九州大学大学院総合理工学報告
第 46 巻 第 2 号 1-5 頁 令和 7 年 2 月

Engineering Sciences Reports, Kyushu University
Vol. 46, No. 2, pp. 1-5, FEB, 2025

Theoretical Study on the Geometry Optimization of the Large Soft Polymers with the Elongation Method

Xiao HUANG,[*1] Yuuichi ORIMOTO,[*2] Yuriko AOKI,[*2,3†]
[†]E-mail of corresponding author: aoki.yuriko.397@m.kyushu-u.ac.jp

(Received November 6, 2024, accepted November 18, 2024)

The elongation optimization (ELG-OPT) method was employed to perform the geometry optimization on a large soft polymer cellulose tris(4-methylbenzoate) (CMB) model. The total energy of each step and optimized geometry, calculated by the ELG-OPT method, were compared to those by the conventional (CONV) method. The results showed that the ELG-OPT method can achieve a more stable structure than the CONV method. It proves more efficient than the CONV method in optimizing the large soft CMB model. Because as the calculation progresses, the CPU time required by the ELG-OPT method gradually becomes less than that by the CONV method and could remain constant. Hence, the ELG-OPT method could be a valuable tool in optimizing the geometries of the large flexible systems.

Key words: *Elongation method, Large soft polymer, Geometry optimization*

1. Introduction

Due to the high recognition ability, durability, and large sample loading capacity, cellulose derivative-based chiral stationary phases (CSPs) are the most studied and widely used in the experiments of high-performance liquid chromatography (HPLC) for chiral separation of enantiomers.[1,2)] They can effectively resolve various racemates, such as aliphatic and aromatic compounds, and numerous drugs. A series of CSPs were prepared by Okamoto group with one or two electron-donating or electron-withdrawing groups on the aryl ring, in which some cellulose tribenzoate-based CSPs coated on silica gel were useful in chiral separations.[3)] Particularly, cellulose tris(4-methylbenzoate) (CMB), which includes an electron-donating methyl group on the phenyl ring, exhibits a significant ability of chiral recognition.[3,4)] This may be because the carbonyl groups of the benzoate derivatives exhibit the high electron charge density, while the methyl groups on the phenyl rings stabilize it through an inductive effect.[3)] Furthermore, a lot of theoretical studies have concentrated on the mechanisms of chiral discrimination with the CMB serving as the CSP, by analyzing the interactions between the CMB and small chiral molecules.[1,2,5,6)]

Given the significance of the CMB in understanding chiral separation mechanism, this study concentrates on examining the CMB structure from a theoretical perspective. The purpose is to examine whether the elongation optimization (ELG-OPT)[7-9)] method can be applied to perform the geometry optimization of the large flexible CMB model by comparing with the results from the conventional (CONV) method.

*1 Department of Interdisciplinary Engineering Science, Chemistry and Materials Science, Interdisciplinary Graduate School of Engineering Sciences, Graduate Student
*2 Department of Material Sciences, Faculty of Engineering Sciences, Kyushu University
*3 Department of Interdisciplinary Engineering Science, Chemistry and Materials Science, Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

2. Methods

2.1 ELG Method

Figure 1 shows the ELG method procedure. ELG method established a computational framework that is analogous to the polymer chain growth in experiment. The details of ELG method are shown in previous studies.[7,8)] Initially, the calculation of the oligomer-sized fragment (hereafter referred to as the starting cluster), determines the canonical molecular orbitals (CMOs) spread over the whole cluster. Second, the CMOs are transformed into the

regional localized MOs (RLMOs) which are situated in the frozen A and active B regions, to proceed the calculations. B region represents the reactive end of the molecular chain, whereas A region is the non-reactive region positioned away from the reactive end. Subsequently, the attacking monomer (M) interacts with the growing end of active B region in the starting cluster. Newly obtained CMOs are re-localized to yield the new frozen and active RLMOs for the next ELG step. The localization and elongation steps as described above are systematically repeated in a successive manner until the polymer chain reaches the target length.

