

國立臺灣大學工學院化學工程學系

碩士論文

Department of Chemical Engineering

College of Engineering

National Taiwan University

Master Thesis



探討單點突變與分子對接模擬對人類 γ D型水晶體蛋白不穩定
區域之影響

Exploring the Influence of Unstable Regions of Human
 γ D-Crystallin by Single-Point Mutations and Molecular Docking

羅峻賢

Chun-Hsien Lo

指導教授：王勝仕 博士

吳宛儒 博士

Advisor: Sheng-Shih Wang, Ph.D.

Josephine W. Wu, O.D., Ph.D.

中華民國 105 年 7 月

July 2016

誌謝



「回首向來蕭瑟處，歸去，也無風雨也無晴。」兩個年頭，將近 70 個學分以及一本碩士論文，總算完成了！在此必須感謝這兩年中，陪伴我走過這些日子的老師、朋友以及家人。有你們的支持，才有現在的我。

首先最要感謝的一定是恩師 王勝仕老師，除了在研究方面給予我相當多的指導之外，同時也教會我許多待人處事的道理，讓我未來在社會可以更圓融地處理未知的的事情。另外，感謝 吳宛儒博士對於我研究內容細心地指導，常常能找到我邏輯論述的盲點並即時指正我，讓我真的受益良多。除此之外，也要感謝口試委員賴進此博士、林達顯教授、侯劭毅教授及廖思婷博士，在席不暇暖之餘還要風塵僕僕地趕來給予我寶貴的建議，讓我的碩士論文可以蓬蓽生輝。

回想兩年半前剛進入 Bio308 時，我還只是一隻不會思考的草履蟲，感謝當時帶領我開始學習實驗的偉安學長，在學長的指導下，總能讓我對實驗的原理與設計有一番新的見解，也開啟我碩士研究生涯的第一扇門。感謝偉大的博班素素學姐，不管我在實驗上有任何的疑難雜症或是在實驗室有天大的事情，遇到學姐都能夠迎刃而解，並時時耳提面命提醒我所有該注意的事情。感謝引領我進入模擬領域的智凱學長，在研究的過程中總是不厭其煩地回答我的問題，甚至在畢業後仍常關心我的模擬進度，令我倍感窩心；而且隨著越深入的研究，越能感受到自己原來真的是站在巨人的肩膀上，許多看似簡單的問題，在過去都是費盡千辛萬苦才得出答案的。

感謝一起度過這兩年的欣怡，有你在的實驗室總是充滿個奇怪的話題；也感謝其他在實驗室共同努力的大家們，包括怡潔學姊、耀宗學長、博裕學長、仁偉學長、瑀鴻學姐、辛瓊、斯閔、旻芝、書舜、牧月、郭董、貢丸、翊誠、泓曄、宜琳、老戴、冠宇、郁涵、Emily、念臻、郡蘋、繳倫、綺貞、張穎、威成、柏竹等，感謝你們在實驗室的相伴。我們承接著 Bio308 的過去，也延續著 Bio308 的未來。



另外，我還要感謝兩年來教過我的老師們，每當我在課堂上學到新知識時，我總在想這些東西是不是也能用在這本論文裡，然後常常害我又少了兩個小時的睡眠時間…。感謝以下老師：徐振哲老師—高等化工動力學、藍崇文老師—高等化工應用數學一、陳平老師—生物化學、陳伊琳老師—教育哲學、張文章/徐駿森老師—蛋白質結構與功能、林英智老師—無機化學、鄭光成老師—食品發酵學、雷庚玲老師—青少年心理學、佘瑞琳老師—化學科教材教法&教學實習、陳俊顯老師—分析化學乙下&實驗。

最後，我要感謝在我求學過程中，總是無條件給我支持的家人們，有你們當我的後盾，讓我更能心無旁騖地在我的學業上。謝謝你們！

中文摘要



白內障是一種與蛋白質聚集相關的眼疾，也是導致失明的主要原因，而引發白內障的主要原因包括受日光的照射、不正常的飲食、藥物使用或是糖尿病等，其機制主要是來自於水晶體蛋白的聚集，因此為了能有效避免這種眼睛相關疾病，了解水晶體蛋白的結構穩定性以避免聚集的形成是非常重要的。人類 γ D 型水晶體蛋白(Human γ D-crystallin, H γ DC)是一種由 173 個殘基組成的蛋白質，其在人類水晶體核中的含量極為豐富，也是與老化型白內障有高度相關的水晶體蛋白之一。

過去研究指出，H γ DC 在殘基 115-118 的位置為一不穩定的區域，在本研究中，吾人利用多種生物資訊工具，就分子層面來預測 H γ DC 與其不同突變種的結構穩定性與聚集傾向，並利用一致性排序的方式挑選出每個殘基上結構最穩定與最不穩定的突變種，接著使用分子動力學模擬來印證。根據 RMSD、原態接觸分析以及聚集傾向的預測，吾人發現所有的突變種中，R116I 和 F117A 分別擁有最高和最低的聚集傾向，然而其結構穩定性都與原態沒有太大的差異。接著，吾人更進一步發現聚集傾向與其不穩定區域 115-118 的擾動情形與疏水性有相當程度的關聯。

除此之外，吾人利用分子對接的方式嘗試找出最有可能避免 H γ DC 聚集形成的抑制劑。在分析 H γ DC 與多種小分子間的交互作用力後，因為迷迭香酸(rosmarinic acid)與殘基 115-118 能產生最多的作用力，吾人推測迷迭香酸應為本研究中最具有治療白內障的潛力。吾人相信此研究結果可以幫助研究者更了解 H γ DC 的結構完整性，並發展未來能治療白內障的可能途徑。

關鍵字：白內障、人類 γ D 型水晶體蛋白、分子動力學模擬、分子對接、蛋白質聚集、突變

Abstract

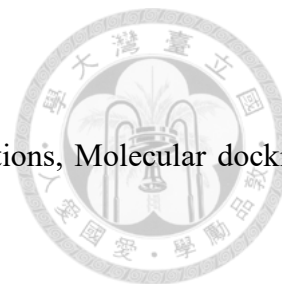


Cataract, an eye disease related to protein aggregation, is the leading cause of blindness worldwide. Factors that cause cataract include exposure of daylight, abnormal diet, drug intake, and diabetes. The underlying mechanism that leads to cataract is related to the aggregation of crystallin proteins. Therefore, understanding the structural stability of crystallin proteins to avoid aggregation is of foremost important for the prevention of this sight-threatening disease. Human γ D-crystallin (H γ DC), a 173-residue protein, is abundant in the nucleus of human eye lens and is one of the major crystallin proteins involved in the age-related nuclear cataract.

Previous studies showed the unstable regions of H γ DC are residues 115-118. In the first part of this thesis, the structural stabilities and aggregation propensity of various H γ DC mutants from the unstable regions were investigated at the molecular level by multiple bioinformatics tools. The most stable and unstable mutants were chosen through a consensus ranking method for further verification by molecular dynamics simulations (MDS). By combining the results of RMSD, native contact fraction, and prediction of aggregation propensity, we found that R116I and F117A have the highest and lowest aggregation propensity among the mutants studied, respectively; however, their structural stabilities were almost the same as the wildtype. Moreover, the degree of aggregation propensity was related to the fluctuation and hydrophobicity in the residues 115-118.

In the second part of this thesis, attempts were made to seek the possible inhibitors to prevent the H γ DC from forming aggregates by molecular docking. According to the interaction analysis between H γ DC and small molecules, we speculated that rosmarinic acid was the best inhibitory molecule in our study because of its binding to the residues 115-118 of H γ DC. We believe the outcome from this research may help to achieve a better understanding of the structural integrity of H γ DC and pave way for future work in the development of therapeutic strategies against cataract.

Key words: Cataract, Human γ D-crystallin, Molecular dynamics simulations, Molecular docking, Protein aggregation, Mutant

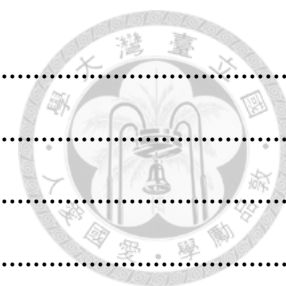


目錄



誌謝.....	I
中文摘要.....	III
Abstract.....	IV
目錄.....	VI
圖目錄.....	IX
表目錄.....	XII
第一章 緒論.....	1
1.1 研究動機.....	1
1.2 章節概述.....	2
第二章 文獻回顧.....	3
2.1 白內障(Cataract).....	3
2.1.1 白內障成因.....	3
2.1.2 白內障種類.....	6
2.1.3 白內障治療方法.....	10
2.2 水晶體蛋白(Crystallin).....	11
2.2.1 α 型水晶體蛋白.....	12
2.2.2 $\beta\gamma$ 型水晶體蛋白.....	13
2.2.3 人類 γ D 型水晶體蛋白(Human γ D crystalline, H γ DC).....	15
2.3 電腦模擬(Computer Simulation).....	24
2.3.1 生物資訊工具(Bioinformatics tools).....	25
2.3.2 分子動力學模擬(Molecular Dynamics Simulations).....	31
2.3.3 分子對接(Molecular Docking).....	36
2.4 模擬分析方法原理介紹.....	42
2.4.1 方均根變異(Root mean square deviation, RMSD)分析.....	42
2.4.2 方均根擾動(Root mean square fluctuation, RMSF)分析.....	42
2.4.3 原態接觸(Native contact)率分析.....	43

2.4.4	蛋白質二級結構分析(Dictionary of protein secondary structure, DSSP)	43
第三章	實驗儀器與步驟.....	45
3.1	模擬系統與軟體.....	45
3.1.1	模擬系統.....	45
3.1.2	生物資訊預測工具.....	45
3.1.3	模擬軟體.....	46
3.2	生物資訊預測工具(Bioinformatics prediction tools)	47
3.2.1	Istable[141]	47
3.2.2	DUET[142]	48
3.2.3	Eris[143-145]	49
3.2.4	EASE-MM[146, 147]	51
3.2.5	PASTA2.0[148]	52
3.2.6	TANGO[149-151]	52
3.2.7	AGGRESCAN[152].....	53
3.2.8	WALTZ[153, 154].....	54
3.2.9	FoldAmyloid[155]	55
3.2.10	Zygggregator[156-159].....	56
3.2.11	AGGRESCAN3D	57
3.2.12	Discovery Studio 2016 (Calculate Aggregation Scores).....	58
3.3	分子動力學模擬及數據分析.....	60
3.3.1	H _γ DC 原態結構取得	60
3.3.2	Gromacs 模擬流程	60
3.3.3	Root Mean Square Deviation (RMSD)分析	61
3.3.4	Root Mean Square Fluctuation (RMSF)分析	61
3.3.5	Native Contact 分析.....	62
3.3.6	氫鍵作用力分析.....	63
3.3.7	距離分析.....	63
3.3.8	計算溶劑接觸面積(solvent accessible surface area)	63
3.3.9	DSSP 二級結構預測分析	63
3.4	蛋白質－小分子對接模擬(Protein-Ligand docking simulations)	64
3.4.1	CDOCKER	64



3.4.2	分子對接作用力分析.....	66
3.4.3	分子對接動力學模擬.....	66
第四章	實驗結果與討論.....	68
4.1	預測 HyDC 及其突變種的穩定性與聚集傾向	68
4.1.1	HyDC 及其突變種的穩定性預測結果	68
4.1.2	HyDC 及其突變種的聚集傾向預測結果	71
4.2	探討 HyDC 及其突變種於動力學下的結構變化	73
4.2.1	RMSD 分析結果	73
4.2.2	原態接觸(Native contact)分析結果.....	75
4.2.3	聚集傾向(Aggregation Propensity)分析結果.....	76
4.2.4	RMSF 分析結果	81
4.2.5	二級結構分析結果.....	82
4.2.6	親疏水程度分析結果.....	84
4.3	探討小分子抑制 HyDC 聚集之能力與機制	87
4.3.1	分子對接結果.....	87
4.3.2	質心距離與最短距離分析結果.....	91
4.3.3	RMSD 分析結果	92
4.3.4	凡德瓦力與氫鍵作用力分析結果.....	93
4.3.5	RMSF 分析結果	95
4.3.6	對接作用力分析結果.....	96
第五章	結論與建議.....	103
5.1	結論.....	103
5.2	建議.....	105
Reference		106

圖目錄



圖 2.1.1 正常水晶體(左)與混濁水晶體(右)折射入射光之示意圖	4
圖 2.1.2 天冬醯胺(Asn)經脫醯胺反應形成天門冬胺酸(Asp)之機制圖[15].....	5
圖 2.1.3 不同型態白內障示意圖[12].....	6
圖 2.1.4 白內障手術更換人工水晶體之示意圖.....	10
圖 2.2.1 人類 α B 型水晶體蛋白結晶結構圖(A)單體(pdb code: 2YGD); (B) 24 聚體(24-meric, pdb code: 2YGD)[52]	12
圖 2.2.2 α 型水晶體蛋白同時具有抑制與引發疾病的示意圖[56].....	13
圖 2.2.3 $\beta\gamma$ 型水晶體蛋白結構示意圖(A)平面圖 (B) 3D 立體圖(以 γ D 型為例, pdb code: 1hk0)[57, 58].....	14
圖 2.2.4 HyDC 去摺疊與聚集的可能途徑[82].....	17
圖 2.2.5 HyDC 各殘基在原態與 1M 胍鹽酸環境下的 RDC 值.....	22
圖 2.2.6 HyDC 各殘基二級結構於模擬過程中的存在百分率[123].....	23
圖 2.3.1 電腦運算能力進展的趨勢圖.....	25
圖 2.3.2 分子動力學模擬演算流程圖.....	32
圖 2.3.3 量子/分子模擬示意圖	33
圖 2.3.4 Gromacs 模擬流程圖	34
圖 2.3.5 Gromacs 模擬流程示意圖(A)建立 $2*2*2\text{ nm}^3$ 的盒子(B)在盒中加入溶劑模型(C)在盒中加入離子(藍:鈉離子,綠:氯離子).....	35
圖 2.3.6 週期性邊界條件示意圖[200].....	35
圖 2.3.7 分子對接模型示意圖.....	36
圖 3.2.1 Istable 界面圖	47
圖 3.2.2 Istable 選擇突變胺基酸頁面示意圖	47
圖 3.2.3 Istable 結果示意圖	48
圖 3.2.4 DUET 選單頁面圖	48
圖 3.2.5 DUET 結果示意圖	49
圖 3.2.6 Eris 首頁圖	49
圖 3.2.7 Eris 選擇突變胺基酸頁面示意圖	50
圖 3.2.8 Eris 結果示意圖	50
圖 3.2.9 EASE-MM Protein 選單頁面圖.....	51
圖 3.2.10 EASE-MM 結果示意圖	51
圖 3.2.11 PASTA2.0 首頁示意圖	52
圖 3.2.12 TANGO 輸入頁面圖	53
圖 3.2.13 TANGO 結果示意圖	53
圖 3.2.14 AGGRESCAN 首頁示意圖	54
圖 3.2.15 AGGRESCAN 計算結果示意圖	54
圖 3.2.16 Waltz 首頁示意圖	55

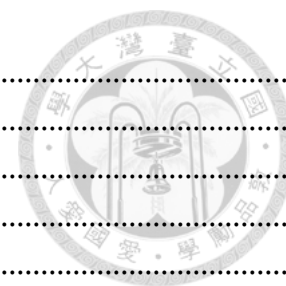


圖 3.2.17 WALTZ 結果示意圖	55
圖 3.2.18 FoldAmyloid 首頁示意圖	56
圖 3.2.19 FoldAmyloid 結果示意圖	56
圖 3.2.20 Zygggregator 頁面圖	57
圖 3.2.21 Zygggregator 結果示意圖	57
圖 3.2.22 AGGRESCAN3D 首頁示意圖	58
圖 3.2.23 AGGRESCAN3D 結果示意圖	58
圖 3.2.24 DS 預測聚集程度結果圖	59
圖 3.4.1 蛋白質與小分子結構示意圖	64
圖 3.4.2 蛋白質可能發生鍵結的區域示意圖	65
圖 3.4.3 CDOCKER 結果示意圖	65
圖 3.4.4 分子對接作用力分析圖	66
圖 3.4.5 PRODRG 執行頁面圖	67
圖 3.4.6 小分子力場檔示意圖	67
圖 4.1.1 以不同生物資訊工具預測 HyDC115-118 所有可能突變之能量結果(A)Istable (B)DUET (C)Eris (D)EASE-MM	69
圖 4.1.2 蛋白質多種形態之能量漏斗圖[277]	72
圖 4.2.1 HyDC 與其突變種在 343K 下 RMSD 值隨時間變化圖	74
圖 4.2.2 HyDC 與其突變種在 425K 下 RMSD 值隨時間變化圖	74
圖 4.2.3 HyDC 與其突變種在 343K 下原態接觸分率隨時間變化圖	75
圖 4.2.4 HyDC 與其突變種在 425K 下原態接觸分率隨時間變化圖	76
圖 4.2.5 以 AGGRESCAN3D 計算 HyDC 與其突變種在 343K 下之聚集分數隨時間變化圖	77
圖 4.2.6 以 AGGRESCAN3D 計算 HyDC 與其突變種在 425K 下之聚集分數隨時間變化圖	78
圖 4.2.7 以 DS 2016 計算 HyDC 與其突變種在 343K 下之聚集分數隨時間變化圖	79
圖 4.2.8 以 DS 2016 計算 HyDC 與其突變種在 425K 下之聚集分數隨時間變化圖	79
圖 4.2.9 HyDC 與其突變種在 343K 下 RMSF 值結果圖	81
圖 4.2.10 HyDC 與其突變種在 425K 下 RMSF 值結果圖	82
圖 4.2.11 HyDC 在 343K 下各殘基二級結構於模擬過程中之存在百分率	83
圖 4.2.12 HyDC 在 425K 下各殘基二級結構於模擬過程中之存在百分率	83
圖 4.2.13 HyDC 與其突變種之殘基 115-118 在 343K 下疏水面積隨時間變化圖	84
圖 4.2.14 HyDC 與其突變種之殘基 115-118 在 425K 下疏水面積隨時間變化圖	85
圖 4.2.15 胺基酸與對應之基因密碼子(codon)一覽	86
圖 4.3.1 HyDC 與小分子對接之位置示意圖	87
圖 4.3.2 HyDC 與小分子對接作用力分析結果(A)薑黃素(Curcumin) (B)兒茶素(Epigallocatechin gallate) (C)楊梅素(Myricetin) (D)槲皮素(Quercetin) (E)迷迭香酸(Rosmarinic acid)	90
圖 4.3.3 HyDC 與小分子的質心距離隨時間變化圖	91
圖 4.3.4 HyDC 與小分子的最短距離隨時間變化圖	92
圖 4.3.5 HyDC 與小分子對接之 RMSD 值隨時間變化圖	93

圖 4.3.6 H γ DC 與小分子對接之凡德瓦力隨時間變化圖	94
圖 4.3.7 H γ DC 與小分子對接之 RMSF 值結果圖	95
圖 4.3.8 H γ DC 與薑黃素(Curcumin)交互作用力之分析圖(A)全圖 (B)局部放大圖	97
圖 4.3.9 H γ DC 與兒茶素(Epigallocatechin gallate)交互作用力之分析圖	98
圖 4.3.10 H γ DC 與楊梅素(Myricetin)交互作用力之分析圖(A)全圖 (B)局部放大圖	99
圖 4.3.11 H γ DC 與槲皮素(Quercetin)交互作用力之分析圖(A)全圖 (B)局部放大圖	100
圖 4.3.12 H γ DC 與迷迭香酸(Rosmarinic acid)交互作用力之分析圖	101

表目錄



表 2.1.1 不同地區對白內障之橫斷研究的綜合分析[32]	7
表 2.1.2 白內障潛在風險因素相關研究	9
表 2.2.1 水晶體蛋白分類一覽[47]	11
表 2.2.2 近年與 HyDC 相關之研究	18
表 2.3.1 FoldAmyloid 演算法中 20 個胺基酸的分數[155]	28
表 2.3.2 多種生物資訊工具整理一覽表	30
表 2.3.3 常見的對接模擬軟體	37
表 2.3.4 抑制蛋白質聚集之小分子一覽表	39
表 2.4.1 DSSP 二級結構定義相關資料[176]	43
表 4.1.1 綜合分析不同生物資訊工具預測突變能量變化之結果	70
表 4.1.2 以多種生物資訊工具預測 HyDC 的聚集傾向一覽表	71
表 4.2.1 HyDC 突變種之結構穩定性與聚集傾向與原態比較一覽表	80
表 4.3.1 HyDC 與小分子對接能量一覽	88
表 4.3.2 HyDC 與小分子在模擬過程中出現氫鍵之平均次數	94
表 4.3.3 對接作用力計算標準一覽表	96
表 4.3.4 HyDC 重原子編號與殘基對照表	96

第一章 緒論



1.1 研究動機

白內障(Cataract)是一種人類水晶體發生混濁的疾病，由於混濁的水晶體會減弱與散射進入眼睛的光線，進而導致光線無法聚焦於視網膜上造成失明。根據歷史記載，白內障的發現與診斷已經有一千多年，雖然到目前為止，以手術更換人工水晶體來治療白內障的技術已經相當成熟，但對於白內障形成的機制以及如何預防仍有許多不清楚的部分。

許多臨床研究指出，混濁的水晶體主要源於水晶體蛋白(crystallin)的聚集(aggregate)，而引發其聚集的原因有多種可能，包括因天生基因缺陷所造成的遺傳性白內障，或是因長期暴露在紫外光照射而形成的老化型白內障等。針對不同的成因，近年來已經有許多研究致力於水晶體蛋白的聚集途徑，以及具有潛力的治療藥物進行探討。

為了更了解水晶體蛋白在白內障形成的過程中扮演的角色，吾人使用電腦模擬的方式進行分子層級的研究。在本研究中，吾人使用單點突變改變水晶體蛋白的結構，並以生物資訊工具預測水晶體蛋白的穩定性以及聚集傾向，後以分子動力學模擬蛋白質結構的改變。除此之外，更以分子對接模擬找出具有抑制效果的小分子，並探討其作用的機制。希望藉由此微觀的研究，使我們對於水晶體蛋白有更深入的認識，同時期待能對於未來預防白內障形成相關的研究有更進一步的幫助。

1.2 章節概述

本論文的第一章旨在說明本研究的動機與對章節進行概述；第二章是對近年來的相關文獻做簡單介紹，包括白內障、水晶體蛋白以及電腦模擬的原理等；第三章則是彙整本研究所使用的材料與方法，包括生物資訊工具、分子動力學模擬與分子對接模擬等；第四章則是論述本研究的結果，分為三個小節討論之；最後於第五章總結本研究的結果。

第二章 文獻回顧



2.1 白內障(Cataract)

白內障是一種會導致視障或失明的疾病[1]，其症狀包括色弱、視力模糊、視野周圍有光暈、對於明亮的光或是在夜間視力都會降低等，這可能會造成駕駛、閱讀或是面孔辨識的問題[2]，且視力不良也會增加得到憂鬱症的機會[3]。根據世界衛生組織(WHO)的調查結果發現，約有 51 % 的失明與 33 % 的視障是由白內障所造成，且在這些視力受損的案例中，有非常高的比例發生在年齡大於 50 歲的人身上[4]。隨著人口結構的老化，得到白內障的人數與日俱增，也因此有越來越多的研究投入於此。在本節將介紹白內障的成因、種類以及目前較主流的治療方法。

2.1.1 白內障成因

白內障發生位置在水晶體(Lens)中，其是眼球中主要的屈光(refractive)結構，也是唯一有調節能力的屈光間質[5, 6]，呈現雙凸形扁圓體。人類可以透過調整水晶體的曲率(curvature)，對於距離不同的物體進行對焦(focus)，使入射光順利匯於視網膜(retina)上[5]。一般而言，人類水晶體的折射率約在 1.386~1.406 間[7]，如此的梯度折射率(index gradient)有助於提升影像的解析度，並有效減少短距物體和長距物體的像差(aberration)[8]。水晶體是由高濃度的水晶體蛋白(crystallin)溶液所組成，提供了良好的折射率與穿透度(transparency)[9, 10]，而當水晶體蛋白產生聚集(aggregation)造成水晶體混濁時，會導致入射光無法匯於視網膜上，造成視力的障礙，即稱為白內障[11]，見示意圖圖 2.1.1。

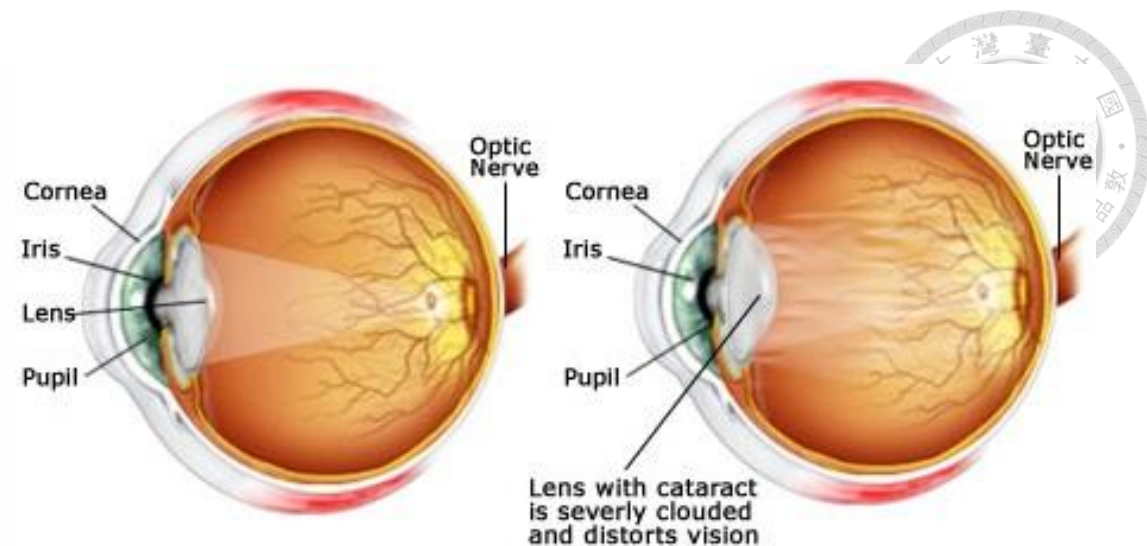


圖 2.1.1 正常水晶體(左)與混濁水晶體(右)折射入射光之示意圖

<http://www.improveeyesighthq.com/cataracts.html>

研究顯示，造成水晶體產生混濁，進而發生白內障的原因主要分為四種，包括先天型 (congenital)、老化型 (aged-related)、代謝型 (metabolic) 以及外傷型 (traumatic)[12]。

先天型白內障主要是由於先天基因的缺陷，造成水晶體蛋白突變，導致其聚集或沉澱。在 Santana, A. 等人的研究指出，當編碼人類水晶體中細胞質蛋白 (cytoplasmic protein) 的基因 (gene) 發生突變時，與許多不同型態 (morphologies) 的白內障有高度相關，包括基因編碼的水晶體蛋白 (genes encoding crystallins, CRYA, CRYB, and CRYG)、特定的水晶體連接蛋白 (lens specific connexins, Cx43, Cx46, and Cx50)、主要的內有蛋白 (major intrinsic protein, MIP)、水通道蛋白 (aquaporin)、細胞骨架結構蛋白 (cytoskeletal structural proteins)、配對同源轉錄因子 3 (paired-like homeodomain transcription factor 3, PITX3) 以及熱休克轉錄因子 4 (heat shock transcription factor 4) 等[13]。

老化型白內障是最大宗的白內障成因[12]，顧名思義即為水晶體因為老化而發生結構的改變，進而形成白內障。在人體中，脫醯胺 (deamidation) 反應是常見的自發性 (spontaneous) 非酵素分解蛋白質反應[14]，其反應會將蛋白質上天冬醯胺 (asparagine, Asn) 以及麩醯胺酸 (glutamine, Gln) 脫去胺基，形成天門冬胺酸 (aspartic

acid, Asp)與穀氨酸(glutamic acid, Glu)，此反應機制通常會伴隨著鄰近的甘氨酸(glycine)，如圖 2.1.2[15]。脫醯胺反應隨著老化，對於蛋白質結構與生理活性(biological activity)有非常重要的影響[15]，經 Ray, N. J.等人研究指出，人類水晶體蛋白會因脫醯胺反應促使二聚體(dimer)的形成[16]。除了年齡增長之外，仍有許多罹患白內障的潛在風險，包括日光的照射、食品與藥品的攝入(intake)或特定疾病的影響，如糖尿病(diabetes)等[17]。

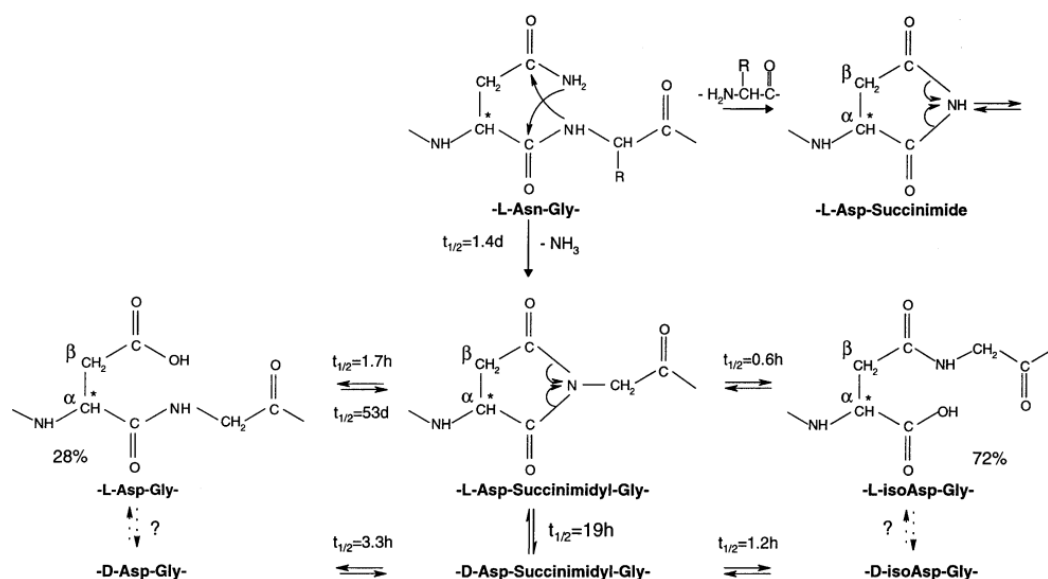


圖 2.1.2 天冬醯胺(Asn)經脫醯胺反應形成天門冬胺酸(Asp)之機制圖[15]

