

蛋白質之分子模擬：蛇毒蛋白

吳姿嫻[†]、李育銘[†]、程麗梅[†]、孫英傑^{*}
臺灣師範大學化學系（[†]碩士生，^{*}副教授）

摘要

本文介紹計算化學中原子層次蛋白質模擬的基本方法，並藉由兩個應用例子，討論蛋白質模擬在生物物理化學研究中的一些課題。

一、簡介

由於電腦的快速發展，已使吾人得以藉助其快速、大量的數值計算能力，在分子模擬領域中，處理較大的物理、化學系統。其中，在生命科學的應用方面，生化分子系統的模擬已有相當的進展，得以檢視生化系統的平衡及動力學性質，成為與實驗相輔相成，探討生物物理化學系統的重要工具。在本文中，筆者將以蛋白質模擬為例，介紹生化系統模擬的一些發展，筆者將簡介蛋白質模擬的基本原理，並以兩個應用的例子—蛋白質反折疊（unfolding）、骨架動力學（backbone dynamics），討論蛋白質模擬的幾個課題。在此說明，本文中的蛋白質模擬，侷限於原子層次的分子模擬，不包含更簡化模型的蛋白質模擬，以下簡介這一類型蛋白質模擬的基本方法。

二、方法

在分子模擬中，分子動力學（molecular dynamics/MD）和蒙地卡羅（Monte Carlo/MC）模擬

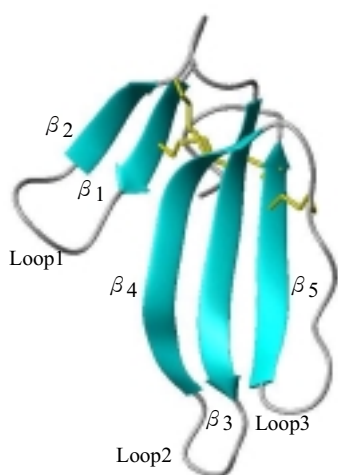
是探討物理、化學系統動態及平衡性質的重要方法，對蛋白質這類大分子，描述原子與原子間的作用力時，通常將原子視為一質點（不外顯性地（explicitly）包含電子），原子與原子間作用位能，使用一經驗（empirical）的位能函數來描述，最常用的位能形式為分子力學力場（molecular mechanics force field），其表示式如式(1)：

$$V = \sum_{bonds} \frac{k_r}{2} (r - r_0)^2 + \sum_{angles} \frac{k_\theta}{2} (\theta - \theta_0)^2 + \sum_{dihedrals} \frac{V_n}{2} [1 - \cos(n\phi - \gamma)] + \sum_{ij} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (1)$$

式(1)中前三項屬於分子中鍵結的位能參數， r 為鍵長， θ 代表鍵角， ϕ 為扭動角， k 為常數。這三項的參數通常使用小且簡單的相似分子進行高層次（level）的 *ab initio* 計算而得；第四項為凡得瓦作用力，第五項為庫侖靜電作用力，力場中的參數可由實驗數據或理論計算推得。有此分子力場在手，吾人即可進行分子動力學或蒙地卡羅模擬。

以下，我們討論兩個運用蛋白質模擬探討蛋白

質反折疊及其骨架動力學的例子，及其中相關蛋白質模擬的課題。我們所模擬的蛋白質是台灣眼鏡蛇毒液中第三號 (CTX3) (見圖一) 及第四號 (CTX4) 心臟毒蛋白 (Cardiotoxin/CTX)；CTX 蛋白的結構、折疊 (folding) 與其他生化分子結合，及其他生化作用機制乃 CTX 蛋白研究的重要課題^[1,2]；這些蛋白質為 60-62 個胺基酸所組成，在三度空間成平面狀，具有五個 β 股的高分子^[3-12]。實驗方面，重要的發現使吾人更深入了解心臟毒蛋白^[1-18]。在下一個章節中，我們將討論運用蛋白質模擬方法，探討 CTX3 的反折疊和 CTX4 骨架動力學的計算程序結果。



