

Supplementary Information:

Identifying molecular structural features by
pattern recognition methods

*Qing Lu**

Beijing National Laboratory for Molecular Sciences, Institute of Chemistry, Chinese

Academy of Sciences, Beijing, 100190, China

Table of Contents

1 Fig. S1. Different isomers with their convex hulls. Top two panels: the convex hulls of cis-2-butene, and trans-2-butene; bottom two panels: the convex hulls of cyclohexane in armchair conformation and boat conformation.

2 Fig. S2 The comparison of protein and its distorted counterpart. The random noise is set as $\pm 1.5 \text{ \AA}$.

3 Fig. S3 The atomic geometries of CH_4 , SiH_4 and GeH_4 and their convex hulls

4 Fig. S4 Comparison of convex hull area, TPSA area and SASA area.

5. Fig. S5 Comparison of convex hull volume and vdw volume.

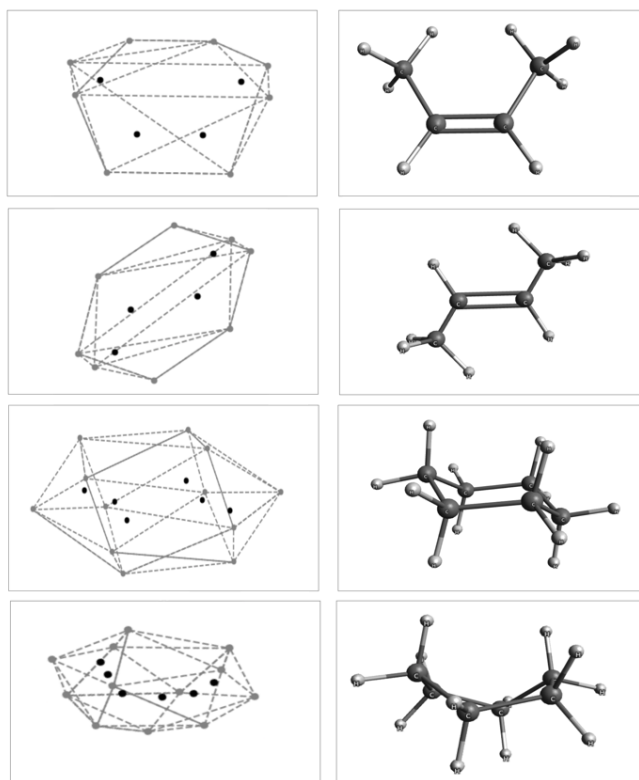


Figure S1. Different isomers with their convex hulls. Top two panels: the convex hulls of cis-2-butene, and trans-2-butene; bottom two panels: the convex hulls of cyclohexane in armchair conformation and boat conformation.

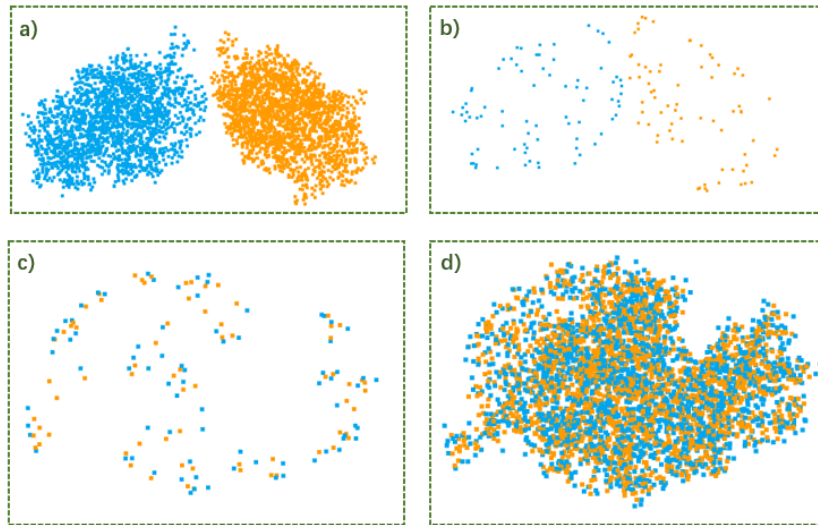


Figure S2. The comparison of proteins and its distorted counterpart. The random noise is set as $\pm 1.5 \text{ \AA}$

a) Structures of protein (Protein Database code, 1AKI, in orange) and distorted, shuffled protein (blue);

b) the convex hulls for protein and distorted, shuffled protein c) the convex hulls after orientation and ICP iterations; d) the structure superposition of protein 1aKI and its distorted counterpart. The distorted, shuffled protein geometry is multiplied with the rotation matrix constructed by matching convex hulls. See text for details.

