

รายงานวิจัยฉบับสมบูรณ์

สัญญาเลขที่ GRB_BSS_๖๗_๕๘_๐๗

โครงการวิจัย เรื่อง โครงสร้างที่หลากหลายของไซโคลเดกซ์ทริน เพื่อประยุกต์ใช้ใน
อุตสาหกรรมอาหารและยา

ผู้วิจัย

ศาสตราจารย์ ดร.สุพจน์ หารหนองบัว

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the Graduate School.

รายงานฉบับสมบูรณ์โครงการวิจัย

1. รายละเอียดเกี่ยวกับแผนงานวิจัย/โครงการวิจัย

ชื่อเรื่อง (ภาษาไทย) โครงสร้างที่หลากหลายของไซโคลเดกซ์ทรินเพื่อการประยุกต์ใช้ในอุตสาหกรรมอาหาร
และยา

(ภาษาอังกฤษ) Conformational diversity of cyclodextrins for applications in food and
pharmaceutical industries

ชื่อผู้วิจัย ศ.ดร. สุพจน์ หารหนองบัว

หน่วยงานที่สังกัด ภาควิชาเคมี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

หมายเลขโทรศัพท์ 02-218-7602 โทรสาร 02-218-7603 e-mail supot.h@chula.ac.th

ได้รับอนุมัติงบประมาณ ประจำปีงบประมาณ พ.ศ. 2558

งบประมาณที่ได้รับ1,000,000..... บาท ระยะเวลาทำการวิจัย1..... ปี

เริ่มทำการวิจัยเมื่อ (เดือน ปี) 1 ตุลาคม 2557 ถึง (เดือน ปี) 29 กันยายน 2558

2. การดำเนินงาน: ได้ดำเนินงานตามแผนงานที่ได้วางไว้ทุกประการ
 ได้เปลี่ยนแปลงแผนงานที่ได้วางไว้ดังนี้คือ

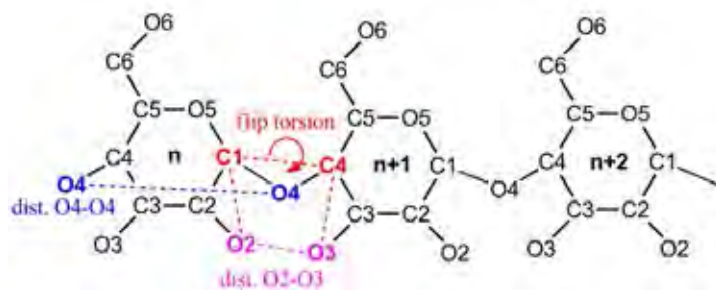
3. วัตถุประสงค์ของแผนงานวิจัย/โครงการวิจัย (โดยสรุป)

เพื่อศึกษาความเสถียรเชิงโครงสร้างของไซโคลเดกซ์ทรินในรูปแบบต่างๆ ทั้งไซโคลเดกซ์ทรินวงใหญ่และ
อนุพันธ์ของปีตาไซโคลเดกซ์ทรินในสถานะอุณหภูมิต่างๆ โดยอาศัยระเบียบทางเคมีคอมพิวเตอร์

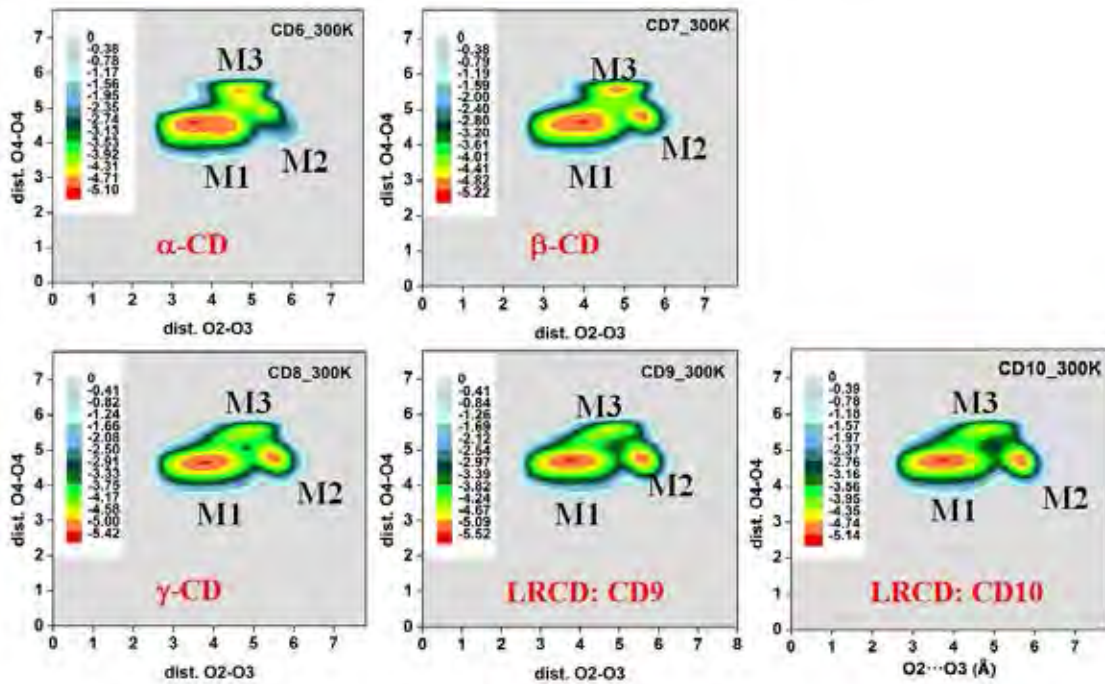
4. สรุปผลการดำเนินงาน

4.1 การศึกษาโครงสร้างของ small-ring CDs: CD6, CD7 และ CD8 เปรียบเทียบกับ large-ring CDs: CD9
และ CD10 ที่อุณหภูมิ 300K จากผลการจำลองพลวัตเชิงโมเลกุล REMD เมื่อพิจารณาแผนภาพ potential
energy surface ของระยะทาง O4-O4' และ O2-O3' (รูปที่ 1) พบว่า ไซโคลเดกซ์ทรินมีโครงสร้าง 3
conformations คือ M1, M2 และ M3 เรียงตามลำดับพลังงานต่ำไปสูง พันธะไฮโดรเจนระหว่างอะตอม
O2 และ O3' ที่ด้านกว้างของไซโคลเดกซ์ทริน (secondary rim) ทำให้โครงสร้าง M1 มีความเสถียรมาก
ที่สุด เมื่อไซโคลเดกซ์ทรินมีขนาดเล็กจะพบโครงสร้าง M2 ที่เป็น turn conformation จำนวนน้อยลง
เนื่องจากโครงสร้างของไซโคลเดกซ์ทรินขนาดเล็กมีความ rigidity มากขึ้นนั่นเอง

(ก)



(ข)



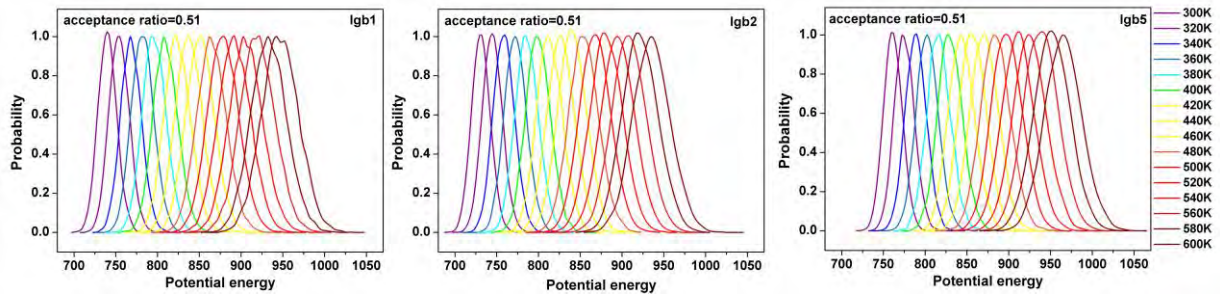
(ค)



รูปที่ 1 (ก) ระยะทาง O4-O4' และ O2-O3' (ข) 2D contour ระหว่าง d_1 [O2-O3'] และ d_2 [O4-O4'] ของไซโคลเดกซ์ทรินขนาดต่างๆ (ค) โครงสร้าง M1, M2 และ M3

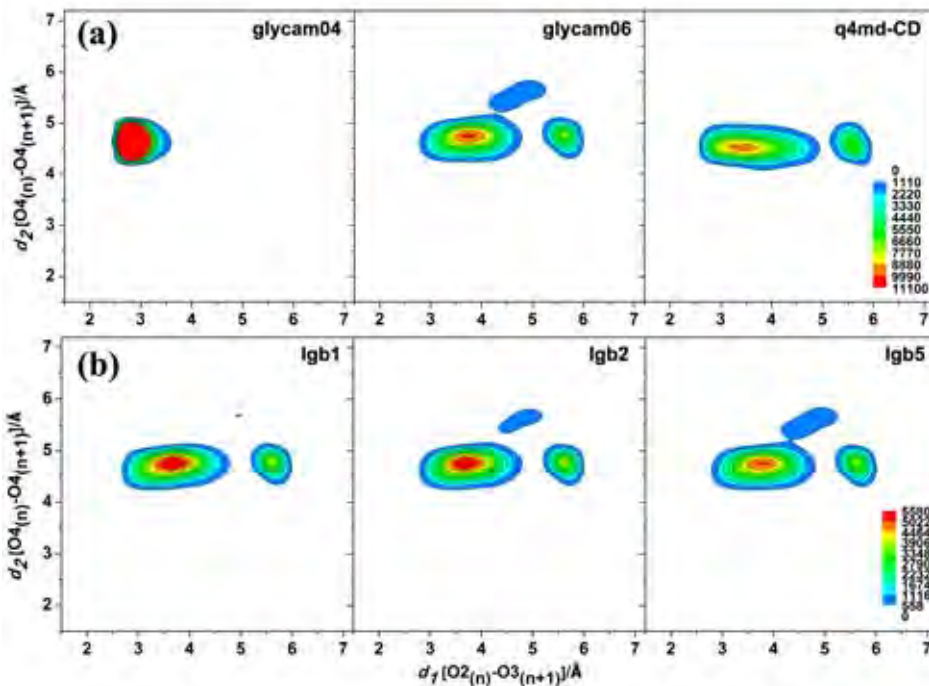
4.2 จากผลการจำลองพลวัตเชิงโมเลกุล REMD ของ large-ring CD: CD10 พบว่า

- สภาวะที่เหมาะสมในการทดลองคือ ช่วงอุณหภูมิ 300K-600 K และจำนวน replica ที่ใช้คือ 16 โดยแต่ละ replica มีอุณหภูมิต่างกัน 20K ดังแสดงในรูปที่ 2 โดยโมเดล GB implicit solvent ที่ต่างกัน มีผลเพียงเล็กน้อยต่อโครงสร้างของ CD10 จากการจำลองของโมเดล GB สามชนิดคือ Igb1, Igb2 และ Igb5



รูปที่ 2 แสดงการ overlap ของค่าพลังงานศักย์เมื่อใช้ โมเดล GB implicit solvent ที่ต่างกัน Igb1 Igb2 และ Igb5

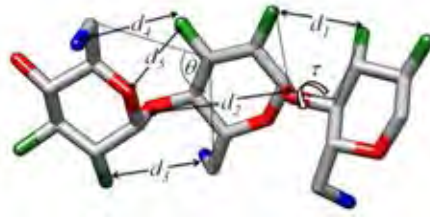
- ในงานวิจัยศึกษา force field สามชนิดคือ Glycam04 Glycam06 และ q4MD-CD พบว่า Glycam06 และ q4MD-CD สามารถอธิบายรูปแบบการเกิดโครงสร้างแบบ trans ได้ดีกว่า Glycam04 สำหรับระบบ CD10 ดังแสดงในรูปที่ 3



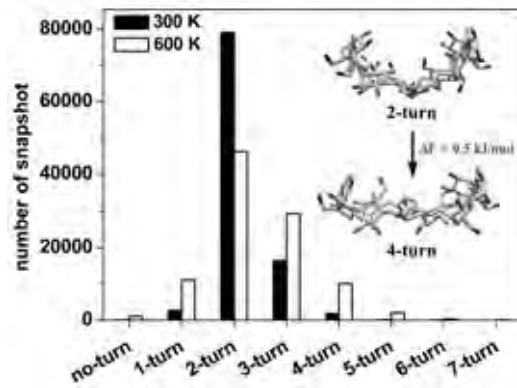
รูปที่ 3 2D contour ระหว่าง $d_1[O2-O3']$ และ $d_2[O4-O4']$

- เมื่อพิจารณาค่ามุม θ ระหว่างคู่ α -D-glucopyranose units โดยกำหนดให้ $\theta > 60^\circ$ แสดงถึงการหมุนของโครงสร้างระหว่างคู่ α -D-glucopyranose units หรือเรียกว่า turn conformation เมื่อพิจารณาจากจำนวนที่เกิด turn conformation ของคู่กลูโคส (รูปที่ 4) พบว่าโครงสร้าง CD10 สามารถเกิดหมุนของโครงสร้างได้หลายรูปแบบ ตั้งแต่ no-turn จนถึง 7-turn โดยรูปแบบที่พบมากที่สุดคือ 2-turn conformation ซึ่งรูปแบบดังกล่าวมีความสอดคล้องกับโครงสร้าง X-ray ของ CD10 เมื่อมีการหมุนของระหว่างคู่ α -D-glucopyranose units มากขึ้น ทำให้โครงสร้าง CD10 มีความเป็นระนาบมากขึ้น โดยอุณหภูมิที่เพิ่มขึ้นจะส่งผลให้เกิดการหมุนของโครงสร้างมากขึ้น

(ก)



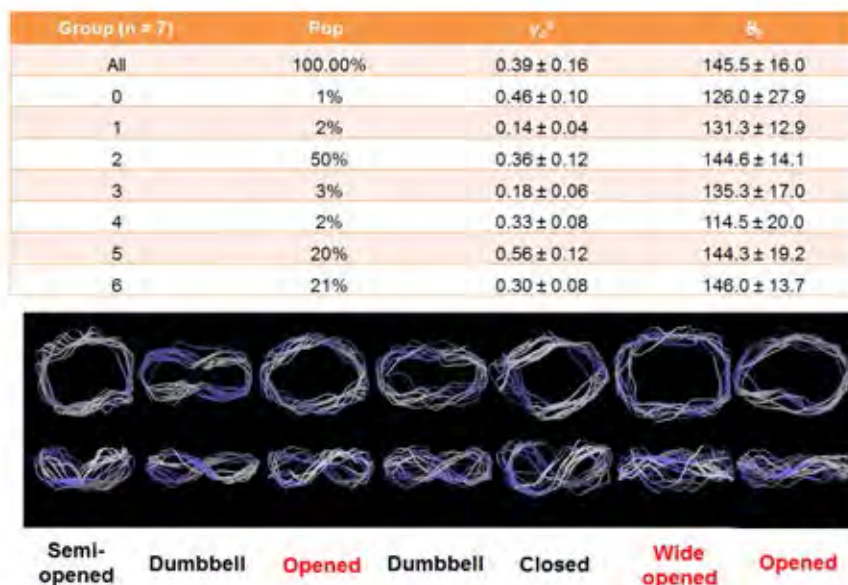
(ข)



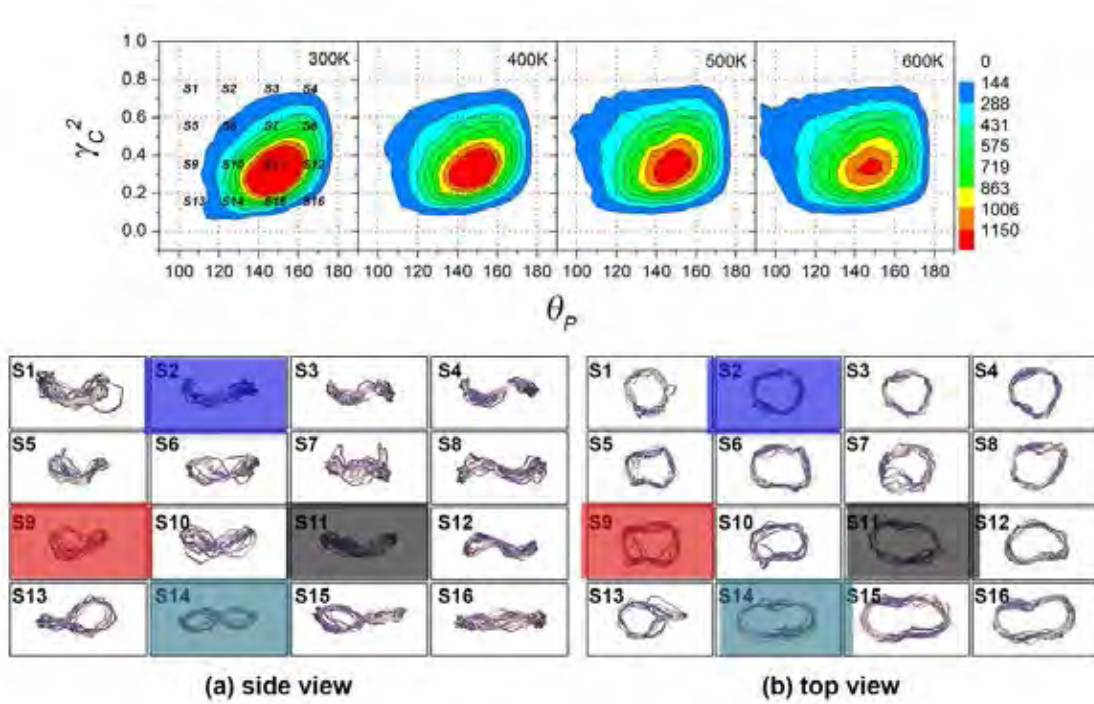
รูปที่ 4 (ก) แสดงมุม $\theta = C6-C2'-C6'$ (ข) แสดงการกระจายตัวของโครงสร้าง CD10 ในรูปแบบต่างๆ

4.3 จากผลการจำลองพลวัตเชิงโมเลกุล REMD ของ large-ring CD: CD14 พบว่า

- ไซโคลเดกซ์ทรินวงใหญ่ CD14 มีโครงสร้างหลากหลายมากขึ้น ประกอบด้วย semi-opened, dumbbell, opened, closed and wide opened conformations โดย dumbbell conformation กลุ่มที่ 2 พบได้มากที่สุด ดังแสดงในรูปที่ 5
- จากการวิเคราะห์พารามิเตอร์ circularity (γ_c^2) และ biplanar angle (θ_p) เทียบกับฟังก์ชันของ อุณหภูมิ พบว่าอุณหภูมิส่งผลต่อโครงสร้างไซโคลเดกซ์ทรินวงใหญ่ CD14 อย่างเห็นได้ชัด อย่างไรก็ตาม S1 conformation ยังพบได้มากที่สุด (รูปที่ 6)