代謝型白內障是伴隨著代謝性疾病而來的，主要是由糖尿病[18]、半乳糖血症(galactosemia)[19, 20]以及半乳糖激酶缺乏症(galactokinase deficiency)[21]等與血糖有關的疾病所引發所致。其機制為血糖在水晶體中累積，進而導致滲透壓不平衡，使得水晶體混濁[22]。近年研究指出，這些代謝症候群(metabolic syndrome)相關的疾病也與老化型的白內障有高度相關[23, 24]。

外傷型白內障是眼睛受到鈍器或是穿透性的傷害所造成，包括紅外線、電擊(electric shock)以及游離輻射(and ionizing radiation)等[25]。於 2002 年，Kuhn 等人提出眼睛創傷指數(ocular trauma score, OTS)的預測模型，來預測患者受到眼外傷



後的視力程度[26]；於 2011 年，Acar 等人修改此模型，來針對兒童進行預測[27]，且在後續的研究中，也證實此模型具有相當的可信度[28]。

2.1.2 白內障種類

根據晶體混濁的位置不同，可以將白內障分成三個種類，包括核型(nuclear)、皮質型(cortical)以及後囊下型(posterior)[12]，如圖 2.1.3 所示。

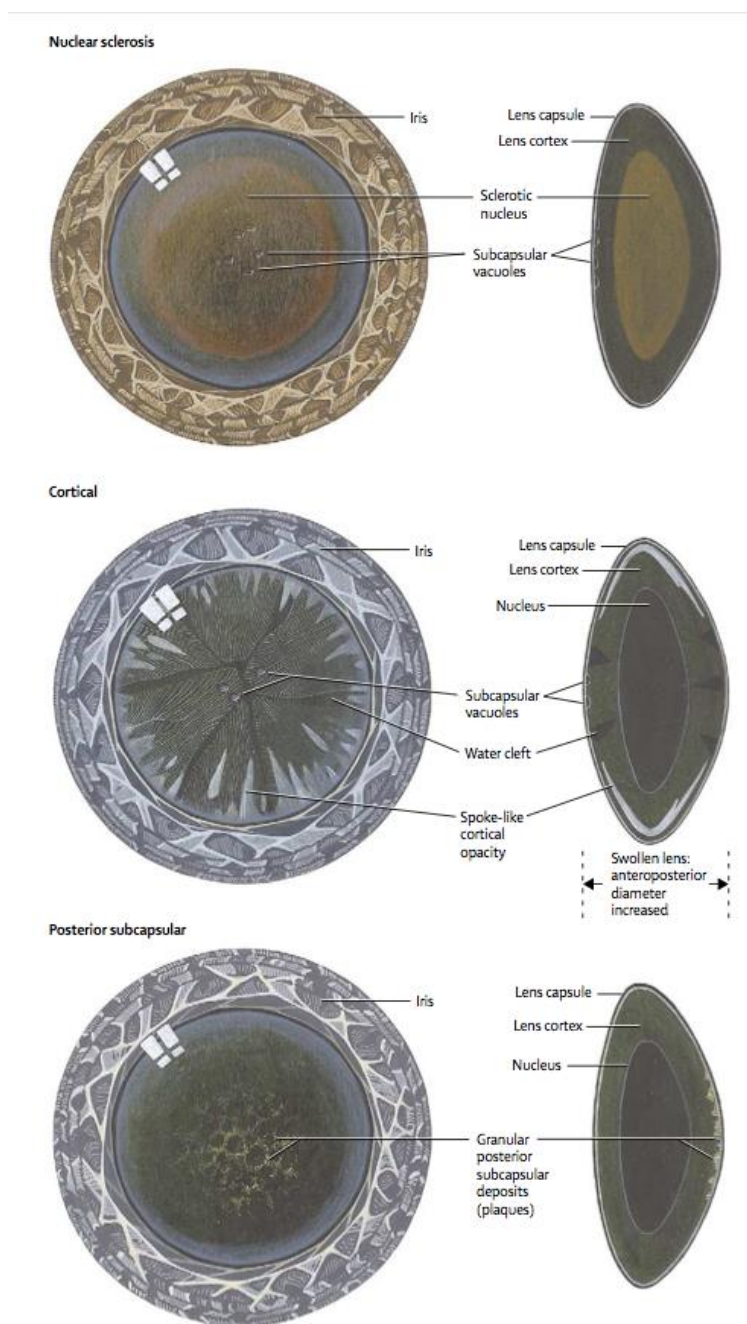


圖 2.1.3 不同型態白內障示意圖[12]

核型白內障發生於水晶體的中心，成因是隨著年齡的增加，水晶體核(lens nucleus)被新生成的纖維細胞擠壓導致變硬變黃，此類型約占 60%[29]。皮質型白內障發生於水晶體的皮質層中，成因是皮質內離子濃度變化導致水晶體混濁，此類型約占 30%[12, 30]。後囊下型白內障發生於皮質層的中後方，成因是上皮細胞增生且異常增大，進而造成水晶體混濁，此類型約占 10%[29]，通常發生在年輕患者身上[31]。

不同種類的白內障潛在風險有許多種，包括年齡、性別、飲酒、紫外光或是遺傳基因等[12]。由 Yu, X. N.等人的文獻整理出 16 個與白內障相關的橫斷研究(Cross-Sectional Study)並綜合分析如表 2.1.1，並指出白內障可能與高血壓有高度相關[32]；另外，將白內障成因相關文獻結果整理如表 2.1.2。由於造成白內障的因素過多，且各因素之間會相互影響，對於白內障發生機制的研究仍面臨相當程度的挑戰，更細節的部分將於 2.2.3 進行探討。

表 2.1.1 不同地區對白內障之橫斷研究的綜合分析[32]

來源	研究類型	種族	樣本數	年齡	白內障種類	有無高血壓	變因
Wang (2009, 中國)	人群研究 橫斷研究	黃種人	3222	≥ 40	核型、皮質型、後囊下型	有	年齡、性別、地區、血糖、收入、身體質量指數、教育等
Delcourt (2000, 法國)	人群研究 橫斷研究	白人	2584	60-95	任何型態	有	年齡、性別、糖尿病、抽菸
Galeone (2010, 義大利)	醫院研究 病例對照	白人	2283	35-79	任何型態	有	年齡、中央型肥胖、性別、糖尿病、抽菸等
Sabanayagam (2011, 新加坡)	人群研究 橫斷研究	黃種人	2794	40-80	核型、皮質型、後囊下型	有	年齡、性別、身體質量指數、糖尿病、維生素使用、抽菸等
Leske (1999, 巴巴多斯)	人群研究 病例對照	黑人	4314	40-84	核型、皮質型、後囊下型	有	年齡、性別、腰臀比、糖尿病

Shah (2007, 巴基斯坦)	人群研究 橫斷研究	黃種人	16402	≥30	任何型態	有	年齡、性別
Rim (2014, 韓國)	人群研究 橫斷研究	黃種人	11591	≥40	任何型態	有	年齡、性別、教育程度、糖尿病、暴露在陽光下的程度
Nirmalan (2004, 印度)	人群研究 橫斷研究	黃種人	5140	≥40	核型、皮質型、後囊下型	有	年齡、糖尿病、尿酸、腰臀比、抽菸
Tavani (1995, 義大利)	醫院研究 病例對照	白人	1541	19-75	任何型態	沒有特別註明	年齡、教育程度、糖尿病、身體質量指數、中央型肥胖、抽菸等
Tsai (2003, 中國)	人群研究 橫斷研究	黃種人	1361	≥65	核型、皮質型、後囊下型	有	年齡、性別、糖尿病、腰臀比、抽菸
Ughade (1998, 印度)	醫院研究 病例對照	黃種人	524	51-70	任何型態	有	年齡、糖尿病、抽菸、青光眼、近視
McCarty (2000, 澳洲)	人群研究 病例對照	白人	4744	40-98	後囊下型	有	年齡、居住地、眼睛屈光程度
Chen (2011, 中國)	人群研究 橫斷研究	黃種人	661	≥65	任何型態	有	年齡、糖尿病、血壓、高胱胺酸尿症、葉酸等
Chen (2011, 中國)	人群研究 橫斷研究	黃種人	645	≥65	任何型態	有	年齡、糖尿病、血壓、高胱胺酸尿症、葉酸等
Goodrich (1999, 澳洲)	人群研究 橫斷研究	白人	3654	49-97	核型、皮質型、後囊下型	有	年齡、性別、身體質量指數、抽菸、血清膽固醇、糖尿病等
Machan (2012, 加拿大)	醫院研究 橫斷研究	白人	6397	1-95	任何形態	有	年齡、性別、抽菸、藥物史他汀的使用、第II型糖尿病

表 2.1.2 白內障潛在風險因素相關研究

潛在風險	成因	說明	文獻
老化	脫醯胺(deamidation)	羧基離子有離子作用力使蛋白質構形改變	[33]
老化	甲基化(methylation)	甲基化可避免雙硫鍵形成，減少白內障出現的可能	[34]
老化	雙硫鍵	形成多組蛋白質內的雙硫鍵使構形改變	[33]
UV 光	色胺酸(tryptophan)的光化學反應	蛋白質的原態結構變得不穩定	[35]
突變	R14C	蛋白質表面性質發生改變	[36]
突變	P23T	降低蛋白質溶解度	[37]
突變	R36S	使蛋白質結晶化(Crystallisation)	[38]
突變	R58H	蛋白質摺疊受損	[39]
突變	W158X	第四希臘鑰匙模組(Greek key motif)結構改變	[40]

2.1.3 白內障治療方法

到目前為止，並沒有藥物能夠有效治療白內障或是延緩發病時間，唯一的方
式是進行白內障手術[41]，其法是將已發病的混濁水晶體更換成人工水晶體
(intraocular lens, IOL)，如圖 2.1.4。研究指出，每 10 位經手術後的病人，有 9 位
能回復到 20/40 或更加的視力[42]。20/40 視力是指站在距離視力表 20 英尺(foot)
的位置，能看清楚一般人在離視力表 40 英尺所能看見之物，換算台灣常用的單位
即為視力 0.5 或是近視 150 度。原則上，在發病的任何時期都可以進行更換手術，
不過一般而言，只要在影響到日常工作與生活習慣時再行更換即可。雖然白內障
手術有著相當高的成功率，但在術後也可能引發前囊混濁化(posterior capsular
opacification)的併發症等[43]。

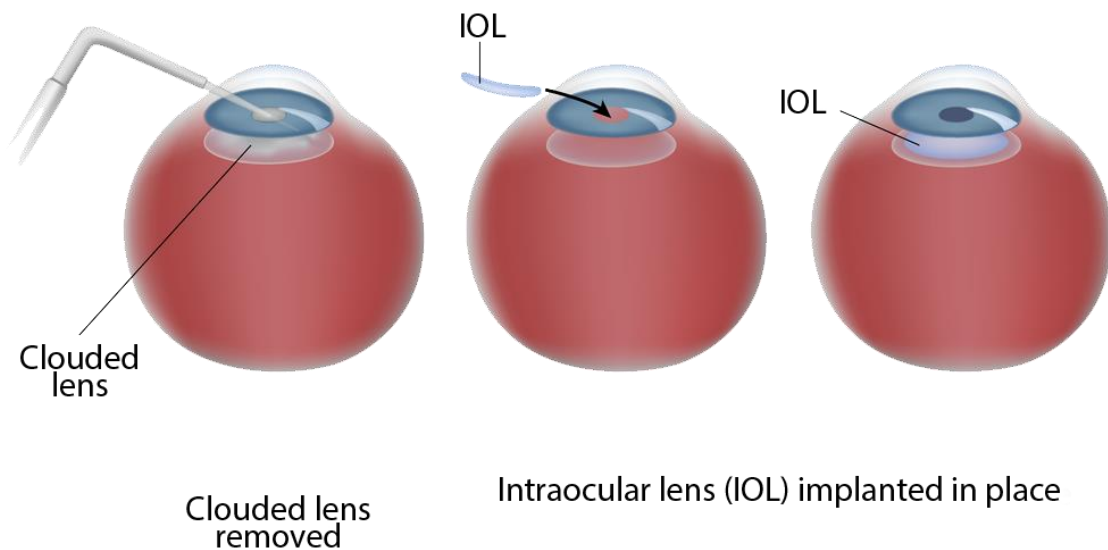


圖 2.1.4 白內障手術更換人工水晶體之示意圖

<http://www.ncascade.com/services/cataract-surgery/>

隨著技術不斷地進步，白內障手術已經發展成熟，但進行手術需耗費相當的
醫療資源以及需有一定的技術，並非所有地區都能做到，因此發展非手術性
(nonsurgical)的治療法是目前研究的目標。為了達成這個目標，必須得先了解白內
障形成的機制，如此才能對症下藥，找出解決之道。



2.2 水晶體蛋白(Crystallin)

前一節提到，水晶體提供良好的折射率與穿透度，使光線可以順利聚焦於視網膜上，這主要是來自於高濃度的水晶體蛋白的功能。在水晶體內的可溶性蛋白中，有超過 90% 為水晶體蛋白[44]，它是一種具有高度穩定性的結構蛋白(structural protein)[45]。其主要分三大類，包括 α 型、 β 型以及 γ 型，是依據膠體層析管柱(gel filtration chromatography column)所沖提(elute)出來的順序所命名。其中後兩者的序列(sequence)以及結構非常相似，被合稱為是超族(superfamily)的 $\beta\gamma$ 型水晶體蛋白[46]，詳細水晶體蛋白的分類如表 2.2.1[47]。本節將對 α 型、 $\beta\gamma$ 型以及本研究的目標 γ D 型水晶體蛋白進行介紹。

表 2.2.1 水晶體蛋白分類一覽[47]

水晶體蛋白	殘基數 (Residues)	分子量 (kDa)	等電點 (pI)	蛋白質序列來源(SwissProt)	PDB Code
α A	173	19.909	5.6	P02489	-
α B	175	20.159	6.8	P02511	2WJ7 (residue 67-157)
β A1	198	23.191	6.4	P05813	-
β A2	196	21.964	5.9	P53672	-
β A3	215	25.150	5.7	P05813	-
β A4	195	22.243	5.8	P53673	-
β B1	251	27.892	8.6	P53674	10ki
β B2	204	23.249	6.5	P43320	-
β B3	211	24.230	5.9	P26998	-
γ S	177	20.875	6.4	P22914	1ha4 (C-td)
γ A	173	20.761	7.8	P11844	-
γ B	174	20.776	7.0	P07316	-
γ C	173	20.747	7.0	P07315	-
γ D	173	20.607	7.2	P07320	1hk0

2.2.1 α 型水晶體蛋白

在人類水晶體蛋白中，大約有 50% 為 α 型，是含量最多的水晶體蛋白[46]。單體(monomer)分子量約為 20 kDa[48]，較常以多聚體(multimers)的形式存在。除了擁有結構蛋白的功能之外，同時也有伴護蛋白(chaperone)的功能[49]。其可分成 α A 型與 α B 型，分別具有 173 個和 175 個胺基酸，兩者序列約有 60% 的相似度[50]，且皆具有小熱休克蛋白(small heat shock protein, sHsp)的特性[51]。

由於 α 型水晶體蛋白具有許多不同的多聚體，所以尚未有完整的晶體結構被 X 射線繞射(X-ray diffraction)或核磁共振光譜(Nuclear Magnetic Resonance spectroscopy, NMR)解出來。在蛋白質資料庫(protein data bank, PDB)中，目前僅有 α B 型的結構，如圖 2.2.1[52]。另外，在二級結構(secondary structure)方面， α 型水晶體蛋白約有 10% 是 α -螺旋(α -helix)，50% 是 β -摺板(β -sheet)[53]。

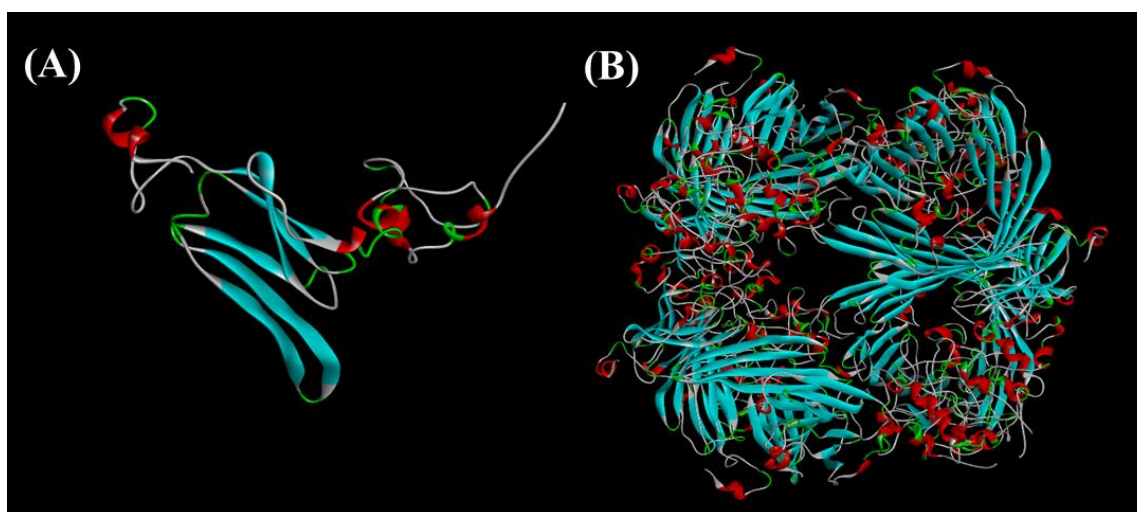


圖 2.2.1 人類 α B 型水晶體蛋白結晶結構圖(A)單體(pdb code: 2YGD); (B) 24 聚體 (24-meric, pdb code: 2YGD)[52]

目前有許多研究會針對其伴護蛋白以及小熱休克蛋白的特性來分析，例如 α 型水晶體蛋白的伴護性質(chaperone-like properties)包括防止蛋白質變性(denature)以及沉澱(precipitation)，並增加細胞對於外壓的耐受性[54]，且此功能對於維持水晶體穿透度與防止白內障發生都有相當程度的關聯[55]。Ram H. Nagaraj 等人在文

獻中整理出， α 型水晶體蛋白能抑制不同的疾病，包括蛋白質聚集、細胞凋亡 (apoptosis) 或是神經炎症(neuroinflammation)等，但同時也可能具有引發癌症(cancer)等能力[56]，如圖 2.2.2 所示。它的功能就像是一把雙面刀，在未來發展相關療法時得更加小心。

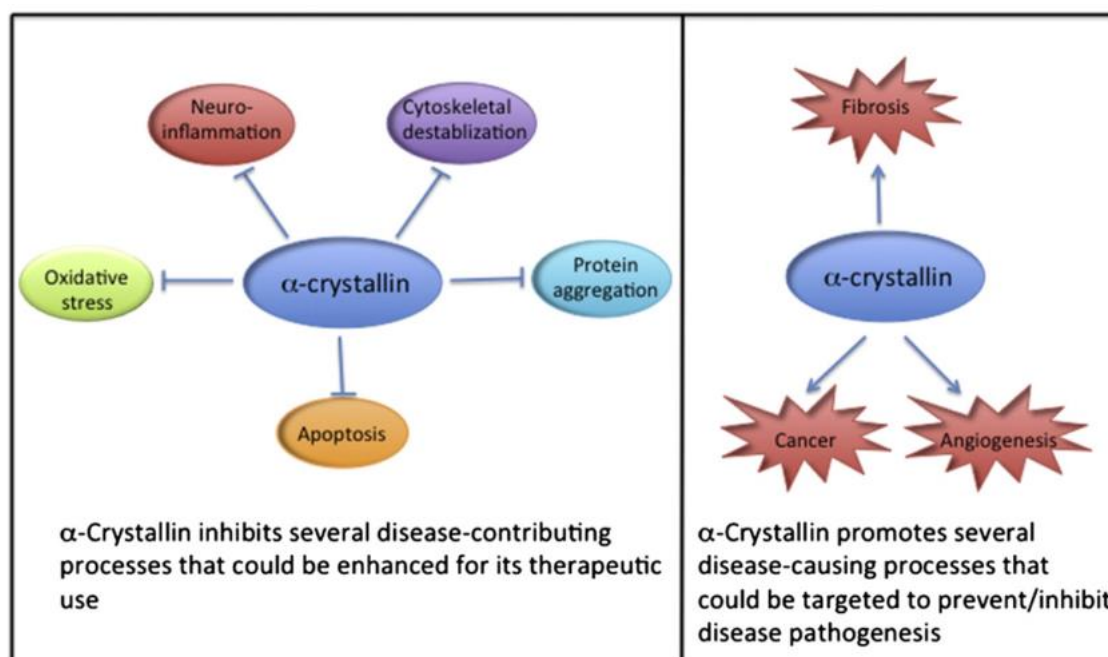
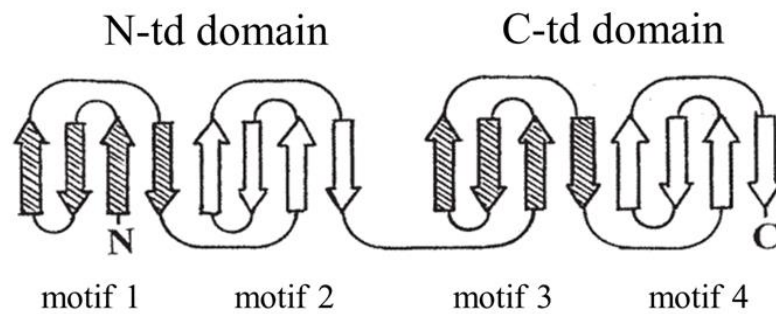


圖 2.2.2 α 型水晶體蛋白同時具有抑制與引發疾病的示意圖[56]

2.2.2 $\beta\gamma$ 型水晶體蛋白

$\beta\gamma$ 型水晶體蛋白皆為結構蛋白，是由兩個相似的區域(domain)組成，根據序列順序分別命名為 N 端區域(N-terminal domain)與 C 端區域(C-terminal domain)，每個區域再由兩個模組(motif)所構成，而每個模組是由四個 β -strands 經摺疊形成希臘鑰匙模組(Greek key motifs)，如圖 2.2.3 所示[57, 58]。一般而言，每個模組大約由 40 個胺基酸所組成，模組與模組間以胜肽鍵(peptide bond)連接。 $\beta\gamma$ 型水晶體蛋白可以區分成 β 型與 γ 型，雖然兩者的結構相似，但序列的相似性僅有約 30%[47]，最主要的差別在於 β 型在 N 端或 C 端序列的延伸[59]。另外， β 型大都由多聚體形式存在，而 γ 型則以單體狀態存在[60]。

(A)



(B)

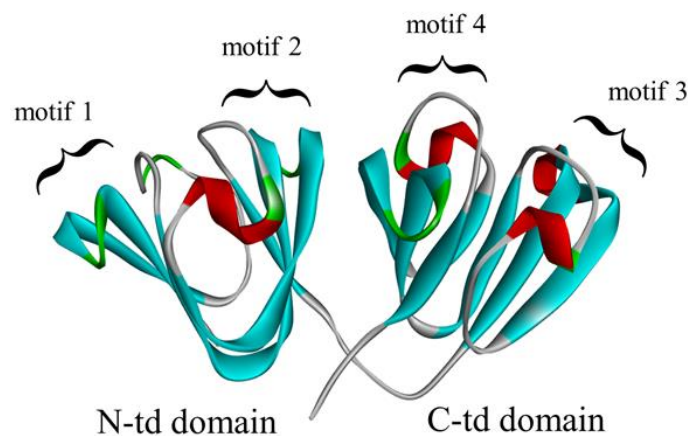


圖 2.2.3 $\beta\gamma$ 型水晶體蛋白結構示意圖(A)平面圖 (B) 3D 立體圖(以 γ D 型為例, pdb code: 1hk0)[57, 58]

β 型水晶體蛋白的單體分子量約為 22~28 kDa，根據性質可再分為酸性的 β A1~ β A4 型，具有 N 端序列的延伸，以及鹼性的 β B1~ β B3 型，同時具有 N 端和 C 端的延伸[61]。 β 型大都由多聚體形式存在，由不同比例的次單元(subunits)所組成，與溫度、離子強度等外在條件有關。在人類水晶體中， β 型水晶體蛋白一般以三種組裝(assemblies)形式存在， β H 型(六聚體/八聚體)、 β L₂ 型(二聚體)以及 β L₁ 型(中間體型)[61, 62]。另外，也有研究顯示，若缺乏 β A3/A1 型水晶體蛋白，會使得星狀細胞(astrocytes)的功能喪失，引發殘留型增殖性初級玻璃體症(Persistent fetal vasculature, PFV)[63]，進而導致視力的喪失。

γ 型水晶體蛋白單體分子量大約為 20~21 kDa，可再細分為 γ A、 γ B、 γ C、 γ D、 γ E、 γ F 以及 γ S 型，共七種[64]。其中，比起 γ A~F 型水晶體蛋白， γ S 型在 N 端區域的多了 4 個殘基(residue)的長度，在 C 端區域則少了 2 個殘基的長度，另外在 N 端與 C 端連接處多了 2 個殘基的長度；而 γ B 在連接處則比其他多了 1 個殘基的長度[47]，詳細的殘基數目見表 2.2.1。在人類水晶體中， γ C 與 γ D 型有較高的表現(expression)量[33]，而雖然人類具有 γ E 和 γ F 型的基因，但並沒有表現出來[65]。 γ 型水晶體蛋白是非常穩定的蛋白，能保存於尿素(urea)或是 2~3 M 的胍鹽酸(guanidinium chloride)溶液中[66]，主要分布於水晶體核心處[47]。由於在老化型白內障的分類中，屬於核型白內障的比例是最高的，因此本研究將針對在水晶體核心含量最多的 γ D 型水晶體蛋白[67]，探討單點突變對其結構穩定性的影響。在下節將介紹 γ D 型水晶體蛋白的基本特性與近年來的研究。

2.2.3 人類 γ D 型水晶體蛋白(Human γ D crystalline, H γ DC)

人類 γ D 型水晶體蛋白是水晶體核心含量最多的蛋白，在 γ 型水晶體蛋白中含量屬第二高[68]，以下以 H γ DC 簡稱之。H γ DC 是一種結構蛋白，由 173 個胺基酸組成，分子量為 20.6 Da，等電點為 7.2，其結構如圖 2.2.3，主要分為 N 端區域與 C 端區域，每個區域由兩個模組(motif)所組成。H γ DC 的結構具有高穩定性，主要源於六對芳香族殘基(F11/Y6, Y16/Y28, Y50/Y45, Y92/Y97, Y138/Y133, F117/F115)散佈在整個蛋白質中[69]，還有疏水的區域界面(domain interface)[70]，以及在希臘鑰匙模組常見的酪氨酸轉角(tyrosine corners)的貢獻[71]。而這組酪氨酸轉角位於 N 端區域的 Tyr62 以及 C 端區域的 Tyr150，它透過氫鍵作用力連接兩對 β -strands[72]。另外有趣的是，比起任何 α 型和 β 型的水晶體蛋白，H γ DC 具有較多個數的半胱胺酸(cysteine)，但卻沒有形成任何一對雙硫鍵(disulphide bond)[73]。

承 2.1.1 節所述，暴露在紫外光下對於白內障的形成有高度相關。研究發現，脊椎動物(vertebrates)具有高保守性的四對色胺酸(tryptophan)[47]，分別位於 N 端區域的 W42、W68 與 C 端區域的 W130、W156，能夠藉由分散 UV 光的輻射來穩

定蛋白質的結構[74, 75]。在 Chen, J.等人的研究同樣指出，蛋白質內的色胺酸能有效過濾 UV 光，透過淬滅(quenching)的機制來避免光化學的傷害[76]。另外，老化也是白內障主要成因，許多相關的後轉譯(post-translational)修飾的研究如 Q12、N49、N160 的脫醯胺(deamidation)反應[77, 78]，Y45、W156 的氧化(oxidation)反應[77, 79]，C110 的甲基化(methylation)反應[34]，G1 的胺甲醯化(carbamylation)反應[34, 78]，以及 C18-C32、C108-C110 的雙硫鍵結反應[80]，都與老化形成的白內障有關聯[33]。

白內障形成的機制是因水晶體蛋白在不正常的情況下產生聚集，造成水晶體混濁，使得入射光無法匯於視網膜上，進而衍生成視力的障礙。這些蛋白質的聚集源自於部分去摺疊(partially unfolded)的中間體(intermediate)，雖然已經廣泛的被各界人士研究，但詳細的路徑尚未被完全解開。大部分的研究使用熱去摺疊法(thermal unfolding)或是化學去摺疊法(chemical unfolding)來探討摺疊機制，Kosinski-Collins 等人發現 HyDC 在去摺疊的過程中僅有單一的去摺疊轉變態(unfolding transition)[66]；而後根據動力學分析三組突變的 HyDC，發現過程中有部分去摺疊的中間體(partially unfolded intermediate)[81]；最後，原態(wildtype)去摺疊的路徑被修正為三態模型(three-state model)[9]。在去摺疊的過程中，HyDC 的 N 端區域會先進行去摺疊，再來才是 C 端區域的去摺疊，如圖 2.2.4 所示[82]。除此之外，還有以突變(mutation)的方式來對蛋白質的穩定性與去摺疊進行研究，主要分兩種類型。第一種是以突變的方式改變蛋白質的穩定性，包括針對高保守性芳香族殘基對(F11/Y6, Y16/Y28, Y50/Y45, Y92/Y97, Y138/Y133, F117/F115)的突變[69]，四個色胺酸(W42, W68, W130, W156)的突變，以及疏水區域界面(Q54/Q143, R79/M147)的突變等[9]。第二種根據與白內障相關的體內(*in vivo*)實驗研究結果，模擬蛋白質結構的改變，包括家族性遺傳疾病的突變種(R114, R38, R36, R14, P23, R36, R58, W42)[38-40, 83-85]，可能發生麩醯胺酸(glutamine)脫醯胺(deamidation)反應的位置(Q54, Q143)[86]，以及遺失型的突變(deletion mutant)(W156X)[87]等。



圖 2.2.4 HyDC 去摺疊與聚集的可能途徑[82]

