2017.03.08 meeting 2

Sunday, October 8, 2017 10:15 AM

Wikipage

1. We can't start wikipage editing yet until the design group has designed the team logo and decided the theme.

Simulation tool

1. AutoDock
   A simulation tool that use for molecular docking simulation, but this tool can just show a series of data, and needs another tool called PyRx in order to visualize the molecules.
2. PyRx
   It is able to visualize the molecules those simulated by AutoDock, so that we can observe the figure of molecular docking situation.
3. Protein Data Bank (PDB)
   Provides detailed protein information by input protein ID.
4. Discovery Studio
   Use for protein structure design, enzyme optimization, and animation making. It can also repair the defective protein sequence by performing some calculation.

Decisions

1. Share the notebook contents on Google Drive in order to confirm working process, and also collect more information.
2. Notebook sharing: every Thursday

2017.03.15 meeting 3

Thursday, October 5, 2017 7:51 PM

Wikipage

1. Team logo had tentatively designed out.
2. The problem on footer has to be fixed.
3. The temporary wikipage components are video and team member introduction.
4. The image background should be united, and the background color should also be bright but not too dark.

Simulation tool

AutoDock:

1. We've realized that the function of this tool is limited, and it is suitable for ranking of numerous proteins to select the proteins those are best fix to our need.
2. We plan to find out the xylanase and glucanase from PDB, and then test for their binding affinity with their ligand.
3. Start from the glucanase-cellulose binding affinity test.

Discovery studio:

1. This is a professional application program that provides the function of simulation and prediction on life sciences research.
2. However, this software just provide for people and organization with having a professional license to use, so students can just borrow it from the authorized facilities.
3. Functions:
   Simulation
   macromolecule design & analysis
   antibody modeling
   drug design
   toxicity prediction
   x-ray
   visualization (free)
4. Discovery Studio Visualizer (free version):

A free multifunctional molecular simulator, just need to sign in

5. Basic functions of this simulator are visualization, macromolecule design, ligand-based design, structure-based design, etc.
6. We have already signed in to this free simulator by using iGEM_CGU_2017 in name.

Plans for next week

1. Test each function for Discovery Studio Visualizer.
2. The team logo and footer on wikipage need to be fixed.
3. Test for the binding affinity between xylanase and cellulose by using AutoDock.

2017.03.21 meeting 4

Wikipage

1. Hold on wikipage editing until deciding the team logo theme.

Discovery Studio & AutoDock

1. See whether the molecular medicine research center has bought the full version discovery studio.
2. Discovery Studio can be used to perform molecular docking, test binding affinity between two molecules, and also make animations.

Problems

1. The mechanism of reaction between cellulose and cellulase.
2. Cellulase breaks the (β1-4) glycosidic linkage in cellulose
   → how to break? Is the reaction need to be worked with other cofactors?
3. Which cellulase should we select?

Plan for next week

1. Improvement of wikipage
2. Dr. Chun-Hsien Chen from department of information management has published a paper about the research of post-translational modification and kinase analysis of bladder cancer
   → ask the protein simulation problems
3. Continue to learn Discovery Studio Visualizer
4. Understand the reaction mechanism between cellulose and cellulase.

2017.04.12 meeting 6

Current work (complete on April)

1. Look for proteins with high amino acid sequence similarity by using SWISS-MODEL, and then download PDB files.
2. Work on affinity ranking by using AutoDock
   → problems: origins, procedures, mehanism
3. Confirm operation procedures for each software, and also understand the mechanism of cellulase reaction.
4. Find related papers
   → simulation of reaction between cellulose and cellulase (if no enough information, we can start from finding starch decomposition mechanism.)

Future plans

1. Search for more information
2. Ask professor form NCHU the detailed mechanism of cellulose decomposition.
3. Complete modeling before summer vacation
4. During summer vacation, we focus on making growth curve and wikipage (need to adjust the working progress with bio-development group).