圖一、由台灣眼鏡蛇蛇毒得到的 CTX3 蛇毒蛋白之二級結構的帶狀 (Ribbon) 圖，四個主鏈間互相連結的雙硫鍵亦顯示在圖上。

三、心臟毒蛋白之模擬

蛋白質展開的模擬 蛋白質模擬中，檢視蛋白質展開路徑一種常用的方法，乃將模擬系統增溫以檢

視蛋白質展開過程結構的變化；展開過程中的過渡態 (transition state/TS)、中間態 (intermediate state) 及展開態 (unfolded state) 的自由能差，決定蛋白質展開的動力學 (kinetics) 過程，但準確計算蛋白質各種自由能差仍是一個發展中的課題 (見下面段落之討論)^[19]；除了檢視蛋白質展開的自由能外，以結構為基礎，檢視展開過程中結構相似性的構形分析方法，亦可被用來找蛋白質展開過程中的過渡態及中間態^[20-22]；先前運用此方法於許多蛋白質的計算研究中，已得到與實驗資料一致的結果^[22,23]，以下筆者敘述 CTX3 的模擬程序，構形分析方法^[20,24]，及計算結果。

CTX3 的展開

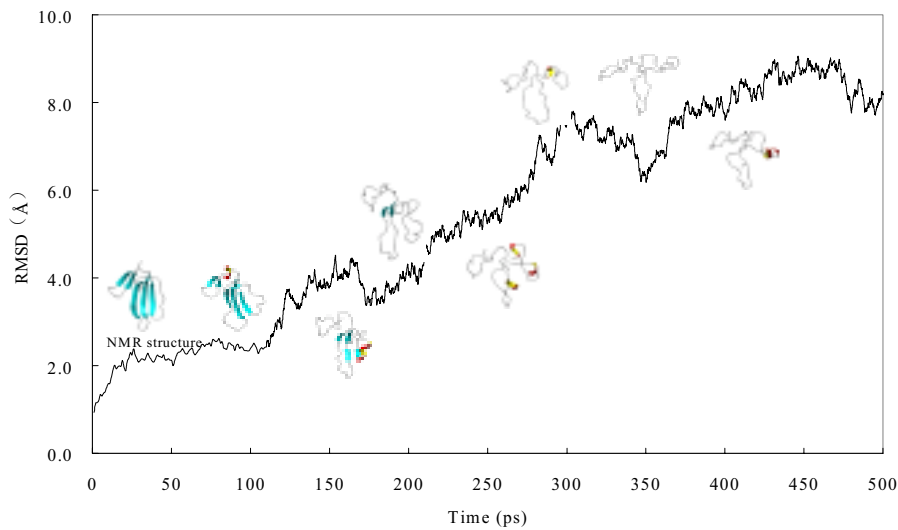
蛋白質展開的模擬，首先從折疊態開始。CTX3 的結構得自於蛋白質資料庫 (Protein Data Base, <http://www.rcsb.org/pdb/>)。組成天然蛋白質的 20 個胺基酸在 pH=7 下，Lys 和 Arg 帶正電，Asp 和 Glu 帶負電；在蛋白質中，則依其 pKa 值決定其帶電狀態。例如 NMR 光譜的測定得到 CTX3 中 Asp40 的 pKa 值約為 3.2，Asp57 則小於 2.3^[14]，若模擬 pH=3 的溶液；則前者的帶電狀態呈中性，後者則帶負電。由於在蛋白質模擬中使用一經驗力場，從蛋白質資料庫中得到實驗的結構，在模擬中，首先將蛋白質用最陡梯度極小化法 (steepest descent method) 做 1000 步的能量極小化，之後把蛋白質放到一個具水分子盒子裡，完成蛋白質在水中的溶液模擬系統設定，(蛋白質的原子和盒子 x, y, z 方向邊緣的最短距離不小於 10 Å，而使得盒子的大小約接近為 60x60x60 Å³)^[25]。加完水後的溶液系統做 8000 步的最陡梯度極小化法作能量極小化，再接著做 2000 步的共軛梯度法 (conjugate gradient method)

