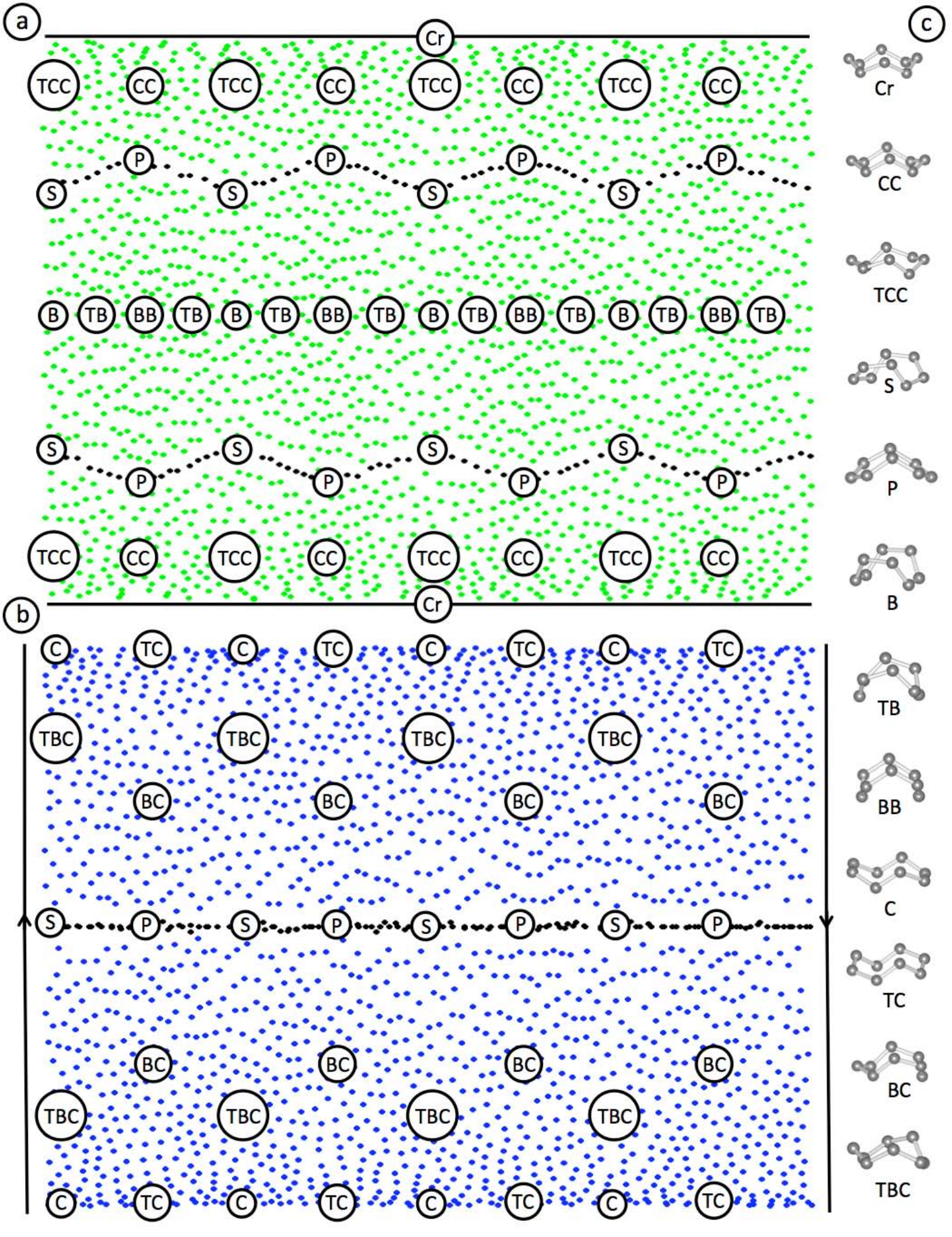
FIG. 4. Hendrickson's cyclooctane conformations. Here we show the positions of Hendrickson's 10 conformations of cyclooctane in our reduced dimensional representation. Shown in (a) are six of the conformations as they appear on the spherical (green) component of the conformation space, including Cr, CC, TCC, B, BB, and TB, as well as our S and P conformations, corresponding to the intersection ring energy minima (S) and maxima (P) seen in Figure 5. Shown in (b) are the C, TC, BC, and TBC conformations occurring on the blue Möbius strip. Conformations on the red Möbius strip are identically distributed. Shown in (c) are instances of the ten conformations, as well as S and P.