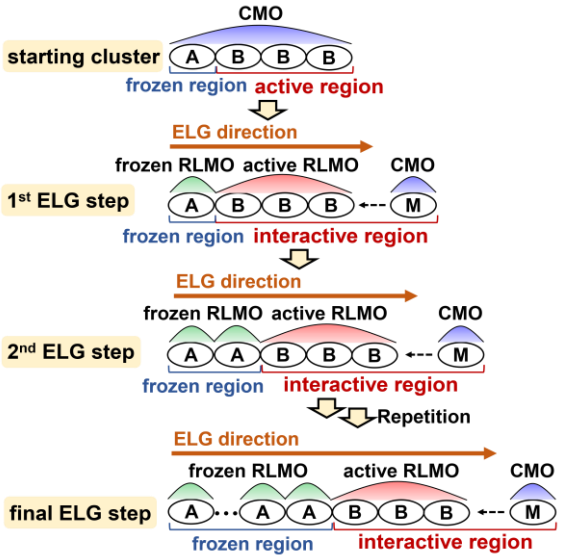

Fig. 1 ELG method procedure.

2.2 ELG-OPT Method

For Hartree-Fock (HF) calculation, at the atomic orbital (AO) basis, the first derivative of total energy E regarding to the nuclear coordinate X_A is given as in equation (1). The meaning of each quantity is displayed in the dashed box below equation (1).

$$\frac{\partial E}{\partial X_A} = \sum_{\mu\nu} \mathbf{D}_{\nu\mu} \frac{\partial \mathbf{H}^{\text{core}}_{\mu\nu}}{\partial X_A} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} \mathbf{D}_{\nu\mu}\mathbf{D}_{\lambda\sigma} \frac{\partial (\mu\nu||\sigma\lambda)}{\partial X_A} + \frac{\partial \mathbf{V}_{NN}}{\partial X_A} - \sum_{\mu\nu} \mathbf{Q}_{\nu\mu} \frac{\partial S_{\mu\nu}}{\partial X_A} \quad (1)$$

E: total energy; $\mathbf{X_A}$: nuclear coordinates
D: density matrix; $\mathbf{H}^{\text{core}}$: core Hamiltonian matrix
S: overlap matrix; $\mathbf{V}_{NN}$: nuclear-nuclear repulsion
$(\mu\nu||\sigma\lambda)$: two-electron integral
Q: energy-weighted density matrix

The HF equations are solved for the whole system in the CONV method, while only for the interactive BM region in the ELG method. Relative to the CONV method, only the last term in equation (1) highlighted by the yellow background has changed in the ELG method. As for the CONV method, the energy-weighted density matrix Q is specified in equation (2). However, for the ELG method, the eigenvalue matrix ε_{ABM} (ABM represents the entire system) in equation (2) becomes the non-diagonal matrix. In the ELG-OPT method, ε_{ij} is calculated by $C_{ui}^{MO\dagger} F_{uv}^{AO} C_{vj}^{MO}$. F_{uv}^{AO} is the Fock matrix, and the subscripts $\mu\nu$ denote the indices of two AOs. The MO coefficients of the whole system C_{ABM}^{MO} include two parts: coefficients of the frozen A region C_A^{LMO} and coefficients of the interactive BM region C_{BM}^{MO}. Then, for the ELG-OPT method, the energy-weighted density matrix Q is shown in equation (3).

CONV method

$$\mathbf{Q}_{\nu\mu} = \sum_{i}^{N/2} n_i \, \varepsilon_i C_{\mu i} C_{\nu i} \quad (2)$$

ELG-OPT method

$$\mathbf{Q}_{\nu\mu} = \sum_{ij}^{N/2} n_i \, \varepsilon_{ij} C_{\mu i} C_{\nu j} \quad (3)$$

n: occupancy number matrix; ε: eigenvalue matrix
C: MO coefficient

The differences between equations (2) and (3) are shown in below. For the CONV method as in equation (2), only diagonal elements of the eigenvalue matrix ε_i are used, meaning that it focuses on individual MO contributions without considering the interactions between different MOs. Unlike the CONV method, for the ELG-OPT method as in equation (3), the non-diagonal elements of the eigenvalue matrix ε_{ij} are used, meaning that it considers the interactions between different MOs.

The ELG-OPT[9] method procedure begins by optimizing the starting cluster with the CONV optimization algorithm. Once the equilibrium geometry is attained, the MOs are localized via the ELG localization procedure. Afterwards, an attacking monomer M is attached, dividing the system into A, B, and M regions. The geometry of A region is frozen, while the geometry of BM region is optimized by the ELG-OPT method.

3. Computational Details and Model

Both the ELG-OPT and CONV methods in GAMESS[10] program package were used to fully optimize the CMB geometry. The 3-21G basis set was used during the optimizations. Obtaining the optimized geometry with the ELG-OPT (or CONV) method includes two steps, (a) a potential geometry was obtained with the optimization convergence criteria of

1.0×10^{-2} Hartree/Bohr, and (b) based on the potential geometry from step (a), the target geometry was obtained with the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr, as shown in Fig. 2. Only the results of further optimization (step (b)) were focused on as shown below.