的能量極小化。之後進行蛋白質 MD 模擬。MD 模擬中起始溫度為 30 K，而後在 10 ps 內漸升到 300 K，之後再接下去做 300 K 的平衡模擬^[26]。平衡了 100 ps 之後，爲了讓 CTX3 展開，整個系統的溫度調升到 600 K，維持數百 ps 的高溫，讓 CTX3 展開使其 RMSD 值上升到約 10 Å（見圖二）。除了上述分子力場外，在蛋白質溶液模擬系統中，須考慮與水分子的作用力，對於非共價作用力的計算方法我們採用的是 8 Å cutoff 方法^[27,28]。此外，所有和氫原子鍵結的鍵都用 SHAKE 方法計算^[27,28]；在 MD 模擬的計算中，我們採用的時間間隔是 2 fs^[27,28]。對水分子的模型計算則爲 TIP3P 水模型^[29]。本篇的模擬使用 AMBER4.1 套裝軟體^[25]；使用的分子力場則爲 Cornell 等人所發展出來的力場^[30]。接下來描述、分析展開過程中結構的構形簇分析方法。

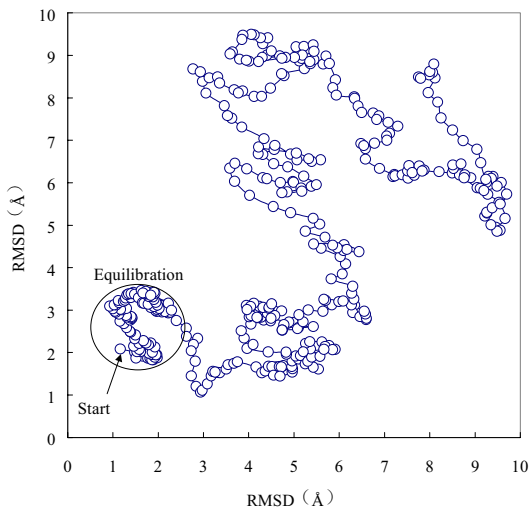
構形簇分析 爲了確認 CTX3 分子在展開過程中的過渡態（transition state），我們使用了構形簇分析方法^[20,24]。這個方法可以用來分辨結構的相似性。以下做一簡介：假設在 MD 的反折疊軌跡中，有 N 個 CTX3 結構資料，則在 N 個結構中，兩兩間可以有 $N(N-1)/2$ 個 RMSD 值，記爲 $\{r_{ij}\}_{N_{i,j=1}}$ 。再考慮一

個 3 度空間的 N 個點，同樣地，兩兩個點之間也會有 $N(N-1)/2$ 個距離，把這些點間距離的資料記爲 $\{d_{ij}\}_{N_{i,j=1}}$ 。我們選定適當的平面大小，考慮 r_{ij} 和 d_{ij} 間的差值，會得到一個方程式： $\delta = \sum_{N_{i=1}} \sum_{N_{j=1}} (r_{ij} - d_{ij})^2$ ，之後去找 δ 的最小值。尋找的方法如下：首先用亂數產生器在一個平面上打 N 個點，然後再找 δ 的最小值。 δ 最小值的大小會隨著開始所選用的起始亂數的位置有關。因此，本篇論文尋找 TS 的過程中，我們重複地選用了 100 個亂數起始值，亦即會有 100 個不同的 δ 最小值，我們再從這 100 個 δ 最小值中挑一個最小的值來作爲想要的 2-D 構形簇分析圖。CTX3 模擬之結果見圖三。

此類增溫的蛋白質展開模擬，在幾百 ps 到幾個 ns 的模擬中，急速將蛋白質展開，此表示在蛋白質的構形空間（Configuration Space）中模擬將取樣（sampling）較多“高溫”區域^[31]。先前此類計算結果得到的過渡態結構，與摺疊（folded）結構相去不遠^[32,33]，其他與過渡相關計算亦得到與實驗資料一致的結果，這些結果建議，此類增溫的蛋白質展開模擬至少在摺疊結構附近可得到良好的蛋白質展開模擬結果^[23]。