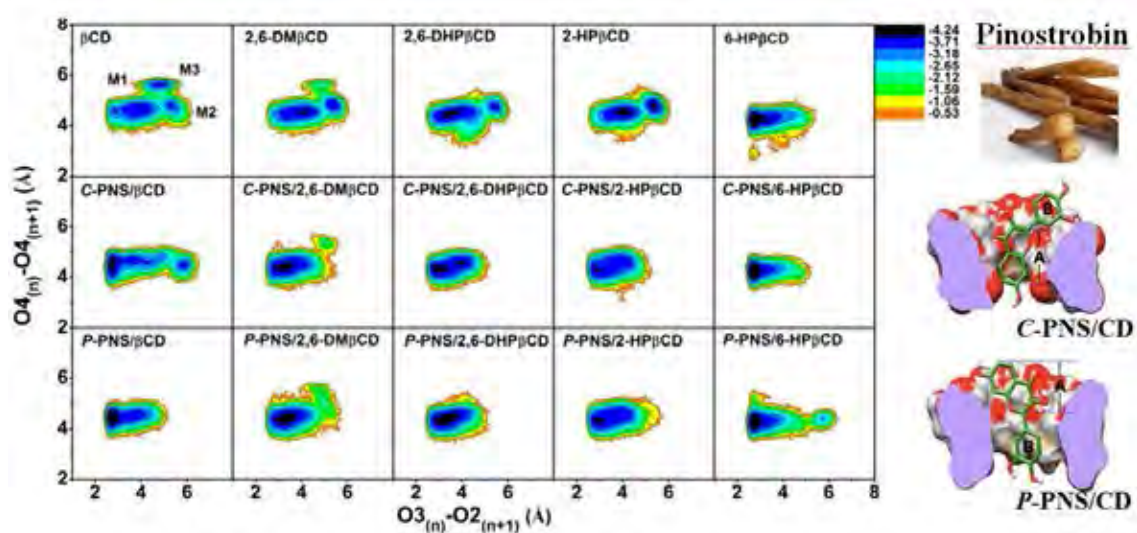


รูปที่ 5 โครงสร้างแบบต่างๆ ของไซโคลเดกซ์ทรินวงใหญ่ CD14 ที่อุณหภูมิ 300 K



รูปที่ 6 ผลของอุณหภูมิต่อโครงสร้างของไซโคลเดกซ์ทรินวงใหญ่ CD14

4.4 จากผลการจำลองพลวัตเชิงโมเลกุล REMD ของ CD7 (β CD) และอนุพันธ์ชนิด 2,6-dimethyl- β CD และ hydroxylpropyl- β -CD พบว่า การสร้างสารประกอบเชิงซ้อนอินคลูชันระหว่าง β CDs กับพินอสโตรบินซึ่งเป็นสารฟลาโวนอยด์ชนิดหนึ่งที่มีสมบัติการละลายน้ำต่ำ ส่งผลให้ conformation ของ β CDs เปลี่ยนแปลงจาก M3 open conformation เป็น M1 conformation มากขึ้น เนื่องจากเกิดโพรงด้านในของไซโคลเดกซ์ทรินเกิดอันตรกิริยาชนิด van der Waals กับสารฟลาโวนอยด์ (รูปที่ 7)



รูปที่ 7 2D contour ระหว่าง $d_1[O_2-O_3']$ และ $d_2[O_4-O_4']$ ของ β CD และอนุพันธ์ชนิด 2,6-dimethyl- β CD (2,6-DM β CD) และ hydroxylpropyl- β -CD (2,6-DHP β CD, 2-HP β CD และ 6-HP β CD)

5. งบประมาณที่ได้ใช้จ่ายไปแล้วนับตั้งแต่เริ่มโครงการเป็นเงินจำนวนทั้งสิ้น 1,000,000 บาท

6. ผลผลิตที่เกิดขึ้น

6.1 ผลงานที่ตีพิมพ์แล้ว (published) หรือได้รับการยอมรับให้ตีพิมพ์ (accepted) จำนวน 2 บทความดังนี้
(กรุณาระบุ authors, year, title, journal name, volume & pages)

1. Khuntawee, W.; Rungrotmongkol, T.; Wolschann, P. ; Pongsawasdi, P.; Kungwan, N.; Okumura, H.; Hannongbua, S., Conformation study of ϵ -cyclodextrin: Replica exchange molecular dynamics simulations. *Carbohydrate Polymers* 2016; 141, 99-105
2. J. Kicuntod, W. Khuntawee, P. Wolschann, P. Pongsawasdi, N.Kuawan, T. Rungrotmongkol, Inclusion complexation of pinostrobin with various cyclodextrin derivatives, *Journal of Molecular Graphics and Modelling* 2016; 63, 91-98

6.2 ผลงานต้นฉบับที่อยู่ระหว่างการพิจารณาเพื่อตีพิมพ์ จำนวน 1 เรื่องดังนี้

1. Khuntawee, W.; Kunaseth, M.; Rungrim, C.; Intagorn, S.; Wolschann, P.; Kungwan, N.; Rungrotmongkol, T.; Hannongbua, S., Ring conformation analytical methods of iotacyclodextrin from replica exchange molecular dynamics simulation, *Journal of Chemical Information and Modeling* 2016; revised

7. การเผยแพร่ผลงานวิจัย

7.1 การเผยแพร่ผลงานวิจัยในงานประชุมทางวิชาการ

1. “A molecular study of small- and large-ring cyclodextrins”, Joint IMS-KU Workshop on Molecular Sciences towards Green Sustainability, 45th Anniversary Building, Faculty of Science, Kasetsart University, Bangkok, Thailand, January 6-7, 2014 (*Invited speaker*)
2. “Investigation of the inclusion complex of flavanones/cyclodextrin based on molecular dynamics simulation”, The 17th International Cyclodextrin Symposium, Saarbrücken, Germany, May 29-31, 2014 (*poster*)
3. “Molecular Dynamics Study on Cyclodextrin Based Drug Delivery System”, Asian International Symposium –Interplay between Theory and Experiment in Physical Chemistry–, 95th Annual Meeting of the Chemical Society of Japan (CSJ), Funabashi, Chiba prefecture, Japan, March 26-29, 2015 (*Invited speaker*)
4. “A Molecular study of cyclodextrin-based drug delivery system”, JSPS-DST Asian Academic Seminar and School 2015, Kolkata, India, March 6-10, 2015 (*poster*)

7.2 การให้สัมมนา

1. “Dynamic behaviors of flavoniod-cyclodextrin complex” Pharmacoinformatics Research Group, Department of Medicinal Chemistry, Faculty of Life Sciences, University of Vienna, Austria, March 6, 2013
2. “A molecular study of cyclodextrins” Centre for Computational Chemistry, School of Chemistry, University of Bristol, UK, January 16, 2014
3. “Theoretical and experimental studies on inclusion complex of flavonoids and β -cyclodextrins” BIE Seminar, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Thailand, October 20, 2014

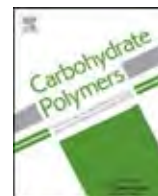
4. “Computational study on cyclodextrin and its inclusion complexes”, Theoretical Chemistry Colloquium, ITbM, Nagoya University, Nagoya, Japan, September 23, 2016

(ลงชื่อ)

(ศ.ดร. สุพจน์ หารหนองบัว)

หัวหน้าโครงการวิจัย

วันที่ เดือน พ.ศ.



Conformation study of ϵ -cyclodextrin: Replica exchange molecular dynamics simulations

Wasinee Khuntawee^a, Thanyada Rungrotmongkol^{b,c}, Peter Wolschann^{d,e}, Piamsook Pongsawasdi^c, Nawee Kungwan^f, Hisashi Okumura^{g,h,*}, Supot Hannongbua^{i,**}

^a Nanoscience and Technology Program, Graduate School, Chulalongkorn University, 254 Phayathai Road, Bangkok 10330, Thailand

^b Ph.D. Program in Bioinformatics and Computational Biology, Faculty of Science, Chulalongkorn University, 254 Phayathai Road, Bangkok 10330, Thailand

^c Department of Biochemistry, Faculty of Science, Chulalongkorn University, 254 Phayathai Road, Bangkok 10330, Thailand

^d Department of Pharmaceutical Technology and Biopharmaceutics, University of Vienna, Vienna 1090, Austria

^e Institute of Theoretical Chemistry, University of Vienna, Vienna 1090, Austria

^f Department of Chemistry, Faculty of Science, Chiang Mai University, 239 Huay Kaew Road, Muang District, Chiang Mai 50200, Thailand

^g Research Center for Computational Science, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan

^h Department of Structural Molecular Science, The Graduate University for Advanced Studies, Okazaki, Aichi 444-8585, Japan

ⁱ Computational Chemistry Unit Cell, Department of Chemistry, Faculty of Science, Chulalongkorn University, 254 Phayathai Road, Bangkok 10330, Thailand

ARTICLE INFO

Article history:

Received 23 January 2015

Received in revised form 4 September 2015

Accepted 8 October 2015

Available online 13 October 2015

Keywords:

Large-ring cyclodextrin

ϵ -Cyclodextrin

Replica exchange molecular dynamics

Conformational changes

ABSTRACT

There is growing interest in large-ring cyclodextrins (LR-CDs) which are known to be good host molecules for larger ligands. The isolation of a defined size LR-CD is an essential prerequisite for studying their structural properties. Unfortunately the purification procedure of these substances turned out to be very laborious. Finally the problem could be circumvented by a theoretical consideration: the highly advantageous replica exchange molecular dynamics (REMD) simulation (particularly suitable for studies of conformational changes) offers an ideal approach for studying the conformational change of ϵ -cyclodextrin (CD10), a smaller representative of LR-CDs. Three carbohydrate force fields and three solvent models were tested. The conformational behavior of CD10 was analyzed in terms of the flip (turn) of the glucose subunits within the macrocyclic ring. In addition a ranking of conformations with various numbers of turns was performed. Our findings might be also helpful in the temperature controlled synthesis of LR-CDs as well as other experimental conditions, in particular for the host–guest reaction.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Cyclodextrins (CDs) are produced from cyclodextrin glucanotransferase or amyloamylase from starch or similar substrates. They are cyclic α -1,4 linked oligosaccharides of D-glucopyranose units (Fig. 1a) with a hydrophobic nanoscale cavity and a hydrophilic outer surface. Generally, cyclodextrins and their derivatives are used to enhance water-solubility, stability and bioavailability of compounds of interest in pharmaceutical and

food industries. The regular-ring CDs contain six, seven and eight α -D-glucopyranoside units known as α -CD, β -CD, and γ -CD, respectively, whereas the large-rings CDs (LR-CDs) are comprising more than eight units. The small-ring CDs prefer the doughnut shape, whereas LR-CDs may exist in various conformations. For example, based on the crystal structures ϵ -CD (CD10) (Jacob et al., 1998, 1999; Ueda, Endo, Nagase, Kobayashi, & Nagai, 1996) and ι -CD (CD14) (Harata, Endo, Ueda, & Nagai, 1998; Jacob et al., 1998, 1999) containing 10 and 14 D-glucose units, respectively, these CDs are stable in bent forms while CD26 with 26 D-glucose units shows a somewhat twisted form (Gessler et al., 1999; Nimz, Geßler, Usón, & Saenger, 2001). For small-ring CDs, the dihedral angles $\tau(O2_{(n)}-C1_{(n)}-C4_{(n+1)}-O3_{(n+1)})$ (Fig. 1a), determined for all glucose units are in the range from -60° to 60° . Contrary to these findings the τ angle of some glucose units which are connected in the larger CD moieties is found within much larger ranges.

* Corresponding author.

** Corresponding author.

E-mail addresses: w.khntawee@gmail.com (W. Khuntawee), t.rungrotmongkol@gmail.com (T. Rungrotmongkol), karl.peter.wolschann@univie.ac.at (P. Wolschann), piamsook.p@chula.ac.th (P. Pongsawasdi), naweekung@gmail.com (N. Kungwan), hokumura@ims.ac.jp (H. Okumura), supot.h@chula.ac.th (S. Hannongbua).

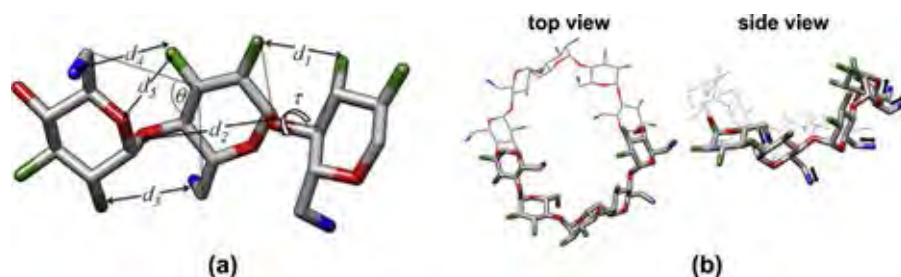


Fig. 1. (a) CD fragment showing the atomic labels and focused structural parameters, d_1 [O2–O3'], d_2 [O4–O4'], d_3 [O2–O6'], d_4 [O3–O6'], d_5 [O3–O5'] and τ [O2–C1–C4'–O3']. A turn between two glucose units is shown for the left glucose unit, whereas the two other glucose rings are within the band structure in cis position. (b) A helical band subunit is shown with five glucose rings (from the X-ray structure of CD10).

The crystal structure of CD10, C2 symmetry is observed. Two slightly twisted band subunits (Fig. 1b) are connected by two flips of glucose rings leading to a saddle-like conformation of ϵ -CD (Jacob et al., 1998, 1999; Ueda et al., 1996).

Besides a variety of researches on the physicochemical properties and host–guest interactions of small ring CDs (Cao & Wu, 2015; Fernandes et al., 2014; Hu, Tang, & Chu, 2014; Khuntawee, Wolschann, Rungrotmongkol, Wong-Ekkabut, & Hannongbua, 2015; Ogawa, Takahashi, & Yamamoto, 2015; Sangpheak, Khuntawee, Wolschann, Pongsawasdi, & Rungrotmongkol, 2014), the studies on LR-CDs (Dodziuk et al., 2003; Larsen, 2002; Larsen, Endo, Ueda, & Zimmermann, 1998; Wang et al., 2015) have been less mentioned. The larger ring size molecules are enabled to be good hosts, thus enhancing the solubility and efficiency of relatively large ligands (Ueda, 2002, 2004). In order to understand this behavior more clearly, the physicochemical and structural properties at a molecular level were investigated. In several publications, the dynamical behavior was described using MD simulation techniques starting from the crystal structures of CD10, CD14, and CD26 (Gotsev & Ivanov, 2007a) and larger macro-ring sized models (CD27–30, 35, 40, 48, 55, 70, 85, and 100) including intermediate conformations of available X-ray structures (CD11–13, CD15–18, and CD20–25) in either gas phase or in solution (Gotsev & Ivanov, 2007b, 2009; Gotsev, Ivanov, & Jaime, 2007; Ivanov, 2011, 2012; Ivanov & Jaime, 2004; Maestre, Beà, Ivanov, & Jaime, 2007; Shimada, Handa, Kaneko, & Takada, 1996; Shimada, Kaneko, Takada, Kitamura, & Kajiwara, 2000). The MD results of CDs with a degree of polymerization higher than 13 support the hypothesis for the existence of more than one cavity in LR-CDs.

In the process of enzymatic synthesis of CD, the CD conformations could be controlled by the number of glucopyranose monomers and experimental conditions (temperature and pH). However, the single conventional simulation of biological systems at low temperature condition trends toward a larger number of local minima. To overcome this multiple-minima problem, the replica exchange molecular dynamics simulation (REMD) which allows the exchange of non-interacting replicas of the system at several temperatures has been applied (Cheng, Cui, Hornak, & Simmerling, 2005; Hansmann, 1997; Nymeyer, Gnanakaran, & Garcia, 2004; Sugita & Okamoto, 1999). Therefore, to study the conformational changes of LR-CDs as well as the temperature effect on their structures, we apply REMD technique on one representative of LR-CDs, CD10, which has two flips (antitype) of glucose subunits at opposite sites in the macrocyclic ring. Additionally, REMD simulations have been carried out using different force fields as well as various solvation models based on generalized born (GB) implicit solvent. The replica exchange based conformational study of CD10 affords the instrument to control the physicochemical properties of this substance and subsequently its application.

2. Methods

A detailed description of the REMD method is given elsewhere (Sugita & Okamoto, 1999). The structure preparation and REMD simulations were performed by Amber10 package (Case et al., 2008) with the carbohydrate force field, glycam06 (Kirschner et al., 2008; Nutho et al., 2014; Tessier, Demarco, Yongye, & Woods, 2008). In addition, the glycam04 (Basma, Sundara, Calgan, Vernali, & Woods, 2001; Kirschner & Woods, 2001a,b) and hybrid force field, q4md-CD (Cezard, Trivelli, Aubry, Djedaini-Pilard, & Dupradeau, 2011) were also tested for reasons of comparison. The crystallographic structure of CD10 was taken from the Cambridge Crystallographic Data Centre CCDC (CCDC) entry code CCDC100656 as the starting structure (Jacob et al., 1998). For taking into account the solvent model effect, 50 ns REMD simulations were performed under three different GB implicit solvent models Igb1, Igb2, and Igb5 (Rungnim, Rungrotmongkol, Hannongbua, & Okumura, 2013). The initial structure of CD10 was fully minimized with 2000 steps of steepest descent (SD) method, followed by 1000 steps of conjugated gradient (CG) method. Here, all REMD simulations were performed with 16 replicas at temperatures ranging from 300K to 600K with interval steps of 20K. This temperature protocol was found to be the most promising one for the present case. To equilibrate the systems at the assigned temperatures, a short MD simulation of 5 ns was performed prior to the REMD simulation. The REMD simulation was then done for 100 ns using the solvent model Igb5. Conformations at all temperatures were sampled at every 2 ps. The conformational change of CD10 was monitored through contour plots of the probability by the distances between the secondary hydroxyl groups of adjacent glucoses, d_1 [O2_(n)–O3_(n+1)], and the distances between the glycosidic oxygen atoms, d_2 [O4_(n)–O4_(n+1)], using the ptraj module in Amber10 package. In addition the conformational and temperature change correlation was investigated and the probabilities of each state were calculated.

3. Results and discussion

3.1. Crystal structure analysis

The basic structural properties of CD10 in the crystalline state are well understood. The distances of adjacent glucopyranose units d_1 [O2–O3'], d_2 [O4–O4'], d_3 [O2–O6'], d_4 [O3–O6'] and d_5 [O3–O5'] as well as the dihedral angles τ (O2_(n)–C1_(n)–C4_(n+1)–O3_(n+1)) are summarized in Table 1. As a consequence of the C2 symmetry the distances of the second half of the CD rim are identical. The distance R5–R6 (as well as R10–R1) is much larger because of the flip of the glucopyranose ring.

Table 1

The focused structural properties, d_1 [O2–O3], d_2 [O4–O4'], d_3 [O2–O6'], d_4 [O3–O6'], d_5 [O3–O5'] and τ [O2–C1–C4'–O3'] measured from the crystal structure of CD10 (Jacob et al., 1998).

Ring connecting	d_1 (Å)	d_2 (Å)	d_3 (Å)	d_4 (Å)	d_5 (Å)	τ (°)
R1–R2	2.89	4.60	6.39	7.35	4.47	–38.1
R2–R3	3.03	4.46	6.01	7.07	4.42	–39.1
R3–R4	2.88	4.36	5.79	7.62	4.82	–12.9
R4–R5	3.91	4.48	6.05	6.80	3.97	–75.2
R5–R6	5.53	4.47	3.93	4.79	3.20	143.4

3.2. Replica exchange simulation of CD10

So as to obtain a reasonable replica exchange simulation initially the temperature distribution and the number of replica were tested. The potential energy distributions of neighboring replicas were shown in Fig. S1 of the supplementary materials. It can be seen that the temperature from 300 K to 600 K with 20 K interval is suitable for CD10 system. 50 ns REMDs in connection with different implicit solvent models were conducted.

In order to investigate the conformational changes of CD10, the distances between the secondary hydroxyl groups of adjacent glucose moieties, d_1 [O2_(n)–O3_(n+1)], and the distances between the adjacent glycosidic oxygen atoms, d_2 [O4_(n)–O4_(n+1)], are plotted for all the snapshots which were taken from the simulation at the lowest temperature. The distance d_2 [O4_(n)–O4_(n+1)] is less sensitive for the glucose ring flips and describes the ellipticity of the rim. The contour lines in the diagrams are giving the probability for the distance values: the highest probabilities are assigned by the darkest color (red to blue). Probability diagrams (Fig. 2) are showing an accumulation of the distance pairs around 3.5 Å/4.5 Å which are correlating satisfactory with glucopyranose ring connections of the crystal structure, wherein the 2- and 3-hydroxy groups remain in cis position (at the same side of the rim). A less populated second accumulation point appears at 5.5 Å/4.5 Å. These distances are

correlated with the turn of the glucopyranose rings from the cis arrangement to a trans position of the six membered rings. Due to the flexibility of the ring systems some spreads of the distance values can be observed.