Others

1. Change meeting date
   4/26 → 4/25 5:00 p.m.
   5/10 → 5/9 5:00 p.m.
2. Experiment principle class on 4/28 (Friday) 4:00 p.m.

2017.04.25 meeting 8

Energy minimization of domain

1. The energy of cellulose needs to be minimized when working on modeling.
2. In the modeling process, the system might automatically modifies domain and affects the actual efficiency of docking, so the energy of domain has to be minimized in order to increase accuracy of docking test.
3. The common energy minimization methods are: steepest descent method, conjugate gradient method).
4. Software: MMC from ChemOffice

Xylanase binding affinity test in AutoDock

1. Although the proteins those found in NCBI are produced by fungi, the protein file format is not PDB format, and the file conversion process is complicated, so we decide to adopt proteins those found from SWISS-MODEL to work on binding affinity test.
2. In the first test, we've selected 5 proteins with different amino acid sequence similarity:
   i. 5K9Y – original Xylanase (from *bacillus subtilis*)
   ii. 2DCZ – highest similarity (from prokaryote)
   iii. 3ZSE – 69% similarity (from prokaryote)
   iv. 1PVX – 50% similarity (form eukaryote)
   v. 1XYO – 55% similarity (from eukaryote)
3. 20 minutes for each protein docking

Problems

1. In the first docking test in AutoDock, the 3 proteins' docking are failed out of 5 (5K9Y, 3ZSE, 1PVX are failed). Why?
2. In the docking test, we realized that a macromolecule can be used to test with many ligands, but a ligand can't be used to test with many macromolecules.
   → can we set enzymes to be ligand, cellulose to be macromolecule?
3. In AutoDock, the system has performed modification such as energy minimization for ligand and macromolecule before docking or not?

Plan for next week

1. Test for the binding affinity of 10 different eukaryotic proteins with high sequence similarity, and we will have Dr. Po-Jung Huang confirm the result.
2. Solve the above problems

2017.05.04 meeting 9

Progress of this week

1. The proteins those found from SWISS-MODEL have lots of overlapping, but later we've realized that the overlapped proteins are the contrast of two different system (BLAST & Hhblist), so we decide to select proteins from Hhblist (total related proteins of xylanase are 96).
2. We'll try to use a free software called Rosetta ddg monomer for cellulose energy minimization.
3. We have not found the method for mass protein docking.

Problems

1. In AutoDock, the conversion of enzyme into macromolecule always fail due to the amino acid selection.
   → ask Dr. Po-Jung Huang

Plans for next week

1. We start to work from proteins those found from SWISS-MODEL, search the origin organisms for each protein (search whether prokaryotic or eukaryotic).
2. Organize the modeling process to power point format, and then we'll have Dr. Huang confirm the procedures.

2017.05.09 meeting 10

Mass protein docking

1. Found a softeare for mass protein docking: Raccoon AutoDock
2. Downloaded an auxiliary software called MGL Tools 1.5.4, but failed to download it because the system does not support.
3. Downloaded old version and succeed to install.
4. Raccoon needs another programming language software called Python, and we've already succeed to install, but we have not operate it.

Energy minimization by Rosetta

1. We have to sign in before using, but failed to sign in.

Conversation with Dr. Po-Jung Huang

1. Why the conversion of protein to macromolecule always fail?
2. Why some proteins need to select amino acid conformation A/B but some proteins don't need?

Problems

1. When the Raccoon start to perform mass protein docking, we realized that the files will automatically converted, but need another docking parameter file
   → what is this 'docking parameter file'?

Plans for next week

1. Continue to look for the origin organisms for each xylanase related protein.
2. Search information about docking parameter file.

2017.05.17 meeting 11

Macromolecule conversion

1. No Gasteiger parameter (failed to produce Gasteiger for PDB).
2. We have to look for AutoDock repair function in order to repair PDB file.