圖二、CTX3 MD 升溫模擬軌跡在 pH=7 下的 RMSD 曲線圖



圖三、CTX3 MD 升溫模擬軌跡在二維 C α RMSD 平面的構形簇分析圖。

蛋白質的摺疊時間 (folding time) 一般落在 1 ms 到 100 sec 間^[34]，一個含 25 胺基酸，天然結構為 α 螺旋的 peptide，由 random coil 形成 α -螺旋 (α -helix) 的摺疊時間約為 100 ns^[31]，形成 β -hairpin 的折疊時間約為 100 ns~1 μ s^[31]，依目前大部分模擬實驗室具有在 ns 時間區段的蛋白質模擬計算能力，模擬蛋白質摺疊仍是一個困難的問題，即使如此，幾個

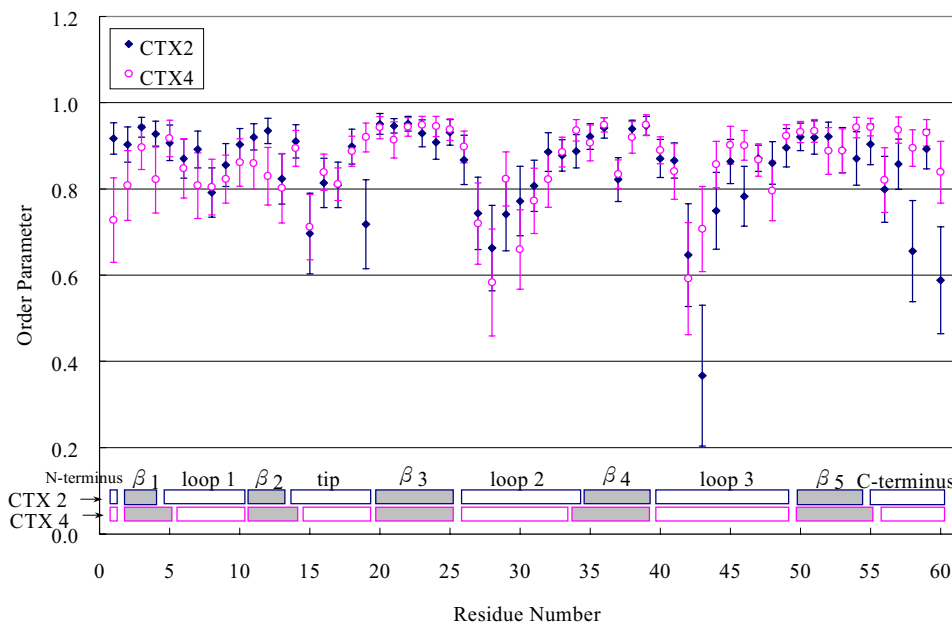
蛋白質摺疊的模擬，由高溫展開態開始，經過降溫過程，可檢視蛋白質摺疊過程中，疏水效應在摺疊過程中的角色^[35]。

蛋白質的主鏈骨架動力學 除了蛋白質結構及上述的蛋白質應用的探討外，蛋白質在水溶液中的動力學性質亦是一重要的課題^[36]。其中，蛋白質主鏈骨架的擾動程度反映蛋白質結構該部份的易曲性 (flexibility)；實驗上，藉由 NMR 光譜中檢視 ^1H 、 ^{13}C 及 ^{15}N 等原子間核自旋的耦合 (coupling) 強度及動力學等測驗，可檢視主鏈上 C α -H 或 N-H 鍵的擾動程度，此結構擾動程度可用一介於 0 與 1 之間的秩序參數值表示，實驗結果可得蛋白質中多個胺基酸的秩序參數值。在蛋白質模擬中，根據理論^[37,38]，可計算得胺基酸的秩序參數值，心臟毒蛋白 CTX2 實驗^[39]和計算^[40]及 CTX4 計算的結果如圖四，此模擬為室溫下的計算結果，模擬系統和程序與上述蛋白質展開類似，模擬時間長度為 600 ps。此 CTX4 與計算的一蛋白質 CTX2 的胺基酸序列 (sequence) 僅在第一個胺基酸不同，其餘均相同；CTX4 秩序參數計算結果顯示，此由於第一個胺基酸