The results of the different tested carbohydrate force fields (Fig. 2a) gave rise to the assumption that glycam06 answers best our purposes. After testing the different solvent modes, lgb1, lgb2, and lgb5 turned out to be the most reliable one which was used for further investigations.

REMD enables to consider the influence of higher temperatures during the simulation. In Fig. 3 the related contour plots according to Fig. 2 are shown for various temperatures. Generally, no principal change of the pattern can be observed, but the spread of the distance distribution increases. This is not unexpected as higher temperatures will effectuate higher amplitudes of the motion of the molecule. No additional conformation can be detected, except for a less pronounced accumulation point at 5.0 Å/5.0 Å.

3.3. The conformation of CD10

The conformational change of CD10 in aqueous solution is determined by several processes. First of all, the ring formed by the O4 oxygen atoms is not planar and there are some changes of the geometry of this larger ring system. Secondly, the occurring movements

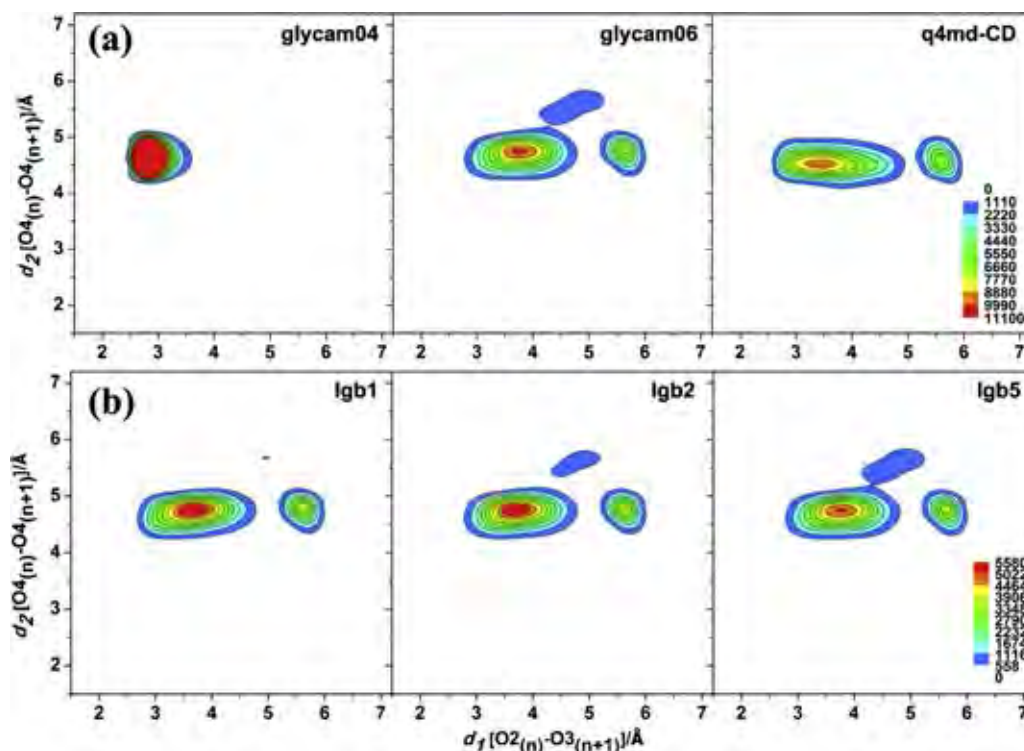


Fig. 2. Contour plots of the probability distribution of overall snapshots of CD10 with (a) the three different force fields (glycam04, glycam06 and q4md-CD force fields) and (b) the three different solvation models (lgb1, lgb2, and lgb5).

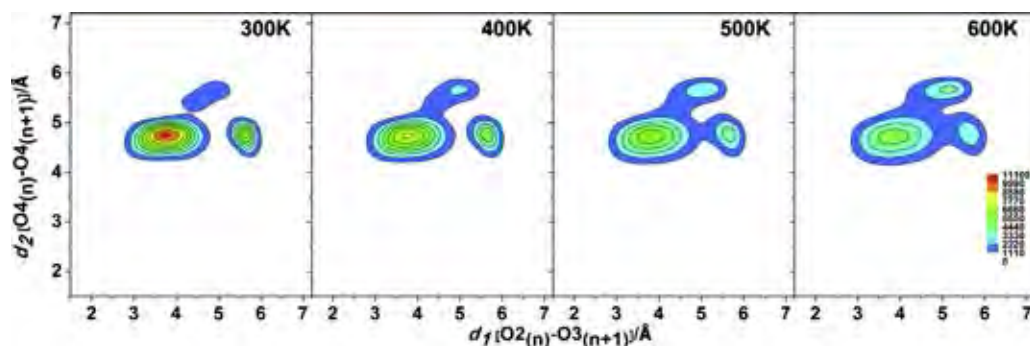


Fig. 3. Contour plots of the probability distribution of overall snap-shots of CD10 at various temperatures using Igb5 solvent model and glycam06 force field.

of the glucose units against each other, including sudden turns leading to a trans arrangement can be also seen from the X-ray crystal structure. Another mode of motion is the conformational change of the glucose units itself, and at least the formation and the breaking of intramolecular hydrogen bonds followed by rotations of the hydroxyl groups and the carbon–carbon bond attached to the primary hydroxyl group. These movements are partly concerted and coupled to a high extent. In particular, the turns of the glucose units are thoroughly correlated with the conformational changes of the large ring system.

In order to estimate the number of turns in various conformations, structural parameters have been sampled for 100 000 snapshots. The glucose subunit is defined to turn when the $\theta[\text{C6}–\text{C2}'–\text{C6}']$ is larger than 60° . Some selected values are shown in Table 2. The snapshots for each type of conformation have been selected randomly. The angle $\theta[\text{C6}–\text{C2}'–\text{C6}']$ proved to be the most sensitive as well as the most significant parameter for the findings of turns. Values around 90° indicate clearly a turn position of a

glucose in the large ring system. Finally, the analysis of the sampled snapshots according to these parameters show a variety of conformations with different numbers of turns in the macrocyclic ring.

Moreover, the shape of CD10 was investigated in terms of the radius of gyration (R_g), the cavity area, and the number of O2–O3 intramolecular hydrogen bonds as shown in Table 3. The cavity area (A) was calculated by the equation given below, where r_i is the distance between the center of mass (COM) of CD10 molecule and each glucose subunit. The number of intramolecular hydrogen bonds was monitored between the –OH at 2-position and the –OH at 3-position according to the following two criteria: (i) a proton donor–acceptor distance ≤ 3.5 Å, and (ii) a donor–H–acceptor bond angle $\geq 120^\circ$. All analyses were computed by the ptraj module in Amber10.

$$A = \frac{\pi}{10} \sum_{i=1}^{10} r_i^2$$

Table 2
Structure properties of classified CD10 conformations as considering the number of turn angle (θ), the $d_1[\text{O2}–\text{O3}']$, $d_2[\text{O4}–\text{O4}']$, $\theta[\text{C6}–\text{C2}'–\text{C6}']$ and $\tau[\text{O2}–\text{C1}–\text{C4}'–\text{O3}']$ of each connecting unit are compared with X-ray and also the previous classical MD simulation.

	R1–R2	R2–R3	R3–R4	R4–R5	R5–R6	R6–R7	R7–R8	R8–R9	R9–R10	R10–R1
$d_1[\text{O2}–\text{O3}']$										
X-ray	2.89	3.03	2.88	3.91	5.53	2.89	3.03	2.88	3.91	5.53
1-Turn	3.90	3.85	3.67	4.85	5.52	3.70	3.65	3.38	3.54	4.77
2-Turn(1–5)	4.00	3.56	5.68	4.03	2.75	3.53	3.91	5.68	3.50	3.30
2-Turn(1–4)	4.05	4.54	3.87	5.40	4.38	3.89	5.12	5.55	2.94	3.13
3-Turn	4.16	4.34	4.71	5.78	3.54	3.93	4.18	4.82	5.85	3.29
4-Turn	4.60	4.85	4.83	3.90	3.12	3.77	4.42	5.88	4.65	3.89
Reference	3.9 ± 0.4^a									
$d_2[\text{O4}–\text{O4}']$										
X-ray	4.60	4.46	4.36	4.48	4.47	4.60	4.46	4.36	4.48	4.47
1-Turn	4.95	4.61	4.49	5.66	4.43	4.51	4.54	4.66	4.80	5.55
2-Turn(1–5)	4.90	4.32	4.93	4.81	4.69	4.78	4.45	4.73	4.65	4.47
2-Turn(1–4)	5.00	5.39	4.72	5.54	4.56	4.62	4.73	4.59	4.55	4.67
3-Turn	4.42	4.52	5.60	4.76	4.63	4.58	4.57	5.44	4.55	4.58
4-Turn	5.51	5.51	5.69	4.42	4.25	4.32	5.61	5.04	5.26	4.54
Reference	4.5 ± 0.2^a									
$\theta[\text{C6}–\text{C2}'–\text{C6}']$										
X-ray	33.0	32.4	26.6	38.7	88.0	33.0	32.4	26.6	38.7	88.0
1-Turn	35.9	29.3	35.0	38.1	86.1	32.7	33.0	30.6	45.2	50.9
2-Turn(1–5)	30.5	29.7	85.7	32.6	28.7	28.7	35.0	87.6	34.9	25.0
2-Turn(1–4)	30.6	57.8	25.6	95.8	34.9	31.6	56.3	88.9	31.5	33.3
3-Turn	35.3	36.0	64.6	88.7	37.0	33.8	44.9	57.6	93.5	36.1
4-Turn	64.5	66.6	64.7	31.9	29.4	34.9	56.5	105.9	56.3	37.6
$\tau[\text{O2}–\text{C1}–\text{C4}'–\text{O3}']$										
X-ray	–38.1	–39.1	–12.9	–75.2	143.4	–38.1	–39.1	–12.9	–75.2	143.4
1-Turn	–61.8	–64.4	1.3	–89.8	101.3	–56.9	–60.2	–46.8	14.0	75.0
2-Turn(1–5)	–78.0	–51.5	113.4	–64.4	–20.5	–53.0	–64.4	131.0	–49.3	–31.3
2-Turn(1–4)	–40.1	–87.7	–31.7	134.3	–73.8	–69.4	–125.9	145.4	–13.4	–41.9
3-Turn	–81.5	–4.4	–101.4	122.2	–58.8	–73.9	16.4	–101.5	131.2	–57.0
4-Turn	–94.5	–105.7	–116.5	–68.9	–37.8	–59.8	–86.1	151.3	–88.1	–9.4

^a MD data of CD10 in solution at 298 K using the force field Glycam04 (Gotsev & Ivanov, 2007a,b).

Table 3
Radius of gyration, cavity area and number of intramolecular hydrogen bonds of CD10.

	Radius of gyration (R_g , Å)	Cavity area (Å ²)	No. of intramolecular H-bond
X-ray	7.6	150.2	ND
1-Turn	8.0	171.2	3
2-Turn(1–5)	7.9	160.7	3
2-Turn(1–4)	7.8	166.2	3
3-Tturn	8.1	179.1	1
4-Turn	8.3	194.1	2
Reference	11.4 ^a	ND	ND

^a MD data of CD10 in solution at 298 K using force field Glycam04 (Gotsev & Ivanov, 2007a).

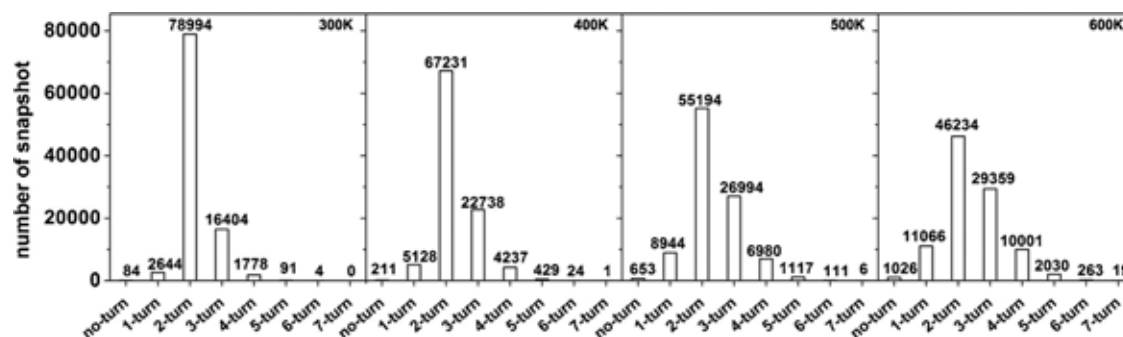


Fig. 4. Probability of turn rotation of glucose units in CD10 ring.

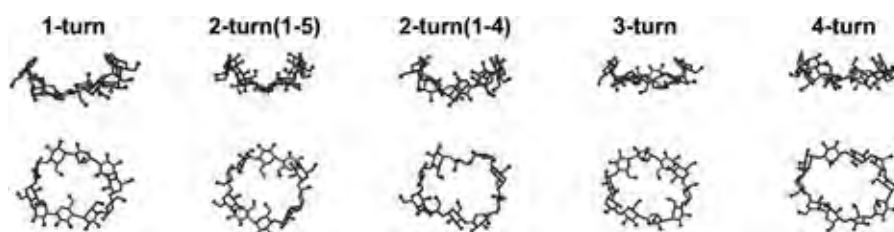


Fig. 5. Examples for CD10 conformations with different numbers of turns, extracted from REMD simulation at 300 K.

In Tables 2 and 3 the various conformations are classified. For reason of comparison the X-ray data are given as well. During the REMD simulation conformations with different numbers of glucose unit flips (turns) can be observed. As already mentioned, two turns could be observed in the crystal structure, whereas the simulations show that there are more possibilities for the appearance of turns. The model 2-turn(1–5) represents a conformation similar to the crystal structure, whereas the 2-turn(1–4) conformation shows a larger and a shorter helical band of glucose units which are located in the macrocyclic ring. Higher numbers of turns are also found during the simulation and no differences between the various arrangements could be seen. In Fig. 4 the probabilities of conformations with different numbers of turns are given both, for the lowest simulation temperature (300 K) as well as for some higher temperatures. At 300 K conformations with 2 glucose flips are dominating to a high extent followed by 3-turn conformations. Only a very small number of conformations without a turn as well as conformations with more than 4 turns can be observed. Thus obviously it can be assumed that these conformations exist in a rapid dynamical equilibrium. From Table 3, it can be seen that the cavity of CD10 was somewhat enlarged at higher number of turns because of the decreasing number of intramolecular hydrogen bonds. Moreover, the radii of gyration of the REMD simulations are close to the value of the crystal structure (Jacob et al., 1998) in contrary to the value of the previous simulation (Gotsev & Ivanov, 2007a).

The temperature dependence of the distance probabilities (Fig. 3) is in accordance with the distribution scattering increase due to temperature changes. These findings are correlating with the population of 0- and 1-turn conformations as well as conformations with higher number of turns. This means that conformations with other turn numbers than 2 are energetically less favored, and moreover, higher temperatures lead to higher populations of these energetically less favored conformations. In Fig. 5 some examples of snap-shot geometries are given for various types of turn-conformations. It can be seen that the conformation of a saddle-point structure, as observed for the crystal structure, is existing for 1-turn, 2-turn and to some extent for 3-turn geometries, whereas higher numbers of turns lead to more flat conformations of the large ring system.

In Table 4 the probabilities of the different conformations obtained from the snap-shots of the LR-system classified by the various turns are given. In REMD simulations, we could calculate differences in free-energy and other thermodynamic properties among these conformations from these probabilities (Okumura, 2012; van der Spoel & Seibert, 2006). Because the 2-turn model is the most dominant conformation, the free-energy difference of other n -turn models from this model is given by

$$\Delta F = F_{n\text{-turn}} - F_{2\text{-turn}} = -RT \ln \frac{f_{n\text{-turn}}}{f_{2\text{-turn}}}$$

Table 4The probability distribution of CD10 conformation with different number of turns and the thermodynamics quantities ΔF , ΔE , and $T\Delta S$ at 300 K.

	Probability				ΔF at 300 K (kJ/mol)	ΔE at 300 K (kJ/mol)	$T\Delta S$ at 300 K (kJ/mol)
	300 K	400 K	500 K	600 K			
No-turn	0.001	0.002	0.007	0.010	17.1	15.4	-1.7
1-Turn	0.026	0.051	0.089	0.111	8.5	9.9	1.5
2-Turn	0.790	0.672	0.552	0.462	0.0	0.0	0.0
3-Turn	0.164	0.227	0.270	0.294	3.9	5.5	1.6
4-Turn	0.018	0.042	0.070	0.100	9.5	11.2	1.7
5-Turn	0.001	0.004	0.011	0.020	16.9	18.1	1.3
6-Turn	0.000	0.000	0.001	0.003	24.7	23.6	-1.0
7-Turn	-	0.000	0.000	0.000	-	-	-
8-Turn	-	-	-	0.000	-	-	-

where $f_{n\text{-turn}}$ is the probability of n -turn conformation and R is the gas constant: $R=8.3145\text{ J}/(\text{mol}\cdot\text{K})$. The energy difference is given from the temperature dependence by

$$\Delta E = E_{n\text{-turn}} - E_{2\text{-turn}} = \frac{\partial(\Delta F/T)}{\partial(1/T)} = -R \frac{\partial(\ln(f_{n\text{-turn}}/f_{2\text{-turn}}))}{\partial(1/T)}$$

Energy difference was calculated from linear fitting of the data. Entropy term $T\Delta S$ can be calculated by

$$T\Delta S = \Delta E - \Delta F$$

These thermodynamic quantities are also listed in Table 4. It should be mentioned that the 7- and 8-turn conformations are only observed at high temperature with a very small probability. The corresponding thermodynamic properties were not calculated.

Referring to the 2-turn conformations several arrangements are possible. The preferred conformation are given by 2-turn(1–5) and 2-turn(1–4) models showing also comparable probabilities. Other models (e.g. 2-turn(1–3)) show a much less probability with energy differences around 2 kcal/mol compared with the lowest energy model.

4. Conclusions

In this study the structural changes of CD10 were simulated by 100 ns REMD. With the glycam06 force field together with the lgb5 solvent model a good description of CD10 could be obtained. The advantage of REMD simulations consists in the fact that a complete picture of the conformational changes of large ring systems with several coupled modes of motion can be achieved. Additionally, it is possible to get some insight into the temperature dependence of conformational changes. In the case of CD10, in turn for which an X-ray crystal structure exists, several molecular descriptors have been selected and tested to explain in particular the flip or the turn of glucose ring subunits within the frame of the macrocyclic ring system. There is some evidence that geometries with two turns of glucose subunits are thermodynamically more favored than other conformations. These findings are in accordance with the crystal structure of CD 10 which shows two turns and two slightly distorted band helices. Increase of energy content due to elevated temperatures is tightly correlated with an increase of diversity of various conformations of the macrocycle. Concluding REMD can be ascertained to give a complete picture of the dynamical behavior of large ring cyclodextrins.

Acknowledgments

We would like to give special thanks to Prof. Heinz Berner for editing the manuscript. The authors would like to thank National Research Council of Thailand and the 90th anniversary of Chulalongkorn University fund for financial support. The Computer Chemistry Unit Cell (CCUC), and the Vienna Scientific Cluster

(VSC-2) are also thanked for facilities and computing resources. T.R. acknowledges the TRF-CHE Research Grant for New Scholars (MRG5580223). Chiang Mai University is also acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.carbpol.2015.10.018.