Mass protein docking

1. The online protein docking is unable to perform mass docking, and it is also take a long time.
2. AutoDock Raccoon needs to be operated with workstation server / Linux Cluster.
3. Workstation server can't activate Linux Cluster. The cost is very high, and we also cannot assure that we can succeed.

Energy minimization

1. Still no reply from Rosetta.
2. Our school might has someone who bought ChemOffice, so we have already sent email to Dr. Siao.

2017.05.25 meeting 12

Progress of this week

Discovery Studio:

1. Can use BLAST to search templates.
2. The macromolecule and ligand can be fixed automatically in the binding affinity testing.
3. Still cannot do batch docking, have to put the macromolecule one by one.

Problems

1. The docking program could only be operated by Windows 7 instead of Windows 10.
2.

Plans for next week

1. Consider how to arrange the result.
2. After the problem be solved, try to run the 96 similar templates of xylanase for the affinity test on Discovery studio.

2017.05.31 meeting 13

The result selection

1. Which data should be selected for the final result present of the ZDock Score?
2. The ZDock Score of a pose maybe vary from different reference point of the others pose's ZDock score.
3. Solution: We came out with an idea to take the average score from every pose of the protein, but we're not sure if this makes sense. We still need to consider the rationality.

Select how many data as the result

1. Merely one docking result of a macromolecule 5k9y has 2000 data from different poses.
2. Take Xylanase as example, there are 96 similar proteins with 2000 data for each, the result will be too huge to deal with.

Problems

1. Do we have to define the RMSD reference value?
2. How to unify and present the result data.

Plans for next week

1. To understand the denotation of ZDock Score in Discovery Studio.
2. To consult Dr. Po-Jung Huang on how to present the final result of the docking score. (Take the average value, maximum value or minimum value?)
3. Complete the docking test result of xylanase.

06.03

- We can't do the comparison of different enzyme's ZDOCK result and RMSD.
- There are "Binding Energy testing" in Discovery Studio, but the result is weird. (Give up this method)
- Solution of protein conformation problem
    - Initial setting:
        - Edit -> Preferences -> Protein Utilizes -> Clean Protein -> select Alternate conformations (retain one set)

Xylanase - Cellulose 對接測試

PAGE No.0770230
CONTINUED FROM PAGE
DATE 2017.06.03

- 使用 ZDOCK 無法做不同酵素的對接結果相互評比
- RMSD 無法做不同酵素的對接結果相互評比
- Discovery Studio 有 Binding Energy testing
  - 遇到有 Conformation 問題的酵素，可利用 Clean Protein 功能去 fix
  - 實用
    - 初次設定 Clean Protein：上方 Edit → Preferences
      → 左側 Protein Utilies → Clean Protein
      → 勾選 Alternate conformations (retain one set)
    - Clean Protein 功能在左側 Protein Report and Utilities 裡
  - 取幾個酵素做計算後得出的數值怪異，因此放棄使用此對接方法
- 改回使用 AutoDock 做對接，如遇 Conformation 問題就使用 Discovery Studio 來 fix (Clean Protein 功能)
- 單一個 AutoDock 無法一次進行多個 Receptor 對接，以及使用的 Ligand 要使用 標記為不同的名字，不然會有對接結果覆蓋的問題
- 把各別的對接結果 (Binding Affinity) 用 excel 做整理即可進行排序
- 有些酵素還是無法做對接

賴威安

CONTINUED ON PAGE
DATE 2017.06.05

Description

In the enzymatic deinking process, our project mainly focus on three deinking enzymes, such as xylanase, glucanase, and lipase. However, there are many different kinds of organism producing these enzymes, and the amino acid sequences are slightly different from each other. In order to find out the enzymes which have the best deinking efficiency, Information and Simulation group has utilized several software and technique to analyze hundreds of enzymes' binding affinity with their ligand, and provided the result to bio-development group for constructing biobricks.