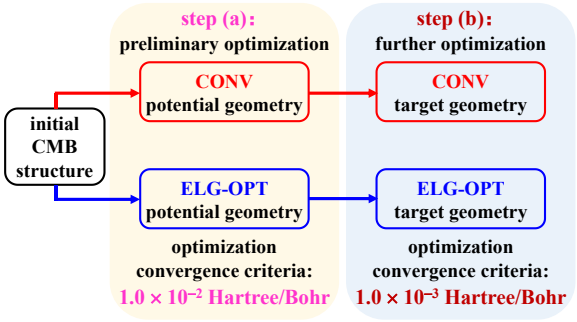

Fig. 2 Process of obtaining the target geometry.

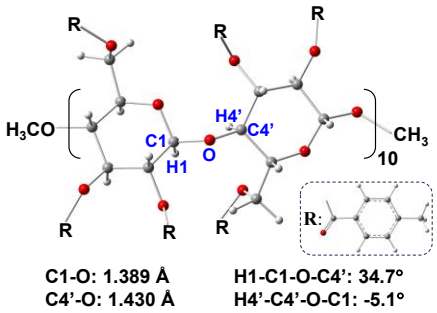

Fig. 3 CMB model and its initial structural parameters.

In this study, a 20-unit CMB model with 1329 atoms was chosen, as shown in Fig. 3. This model was built by stacking the unit structures. The CMB unit was constructed based on the crystal structure of cellulose tribenzoate (CTB) unit[11], in which a hydrogen atom on the phenyl ring is substituted with a methyl group. Our making FORTRAN code was used to stack the CMB unit structures.

4. Results and Discussion

Geometry optimization of CMB model with 20 units was carried out using both the ELG-OPT and CONV methods at the HF/3-21G level. The starting cluster consists of one unit in frozen A region and three units in active B region when doing the geometric optimization by the ELG-OPT method. The equilibrium structure was obtained via the ELG-OPT geometry optimization after a total of 17 steps. In Fig. 3, the CMB model and its initial C-O bond lengths and dihedral angles between the two adjacent units are presented.

Using the ELG-OPT and CONV methods, the total energies for each step of the 20-unit CMB model were compared under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr, as shown in Table 1. Correspondingly, Figure 4 displays a clearer view of the total energy differences (in Hartree) for each step. Except for the steps 1-3 and 6, the total energies obtained by the ELG-OPT method are lower than those by the CONV method. In the final step, the equilibrium structure is achieved, and the total energy using the ELG-OPT method is lower by -8.78×10^{-6} Hartree than that using the CONV method. These results show that the ELG-OPT method could achieve a more stable geometry than the CONV method under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr. The reason could be that the ELG-OPT method only centers on the local reaction part while fixing the irrelevant part that is already well-optimized. This strategy reduces the unnecessary fluctuations during the optimization of the whole system, which are reflected in the small number of the structural parameters. Consequently, it allows for an efficient search for the minimum point without causing unnecessary structural movements.

Table 1 Comparison of the total energies for each step and the error/atom for the final step using the ELG-OPT and CONV methods at the HF/3-21G level under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr.

step	total energy (Hartree) ELG-OPT	ΔE[b] (Hartree)	atom numbers	ΔE/atom[c] (Hartree)	error/atom[d] (Hartree)
1[a]	–7080.876650	0.003494	269	1.30E-05	---
2	–8822.498604	0.002079	335	6.21E-06	---
3	–10564.119523	0.001061	401	2.65E-06	---
4	–12305.742367	–0.003257	467	–6.98E-06	---
5	–14047.363053	–0.001333	533	–2.50E-06	---
6	–15788.982434	0.001549	599	2.59E-06	---
7	–17530.604170	–0.000345	665	–5.19E-07	---
8	–19272.227299	–0.003899	731	–5.33E-06	---
9	–21013.847799	–0.003122	797	–3.92E-06	---
10	–22755.469446	–0.003352	863	–3.88E-06	---
11	–24497.091899	–0.008698	929	–9.36E-06	---
12	–26238.712231	–0.008558	995	–8.60E-06	---
13	–27980.334506	–0.006191	1061	–5.83E-06	---
14	–29721.954510	–0.009359	1127	–8.30E-06	---
15	–31463.577267	–0.006680	1193	–5.60E-06	---
16	–33205.198576	–0.002050	1259	–1.63E-06	---
17	–34985.628384	–0.011667	1329	–8.78E-06	4.25E-09

[a]step 1: the calculation of starting cluster
[b]$\Delta E = E_{(ELG\text{-}OPT)} - E_{(CONV)}$
[c]ΔE/atom is obtained by ΔE/atom numbers.
[d]error/atom is obtained by $[E_{(ELG)} - E_{(CONV)}]$/atom numbers.