這些突變若發生在 N 端區域，會使得 N 端區域在低變性劑濃度就發生去摺疊，但不會影響 C 端區域的結構；若突變發生在 C 端區域，則會同時使 N 端區域與 C 端區域發生改變[69]；而若突變發生在疏水區域界面，僅在 N 端區域的結構發生改變，表示疏水區域界面的功能主要是作為穩定 N 端區域的鷹架(scaffold)[9, 86]。

有部分研究針對外在因素如 UV 光的照射或是 pH 值的改變，來探討 HyDC 的去摺疊機制。如 HyDC 暴露 UV 光下一定時間後，會產生聚集的行為[41, 88]，或是使得骨架(backbone)上的特定殘基被切斷造成損害[89]，其中色胺酸扮演重要角色如上所述。在 2013 年，Xia, Z. 等人使用電腦模擬 HyDC 受光照後的影響，將色胺酸突變成犬尿胺酸(kynurenine)，發現 HyDC 的去摺疊速度變快，且 N 端區域與 C 端區域同時發生結構改變[74]。而 HyDC 在酸性條件下，會形成類澱粉纖維的聚集，且其二級結構與三級結構都會有大量的變化[90]。同時，根據 Das 等人的研究，發現 HyDC 可能藉由區域置換(domain swapping)的方式發生聚集[67]。綜合上述，將 HyDC 去摺疊與聚集的可能途徑整理於圖 2.2.4[82]。

為了治療因 H_γDC 聚集所引起的白內障，許多不同的化合物被研究作為預防白內障形成的療方。例如，咖啡因(caffeine)能有效避免 UV 光的輻射傷害[91]，白當歸素(byakangelicin)能夠抑制半乳糖誘發(galactose-induced)的白內障[92]，檸檬酸(citric acid)能抑制患糖尿病大鼠的白內障[93]，結合阿斯匹靈(aspirin)與肌肽(carnosine)的眼藥水能抑制鏈脲佐菌素(streptozotocin)所引起的白內障[94]，維生素 K(vitamin K)能幫助鈣離子的 ATP 酶保持活性並抑制糖尿病所引發的白內障[95]，檸檬酸鈉(sodium citrate)能夠鍵結並穩定 H_γDC 的去摺疊中間體[96]，4-苯基丁酸鈉鹽(sodium 4-phenyl butyrate)能增加缺陷型 H_γDC 的溶解度[97]，白藜蘆醇(resveratrol)能抑制 H_γDC 受光照產生的聚集行為[98]。另外，許多抗氧化劑也能作為預防白內障的藥物，如 2,6-二第三丁基對甲酚(butylated hydroxytoluene, BHT)[99, 100]、類維生素 E (Trolox)[101]與維生素 E(vitamin E)[102]等。

雖然到目前為止關於 H_γDC 的研究已有許多，但對於其形成白內障的機制以及有效治療的方式仍沒有一個完善的解答，吾人將近年來 H_γDC 相關的研究整理於表 2.2.2。

表 2.2.2 近年與 H_γDC 相關之研究

文獻	蛋白質種類	研究方法	研究結果
Chen, J., et al. (2006)[75]	WT、W42F、W68F、W130F、W156F	Fluorescence spectroscopy(intrinsic, tryptophan)、QM-MM simulations	實驗與模擬結果說明 H _γ DC 上的色氨酸透過能量轉移與電子轉移的機制，保護水晶體以避免受到紫外線的光化學傷害。
Chen, J., et al. (2009)[103]	WT、W42F、W68F、W130F、W156F	Fluorescence spectroscopy(intrinsic, tryptophan)、QM-MM simulations	研究結果指出經演化以希臘鑰匙方式摺疊的結構，能有快速電荷轉移的淬滅現象，進而可以減少因紫外光誘導造成的光損傷。
Moreau, K. L., et al. (2009)[104]	WT、L5S、V75D、I90F	Circular dichroism、Fluorescence spectroscopy、Equilibrium unfolding/refolding、Thermal dena-	H _γ DC 是以三態模型進行去摺疊，L5S 及 V75D 會使 N 端區域不穩定，而 I90F 則是使 C 端區域不穩定，且這些部分

		turation、Turbidity	去摺疊態的 HyDC 是造成聚集形成的前驅物。
Das, P., et al. (2010)[105]	WT、E135A	MD simulations (NAMD, 8M urea, 425K)	模擬結果發現 Glu135-Arg14 形成的鹽橋能穩定疏水區的界面，在突變為 E135A 時會導致 HyDC 開展與聚集。
Banerjee, P., et al. (2011) [106]	WT、P23T	NMR spectroscopy、Bis-ANS binding、modeling	P23T 能使 HyDC 的表面疏水面積增加並降低蛋白質骨架的移動性，以此解釋 HyDC 的聚集機制。
Das, P., et al. (2011)[67]	WT	MD simulations (NAMD)	模擬結果發現 HyDC 的 N 端區域會與 C 端區域的 motif-4 以區域置換方式形成聚集，進而導致老化型白內障。
Goulet, D. R., et al. (2011) [107]	WT、L5S、I90F	Turbidity、Equilibrium unfolding/refolding	檸檬酸鈉可以有效降低 L5S 及 I90F 在高濃度胍鹽酸環境下的聚集速率，也能有效減緩原態及 I90F 的復性速率，但對 L5S 沒有顯著影響。
Sahin, E., et al. (2011)[108]	WT、M69Q、S130P	Computational predictions、Equilibrium unfolding/refolding (urea)、DSC、Size exclusion chromatography	結果指出，比起原態能量較不穩定的 S130P，有較低的自身聚集傾向(Intrinsic Aggregation Propensity)，與一般的認知不同。
Wang, L., et al. (2011)[109]	WT、R36P	Clinical evaluation、Mutation detection、Bioinformatics analysis	在中國裔白內障患者發現 R36P(condon:110G-->C)，此突變會增加 HyDC 的局部親水性，但結構上與原態沒有太大的差異。
Ji, F., et al. (2012)[110]	WT、R76S	Circular Dichroism、Fluorescence spectroscopy、Thermal and chemical denaturation、Turbidity、NMR spectroscopy	在印地安裔家族，R76S 與早年白內障的形成有關，但其結構穩定性與原態無明顯差異，只有等電點極表面帶電分布有所不同。
Mishra, S., et al. (2012)[111]	WT、I4F、V76D	Circular Dichroism、Size exclusion chromatography、Denaturant-unfolding(GdnHCl)	雖然 I4F 與 V76D 能穩定 HyDC 的部分去摺疊中間體，但其卻使其無法受到 α 型水晶體蛋白的伴護，進而引

Moreau, K. L., et al. (2012)[112]	WT、V75D、I90F	Native interaction assays	發聚集並導致白內障的出現。I90F 與 V75D 是因突變而無法受到 α 型水晶體蛋白的伴護，進而產生聚集，與造成先天型白內障。
Ji, F., et al. (2013)[113]	W42R	X-ray diffraction、Chemical denaturation (GdnHCl, urea)、Trypsin digestion、UV-C irradiation、NMR spectroscopy	W42R 與中國裔家族先天型白內障有關，其溶解度與穩定性都比原態來得低，且可作為因光受損的老化型白內障相關之水晶體蛋白模型
Ji, F.L., et al. (2013)[68]	P23T	X-ray diffraction	P23T 為最常出現的突變，根據結果發現其結晶結構與原態差異不大，但溶解度變小，可能與蛋白質聚集有關。
Moran, S. D., et al. (2013)[114]	S84C	Mass spectrometry、UV-Vis spectroscopy、TEM、ThT binding assay、2D IR	UV-B 照射會切斷 H γ DC 的骨架，並氧化區域界面的側鏈，產生類澱粉纖維的聚集，此聚集主要由 C 端區域所組成。
Schafheimer, N., et al. (2013)[115]	WT、W42F、W68F、W130F、W156F	Photoaggregation experiments、TEM、Absorbance spectra measurements、Circular dichroism、Thermal unfolding	H γ DC 中，高保守性的色氨酸殘基對，能透過能量轉移機制，來保護 H γ DC 避免受到 UV 光的損害。
Wu, J. W. R., et al. (2013)[98]	WT	Turbidity、TEM、ANS fluorescence spectroscopy、SDS-PAGE、Quantification of thiols groups	白藜蘆醇具有抑制 H γ DC 因 UV-C 照射所形成聚集的能力，可能與白藜蘆醇的自由基清除能力有關。
Xia, Z., et al. (2013)[74]	WT、KN mutant	MD simulations (NAMD)	使用分子動力學模擬色胺酸受光照氧化成大尿胺酸(kynurenine)的行為，發現其更能吸引水分子以及極性側鏈，使得鄰近的兩組希臘鑰匙模組間形成疏水中心。
DiMauro, M.A., et al.(2014)[116]	WT	Mass Spectroscopy、Circular Dichroism、Intrinsic Tryptophan Fluorescence、Bis-ANS Fluorescence、Thermal and chemical unfolding、DLS、APBS、MD simulations	結果指出將殘基 Gly1 和 Lys2 乙醯化(acetylation)會形成球狀結構，進而引發聚集，與老化型的白內障有關。

Schafheimer, N., et al. (2014)[117]	WT、Y:A mutants、F:A mutants、C:S mutants	Photoaggregation experiments、Mass spectroscopy、Circular dichroism	Y16 會因照紫外光而被激發，並將電子轉移至相鄰的 C18 上，進而啟動自由基的聚合反應。在突變為 Y116A 後，會使因光照而形成聚集的速率降低。
Wu, J. W. R., et al. (2014)[90]	WT	Turbidity、ThT、TEM、Circular dichroism、ANS Fluorescence、Intrinsic Tryptophan Fluorescence、Thermally denaturation、DLS、Bioinformatics prediction、MD simulations	HyDC 在酸性條件下會形成纖維狀的聚集，且其二級結構與三級結構都會有大量的變化，且預測其類澱粉纖維的生成是由 C 端區域所主導。
Yang, Z., et al. (2014)[118]	WT、Y45A、Y50A、Y133A、Y138A	MD simulations (NAMD)	利用分子動力學模擬在 N 端區域與 C 端區域 β -hairpin 上的酪氨酸對，其對於 HyDC 結構的穩定性有重要影響。
Serebryany, E., et al.(2015)[119]	W42Q	Chemical denaturation (GdnHCl)、Turbidity、SDS-PAGE、Bis-ANS、ThT、Filter Trapping、TEM、Mass Spectrometry	W42Q 是模擬蛋白質氧化態的突變型，研究結果發現原態 HyDC 可作為 W42Q 藉由形成分子內雙硫鍵產生聚集反應的催化劑。
Eugene S., et al. (2016)[120]	WT、W42E、W68E、W130E、W156E	Equilibrium unfolding/refolding(GdnHCl)、Circular dichroism、ThT、TEM	結果發現 W42E 與 W130E 會產生聚集，但 $\alpha\beta$ 型水晶體僅能有效抑制 W130E 的聚集，表示兩者聚集路徑因突變的區域不同而有所不同。此聚集方式與區域置換是類似的。
Garcia-Manyes, S., et al. (2016)[121]	WT	Force spectroscopy、Light Scattering	結果發現在 N 端區域上鄰近的 β 1 與 β 2 發生置換，進而產生聚集。
Quintanar, L., et al. (2016)[122]	WT	Turbidity、SDS-PAGE、TEM、Circular dichroism、Thermal Stability、NMR Spectroscopy	結果指出銅離子與鋅離子能夠誘導 HyDC 產生非纖維狀的聚集，且前者不僅能降低 HyDC 的熱穩定性，更可以引發雙硫鍵的生成。

本研究將採用單點突變(single point mutation)以及分子對接(molecular docking)的方式來探討 H_γDC 結構的改變。根據陽明大學生化暨分子生物研究所林達顯教授實驗室的實驗結果，以核磁共振光譜(Nuclear Magnetic Resonance Spectroscopy, NMR)分別測得 H_γDC 在原態與 1M 胍鹽酸環境下的結構，並使用殘留偶極偶合(Residual dipolar coupling, RDC)分析其結構差異，結果如圖 2.2.5，發現於殘基 20-35、114-122 以及 130-150 的 RDC 值有較大的差異。另外，根據本實驗室過去的研究發現，相對於中性環境下，H_γDC 處於酸性環境會有四組類澱粉纖維中含量豐富的 β -sheet 結構生成，分別位於殘基 28-32、71-74、115-118、160-164，如圖 2.2.6[123]。結合上述兩者的研究結果，殘基 115-118 可能為 H_γDC 在變性(denature)過程中變動較大的區域，因此本研究將針對此區域進行深入探討。

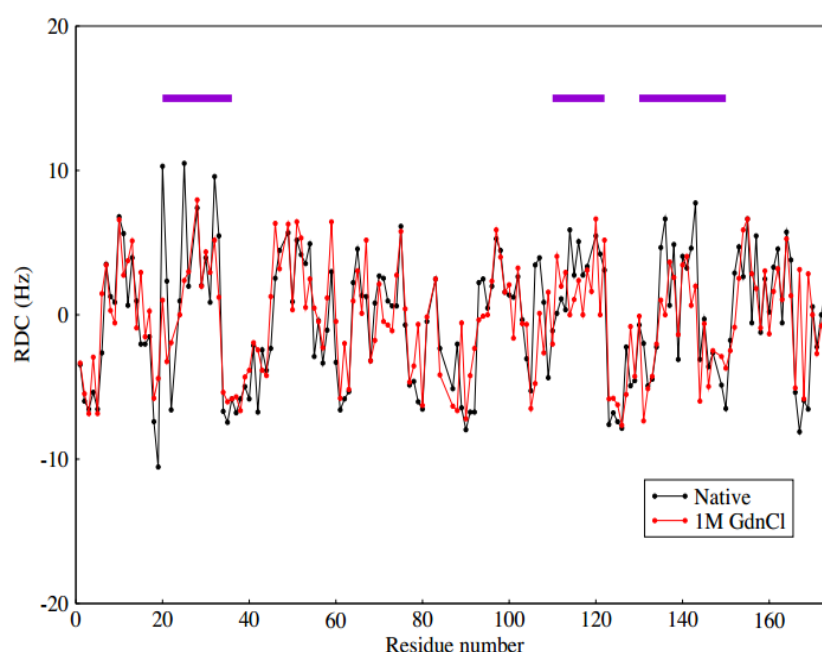


圖 2.2.5 H_γDC 各殘基在原態與 1M 胍鹽酸環境下的 RDC 值
(數據取自陽明大學生化暨分子生物研究所林達顯教授實驗室的實驗結果)

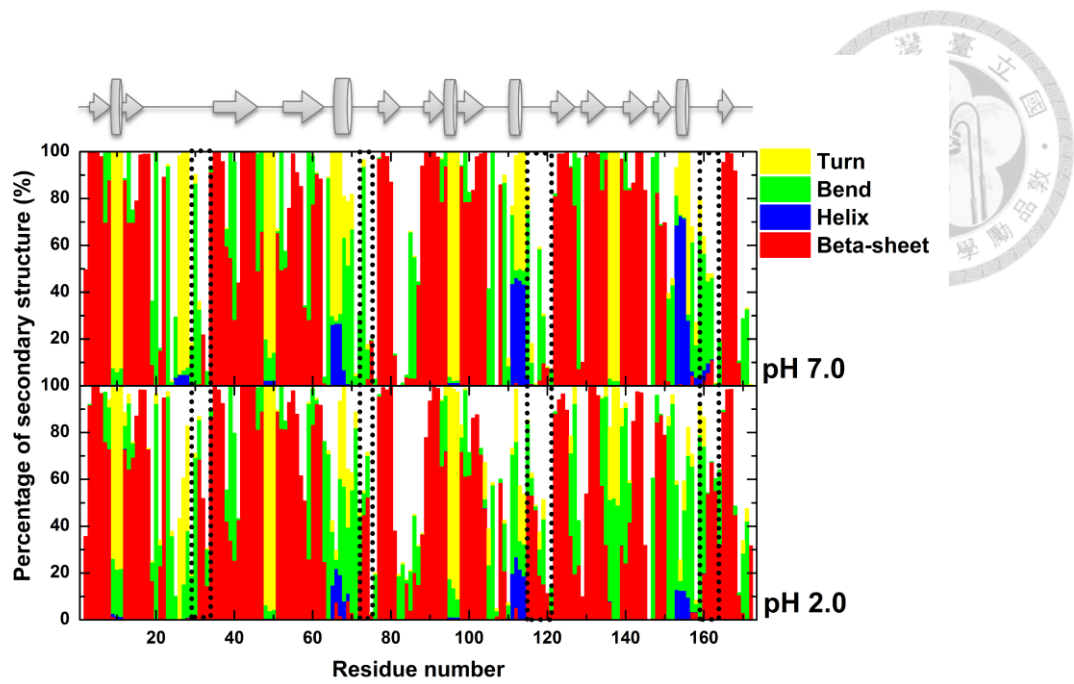


圖 2.2.6 H γ DC 各殘基二級結構於模擬過程中的存在百分率[123]

2.3 電腦模擬(Computer Simulation)

電腦模擬是利用電腦科學與技術來建構模型(model)，給予必要的條件(通常要與實驗的條件相符)，並加以運算得出結果的預測工具。它的主要優點是能夠有效地減少人力與時間成本，且不受限於環境條件的影響，得到最符合目標的結果；但缺點是每一種模型必帶有其本身特定的假設，而當此假設與真實情況相差過大時，即使信度(reliability)很高，其效度(validity)也可能是非常低的。一般而言，電腦模擬主要分為兩種類型，第一種是搭配既有的實驗結果，以微觀(microcosmic)的角度來模擬反應過程，同時印證巨觀(macroscopic)的實驗結果，並藉此推測反應機制(mechanism)；第二種是透過統計學(statistics)或其他特定的模擬方式，來預測未知的實驗結果，使用此法能有效地減少實驗可能所耗費的人力與時間。

近年來，電腦運算能力(computing power)的進展一日千里，呈現指數型(exponential)的遞增(圖 2.3.1)，這意味著利用電腦模擬能夠處理越來越龐大的系統，且隨著演算法(algorithm)不斷改良，其模擬的品質與效率也隨之提升。自 2012 年以來，大數據(big data)一詞蔚為流行，其是指某特定領域所涉及的資料量規模巨大到無法透過人工，在短時間內處理並整理成為人們所能解讀的資訊[124]。透過數十、數百甚至數千台伺服器(server)上同時平行運算(parallel computing)的軟體，進行資料庫系統的管理[125]。目前已經用於許多領域之中，包括氣候預測[126]、基因體學[127]、神經網路體學[128]以及生物和環境研究[129]等。於 2016 年三月，由 google 開發的人工智慧軟體 AlphaGo，在公認複雜度(complexity)最高的圍棋領域，打敗世界頂尖的職業選手，如此得以證明電腦輔助決策的時代即將來臨。

目前在生物分子領域中也有大量模擬相關的研究，包括藥物開發以及鑑定結構等[130, 131]。本節將針對本研究所使用的生物資訊工具(bioinformatics tools)、分子動力學模擬(molecular dynamics simulations)以及分子對接(molecular docking)進行介紹。

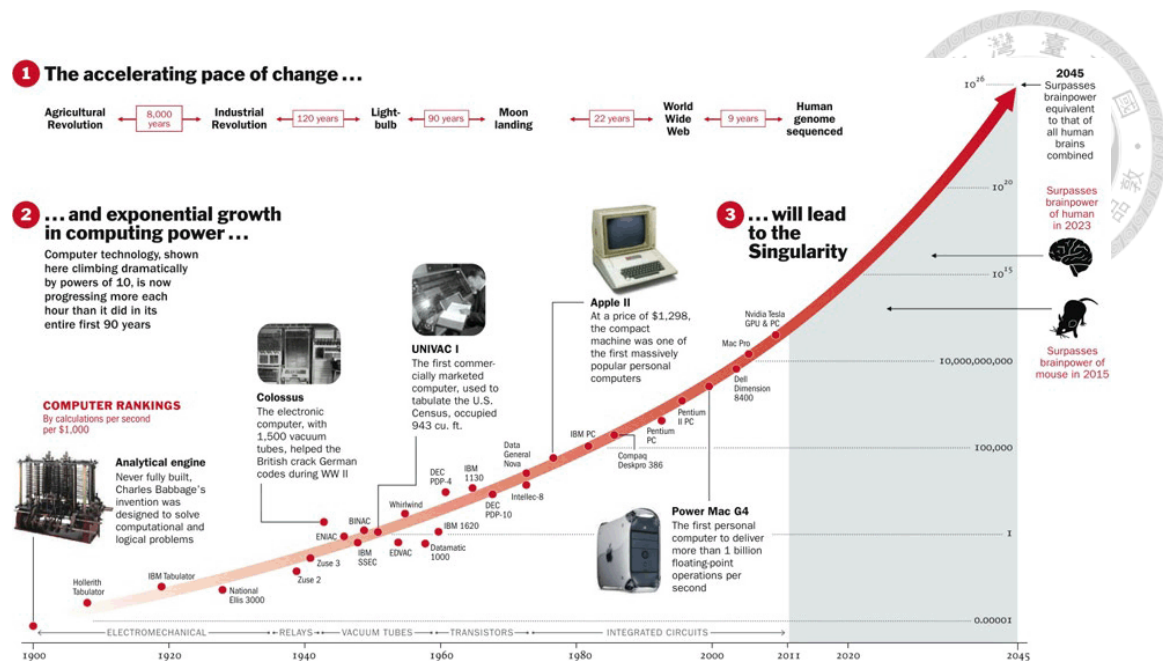


圖 2.3.1 電腦運算能力進展的趨勢圖

<https://www.quora.com/Where-can-I-see-a-graph-of-computing-power-growth-for-the-next-38-years-2050>

2.3.1 生物資訊工具(Bioinformatics tools)

生物資訊學(bioinformatics)發展於 1980 年代後期，是一門結合應用數學、資訊學、統計學和計算機科學等方法來研究生物學的學問。簡單來說，就是分子生物學(molecular biology)與資訊技術(information technology)的結合體。其原理是針對各式各樣的生物學資料，採以蒐集、篩選、歸類、整理、計算和模擬等一系列的流程，來完成使用者欲得到之結果。近年主要的研究方向包括序列比對(sequence alignment)[132, 133]、基因預測(gene prediction)[134]、基因重組(gene recombination)[135]、蛋白質結構預測[136, 137]、基因表現(gene expression)[138]、標靶藥物(target-drug)開發[139]，以及建立演化模型(evolution model)[140]等。生物資訊工具利用生物資訊學的理論，搭配簡單易懂的使用者介面(user interface)，讓使用者能快速獲得分析結果，常以軟體(software)或是網頁伺服器(web server)的方式呈現。本研究著重於探討蛋白質在突變(mutation)前後結構的穩定性以及聚集傾向(aggregation potential)，藉以得知特定區域或特定殘基是否對於蛋白質結構扮演關鍵性的角色，而以下將針對本研究所使用的生物資訊工具進行介紹，包括

Istable[141]、DUET[142]、Eris[143-145]、EASE-MM[146, 147]、PASTA2.0[148]、TANGO[149-151]、AGGRESCAN[152]、WALTZ[153, 154]、FoldAmyloid[155]、Zygggregator[156-159]以及 AGGRESCAN3D[160]。

Istable 是一款預測經單點突變(single mutation)之蛋白質穩定度的工具，其使用的訓練資料(training dataset)來自線上資料庫 ProTherm[161, 162]，內含有不同蛋白質在特定條件下，進行突變後之熱力學相關的實驗結果，如吉布斯自由能(Gibbs free energy)、焓(enthalpy)或是熱容(heat capacity)的變化量等。除此之外，Istable 使用支持向量機(Support Vector Machine, SVM)將多種線上預測工具進行整合，包含以下五種 I-Mutant 2.0[163]、AUTO-MUTE[164]、MUPRO[165]、PoPMuSiC 2.0[166]和 CUPSAT[167]，透過最小化經驗誤差(Empirical risk minimization)與最大化幾何邊緣區(Maximize geometric margins)的方式，取得可信度較高的結果。此法定義 $\Delta\Delta G$ 為正值時，表示蛋白質在突變之後會變得更穩定，而負值則是不穩定。

DUET 同樣是一款整合型預測蛋白質突變穩定性的工具，其訓練資料的來源也是 ProTherm，但選用的資料與 Istable 不同，主要參考自 SDM[168, 169]與 mCSM[170, 171]兩種方式。前者根據同源蛋白家族(homologous protein families)的統計數據，計算蛋白質在原態(wildtype)與突變後的能量差值；後者則是根據蛋白質的結構特徵(structural signatures)來作為預測的方法。接著，DEUT 使用支持向量機(SVM)之序列最小優化法(Sequential Minimal Optimization)將上述兩者整合，並經測試後得出整合後結果比單獨使用 mCSM 或 SDM 結果來得更佳的結論。此法定義 $\Delta\Delta G$ 為正值時，表示蛋白質在突變之後會變得更穩定，而負值則是不穩定。

Eris 使用美杜莎力場(Medusa force field)[172]，透過蒙特卡羅法(Monte Carlo simulations)針對蛋白質剛性(fixed)與彈性(flexible)的骨架(backbone)、側鏈(side chain)的位向、轉動能量以及鄰近殘基的相互作用力，計算能量的變化值($\Delta\Delta G$)。根據文獻，使用 Eris 預測超過 500 種蛋白質突變的穩定性，結果與實驗數據相比有非常高的準確性[173, 174]。此法定義 $\Delta\Delta G$ 為負值時，表示蛋白質在突變之後會變得更穩定，而正值則是不穩定。

EASE-MM (Evolutionary, Amino acid, and Structural Encodings with Multiple Models)同樣是取自線上資料庫 ProTherm 作為訓練資料的來源，搭配兩種模型並使用支持向量機(SVM)整合之。一種是 EASE-ASA (accessible surface area)，根據蛋白質可以接觸到溶劑的程度分為暴露型(exposed)與包埋型(buried)；另一種則是 EASE-SS (secondary structure)，根據蛋白質的二級結構的比例分為 α -螺旋(alpha-helix)、 β -摺板(beta-sheet)、和無規則捲曲(random coil)等三種類型。當輸入一個單點突變的蛋白質序列時，EASE-MM 會同時使用上述兩種方式，並預測其穩定或不穩定的可能性(probability)，最後計算平均如下公式(1)所示。此法定義 $\Delta\Delta G$ 為正值時，表示蛋白質在突變之後會變得更穩定，而負值則是不穩定。

$$P_{EASE-MM} = \frac{1}{2}(P_{EASE-ASA} + P_{EASE-SS}) \quad (1)$$

PASTA2.0 能根據蛋白質序列來計算蛋白質核心(core)的配對能量勢(pairwise energy potential)，進而預測可能形成類澱粉纖維的區域。在 2.0 版中，加入二級結構以及無規則構形(intrinsic disorder)的計算，提高指認(assignment)出形成類澱粉纖維區域的能力。其原理是計算在 β -strand 上氫鍵(hydrogen bonding)的能量[175]，並使用 DSSP 演算法提供更嚴謹的閾值(threshold)[176]，而參數則是校正自網路伺服器 TESE[177]。

TANGO 是依據 TANGO 演算法計算蛋白質或胜肽(peptide)的 β -聚集(β -aggregation)傾向，包含相空間(phase-space)中所有蛋白質的原態構形與無規則捲曲(random coil)，且每一段序列可以根據波茲曼分布(Boltzmann distribution)來改變狀態。此外，TANGO 同時也將生化反應的環境條件納入考慮，包括 pH 值、溫度與離子強度(ionic strength)的設定，能依據熵值(entropy)將氫鍵作用力與疏水交互作用力(hydrophobic interaction)一起加入計算。雖然本法預測結果與類澱粉纖維不同，但具高度相關。

AGGRESKAN 是利用體內(*in vivo*)實驗結果計算天然胺基酸的聚集傾向，並假設一些短鏈的特定序列容易調控蛋白質聚集的生成。此演算法能找出蛋白質上會因聚集而引發疾病的特定片段，並同時預測突變對蛋白質沉積傾向(*deposition propensity*)的影響。另外，較佳預測結果會依據胺基酸序列的長度，設定不同窗口大小(*window size*)，若胺基酸序列太大而窗口大小不夠大時，則會有很明顯的誤判出現。

WALTZ 使用特殊位置的計分矩陣(*position-specific scoring matrix*)，包括 19 種最適合用來描述形成類澱粉傾向(*amyloid propensity*)的物理性質，搭配 20 種胺基酸在常規化(*normalized*)後的性質，以及類澱粉骨架(*amyloid backbone*)的結構等，藉此預測真實能形成類澱粉纖維的序列，並且避開可能產生不規則聚集(*amorphous aggregate*)的位置。

FoldAmyloid 能根據蛋白質的序列，預測形成類澱粉纖維的可能性，當骨架(*backbone*)間具有氫鍵時，會有比較高的填充密度(*packing density*)，此與類澱粉纖維的形成有高度相關。此計算方式是依據形成氫鍵的能力，分別給予 20 個胺基酸一個特定分數，當連續五個胺基酸的分數超過設定的閾值時，就會被判定有形成類澱粉纖維的可能。各胺基酸的分數如表 2.3.1[155]。

表 2.3.1 FoldAmyloid 演算法中 20 個胺基酸的分數[155]

胺基酸名稱	分數	胺基酸名稱	分數	胺基酸名稱	分數	胺基酸名稱	分數
丙胺酸 (alanine)	19.89	甘胺酸 (glycine)	17.11	甲硫胺酸 (methionine)	24.82	絲胺酸 (serine)	18.19
半胱胺酸 (cysteine)	23.52	組胺酸 (histidine)	21.72	天門冬醯胺 (asparagine)	18.49	蘇胺酸 (threonine)	19.81
天門冬胺酸 (aspartic acid)	17.41	異白胺酸 (isoleucine)	25.71	脯胺酸 (proline)	17.43	纈胺酸 (valine)	23.93
穀胺酸 (glutamic acid)	17.46	離胺酸 (lysine)	17.67	麩醯胺酸 (glutamine)	19.23	色胺酸 (tryptophan)	28.48
苯丙胺酸 (phenylalanine)	27.18	白胺酸 (leucine)	25.36	精胺酸 (arginine)	21.03	酪胺酸 (tyrosine)	25.93

Zygggregator 會根據位置給予每個胺基酸一個聚集分數，並計算每個胺基酸本身與前後三個胺基酸的分數，藉以得出自身聚集傾向(intrinsic aggregation propensity)的值，公式如下式(2)。其中，每個胺基酸的分數計算包括 α -螺旋(α -helix)、 β -摺板(β -sheet)、親疏水性以及帶電性來綜合計算之，如下式(3)。此法能預測蛋白質聚集或是去摺疊的區域，與造成疾病的類澱粉纖維結構有關聯。

$$P_i^{agg} = \frac{1}{7} \sum_{j=-3}^3 p_{i+j}^{agg} + \alpha_{pat} l_i^{pat} + \alpha_{gk} l_i^{gk} \quad (2)$$

$$p_i^{agg} = \alpha_h p_i^h + \alpha_s p_i^s + \alpha_{hyd} p_i^{hyd} + \alpha_c p_i^c \quad (3)$$