References

- Basma, M., Sundara, S., Calgan, D., Vernali, T., & Woods, R. J. (2001). Solvated ensemble averaging in the calculation of partial atomic charges. *Journal of Computational Chemistry*, 22(11), 1125–1137.
- Cao, R., & Wu, S. (2015). In silico properties characterization of water-soluble γ -cyclodextrin bi-capped C60 complex: Free energy and geometrical insights for stability and solubility. *Carbohydrate Polymers*, 124, 188–195.
- Case, D., Darden, T., Cheatham Iii, T., Simmerling, C., Wang, J., Duke, R., Luo, R., Crowley, M., Walker, R. C., & Zhang, W. (2008). *AMBER 10*. pp. 32. San Francisco: University of California.
- Cezard, C., Trivelli, X., Aubry, F., Djedaini-Pilard, F., & Dupradeau, F.-Y. (2011). Molecular dynamics studies of native and substituted cyclodextrins in different media: 1. Charge derivation and force field performances. *Physical Chemistry Chemical Physics*, 13(33), 15103–15121.
- Cheng, X., Cui, G., Hornak, V., & Simmerling, C. (2005). Modified replica exchange simulation methods for local structure refinement. *Journal of Physical Chemistry B*, 109(16), 8220–8230.
- Dodziuk, H., Ejchart, A., Anczewski, W., Ueda, H., Krinichnaya, E., Dolgonos, G., & Kutner, W. (2003). Water solubilization, determination of the number of different types of single-wall carbon nanotubes and their partial separation with respect to diameters by complexation with η -cyclodextrin. *Chemical Communication (Cambridge, U.K.)*, 8, 986–987.
- Fernandes, A., Ivanova, G., Brás, N. F., Mateus, N., Ramos, M. J., Rangel, M., & Freitas, V. (2014). Structural characterization of inclusion complexes between cyanidin-3-O-glucoside and β -cyclodextrin. *Carbohydrate Polymers*, 102, 269–277.
- Gessler, K., Usón, I., Takaha, T., Krauss, N., Smith, S. M., Okada, S., Sheldrick, G. M., & Saenger, W. (1999). V-Amylose at atomic resolution: X-ray structure of a cycloamylose with 26 glucose residues (cyclomaltohexaicosaoase). *Proceedings of the National Academy of Science of the United States of America*, 96(8), 4246–4251.
- Gotsev, M. G., & Ivanov, P. M. (2007a). Large-ring cyclodextrins. A molecular dynamics study of the conformational dynamics and energetics of CD10, CD14 and CD26. *ARKIVOC*, 13, 167–189.
- Gotsev, M. G., & Ivanov, P. M. (2007b). Large-ring cyclodextrins. Further support for the preferred conformations of CD26 in water solution: Molecular dynamics studies on CD26-derived conformations of CDn ($n=24, 25, 27, 28, 29, 30$). *International Journal of Quantum Chemistry*, 107(8), 1657–1672.
- Gotsev, M. G., & Ivanov, P. M. (2009). Molecular dynamics of large-ring cyclodextrins: Principal component analysis of the conformational interconversions. *Journal of Physical Chemistry B*, 113(17), 5752–5759.
- Gotsev, M. G., Ivanov, P. M., & Jaime, C. (2007). Molecular dynamics study of the conformational dynamics and energetics of some large-ring cyclodextrins (CDn, $n=24, 25, 26, 27, 28, 29$). *Chirality*, 19(3), 203–213.
- Hansmann, U. H. (1997). Parallel tempering algorithm for conformational studies of biological molecules. *Chemical Physics Letters*, 281(1), 140–150.
- Harata, K., Endo, T., Ueda, H., & Nagai, T. (1998). X-ray structure of α -cyclodextrin. *Supramolecular Chemistry*, 9(2), 143–150.
- Hu, Q.-D., Tang, G.-P., & Chu, P. K. (2014). Cyclodextrin-based host-guest supramolecular nanoparticles for delivery: From design to applications. *Accounts of Chemical Research*, 47(7), 2017–2025.
- Ivanov, P. (2012). Conformations of some lower-size large-ring cyclodextrins derived from conformational search with molecular dynamics and principal component analysis. *Journal of Molecular Structure*, 1009, 3–10.

- Ivanov, P. M. (2011). Computational studies on the conformations of some large-ring cyclodextrins (CDn, n = 20, 21, 22, 23). *Chirality*, 23(8), 628.
- Ivanov, P. M., & Jaime, C. (2004). Insights into the structure of large-ring cyclodextrins through molecular dynamics simulations in solution. *Journal of Physical Chemistry B*, 108(20), 6261–6274.
- Jacob, J., Ge ler, K., Hoffmann, D., Sanbe, H., Koizumi, K., Smith, S. M., Takaha, T., & Saenger, W. (1999). Band-flip and kink as novel structural motifs in a-(1-4)-D-glucose oligosaccharides. Crystal structures of cyclodeca- and cyclotetradecaamylose. *Carbohydrate Research*, 322(3), 228–246.
- Jacob, J., Geßler, K., Hoffmann, D., Sanbe, H., Koizumi, K., Smith, S. M., Takaha, T., & Saenger, W. (1998). Strain-induced “Band Flips” in cyclodecaamylose and higher homologues. *Angewandte Chemie International Edition*, 37(5), 605–609.
- Khuntawee, W., Wolschann, P., Rungrotmongkol, T., Wong-Ekkabut, J., & Hannongbua, S. (2015). Molecular dynamics simulations of the interaction of beta cyclodextrin with lipid bilayer. *Journal of Chemical Information Modeling (in press)*
- Kirschner, K. N., & Woods, R. J. (2001a). Quantum mechanical study of the nonbonded forces in water–methanol complexes. *Journal of Physical Chemistry A*, 105(16), 4150–4155.
- Kirschner, K. N., & Woods, R. J. (2001b). Solvent interactions determine carbohydrate conformation. *Proceedings of the National Academy of Sciences of the United States of America*, 98(19), 10541–10545.
- Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., & Woods, R. J. (2008). GLYCAM06: A generalizable biomolecular force field. *Carbohydrates. Journal of Computational Chemistry*, 29(4), 622–655.
- Larsen, K. (2002). Large cyclodextrins. *Journal of Inclusion Phenomena and Macrocyclic Chemistry*, 43(1–2), 1–13.
- Larsen, K. L., Endo, T., Ueda, H., & Zimmermann, W. (1998). Inclusion complex formation constants of α -, β -, γ -, δ -, ϵ -, ζ -, η - and θ -cyclodextrins determined with capillary zone electrophoresis. *Carbohydrate Research*, 309(2), 153–159.
- Maestre, I., Beà, I., Ivanov, P. M., & Jaime, C. (2007). Structural dynamics of some large-ring cyclodextrins. A molecular dynamics study: An analysis of force field performance. *Theoretical Chemistry Accounts*, 117(1), 85–97.
- Nimz, O., Geßler, K., Usón, I., & Saenger, W. (2001). An orthorhombic crystal form of cyclohexacosaoose, CA26-32.59 H₂O: Comparison with the triclinic form. *Carbohydrate Research*, 336(2), 141–153.
- Nutho, B., Khuntawee, W., Rungnim, C., Pongsawasdi, P., Wolschann, P., Karpfen, A., Kungwan, N., & Rungrotmongkol, T. (2014). Binding mode and free energy prediction of fisetin/ β -cyclodextrin inclusion complexes. *Beilstein Journal of Organic Chemistry*, 10, 2789–2799.
- Nymeyer, H., Gnanakaran, S., & Garcia, A. E. (2004). Atomic simulations of protein folding, using the replica exchange algorithm. *Methods in Enzymology*, 383, 119–149.
- Ogawa, N., Takahashi, C., & Yamamoto, H. (2015). Physicochemical characterization of cyclodextrin–drug interactions in the solid state and the effect of water on these interactions. *Journal of Pharmaceutical Sciences*, 104(3), 942–954.
- Okumura, H. (2012). Temperature and pressure denaturation of chignolin: Folding and unfolding simulation by multibaric-multithermal molecular dynamics method. *Proteins: Structure, Function, and Bioinformatics*, 80(10), 2397–2416.
- Rungnim, C., Rungrotmongkol, T., Hannongbua, S., & Okumura, H. (2013). Replica exchange molecular dynamics simulation of chitosan for drug delivery system based on carbon nanotube. *Journal of Molecular Graphics and Modelling*, 39, 183–192.
- Sangpheak, W., Khuntawee, W., Wolschann, P., Pongsawasdi, P., & Rungrotmongkol, T. (2014). Enhanced stability of a naringenin/2,6-dimethyl beta-cyclodextrin inclusion complex: Molecular dynamics and free energy calculations based on MM- and QM-PBSA/GBSA. *Journal of Molecular Graphics and Modelling*, 50, 10–15.
- Shimada, J., Handa, S., Kaneko, H., & Takada, T. (1996). Conformation of novel cycloamylose: Topological aspects and simulations. *Macromolecules (Washington, DC, U.S.)*, 29(20), 6408–6421.
- Shimada, J., Kaneko, H., Takada, T., Kitamura, S., & Kajiwara, K. (2000). Conformation of amylose in aqueous solution: Small-angle x-ray scattering measurements and simulations. *Journal of Physical Chemistry B*, 104(9), 2136–2147.
- Sugita, Y., & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1–2), 141–151.
- Tessier, M. B., Demarco, M. L., Yongye, A. B., & Woods, R. J. (2008). Extension of the GLYCAM06 biomolecular force field to lipids, lipid bilayers and glycolipids. *Molecular Simulation*, 34(4), 349–363.
- Ueda, H. (2002). Physicochemical properties and complex formation abilities of large-ring cyclodextrins. *Journal of Inclusion Phenomena and Macrocyclic Chemistry*, 44(1–4), 53–56.
- Ueda, H. (2004). Large ring cyclodextrins – Recent progress. *FABAD Journal of Pharmaceutical Science*, 29, 27–38.
- Ueda, H., Endo, T., Nagase, H., Kobayashi, S., & Nagai, T. (1996). Isolation, purification, and characterization of cyclomaltodecaose (ϵ -CD). *Journal of Inclusion Phenomena and Molecular Recognition in Chemistry*, 25(1–3), 17–20.
- van der Spoel, D., & Seibert, M. M. (2006). Protein folding kinetics and thermodynamics from atomistic simulations. *Physical Review Letters*, 96(23), 238102.
- Wang, J., Wei, R., Tian, Y., Yang, N., Xu, X., Zimmermann, W., & Jin, Z. (2015). Multi-wavelength colorimetric determination of large-ring cyclodextrin content for the cyclization activity of 4- α -glucanotransferase. *Carbohydrate Polymers*, 122, 329–335.



Inclusion complexation of pinostrobin with various cyclodextrin derivatives



Jintawee Kicuntod^{a,b}, Wasinee Khuntawee^c, Peter Wolschann^{a,d,e},
Piamsook Pongsawasdi^b, Warinthorn Chavasiri^f, Nawe Kungwan^g,
Thanyada Rungrotmongkol^{a,h,*}

^a Structural and Computational Biology Unit, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

^b Starch and Cyclodextrin Research Unit, Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

^c Nanoscience and Technology Program, Graduate School, Chulalongkorn University, Bangkok 10330, Thailand

^d Department of Pharmaceutical Technology and Biopharmaceutics, University of Vienna, Vienna 1090, Austria

^e Institute of Theoretical Chemistry, University of Vienna, Vienna 1090, Austria

^f Natural Products Research Unit, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

^g Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

^h Program in Bioinformatics and Computational Biology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

ARTICLE INFO

Article history:

Received 29 June 2015

Received in revised form 3 November 2015

Accepted 6 November 2015

Available online 12 November 2015

Keywords:

Pinostrobin

Cyclodextrin

Inclusion complex

Molecular dynamics simulations

Phase solubility

ABSTRACT

Pinostrobin (PNS) is one of the important flavonoids and can be abundantly found in the rhizomes of fingerroot (*Boesenbergia rotunda*) and galangal (*Alpinia galangal* and *Alpinia officinarum*), the herbal basis of Southeast Asian cooking. Similar to other flavonoids, PNS exhibits anti-oxidative, anti-inflammatory and anti-cancer properties. However, this compound has an extremely low water solubility that limits its use in pharmaceutical applications. Beta-cyclodextrin (β CD) and its derivatives, 2,6-dimethyl- β CD (2,6-DM β CD) and the three hydroxypropyl- β CDs (2-HP β CD, 6-HP β CD and 2,6-DHP β CD), have unique properties that enhance the stability and solubility of such low-soluble guest molecules. In the present study, molecular dynamics simulations were applied to investigate the dynamics and stability of PNS inclusion complexes with β CD and its derivatives (2,6-DM β CD, 2,6-DHP β CD, 2-HP β CD and 6-HP β CD). PNS was able to form complexes with β CD and all four of its derivatives by either the chromone (C-PNS) or phenyl (P-PNS) ring dipping toward the cavity. According to the molecular mechanics-generalized Born surface area binding free energy values, the stability of the different PNS/ β CD complexes was ranked as 2,6-DHP β CD > 2,6-DM β CD > 2-HP β CD > 6-HP β CD > β CD. These theoretical results were in good agreement with the stability constants that had been determined by the solubility method.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Flavonoids are a large group of heterocyclic compounds that are commonly found in fruits, vegetables and herbs as plant secondary metabolites [1,2]. A fairly diversified range of bioactivities, such as anti-bacterial, anti-allergic and anti-oxidative activities, have been reported for flavonoids [3,4]. They are widely used as drug and dietary supplements due to their potential pharmacological properties and their rather low toxicity. Pinostrobin (PNS) belongs to the flavanone subclass of flavonoids (Fig. 1A) and can be extracted from the rhizomes of Thai galangal (*Alpinia rotunda* and *Alpinia*

officinarum) and Chinese ginger (*Boesenbergia rotunda*). It has several important biological activities, such as an anti-inflammatory role on the cyclooxygenase activity [5], anti-aromatase activity [6], decreased growth rate of the MCF-7 breast cancer cell line [6] and inhibition of HIV-1 protease [7]. Furthermore, the activity of the β -amyloid peptide related to Alzheimer's disease can be inhibited by PNS through reducing the oxidative damage and calcium overload as well as suppressing the mitochondrial pathway that is involved in the cellular apoptosis [8]. Similar to many flavonoids, PNS exhibits a relatively low water solubility leading to a significant limitation in its use in pharmaceutical applications. Consequently, suitable drug delivery carriers are of interest to solve this problem.

Cyclodextrins (CDs) are macrocyclic oligosaccharides of α -1,4 linked D-(+)-glucopyranose produced from starch by cyclodextrin glycosyl transferase catalysis [9]. CD structures have a trun-

* Corresponding author. Fax: +66 2218 5418.

E-mail addresses: t.rungrotmongkol@gmail.com, thanyada.r@chula.ac.th (T. Rungrotmongkol).

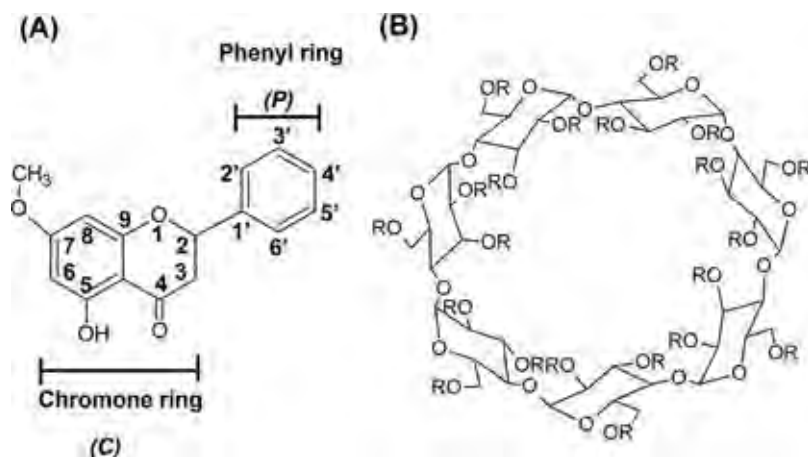


Fig. 1. Two dimensional structures of (A) pinostrobin (PNS) and (B) β -cyclodextrin (β CD) and its derivatives, where $-R$ is $-H$, $-CH_3$ and $-C_3H_7O$ for β CD, 2,6-DM β CD and 2,6-DHP β CD, respectively.

cated cone shape with a relatively hydrophilic outer surface and a hydrophobic inner cavity. The wider rim of β CD consists of hydroxyl groups at the 2- and 3-position of each glucose subunit, while the other narrower rim contains hydroxyl groups at the 6-position of each glucose subunit.

The inclusion of poorly water-soluble drugs into β CD can be used to enhance the solubility, stability and bioavailability of these drugs [10–12], where the hydrophobic drugs prefer to insert their hydrophobic motif(s) inside the CD cavity [13]. Naturally, there are three major types of small-ring CDs; namely α CD, β CD (Fig. 1B) and γ CD that are composed of six, seven and eight glucose subunits, respectively [14]. Due to the high yield synthesis and low price of β CD it has been the main form used in the pharmaceutical, food and cosmetic industries [15]. Albeit, β CD shows a relatively lower water solubility (18.5 mg/mL) than other CDs [16,17], which limits its applications. Derivatives of β CD, for example 2,6-dimethyl β CD (2,6-DM β CD) and 2-hydroxypropyl β CD (HP β CD) that have a higher water solubility (570 and >600 mg/mL, respectively) [18,19], have been shown to improve the bioavailability and bioactivity of the encapsulated molecule in a much more efficient way [20–22].

Experimental studies to determine the stability and phase solubility of inclusion complexes between CDs and flavonoids have been reported [23,24], and have shown that a higher solubility of the CD derivatives and better affinities between host and guest molecules are advantages for pharmaceutical applications. For example, phase solubility studies on rutin forming 1:1 molar ratio complexes with HP α CD, HP β CD and HP γ CD revealed that the complex stability constants with HP β CD and HP γ CD were significantly increased [25]. Accordingly, the solubility of artesunate, a low water soluble antimalarial drug, can be improved by a 1:1 molar ratio formation with methyl- β CD more than with HP β CD or β CD [26].

In addition to experimental investigations, computational tools are useful to determine the preferable binding mode of drugs, such as flavonoids, and for the prediction of guest-host interactions as well as the stability of inclusion complexes at a molecular level [27]. For example, the molecular dynamics simulations (MDs) of quercetin and myricetin complexed with different CD derivatives provided the host-guest orientation in a good agreement with the 1H NMR results [28]. In addition, although there are various possible inclusion geometries of fisetin/ β CD complexes, it was found by molecular simulations that the insertion of the phenyl ring inside the cavity of β CD was more favorable [29]. Moreover, the molecular mechanics-Poisson-Boltzmann surface area/-generalized Born surface area (MM-PBSA/GBSA) and quantum mechanics (QM)-PBSA/GBSA binding free energy calculations

predicted that naringenin would bind to 2,6-DM β CD better than to natural β CD [30].

To date, theoretical and experimental studies of PNS in inclusion complexes with various β CDs have not been reported. Thus, the main aim of this study was to predict computationally the most suitable β CD derivative for PNS encapsulation, and then the PNS/ β CD complexes were formed experimentally for further pharmaceutical applications. In addition, the molecular details of the PNS binding mode and orientation, stability and solvation in the inclusion complex with β CD and its derivatives, 2,6-DM β CD and the three HP β CDs (2- and 6-HP β CD and 2,6-DHP β CD), were discussed and compared.

2. Methods

2.1. Computation

2.1.1. System preparation

The optimized structures of β CD and 2,6-DM β CD were taken from our previous studies [31]. For the three HP β CD derivatives, the structures of 2-HP β CD, 6-HP β CD and 2,6-DHP β CD were prepared by 2-hydroxypropyl substitutions on all 2-, 6- and both 2- and 6-hydroxyl positions of the natural β CD, respectively. The PNS geometry was built and then optimized by the HF/6-31(d) level of theory using the Gaussian09 program to obtain well-adjusted bond lengths and angles [32]. The inclusion complex between PNS and each respective (modified or not) β CD was constructed by the CDOCKER module in the Discovery Studio 2.5 (Accelrys, Inc.) with 500 independent docking runs. The complex with the best ranked interaction energy and highest hydrogen bond (H-bond) formation at each PNS binding mode, chromone (C-PNS) or phenyl ring (P-PNS) dipping into the β CD cavity, was chosen as the representative structure. Where only one binding orientation was obtained by the docking procedure, another was manually generated for comparison. In total, there were 15 systems for further MD studies; five free CDs (β CD, 2,6-DM β CD, 2,6-DHP β CD, 2-HP β CD and 6-HP β CD) plus these five different CDs complexed with either P-PNS or C-PNS. Note that the docking manner was used to determine the starting position of PNS in β CDs. Hence, the molecular dynamics (MD) simulation was performed to clarify the importance of solvent effects in the forming inclusion complexes.