的不同，沒有造成與 CTX2 計算結果明顯的差別。其中一個原因，乃模擬時間僅數百 ps，但蛋白質主鏈骨架動力學，在迴圈 (loop)，轉折 (turn) 部份的相關時間 (correlation time) 常較此長許多，以致模擬計算未能對這部份蛋白質結構得到準確結果。即使如此，目前幾百 ps 到幾十 ns 的計算結果仍提供蛋白質相當多的動力學資訊，並可作為探討蛋白質與受體 (substrate) 結合時，結合位置 (binding site) 的參考。

由上述兩個例子的討論可知，在蛋白質模擬中，取樣構形空間 (conformational space) 是檢視計

算結果的一重要課題。最近，一個從偏離折疊結構不遠的結構，當作起始結構的蛋白質折疊模擬，即使經過 1 μ s 的模擬，此一含 36 胺基酸的小蛋白質仍未折疊回折疊態的天然 (native) 結構^[21]；另外一個室溫下平衡的蛋白質模擬結果顯示^[41]，在 18 ns 的模擬軌跡中，即使經過 15 ns，蛋白質的 RMSD 質仍會有明顯變化，顯示在天然結構為起始結構的蛋白質模擬中，天然結構附近仍有許多能量相近之構形空間，在幾十個 ns 的模擬時間區段中，模擬中的取樣對探討大部分蛋白質的平衡及動力學性質



圖四、CTX2 & CTX4 平衡後 600 ps 分子模擬軌跡之秩序參數值 (Order Parameter) 比較。

仍是一個須仔細探討的問題。幾個減輕取樣問題的模擬方法已陸續被提出^[42,43]，相信在近期內對蛋白質模擬計算將會有明顯助益。

四、結語

近來生化系統模擬的發展，已從上述使用經驗

函數的分子動力學模擬進展到包含部份電子結構於計算中，以檢視生化系統中的化學反應，特別是蛋白酶的催化機制。一些模擬亦由本文所描述之古典模擬，進步至包含原子的量子效應一如氫原子量子效應，在生化系統中的角色。綜合言之，本文中敘述的生化系統模擬，可使吾人計算檢視生化分子系統的各種物理量至原子層次，除了上述蛋白質本身