AGGRESCAN3D 的理論建立於 AGGRESCAN 的演算法上，但與上述幾種預測聚集或類澱粉纖維的方式有很大的不同，它使用蛋白質的三級結構來計算聚集分數，即需要給予所有原子在空間中的座標才得以計算。另外，此法也將親疏水面積以及蛋白質在溶液中的動態擾動(dynamic fluctuations)納入考量，藉以更精準地預測蛋白質的行為。

吾人將上述多種生物資訊工具之計算方式整理如表 2.3.2。雖然利用生物資訊工具可以有效快速地預測蛋白質在突變前後的穩定性以及聚集傾向，但因不同工具所採用的資料庫(database)以及計算能量來輔助判斷蛋白質穩定性的演算法(algorithm)都不同，使得在結果的預測上可能會有所不同。常見的解決方式除了上述提到的支持向量機(SVM)之外，還能利用分子動力學的模擬，以求得更精確的結果。比起預測熱力學性質的生物資訊工具而言，分子動力學提供了分子真實運動情形的模擬，更有助於了解蛋白質結構的改變，細節於 2.3.2 介紹。不過，由於分子動力學的模擬計算量龐大，為了節省研究時間與經費，在進行分子動力學的計算之前，使用生物資訊工具進行初步的預測仍是不可或缺的。

表 2.3.2 多種生物資訊工具整理一覽表

生物資訊工具	計算方式描述
Istable	透過計算吉布斯自由能(Gibbs free energy)、焓(enthalpy)或是熱容(heat capacity)的變化量來分析蛋白質的穩定性
DUET	根據同源蛋白家族(homologous protein families)的統計數據，以及蛋白質的結構特徵(structural signatures)來預測的穩定性
Eris	針對蛋白質剛性(fixed)與彈性(flexible)的骨架(backbone)、側鏈(side chain)的位向、轉動能量以及鄰近殘基的相互作用力，計算能量的變化值
EASE-MM	根據蛋白質可以接觸到溶劑的程度以及二級結構的比例來分析蛋白質的穩定性
PASTA2.0	根據蛋白質序列來計算蛋白質核心(core)的配對能量勢(pairwise energy potential)，預測可能形成類澱粉纖維的區域
TANGO	依據熵值(entropy)將氫鍵作用力與疏水交互作用力(hydrophobic interaction)計算 β -聚集(β -aggregation)傾向
AGGRESCAN	利用體內(in vivo)實驗結果計算天然胺基酸的聚集傾向，並找出蛋白質上會因聚集而引發疾病的特定片段
WALTZ	使用特殊位置的計分矩陣(position-specific scoring matrix)，包括 19 種最適合用來描述形成類澱粉傾向(amyloid propensity)的物理性質
FoldAmyloid	依據形成氫鍵的能力，分別給予 20 個胺基酸一個特定分數，當連續五個胺基酸的分數超過設定的閾值時，就會被判定有形成類澱粉纖維的可能
Zygggregator	根據位置給予每個胺基酸一個聚集分數，並計算每個胺基酸本身與前後三個胺基酸的分數，計算方式包括 α -螺旋(α -helix)、 β -摺板(β -sheet)、親疏水性以及帶電性等
AGGRESCAN3D	根據蛋白質原子在空間中的座標來計算，並同時考慮親疏水面積以及在溶液中的動態擾動(dynamic fluctuations)，進而分析聚集傾向

2.3.2 分子動力學模擬(Molecular Dynamics Simulations)

分子動力學理論於 1950 年由 Irving 與 Kirkwood 首先提出[178]，其主要是以分子力學為基礎，透過高效能的計算來處理分子的運動，以時間為變項將系統內所有原子的軌跡進行詳細的紀錄，藉以了解微觀情況下，分子真實的運動行為與性質。分子動力學已廣泛應用於生物分子的領域中[179]，如探討蛋白質摺疊行為[180, 181]，酵素的催化行為與機制[182]，或是小分子抑制蛋白質的藥物設計[183]等。

分子動力學模擬的理論主要是建立在古典力學下，根據牛頓第二運動定律來計算系統內原子受力的運動行為[184]，簡要流程如圖 2.3.2。首先，需確定分子的起始構形，即需有每個原子的起始位置座標，並設定起始時間的加速度為 0，另以馬克士威-波茲曼分布(Maxwell - Boltzmann distribution)隨機給予原子起始速度。接著，計算原子間的交互作用力，進一步取得原子在經過特定間隔時間 Δt 後的新位置與新速度。過程中，須符合使用者起始設定的邊界條件(boundary conditions)，且整個系統在定莫耳數、定壓及定溫的環境之下。最後，將時間從特定時間點 t 移至下個時間點 $t+\Delta t$ ，並重複上述步驟直到模擬跑到設定的時間為止。

在上述流程中，最關鍵的步驟是計算原子間的交互作用力，若能越完整描述原子間相互影響的情形，此模擬預測的準確性就可以越高。在分子動力學模擬中，稱為力場(forcefield)的建立。一般而言，原子間的交互作用力可以分為共價鍵結(covalent bonds)作用力與非共價鍵結(non-covalent bonds)作用力，如下式(4)。前者以二次能量函數(quadratic energy function)或是莫耳斯能量勢(Morse potential)來描述鍵能和鍵角，以正弦函數描述兩面角(dihedral angle)；後者以庫倫定律(Coulomb's law)計算靜電作用力，以蘭納-瓊斯勢(Lennard-Jones potential)計算凡得瓦力(van der Waals)。

$$E_{total} = E_{bonded} + E_{nonbonded}$$

$$= (E_{bond} + E_{angle} + E_{dihedral}) + (E_{electrostatics} + E_{van\ der\ Waals})$$

(4)



Simplified schematic of the molecular dynamics algorithm

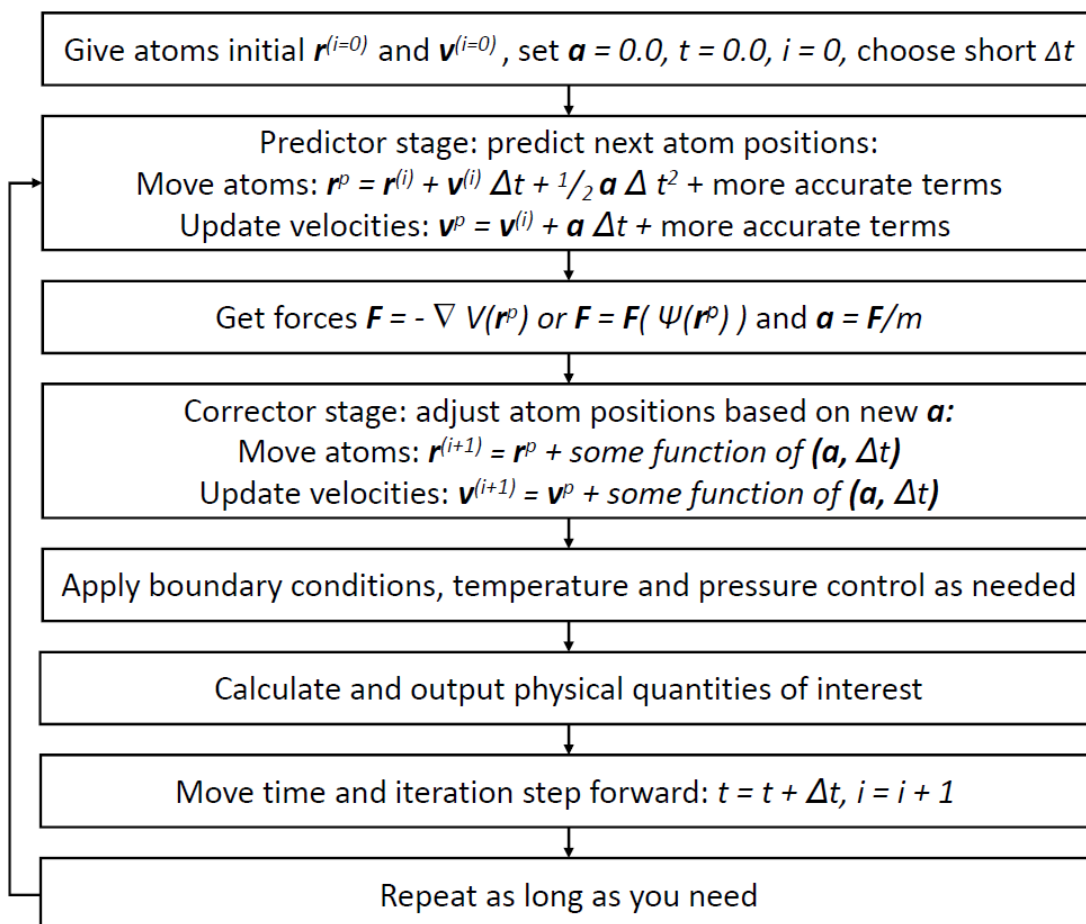


圖 2.3.2 分子動力學模擬演算流程圖

https://en.wikipedia.org/wiki/Molecular_dynamics#/media/File:Molecular_dynamics_algorithm.png

為了更精確描述原子間的作用力，部分領域會將量子力學(quantum mechanics)一併納入計算[185, 186]。此法使用第一原理(*ab initio*)進行計算，根據量子力學的理论，將系統內所有交互作用力，包括原子核與電子等，完全由理論推導而得，不做任何的假設與經驗式的帶入，是一套非常準確的計算方法；而使用半經驗(semi-empirical)的薛丁格方程式(Schrödinger equation)，透過波函數(wave function)來描述電子運動狀態的方法也很常見於模擬相關的研究[187, 188]。但由於上述的

計算需耗費龐大的時間，因此近年來發展出結合量子與分子(quantum mechanics/molecular mechanics, QM/MM)的計算方式，在欲研究的重點區域使用量子計算，而其他部分則使用分子計算，如此能較有效率且又精確地完成模擬，如圖 2.3.3。在 2013 年，Karplus 等人因此研究得到該年的諾貝爾化學獎[189]。

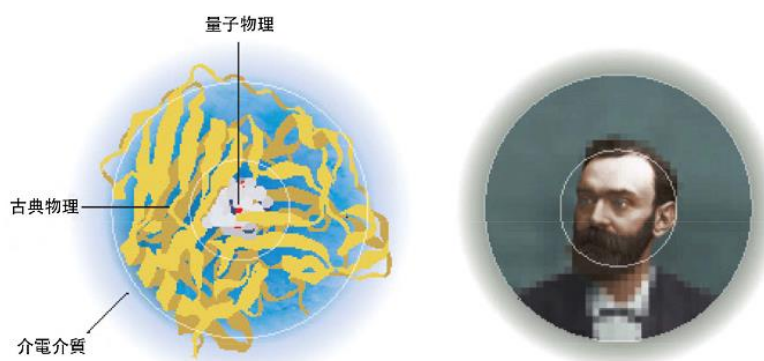


圖 2.3.3 量子/分子模擬示意圖

<http://www.ch.ntu.edu.tw/nobel/2013.html>

目前常見的力場，包括廣泛用於蛋白質或是去氧核糖核酸(DNA)的 AMBER[190]，適用於大分子與小分子的 CHARMM[191]與 CVFF[192, 193]，使用量子/分子(QM/MM)計算各種無機(inorganic)、有機化合物的 COSMOS[194]，以及應用於生物分子系統的 GROMOS[195]。由於不同力場所適用條件以及計算方式都不盡相同，因此根據使用者的需求挑選合適的力場是非常重要的。

為了更接近真實的情況，溶劑模型的加入是必須的，主要計算包括模型產生作用力位置(site)的數目，模型為剛性(rigid)或彈性(flexible)，以及模型極化性(polarization effects)的影響。溶劑模型根據計算方式分為顯性溶劑(explicit solvent)與隱性溶劑(implicit solvent)兩種[130]，前者透過外加溶劑分子至系統中模擬真實溶劑的影響，不過由於溶劑分子的數目太多可能會使得計算速率下降；後者則是使用波松-波茲曼(Poisson-Boltzmann)方程式[196]直接估計溶劑效應，避免在系統中加入溶劑分子，藉此可以使得計算時間縮短。但就準確度而言，還是以顯性溶劑較為準確，也因此較為廣泛地被使用。

本研究使用 Gromacs 套裝軟體來進行分子動力學模擬，由荷蘭格羅寧根大學 (University of Groningen) 所開發，全名為 GRONingen MACHine for Chemical Simulations，主要用來研究蛋白質與核酸等生物分子[197]。此軟體由 C 語言以及 C++ 所撰寫而成，可以在 Linux、OS X、Windows 以及任何 Unix 的作業系統下操作，可以利用平行運算的方式來減少運算的時間，目前最新的版本為 2015 年 11 月釋出的 5.1.1 版。Gromacs 須以輸入指令來執行功能，模擬流程如圖 2.3.4。先取得蛋白質的結構檔，接著以建盒子、建力場、加溶劑、加入離子、能量最小化、系統平衡以及分子動力學模擬的流程進行，示意圖如圖 2.3.5。其中，本研究的力場選用 GROMOS[195]，且配合使用顯性溶劑模型。

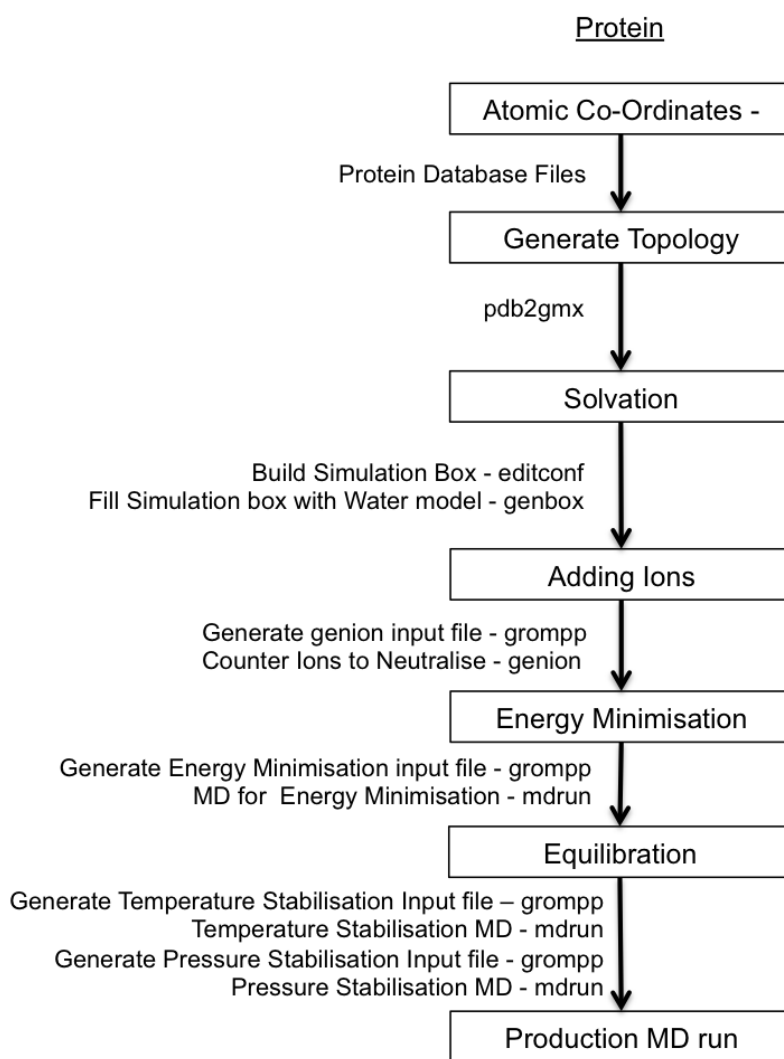


圖 2.3.4 Gromacs 模擬流程圖

<http://2013.igem.org/Team:Manchester/FabProteinModel>

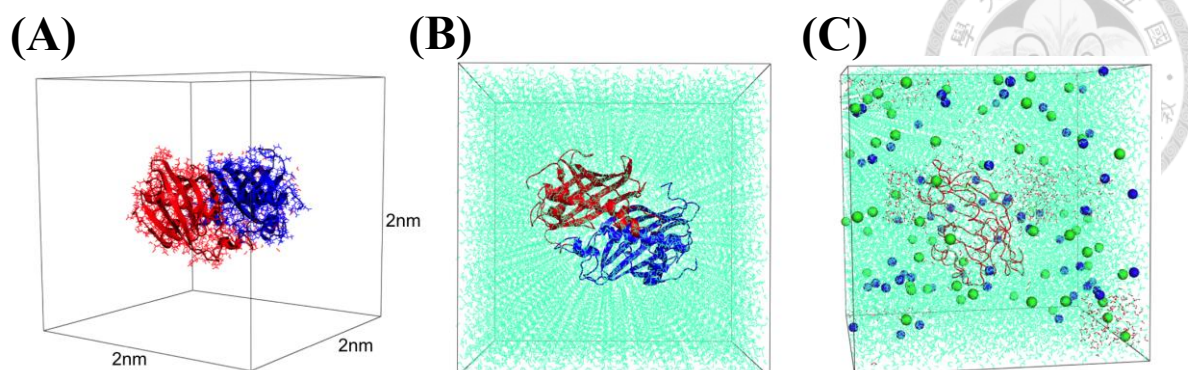


圖 2.3.5 Gromacs 模擬流程示意圖(A)建立 2*2*2 nm³ 的盒子(B)在盒中加入溶劑模型(C)在盒中加入離子(藍:鈉離子,綠:氯離子)

<http://2013.igem.org/Team:Manchester/FabProteinModel>

另外，Gromacs 為了減少計算的複雜度，使用 LINCS (Linear Constraint Solver) 演算法對鍵長做限制[198]，以及 PME (Particle-mesh Ewald) 演算法分類並計算長程靜電作用力[199]。除此之外，由於在分子動力學的系統建立在奈米尺度，當分子在模擬過程中靠近系統邊界時，因受到周圍粒子的作用力較小，而產生邊際效應 (edge effects)，失去其準確性；但若在大系統中，此邊際效應則會非常小。為了減少邊際效應的影響，Gromacs 使用週期性邊界條件 (Periodic Boundary Conditions, PBC) 將邊界複製並無限延伸，使分子在移動到邊界時能完整考慮周圍的作用力 [197]。週期性邊界條件示意圖如圖 2.3.6。

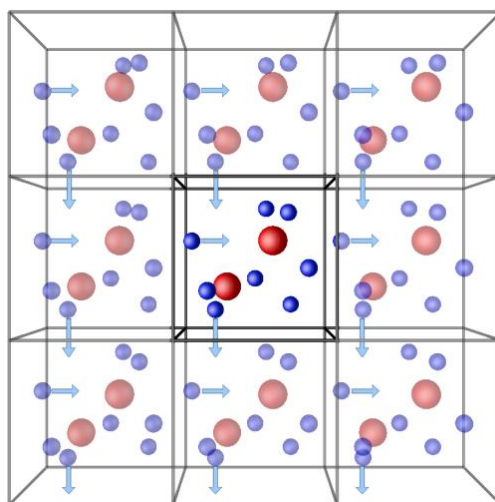


圖 2.3.6 週期性邊界條件示意圖[200]

2.3.3 分子對接(Molecular Docking)

在分子建模(molecular modeling)的領域中，分子對接(molecular docking)能夠預測兩個分子相互鍵結時，產生穩定複合物(complex)的位向(orientation)[201]，如圖 2.3.7。在預測出可能的位向後，可以更進一步探討兩分子間作用力的強度，即兩者鍵結的穩定程度。常見的對接分為三種，第一為有機小分子與蛋白質對接，大多用來進行標靶藥物(target-drug)的研究以及藥物設計(drug design)等[202]；第二種是蛋白質與蛋白質對接[203]，除了可以預測酵素(enzyme)與抗原抗體的複合物之外[204, 205]，也有研究用來模擬蛋白質的聚集行為[206]；第三種是蛋白質與 DNA 對接，用來探討基因調節或是 DNA 修復等議題[207]。本研究主要採用上述第一種小分子與蛋白質對接的方式，來分析多種小分子與 HyDC 對接的位向與作用力。

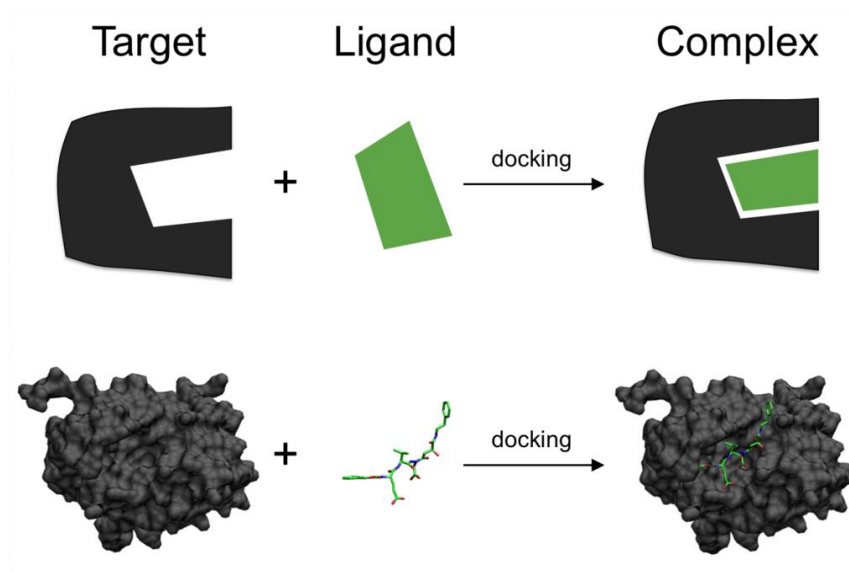


圖 2.3.7 分子對接模型示意圖

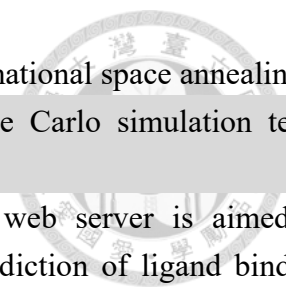
[https://en.wikipedia.org/wiki/Docking_\(molecular\)#/media/File:Docking_representation_2.png](https://en.wikipedia.org/wiki/Docking_(molecular)#/media/File:Docking_representation_2.png)

使用電腦軟體來進行對接預測的方式主要分為兩種，第一種是根據蛋白質與小分子的形狀，使用配對技術來分析兩者形狀的互補性(complementary)[208]，雖然此法快速且健全(robust)，但無法準確地預測結構的動態變化；第二種是計算對接後的配對能量(pairwise interaction energies)[209]，以彈性(flexibility)模型取代剛性

(rigid)模型，使預測更為準確，但缺點是較為耗時且昂貴，以及需要較高的技術條件，如網格導向(grid-based)的技術、最佳化的方法(optimization methods)與電腦的執行速度等。在預測出大量可能對接的結構後，不同的軟體有不同的評分方式(score function)來進行排序，提供使用者挑選較有可信度的對接結構。常見的對接模擬軟體整理如表 2.3.3。

表 2.3.3 常見的對接模擬軟體

軟體名稱	出版年份	軟體描述
AADS[210]	2011	Automated active site detection, docking, and scoring(AADS) protocol for proteins with known structures based on Monte Carlo Method
AutoDock[211]	1990	Automated docking of ligand to macromolecule by Lamarckian Genetic Algorithm and Empirical Free Energy Scoring Function
AutoDock Vina[212]	2010	New generation of AutoDock
BetaDock[213]	2011	Based on Voronoi Diagram
Blaster[214]	2009	Combines ZINC databases with DOCK to find ligand for target protein
BSP-SLIM[215]	2012	A new method for ligand-protein blind docking using low-resolution protein structures
DARWIN[216]	2000	Prediction of the interaction between a protein and another biological molecule by genetic algorithm
EADock[217]	2007	Based on evolutionary algorithms
EUDOC[218]	2001	Program for identification of drug interaction sites in macromolecules and drug leads from chemical databases
FlexAID[219]	2015	Target side-chain flexibility and soft scoring function, based on surface complementarity
FLIPDock[220]	2007	Genetic algorithm based docking program using FlexTree data structures to represent a protein-ligand complex
GEMDOCK[221]	2004	Generic Evolutionary Method for molecular docking
HADDOCK[222]	2003	Makes use of biochemical and/or biophysical interaction data such as chemical shift perturbation data resulting from NMR titration experiments, mutagenesis data or bioinformatic predictions. Developed for protein-protein docking, but can also be applied to protein-ligand docking.
LigandFit[223]	2003	CHARMm based docking program



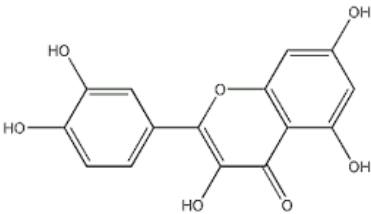
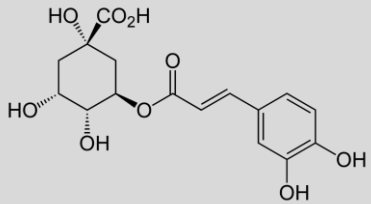
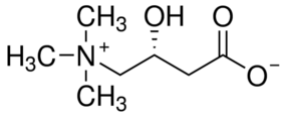
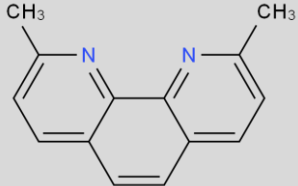
LigDockCSA[224]	2011	Protein-ligand docking using conformational space annealing
MCDOCK[225]	1999	Based on a non-conventional Monte Carlo simulation technique
MEDock[226]	2005	Maximum-Entropy based Docking web server is aimed at providing an efficient utility for prediction of ligand binding site
MolDock[227]	2006	Based on a new heuristic search algorithm that combines differential evolution with a cavity prediction algorithm
MS-DOCK[228]	2008	Multi-stage docking/scoring protocol
ParDOCK[229]	2007	All-atom energy based Monte Carlo, rigid protein ligand docking
PatchDock[230]	2005	The algorithm carries out rigid docking, with surface variability/flexibility implicitly addressed through liberal intermolecular penetration
PRODOCK[231]	1999	Based on Monte Carlo method plus energy minimization
PSI-DOCK[232]	2006	Pose-Sensitive Inclined (PSI)-DOCK
PythDock[233]	2011	Heuristic docking program that uses Python programming language with a simple scoring function and a population based search engine
Q-Dock[234]	2008	Low-resolution flexible ligand docking with pocket-specific threading restraints
rDock[235]	2013	HTVS of small molecules against proteins and nucleic acids
SODOCK[236]	2007	Swarm optimization for highly flexible protein-ligand docking
Surflex-Dock[237]	2007	Based on an idealized active site ligand (a protomol)
SwissDock[238]	2011	Webservice to predict interaction between a protein and a small molecule ligand
VoteDock[239]	2011	Consensus docking method for prediction of protein-ligand interactions

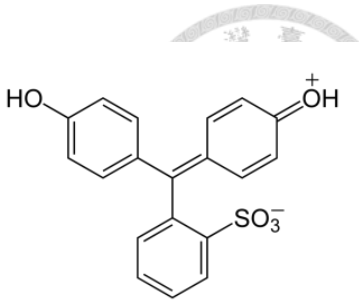
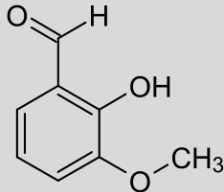
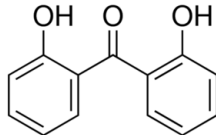
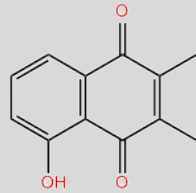
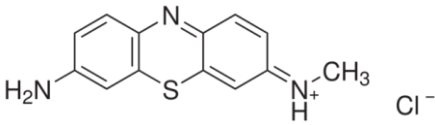
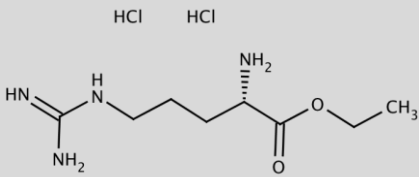
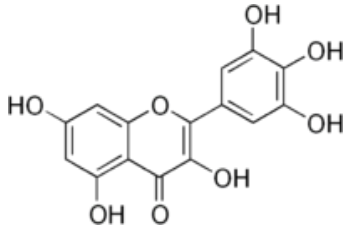
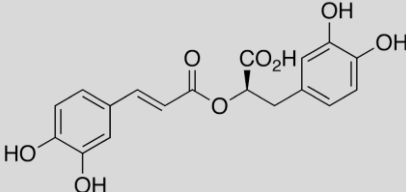
本研究使用 Discovery Studio 2016 版內建分子對接功能 CDOCKER[240]，能精細地進行多種小分子與蛋白質的對接模擬。CDOCKER 的原理是將系統切割成許多網格點，並使小分子在其中進行隨機翻轉，最後將所有可能性以能量(CDocker Energy)排序之。因此，在對接前並不一定要先知道蛋白質鍵結位置(binding site)。但若先預測出蛋白質的活性位置(active site)，在後續的對接預測就能更為快速且

準確。其中，CDOCKER 的力場是採用 CHARMM，並假設蛋白質受體(receptor)是剛性的，而小分子配體(ligand)為彈性的。

承 2.2.3 節所提，本研究用分子對接的方式預測多種小分子與 H γ DC 對接的位向與作用力，進一步使用分子動力學模擬探討小分子的鍵結是否能穩定 H γ DC 之結構，藉以判斷該小分子是否具有作為抑制白內障藥物的潛力。本研究所挑選的小分子都被過去的研究指出有具有抑制蛋白質聚集的能力，且其中部分研究與白內障有關，包括槲皮素(Quercetin)[241, 242]、綠原酸(Chlorogenic acid)[243, 244]與楊梅素(Myricetin)[245, 246]等，詳細資料如表 2.3.4。

