

淺談電腦輔助藥物設計

1 楊進木

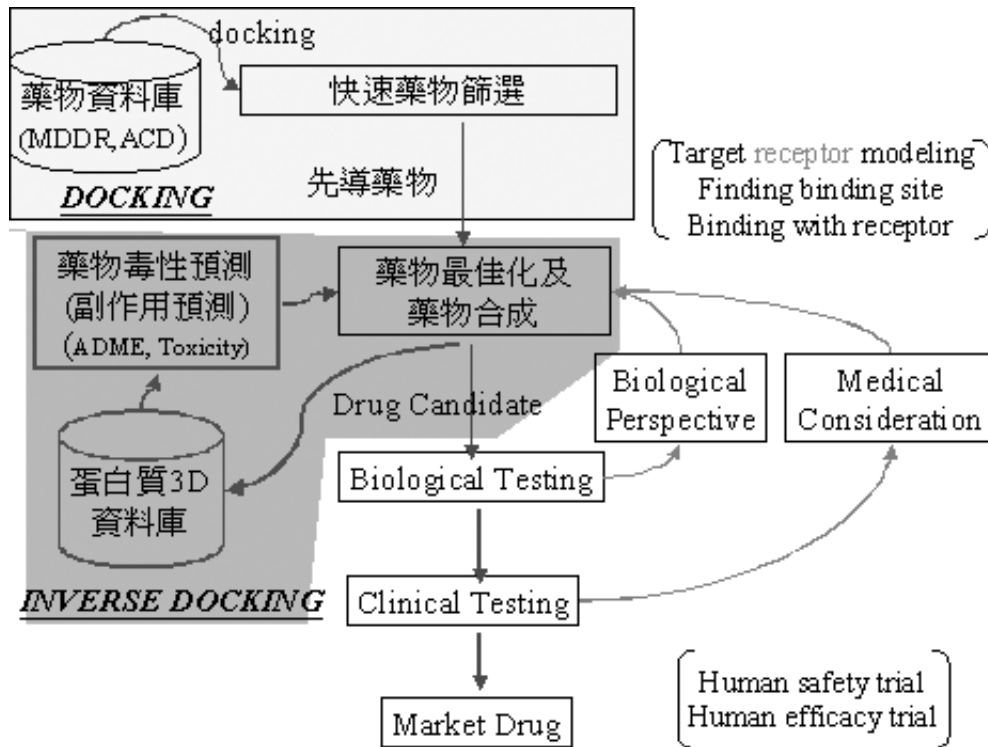
傳統開發藥物的方法常需耗費大量時間與金錢，發展一種新藥平均約需十二年，其成本約為 4.5 億美金（超過一百億台幣），如此耗時及高成本的發展方法，無法滿足人類對抗疾病所需新藥的需求。幸運地，因人類基因體計畫及電腦輔助藥物設計的發展，預期可縮短藥物開發時間及降低其開發成本，提供新藥開發的契機。

藥物開發的概要流程如圖一所示。電腦輔助藥物設計在藥物開發之核心技術包含：一、快速藥物篩選（**virtual drug screening**）、二、藥物最佳化、三、藥物的毒性及副作用預測（**ADMET: absorption, distribution, metabolism, elimination and toxicity**）。我們相信電腦輔助藥物設計可與傳統的藥物開發方式互補，組合成一個更完善的藥物開發體系，並極有可能壓縮臨床實驗所需的時間與經費。

本實驗室已成功建構一個藥物篩選系統（**GEMDOCK**：發表在國際期刊 **Proteins** 上^{1,2}），**GEMDOCK** 已實際測試於亞硫酸轉氨酶（**sulfotransferase**）、登革熱病毒（**Dengue virus**）NS3 水解酶及 **SARS coronavirus 3C-like protease** 之篩選模擬。

快速藥物篩選（**virtual drug screening**）之核心技術為 **protein-ligand docking**，此技術的中心概念在於藉由電腦高速運算的功能，由已知藥物之立體構形（**conformation**）的彈性配體（**flexible ligand**）與蛋白質巨分子兩者來預測其 **protein-ligand** 複合體（**complex**）的結構。近年來由於高解析度的蛋白質結構較早年易於取得，以及 **docking process** 可藉由電腦模擬（**computer-based simulation**）進行全自動化演算之故，**protein-ligand docking** 技術愈來愈受重視，應用範圍亦愈加廣泛。此技術可用於篩選或設計能與蛋白質上特定位置達成良好的結構上與化學上互補（**complementarity**）的 **ligand**，此即能幫助我們尋得具備藥物潛力的小分子物質。而過去的諸多研究顯示，**protein-ligand docking** 程式所

