

การนำระเบียบวิธี PaDSAR ไปใช้บนโปรแกรม NAMD



นายจิระยุทธ สุปัญญาบุตร



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเคมี ภาควิชาเคมี
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2556
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย



5471932323

IMPLEMENTATION OF PaDSAR METHOD ON NAMD

Mr. Chirayut Supunyabut



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Chemistry
Department of Chemistry
Faculty of Science
Chulalongkorn University
Academic Year 2013
Copyright of Chulalongkorn University

จิระยุทธ สุปัญญาบุตร : การนำระเบียบวิธี PaDSAR ไปใช้บนโปรแกรม NAMD. (IMPLEMENTATION OF PaDSAR METHOD ON NAMD) อ.ที่ปรึกษาวิทยานิพนธ์
หลัก: รศ. ดร.พรเทพ สมพรพิสุทธิ์, 57 หน้า.

Pseudoatoms Driven Solvent Accessibility Refinement หรือ PaDSAR เป็นวิธีการทางคอมพิวเตอร์สำหรับการสร้างแบบจำลองโครงสร้างความละเอียดต่ำของเมมเบรนโปรตีน โดยอาศัยเงื่อนไขการบังคับซึ่งมีที่มาจากข้อมูลโซลเวนต์แอกเซสซิบิลิตีจากเทคนิคการติดสปีนที่ตำแหน่งจำเพาะและอิเล็กทรอนิกส์เรโซแนนซ์สเปกโทรสโกปี โดยวิธีการของ PaDSAR นั้น โครงสร้างของโปรตีนจะถูกขับเคลื่อนโดยอันตรกิริยาทั้งแรงผลักและแรงดึงดูดระหว่างซูดอะตอมที่ติดสปีนกับสิ่งแวดล้อมรอบข้างในระหว่างที่ทำการจำลองพลวัตเชิงโมเลกุล งานวิจัยนี้ได้แสดงให้เห็นถึงการนำเอาวิธี PaDSAR ไปใช้บนโปรแกรมการจำลองพลวัตเชิงโมเลกุลที่ชื่อว่า NAMD เพื่อการขัดเกลาโครงสร้างสามมิติของโพแทสเซียมแชนแนลจากแบคทีเรีย *Streptomyces lividans* (KcsA) กระบวนการทำงานของ NAMD-PaDSAR ซึ่งจะดำเนินการผ่านชุดคำสั่งสคริปต์ภาษา Tcl เป็นหลักนั้น ประกอบด้วย i) การเตรียมคอนฟิกูเรชันเริ่มต้นของระบบที่ประกอบด้วยโคออร์ดิเนตของโปรตีน การติดซูดสปีนไปที่อะตอม $C\alpha$ และซูดอะตอมของอนุภาคแวดล้อมที่มีการสร้างตำแหน่งมาก่อน และ ii) การจำลองพลวัตเชิงโมเลกุลที่มีเงื่อนไขบังคับ ในการเตรียมคอนฟิกูเรชันเริ่มต้นนั้น อนุภาคแวดล้อมอย่าง OXY กับ NIC จะถูกวางแบบสุ่มตำแหน่งภายใต้ขอบเขตที่กำหนดของบริเวณภายในและภายนอกเมมเบรนตามลำดับ NAMD-PaDSAR จะมีวิธีการปฏิบัติเช่นเดียวกับ CHARMM-PaDSAR ซึ่งเป็นต้นแบบที่เคยอธิบายไปก่อนหน้านี้ ยกเว้นเงื่อนไขบังคับของพลังงานศักย์ชนิดที่สองนั้นไม่สามารถนำมาใช้ได้ เนื่องจากเป็นปัญหาเฉพาะทางเทคนิค งานวิจัยนี้ยังแสดงให้เห็นว่าทางออกที่ดีที่สุดของปัญหาดังกล่าวคือการคำนึงถึงความยืดหยุ่นของซูดสปีนโดยปรับลดค่าคงที่ของแรงของมูมิมพรอพเพอร์ ไดฮีดรัล ผลการทดลองแสดงให้เห็นว่าโครงสร้างดีคอยซึ่งเป็นโครงสร้างที่ถูกบิดไปจากโครงสร้างเอ็กเซอร์โดยมีค่า RMSD อยู่ในช่วง 3 – 13Å จำนวน 433 จาก 524 โครงสร้างสามารถหมุนตัวกลับเข้าสู่การหมุนตัวที่ถูกต้องโดยมีค่า RMSD น้อยกว่า 3Å