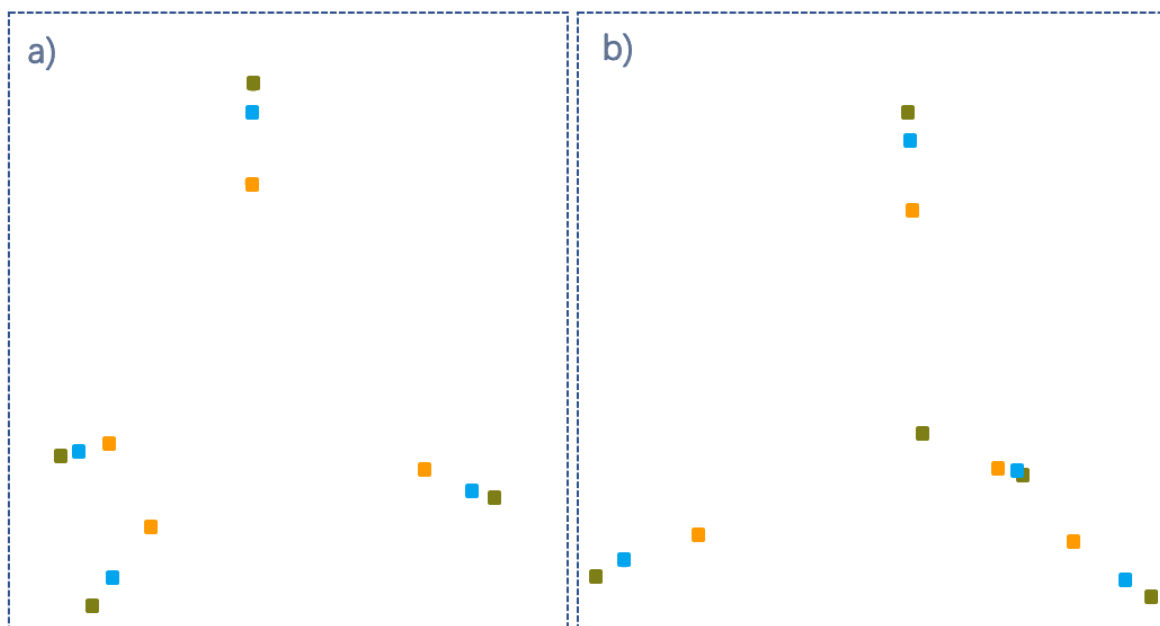


Figure S3. a) Convex hulls of CH₄ (orange), SiH₄ (blue) and GeH₄ (dark green) after orientation and ICP matching; b) the aligned atomic geometry of CH₄ (orange), SiH₄ (blue) and GeH₄ (dark green) after orientation and ICP matching;

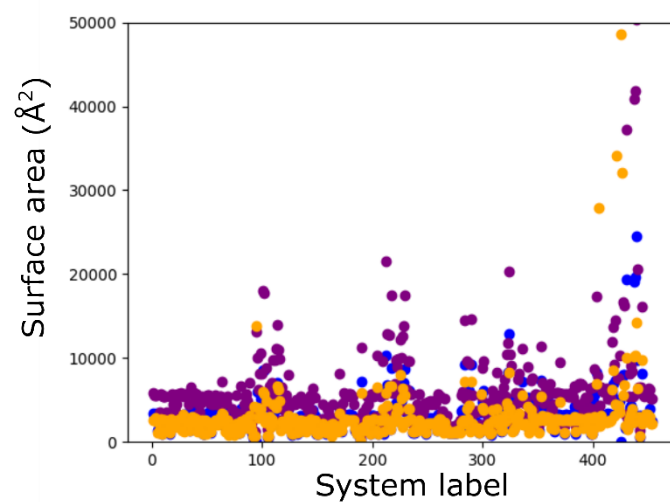


Figure S4. Comparison of convex hull area (orange dots), TPSA area (blue dots) and SASA area (purple dots).

The correlation coefficient between convex hull area and TPSA area is 0.73; the correlation coefficient between convex hull area and SASA area is 0.84.

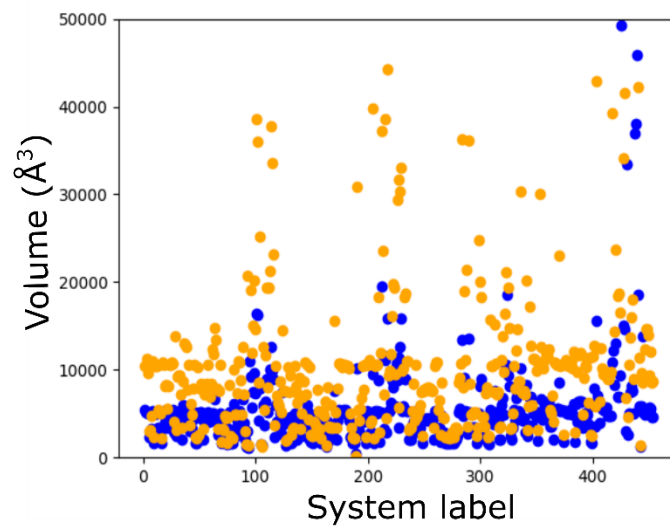


Figure S5. Comparison of convex hull volume (orange dots) and vdw volume (blue dots).

The figure below shows the corresponding convex hull volume (orange) and the vdw volume (blue).
The correlation coefficient is 0.81