圖一、藥物開發流程概念圖



預測的 *protein-ligand complex* 構形實際上已相當接近實驗所得的 *complex* 真實結晶結構。

藥物最佳化之典型的方法就是 QSAR (*quantitative structure-activity relationship*) and QSPR (*quantitative structure-property relationship*)，QSAR 與 QSPR 建立在 1960 年代，利用統計方法來找出一群生化與物化性質已知之分子特性與結構敘述子 (*structure descriptors*) 之間的關係³。當一個 QSAR 模型被含有實驗資料的分子資料組訓練出實驗值與特性之間的關係之後，就能夠用來預測可能的潛在分子。在 QSAR 模型中使用適當的敘述子，可讓簡單的線性回歸關係被應用統計學的多變數分析技術來加以分析與利用。而許多資料探勘 (*data mining*) 與機器學習 (*machine learning*) 技術也成功的應用在 QSAR 的建立上，例如類神經網路 (*neural networks*)⁴，SOM (*self-organizing maps*) 以及 SVM (*support vector machines*)⁵。

早期預測藥物毒性對於藥物開發是非常重要的，因為可大幅縮減後續的生體實驗以及藥物毒性測試的時間及金錢開銷。過去傳統藥物發展過程中，經常使用實驗高速篩選(*in vitro high-throughput screening*)平台來尋找可能的候選藥物。這樣的過程不僅耗費大量金錢與時間，而且所找出的候選者(*candidates*)也無法應用實驗過程獲得的數據，輔助後續設計藥物結構的最佳化與評估。一九九〇年代的研究發現，許多藥物開發案例在最後階段的失敗常肇因於藥物動力特性不良與其毒性⁶。因此目前藥物發展的趨勢傾向利用電腦高速運算特性，加速藥物開發與篩選。

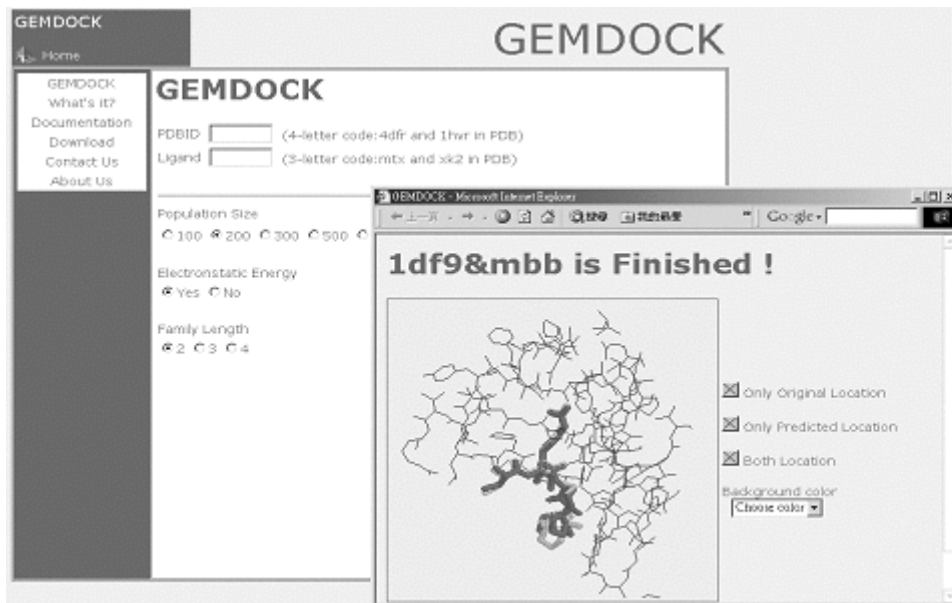
組合性化學(*combinatorial chemistry*)與電腦高速篩選(*in silico high-throughput screening*)技術⁷，在近幾年經常被應用到藥物開發的早期階段上及評估潛在藥物的 ADMET(*absorption, distribution, metabolism, elimination and toxicity*)特性是否符合，進一步提高與加速實驗高速篩選的效率與準確率⁶。電腦 ADMET 模擬技術目前朝向兩大發展方向建立：資料模擬與分子模擬。分子模擬包含蛋白質模擬，例如利用量子力學方法來考慮涉及 ADMET 過程的的蛋白質與小分子間的潛在作用力，例如 *cytochrome P450s*，但是這樣的方法需要利用已知蛋白質結構。當結構未知的時候需要以 *homology modeling* 來建立結構，或者是利用此一蛋白質已知的受質或抑制劑結構疊合，找出和蛋白質作用的特性來建立 *Pharmacophore model*⁵。*Protein-ligand inverse docking* 技術應用於藥物開發初期，*Protein-ligand inverse docking* 技術的觀念在於搜尋可與單一 *ligand* 結合的多個蛋白質，與 *protein-ligand docking* 著重由單一蛋白質尋找可與該蛋白質結合的多個 *ligands* 恰好相反，而依照該 *ligand* 可與哪些蛋白質標的(*protein target*)結合，結合力強或弱，我們即可依之推測該藥物可能產生的副作用或毒性。