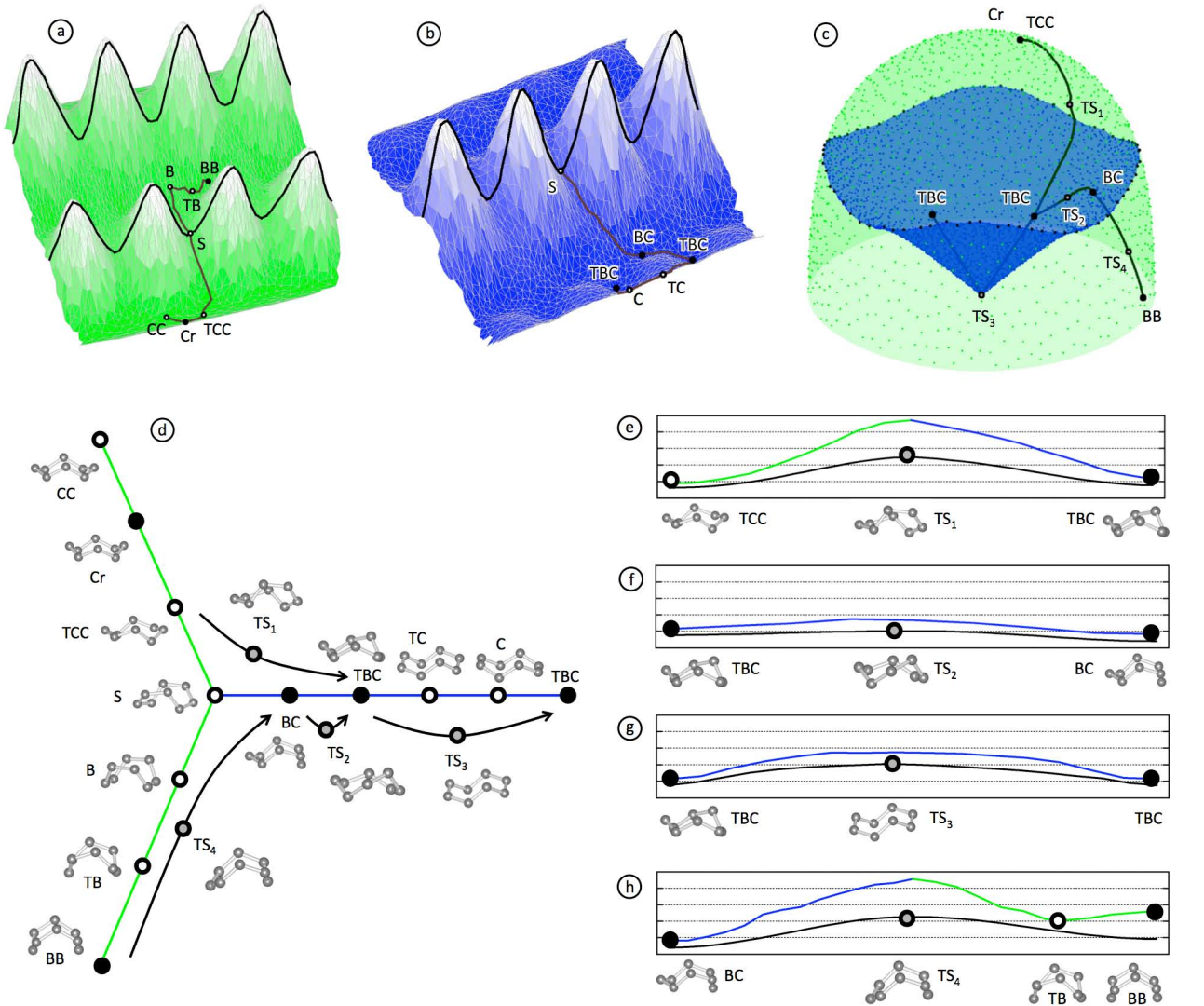
FIG. 5. Energy landscape for cyclooctane. Here we show Hendrickson's 10 conformations and 4 known transition states for cyclooctane relative to the MM3 energy landscape using the parameterizations described in Figure 3. In (a) we plot $(\theta, \mathbf{c}_r, \text{MM3 energy})$ for the spherical (green) component of the cyclooctane conformation space. The intersection rings are shown using black lines and Hendrickson's conformations are labeled using circles. The circles are colored black if they correspond to predicted energy minima[16,28] or white otherwise. Transition pathways are shown in brown between conformers. In (b) we show the energy landscape in $(\theta, r_1, \text{MM3 energy})$ coordinates for the blue Möbius strip. In (c) we show the 4 transition states and pathways computed using the EK algorithm in $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{c}_r)$ coordinates. Note that the transition states and pathways are unconstrained (in terms of bond length/angle) and therefore lie near, but not on the constrained surface described in Figure 3. In (d) we show a network of transition pathways with all of the paths drawn from (a-c). The paths are colored according to the component (green or blue) that contains each one, or using black lines for transition state pathways. A grey colored circle indicates a transition state. In (e-h) we show transition pathway coordinates vs. MM3 energy for the 4 transition states and the closest corresponding paths on the constrained surface. The MM3 energy ranges from 20-45 kcal/mol, and the same energy scale is used for each plot. It is interesting to note that interconversions between the sphere and Klein bottle are less likely than interconversions within the Klein bottle. In (e,h), for example, we see that the transition states are relatively high energy, while in (f,g) the transition states are low energy. These differences are also reflected in the constrained vs. unconstrained pathways: in (e,h), the transition states are significantly lower in energy than the constrained saddle S, and in (f,g), the constrained vs. unconstrained paths are not much different. (We also note that TB appears to be a minima using the constrained MM3 energy surface (h), but in reality this is not the case[16,28].)