2.1.2. Molecular dynamics (MD) simulation

The structure of β CD and its dimethyl and hydroxypropyl derivatives alone and complexed with PNS in aqueous solution were simulated with three different initial velocities using the

Amber 12 software package [33]. The parameters of β CD and its derivatives were taken from the Glycam06 carbohydrate force field [34], while the partial charges and empirical parameters of PNS were obtained by the standard procedures described elsewhere [35–39]. To relax the structure prior to simulation, the hydrogen atoms of each system were minimized with 1000 steps of steepest descents (SD) and continued by 3000 steps of conjugated gradient (CG). The simple point charge (SPC) water molecules were then used to solvate around the inclusion complexes with the minimum distance of 12 Å from the system surface. As a result, the truncated octahedron periodic water box size of the simulated systems was approximately $45 \times 45 \times 45 \text{ \AA}^3$ and consisted of 1400 ± 42 water molecules. The SD (1000 steps) and CG (3000 steps) minimizing procedure was applied to the whole system. The periodic boundary condition with *NPT* ensemble was applied for all simulations. The cutoff distance for electrostatic interactions was set at 12 Å according to the particle mesh Ewald approach [40]. To constrain all bonds involving hydrogen atoms, the SHAKE algorithm was used. The systems were heated up to 298 K for 100 ps, and then were continuously held at constant temperature of 289 K for 80 ns. The integration time was set at 2 fs and the MD trajectories were collected every 2 ps. The equilibrium state of simulated systems was considered by the root mean square displacement (RMSD). The conformation changes in β CD and its derivatives in the free form and when complexed with PNS were discussed in terms of the potential energy surface (PES). The mobility and water accessibility of PNS in the cavity of β CD and each derivative were determined by measurement of the distances between PNS and the respective β CD or derivative molecule and the radial distribution functions (RDF) of water around the PNS hetero atoms, respectively. Using the same trend of binding free energy prediction reported previously using the MM- and QM-PBSA/GBSA methods in the flavonoid/CD complexes [29,30], the MM-GBSA approach was adopted to estimate the binding affinity of PNS and β CDs in this study. All structural data were analyzed by the ptraj module, while the binding free energy was calculated by MM-GBSA module implemented in Amber program.

In addition, to test the simulation convergence, the two starting structures of C-PNS inserted, (A1) its chromone ring in β CD cavity and (A2) phenyl ring in the cavity instead, were separately simulated for 20 ns. The same procedure was prepared for P-PNS (B1 and B2). The results of distance and distribution plots of the centers of gravity between pinostrobin and β CD are newly given in Fig. S1 in supplementary materials. It can be seen that after 2-ns simulation the (A2) C-PNS has changed its position from $\sim -2 \text{ \AA}$ to $\sim 2 \text{ \AA}$ as found for (A1) C-PNS. The similar phenomenon was observed for P-PNS. This finding indicates the convergence of simulation and the intermediate position of pinostrobin in the hydrophobic cavity.

2.2. Experiment

2.2.1. Materials

The PNS was extracted from the co-author synthesis, while β CD and its derivatives (2,6-DM β CD and HP β CD) were purchased from Wako pure chemical industries, Ltd. and Sigma-Aldrich, respectively, and were of technical grade only.

2.2.2. Phase solubility study

An excess amount of PNS was dissolved in 0–10 mM β CD, 2,6-DM β CD and HP β CD, respectively. The mixtures were shaken at 25 °C for 72 h in a water bath. The dissolved PNS concentration was determined by UV–vis spectrophotometer at 290 nm. Each experiment was performed in triplicate. Additionally, the solubility of PNS in pure water was also determined in the same manner. The phase solubility diagrams of PNS in the presence of various concentrations of β CD and its derivatives were obtained using the method of

Higuchi and Connors [41]. The stability constant (K_C) of PNS when complexed with the different β CDs was calculated from the slope data of the phase solubility linear diagrams according to Eq. (1);

$$K_C = \frac{\text{Slope}}{S_0(1 - \text{Slope})} \quad (1)$$

where S_0 is the saturation concentration of PNS in water at 25 °C.

3. Results and discussion

The three independent simulations on each system provided relatively similar results. Thus, for simplification the data from only one simulation are presented here.

3.1. System stability

To monitor the system stability, the RMSD with respect to the initial structure of the free CDs (β CD, 2,6-DM β CD and the three HP β CDs), and their complexes with PNS (as C-PNS and P-PNS) was calculated. The results are illustrated in Fig. 2, where, as expected, the averaged RMSD of the parent β CD was smaller than those of all the modified β CDs. Among the uncomplexed CDs, a noticeable fluctuation in the RMSD was observed in 6-HP β CD, which contains the hydroxypropyl substitutions at the 6-position, as found previously [42]. Complexation with PNS (both P-PNS and C-PNS conformations) led to a reduction in the RMSD values in all inclusion complexes. For instance, the RMSD of the C-PNS and P-PNS in complex with 2,6-DHP β CD was observed in a range of 3.0–3.5 Å, whilst the RMSD of the free 2,6-DHP β CD was larger than 4 Å. The differences in the RMSD fluctuations between the two relative complexes resulted from the structural and dynamic properties of PNS binding inside the host cavity (discussed later in Fig. 4). From the RMSD plots, it seems that all the simulated systems had reached equilibrium by 40 ns (except for the free 6-HP β CD), and so for each system the MDs trajectories from the 60 ns were extracted for further analysis.

3.2. Conformation of β CDs in free and complex forms

As seen from the RMSD plots, PNS might affect the conformation of the host β CDs molecule. To investigate such conformational changes, the distance between the glycosidic oxygen atoms ($O4_{(n)} - O4_{(n+1)}$) and the distances between the hydroxyl groups on the secondary rim of neighboring glucoses ($O3_{(n)} - O2_{(n+1)}$) were monitored from the last 40-ns snapshots. It has to be noticed that the second distance related to intramolecular hydrogen bond formation refers to the structural characteristic feature of CDs [43]. Using Eq. (2), the distributions of these two distances were calculated and are presented in terms of free energy in Fig. 3.

$$F(x, y) = -k_B T \log[P(x, y)] \quad (2)$$

where k_B is the Boltzmann constant and T is the absolute temperature. $P(x, y)$ is the probability of the parameters distance ($O3_{(n)} - O2_{(n+1)}$) and distance ($O4_{(n)} - O4_{(n+1)}$) as x and y , respectively.

In Fig. 3, the three minima M1, M2 and M3 were detected from the PES of parent β CD, with a free energy ranked in the order of $M1 < M2 < M3$. For the most probable conformation of β CD (M1), the $O3_{(n)} - O2_{(n+1)}$ distance was distributed in range of 3–5 Å with a possible formation of intramolecular hydrogen bonds on the wider rim ($\leq 3.5 \text{ \AA}$), while the $O4_{(n)} - O4_{(n+1)}$ distance was likely presented at $\sim 4.5 \text{ \AA}$. At the same $O4_{(n)} - O4_{(n+1)}$ distance, an outspread of the $O3_{(n)} - O2_{(n+1)}$ distance was found for M2. In contrast, the $O4_{(n)} - O4_{(n+1)}$ distance for M3 was lengthened over 5 Å and the $O3_{(n)} - O2_{(n+1)}$ distance was in between those of M1 and M2.

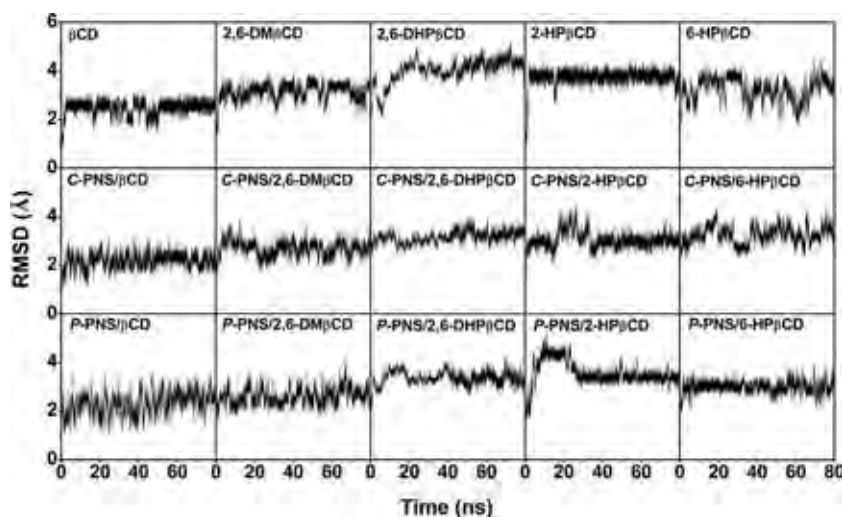


Fig. 2. RMSD plots of all atoms for the 15 simulated systems (β CD and its four derivatives alone and each complexed with either C-PNS or P-PNS) versus the simulation time.

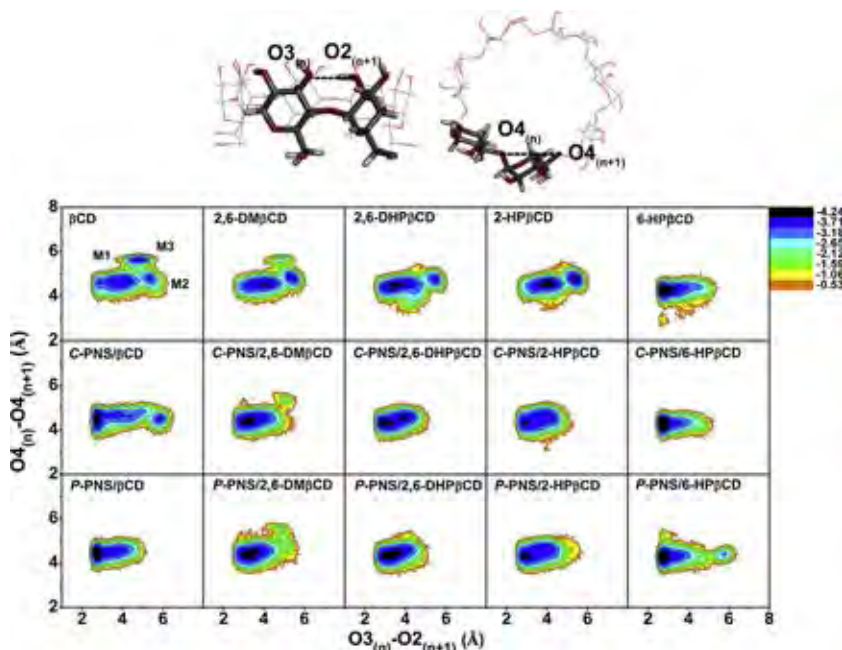


Fig. 3. The potential energy surface (PES) plots of the distance between the adjacent glycosidic oxygens, $O4_{(n)} - O4_{(n+1)}$, versus the intramolecular hydrogen bond distance, $O3_{(n)} - O2_{(n+1)}$, for the different β CDs (top row) alone or complexed with (middle row) C-PNS or (bottom row) P-PNS.

Among all β CD derivatives in free form, only 2,6-DM β CD shared a similar PES to the unmodified β CD, while the M3 had totally disappeared within the others. Upon complexation with PNS the binding of PNS forced the β CDs structure to adopt the most stable conformation (a more dominant M1) with stronger intramolecular hydrogen bonds (higher proportion of $O3_{(n)} - O2_{(n+1)}$ with a distance ≤ 3.5 Å). In addition, the shape of β CDs also depended on the PNS binding mode, in particular for β CD and 6-HP β CD. Both M1 and M2 were detected in C-PNS/ β CD and P-PNS/6-HP β CD complexes, whereas in the other cases only an M1 conformation was observed.

3.3. PNS mobility and preferential displacement in CD cavity

The dynamic behavior of the C-PNS- or P-PNS binding toward the five different β CDs were determined by monitoring the distance from the center of gravity of PNS molecule to the center of gravity of each different CD without considering the functional

modifications of the respective modified β CD molecule. The results of the 10 inclusion complexes were plotted and compared in Fig. 4, where the horizontal dashed lines at -3.95 Å and 3.95 Å represent the positions of the primary and secondary rims of β CD, respectively [44]. It can clearly be seen from the distance plots that PNS in both the C-PNS and P-PNS orientations can form inclusion complexes well with all five different CDs. The two important factors, (i) the substitutions on 2- and 6-positions of β CD and (ii) the mode of PNS binding, are likely to be involved in the mobility and preferential displacement of PNS inside the hydrophobic cavity of CDs. Where the same substitutions on both 2- and 6-positions of β CD (i.e., β CD, 2,6-DM β CD, and 2,6-DHP β CD), the C-PNS stably positioned nearby the secondary rim (~ 3 Å), while P-PNS preferred a deeper location and showed a higher movement along the simulation due to a reduced fitting of its benzene ring within the cavity. For 2-hydroxypropyl substitutions on the 2- or 6-position of β CD (2-HP β CD or 6-HP β CD), the hydrophilic functional groups signif-

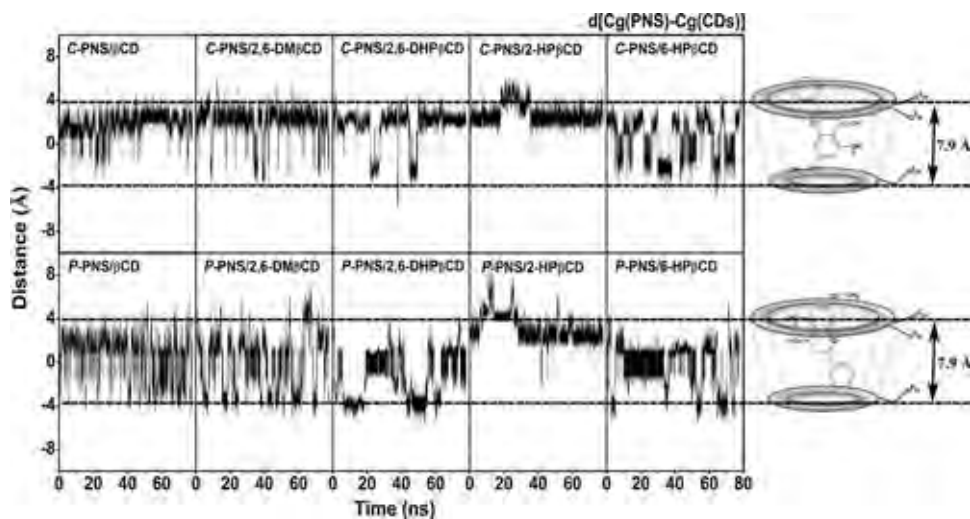


Fig. 4. Distance between the centers of gravity of PNS and CD excluding the modified functional groups for all systems. Dashed lines represent the cavity height of regular β CD.

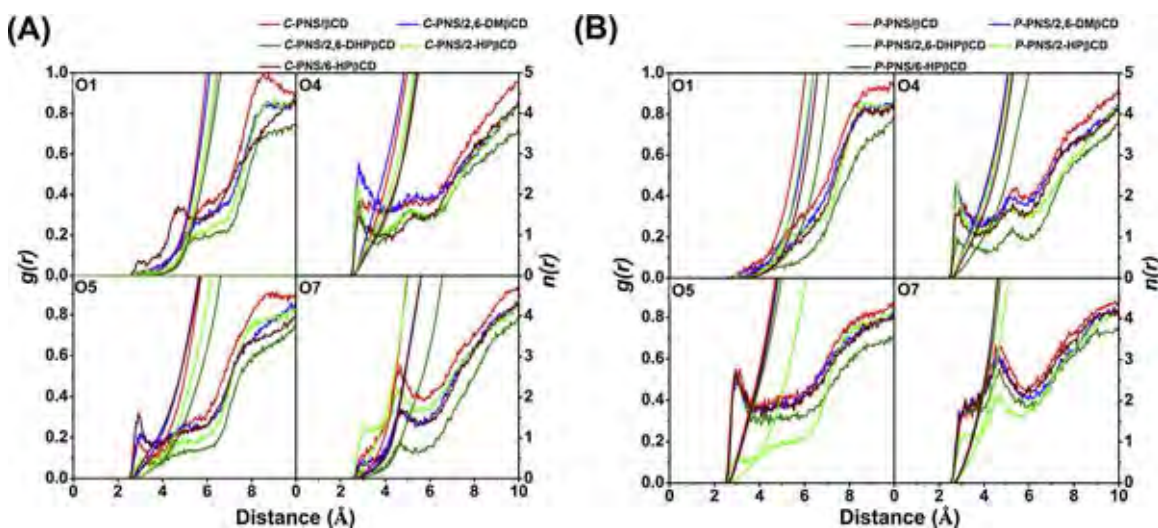


Fig. 5. The radial distribution function (RDF) of water molecules around the PNS oxygen atoms (as defined in Fig. 1A) for (A) C-PNS/CDs and (B) P-PNS/CDs inclusion complexes.

icantly induced PNS displacement to be relatively nearer to the substituted rim. In 2-HP β CD, the PNS molecule initially moved up over 4 Å and then moved down to be steadily located at \sim 3 Å until the end of simulations. Although the large fluctuation of both the C-PNS and P-PNS binding modes were observed in 6-HP β CD, a more than 50% occupation of PNS was found at the position below the center of gravity of 6-HP β CD.

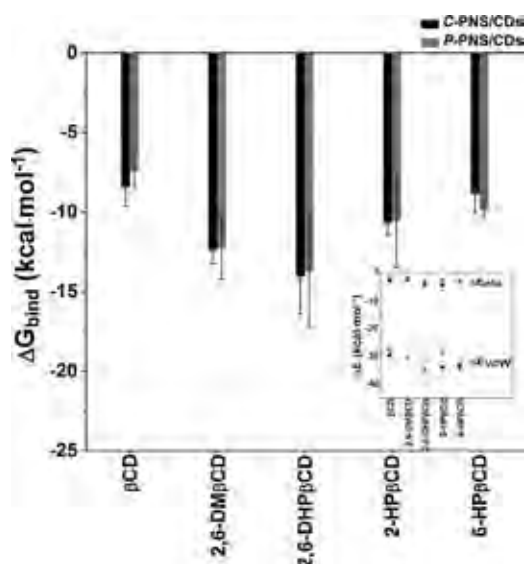
3.4. Water accessibility to PNS in complex with CDs

The radial distribution functions (RDFs) were used to ascertain the water distribution around the four oxygen sites on the chromone ring of the PNS molecule (O1, O4, O5 and O7) (Fig. 1A) in complex with the different CDs. The RDFs of these PNS oxygen atoms together with the integration number, $n(r)$, for the different C-PNS/CD and P-PNS/CD inclusion complexes are plotted in Fig. 5. The $n(r)$ value at the first determined minima was used to estimate the water accessibility toward the focused oxygen, and the results are summarized in Table 1. Among the four sites, no peak was detected within the 3 Å sphere of the O1 site in all systems (except for only a rather small peak found in C-PNS/6-HP β CD). These data suggested that water hardly solvates these oxygen atoms of the

chromone ring reflecting that it mostly stayed inside the center of the hydrophobic cavity of the respective CDs. An unpronounced peak was observed at around 2.7–4.0 Å for the O7 methoxy oxygen atom indicating a weak solvation or an O7-water interaction. In contrast, the two other sites exhibited the first intense peak at \sim 3 Å, corresponding to possible hydration, while the first minima accounted for a number of water molecules maintained on the first hydration shell. Since the first minima never reached a $g(r)$ value of zero, water exchange between the first and secondary hydration shells was feasible. In C-PNS/CDs inclusion complexes, the O4 oxygen was more accessible to water than the O5 and O7 ones, in agreement with the C-PNS displacement nearby the secondary rim where O5 and O7 atoms of the chromone ring were located a little bit deeper in the cavity than O4. Interestingly, lower water accessibility was seen in the C-PNS/2,6-DHP β CD complex due to 2-hydroxypropyl substitutions on both CD rims. For the P-PNS binding mode, the chromone ring of P-PNS stayed close to the wider rim and so a \sim two-fold higher water accessibility was observed (Table 1). Since both C-PNS and P-PNS/CDs complexes with the different β CD could be formed (evidently from the distance plot) with an almost equal binding strength (discussed later in Fig. 6), the total number of accessible water molecules could be consid-

Table 1
Integral number, $n(r)$, up to the first minimum (derived from Fig. 5) for the PNS oxygen atoms in the different PNS/CDs complexes.