ภาควิชา เคมี
สาขาวิชา เคมี
ปีการศึกษา 2556

ลายมือชื่อนิสิต จิระยุทธ สุปัญญาบุตร
ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก พรเทพ

5471932323 : MAJOR CHEMISTRY

KEYWORDS: NAMD / PADSAR / MD SIMULATION / RMSD

CHIRAYUT SUPUNYABUT: IMPLEMENTATION OF PADSAR METHOD ON NAMD. ADVISOR: ASSOC. PROF. PORNTHEP SOMPORNPIST, Ph.D., 57 pp.

Pseudoatoms Driven Solvent Accessibility Refinement or PaDSAR is a computational approach for modeling low-resolution membrane protein structure using restraints derived from the solvent accessibility data from site-directed spin labelling and electron paramagnetic resonance spectroscopy. In the PaDSAR method, protein conformations were driven by repulsive and attractive interactions between the spin-label pseudoatoms and the surrounding environment during molecular dynamics (MD) simulations. This work has demonstrated an implementation of PaDSAR on the molecular dynamics simulation program, NAMD for refining the 3D structure of a potassium channel from *Streptomyces lividans* (KcsA). Working steps in NAMD-PaDSAR that are mainly conducted through the tcl script commands, include i) preparing the initial configuration of the system consisting of protein coordinates, pseudo-spins attaching to the C α atom and the surrounding particles with pre-generated positions, and ii) performing MD simulations with the restraints. For the initial configuration, the surrounding OXY and NIC particles were randomly placed in the simulation within the specified boundary of membrane and non-membrane regions, respectively. NAMD-PaDSAR was treated the same way as that previously described in the original CHARMM-PaDSAR, except that the incorporation of the type II restraint potential was not successful due to a technical problem. This study also demonstrated that the best solution for the present problem was to take a flexible pseudo-spin into account via modifying the force constant of improper dihedral angle. The results showed that 433 out of 524 decoys which has the structure distortion based on RMSD to the x-ray structure in a range of 3 – 13Å, can refold back to the correct fold with RMSD < 3Å.

Department: Chemistry

Field of Study: Chemistry

Academic Year: 2013

Student's Signature CHIRAYUT SUPUNYABUT

Advisor's Signature



ACKNOWLEDGEMENTS

I would like to express my grateful thanks to Associate Professor Dr. Pornthep Sompornpisut, my research supervisor, for his useful advice, encouragement and patient guidance. My special thanks are also extended to the exam committee, Assistant Professor Dr. Warinthorn Chavasiri, Assistant Professor Dr. Somsak Pianwanit and Dr. Arthorn Loisruangsin for their constructive suggestions and I wish to thank for the help provided by Pornthep research group's members. I would like to acknowledge the Computational Chemistry Unit Cell (CCUC) at Department of Chemistry, Faculty of Science, Chulalongkorn University for research facilities and resources. And special thanks for following funding agencies for their financial support, the National Research University Project of CHE (HR1155A), Ratchadaphiseksomphot Endowment Fund (RES560530217-HR).

Finally, I would like to offer my very great appreciation to my family for their support and encouragement throughout my research.