表 2.3.4 抑制蛋白質聚集之小分子一覽表

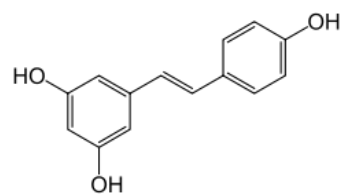
小分子名稱	分子式	分子量 (g/mol)	結構
槲皮素 (Quercetin) [241, 242]	C ₁₅ H ₁₀ O ₇	302.24	
氯原酸 (Chlorogenic acid) [243, 244]	C ₁₆ H ₁₈ O ₉	354.31	
肉鹼 (L-carnitine inner salt) [247, 248]	C ₇ H ₁₅ NO ₃	161.20	
2,9-二甲基-4,10-菲羅 (Neocuproine free base) [246]	C ₁₄ H ₁₂ N ₂	208.25	

酚紅 (Phenol red) [246]	$C_{19}H_{14}O_5S$	354.38	
鄰香蘭素 (Ortho-vanillin) [246, 249]	$C_8H_8O_3$	152.15	
2,2-,二羥基二苯甲酮 (2,2-dihydroxy-Benzophenone) [246, 250]	$(HOC_6H_4)_2CO$	214.22	
5-羥基-1,4-萘醌 (5-hydroxy-1,4-naphthoquinone) [246, 251]	$C_{10}H_6O_3$	174.15	
天青藍 C (Azure C) [246, 250]	$C_{13}H_{12}ClN_3S$	277.77	
精氨酸乙基醚 (L-arginine ethyl ester dihydrochloride) [252, 253]	$C_8H_{18}N_4O_2 \cdot 2HCl$	275.18	
楊梅素 (Myricetin) [245, 246]	$C_{15}H_{10}O_8$	318.24	
迷迭香酸 (Rosmarinic acid) [254, 255]	$C_{18}H_{16}O_8$	360.32	

白藜蘆醇
(Resveratrol)
[98]

$C_{14}H_{12}O_3$

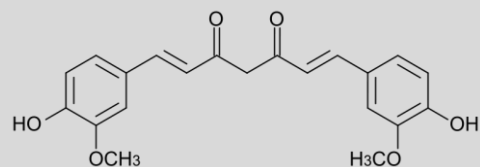
228.25



薑黃素
(Curcumin)
[255, 256]

$C_{21}H_{20}O_6$

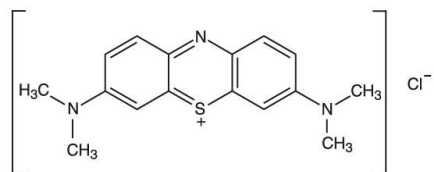
368.39



亞甲基藍
(Methylene blue)
[257-259]

$C_{16}H_{18}ClN_3S$

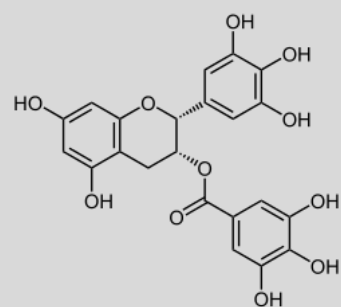
319.85



兒茶素
(Epigallocatechin gallate)
[260-262]

$C_{22}H_{18}O_{11}$

458.37



2.4 模擬分析方法原理介紹

在分子動力學模擬的過程中，蛋白質於特定溫度與壓力的條件下，因自身運動與受周圍分子作用力的影響，其結構會隨時間而有所改變，因此分析需由隨時間改變之各原子座標點所構成的軌跡檔來進行。本節主要介紹本研究中所用到的方法，如方均根變異(Root mean square deviation, RMSD)、方均根擾動(Root mean square fluctuation, RMSF)、原態接觸(Native contact)、蛋白質二級結構分析(Dictionary of protein secondary structure, DSSP)等。

2.4.1 方均根變異(Root mean square deviation, RMSD)分析

為了比較蛋白質結構的差異，RMSD 是最普遍且常見的方式。此法是將兩組欲比較結構中的每個原子，計算其座標的偏移量，並平方加總再開根號所得，如式(5)所示，其中 M 為總分子量， m_i 為分子內 i 的原子量， $r_i(t_1)$ 為 i 原子在時間 t_1 時的位置向量， $r_i(t_2)$ 為 i 原子在時間 t_2 時的位置向量。又因每個原子的質量不同，對蛋白質的影響也不同，因此需要以質量加權進行計算。計算結果的數值越大，表示兩組的結構相差越大，反之亦然。另外，由於分子在運動過程中有三個方向的平移(translation)與轉動(rotation)，所以在計算前需將結構進行最適化疊合(best fit)。

$$RMSD(t_1, t_2) = \left(\frac{1}{M} \sum_{i=1}^N m_i |r_i(t_1) - r_i(t_2)|^2 \right)^{\frac{1}{2}} \quad (5)$$

2.4.2 方均根擾動(Root mean square fluctuation, RMSF)分析

RMSF 與 RMSD 的原理類似，但並非探討蛋白質結構隨時間變化的情形，而是分析每個殘基(或特定組別的原子)在模擬過程中的擾動情形，如式(6)所示，其中 T 為模擬總時間， t_i 為時間， $r_i(t_i)$ 為時間 t_i 時的位置向量。計算結果的數值越大，表示在模擬過程中此殘基的擾動最大，反之則小。同樣地，在分析之前需將蛋白質的結構進行最適化疊合。



$$RMSF = \left(\frac{1}{T} \sum_{t_i=1}^T r_i(t_i) - r_i(t_2) \right)^{\frac{1}{2}} \quad (6)$$

2.4.3 原態接觸(Native contact)率分析

原態接觸能用來分析蛋白質摺疊與去摺疊的情形，先計算蛋白質原態的原子接觸組數，並將此接觸情形定義為 1，再分別計算在不同時間點下蛋白質的原子接觸組數，兩者相除可得到接觸百分率[263]，其值會落在 0~1 間，其中接觸的定義為兩原子(i,j)的距離小於 d_{ij} 小於截斷半徑 R_c 。當值越接近 1 時，表示其結構與原態較相近；若值越接近 0，表示結構與原態差異較大，即其結構較為開展。

2.4.4 蛋白質二級結構分析(Dictionary of protein secondary structure, DSSP)

此法能根據蛋白質特定區間的氫鍵有無與幾何結構來判斷蛋白質二級結構的型態[176]，其中氫鍵的有無是由靜電作用力來計算，計算公式如(7)，設定碳、氧、氫、氮的部分電荷分別為 $q_1=0.42e$ 、 $-q_1=-0.42e$ 、 $q_2=0.2e$ 、 $-q_2=-0.2e$ ，且因次係數 $f=332$ 。一般正常的氫鍵的鍵能約為-3 kcal/mol，當計算出的鍵能小於-0.5 kcal/mol 時，表示有氫鍵存在。

$$E = q_1 q_2 \left(\frac{1}{r_{ON}} + \frac{1}{r_{CH}} - \frac{1}{r_{OH}} - \frac{1}{r_{CN}} \right) f \quad (7)$$

DSSP 根據上述的計算方式，可以將二級結構分成八類，並使用特定的英文字母與顏色來表示，如表 2.4.1。

表 2.4.1 DSSP 二級結構定義相關資料[176]

二級結構名稱	DSSP 代碼	顏色	定義
3_{10} -helix	G	灰色	第 i 個殘基與第 i+3 個殘基形成氫鍵並重複序列。
α -helix	H	藍色	第 i 個殘基與第 i+4 個殘基形成氫鍵並重複

序列。			
5-helix	I	紫色	第 i 個殘基與第 i+5 個殘基形成氫鍵並重複序列。
β -bridge	B	黑色	兩殘基鏈 i-1、i、i+1 及 j-1、j、j+1 間存在氫鍵。
β -sheet	E	紅色	一個以上的 β -bridge 結構以相同殘基組合所形成。
turn	T	黃色	殘基 i 的 CO(i)與 NH(i+n)之間存在有氫鍵則定義為轉角，其中 n 可為 3、4 或 5。
bend	S	綠色	Ca(i)-Ca(i-2)與 Ca(i+2)-Ca(i)之間的夾角大於 70°。
random coil	L	白色	不屬於上述七種定義方式即屬此種。

第三章 實驗儀器與步驟



3.1 模擬系統與軟體

3.1.1 模擬系統

Gromacs 4.5.3: Ubuntu 10.04.4 LTS

Discovery studio 2016: Windows 7

3.1.2 生物資訊預測工具

名稱	用途
Istable[141]	蛋白質穩定性預測
DUET[142]	蛋白質穩定性預測
Eris[143-145]	蛋白質穩定性預測
EASE-MM[146, 147]	蛋白質穩定性預測
PASTA2.0[148]	易形成類澱粉纖維區域之預測
TANGO[149-151]	蛋白質 β -聚集預測
AGGRESCAN[152]	蛋白質聚集傾向預測
WALTZ[153, 154]	易形成類澱粉纖維區域之預測
FoldAmyloid[155]	易形成類澱粉纖維區域之預測
Zygggregator[156-159]	蛋白質聚集傾向預測
Aggrescan3D[160]	蛋白質聚集傾向預測

3.1.3 模擬軟體



軟體名稱	用途
Gromacs 4.5.3[264-267]	1. 分子動力學模擬 2. 模擬數據分析
BIOVIA Discovery studio 2016	1. 分子視覺化圖形呈現 2. 蛋白質聚集分數(aggregation score) 預測
Visual Molecular Dynamics (VMD) 1.9.1[268]	1. 分子視覺化圖形呈現 2. 分子動力學模擬軌跡影片製作
MDAnalysis 0.7.4[269]	1. 模擬數據分析(native contact)

3.2 生物資訊預測工具(Bioinformatics prediction tools)

本研究使用多種生物資訊軟體預測蛋白質的穩定性與聚集傾向，以下分別介紹其操作步驟。



3.2.1 Istable[141]

1. 進入 Istable 首頁(<http://predictor.nchu.edu.tw/iStable/>)，界面如圖 3.2.1 所示。

Running iStable from Protein [Structure | **Sequence** | About iStable | Help]

iStable

Integrated predictor for protein **stability** change upon single mutation

PDB ID	Chain	Wild-type	Position	Mutant	Temperature (°C, 0-100)	pH (0-14)

Protein sequence(no headers)

圖 3.2.1 Istable 界面圖

2. 選擇「Structure」的模式，輸入蛋白質的 PDB ID「1hk0」，以及溫度與 pH 值後，點選 next。
3. 選擇胺基酸編號，並設定欲突變的目標胺基酸，如圖 3.2.2，點選提交。

iStable

Integrated predictor for protein **stability** change upon single mutation

Chain	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
SEQuence	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
Residue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
PDBIndex	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

Chain	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
SEQuence	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
Residue	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
PDBIndex	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60

Chain	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
SEQuence	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
Residue	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
PDBIndex	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90

Chain	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
SEQuence	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
Residue	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
PDBIndex	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120

Chain	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
SEQuence	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
Residue	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
PDBIndex	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150

Chain	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
SEQuence	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
Residue	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
PDBIndex	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180

圖 3.2.2 Istable 選擇突變胺基酸頁面示意圖

4. 計算完畢後，系統會輸出如圖 3.2.3 之結果，自行儲存 $\Delta\Delta G$ 變化的數值。

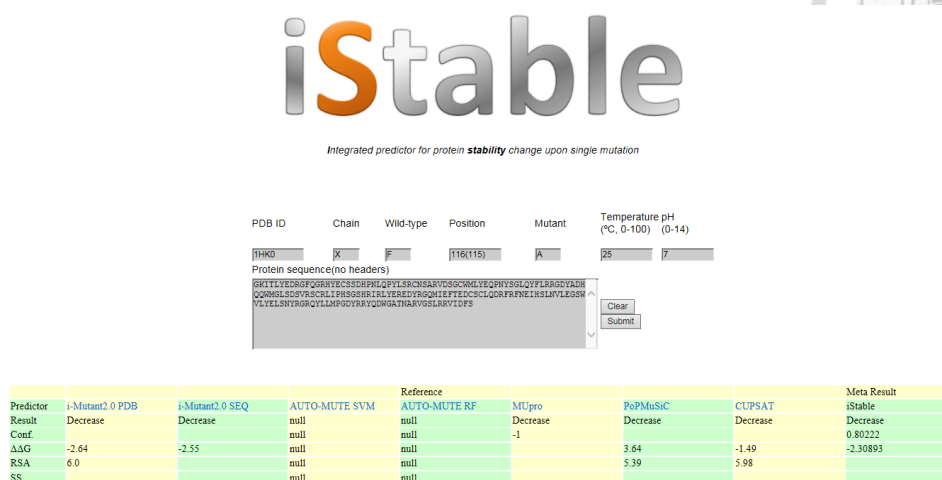


圖 3.2.3 Istable 結果示意圖

3.2.2 DUET[142]

1. 進入 DUET 首頁(<http://structure.bioc.cam.ac.uk/duet>)，點選「Run Predictions」進入選單頁面，如圖 3.2.4 所示。

圖 3.2.4 DUET 選單頁面圖

2. 首先輸入蛋白質的結構檔或是 PDB code，接著在「Single mutation」處輸入欲突變位置與鍵的名稱，最後點選提交。
3. 計算完成後，結果包括 $\Delta\Delta G$ 變化值與結構圖，如圖 3.2.5 所示。

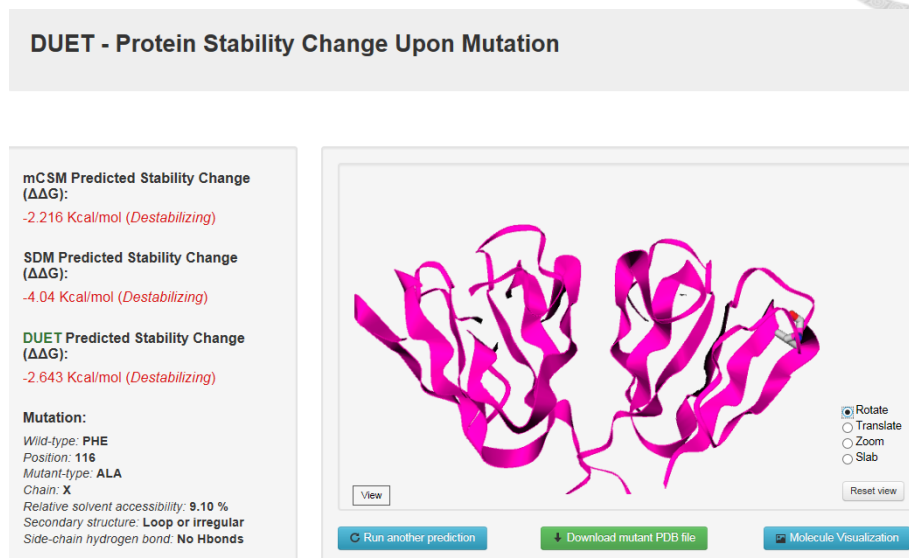


圖 3.2.5 DUET 結果示意圖

3.2.3 Eris[143-145]

1. 進入 Eris 首頁(<http://troll.med.unc.edu/eris/login.php>)，如圖 3.2.6 所示。

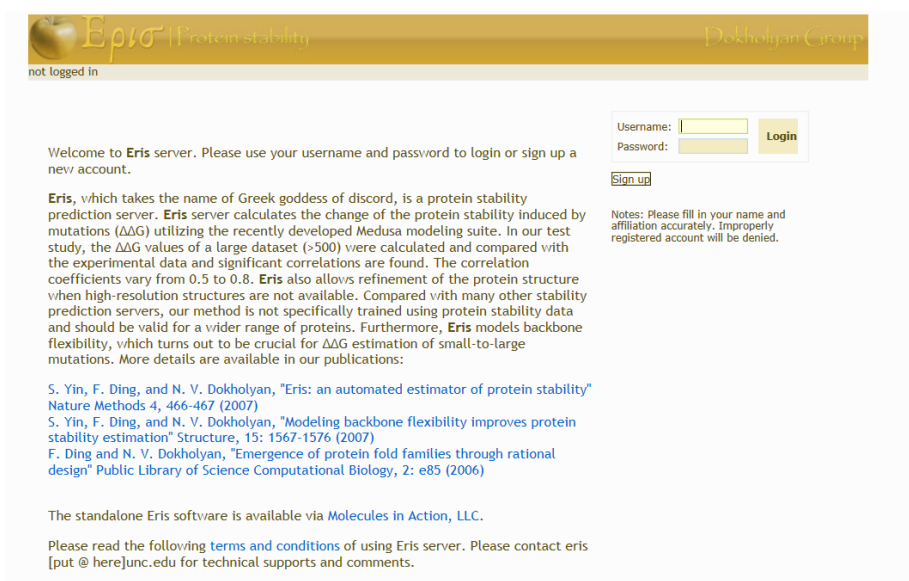


圖 3.2.6 Eris 首頁圖

2. 使用需先註冊，註冊後便可登入，登入後點選「Submit a Task」，輸入 PDB code 或是蛋白質結構檔即可進入下一步。
3. 選擇欲突變之目標胺基酸，點選下一步，如圖 3.2.7。



Submit a task - Step 2.

Click on residue sites and choose mutations.

1 G K I T L Y E D R G F Q G R H Y E C S S

21 D H P N L Q P Y L S R C N S A R V D S G

41 C W M L Y E Q P N Y S G L Q Y F L R R G

61 D Y A D H Q Q W M G L S D S V R S C R L

81 I P H S G S H R I R L Y E R E D Y R G Q

101 M I E F T E D C S C L Q D R F R F N E I

121 H S L N V L E G S W V L Y E L S N Y R G

141 R Q Y L L M P G D Y R R Y Q D W G A T N

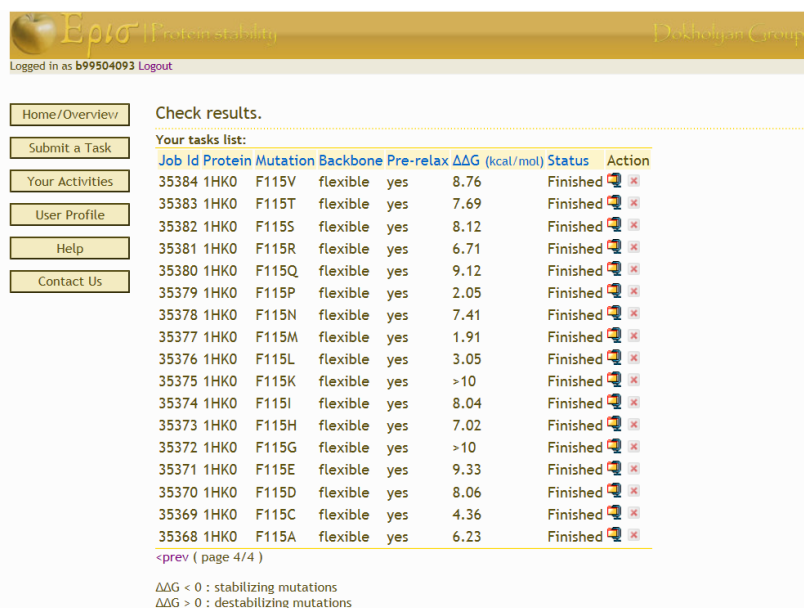
161 A R V G S L R R V I D F S

mutation: F115A

Next >>

圖 3.2.7 Eris 選擇突變胺基酸頁面示意圖

4. 在預測方法選擇「Flexible Backbone」並勾選「Backbone Pre-relaxation」，以得更精準的預測結果，點選「Next」開始預測。
5. 在「Your Activities」可以查看目前進度，如已經預測完成會顯示「Finished」，並顯示 $\Delta\Delta G$ 值，如圖 3.2.8 所示。



Check results.

Your tasks list:

Job Id	Protein	Mutation	Backbone	Pre-relax	$\Delta\Delta G$ (kcal/mol)	Status	Action
35384	1HK0	F115V	flexible	yes	8.76	Finished	
35383	1HK0	F115T	flexible	yes	7.69	Finished	
35382	1HK0	F115S	flexible	yes	8.12	Finished	
35381	1HK0	F115R	flexible	yes	6.71	Finished	
35380	1HK0	F115Q	flexible	yes	9.12	Finished	
35379	1HK0	F115P	flexible	yes	2.05	Finished	
35378	1HK0	F115N	flexible	yes	7.41	Finished	
35377	1HK0	F115M	flexible	yes	1.91	Finished	
35376	1HK0	F115L	flexible	yes	3.05	Finished	
35375	1HK0	F115K	flexible	yes	>10	Finished	
35374	1HK0	F115I	flexible	yes	8.04	Finished	
35373	1HK0	F115H	flexible	yes	7.02	Finished	
35372	1HK0	F115G	flexible	yes	>10	Finished	
35371	1HK0	F115E	flexible	yes	9.33	Finished	
35370	1HK0	F115D	flexible	yes	8.06	Finished	
35369	1HK0	F115C	flexible	yes	4.36	Finished	
35368	1HK0	F115A	flexible	yes	6.23	Finished	

<prev (page 4/4)

$\Delta\Delta G < 0$: stabilizing mutations
 $\Delta\Delta G > 0$: destabilizing mutations

圖 3.2.8 Eris 結果示意圖

3.2.4 EASE-MM[146, 147]

1. 進入 EASE-MM 首頁(<http://sparks-lab.org/server/ease/protein/>)，點選「EASE-MM Protein」進入選單頁面，如圖 3.2.9。

圖 3.2.9 EASE-MM Protein 選單頁面圖

2. 輸入電子信箱與工作名稱，並照規定格式輸入蛋白質序列與欲突變之胺基酸編號，點擊提交。
3. 模擬過程需等待 30 分鐘以上，結果會直接顯示於頁面，如圖 3.2.10 所示。

EASE-MM results for job: p0310

Results in the EASE-MM query format: [results_query.txt](#)

Results in a tab-delimited format: [results.txt](#)

The query submitted for this job: [query](#)

Please see the description of EASE-MM's [output format](#).

Sequence name	Mutation	$\Delta\Delta G_u$ [2]	Stability class [2]	SS [2]	rASA [2]			
my_short_seq	G4A	-0.6435	likely destabilising	C	0.58	PSSM	SPIDER	SPINE-D
my_short_seq	A40G	-1.7756	destabilising	E	0.19	PSSM	SPIDER	SPINE-D
my_short_seq	L16V	-0.9270	likely destabilising	E	0.12	PSSM	SPIDER	SPINE-D

圖 3.2.10 EASE-MM 結果示意圖

3.2.5 PASTA2.0[148]

1. 進入 PASTA2.0 首頁(<http://protein.bio.unipd.it/pasta2/index.html>)，頁面顯示如圖 3.2.11 所示。

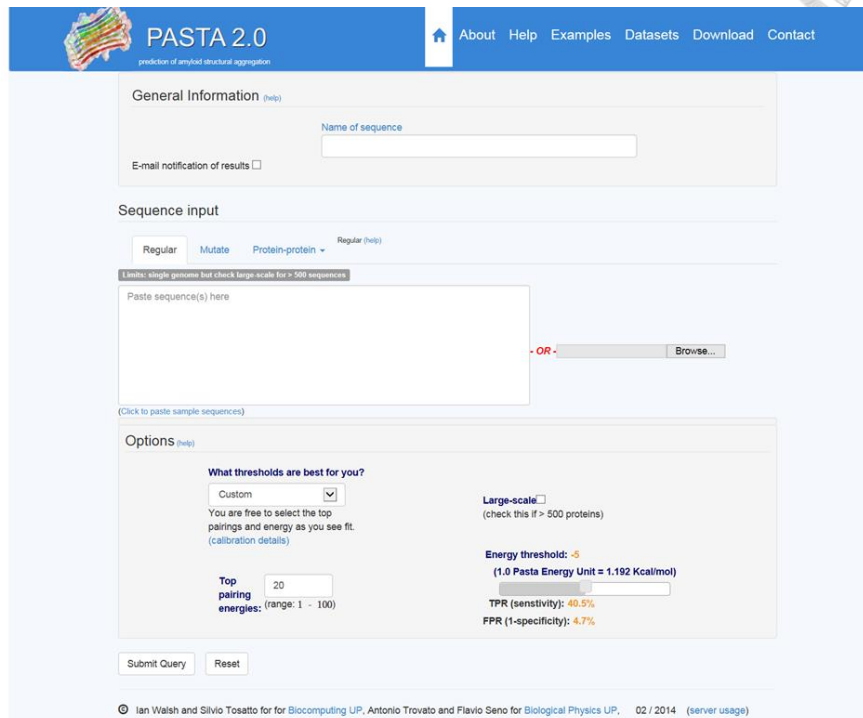


圖 3.2.11 PASTA2.0 首頁示意圖

2. 選擇欲執行的模式，若為原態蛋白質則選「Regular」，若為突變種則選擇「Mutate」。輸入蛋白質序列並設定閾值為「Region (90% spec)」，點選提交。
3. 計算完成後，結果會顯示在頁面。

3.2.6 TANGO[149-151]

1. 進入 TANGO 首頁(<http://tango.crg.es/>)，需先註冊並登入才能使用。在「Toolbox」功能區選擇「Calculation」進入頁面，如圖 3.2.12。
2. 依序輸入蛋白質序列，選擇 N 端與 C 端的保護情形，pH 值、溫度、離子強度以及蛋白質的濃度，並點選計算。
3. 計算完成後，結果呈現如圖 3.2.13，如有需要，則可以在頁面直接下載結果的數據。



A computer algorithm for prediction of aggregating regions in unfolded polypeptide chains

b99504093

Tango

About

References

Handbook

Examples

Support

Contact

Related Sites

Toolbox

Calculation

Download

User data

Calculation

Amino acid sequence(s) :
(Capital letters, separated by carriage returns.
Max. 500 characters and 10 sequences)

Parameters and conditions :
Nterm protected :
Cterm protected :
pH (0 - 12) :
Temperature [K] (273 - 400) :
Ionic strength [M] (0 - 5) :
Concentration [M] (0.00001 - 5) :

圖 3.2.12 TANGO 輸入頁面圖

Tango Result :

(Note : If you download a file : first save the file to disk, then open the file.)

The tendency for B-sheet aggregation is given by the Agg parameter.
The tendency for a-helix aggregation is given by Helagg.
However, the user must know that the parameter that is more reliable is the Agg one.

Sequence 1									
sequence1.xls		sequence1.csv		sequence1.txt					
AGG	26.0497	AMYLO	448324	TURN	78.9706	HELIX	31.4953	HELAGG	0
BETA	552.845								
Position	Amino acid	Beta sheet	Beta turn	Alpha helix	Beta aggregation	Helix aggregation			
01	G	0.0	0.0	0.125	0.000	0.000			
02	K	5.6	0.0	0.125	0.000	0.000			
03	I	16.3	0.0	0.125	0.000	0.000			
04	T	17.3	0.0	0.125	0.000	0.000			
05	L	12.2	0.0	0.125	0.000	0.000			
06	Y	2.2	0.3	0.125	0.000	0.000			
07	E	0.9	0.3	0.000	0.000	0.000			
08	D	0.7	0.5	0.000	0.000	0.000			
09	R	0.1	0.5	0.000	0.000	0.000			
10	G	0.2	0.3	0.000	0.000	0.000			
11	F	0.7	0.5	0.000	0.000	0.000			
12	Q	0.6	0.4	0.000	0.000	0.000			
13	G	0.5	0.5	0.000	0.000	0.000			
14	R	1.8	0.5	0.000	0.000	0.000			
15	H	5.6	0.2	0.000	0.000	0.000			
16	Y	13.3	0.1	0.000	0.000	0.000			
17	E	12.2	0.0	0.000	0.000	0.000			

圖 3.2.13 TANGO 結果示意圖

3.2.7 AGGRESCAN[152]

1. 進入 AGGRESCAN 首頁 (<http://bioinf.uab.es/aggrescan/>)，如圖 3.2.14。依照 FASTA 的格式，輸入蛋白質序列，點選提交，即可開始計算熱點區域。
2. 計算完成結果如圖 3.2.15，紅色字樣表示預測出的熱點區域，如有需要，可以直接下載結果的數據。

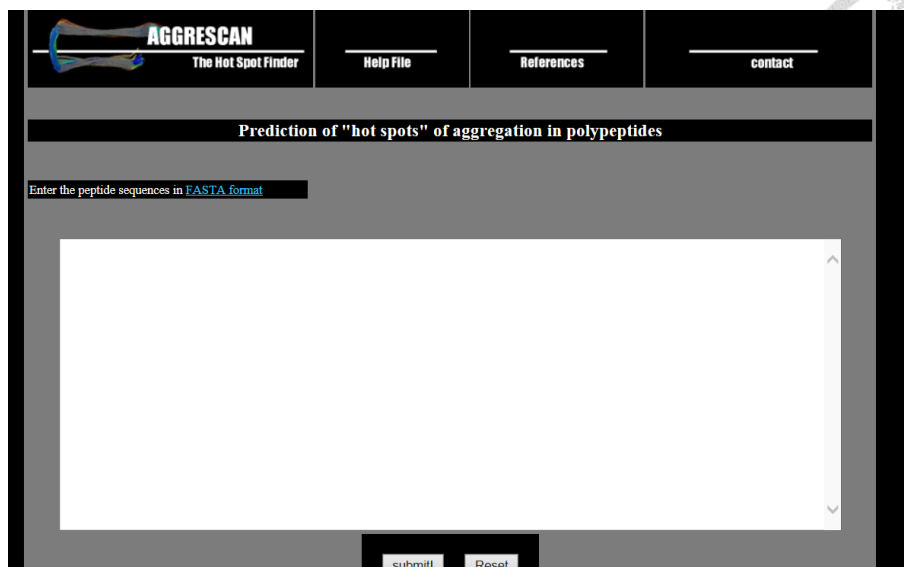


圖 3.2.14 AGGRESCAN 首頁示意圖

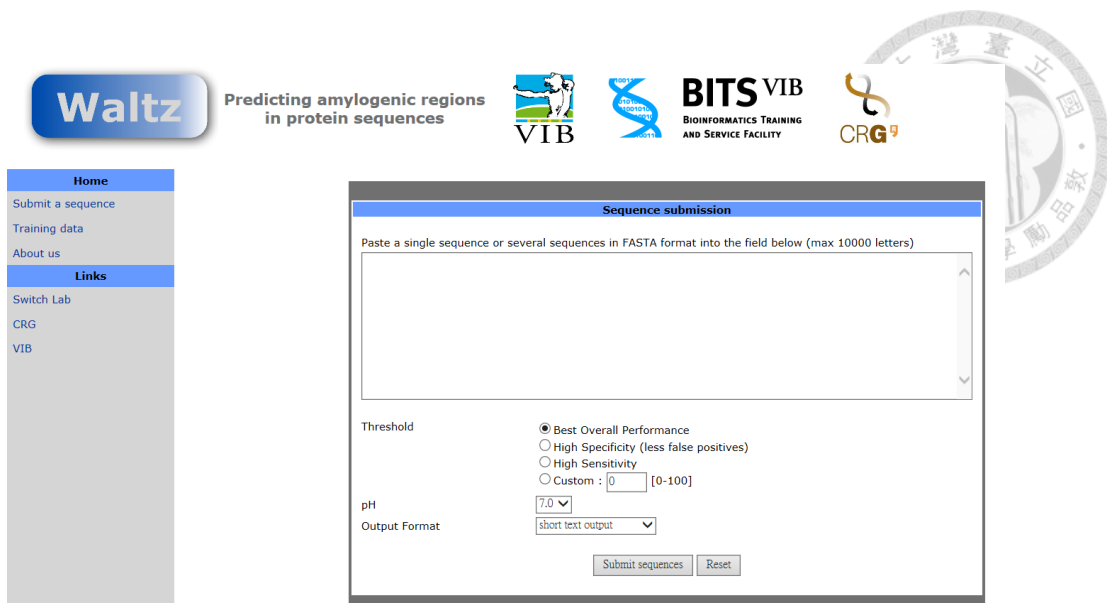
Sequence Name:	HGDC	Average over all sequences
Graphics:		
a3v Sequence Average (a3vSA):	-0.135	-0.135
Number of Hot Spots (nHS):	6	6.000
Normalized nHS for 100 residues (NnHS):	3.468	3.468
Area of the profile Above Threshold (AAT):	14.713	14.713
Total Hot Spot Area (THSA):	12.064	12.064
Total Area (TA):	-20.734	-20.734
AAT per residue (AATr):	0.085	0.085
THSA per residue (THSAr):	0.070	0.070
Normalized a3v Sequence Sum for 100 residues (Na3vSS):	-13.9	-13.900
Sorted by Na3vSS		
HGDC -13.90		