The error from the CONV method for the final step of the CMB model is shown in Table 1. To assess the error from the CONV method, the single point energy with the ELG method is calculated using the optimized geometry with the CONV method. It was found that the error/atom is 4.25×10^{-9} Hartree, indicating that the ELG method could match the results of the CONV method with a high precision for this large soft CMB model.

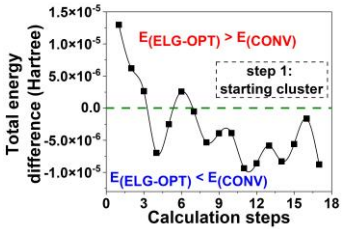

Fig. 4 Total energy difference (in Hartree) for each step between the ELG-OPT and CONV methods at the HF/3-21G level under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr.

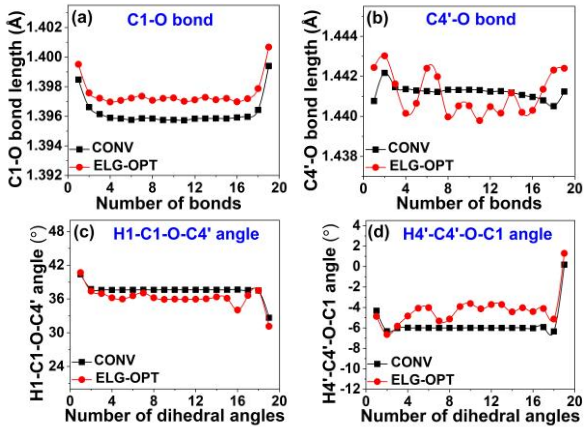

Fig. 5 Comparisons of bond lengths (a) and (b), and dihedral angles (c) and (d) between adjacent units of the geometries calculated by the ELG-OPT and CONV methods at the HF/3-21G level under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr.

Figure 5 shows the comparisons of the C-O bond lengths and dihedral angles between the two neighboring units, based on the calculations using the ELG-OPT and CONV methods. The structural parameters of each step obtained by the ELG-OPT method exhibit the greater fluctuations than those from the CONV method, indicating that the ELG-OPT method may be more suitable for optimizing such soft systems. This is because the larger fluctuations during the geometry optimization might reveal that the ELG-OPT method could allow for a nearly thorough exploration of the potential energy surface within the active region, facilitating the search for a structure with lower energy. However, it was found that there are small differences for both the bond lengths (see Fig. 5a and 5b) and dihedral angles (see Fig. 5c and 5d) between these two methods. The small discrepancies of the above bond lengths and dihedral angles indicate that the ELG-OPT method can align well with the CONV method in the geometric optimizations of this large soft CMB system. Figure 6 displays a comparison of the optimized geometries calculated by the ELG-OPT and CONV methods. It was found that a very small root mean square gradient (RMSD) of 0.16 was obtained between the geometries optimized by these two methods, indicating that the geometry optimized by the ELG-OPT method has the small differences relative to that by the CONV method. Figure 6 was generated by VMD[10] software, which was also used to calculate the RMSD value.

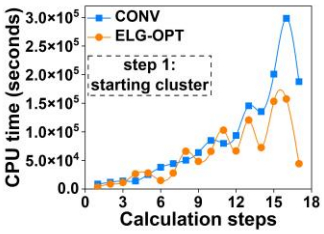

Fig. 7 Comparison of CPU time for each step between the ELG-OPT the CONV methods under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr.

Figure 7 shows the comparisons of CPU time for each step between the ELG-OPT and CONV methods. It was found that the fluctuations occur for the CPU time, which may be caused by the variations in the numbers of optimization cycles and SCF cycles in each step.