性質的探討外，生化分子模擬亦已被廣泛運用於製藥、生物技術等產業中，它的發展將繼續提供更多更廣泛的理論分析，更深入了解生化系統的物理化學性質。

五、參考資料

1. C. Yu, R. Bhaskaran, and C. C. Yang., *J. Toxicol. Toxin Reviews.* 13, 291(1994) and references therein.
2. W. Wu., *J. Toxicol. Toxin Reviews.* 16, 115(1997) and references therein.
3. R. Bhaskaran, C. C. Huang, D. K. Chang, and C. Yu., *J. Mol. Biol.* 235, 1291(1994)
4. R. Bhaskaran, C. C. Huang, Y. C. Tsai, G. Jayaraman, D. K. Chang, and C. Yu., *J. Biol. Chem.* 269, 23500(1994)
5. A. K. Singhal, K. Y. Chien, W. Wu, and G. S. Rule., *Biochemistry.* 32, 8036(1993)
6. T. K. S. Kumar, C. S. Lee, and C. Yu., *Natural Toxins.* 115(1996)
7. C. Yu, R. Bhaskaran, L. C. Chuang, and C. C. Yang., *Biochemistry.* 32, 2131(1993)
8. R. Bhaskaran, C. C. Huang, Y. C. Tsai, G. Jayaraman, D. K. Chang, and C. Yu., *J. Biol. Chem.* 38, 23500(1994)
9. Y. J. Sun, W. Wu, C. M. Chiang, A. Y. Hsin, and C. D. Hsiao., *Biochemistry.* 36, 2403(1997)
10. J. Y. Jang, T. K. S. Kumar, G. Jayaraman, P. W. Yang, and C. Yu., *Biochemistry.* 36, 14635(1997)
11. T. Sivaraman, T. K. S. Kumar, P. W. Yang, and C. Yu., *Toxicon.* 35, 1367(1997)
12. J. F. Oconnell, P. E. Bougis, and K. Wuthrich., *Eur. J. Biochem.* 213, 891(1993)
13. T. K. S. Kumar, G. Jayaraman, C. S. Lee, A. I. Arunkumar, T. Sivaraman, D. Samuel, and C. Yu., *J. Biolmol. Struc. Dyna.* 15, 431(1997) and references therein.
14. C. M. Chiang, S. L. Chang, H. J. Lin, and W. Wu., *Biochemistry.* 35, 9177(1996)
15. C. C. Lo, J. H. Hsu, Y. C. Sheu, C. M. Chiang, W. G. Wu, W. Fann, and P. H. Tsao., *Biophys. J.* 75, 2382(1998)
16. T. Sivaraman, T. K. S. Kumar, and C. Yu., *Biochemistry.* 38, 9899(1999)
17. T. Sivaraman, T. K. S. Kumar, D. K. Chang, W. Y. Lin, and C. Yu., *J. Biol. Chem.* 273, 17, 10181(1998)
18. T. Sivaraman, T. K. S. Kumar, G. Jayarman, C. C. Han, and C. Yu., *Biochem. J.* 321, 457(1997)
19. P. A. Kollman, *Chem. Rev.* 93, 2395(1993)
20. V. Daggett and M. Levitt., *Annu. Rev. Biophys. Biomol. Struct.* 22, 353(1993) and references therein.
21. Y. Duan and P. A. Kollman., *Science.* 282, 740(1998) 此篇蛋白質模擬中，約有 10000 個原子（含水）於模擬系統中，乃於近年來，第一篇超過 1 μ s 模擬之報告。
22. A. Li and V. Daggett., *J. Mol. Biol.* 257, 412(1996)
23. K.F. Fulton, E.R.G. Main, V. Daggett, and S.E. Jackson., *J. Mol. Biol.* 291, 445(1999).
24. M. Levitt., *J. Mol. Biol.* 168, 621(1983)
25. D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, and P. A. Kollman., *Univ. of Cal., San Francisco, San Francisco, 1995.*
26. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak., *J. Chem. Phys.* 81, 3684(1984)

27. A.R. Leach. Addison Wesley Longman Limited, England, (1996).
28. M.P. Allen and D.J. Tildesley. Oxford Science Publications, Oxford, (1989).
29. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein., J. Chem. Phys. 79, 926(1983)
30. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Jr. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman., J. Am. Chem. Soc. 117, 5179(1995)
31. S. Doniach and P. Eastman., Curr. Opinion Strut. Biol. 2,157(1999).
32. A. Li and V. Daggett., Proc. Natl. Acad. Sci. USA. 91,10430(1994).
33. J.N. Onuchic, Z. Luthey-Schulten, and P.G. Wolynes. Ann. Rev. Phys. Chem. 48, 545(1997).
34. M. Karplus and D. L. Weaver., Protein Sci. 3, 650(1994).
35. D.O.V. Alonso and V. Daggett., Prot. Sci. 7, 860(1998).
36. A.G. Palmer, M. Rance, and P.E. Wright., J. Am. Chem. Soc. 113, 4371(1991).
37. G. Lipari and A. Szabo., J. Am. Chem. Soc. 104, 4546(1982).
38. G. Lipari and A. Szabo., J. Am. Chem. Soc. 104, 4559(1982).
39. C.S. Lee, T.K.S. Kumar, L.Y. Lian, J.W. Cheng, and C. Yu., Biochemistry. 37, 155(1998).
40. Y.C. Sun, S.F. Yang, I.L. Hwang, and T.H. Wu., J. Comput. Chem. 20, 546(1999).
41. D.C. Chatfield, A. Szabo, and B.R. Brooks., J. Am. Chem. Soc. 120, 5301(1998).
42. A. Kitao, S. Hayward, and N. Go., Proteins. 33, 496(1998).
43. T. Huber and W.F. Vangunsteren., J. Phys. Chem. A. 102, 5937(1998).