#	AA	a3v	HSA	NHSA	a3vAHS
1	G	0.159	0.858	0.172	0.134
2	R	0.159	0.858	0.172	0.134
3	I	0.159	0.858	0.172	0.134
4	T	0.189	0.858	0.172	0.134
5	L	0.003	0.858	0.172	0.134
6	Y	-0.041	0.000	0.000	0.000
7	E	-0.378	0.000	0.000	0.000
8	D	-0.104	0.000	0.000	0.000
9	R	-0.477	0.000	0.000	0.000
10	G	-0.719	0.000	0.000	0.000
11	F	-0.695	0.000	0.000	0.000
12	Q	-0.580	0.000	0.000	0.000
13	G	-0.237	0.000	0.000	0.000
14	R	-0.363	0.000	0.000	0.000
15	H	-0.527	0.000	0.000	0.000
16	Y	-0.393	0.000	0.000	0.000

圖 3.2.15 AGGRESCAN 計算結果示意圖

3.2.8 WALTZ[153, 154]

1. 進入 WALTZ 首頁(<http://waltz.switchlab.org/>)，如圖 3.2.16 所顯示。
2. 按照 FASTA 的格式輸入蛋白質序列，在閾值選擇「Best Overall Performance」，並輸入 pH 值以及輸出的格式，即可點選提交。



Waltz Predicting amylogenic regions in protein sequences

Home
Submit a sequence
Training data
About us

Links
Switch Lab
CRG
VIB

Sequence submission

Paste a single sequence or several sequences in FASTA format into the field below (max 10000 letters)

Threshold
☒ Best Overall Performance
☐ High Specificity (less false positives)
☐ High Sensitivity
☐ Custom : [0] [0-100]

pH [7.0]

Output Format [short text output]

Submit sequences Reset

圖 3.2.16 Waltz 首頁示意圖

3. 計算結束後，會得出與類澱粉纖維形成相關的區域，結果如圖 3.2.17 所示。

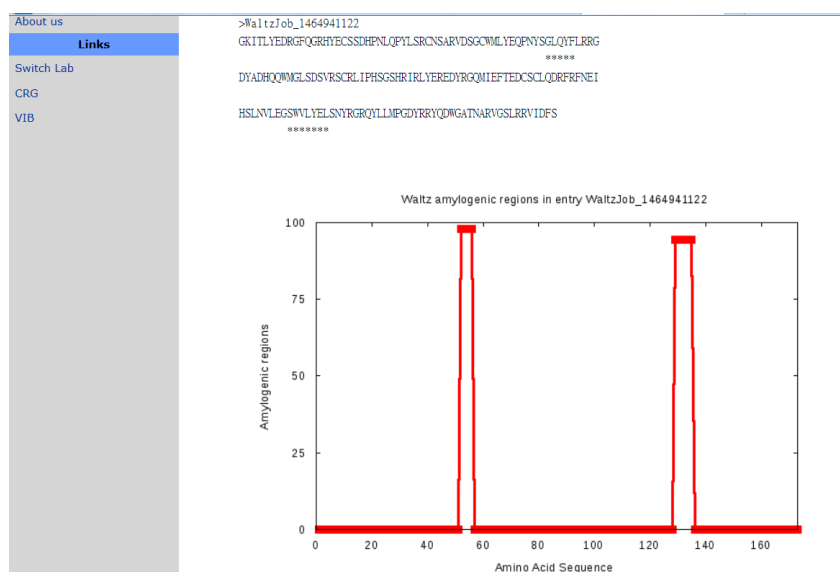


圖 3.2.17 WALTZ 結果示意圖

3.2.9 FoldAmyloid[155]

1. 進入 FoldAmyloid 首頁(<http://bioinfo.protres.ru/fold-amyloid/>)，如圖 3.2.18 頁面所示。
2. 依據 FASTA 格式輸入蛋白質序列，Scale 處選擇「Expected number of contacts 8Å」，閾值為 21.4(各胺基酸數值參考圖右方)，開始進行預測。

Sequence Short result Long result Information

Target sequence in FASTA format [Clear](#)

Scale: Expected number of contacts 8Å Averaging frame: 5 Threshold: 21.4

[Pred...](#)

Profile values

Co...	Na...	Residue	Value
A	ALA	Alanine	19.89
C	CYS	Cysteine	23.52
D	ASP	Aspartate	17.41
E	GLU	Glutamate	17.46
F	PHE	Phenylalanine	27.18
G	GLY	Glycine	17.11
H	HIS	Histidine	21.72
I	ILE	Isoleucine	25.71
K	LYS	Lysine	17.67
L	LEU	Leucine	25.36
M	MET	Methionine	24.82
N	ASN	Asparagine	18.49
P	PRO	Proline	17.43
Q	GLN	Glutamine	19.23
R	ARG	Arginine	21.03
S	SER	Serine	18.19
T	THR	Threonine	19.81
V	VAL	Valine	23.93
W	TRP	Tryptophan	28.48
Y	TYR	Tyrosine	25.93
X	UNK	Unknown	20.73

圖 3.2.18 FoldAmyloid 首頁示意圖

3. 計算結果如圖 3.2.19，其中藍點標示的位置即為類澱粉纖維相關的區域。

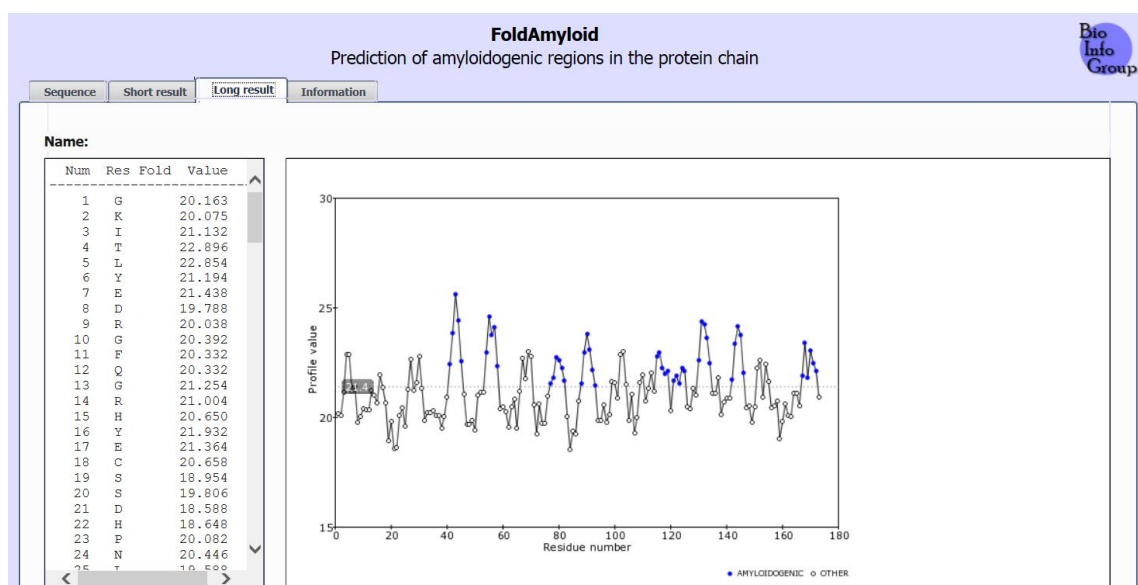


圖 3.2.19 FoldAmyloid 結果示意圖

3.2.10 Zygggregator[156-159]

1. 進入 Zygggregator 首頁，需註冊並登入才可使用。在登入後，頁面如圖 3.2.20。

(<http://www-mvsoftware.ch.cam.ac.uk/index.php/zygggregator>)

2. 輸入蛋白質序列以及 pH 值，點選提交即可。

Zygggregator: Prediction of Protein Intrinsic Aggregation Propensities

A sequence-based method of predicting intrinsic (i.e. sequence-dependent) amyloid aggregation propensities of proteins. If you are more interested in predicting the solubility or a generic aggregation propensity (i.e. non necessarily amyloid) you should use the CamSol calculation of the [intrinsic solubility](#) profile instead, which is a newer improved method.

For use by commercial organisations, please contact Lonza Biologics (contact.lonza@lonza.com)

Input sequence (20 standard amino acid, capital letters)

pH=



圖 3.2.20 Zygggregator 頁面圖

3. 計算完成後，可以得出每個胺基酸的分數，如圖 3.2.21 所示。



Vendruscolo Lab - Software
Department of Chemistry, University of Cambridge

[software home](#)

position	aa	Zygg Residue Score
1	G	0.717443
2	K	0.89062
3	I	1.0257
4	T	0.856876
5	L	0.607011
6	Y	0.487159
7	E	0.320207
8	D	0.370043
9	R	0.158676
10	G	-0.106214
11	F	-0.125313
12	Q	-0.172314
13	G	0.522337
14	R	-0.0760159
15	H	0.187216
16	Y	0.319056
17	E	0.228013
18	C	0.783302
19	S	0.970998
20	S	-0.857156
21	D	-0.161976
22	H	-0.800531
23	P	-0.93237
24	N	-2.2169
25	L	-1.88984
26	Q	-1.60449
27	P	-0.319967
28	Y	-0.846555

圖 3.2.21 Zygggregator 結果示意圖

3.2.11 AGGRESCAN3D

1. 進入 AGGRESCAN3D 首頁(<http://biocomp.chem.uw.edu.pl/A3D/>), 如圖 3.2.22。
2. 輸入 pdb code 或欲分析的 pdb 檔，將「Dynamic mode」以及「Mutate residues」選為「Yes」，並選擇欲突變的胺基酸，點選提交開始計算。

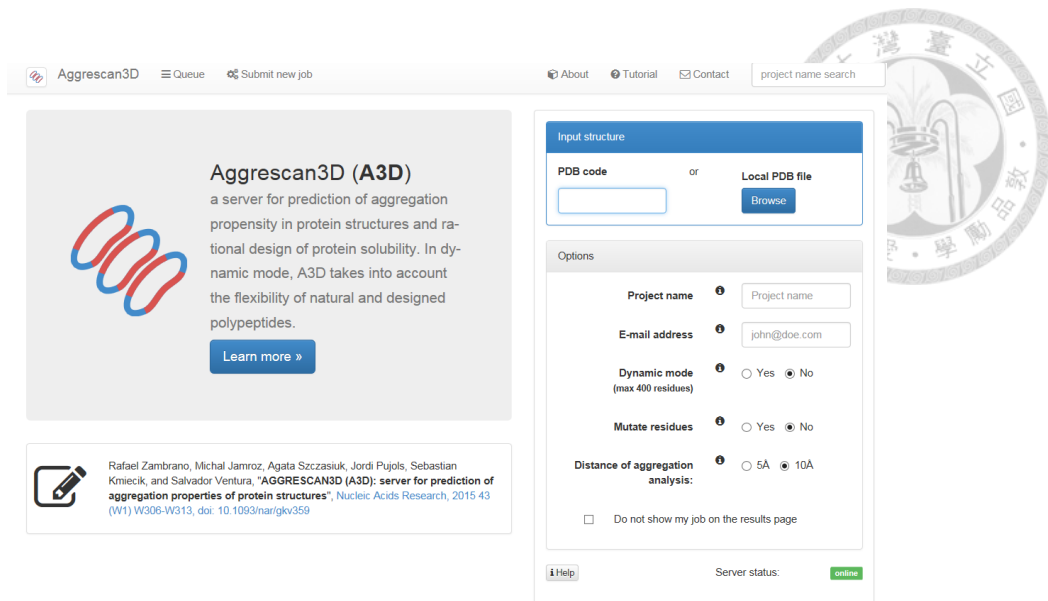


圖 3.2.22 AGGRESKAN3D 首頁示意圖

3. 計算約花費 2 小時，結束後會顯示各殘基的聚集分數，如圖 3.2.23 所示。

residue index	residue name	chain	Aggrescan3D score	mutation
1	G	X	-1.4743	mutate
2	K	X	-2.6578	mutate
3	I	X	0.0000	mutate
4	T	X	-2.3252	mutate
5	L	X	0.0000	mutate
6	Y	X	-1.8909	mutate
7	E	X	-1.9507	mutate
8	D	X	-1.8137	mutate

圖 3.2.23 AGGRESKAN3D 結果示意圖

3.2.12 Discovery Studio 2016 (Calculate Aggregation Scores)

1. 將欲分析之蛋白質結構檔匯入 DS 2016。
2. 進入 Macromolecule→Predict Protein Aggregation 點選 Calculate Aggregation Scores 功能。
3. 選擇欲分析之蛋白質結構檔，設定截斷半徑為 cut-off radius 為 5 和 10 Å，使用 CHARMM 為力場，開始進行分析。

4. 結果能同時顯示每個胺基酸的聚集分數與顏色示意圖，如圖 3.2.24。

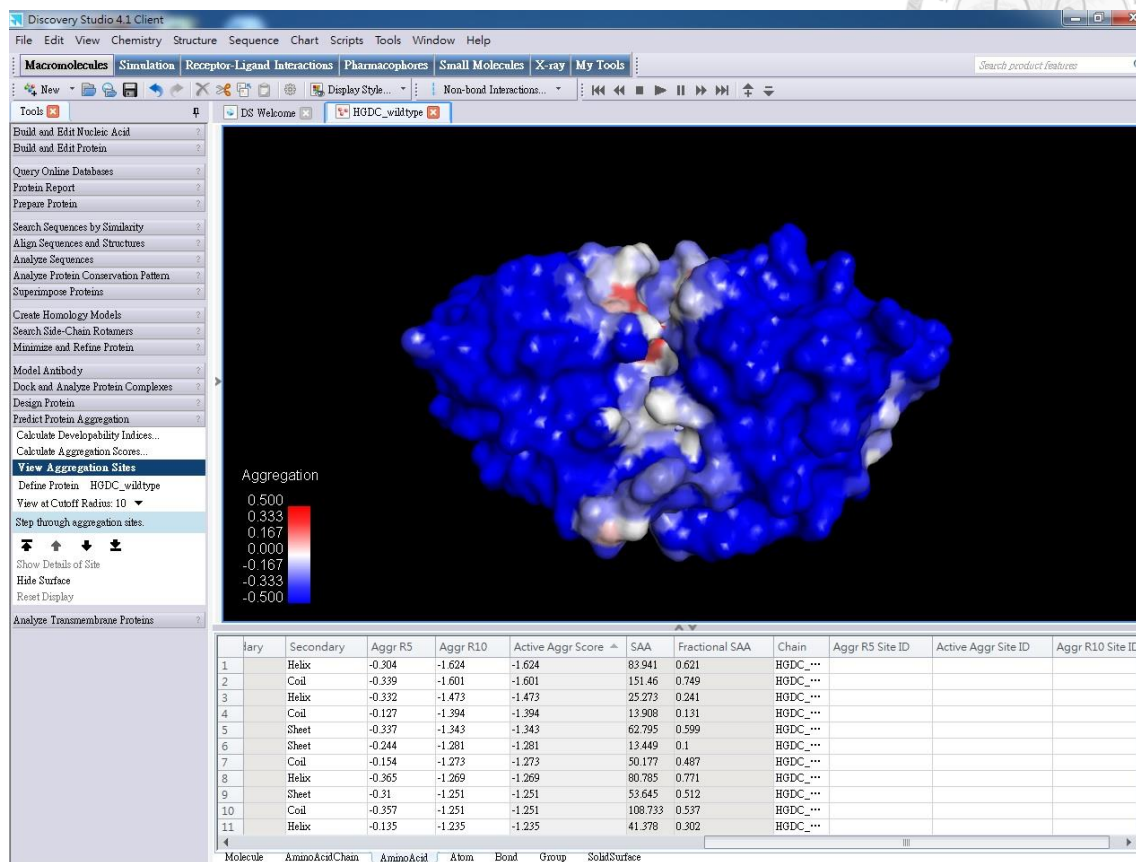


圖 3.2.24 DS 預測聚集程度結果圖



3.3 分子動力學模擬及數據分析

3.3.1 HyDC 原態結構取得

1. 從 Protein data bank 網站下載 HyDC 之結晶結構(pdb code: 1hk0)[57, 58]。
2. 利用軟體 Discovery studio 2016 將水晶體蛋白 HyDC 結構檔中的水去除，並將胺基酸序列重新編號。(原檔案的胺基酸編號不連續)

3.3.2 Gromacs 模擬流程

1. 將水晶體蛋白 HyDC 之 pdb 結構置於 $7.0 \times 7.0 \times 7.0$ (nm³)立方體盒中，並將整個蛋白質的質量中心平移到盒子的中心處。指令如下所示：

editconf_4.5.3 -f 1hk0.pdb -o 1hk0_box.pdb -bt cubic -box 7 -c

2. 本研究選擇 Gromos96 45a3[270]作為蛋白質的力場，系統預設環境為中性(pH 7)的質子化狀態。指令如下所示：

pdb2gmx_4.5.3 -f 1hk0_box.pdb -ignh -o 1hk0_g.pdb -p forcefield.top

3. 在盒子中加入溶劑，選擇 SPC 的水模型，並新增溶劑的力場。指令如下所示：

genbox_4.5.3 -cp 1hk0_g.pdb -cs spc216.gro -o 1hk0_sol.pdb -p forcefield.top

4. 為了配合實驗的離子濃度，將盒子內部分的水分子以離子取代之，同時將系統之淨電荷調整至 0 以達到電中性。在本研究中，於盒子內加入 28 個鈉離子及 28 個氯離子，並考慮 HyDC(以及其突變種)在中性環境的淨電荷數，以鈉離子或氯離子於以平衡之。指令如下所示：

grompp_4.5.3 -f emdob.mdp -c 1hk0_sol.pdb -p forcefield.top -o 1hk0_ion.tpr

**genion_4.5.3 -s 1hk0_ion.tpr -o 1hk0_ion.pdb -pname NA -np (鈉離子數目)
-nname CL -nn (氯離子數目)**

5. 在所有物種都加入系統後，使用最陡坡度法(steepest descent)將系統進行能量最小化(energy minimization)。指令如下所示：

grompp_4.5.3 -f emdob.mdp -c 1hk0_ion.pdb -p forcefield.top -o 1hk0_em.tpr

mdrun_4.5.3 -s 1hk0_em.tpr -o 1hk0_em.trr -c 1hk0_em.pdb -g em.log -e em.edr &

6. 將蛋白質的位置進行限制(position restraint)，並設定系統在 1 ns 內以步階方式上升至目標溫度(343K、425K)。指令如下所示：

grompp_4.5.3 -f pr.mdp -c 1hk0_em.pdb -p forcefield.top -o 1hk0_pr.tpr
mdrun_4.5.3 -s 1hk0_pr.tpr -o 1hk0_pr.trr -c 1hk0_pr.pdb -g pr.log -e pr.edr &

7. 將系統開始進行分子動力學模擬(molecular dynamics simulations)，整個系統的環境設定在定莫耳數、定壓以及定溫的情況。指令如下所示：

grompp_4.5.3 -f md.mdp -c 1hk0_pr.pdb -p forcefield.top -o 1hk0_ds.tpr
mdrun_4.5.3 -s 1hk0_ds.tpr -o 1hk0_ds.trr -c 1hk0_ds.pdb -g ds.log -e ds.edr &

8. 若在模擬過程中，因不可抗力影響使模擬中斷，或是想要進行更長時間的模擬，則需執行下列接跑指令：

tpbconv_4.5.3 -s 1hk0_ds.tpr -f 1hk0_ds.trr -e ds.edr -o 1hk0_ds_2.tpr -until
(時間, 單位 ps)
mdrun_4.5.3 -s 1hk0_ds_2.tpr -o 1hk0_ds_2.trr -c 1hk0_ds_2.pdb -g ds_2.log -e ds_2.edr &

3.3.3 Root Mean Square Deviation (RMSD)分析

Gromacs 即有計算 RMSD 分析的內建指令，原理如 2.4.1 所述。在計算前須將結構進行最適化疊合，並選定參考點為時間 0 時之骨架(backbone)結構，如此可輸出隨時間變化之 RMSD 結果，指令如下。

g_rms_4.5.3 -f 1hk0_ds.trr -s 1hk0_ds.tpr -o rmsd.svg

3.3.4 Root Mean Square Fluctuation (RMSF)分析

Gromacs 即有計算 RMSF 分析的內建指令，原理如 2.4.2 所述。在計算前須將結構進行最適化疊合，並選定參考點為時間 0 時之骨架(backbone)結構，如此可輸



出各殘基在時間內平均擾動的結果，指令如下，其中-res 表示結果會以殘基為單位輸出。

```
g_rmsf_4.5.3 -f 1hk0_ds.trr -s 1hk0_ds.tpr -rmsf.xvg -res
```

3.3.5 Native Contact 分析

在計算蛋白質的 native contact 之前，需先在 Linux 的作業系統內安裝以 python 語言所撰寫的軟體 MDAAnalysis[271]。分析的操作流程如下，其中設定截斷半徑為 6.5Å 並選擇重原子(非氫原子)進行計算[105]。

1. 在 Linux 系統中進入 python 語言模式，並匯入 MDAAnalysis 以及 native contact 分析軟體包。

```
python
```

```
import MDAAnalysis
```

```
import MDAAnalysis.analysis.contacts
```

2. 匯入待分析的結構檔及軌跡檔，其中兩者的原子數必須相同才能進行計算。

```
u=MDAnalysis.Universe("1hk0_ds.pdb", "1hk0_ds.trr")
```

3. 定義與選取欲分析之原子組別。

```
sel_g1="(resid 1:173) and (name C* or name O* or name N* or name S*)"
```

```
sel_g2="(resid 1:173) and (name C* or name O* or name N* or name S*)"
```

```
g1=u.selectAtoms(sel_g1)
```

```
g2=u.selectAtoms(sel_g2)
```

4. 在截斷半徑為 6.5Å 的條件下進行 native contact 的計算。

```
CA1=MDAnalysis.analysis.contacts.ContactAnalysis1(u, selection=(sel_g1,
```

```
sel_g2), regroup=(g1, g2), radius=6.5, outfile="native_contact.dat")
```

```
CA1.run(force=True)
```

5. 計算結束後，會輸出兩個結果，一個是隨時間變化之 native contact fraction，另一個則是 contact map 的矩陣。



3.3.6 氫鍵作用力分析

Gromacs 即有計算氫鍵數目的內建指令，指令如下所示，其中-n 表示能分析使用者自訂的蛋白質特定區間。另外，其氫鍵的定義為 Donor-Acceptor 間的距離小於 0.35 nm 且 Hydrogen-Donor-Acceptor 之夾角小於 30 度。

```
g_hond_4.5.3 -f 1hk0_ds.trr -s 1hk0_ds.tpr -num hbnum.svg -n index.ndx
```

3.3.7 距離分析

Gromacs 即有計算距離的內建指令，其可以計算兩個組別的最短距離或是質心距離隨時間的分布，指令分別如下，其中-n 表示能分析使用者自訂的蛋白質特定區間。

```
g_mindist_4.5.3 -f 1hk0_ds.trr -s 1hk0_ds.tpr -od mindist.svg -n index.ndx
```

```
g_dist_4.5.3 -f 1hk0_ds.trr -s 1hk0_ds.tpr -o distance.svg -n index.ndx
```

3.3.8 計算溶劑接觸面積(solvent accessible surface area)

Gromacs 即有計算溶劑接觸面積的內建指令，包括計算親水面積、疏水面積以及總面積等三種，指令如下，其中-n 表示能分析使用者自訂的蛋白質特定區間。

```
g_sas_4.5.3 -f 1hk0_ds.trr -s 1hk0_ds.tpr -o area.svg -n index.ndx
```

3.3.9 DSSP 二級結構預測分析

Gromacs 即有計算 DSSP 分析的內建指令，指令如下所示。其中-n 表示能分析使用者自訂的蛋白質特定區間，-sc 表示可以繪出每個時間下各種二級結構的數量，-o 則是能夠輸出一個二維矩陣，以便了解各殘基二級結構的變化。

```
do_dssp_4.5.3 -f 1hk0_ds.trr -s 1hk0_ds.tpr -n index.ndx -sc scount.svg -o ss.xpm
```


3.4 蛋白質—小分子對接模擬(Protein-Ligand docking simulations)

除了以突變的方式探討 H γ DC 結構改變之外，本研究還利用蛋白質與小分子對接的方式，嘗試找出能穩定 H γ DC 結構的小分子，以及了解其穩定的機制。本研究選用過去實驗室以酸性條件(pH 2)模擬 H γ DC 的結構，在 196 ns 時有四組新生成的 β -sheet 結構，並與多種小分子對接以及進行分子動力學的模擬。

3.4.1 CDOCKER

CDOCKER 對接模擬是利用 Discovery studio 2016 來進行，步驟如下所示。

1. 分別開啟蛋白質與小分子的結構檔，如圖 3.4.1。

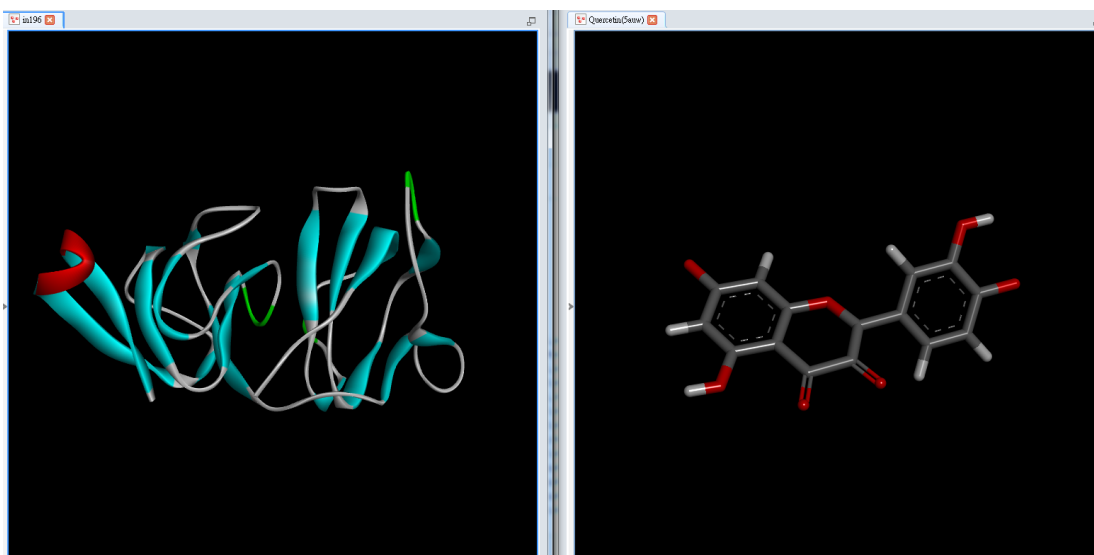


圖 3.4.1 蛋白質與小分子結構示意圖

2. 進入「Receptor-Ligand Interactions」→「Dock Ligands」選項中，分別點選「Prepare Protein」和「Prepare Ligands」，設定 pH 值以及選用 CHARMM 為力場，為進行對接的前置動作。
3. 進入「Receptor-Ligand Interactions」→「Define and Edit Binding Site」選項中，點選「From Receptor Cavities」，即可出現蛋白質可能發生鍵結之位置(binding site)，如圖 3.4.2 的紅色區域。根據需求，可調整此 Cavities 的大小。

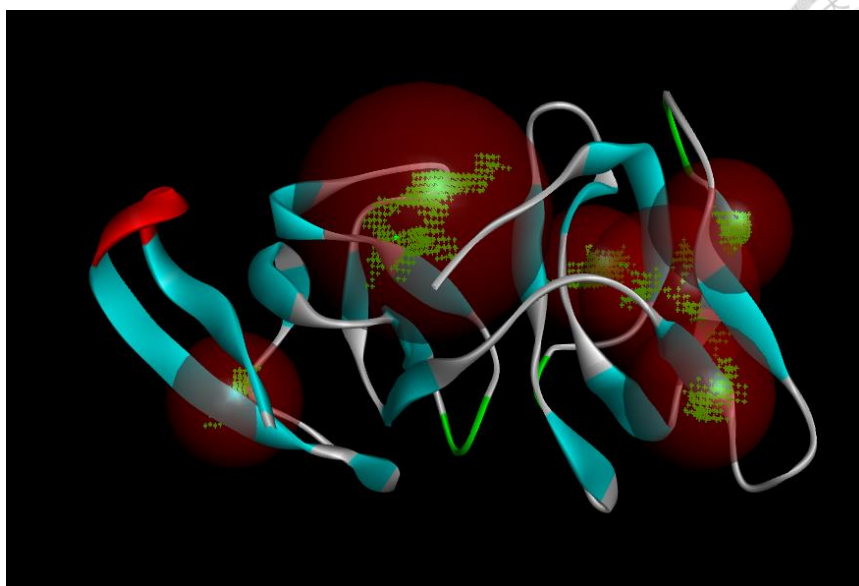


圖 3.4.2 蛋白質可能發生鍵結的區域示意圖

4. 進入「Receptor-Ligand Interactions」→「Dock Ligands」選項中，點選「Dock Ligands (CDOCKER)」。
- 依序設定要執行對接的 Receptor 和 Ligand 結構，以及欲產生鍵結的 binding site 區域，並將「Top hits」下的「Pose Cluster Radius」設定為 0.5，便可開始進行模擬。
5. 模擬結束後，會出現依能量排序前 10 名的 pose 以及其位置，如圖 3.4.3 所示。