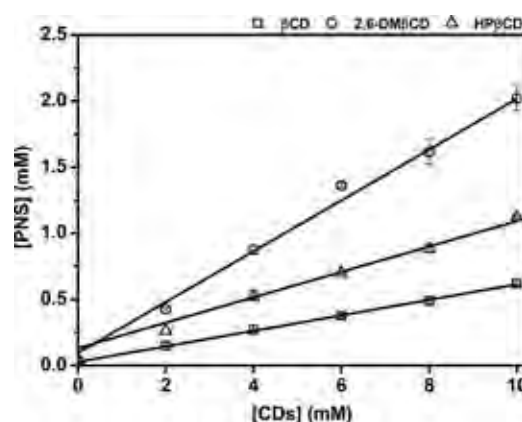
	β CD		2,6-DM β CD		2,6-DHP β CD		2-HP β CD		6-HP β CD	
	C-PNS	P-PNS	C-PNS	P-PNS	C-PNS	P-PNS	C-PNS	P-PNS	C-PNS	P-PNS
O1	–	–	–	–	–	–	–	–	0.003	–
O4	1.5	2.3	3.5	1.4	0.9	1.0	1.4	1.3	0.9	1.5
O5	0.5	2.5	0.8	2.0	–	2.5	–	0.5	0.9	2.7
O7	–	1.3	0.3	2.5	–	1.7	0.8	1.1	0.1	1.7
Total	2.0	6.1	4.6	5.9	0.9	5.2	2.2	2.9	1.9	5.8
Average	4.0		5.2		3.0		2.5		3.9	

**Fig. 6.** MM-GBSA binding free energy (ΔG_{bind}) in kcal/mol of the different C-PNS/ β CD (black) and P-PNS/ β CD (gray) complexes, where the interaction energy components, ΔE_{ele} and ΔE_{vdW} , are given in the inset.

ered for both complexes. Considering the average total number of water molecules solvating the PNS bound to each respective CD, it appears that 2,6-DM β CD introduced one more solvated water, but ~ 1 – 2 less water molecules approached PNS when complexed with 2,6-DHP β CD and 2-HP β CD compared to when complexed with β CD.

3.5. Binding free energy of inclusion complex

To estimate the binding free energy or the absolute free energy of molecules in aqueous solution (ΔG_{bind}) for all PNS/ β CDs inclusion complexes, the MM-GBSA approach [45] was performed on the 30MD snapshots extracted from the last 40 ns. It has to be noticed that four sets of 100, 30, 20 and 10 snapshots were investigated and the 30 snapshots yielded the relative magnitude of mean and standard deviation values of the 100 snapshots. Based on MM-GBSA calculations, ΔG_{bind} is the summation of the electrostatic (ΔE_{ele}) and van der Waals (ΔE_{vdW}) interaction energies in gas phase, GBSA solvation free energy [46], and entropy contribution using normal mode analysis [47]. The ΔG_{bind} , E_{ele} and E_{vdW} values of C-PNS and P-PNS in complex with the different β CDs are presented and compared in Fig. 6. Although the structural and dynamic behaviors, as well as water accessibility, of the two modes of PNS binding in the cavity were somewhat distinguishable, the difference in their binding free energies was less than 2 kcal/mol. Importantly, the averaged ΔG_{bind} values (kcal/mol) suggested that the modification of β CD by both methyl and 2-hydroxypropyl groups could enhance the stability of inclusion complex in order of 2,6-DHP β CD (-13.86) > 2,6-DM β CD (-12.32) > 2-HP β CD (-10.59) > 6HP β CD (-9.34) > β CD (-7.90). As expected for the complex formation of phenolic compounds dip-

**Fig. 7.** Phase solubility diagram for PNS in various concentrations of (\square) β CD, (\circ) 2,6 DM β CD and (Δ) HP β CD at 25 °C.**Table 2**
Slope, correlation coefficient (r^2) and stability constant (K_C) (derived from Fig. 7).

CDs	Slope	r^2	K_C (M^{-1})
β CD	0.058	0.999	1710
2,6-DM β CD	0.200	0.994	6944
HP β CD	0.107	0.996	3328

ping into the hydrophobic inner surface, vdW interaction plays an important role in the complexation process, and they appear to be much stronger than the electrostatic interaction. It turned out to be approximately 6-fold in the case of PNS/CDs complexes.

3.6. Experimental phase solubility study

The phase solubility diagram (Fig. 7) revealed the influence of the different β CDs on the solubility of PNS in water, where the PNS solubility was linearly enhanced with increasing CD concentrations. All the PNS/CD complexes were categorized as A_1 type with the formation of 1:1 molar ratio complexes. The slope and the intrinsic solubility of PNS (S_0) from the diagram were used to calculate the stability constant (K_C) of each complex by Eq. (1), as reported previously [41], with the results summarized in Table 2. As equilibrium constant determined by this method is rather sensitive to the solubility of PNS in pure water, this was measured independently as 3.6×10^{-5} mol/L. In case of low solubility of the compound in pure water the intercept cannot be measured that exactly to ensure an accurate determination of the equilibrium constants. The stability constant of the PNS/2,6-DM β CD complex ($6944 M^{-1}$) was much greater than the complex constants of HP β CD ($3328 M^{-1}$) and β CD ($1710 M^{-1}$), respectively. It is worth noting that the substituted β CD derivatives contained a mixture of various degrees of substitutions as technical products, while the derivative with full substitutions is considered in this MD study. However, the experimental equilibrium constants correlated rather well with the MDs. Thus, this computational tool is able to guide selection of the most

suitable host CD for complexation with the desired flavonoid prior to pharmaceutical application.

4. Conclusions

In the MD simulations the stability of simulated systems of each PNS/ β CD inclusion complex reached equilibrium by 40 ns and was then used to investigate the conformations of β CD and its derivatives, the mobility of the host PNS molecule in the different PNS/ β CD complexes, the water accessibility around PNS and the binding free energies of each inclusion complex. The results revealed that the conformation of β CD and its derivatives complexed with PNS is mostly found in M1 structure of the PES. By means of the PNS mobility, the chromone ring of the PNS molecule preferably stays nearby the secondary rim of β CD and its derivatives, while another orientation of the PNS molecule the phenyl ring shows a higher movement inside the hydrophobic cavity. In addition, the RDF plots show that PNS in the different *P*-PNS/ β CD inclusion complexes interact with \sim two-fold more solvating water molecules than in the different *C*-PNS/ β CD inclusion complexes. It could then be assumed that the *P*-PNS/ β CD inclusion complexes have a better water solubility than the corresponding *C*-PNS/ β CD ones. According to the binding free energy, van der Waals interaction is the most important interaction between the PNS and the β CD or its derivatives. The PNS complex with 2,6-DHP β CD had the highest stability, followed by the inclusion complexes with 2,6-DM β CD, 2-HP β CD, 6-HP β CD and β CD, respectively. Moreover, experimental phase solubility studies supported the computational approach since complexes of PNS/2,6-DM β CD and PNS/HP β CD showed a significant higher stability than those with β CD.

Acknowledgments

This work was financially supported by the National Research Council of Thailand. J. Kicuntod thanks the Scholarship from Graduated School, Chulalongkorn University to commemorate the 72nd anniversary of His Majesty King Bhumibala Aduladeja, and the 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund). We also thank Prof. Dr. Heinz Berner for helpful comments on the manuscript. T. Rungrotmongkol thanks the Thailand Research Fund (MRG5580223 and IRG5780008) and P. Wolschann thanks Chulalongkorn University, The Computer Chemistry Unit Cell, and the Vienna Scientific Cluster (VSC-2) are acknowledged for facility and computing resources.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jmgm.2015.11.005>.

References

- [1] G.R. Beecher, Overview of dietary flavonoids: nomenclature, occurrence and intake, *J. Nutr.* 10 (2003) 3248S–3254S.
- [2] J.V. Formica, W. Regelson, Review of the biology of quercetin and related bioflavonoids, *Food Chem. Toxicol.* 33 (1995) 1061–1080.
- [3] P.C.H. Hollman, M.B. Katan, Absorption, metabolism, and health effects of dietary flavonoids in man, *Biomed. Pharmacother.* 51 (1997) 305–310.
- [4] C.A. Rice-Evans, N.J. Miller, G. Paganga, Structure antioxidant activity relationships of flavonoids and phenolic acids, *Free Radic. Biol. Med.* 20 (1996) 933–956.
- [5] K.A. O'Leary, S. Pascual-Tereasa, Y.P. Needs, N.M. O'Brien, G. Williamson, Effect of flavonoids and vitamin E on cyclooxygenase-2 (COX-2) transcription, *Mutat. Res.* 551 (2004) 245–254.
- [6] J.C. Le Bail, L. Aubourg, G. Habrioux, Effects of pinostrobin on estrogen metabolism and estrogen receptor transactivation, *Cancer Lett.* 156 (2000) 37–44.
- [7] S. Tewtrakul, S. Subhadhirasakul, J. Puripattavong, T. Panphadung, HIV-1 protease inhibitory substances from rhizomes of *Boesenbergia pandurata* Holtz Songklanakarinn, *J. Sci. Technol.* 25 (2003) 503–508.
- [8] Y.F. Xian, S.P. Ip, Z.X. Lin, Q.Q. Mao, Z.R. Su, X.P. Lai, Protective effects of pinostrobin on β -amyloid-induced neurotoxicity in PC12 cell, *Cell. Mol. Neurobiol.* 32 (2012) 1223–1230.
- [9] A. Ribeiro, A. Figueiras, D. Santos, F. Veiga, Preparation and solid-state characterization of inclusion complexes formed between miconazole and methyl- β -cyclodextrin, *AAPS PharmSciTech* 9 (2008) 1102–1109.
- [10] J.L. Atwood, J.E.D. Davies, D.D. MacNicol, F. Vögtle, J.-M. Lehn, Comprehensive supramolecular chemistry, in: J. Szejtli, T. Osa (Eds.), *Cyclodextrins*, third edn, Pergamon, New York, 1996.
- [11] A.L. Laza-Knoerr, R. Gref, P. Couvreur, Cyclodextrins for drug delivery, *J. Drug Target* 18 (2010) 645–656.
- [12] G. Tiwari, R. Tiwari, A. Rai, Cyclodextrins in delivery systems: applications, *J. Pharm. Bioallied. Sci.* 2 (2010) 72–79.
- [13] A.E. Radi, S. Eissa, Electrochemistry of cyclodextrin inclusion complexes of pharmaceutical compounds, *J. Open Chem. Biomed. Methods* 3 (2010) 74–85.
- [14] C. Dass, W. Jessup, Apolipoprotein A-I cyclodextrins and liposomes as potential drugs for the reversal of atherosclerosis, *J. Pharm. Pharmacol.* 52 (2000) 731–761.
- [15] E.M. Martin Del Valle, Cyclodextrins and their uses: a review, *Process Biochem.* 39 (2004) 1033–1046.
- [16] L. Szenté, J. Szejtli, Highly soluble cyclodextrin derivatives: chemistry, properties, and trends in development, *Adv. Drug Deliv. Rev.* 36 (1999) 17–28.
- [17] D.W. Frank, J.E. Gray, R.N. Weaver, Cyclodextrin nephrosis in the rat, *Am. J. Pathol.* 83 (1976) 367–382.
- [18] J. Miranda, T. Martins, F. Veiga, H. Ferraz, Cyclodextrins and ternary complexes: technology to improve solubility of poorly soluble drugs, *Braz. J. Pharm. Sci.* 47 (2011) 665–681.
- [19] T. Loftsson, M. Masson, Cyclodextrins in topical drug formulations: theory and practice, *Int. J. Pharm.* 225 (2001) 15–30.
- [20] C. Rajeswari, A. Alka, A. Javed, R.K. Khar, Cyclodextrins in drug delivery: an updated review, *AAPS PharmSciTech* 6 (2005) E329–E357.
- [21] J. Szejtli, The properties and potential uses of cyclodextrin derivatives, *J. Inclusion Phenom. Mol. Recognit. Chem.* 14 (1992) 25–36.
- [22] R.L. Carrier, L.A. Miller, I. Ahmed, The utility of cyclodextrins for enhancing oral bioavailability, *J. Control. Release* 123 (2007) 78–99.
- [23] S. Tommasini, D. Raneri, R. Ficarra, M.L. Calabrò, R. Stancanelli, P. Ficarra, Improvement in solubility and dissolution rate of flavonoids by complexation with β -cyclodextrin, *J. Pharm. Biomed. Anal.* 35 (2004) 379–387.
- [24] C. Jullian, L. Moyano, C. Yañez, C. Olea-Azar, Complexation of quercetin with three kinds of cyclodextrins: an antioxidant study, *Spectrochim. Acta Part A* 67 (2007) 230–234.
- [25] T.A. Nguyen, B. Liu, J. Zhao, D. Thomas, J. Hook, An investigation into the supramolecular structure, solubility, stability and antioxidant activity of rutin/cyclodextrin inclusion complex, *Food Chem.* 136 (2013) 186–192.
- [26] R. Chadha, S. Gupta, G. Shukla, D.V.S. Jain, R. Pissurlenkar, E. Coutinho, Interaction of artesunate with β -cyclodextrin: characterization, thermodynamic parameters, molecular modeling, effect of PEG on complexation and antimalarial activity, *Results Pharma. Sci.* 1 (2011) 38–48.
- [27] H. Zhang, W. Feng, C. Li, T. Tan, Investigation of the inclusions of puerarin and daidzin with β -cyclodextrin by molecular dynamics simulation, *J. Phys. Chem. B* 114 (2010) 4876–4883.
- [28] Y. Zheng, Albert H.L. Chow, Ian S. Haworth, Molecular modeling of flavonoid- β -cyclodextrin complexes, *Lett. Drug Des. Discov.* 5 (2008) 512–520.
- [29] B. Nutho, W. Khuntawee, C. Rungnim, P. Pongsawasdi, P. Wolschann, A. Karpfen, N. Kungwan, T. Rungrotmongkol, Binding mode and free energy prediction of fisetin/ β -cyclodextrin inclusion complexes Beilstein, *J. Org. Chem.* 10 (2014) 2789–2799.
- [30] W. Sangpheak, W. Khuntawee, P. Wolschann, P. Pongsawasdi, T. Rungrotmongkol, Enhanced stability of a naringenin/2,6-dimethyl β -cyclodextrin inclusion complex: molecular dynamics and free energy calculations based on MM- and QM-PBSA/GBSA, *J. Mol. Graphics Modell.* 50 (2014) 10–15.
- [31] W. Snor, E. Liedl, P. Weiss-Greiler, A. Karpfen, H. Viernstein, P. Wolschann, On the structure of anhydrous β -cyclodextrin, *Chem. Phys. Lett.* 441 (2007) 159–162.
- [32] I. Alecu, J. Zheng, Y. Zhao, D.G. Truhlar, Computational thermochemistry: scale factor database and scale factors of vibrational frequencies obtained from electronic model chemistries, *J. Chem. Theory Comput.* 6 (2010) 2872–2887.
- [33] R.C. Walker, M.F. Crowley, D.A. Case, The Implementation of a fast and accurate QM/MM potential method in Amber, *J. Comput. Chem.* 29 (2008) 1019–1031.
- [34] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, C.R. Gonzalez-Outeirino, B.L. Foley, R.J. Woods, GLYCAM06: a generalizable biomolecular force field carbohydrates, *J. Comput. Chem.* 29 (2008) 622–655.
- [35] W. Khuntawee, T. Rungrotmongkol, S. Hannongbua, Molecular dynamic behavior and binding affinity of flavonoid analogues to the cyclin dependent kinase 6/cyclin D complex, *J. Chem. Inf. Mol.* 52 (2012) 76–83.
- [36] N. Kaiyawet, T. Rungrotmongkol, S. Hannongbua, Effect of halogen substitutions on dUMP to stability of thymidylate synthase/dUMP/mTHF ternary complex using molecular dynamics simulation, *J. Chem. Inf. Model.* 6 (2013) 1315–1323.

- [37] A. Meeprasert, W. Khuntawee, K. Kamlungsua, N. Nunthaboot, T. Rungrotmongkol, S. Hannongbua, Binding pattern of the long acting neuraminidase inhibitor laninamivi towards influenza A subtypes H5N1 and pandemic H1N1, *J. Mol. Graphics Modell.* 38 (2012) 148–154.
- [38] A. Meeprasert, S. Hannongbua, T. Rungrotmongkol, Binding and susceptibility of NS3/4A serine protease inhibitors against hepatitis C virus, *J. Chem. Inf. Model.* 54 (2014) 1208–1217.
- [39] A. Meeprasert, T. Rungrotmongkol, M.S. Li, S. Hannongbua, In silico screening for potent inhibitors against the NS3/4A protease of hepatitis C virus, *Curr. Pharm. Des.* 20 (2014) 3465–3477.
- [40] B.A. Luty, W.F. Gunsteren, Calculating electrostatic interactions using the particle–particle particle-mesh method with nonperiodic long-range interactions, *J. Phys. Chem.* 100 (1996) 2581–2587.
- [41] T. Higuchi, K.A. Connors, Phase solubility techniques, *Adv. Anal. Chem. Instrum.* 4 (1965) 117–122.
- [42] C.W. Yong, C. Washington, W. Smith, Structural behaviour of 2-hydroxypropyl- β -cyclodextrin in water: molecular dynamics simulation studies, *Pharm. Res.* 25 (2007) 1092–1099.
- [43] T. Kozar, C. Venanzi, Reconsidering the conformational flexibility of β -cyclodextrin, *J. Mol. Struct. THEOCHEM* 395–396 (1997) 451–468.
- [44] L.J. Yang, S.X. Ma, S.Y. Zhou, W. Chen, M.W. Yuan, Y.Q. Yin, X.D. Yang, Preparation and characterization of inclusion complexes of naringenin with β -cyclodextrin or its derivative, *Carbohydr. Polym.* 98 (2013) 861–869.
- [45] P. Greenidge, C. Kramer, J.C. Mozziconacci, R. Wolf, MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement, *J. Chem. Inf. Model.* 53 (2013) 201–209.
- [46] V. Zoete, M. Meuwly, M. Karplus, Study of insulin dimerization: binding free energy calculations and per-residue free energy decomposition, *Proteins: Struct., Funct. Bioinf.* 61 (2005) 79–93.
- [47] S. Genheden, O. Kuhn, P. Mikulskis, D. Hoffmann, U. Ryde, The normal-mode entropy in the MM/GBSA method: effect of system truncation, buffer region, and dielectric constant, *J. Chem. Inf. Model.* 52 (2012) 2079–2088.

Ring conformation analytical methods of *iota*-cyclodextrin from replica exchange molecular dynamics simulation

Wasinee Khuntawee^{1†}, *Manaschai Kunaseth*^{2*}, *Chompoonut Rungnim*², *Suradej Intagorn*³, *Peter Wolschann*^{4,5,6}, *Nawee Kungwan*⁷, *Thanyada Rungrotmongkol*^{8,9*}, *Supot Hannongbua*⁶

¹Nanoscience and Technology Program, Graduate School, Chulalongkorn University, 254 Phayathai Road, Bangkok 10330, Thailand

²National Nanotechnology Center (NANOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand

³Department of Mathematics, Statistics and Computer Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand

⁴Department of Pharmaceutical Technology and Biopharmaceutics, University of Vienna, Vienna 1090, Austria

⁵Institute of Theoretical Chemistry, University of Vienna, Vienna 1090, Austria

⁶Computational Chemistry Unit Cell, Department of Chemistry, Faculty of Science, Chulalongkorn University, 254 Phayathai Road, Bangkok 10330, Thailand

⁷Department of Chemistry, Faculty of Science, Chiang Mai University, 239 Huay Kaew Road,
Muang District, Chiang Mai 50200, Thailand

⁸Structural and Computational Biology Unit, Department of Biochemistry, Faculty of Science,
Chulalongkorn University, 254 Phayathai Road, Bangkok 10330, Thailand

⁹Ph.D. Program in Bioinformatics and Computational Biology, Faculty of Science,
Chulalongkorn University, 254 Phayathai Road, Patumwan, Bangkok 10330, Thailand

KEYWORDS. Conformational diversity, *iota*-cyclodextrin, replica-exchange molecular dynamics, Gaussian mixture model.