a
$c_{1,1,x}$ $c_{2,1,x}$
$c_{1,1,y}$ $c_{2,1,y}$
$c_{1,1,z}$ $c_{2,1,z}$
$h_{1,16,x}$ $h_{2,16,x}$
$h_{1,16,y}$ $h_{2,16,y}$
$h_{1,16,z}$ $h_{2,16,z}$
72
1,031,644
b
c


a
b
c
d


a
$\mathbf{c}_r$
$\mathbf{b}_1$
$\mathbf{b}_2$
$\mathbf{c}_1$
$\mathbf{c}_2$
b
c
d
e
f
g
h
θ
A
B
C
D
r_1
r_2


a
Cr
TCC
CC
P
S
B
TB
BB
Cr
b
C
TC
TBC
BC
c
Cr
CC
TCC
S
P
B
TB
BB
C
TC
BC
TBC


a
B
BB
TB
S
CC
Cr
TCC
b
S
BC
TBC
TBC
C
TC
c
Cr
TCC
TS1
BC
TBC
TBC
TS2
TS4
TS3
BB
d
CC
Cr
TCC
S
B
TB
BB
TS1
TC
C
TBC
TBC
BC
TS2
TS3
TS4
e
TCC
TS1
TBC
f
TBC
TS2
BC
g
TBC
TS3
TBC
h
BC
TS4
TB
BB