整合快速藥物篩選(*virtual drug screening*)、藥物最佳化(QSAR)、藥物的毒性及副作用預測的工具是現代藥物發現的關鍵技術之一，此整合技術能應用於辨識潛在藥物，預測這些藥物的副作用以及毒性，因此在藥物發展的初期可作為低成本、快速度藥物測試的技術。我們相信電腦輔助藥物可與傳統的藥物開發方式互補，組合成一個更完善的藥物開發體系，下面簡述本實驗室在藥物開發領域中研究獲得的初步成果：

一． 建構完成通用的電腦輔助藥物快速篩選系統，簡稱為

GEMDOCK^{1,2,8}，其網頁位置為 <http://gemdock.life.nctu.edu.tw>，此系統具更快且精確描述 **protein-ligand interaction** 能量變化的計分函

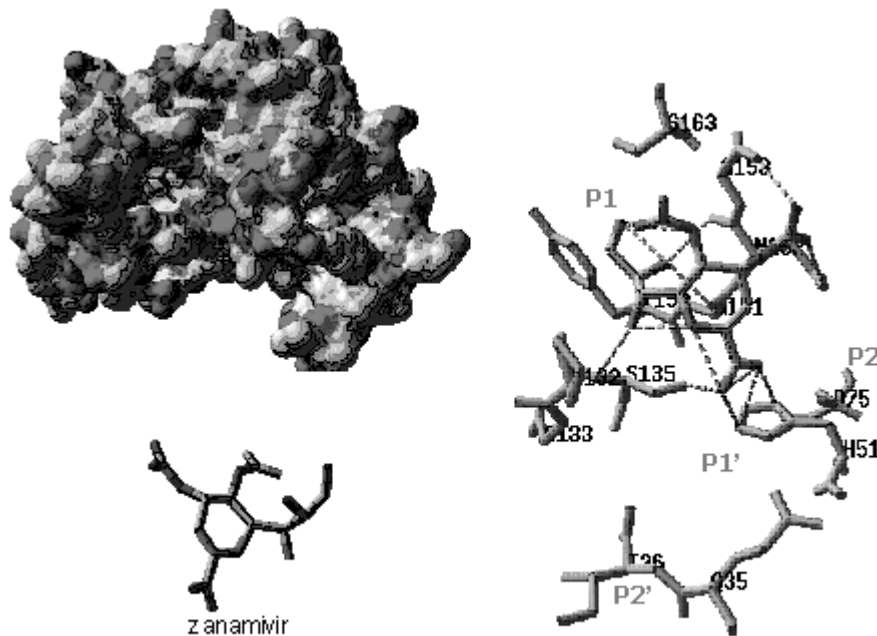
圖二、GEMDOCK web based 遠端服務系統圖示



式。GEMDOCK 可應用於各種藥物開發平台的前端開發，如下圖二所示，綠色為原來 **ligand** 位置，紅色為預測之 **ligand** 位置，此系統可提供 **web based** 的遠端服務。（編按：本刊因灰階印刷，無法顯示顏色，特此致歉。）

二. 已將第一項之工具實際用於發展登革熱蛋白水解酶 NS3 (PDB Code: 1BEF, 1DF9) 抑制劑快速之篩選，此系統會針對這類蛋白水解酶的重要胺基酸，如活性區域中的 Ser 135、Asp 75、His 51 (catalytic triad) 等，這些胺基酸會與潛在的抑制劑間形成作用力，如靜電力、氫鍵 (H) 及凡得瓦爾力 (vW) 等，系統會自動偵測並加重其權重，以提昇藥物快速篩選系統的專一性及正確性。我們以 CMC (Comprehensive medicinal chemistry: about 7,000 Compounds) 為藥物篩選的資料庫，分別針對 1BEF 及 1DF9 的活性區域進行嵌合預測。初步篩選出 23 個 Compounds 對蛋白水解酶 NS3 的潛在抑制劑，我們以 zanamivir 為例說明 Docking 預測結果，zanamivir 本來用作為治療感冒之藥物。圖三是