CONTENTS

	Page
THAI ABSTRACT	iv
ENGLISH ABSTRACT	v
ACKNOWLEDGEMENTS	vi
CONTENTS	vii
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xiii
CHAPTER I INTRODUCTION	1
1.1 NAMD and molecular dynamics simulation	1
1.2 Site-directed spin labeling and electron paramagnetic resonance	4
1.3 Membrane proteins and potassium channels	6
1.4 Pseudoatom-driven solvent accessibility refinement	9
1.5 Rationale and Objectives.....	10
CHAPTER II MATERIALS AND METHODS	11
2.1 Materials.....	11
2.1.1 Hardware	11
2.1.2 Software	11
2.2 Methods	12
2.2.1 Writing the Tcl script file for running the simulation in NAMD	13
2.2.2 Building the system containing pseudoatoms	14
2.2.3 Molecular dynamics simulation	18
2.2.4 Analysis of MD trajectory	23
CHAPTER III RESULTS AND DISCUSSION.....	24
3.1 Protocol for preparation of model system with pseudoatoms.	24
3.2 Amount of CPU time	36
3.2.1 The effect of tclBC and the modification of bubble.tcl	36
3.2.2 Number of OXY and NIC PsDAs.....	38



	Page
3.2.3 The number of simulation step.....	39
3.3 The radial distribution function of pseudoatoms.....	40
3.4 Varying the Improper dihedral force constant (K_{imp}) of EPs atom.....	41
3.5 Refolding the decoy to the native conformation	47
CHAPTER IV CONCLUSION	49
REFERENCES	51
APPENDICES.....	52
VITA.....	57



229682773

LIST OF TABLES

	Page
Table 2.1 KcsA residues with the assignments of the pseudoatom types.....	15
Table 2.2 Pseudoatom interactions defined in terms of Lennard-Jones Like potential	17
Table 2.3 Force field parameters for the pseudoatoms	17
Table 3.1 The comparison of the simulation time spending.	38
Table 3.2 Classification of amino acid residues by the position in environments.	43
Table 3.3 Classification of amino acid residues by the solvent accessibility.....	45



LIST OF FIGURES

Page

Figure 1.1 Atomic connectivity defined for bonds and angles in the bonded potential energy A. The covalence bonding and the distance r between two atoms. B. The angle of three adjacent atoms and θ is the angle defined in <i>Uangle</i> . C. The dihedral angle, φ and D. The improper dihedral angle. φ is controlled by α and γ	2
Figure 1.2 The chemical modification of cysteine residue with the spin labeling, methanethiosulfonate spin label (MTSL).	4
Figure 1.3 The experimental EPR data obtained from SDSL/EPR technique. A. Mobility data B. Solvent accessibility.....	5
Figure 1.4 Models of three type helical proteins and their ΠO_2 , non-polar solvent, $\Pi NiEDDA$, polar solvent, as a function of residue position A. A soluble protein, represents the residues that expose to the solvent B. A transmembrane protein and C. A surface-adsorbed helix.....	6
Figure 1.5 Three classes of membrane proteins.....	7
Figure 1.6 The illustration of potassium channel	8
Figure 1.7 The cartoon represents the KcsA.....	8
Figure 1.8 A. Electron paramagnetic resonance experimental schemes B. Scenario of a membrane protein having different surrounding environment C. Attachment of pseudoatom EP-type into transmembrane segments of KcsA channel D. A snapshot (a tetramer structure of KcsA channel and the environment pseudoatoms (OXY=red, NIC=blue)).....	10
Figure 2.1 Schematic shows the preparation of protein to run PaDSAR on NAMD. 1 Setting initial parameter in the run.tcl. 1.1 An example of run.tcl. 2 Generating the needed files for running the simulation on NAMD 2.1 Generating the protein coordinate file. 1) Initial coordinates of the protein in monomer. 2) In case of symmetry multi-mers proteins. The monomer of protein will be duplicated and rotated to generate the multi-mers. 3) The EPs types pseudoatoms will be attached to an α -carbon in backbone of protein relate to the SDSL/EPR. 4) Several hundred of OXY and NIC pseudoatoms (PsDAs) are distributed to the system representing the environment background. 3 Getting the output files for running PaDSAR. 3.1 examples of output files. 4 Running the PaDSAR on NAMD and obtaining the PaDSAR protein. 12	
Figure 2.2 Tetramer X-ray structure of a potassium channel from <i>Streptomyces lividans</i> (KcsA) TM1 and TM2 helices of a monomer are shown in red and orange helices respectively.....	14