ABSTRACT. Large ring cyclodextrins have become increasingly important for drug delivery applications. In this work, we have performed replica-exchange molecular dynamics simulation to study the conformational diversity of *iota*-cyclodextrin containing 14 α -1,4 glycosidic linked D-glucopyranose units (CD14). The new calculation methods are proposed to analyze the openness and bending conformation of CD14 in terms of circularity and biplanar angle. As a result, CD14 is mostly opened with average circularity of 0.39 ± 0.16 and is slightly bent with average biplanar angle of $145.5 \pm 16.0^\circ$. In addition, classification of CD14 conformations using Gaussian mixture model (GMM) shows that 85.0% of all CD14 trajectories at 300 K are corresponding to the opened conformation. The GMM clustering also reveals the minority conformations of CD14 such as 8-shape and U-shape conformations. The obtained results indicate that CD14 is a suitable candidate for inclusion of drug molecule. This work provides fundamental insights of CD14 conformation and also proposing the new techniques for molecular conformation study in the future.

1. Introduction

Large ring cyclodextrins (LR-CDs) are cyclic oligosaccharide of 1-4 linked alpha-D-glucopyranoses consisting of more than eight glucose subunits. Currently, the several size LR-CDs (up to hundreds subunits) have achieved by enzymatic syntheses of amyloamylase, another 4 α -glucanotransferase, which their unique loop structure has not existed in small ring cyclodextrins (SR-CDs) produced by cyclodextrin glycosyltransferase (CGTase). However, only four crystal structures of δ -CD (CD9),¹ ϵ -CD (CD10),²⁻⁴ ι -CD (CD14),³⁻⁵ and ϕ -CD (CD26)^{6, 7} were reported. Based on crystallographic studies, these LR-CDs can form into various shapes such as the bent shape of CD10 and CD14 and the twisted shape of CD26, whilst SR-CDs as CD6, CD7 and CD8 are found only in a cyclic shape.⁸⁻¹⁰ Mixed sizes of CDs can be synthesized with the polymerization of CGTase. The synthesis conditions *i.e.* temperature and concentration of ethanol in an aqueous solution resulting in the increasing chain-length of LR-CDs.^{11, 12} To investigate the structural properties of individual LR-CD, single-size LR-CD purification and their crystal structures are required via a costly process and sophisticated laboratory analysis. Over the last decade, classical molecular dynamics (MD) method has been applied to study the LR-CDs conformation both in gas and solution phases.¹³⁻¹⁵ The structure stability of LR-CDs has been discussed in terms of their shape, hydrogen bond formation, as well as their energy. The larger ring of LR-CDs has more flexible structure, where multiple cavities of LR-CDs could be frequently observed. Using classical MD simulation of CD14 in solution at room temperature, the deformed conformation from the crystal structure *e.g.* open form, twisted shape containing two loops and dumbbell shape were found.¹⁶ Moreover, the CD26 geometry studied by DFT method in gas phase is in agreement with its crystal structure solvated by waters.¹⁷

Although classical MD is a useful method to study conformation of LR-CDs, it requires long-time dynamics in order to overcome the energy barrier among local minima. In the context of conformational exploration, replica exchange molecular dynamics (REMD) simulation is a potentially powerful technique.¹⁸⁻²¹ REMD method can be used to efficiently solve multiple minima problem using independent MD simulations (*i.e.* replica) with different temperatures and temperature exchange among replicas. Nevertheless, a significant challenge from REMD for conformational searching is the vast amount of conformation data generated from the multiple REMD trajectories. One way to deal with this issue is to visualize each MD trajectory and manually choose “important” snapshots based on visual data and researcher’s experience. This method might be possible and effective for small trajectory, however capturing all of the key snapshots from large trajectory is highly challenging and may introduce sampling bias. In addition, trajectories generated from REMD are subjected to spontaneous and erratic motion due to the temperature/trajectory exchanges between different replicas, which might hamper visual information of CD14 conformation. Therefore, designing systematic routine to quantify structural properties of CD14 is crucial for analyzing large amount of data from REMD simulation.

Recently, machine learning and data analytic techniques have become immensely useful for computational studies.²²⁻²⁴ Unsupervised learning techniques such as clustering can be applied on large amount of data to discover hidden classes of data.^{25, 26} Clustering algorithm based on principal component analysis (PCA) technique is commonly used to reduce high dimensional data into lower dimension but highly significant data, while methods such as *K*-mean clustering, DBSCAN, and Gaussian mixture models (GMM) are more suitable for low-dimensional data.

GMM clustering is a method of interest due to its fast, suitability for density estimation, and non-bias toward zero means and cluster sizes.

In this paper, REMD simulation has been performed to explore the conformational space of CD14 structure. To efficiently analyze CD14 conformation from REMD results, we have combined two approaches: (1) development of new LR-CD structural characterizations; and (2) application of machine learning based clustering algorithm for classification of CD14 structures. In the first approach, we have developed two new LR-CD characterizations including biplanar angle and circularity to quantitatively analyze CD14 conformation from REMD simulation. Here, in-depth computational steps, detailed analysis, and performance evaluation of the new techniques are thoroughly discussed. In the second approach, GMM clustering algorithm is applied to systematically identify the groups of CD14 based on their conformations. Finally, the new characterization techniques are applied on GMM clustering results, revealing the underlying groups of CD14 conformations. Our new bifurcated approach has provided a systematic way to analyze complex conformation of LR-CDs, which is beneficial for broad-range of research areas.

2. Methodology

2.1 Replica exchange molecular dynamics (REMD)

All system preparations and REMD simulations are performed using Amber10.²⁷ The starting structure of CD14 is obtained from the Cambridge Crystallographic Data Centre entry code CCDC124917.²⁸ To remove steric interactions, the initial structure of CD14 is minimized by 2000 steps of steepest descent method, continued by 1000 steps of conjugated gradient method. The glycam06 force field and the GB implicit solvent model, Igb5, are used where more details of the force field testing as well as the solvation model effect have thoroughly discussed

in the previous study.²⁹ Based on REMD simulation, CD14 is performed for 16 replicas under temperature exchanges within the range 300-600 K with 20 K interval. By this REMD condition, the acceptance ratio is more than 40%. These REMD simulation condition had been validated in our previous study.²⁹ A short MD simulation of 5 ns is performed prior to the REMD simulation in order to equilibrate the systems at the assigned temperatures. Afterwards, the REMD simulation is completed for 100 ns using Igb5 solvent model. CD14 snapshot is taken every 1 ps from REMD simulation at 300 K for conformation analysis, resulting in 100,000 snapshots of CD14 in total.

2.2 Structural analysis

We proposed two molecular-structure analytical methods, which are biplanar angle and circularity. The analysis codes were implemented using Python 2.7 incorporated with scikit-learn machine-learning toolkit 0.15.0,³⁰ scipy version 0.14.0 and numpy 1.8.1 for numerical and mathematical supporting subroutines.

2.2.1. Biplanar angle

Previous study reported that CD14 conformations were found bent³⁻⁵ resulting in a folded biplanar configuration. In this type of structure, the degree of bending between the two internal planes significantly affects the inclusibility of drug molecule. Here, biplanar angle, θ_p , is proposed to characterize the bending of internal molecular planes within LR-CD structures, see Figure 1. This calculation method can be summarized in three steps:

- (I) Clustering of atoms in CD14 into two groups (\mathbf{X}_1 and \mathbf{X}_2) based on their coordinates.
- (II) Find the internal planes \mathbf{P}_1 and \mathbf{P}_2 , which are the most spatially correlated 2-dimensional plane of \mathbf{X}_1 and \mathbf{X}_2 embedded in 3-dimensional space.

(III) Calculate θ_P from the angle between normal vector of planes P_1 and P_2 .

In step (I), atoms of CD14 conformation are clustered into two groups using k -mean clustering algorithm.²⁶ Let \mathbf{r}_i be a spatial coordinate of i -th atom in CD14. k -mean clustering algorithm decides whether \mathbf{r}_i is belong to \mathbf{X}_1 or \mathbf{X}_2 by minimizing the following objective function Z :

$$Z^* = \operatorname{argmin}_{\mathbf{X}_1, \mathbf{X}_2} \left\{ \sum_{\mathbf{r}_i \in \mathbf{X}_1} \operatorname{dist}(\mathbf{c}_1, \mathbf{r}_i) + \sum_{\mathbf{r}_i \in \mathbf{X}_2} \operatorname{dist}(\mathbf{c}_2, \mathbf{r}_i) \right\} \dots\dots\dots(1),$$

where $\operatorname{dist}(\mathbf{a}, \mathbf{b})$ is the Euclidian distance between two points, \mathbf{c}_1 and \mathbf{c}_2 denoted the centroids of atoms in \mathbf{X}_1 and \mathbf{X}_2 calculated from: $\mathbf{c}_i = \frac{1}{|\mathbf{X}_i|} \sum_{\mathbf{r}_i \in \mathbf{X}_i} \mathbf{r}_i$. Here, mini-batch k -mean clustering algorithm is used due to its highly efficient computation time.³⁰ The parameters of mini-batch k -mean procedure are as follows: number of clusters (k) = 2, number of mini batches = 100, number of randomized initialization trials = 10, maximum clustering iterations = 300, and stopping criteria of smoothed variance-normalized of the mean-center squared position changes = 10^{-4} . After the mini-batch k -mean process, two sets of atomic coordinates \mathbf{X}_1 and \mathbf{X}_2 are obtained.

In step (II), the most spatially correlated 2-dimensional plane P_i for atoms in \mathbf{X}_i is computed. Let $P_i = (\hat{\mathbf{p}}_i^1, \hat{\mathbf{p}}_i^2)$, where $\hat{\mathbf{p}}_i^1 \perp \hat{\mathbf{p}}_i^2$ are orthonormal basis vectors spanning plane P_i . Here, $\hat{\mathbf{p}}_i^1, \hat{\mathbf{p}}_i^2$ are obtained from PCA as the two-most correlated principle components of atomic coordinate of all atoms in \mathbf{X}_i . In order to find P_i , we performed a principle-component analysis (PCA) on \mathbf{X}_i . The two-most correlated orthonormal basis from PCA (*i.e.* the principle components 1 and 2) spanning P_i . Note that the third orthonormal basis is the normal vector ($\hat{\mathbf{n}}_i$) of plane P_i .

Finally, step (III) is achieved by calculating angle between planes \mathbf{P}_1 and \mathbf{P}_2 based on the cosine of their normal vectors:

$$\theta_p = \cos^{-1}(\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2) \dots \dots \dots (2).$$

For $\theta_p < 180^\circ$, two cosine angles exist, which are θ_p and $\pi - \theta_p$. Here, we assume that $\theta_p > 90^\circ$ because the steric repulsion would prevent \mathbf{P}_1 and \mathbf{P}_2 becoming too close to each other. Thus, we considered the biplanar angle as $\theta_p = \max(\theta_p, \pi - \theta_p)$ hereinafter.

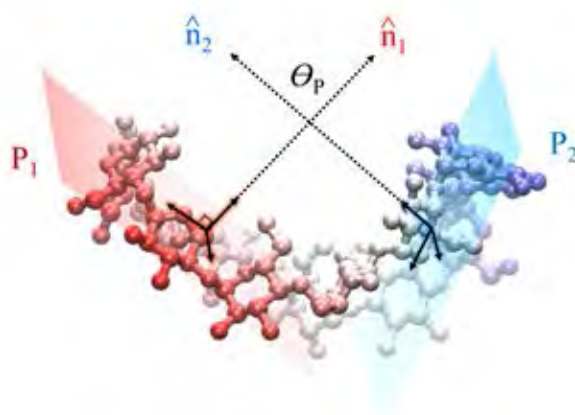


Figure 1. Schematic diagram illustrates the internal plane \mathbf{P}_i for each cluster of atoms. Biplanar angle is noted as an angle between two normal vectors \mathbf{n}_i .

2.2.2. Circularity

Extensive length of CD14 oligosaccharides makes its configuration possibly twisted and deformed. In order to quantify the twistiness degree of CD14 configurations, circularity (γ_c) is defined as a ratio between major and minor diameters in a deformed elliptical configuration of CD14 (see Figure 2). Circularity is computed in the following steps. First, diameters of CD14 are calculated as a distance between two oxygen atoms at glycosidic bond (O4) on the opposite side of the CD14 ring (*i.e.* for 14 oligosaccharides, O4 atom of i -th saccharide molecule and the O4

atom of the $(i+7)$ -th saccharide molecule). Circularity of CD14 is defined as the maximum ratio between the shortest and the longest of CD14 diameter:

$$\gamma_c = \max \left(\frac{d_i}{d_{\left(i+\frac{L}{2}\right) \bmod L}} \right), \quad \dots\dots\dots(3),$$

where L denoted the number of sugar molecules presented in the LR-CDs ring (*i.e.* $L = 14$ for CD14), d_i is the diameter between opposite O4 atoms, and mod denoted the modulo operation. Note that the value of γ_c is the range of $0 < \gamma_c \leq 1$, while $\gamma_c = 1$ is perfect circular shape of LR-CD. Circularity decreases as the structure deforms into elliptical shape.

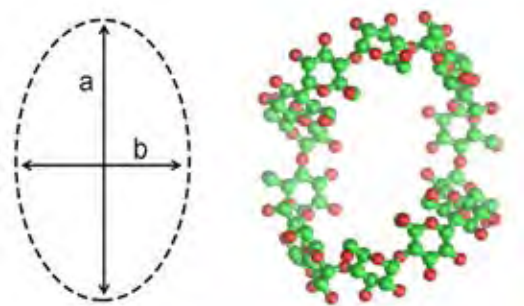


Figure 2. Schematic diagram depicting the ratio between the max and min diameters of ellipse on a 2D plane.

2.3) Clustering algorithm

Clustering algorithm is an unsupervised learning algorithm to classify dataset into multiple classes based on the input properties. Gaussian mixture model (GMM) clustering is an N -class clustering algorithm based on a set of GMMs $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$, where each GMM is a D -dimensional weighted sum of M -component normal densities. The mixture density of n -th GMM is defined as

$$p(\bar{x} | \lambda_n) = \sum_{i=1}^M w_i^n p_i^n(\bar{x}) \dots\dots\dots(4),$$

where M is the number of GMM components, w_i^n denoted the mixture weights of i -th uni-model normal density function $p_i^n(\bar{x})$ for n -th GMM λ_n , where $p_i^n(\bar{x})$ is defined as a D -dimensional Gaussian function centralized at D mean vector $\bar{\mu}_i$ with $D \times D$ covariance matrix

$$p_i^n(\bar{x}) = \frac{1}{\sqrt{2\pi^D |C_i^n|}} \exp\left\{-\frac{1}{2}(\bar{x} - \bar{\mu}_i^n)^T (C_i^n)^{-1} (\bar{x} - \bar{\mu}_i^n)\right\} \dots\dots\dots(5).$$

Thus, a completed description of each GMM is defined as the following

$$\lambda_n = \{w_i^n, \bar{\mu}_i^n, C_i^n | i = 1, \dots, M\} \dots\dots\dots(6).$$

Note that GMM clustering utilizes expectation–maximization algorithm³¹ to determine the most suitable N classes for the dataset.

3. Results and discussion

3.1 CD14 conformational analysis

3.1.1 Biplanar angle

In this section, we analyze CD14 structures in terms of biplanar angle (θ_p) performed on each of the 100,000 CD14 trajectories obtained from REMD at 300 K. Figure 3 shows the superimposition of sample CD14 snapshots in each θ_p range. Here, Figures 3(a) – 3(d) show that CD14 with $\theta_p = 90$ - 130° are clearly bent forming V-shape geometry, as well as the curvature of V-shape is decreasing from Figures 3(a) to 3(d) corresponding to their increasing θ_p . For θ_p in the range of 130 - 150° , CD14 structures are more opened in a wider V-shape or almost U-shape in some cases, signaling the transition from close to open biplanar angle conformations. Finally

for CD14 with high $\theta_p = 150\text{-}180^\circ$, the conformations are mostly flat with a slight wavy edge from the side view.

In terms of population, Figure 4 shows the distribution of CD14 conformations based on their θ_p . The distribution of θ_p is a slightly left-skewed normal distribution with mean $\theta_p = 145.5^\circ$ and $SD = 16.0^\circ$. The most dense CD14 population is in the range of $\theta_p = 140\text{-}150^\circ$, accounting for 24.5% of all CD14 structures. Fewer amount of CD14 is found at low θ_p such that only 7.1% has $\theta_p < 120^\circ$, while 18.6% of overall CD14 population has $\theta_p > 160^\circ$.










Conformation	Sample snapshots			
V-shape	(a)  $90^\circ < \theta_p < 100^\circ$	(b)  $100^\circ < \theta_p < 110^\circ$	(c)  $110^\circ < \theta_p < 120^\circ$	(d)  $120^\circ < \theta_p < 130^\circ$
Wide V-shape	(e)  $130^\circ < \theta_p < 140^\circ$	(f)  $140^\circ < \theta_p < 150^\circ$		
Wavy flat	(g)  $150^\circ < \theta_p < 160^\circ$	(h)  $160^\circ < \theta_p < 170^\circ$	(i)  $170^\circ < \theta_p < 180^\circ$	

Figure 3. Superimposed CD14 structures with various biplanar angles (θ_p), extracted from REMD simulation at 300 K.

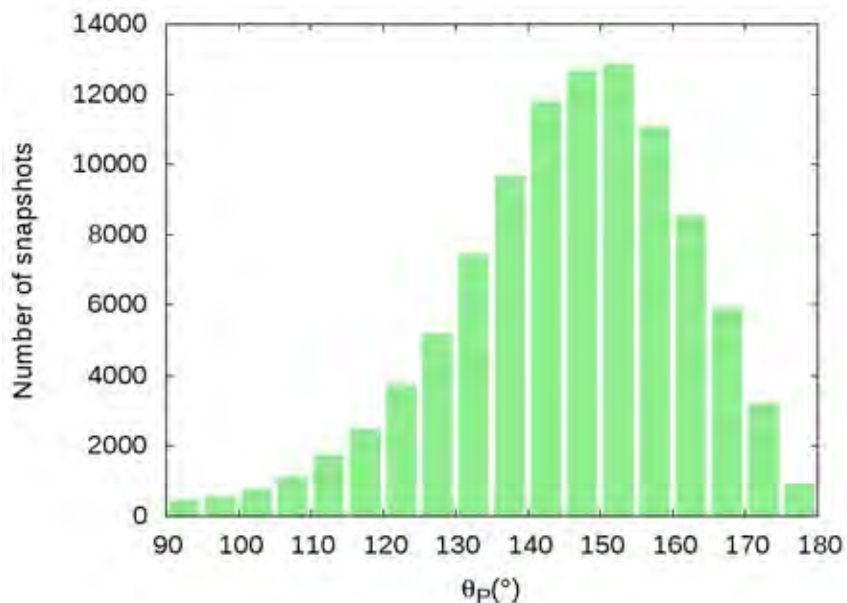


Figure 4. Distribution of CD14 conformations in terms of θ_p .