We have firstly searched the templates of related homologous proteins by using SWISS-MODEL, and downloaded all these protein 3D structure files from RCSB protein data bank. Before performing protein docking, the protein conformation problem is fixed, and the energy minimization of ligands is also done by using Discovery Studio in order to increase the accuracy of result.

Another simulation software, AutoDock can simulate the docking between enzyme and ligand, and helps to calculate the binding affinity. We have ranked the binding affinity of hundreds of enzymes, and visualized the top 10 docking situation by using PyMOL. Consequently, the bio-development group able to construct the plasmid base on this docking result.

Furthermore, Information and simulation group has also built the wiki page to present our project, the enzymatic deinking developed by 2017 iGEM CGU Taiwan team.

06.28

雜項記錄

PAGE No.0770235

CONTINUED FROM PAGE

DATE 2017.06.28

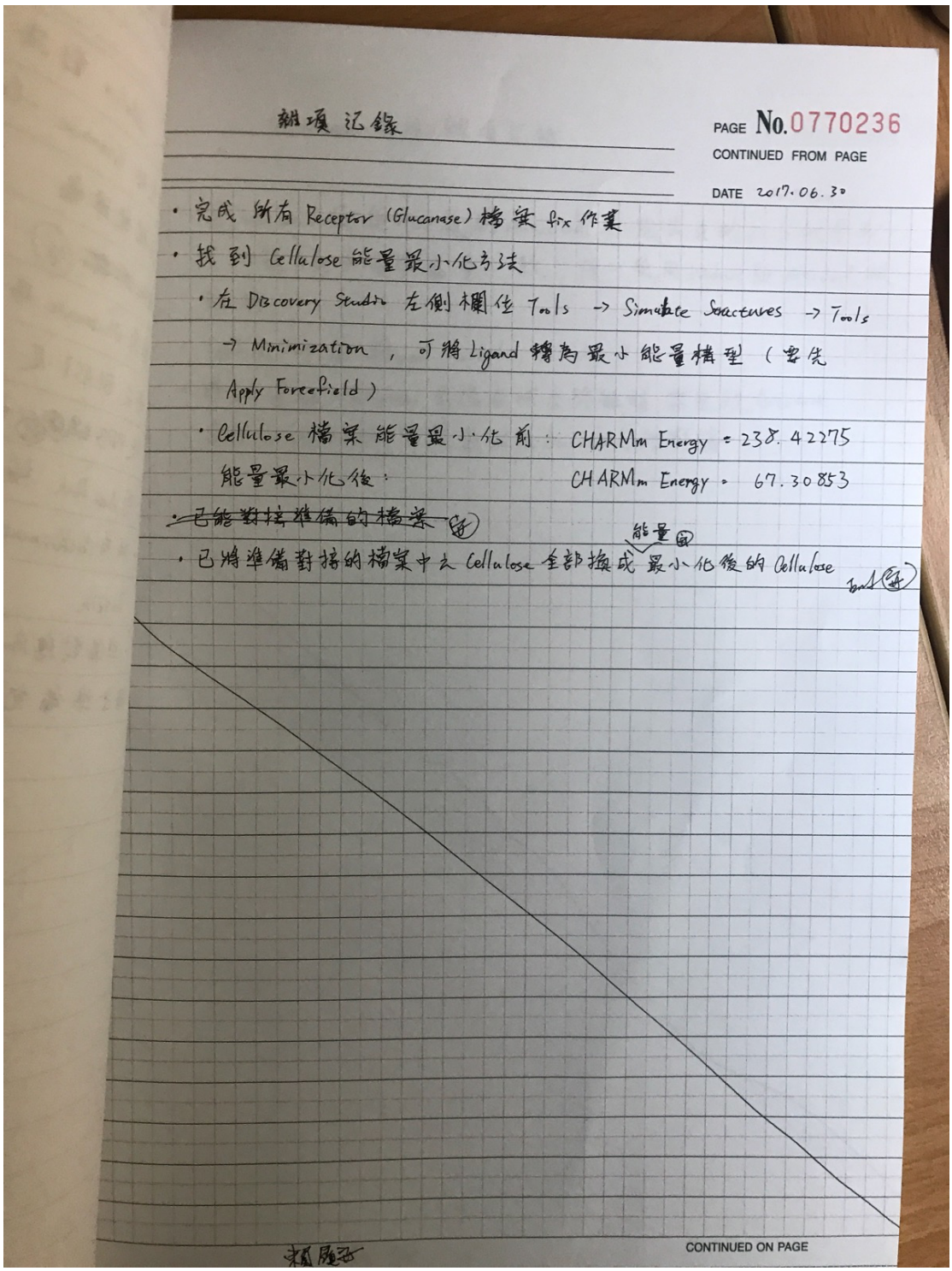
- Glucanase - Cellulose 對接
  - 依照先前 Xylanase -Cellulose 的模式進行對接及排序
    - 對接檔案準備
      - ~~相似 與 Glu~~
      - 與原 Glucanase 序列相似的酵素搜索
      - 刪去 BLAST 及重複的酵素
      - 從 PDB 網站取得所有檔案
      - 用 Auto Dock 檢測每個 Receptor 是否有 Conformation 問題
      - 將有 Conformation 問題的 Receptor 放入 Discovery Studio 進行 Clean Protein
      - 準備對應各別標記的 Cellulose 檔案
    - 檔案準備完成後放入 Auto Dock 做對接