229682773

Figure 2.3 A. Scenario of a membrane protein having different surrounding environment B. Attachment of pseudoatom EP-type into transmembrane segments of KcsA channel.....	15
Figure 2.4 NAMD snapshot (a tetramer structure of KcsA channel and the environment pseudoatoms (OXY=red dot, NIC=blue dot)).....	16
Figure 2.5 Topology of EPR pseudoatoms	16
Figure 2.6 Lennard-Jones potential energy represents three types of pseudoatoms interaction.....	18
Figure 2.7 The graph represents the van der Waals energy with (red dot) and without (green line) using switching function.	20
Figure 2.8 The kinds of non-bonded atoms to be excluded when using exclude scaled1-4 command.	20
Figure 3.1 The parameters for generating the OXY and NIC initial position.....	25
Figure 3.2 Output files from NAMD.....	36
Figure 3.3 The cartoon represents how the <i>cleardrops</i> and <i>dropatom</i> work. The tclBC A. without <i>cleardrops</i> and <i>dropatom</i> commands. All atom in system will be considered in tclBC loop every step. B. with the <i>cleardrops</i> and <i>dropatom</i> command. All atoms will be considered only the step that set as the <i>cleardrops</i> step and the atom labelled <i>dropatom</i> will not be considered for the other.	37
Figure 3.4 The box chart shows the average RMSD of 4 models.....	39
Figure 3.5 The RMSD for 8 data sets A. 1,000,000 steps B. 50,000 steps.....	39
Figure 3.6 The radial distribution function plot, $g(r)$ black line, and its integration, blue line, of nine pairs of pseudoatoms.	41
Figure 3.7 The RMSD of five samples with varying the improper dihedral force constant.....	42
Figure 3.8 The cartoon shows classification of amino acid residues into three types related with environments and defined by the distance in z direction (Å) form the origin coordinate. SOL: solution ($z > 15$ or $z < -15$), MEM: membrane ($-15 < z < 15$) and INT: interface (residue that overlaps both SOL and MEM).	44
Figure 3.9 The graph of the improper dihedral angle of EP atoms varying K values during the simulation, 70 (blue), 55 (red), 25(green) and 10 (black) $\text{kcal mol}^{-1} \text{rad}^{-2}$ in different positions related with the environment (column) and solvent accessibility (row). y-axis is the improper dihedral angle and x-axis is the simulation frame.....	46
Figure 3.10 The graph shows the initial and final RMSD of 524 samples and the orange shade is the acceptable area for the final RMSD.....	47



Figure 3.11 The bar graph shows the percent refolding. The first column is the overall decoys and the others are the percent refolding in an individual range. 48



LIST OF ABBREVIATIONS

MD	molecular dynamics
VMD	visual molecular dynamics
SDSL	site-directed spin labeling
EPR	electron paramagnetic resonance
MTSL	methanethiosulfonate spin label
PaDSAR	pseudoatom-driven solvent accessibility refinement
ΔH_0^1	probe mobility
ΠO_2	oxygen accessibility
NiEDDA	nickel-ethylenediaminediacetate
Π NiEDDA	nickel-ethylenediaminediacetate accessibility
Kv	voltage-gated potassium channels
KcsA	prokaryotic potassium channel from soil bacteria <i>streptomyces lividans</i>
TM	transmembrane
PDB	protein structure coordinate file (protein data bank)
PSF	protein structure file
PsDAs	pseudoatoms
NiC	nickel
OXY	oxygen
RMSD	root-mean square deviation
CPU	central processing unit
RDF	radial distribution function
ns	nano seconds
EP, PVP	pseudoatoms that represent spin labeled atoms



229682773