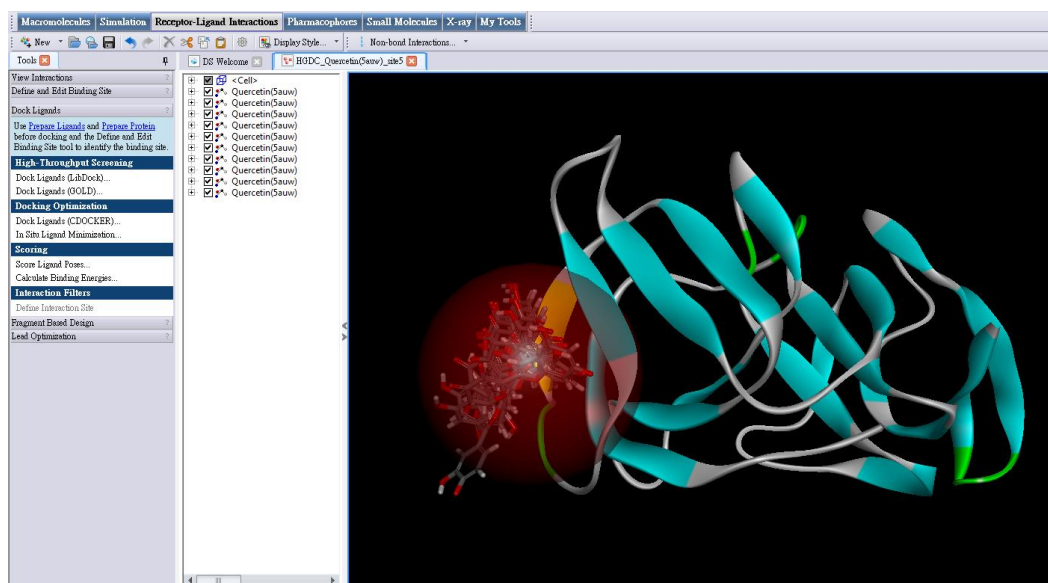


圖 3.4.3 CDOCKER 結果示意圖

3.4.2 分子對接作用力分析

分子對接作用力是利用 Discovery studio 2016 內建的功能來進行分析，步驟如下所示。

1. 開啟已對接完成的結構檔，點選任一個 pose，於「Receptor-Ligand Interactions」→「View Interactions」選單中，執行「Show 2D diagram」。
2. 分析結果會以小分子為中心，顯示有作用力的蛋白質殘基，並以不同的顏色表示不同的作用力，如圖 3.4.4。

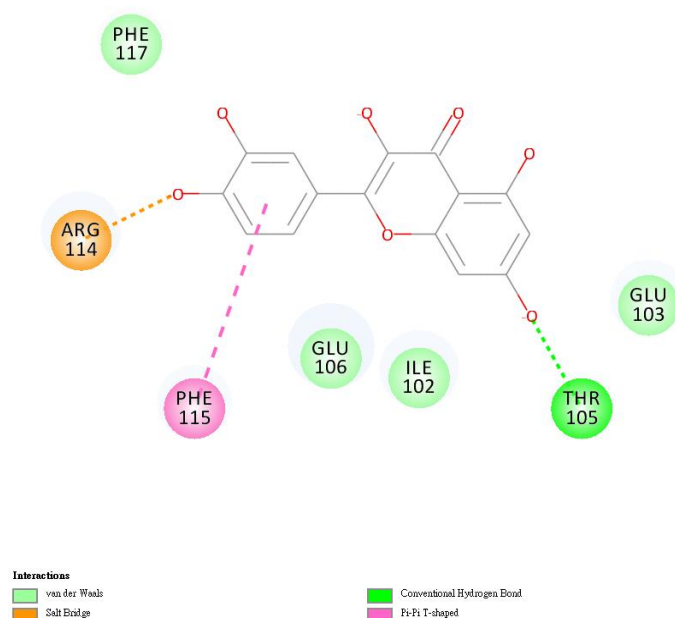


圖 3.4.4 分子對接作用力分析圖

3.4.3 分子對接動力學模擬

將分子對接後的 poses 進行分子動力學模擬，步驟見 3.3.2。其中，Gromacs 預設的力場並不適用於小分子，因此採用 PRODRG[272]產生小分子的力場。步驟如下所示。

1. 進入 PRODRG 首頁(<http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrgr>)，選擇「Run PRODRG」進入頁面，如圖 3.4.5。



[PRODRG Home](#) | [Run PRODRG](#) | [Get PRODRG](#) | [FAQ](#) | [Usage Stats](#)

Compound submission

Before you can submit molecules to the PRODRG server, you will need to have a token. Fill in your email address below and hit submit, and a token (essentially a short text string) valid for five PRODRG runs will be emailed to you.

My email address

If you already have a token, paste it here:

Then either or paste your input (PDB coordinates, MDL Molfile, SYBYL Mol2 file or [text drawing](#)) here:

Chirality Charges EM

圖 3.4.5 PRODRG 執行頁面圖

2. 使用前須先依說明取得 token，並輸入小分子的結構座標。設定「Chirality」為「No」，「Charges」為「Full」以及「EM」為「No」，便可開始執行。
3. 計算完成後，於 GROMACS(topology)處即可取得小分子的力場檔，如圖 3.4.6。

The GROMACS topology

```

;
; This file was generated by PRODRG version AA100323.0717
; PRODRG written/copyrighted by Daan van Aalten
; and Alexander Schüttelkopf
;
; Questions/comments to dava@davapci.bioch.dundee.ac.uk
;
; When using this software in a publication, cite:
; A. W. Schüttelkopf and D. M. F. van Aalten (2004).
; PRODRG - a tool for high-throughput crystallography
; of protein-ligand complexes.
; Acta Crystallogr. D60, 1355--1363.
;
;
[ moleculetype ]
; Name nrexcl
UNK      3
[ atoms ]
nr      type  resnr resid  atom  cgnr  charge  mass
1       O      1  UNK    O20    1    -0.510  15.9994
2       C      1  UNK    C8     1     0.255  12.0110
3       C      1  UNK    C7     1     0.255  12.0110
4       O      1  UNK    O17   2    -0.634  15.9994
5       C      1  UNK    C3     2     0.053  12.0110
6       C      1  UNK    C2     2     0.205  12.0110
7       OA     1  UNK    O22   2    -0.116  15.9994
8       H      1  UNK    H22   2     0.121  1.0080
9       CR1    1  UNK    C1     2     0.067  12.0110
10      HC     1  UNK    H1     2     0.099  1.0080
11      C      1  UNK    C6     2     0.205  12.0110
12      O      1  UNK    O21   3    -0.584  15.9994
13      CR1    1  UNK    C5     3     0.073  12.0110
14      HC     1  UNK    H5     3     0.108  1.0080
15      C      1  UNK    C4     3     0.223  12.0110
16      OA     1  UNK    O16   3    -0.101  15.9994
17      C      1  UNK    C9     3     0.223  12.0110
18      C      1  UNK    C10   3     0.058  12.0110
19      CR1    1  UNK    C15   4     0.057  12.0110
20      HC     1  UNK    H15   4     0.085  1.0080
21      C      1  UNK    C14   4     0.175  12.0110
22      OA     1  UNK    O18   4    -0.136  15.9994
23      H      1  UNK    H18   4     0.103  1.0080
24      C      1  UNK    C13   4     0.175  12.0110

```

圖 3.4.6 小分子力場檔示意圖

4. 將此檔案移入分子動力學的資料夾，並修改蛋白質的力場檔(topology)，以完成小分子力場的加入。

第四章 實驗結果與討論



如 2.2.3 與 2.3.3 所述，本研究主要在探討單點突變對於 HyDC 之不穩定區域(殘基 115-118)的結構改變，以及其與小分子對接的作用機制。吾人將本篇研究的結果分為三個部分，首先以生物資訊工具預測 HyDC 在進行單點突變後，其熱力學上的穩定性與聚集傾向；第二部分，使用分子動力學模擬 HyDC 與其突變種結構隨時間的變化情形；第三部分則是使用分子對接找出具有抑制 HyDC 聚集的小分子，並分析兩者間的作用力。

4.1 預測 HyDC 及其突變種的穩定性與聚集傾向

一般基本的胺基酸共有 20 種(目前最新發現已有 22 種胺基酸[273, 274]，然而因模擬軟體尚無支援，所以不納入考慮)，欲將某一種胺基酸進行單點突變時，會有另外 19 種可能，搭配本研究的 4 個目標殘基(115-118)，總計有 76 個可能的突變種。接著會針對這 76 個突變種進行分析。為了探討單點突變後的結構穩定性，吾人使用 Istable、DUET、Eris 與 EASE-MM 等四種生物資訊工具進行預測；另外使用 PASTA2.0、TANGO、AGGRESCAN、WALTZ、FoldAmyloid、Zygggregator 與 Disconvery Studio 2016 等工具預測其聚集傾向。

4.1.1 HyDC 及其突變種的穩定性預測結果

使用四種不同的生物資訊工具預測結果圖 4.1.1，其能量值為正表示其結構較穩定，而能量值為負則表示較不穩定。從大方向來看，這四種生物資訊工具都有一致性的結果，殘基 F115 與 F117 的能量值都較偏負，且能量的變化較大，表示將這兩個殘基進行突變時，結構都會變得比較不穩定；反之，將 R116 與 N118 進行突變時，結構會變得較穩定。此結果說明 F115 與 F117 對於 HyDC 的結構而言，是較為穩定的殘基，而 R116 與 N118 則是較為不穩定的殘基。

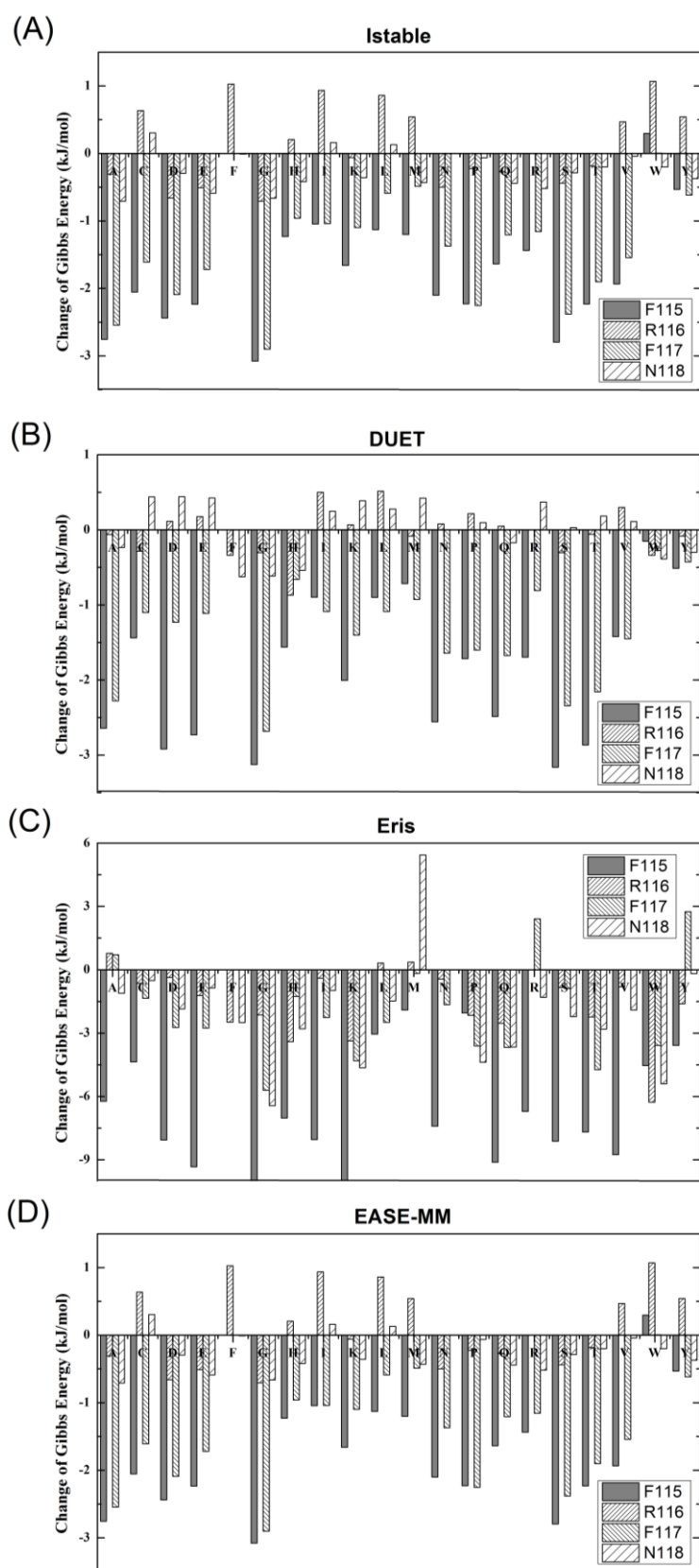
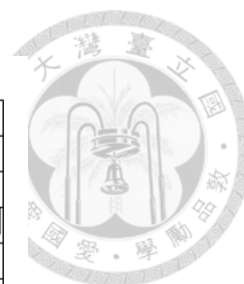


圖 4.1.1 以不同生物資訊工具預測 H_YDC115-118 所有可能突變之能量結果

(A)Istable (B)DUET (C)Eris (D)EASE-MM

(檔案路徑名稱：Sean_Thesis_Figure/4.1.1.tif)

接著，個別探討每種預測工具的結果。在 Istable 中，能量變化大小排名前三的是突變成 Alanine(A)、Glycine(G)或是 Valine(V)時；在 DUET 中，突變成 Alanine(A)、Aspartic acid(D)、Glycine(G)和 Serine(S)的不穩定性是較高的；在 Eris 中，多組的能量變化值都非常高，且差異不大；而在 EASE-MM 中，F116G、F116A 以及 F116S 分別是能量變化的前三名。綜合分析上述結果，雖然可以大致看出何種突變會導致結構的不穩定，但由於每一種工具的預測結果不盡相同，不同突變目標之殘基的能量值都互有勝負，因此無法輕易比較所有突變種的穩定與不穩定性。

為了進一步比較所有突變種的穩定與不穩定性，吾人使用下列計算方式來進行評分。首先，將每種預測工具計算出的自由能變化值都常規化(normalize)至 100 分，即把該殘基原態及其 19 個突變種之能量變化值最大者定義為 100 分，其餘者則等比例縮放，接著把每種預測工具計算出的分數，以其資料庫(database)的數目來做加權計算，排序出每個殘基各自的 1-20 名，其中第 1 名即代表最不穩定的突變種，而第 20 名則表示最穩定的突變種，綜合分析結果如表 4.1.1 所示。

表 4.1.1 綜合分析不同生物資訊工具預測突變能量變化之結果

殘基編號	最穩定突變種	最不穩定突變種
115	F115W	F115A
116	R116I	R116S
117	F117Y	F117A
118	N118I	N118H

*Glycine 與 Proline 因本身即容易破壞原本的二級結構[275]，所以不列入排序中

當將特定殘基進行突變時，造成結構不穩定的因素主要是由於側鏈性質的改變，包括立體障礙從大變成小，極性變成非極性，或是帶電變成不帶電等。依據上述因素的影響程度，能決定突變前後穩定性的改變程度，如在殘基 115 中，當把 Phenylalanine(F)突變成結構相似，同樣是苯環類的 Tryptophan(W)時，整體而言仍較穩定；而當突變成立體大小差異較大的 Alanine 時，其結構就明顯變得不穩定。其中，結構的相似度可以參考 BLOSUM62 陣列[276]的轉換分數。



4.1.2 HyDC 及其突變種的聚集傾向預測結果

本研究使用多種生物資訊工具預測 HyDC 在突變後的聚集傾向，此處將原態 HyDC 的聚集傾向當作標準，若計算出的聚集分數大於原態時，便視為是更容易產生聚集，將所有結果整理如表 4.1.2。從大方向來看，預測出聚集傾向較高的組別大多位於殘基 R116 以及 N118，表示當對於此兩個殘基進行突變時，HyDC 是比較容易發生聚集的。而這兩個殘基的突變種在能量的預測上，是會變得比較穩定的，其理由有二：第一是根據蛋白質的能量漏斗圖[277]，如圖 4.1.2，可以發現蛋白質在原態時，其能量僅是相對左右兩邊來得低，即為暫時穩定的狀態，但若在極端環境(如高溫或 pH 值低)下使之能量提升，則會變成去摺疊(unfolding)的型態，接著當系統要趨向穩定而能量下降時，有比較大的可能會變成不規則聚集(amorphous aggregates)或是形成類澱粉纖維(amyloid fibril)；第二是由於殘基 115-118 原本就被預測出可能發展出 β -sheet 結構的區間，當變成結構類似的突變種時，其可能仍具有高聚集傾向，而其結構改變造成的能量變化並不大。

表 4.1.2 以多種生物資訊工具預測 HyDC 的聚集傾向一覽表

生物資訊工具	聚集傾向較原態高之突變種
PASTA2.0	R116I, R116V, N118I, N118V
TANGO	N118F, N118I, N118L, N118V, N118W, N118Y
AGGRESCAN	將殘基 R116 和 N118 突變為 A, C, F, G, H, I, L, M, P, S, T, V, W, Y 任一種
WALTZ	R116I, R116W, R116Y
FoldAmyloid	將四個殘基突變為 C, F, H, I, L, M, V, W, Y 任一種
Zygggregator	F115W, R116K, F117W, N118T
Discovery Studio 2016	R116F, R116I, R116L, R116V, R116W, R116Y, N118F, N118I, N118L, N118Y

由於不同的生物資訊工具使用不同的資料庫以及演算法，因此預測結果都略有出入。吾人選出上述預測結果中，有最大交集的組別 R116I 以及 N118I，即被多種預測工具判定為聚集程度較高的突變種，而這兩個突變種在表 4.1.1 都是屬於穩

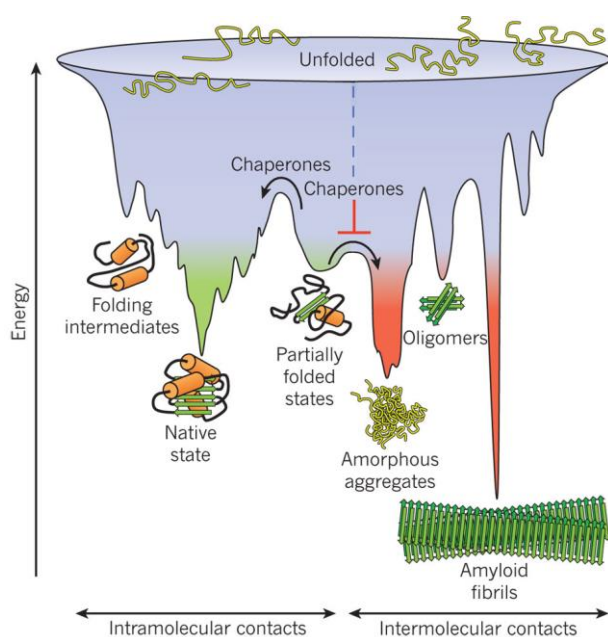


圖 4.1.2 蛋白質多種形態之能量漏斗圖[277]

定的。從結構來分析，當側鏈從帶正電的 Arginine(R)或是具高極性的 Asparagine(N)變成高疏水性的 Isoleucine(I)，搭配鄰近且具疏水性的 Phenylalanine(F)時，會將整個區間的疏水區裸露出來，而 HyDC 為了減少疏水區的裸露，就可能會發生聚集。另外，根據 Cornish 等人的研究指出[278]，當突變成具有 β -branched 的胺基酸(如 Valine、Isoleucine 等)時，容易因為立體結構其較大的使得螺旋(helices)結構瓦解，即較容易使得結構開展(unfolding)；當結構更開展時，蛋白質為了達到較穩定的狀態，會有聚集或形成類澱粉纖維的可能。本研究的結果 R116I 與 N118I 皆為有 β -branched 的突變種，與過去研究有一致的結果。

然而，生物資訊工具大多以蛋白質序列(sequence)以及特定的熱力學性質(thermodynamic properties)進行預測，與真實情況仍有一定程度的不同，因此在本研究中僅先以生物資訊工具進行初步的預測，將原先 76 個突變種縮減為 8 個，再接著使用分子動力學進行模擬，深入並真實模擬蛋白質結構的變動。

4.2 探討 HyDC 及其突變種於動力學下的結構變化

為了探討 HyDC 及其突變種在動力學模擬下的結構變化，此節先利用 RMSD 與原態接觸分析探討結構的穩定性與開展程度，並使用 AGGRESCAN3D 與 Discovery Studio 2016 來計算聚集傾向。在綜合比較之後，進行不穩定區域 115-118 的結構分析，包括 RMSF、DSSP 以及親疏水程度分別探討殘基結構擾動以及二級結構與三級結構的變動。除此之外，吾人將模擬系統分別設定在 pH7、343K 以及 pH7、425K 的環境中，並討論比較之。

4.2.1 RMSD 分析結果

Root mean square deviation (RMSD)常用來探討結構的改變，藉由計算每個原子與參考結構原子的距離方均根隨時間變化，可以了解分子結構在模擬過程中的改變為何，其在 343K 與 425K 的分析結果分別如圖 4.2.1 與圖 4.2.2 所示。

在 343K 的環境下，大多數突變種在 100 ns 的 RMSD 值都與原態差異不大，大約在 0.25~0.3 nm 間，唯一 RMSD 值較大的突變種是 N118H，約在 0.35~0.4 nm 左右，是為較不穩定的突變種；另外，F115W 自 70 ns 起，RMSD 值持續下降至 0.2~0.25 nm 左右，略低於原態，是為較穩定的突變種。在 425K 的環境下，原態的 RMSD 值在 50 ns 時約在 0.5 nm 左右，與大多數的組別沒有明顯差異，其中 RMSD 值變化較大的突變種為 N118H，其值在 40 ns 後上升至約 0.8 nm，是較為不穩定的突變種；而 F115W 與 F117Y 在 50 ns 時的 RMSD 值分別約為 0.42 與 0.30 nm，是比原態較為穩定的突變種。

結合不同溫度下的結果，可以合理推論 N118H 為較不穩定的突變種，而 F115W 與 F117Y 是較穩定的突變種。此結果也與表 4.1.1 所列之最穩定突變種與最不穩定突變種的結果相符合。

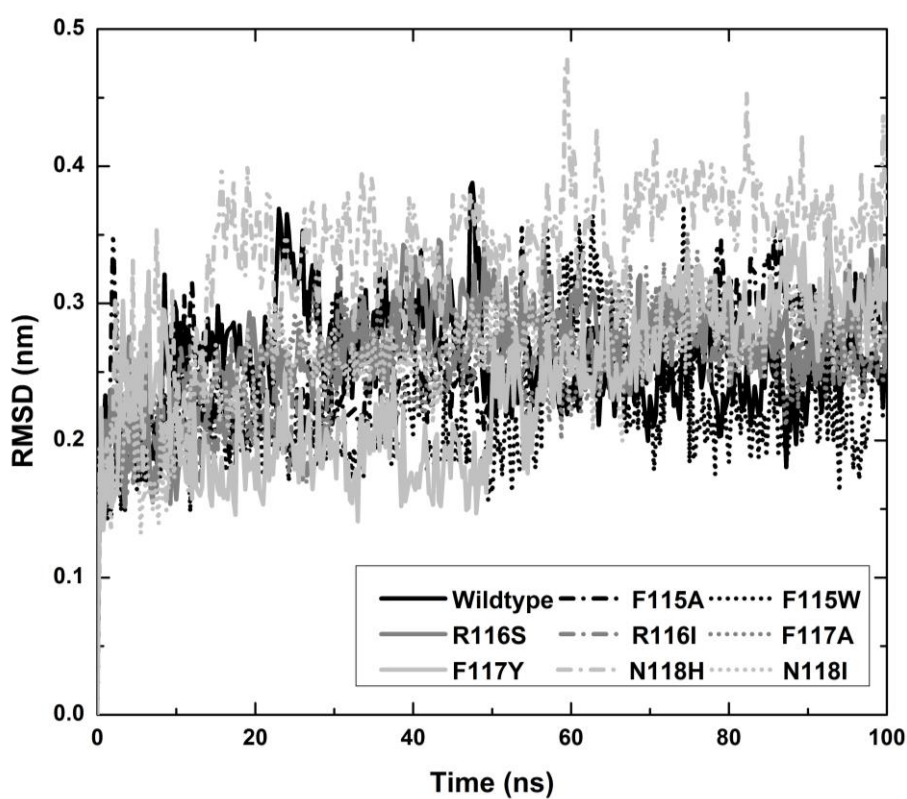


圖 4.2.1 H γ DC 與其突變種在 343K 下 RMSD 值隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.1.tif)

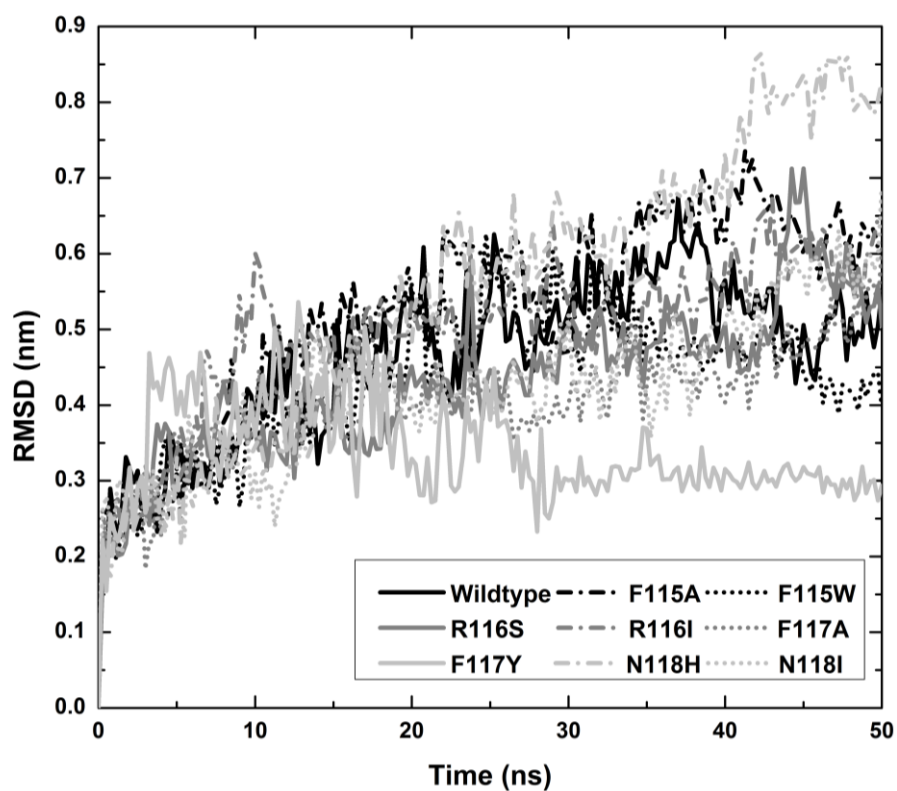


圖 4.2.2 H γ DC 與其突變種在 425K 下 RMSD 值隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.2.tif)

4.2.2 原態接觸(Native contact)分析結果

原態接觸是分析蛋白質摺疊程度的一項指標，透過計算蛋白質中每個原子在原態下的接觸情形，分析計算蛋白質隨時間的開展程度，當原態接觸分率下降時，蛋白質的結構會比較開展，其在 343K 與 425K 的分析結果分別如圖 4.2.3 與圖 4.2.4 所示。

在 343K 環境下，雖然 HyDC 原態的原態接觸分率在模擬過程中是最小的，F117A 的原態接觸分率是最大的，但在 60 ns 後已無明顯的不同，此差異應為模擬前期的未平衡狀態；在 100 ns 時，HyDC 原態與其突變種的原態接觸分率都在約 0.77~0.81 間，沒有太大的差異，即結構並無明顯的開展。在 425K 環境下，F117Y 在 50 ns 時的原態接觸分率約為 0.72，是開展程度最小的突變種，其次是 F115W 與 R116I，其值約為 0.67，而其他突變種的原態接觸分率都在約 0.6~0.63 間，與原態沒有明顯的差異。

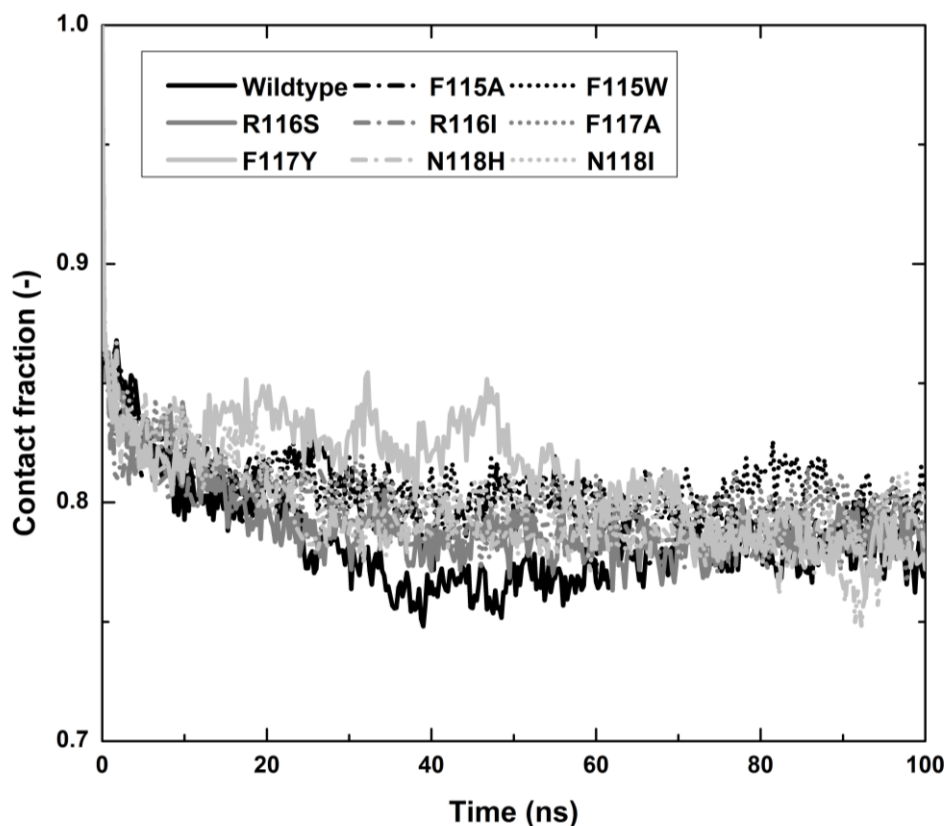


圖 4.2.3 HyDC 與其突變種在 343K 下原態接觸分率隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.3.tif)