SIGNATURE OF INVESTIGATOR

CONTINUED ON PAGE

DATE 2017.06.28

06.30



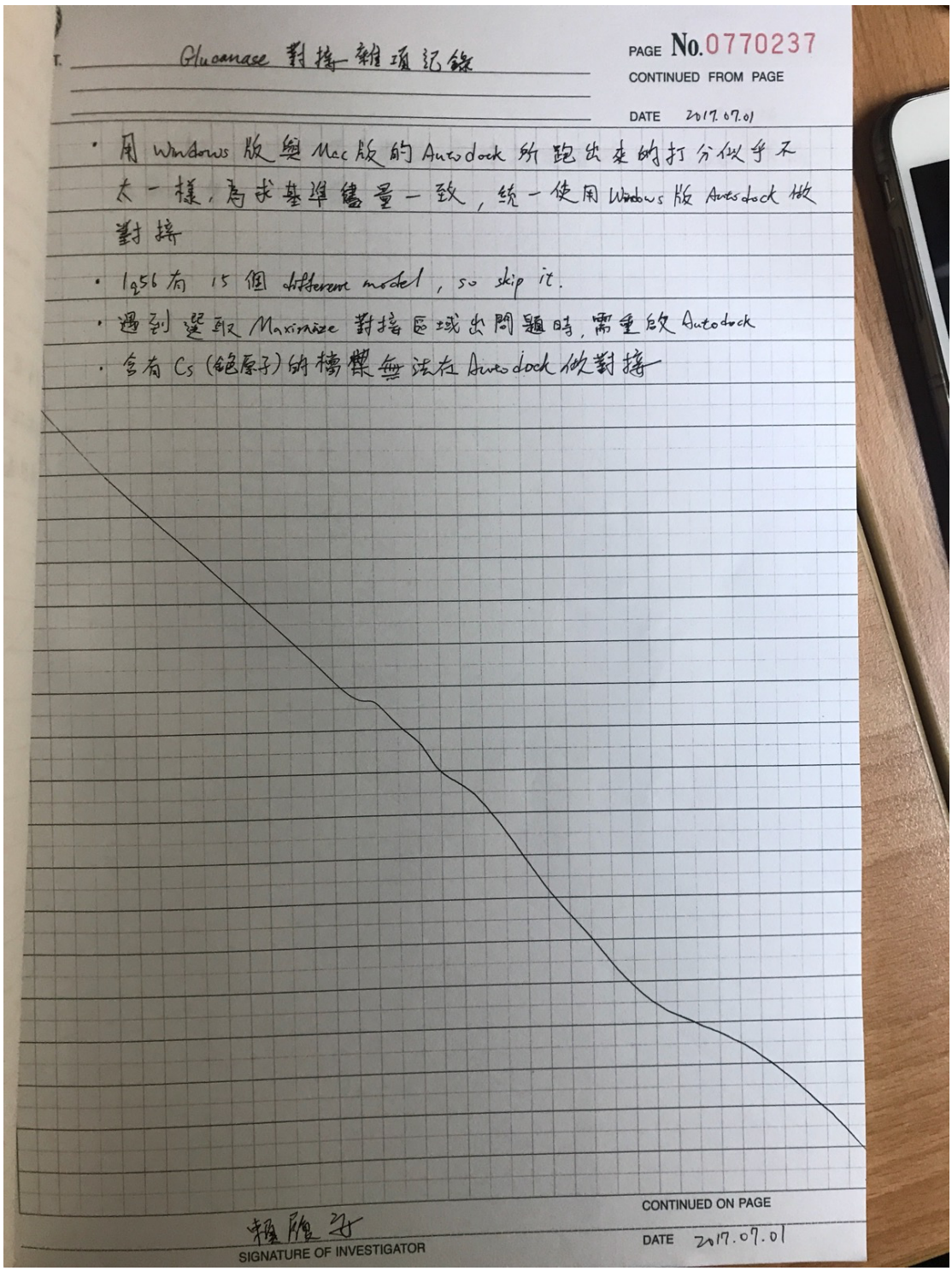
雜項記錄

PAGE No. 0770236
CONTINUED FROM PAGE
DATE 2017.06.30

- 完成 所有 Receptor (Glucanase) 檔案 fix 作業
- 找到 Cellulose 能量最小化方法
  - 在 Discovery Studio 左側欄位 Tools → Simulate Structures → Tools → Minimization，可將 Ligand 轉為最小能量構型（要先 Apply Forcefield）
  - Cellulose 檔案 能量最小化前：CHARMm Energy = 238.42275
    能量最小化後：CHARMm Energy = 67.30853
- ~~已能對接準備的檔案~~
- 已將準備對接的檔案中之 Cellulose 全部換成 最小化後的 Cellulose（能量圖）