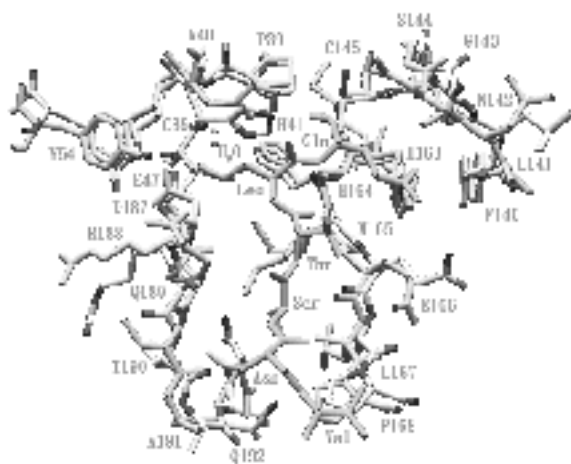
圖三、zanamivir 對革熱蛋白水解酶 1BEF 所做的嵌合預測結果



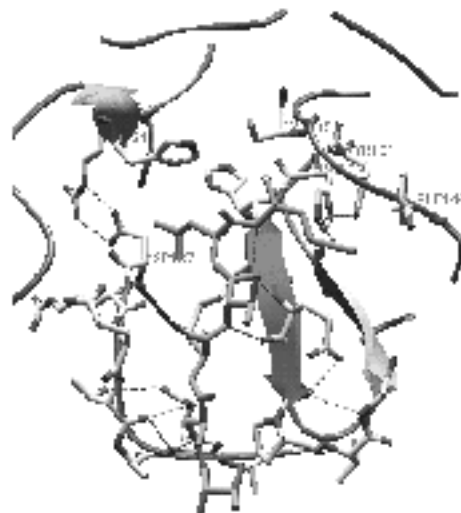
化合物 zanamivir 對 1BEF 所做的嵌合預測結果，zanamivir 穩定地結合在 1BEF 活性中心的位置上，且與重要的胺基酸均有形成氫鍵。

三. 我們也將第一項之工具實際用於發展 SARS coronavirus 3C-like protease 抑制劑之篩選。已初步成功證實本實驗室發展之 docking 工具 GEMDOCK 可應用於 SARS coronavirus 同一 Family 之 3C-like proteases (1P9U, 1P9S) 之 docking 模擬。並成功以 Homology Model 及分子模擬方法 modeling SARS 3CL^{pro}。此蛋白質模擬結構已登錄在 PDB，其 code 為 1Q1X，經與嗣後發表的結晶結構比較後證實兩者甚為相似（如圖四所示），此或可作為本實驗室結構模擬策略正確之一佐證。我們已利用 3CL^{pro} 切割蛋白鏈後的殘剩 peptide fragment (VNSTLQ 與 TSAVLQ) 及 AG7088（用於治療 rhinovirus 引起的感冒症狀，已進入臨床實驗）作為基礎，其 docking 結果如圖五所示，印證我們工具之正確性。

圖四及圖五、本實驗室電腦模擬之蛋白質結構(1Q1X)與結晶結構(1Q2W)之比較示意圖。經比較後發現兩者在 binding site 區域甚為相



圖四



圖五

似，CPK 著色部分為 1Q2W，淺藍色者為 1Q1X，peptide VNSTLQ 標示為綠色。圖五、電腦模擬預測之 Docking Results。（編按：本刊因灰階印刷，無法顯示顏色，特此致歉。）

參考文獻

1. Yang J-M, Chen C-C. GEMDOCK: A generic evolutionary method for molecular docking. to appear in *Proteins: Structure, Function and Genetics*.
2. Yang JM. An evolutionary approach for molecular docking. *Lecture Notes in Computer Science* 2003;2724:2372-2383.
3. Todeschini R, Consonni V. *Handbook of Molecular Descriptors*. Weinheim: Wiley-VCH; 2000.
4. So SS, Karplus M. Three-dimensional quantitative structure-activity relationships from molecular similarity matrices and genetic neural networks. 2. Applications. *Journal of Medicinal Chemistry*

1997;40(26):4360-4371.

5. Hodgson J. ADMET-turning chemicals into drugs. *Nature biotechnology* 2001;19:722-726.

6. Waterbeemd H, Gifford E. ADMET in silico modeling: towards prediction paradise. *Nature reviews drug discovery* 2003;2:192-204.

7. Dearden JC. In silico prediction of drug toxicity. *Journal of Computer-Aided Molecular Design* 2003;17:119-127.

8. Lin ES, Yang J.-M., Yang YS. Modeling the binding and inhibition mechanism of nucleotide and sulfotransferase using molecular docking. *Journal of the Chinese Chemical Society* 2003.



楊進木教授小檔案：

學歷：台灣大學資訊工程博士
中央大學資訊工程碩士
淡江大學管理科學學士

現職：交通大學生物科技系暨生物資訊所助理教授

專長：生物資訊

榮譽：-Win the prize of 2001 Best M.S. and
Ph. D. Dissertation Award hosted by IICM
-2003 Marquis's Who's Who
in Science and Engineering