3.1.2 CD14 circularity (γ_c^2)

In this section, CD14 conformational diversity is analyzed in terms of γ_c^2 . Figure 5 shows superimposition of CD14 snapshots based on their measured γ_c^2 . Here, we found 3 major groups of conformation based on circularity, which are: (1) 8-shape; (2) elliptical; and (3) circular. For $\gamma_c^2 < 0.2$ (see Figures 5(a)-(b)), CD14 is formed into an 8-shaped conformation, in which two sides of the ring are interacted with each other near the center of the ring. Also, two smaller cavities are presented instead of one large cavity in this case. In addition, both cavities aligned in the same and perpendicular planes to each other are observed. For $0.2 < \gamma_c^2 < 0.7$ (Figures 5(c)-

(g)), CD14 are formed in elliptic shape with more band flips observed in $\gamma_c^2 = 0.2 - 0.4$ range than in $\gamma_c^2 = 0.4 - 0.7$ range. Finally, CD14 conformation is almost circular when $\gamma_c^2 > 0.7$ (see Figures 5(h)-(j)).











Conformation	Sample snapshots
8-shape	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>(a)</p>  <p>$0.0 < \gamma_c^2 < 0.1$</p> </div> <div style="text-align: center;"> <p>(b)</p>  <p>$0.1 < \gamma_c^2 < 0.2$</p> </div> </div>
Elliptical	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>(c)</p>  <p>$0.2 < \gamma_c^2 < 0.3$</p> </div> <div style="text-align: center;"> <p>(d)</p>  <p>$0.3 < \gamma_c^2 < 0.4$</p> </div> <div style="text-align: center;"> <p>(e)</p>  <p>$0.4 < \gamma_c^2 < 0.5$</p> </div> <div style="text-align: center;"> <p>(f)</p>  <p>$0.5 < \gamma_c^2 < 0.6$</p> </div> </div> <div style="display: flex; justify-content: center; align-items: center; margin-top: 10px;"> <div style="text-align: center;"> <p>(g)</p>  <p>$0.6 < \gamma_c^2 < 0.7$</p> </div> </div>
Circular	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>(h)</p>  <p>$0.7 < \gamma_c^2 < 0.8$</p> </div> <div style="text-align: center;"> <p>(i)</p>  <p>$0.8 < \gamma_c^2 < 0.9$</p> </div> <div style="text-align: center;"> <p>(j)</p>  <p>$0.9 < \gamma_c^2 < 1.0$</p> </div> </div>

Figure 5. Superimposition of CD14 according to their circularity (γ_c^2) in the range of 0 to 1 with 0.1 interval.

Next, we analyze the population CD14 in terms of γ_c^2 (see Figure 6). The population of CD14 for γ_c^2 is right-skewed normal distribution with mean = 0.39. The most dense CD14 population in the range of $\gamma_c^2 = 0.3 - 0.4$ are elliptical shape (Figure 5(d)), which is accounted for 25.8% of overall CD14 snapshots.

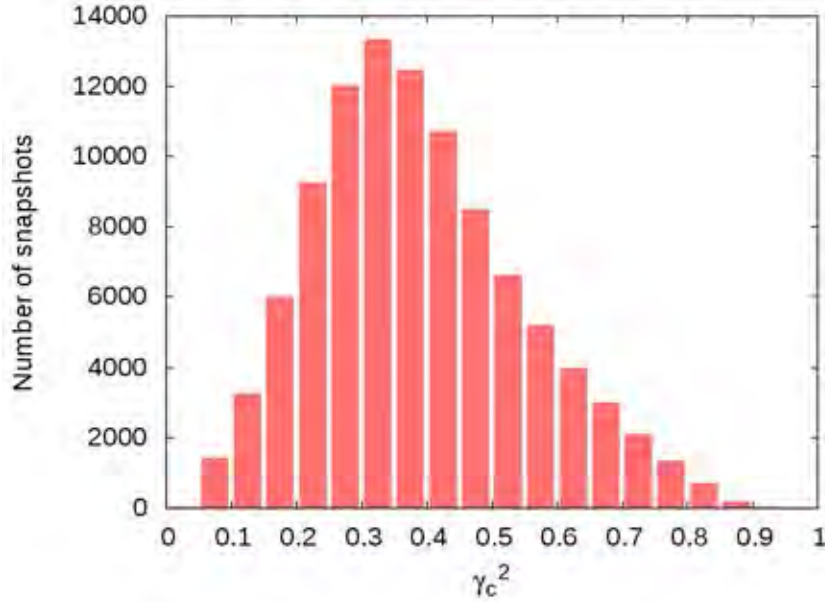


Figure 6. Distribution of CD14 conformation in terms of γ_c^2 .

3.2 Gaussian mixture models clustering

In this section, we applied GMM clustering method to classify CD14 conformation from a set of 21-dimension feature vector $\mathbf{Y} = [d_1, \dots, d_7, \alpha_1, \dots, \alpha_{14}]$ of each CD14 trajectory. Here, the first seven elements, d_i , are the diameters. Since CD14 is cyclic, rotational-invariant ordering of diameter should be considered for every snapshot. This is addressed by defining d_1 as the longest diameter and the order from d_2 to d_7 is cycled either clockwise or counter-clockwise direction

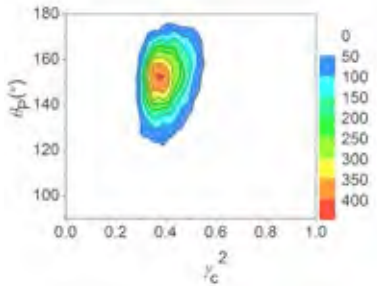


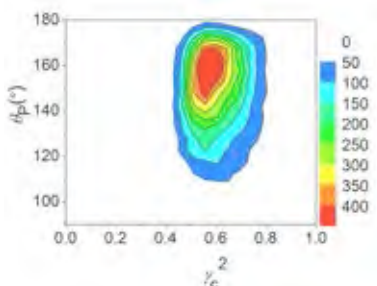


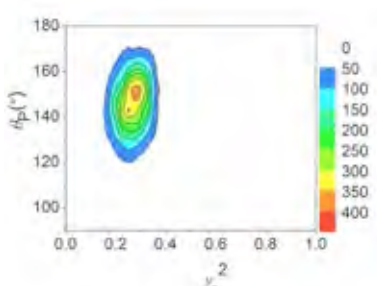


from d_1 through the shortest-diameter-last direction (*i.e.* the shortest diameter is d_4). To use simple attribute to characterize open angle of CD14, the last fourteen elements in \mathbf{Y} are the O4-centroid-O4 angles (α_j) starting from the O4 oxygen of the longest diameter and rotate clockwise/counter-clockwise via the shortest-diameter-last algorithm.

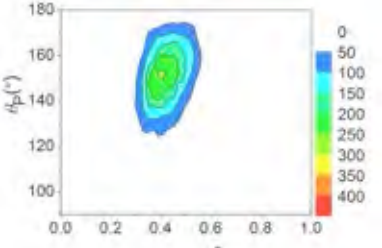


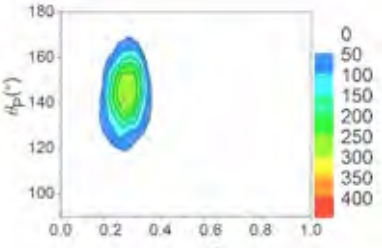


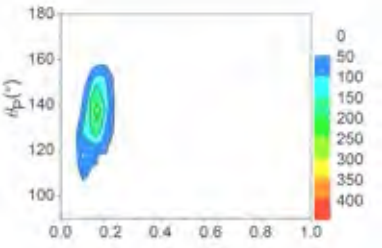


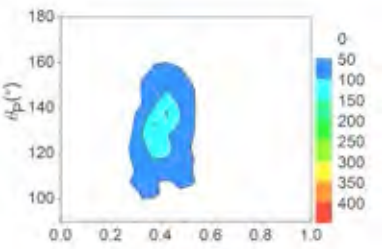


One of the key challenges of clustering process is the unknown hidden number of classes, N . To address this issue, GMM clustering is performed based on the N ranging from 2 to 7 classes. As a result, we found that for $N = 2$, the number of classes tends to be too small since the variations of γ_c^2 are too large. When $N = 3$ to 6, the variations of γ_c^2 become relatively smaller than that for $N = 2$. However, some of the resulting classes are indistinguishable suggesting that there might be an overfitting for larger N . Although the overfitting is presented for larger N , we found that some of the less populated and erratic conformations of CD14 could be captured using the larger N . As a result, we found that $N = 7$ is a good compromising value of N , since it affects only slightly by overfitting but it is able to include atypical conformations in the result. To keep the context concise, we focus the discussion on the clustering result of $N = 7$, while the for clustering results of $N = 2$ to 6 are provided in table S1 in the supplementary.

Let C_k^N denoted a resulting class from GMM clustering, where N is the number of clusters specified in the GMM and k denoted the k -th class in descending order of population size. Considering the clustering result for $N = 7$ in Table 1, we found that CD14 conformations can be classified into 3 categories: (1) circular-opened conformation; (2) elliptical-opened conformation; and (3) closed conformation. First, 19.01% of CD14 are in C_2^7 class, presenting a wide circular conformation with average $\gamma_c^2 = 0.63$. Next, C_1^7 and C_4^7 classes are opened conformations but slightly less circular compared to C_2^7 . These two classes are almost identical

in terms of γ_c^2 and θ_p as the highly resemblance snapshots of both classes also confirmed the similarity (see Table 1). A combined population of C_1^7 and C_4^7 is accounted for 36.3% of all CD14, the largest population group. For the elliptical-opened conformation, C_3^7 and C_5^7 are almost similar in terms of γ_c^2 in the range of 0.26 – 0.27. The snapshots suggest that conformation in this category appears to be a narrower ring, forming an elliptical shape from the top view in accordance to their relatively low γ_c^2 . Due to the lesser cavity presented in this conformation, elliptical-opened conformation might be suitable for smaller drug molecule inclusion. On the contrary to the former classes, 8.54% of all CD14 are in 8-shape conformation of C_6^7 , which are clearly closed with average $\gamma_c^2 = 0.15$. Finally, although the conformation of C_7^7 tends to be opened based on its γ_c^2 alone, the narrowest average $\theta_p = 127.6^\circ$ of C_7^7 suggests that its conformation is formed in a U-shape. The steric hindrance of U-shape CD14 is likely prevented the inclusion of drug molecule in this conformation. In summary, the populations of CD14 conformations are (1) 55.3% circular-opened conformation; (2) 29.7% elliptical-opened conformation; and (3) 15.0% closed conformation. This indicates that CD14 is a suitable candidate for inclusion of drug molecule.

Table 1. The population distribution, sample snapshots, θ_p , and γ_c^2 of each class for GMM clustering result with $N = 7$. Note that each blue band in the top and side views represent one CD14 snapshot. The numerals after \pm denote the standard deviation.

Clustering group	Distribution (%)	Top view	Side view
C_1^7 (circular)	 <p>20.99%</p>	 <p>$\gamma_c^2 = 0.40 \pm 0.07$</p>	 <p>$\theta_p = 150.8 \pm 12.8^\circ$</p>
C_2^7 (circular)	 <p>19.01%</p>	 <p>$\gamma_c^2 = 0.63 \pm 0.09$</p>	 <p>$\theta_p = 148.2 \pm 17.5^\circ$</p>
C_3^7 (elliptical)	 <p>16.11%</p>	 <p>$\gamma_c^2 = 0.27 \pm 0.05$</p>	 <p>$\theta_p = 146.3 \pm 12.1^\circ$</p>

<p>C_4^7 (circular)</p>	 <p>15.29%</p>	 <p>$\gamma_c^2 = 0.42 \pm 0.07$</p>	 <p>$\theta_p = 150.0 \pm 13.5^\circ$</p>
<p>C_5^7 (elliptical)</p>	 <p>13.60%</p>	 <p>$\gamma_c^2 = 0.26 \pm 0.05$</p>	 <p>$\theta_p = 143.5 \pm 12.7^\circ$</p>
<p>C_6^7 (closed)</p>	 <p>8.54%</p>	 <p>$\gamma_c^2 = 0.15 \pm 0.05$</p>	 <p>$\theta_p = 133.5 \pm 16.3^\circ$</p>
<p>C_7^7 (closed)</p>	 <p>6.47%</p>	 <p>$\gamma_c^2 = 0.43 \pm 0.10$</p>	 <p>$\theta_p = 127.6 \pm 19.2^\circ$</p>

4. Conclusion

In this work, we have performed REMD simulation to study the conformational diversity of CD14. To analyze a large amount of trajectory data from REMD simulation, the newly proposed CD14 conformation analysis methods were successfully quantified the openness and the bending degree of CD14 conformations via circularity and biplanar angle. Our analysis indicated that CD14 is mostly opened with average circularity, $\gamma_c^2 = 0.39 \pm 0.16$. In term of bending degree, biplanar angle analysis suggested that conformation of CD14 is dominantly flat with $\theta_p = 145.5 \pm 16.0^\circ$. In addition, classification of CD14 conformation using GMM algorithm revealed that 85.0% of all CD14 trajectory are corresponding to the open conformation, while the minority closed conformations such as 8-shape and U-shape conformations were also identified. To the best of our knowledge, this is the first attempt to systematically classify conformation of cyclodextrin using GMM algorithm. This work provides fundamental insights of CD14 conformation and also proposing the new techniques for molecular conformation study in the future.

ASSOCIATED CONTENT

Supporting Information. The population distribution, θ_p , and γ_c^2 of each class for GMM clustering result with $N = 2 - 6$.

AUTHOR INFORMATION

Corresponding Author

*manaschai@nanotec.or.th, *t.rungrotmongkol@gmail.com

Present Addresses

†Computational Biomodelling Laboratory for Agricultural Science and Technology (CBLAST),
Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

This work was financially supported by the National Research Council of Thailand. T.R. thanks the Thailand Research Fund (RSA5980069 and IRG5780008) and the Structural and Computational Biology Research Group, Special Task Force for Activating Research (STAR), Faculty of Science, Chulalongkorn University. The Computer Chemistry Unit Cell, the Vienna Scientific Cluster (VSC-2), and National e-Science Consortium are acknowledged for facility and computing resources.

REFERENCES

1. Fujiwara, T.; Tanaka, N.; Kobayashi, S., Structure of δ -cyclodextrin•13.75 H₂O. *Chemistry Letters* **1990**, 19, 739-742.
2. Ueda, H.; Endo, T.; Nagase, H.; Kobayashi, S.; Nagai, T., Isolation, purification, and characterization of cyclomaltoheptaose (ϵ -CD). *J. Includ. Phenom. Mol.* **1996**, 25, 17-20.

3. Jacob, J.; Geßler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W., Strain-Induced “Band Flips” in Cyclodecaamylose and Higher Homologues. *Angew. Chem. Int. Edit* **1998**, *37*, 605-609.
4. Jacob, J.; Geßler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W., Band-flip and kink as novel structural motifs in α -(1-4)-D-glucose oligosaccharides. Crystal structures of cyclodeca- and cyclotetradecaamylose. *Carbohydr. Res.* **1999**, *322*, 228-246.
5. Harata, K.; Endo, T.; Ueda, H.; Nagai, T., X-Ray Structure of α -Cyclodextrin. *Supramol. Chem.* **1998**, *9*, 143-150.
6. Gessler, K.; Usón, I.; Takaha, T.; Krauss, N.; Smith, S. M.; Okada, S.; Sheldrick, G. M.; Saenger, W., V-Amylose at atomic resolution: X-ray structure of a cycloamylose with 26 glucose residues (cyclomaltohexaicosaoase). *P. Natl. Acad. Sci. USA.* **1999**, *96*, 4246-4251.
7. Nimz, O.; Geßler, K.; Usón, I.; Saenger, W., An orthorhombic crystal form of cyclohexaicosaoase, CA26-32.59 H₂O: comparison with the triclinic form. *Carbohydr. Res.* **2001**, *336*, 141-153.
8. Chacko, K. K.; Saenger, W., Topography of cyclodextrin inclusion complexes. 15. Crystal and molecular structure of the cyclohexaamylose-7.57 water complex, form III. Four- and six-membered circular hydrogen bonds. *Journal of the American Chemical Society* **1981**, *103*, 1708-1715.
9. Betzel, C.; Saenger, W.; Hingerty, B. E.; Brown, G. M., Topography of cyclodextrin inclusion complexes, part 20. Circular and flip-flop hydrogen bonding in β -cyclodextrin

undecahydrate: a neutron diffraction study. *Journal of the American Chemical Society* **1984**, 106, 7545-7557.

10. Harata, K., The Structure of the Cyclodextrin Complex. XX. Crystal Structure of Uncomplexed Hydrated γ -Cyclodextrin. *Bulletin of the Chemical Society of Japan* **1987**, 60, 2763-2767.

11. Qi, Q.; She, X.; Endo, T.; Zimmermann, W., Effect of the reaction temperature on the transglycosylation reactions catalyzed by the cyclodextrin glucanotransferase from *Bacillus macerans* for the synthesis of large-ring cyclodextrins. *Tetrahedron* **2004**, 60, 799-806.

12. Qi, Q.; Mokhtar, M.; Zimmermann, W., Effect of ethanol on the synthesis of large-ring cyclodextrins by cyclodextrin glucanotransferases. *Journal of Inclusion Phenomena* **2007**, 57, 95-99.

13. Ivanov, P. M., Computational Studies on the Conformations of Some Large-Ring Cyclodextrins (CD_n, n=20, 21, 22, 23). *Chirality* **2011**, 23, 628-637.

14. Ivanov, P., Conformations of some lower-size large-ring cyclodextrins derived from conformational search with molecular dynamics and principal component analysis. *Journal of Molecular Structure* **2012**, 1009, 3-10.

15. Ivanov, P.; Atanassov, E.; Jaime, C., Computational study on the conformations of CD38 and inclusion complexes of some lower-size large-ring cyclodextrins. *Journal of Molecular Structure* **2014**, 1056, 238-245.

16. Gotsev, M. G.; Ivanov, P. M., Large-ring cyclodextrins. A molecular dynamics study of the conformational dynamics and energetics of CD10, CD14 and CD26. *ARKIVOC* **2007**, 13, 167-189.
17. Schnupf, U.; Momany, F. A., DFT Energy Optimization of a Large Carbohydrate: Cyclomaltohexaicosaoose (CA-26). *The Journal of Physical Chemistry B* **2012**, 116, 6618-6627.
18. Hansmann, U. H., Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* **1997**, 281, 140-150.
19. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, 314, 141-151.
20. Nymeyer, H.; Gnanakaran, S.; Garcia, A. E., Atomic simulations of protein folding, using the replica exchange algorithm. *Method Enzymol.* **2004**, 383, 119-149.
21. Cheng, X.; Cui, G.; Hornak, V.; Simmerling, C., Modified replica exchange simulation methods for local structure refinement. *J. Phys. Chem. B* **2005**, 109, 8220-8230.
22. Sittel, F.; Stock, G., Robust Density-Based Clustering To Identify Metastable Conformational States of Proteins. *Journal of Chemical Theory and Computation* **2016**, 12, 2426-2435.
23. Bottegoni, G.; Rocchia, W.; Cavalli, A. Application of Conformational Clustering in Protein, Ligand Docking. In *Computational Drug Discovery and Design*, Baron, R., Ed.; Springer New York: New York, NY, 2012, pp 169-186.

24. Blöchliger, N.; Caflisch, A.; Vitalis, A., Weighted Distance Functions Improve Analysis of High-Dimensional Data: Application to Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation* **2015**, 11, 5481-5492.
25. Bishop, C. M., *Pattern Recognition and Machine Learning*. Springer: 2006.
26. Raschka, S., *Python Machine Learning*. Packt Publishing: 2015.
27. Case, D.; Darden, T.; Cheatham Iii, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W., AMBER 10. *University of California, San Francisco* **2008**, 32.
28. Jacob, J.; Geßler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W., Strain-Induced “Band Flips” in Cyclodecaamylose and Higher Homologues. *Angewandte Chemie International Edition* **1998**, 37, 605-609.
29. Khuntawee, W.; Rungrotmongkol, T.; Wolschann, P.; Pongsawasdi, P.; Kungwan, N.; Okumura, H.; Hannongbua, S., Conformation study of ϵ -cyclodextrin: Replica exchange molecular dynamics simulations. *Carbohydrate polymers* **2016**, 141, 99-105.
30. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E., Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, 12, 2825--2830.
31. Dempster, A. P.; Laird, N. M.; Rubin, D. B., Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* **1977**, 39, 1-38.

Table of Contents graphic