CONTINUED ON PAGE

07.01



Glucanase 對接雜項記錄

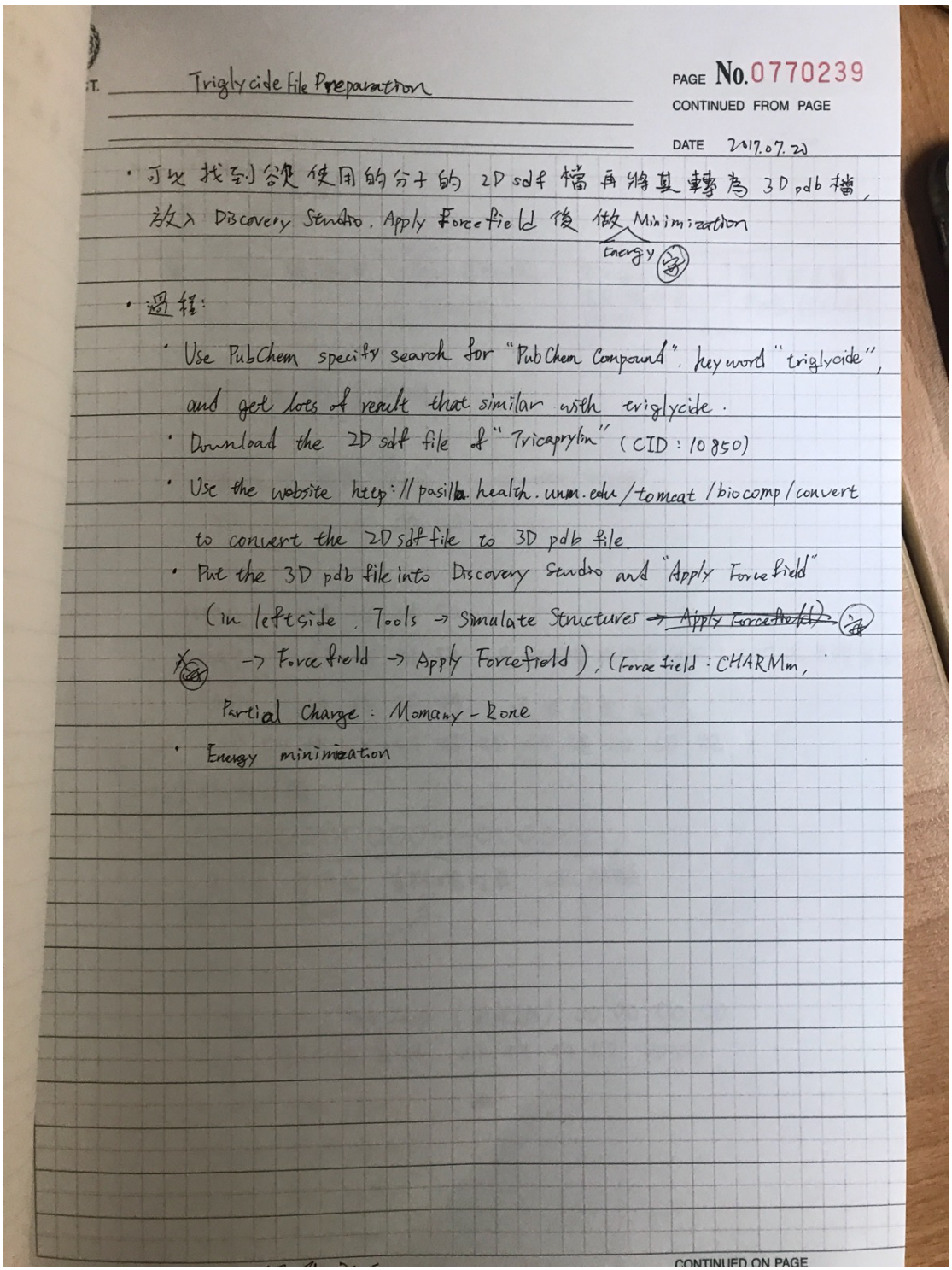
PAGE No.0770237
CONTINUED FROM PAGE
DATE 2017.07.01

- 用 Windows 版與 Mac 版的 Autodock 所跑出來的打分似乎不太一樣，為求基準統一致，統一使用 Windows 版 Autodock 做對接
- 1g56 有 15 個 different model，so skip it.
- 遇到選取 Maximize 對接區域出問題時，需重啟 Autodock
- 含有 Cs (鉋原子) 的檔案無法在 Autodock 做對接

賴順子
SIGNATURE OF INVESTIGATOR

CONTINUED ON PAGE
DATE 2017.07.01

07.20



T. Triglycide file Preparation

PAGE No. 0770239
CONTINUED FROM PAGE

DATE 2017.07.20

- 可以找到會使用的分子的 2D sdf 檔再將其轉為 3D pdb 檔，放入 DScovery Studio. Apply Forcefield 後做 Minimization Energy
- 過程：
  - Use PubChem specify search for "PubChem Compound", keyword "triglycide", and get lots of result that similar with triglycide.
  - Download the 2D sdf file of "Tricaprylin" (CID : 10850)
  - Use the website http://pasilla.health.unm.edu/tomcat/biocomp/convert to convert the 2D sdf file to 3D pdb file.
  - Put the 3D pdb file into Discovery Studio and "Apply Forcefield" (in leftside, Tools → Simulate Structures ~~→ Apply Forcefield~~ → Forcefield → Apply Forcefield), (Forcefield : CHARMm, Partial Charge : Momany-Rone
  - Energy minimization

CONTINUED ON PAGE