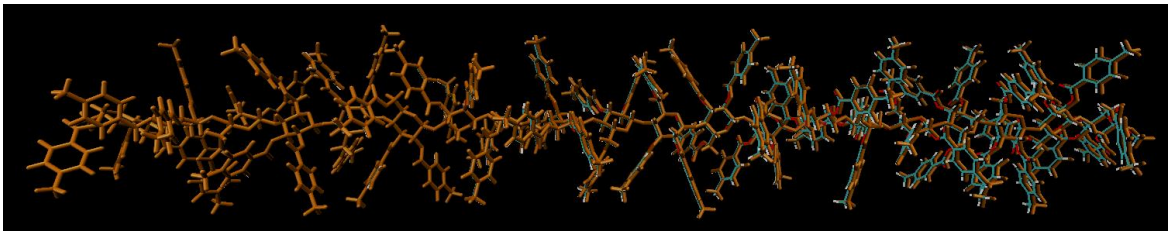

Fig. 6 Comparison of the optimized geometries calculated by the ELG-OPT (C (green), O (red), and H (gray)), and CONV (orange) methods at the HF/3-21G level under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr.

For the steps 1-5, the CPU time of each step using the ELG-OPT method is close to that using the CONV method. Subsequently, with a further increase of calculation steps, the ELG-OPT method gradually requires less CPU time relative to the CONV method. It may be attributed to the shorter time required for optimization and SCF convergences in each step of the ELG-OPT method. Besides, starting from the step 12, the CPU time required by the ELG-OPT method begins to be constant. These results indicate that the ELG-OPT could be more efficient in optimizing this large and soft CMB system relative to the CONV method.

In a word, with the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr, a lower total energy of the final step is obtained using the ELG-OPT method than the CONV method, supporting a more stable geometry.

5. Summary and Conclusion

ELG-OPT method was applied to optimize the geometry of the large flexible CMB model. Under the optimization convergence criteria of 1.0×10^{-3} Hartree/Bohr, the total energy for the final step obtained by the ELG-OPT is lower than that by the CONV method. It means that the ELG-OPT method can reach a more stable geometry. Besides, the ELG-OPT method shows a greater efficiency than the CONV method for the geometry optimization of the large soft CMB model. It is because as the calculation proceeds, the CPU time for the ELG-OPT method tends to decrease compared to the CONV method, and it could become constant. In a word, the ELG-OPT method could be effectively used for the geometric optimizations of the sophisticated large and soft systems.

Acknowledgments

This work was financially supported by a Grant-in-Aid for JSPS Fellows DC1 (KAKENHI: 202321990) from Japan Society for the Promotion of Science (JSPS), and the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) (KAKENHI: JP23245005, JP16KT0059, JP25810103, JP15KT0146, JP16K08321, and JP20H00588), and the Japan Science and Technology Agency (JST), CREST. The computations were carried out by Linux systems in our research group and the high-performance computing systems at the Research Institute for Information Technology at Kyushu University. The computation was performed using Research Center for Computational Science, Okazaki, Japan (Project: 24-IMS-C009).

References

1) R. Dallocchio, A. Dessì, B. Sechi, B. Chankvetadze, G. Jibuti, S. Cossu, V. Mamane, and P. Peluso, *Electrophoresis*, 44, 203 (2023).
2) R. Dallocchio, A. Dessì, B. Sechi, and P. Peluso, *Molecules*, 28, 7419 (2023).
3) Y. Okamoto and E. Yashima, *Angew. Chem. Int. Ed.*, 37, 1020 (1998).
4) B. Chankvetadze, *J. Chromatogr. A*, 1269, 26 (2012).
5) T. O'Brien, L. Crocker, R. Thompson, K. Thompson, P.H. Toma, D.A. Conlon, B. Feibush, C. Moeder, G. Bicker, N. Grinberg, *Anal. Chem.*, 69, 1999 (1997).
6) Y. Murakami, T. Shibata, K. Ueda, *Carbohydr. Res.*, 439, 35 (2017).
7) A. Imamura, Y. Aoki, and K. Maekawa, *J. Chem. Phys.*, 95, 5419 (1991).
8) Y. Aoki and F. L. Gu, *Phys. Chem. Chem. Phys.*, 14, 7640 (2012).
9) K. Liu, T. Inerbaev, J. Korchowiec, F. L. Gu, and Y. Aoki, *Theor. Chem. Acc.*, 131, 1277 (2012).
10) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, GAMESS Version 14 (Iowa State University, Iowa, 2003). "General Atomic and Molecular Electronic Structure System", *J. Comput. Chem.*, 14, 1347 (2003).
11) P. Zugenmaier, "*Crystalline Cellulose and Derivatives*", Springer, Berlin (2008).
12) W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual Molecular Dynamics," *J. Mol. Graphics*, 14, 33 (1996).