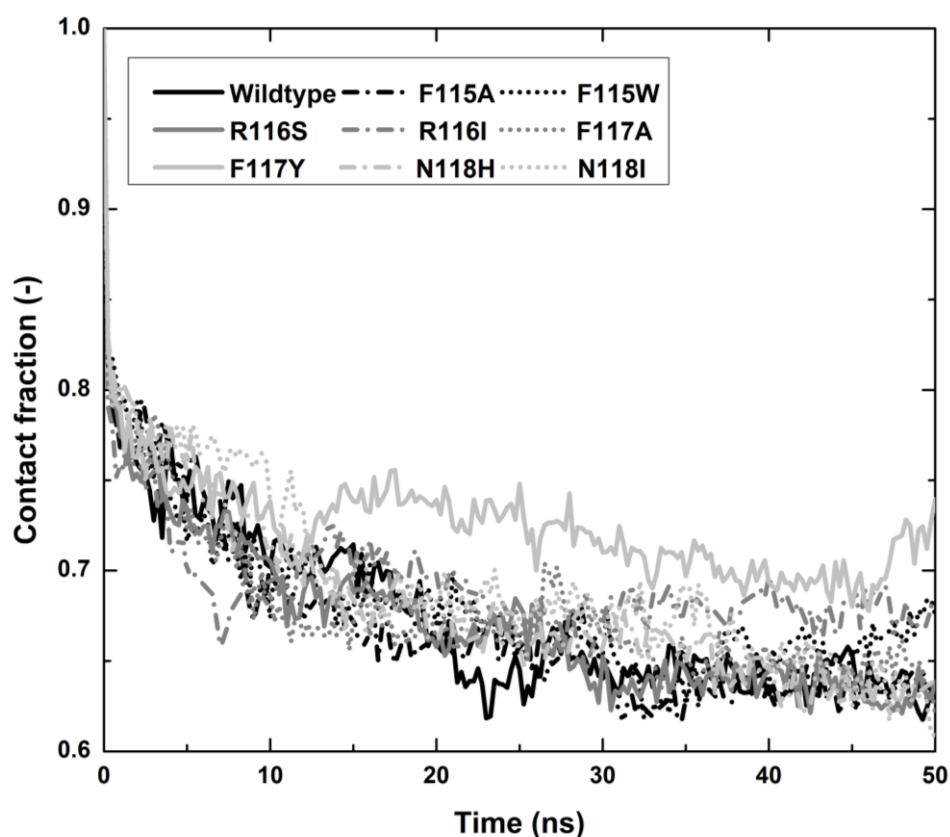


圖 4.2.4 HyDC 與其突變種在 425K 下原態接觸分率隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.4.tif)

綜合比較不同溫度下的結果，雖然在 343K 環境下的開展程度沒有明顯差異，但在 425K 下，可以發現有三個相對於原態較不開展的突變種，分別為 F115W、R116I 與 F117Y。其中 F115W 和 F117Y 在 RMSD 的結果被判定為較穩定的突變種，與此結果相符合。

4.2.3 聚集傾向(Aggregation Propensity)分析結果

雖然現今能預測蛋白質聚集傾向以及可能發生聚集區域的生物資訊工具已越來越多，但大多數都以蛋白質序列作為演算的方法，為了更精確瞭解蛋白質在動力學模擬下的聚集傾向，本研究使用 AGGRESCAN3D 與 Discovery Studio 2016 來計算 HyDC 與其突變種在模擬過程中的聚集分數，其中分數越大表示聚集傾向越高，反之則較低。

AGGRESCAN3D 的計算結果如圖 4.2.5 以及圖 4.2.6，為了觀察方便，將原態 HyDC 在 0 ns 時所計算出的聚集分數平移至 0，若結果為正表示其突變種比原態更容易產生聚集，反之則否。在 343K 環境下，R116I 的聚集分數在所有時間點都大於原態，表示此種突變對於造成蛋白質聚集的影響較為顯著，而 N118I 在 80 ns 後，聚集分數也有變高的傾向；另外，F115A 以及 F117A 的聚集分數則在 60 ns 後開始變小，表示其相對來說比較不容易有聚集的傾向。在 425K 環境下，發現 F115A、R116I 與 R116S 在 40~50ns 時，其聚集分數都能高達接近 40 左右，明顯比原態來得高，是較可能產生聚集的突變種；而 F117A 與 F117Y 在 50ns 時，聚集分數都是較小的，為較不易產生聚集的突變種。綜合比較不同溫度下的結果，可以發現 R116I 是較為容易產生聚集的突變種，而 F117A 則為較不易產生聚集的突變種。其中 R116I 於 4.1.2 節中，也被生物資訊工具預測為較有可能產生聚集的突變種之一。

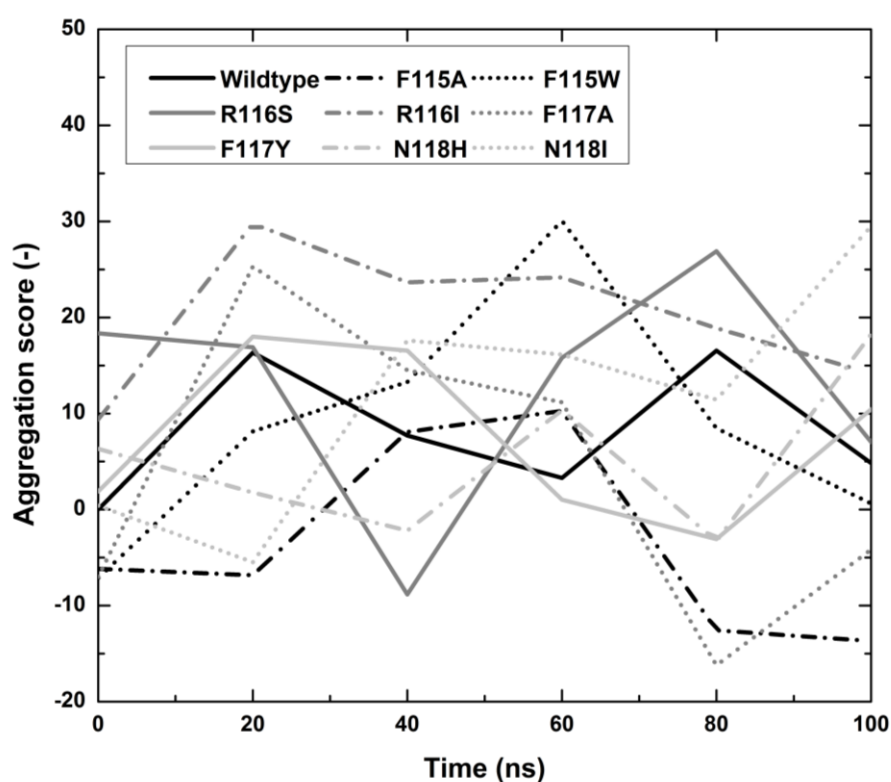


圖 4.2.5 以 AGGRESCAN3D 計算 HyDC 與其突變種在 343K 下之聚集分數隨時間變化圖

(檔案路徑名稱：Sean_Thesis_Figure/4.2.5.tif)

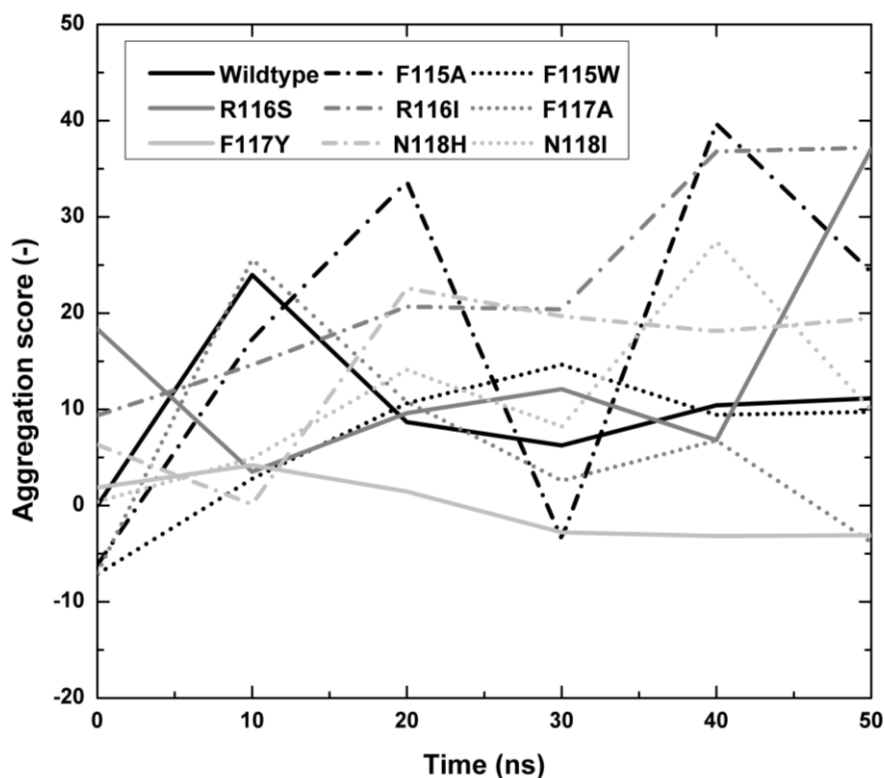


圖 4.2.6 以 AGGRESCAN3D 計算 HyDC 與其突變種在 425K 下之聚集分數隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.6.tif)

以 Discovery Studio 2016 的計算結果如圖 4.2.7 與圖 4.2.8，同樣將原態 HyDC 在 0 ns 的聚集分數平移至 0 進行調整。根據結果可以發現，在 343K 環境下，R116I 的聚集分數隨著時間演進都是最高的，即最有可能產生聚集的突變種；而 F117A 以及 N118H 的聚集分數為最低的兩組，為較不容易產生聚集的突變種，其餘的組別則沒有明顯較大或較小的差異。在 425K 環境下，R116I、R116S 與 N118I 的聚集分數在 20 ns 後都明顯高於原態，是聚集傾向較高的突變種，而 F115W、F117Y 與 N118H 的聚集分數在 50 ns 則是較小的組別，是聚集傾向較低的突變種。綜合比較不同溫度下的結果，可以推知 R116I 應具有較高的聚集傾向，而 N118H 的聚集傾向則是較小的。

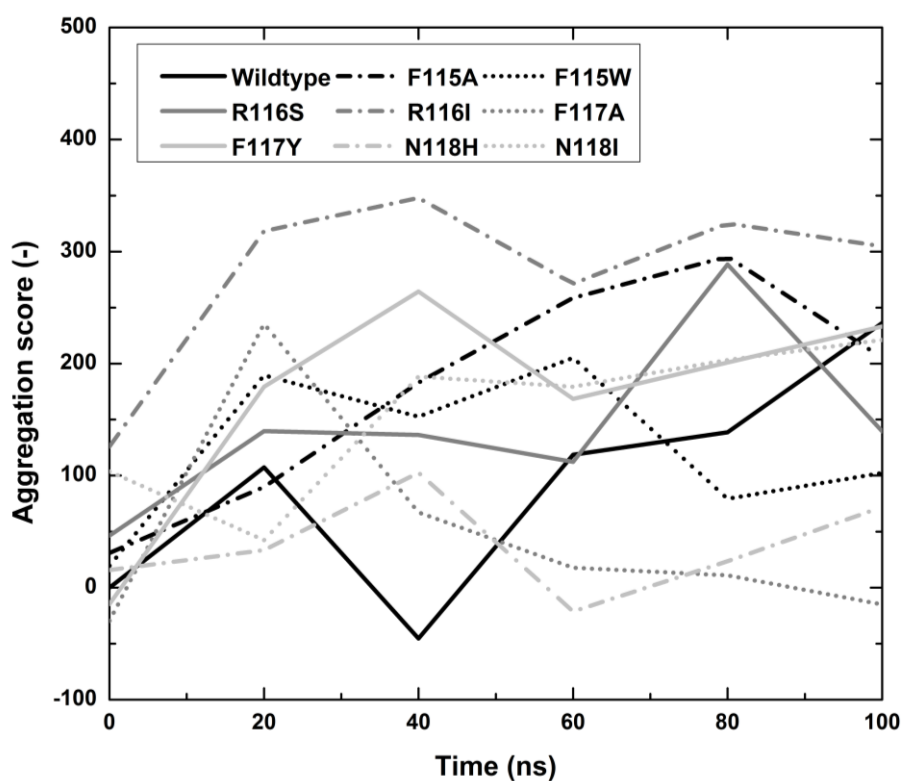


圖 4.2.7 以 DS 2016 計算 H γ DC 與其突變種在 343K 下之聚集分數隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.7.tif)

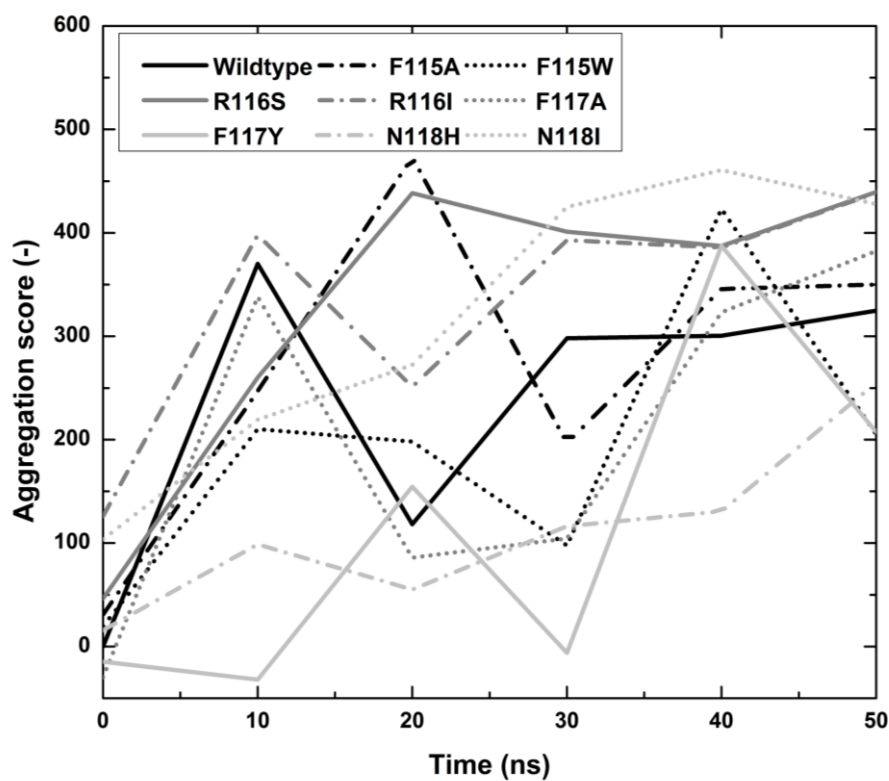


圖 4.2.8 以 DS 2016 計算 H γ DC 與其突變種在 425K 下之聚集分數隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.8.tif)

綜合比較 4.2.1 至 4.2.3 的結果，吾人將各組別最後 20 ns 之 RMSD 值、原態接觸分率以及聚集傾向值進行平均，並加減一個標準差，若其範圍與原態所得出之範圍有重疊，則表示無顯著差異，可能仍在擾動範圍間；若範圍無重疊，再進一步判斷其值為較大或較小，整理結果如表 4.2.1 所示。F115A、F115W 的穩定性與開展程度與原態沒有太大的區別，聚集傾向也是有大有小無法判定；R116I、R116S、F117A、N118I 的結構穩定性和開展程度都與原態差不多，然而 R116I 的聚集傾向極為顯著地大，F117A 的聚集傾向極為顯著地小，另外兩者則是與原態沒明顯差異；F117Y 為結構較穩定且較無開展的突變種，然而聚集傾向是略小的；最後是 N118H 的突變種，其結構會變得比較不穩定，而聚集傾向是略為減少的。過去研究蛋白質結構穩定性與聚集傾向的文獻指出，蛋白質因突變造成結構的不穩定，會降低其聚集傾向[279]。

表 4.2.1 HyDC 突變種之結構穩定性與聚集傾向與原態比較一覽表

突變種	穩定性		開展程度		聚集傾向		聚集傾向	
	(RMSD)		(Native contact)		(AGGRESCAN3D)		(DS2016)	
	343K	425K	343K	425K	343K	425K	343K	425K
F115A	=	=	=	=	-	+	=	+
F115W	=	=	=	=	=	=	-	=
R116I	=	=	=	-	=	+	+	+
R116S	=	=	=	=	=	=	=	+
F117A	=	=	=	=	-	-	-	=
F117Y	=	+	=	-	=	-	=	=
N118I	=	=	=	=	=	=	=	+
N118H	-	-	=	=	=	+	-	-

*穩定性：+表示結構較原態穩定，-表示不穩定，=表示無明顯差異

*開展程度：+表示結構較原態開展，-表示不開展，=表示無明顯差異

*聚集傾向：+表示聚集傾向較原態高，-表示聚集傾向低，=表示無明顯差異

分子動力學的模擬結果與 4.1.1 節與 4.1.2 節的預測結果大多符合，如 F115W 為殘基 115 的最穩定突變種，N118H 為殘基 118 的最不穩定突變種，以及 R116I 與 N118I 為聚集傾向最高的突變種等。然而，結構穩定性與聚集傾向在本研究中，

並沒有一致性的關聯。為了進一步瞭解 HyDC 聚集相關的結構改變，吾人挑選聚集傾向最大的 R116I 以及最小的 F117A 為目標突變種，以 RMSF、DSSP 以及計算疏水面積對 HyDC 之不穩定區域(殘基 115-118)進行分析。



4.2.4 RMSF 分析結果

RMSF 可以用來計算各殘基在模擬過程中的擾動情形，於 343K 以及 425K 環境下的分析結果分別如圖 4.2.9 與圖 4.2.10 所示。在 343K 環境下，原態、R116I 與 F117A 在模擬過程中，殘基 115-118 的擾動值都約 0.23 nm 左右，幾乎沒有差異。在 425K 的環境下，R116I 突變種在殘基 115-118 的 RMSF 值最大約為 0.57 nm，大於原態的 0.47 nm，而 F117A 在殘基 115-118 的 RMSF 值最大為 0.42 nm，略小於原態。雖然在 343K 的分析結果沒有明顯的差異，但就 425K 的結果而言，可以推論為聚集傾向的程度，可能來自於此不穩定區域的擾動情形。

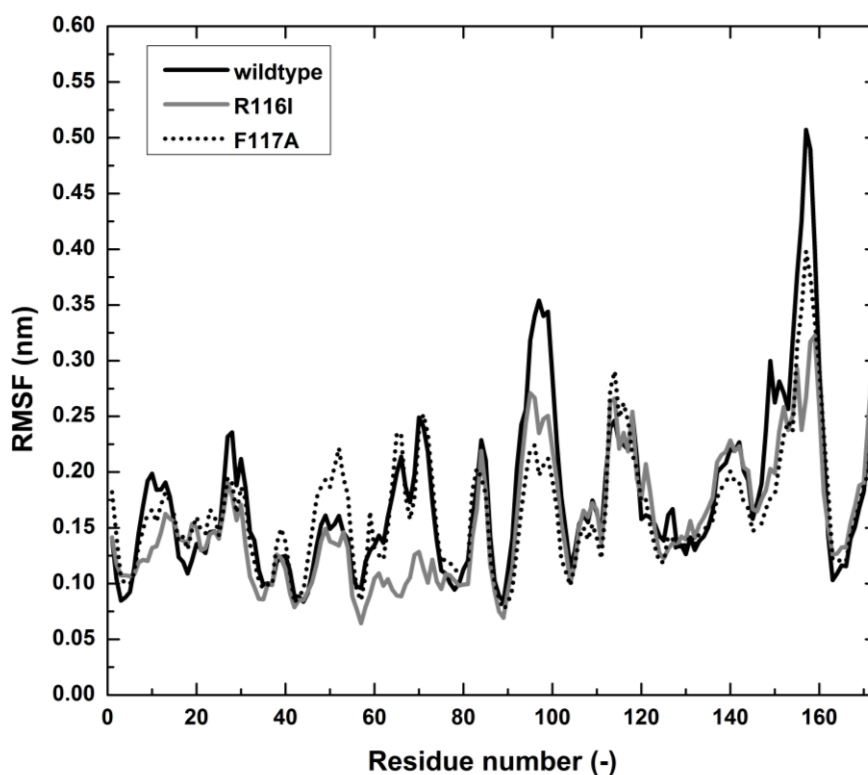


圖 4.2.9 HyDC 與其突變種在 343K 下 RMSF 值結果圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.9.tif)

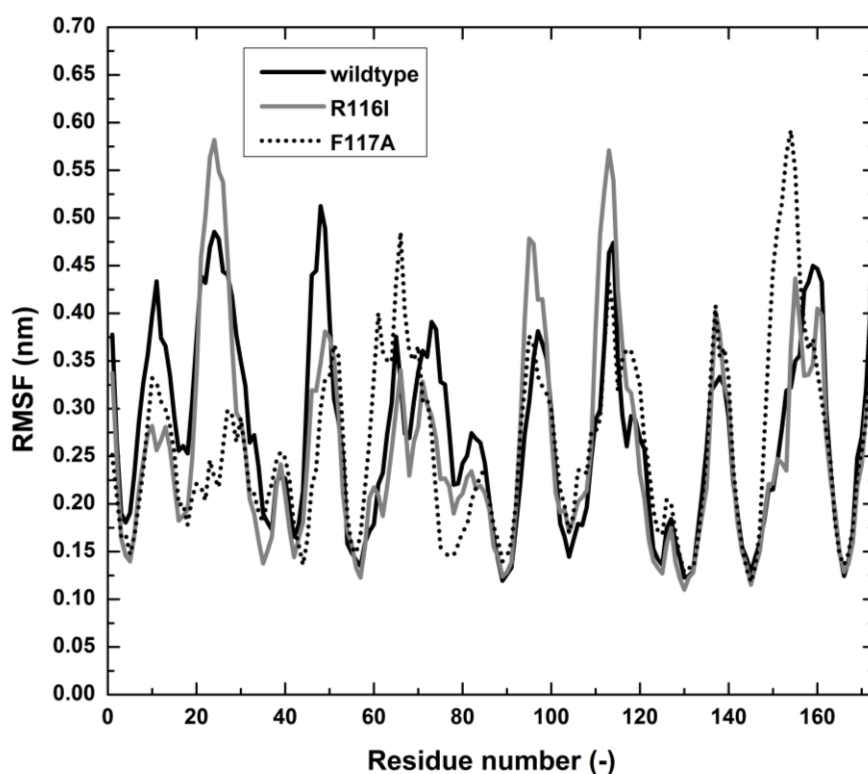


圖 4.2.10 HyDC 與其突變種在 425K 下 RMSF 值結果圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.10.tif)

4.2.5 二級結構分析結果

在蛋白質進行去摺疊或是形成聚集時，通常會伴隨著二級結構的改變，本節使用 DSSP 演算法計算 HyDC 與其突變種在模擬過程中二級結構的變化，並以百分率的方式呈現之，在 343K 與 425K 環境下的結果分別如圖 4.2.11 與圖 4.2.12 所示。在原態結構中，殘基 115-118 的二級結構應屬於 bend，位於 motif-3 上，鄰近 motif-4，連接著前端的 α -helix(殘基 111-114)與後端的 β -sheet(殘基 122-127)。在 343K 環境下，原態、R116I 與 F117A 在殘基 115-118 區域的二級結構沒有明顯的改變，大多仍是以 bend 的形式存在。在 425K 環境下，雖然整體的 β -sheet 減少，且前端的 α -helix(殘基 111-114)有被瓦解的趨勢，但原態、R116I 與 F117A 在殘基 115-118 區域仍然無二級結構的變化。在 RMSF 結果中，三者在此區域雖然有擾動的差異，但可能還沒有達到能影響二級結構的擾動程度。

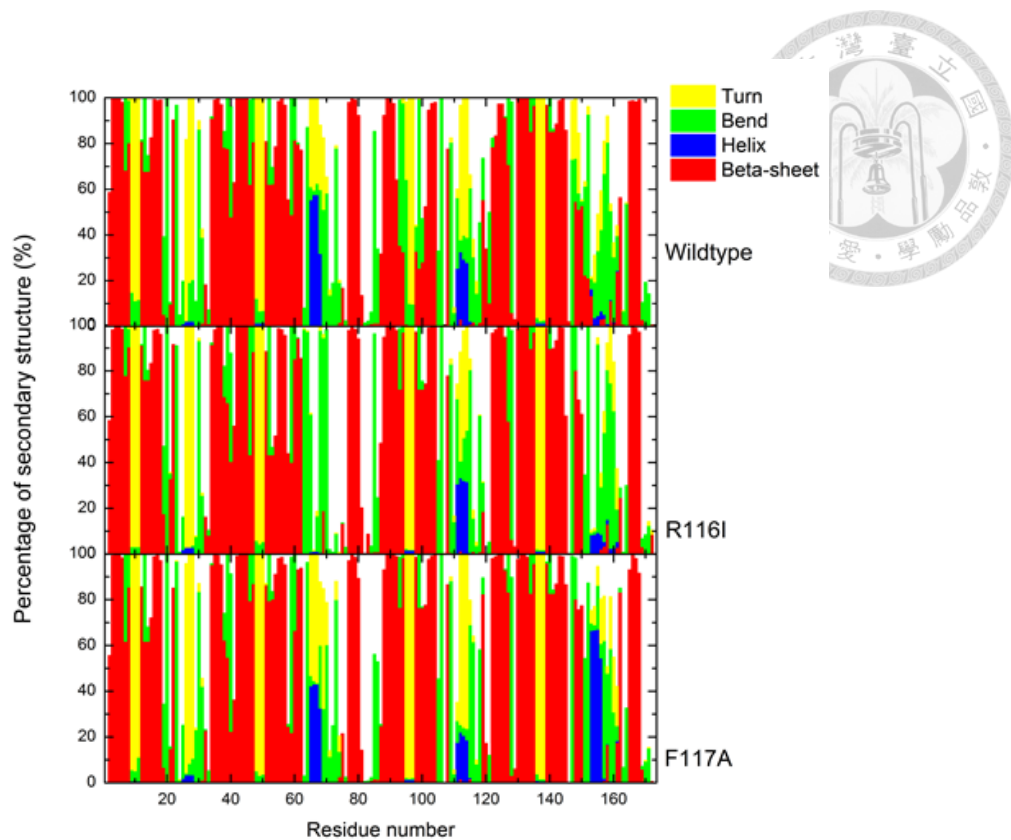


圖 4.2.11 H γ DC 在 343K 下各殘基二級結構於模擬過程中之存在百分率
(檔案路徑名稱：Sean_Thesis_Figure/4.2.11.tif)

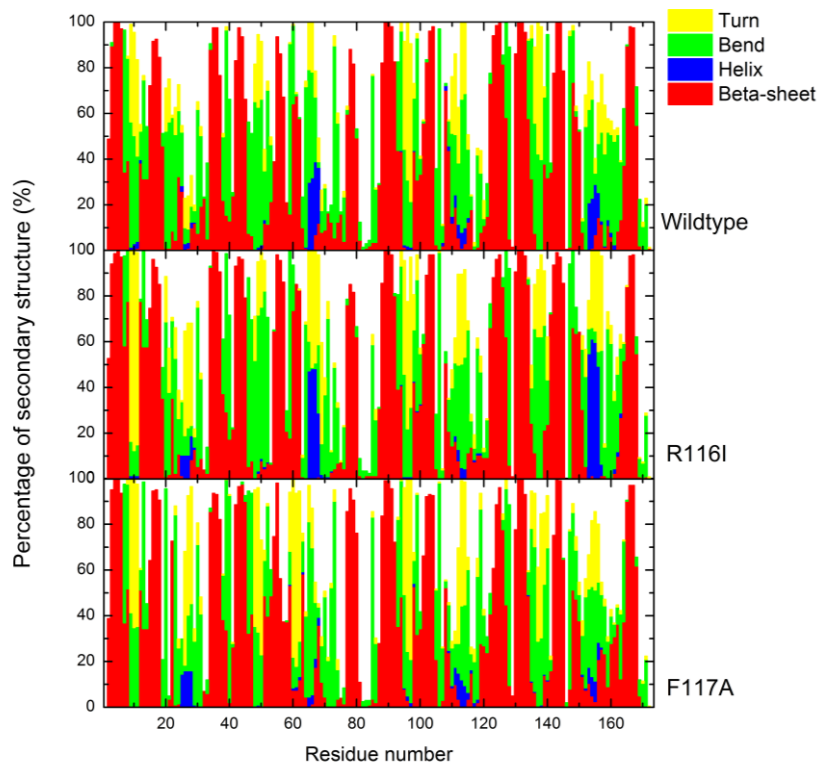


圖 4.2.12 H γ DC 在 425K 下各殘基二級結構於模擬過程中之存在百分率
(檔案路徑名稱：Sean_Thesis_Figure/4.2.12.tif)

4.2.6 親疏水程度分析結果

為了瞭解 HyDC 在模擬過程中三級結構的變化，吾人計算殘基 115-118 之疏水面積隨時間的變化，於 343K 與 425K 的結果分別如圖 4.2.13 與圖 4.2.14。結果發現在 343K 的環境下，R116I 之殘基 115-118 的疏水面積約在 4.5~5.0 nm²，大於原態的 4.0~4.5 nm²，而 F117A 之殘基 115-118 的疏水面積則約為 3.0~3.5 nm²，明顯小於原態，且在 343K 與 425K 有一致性的結果。根據疏水面積的計算結果，可以推論突變種的聚集傾向可能也與此不穩定區域的親疏水程度有所關聯，當疏水面積越大時，其聚集傾向可能就越高，反之亦然。

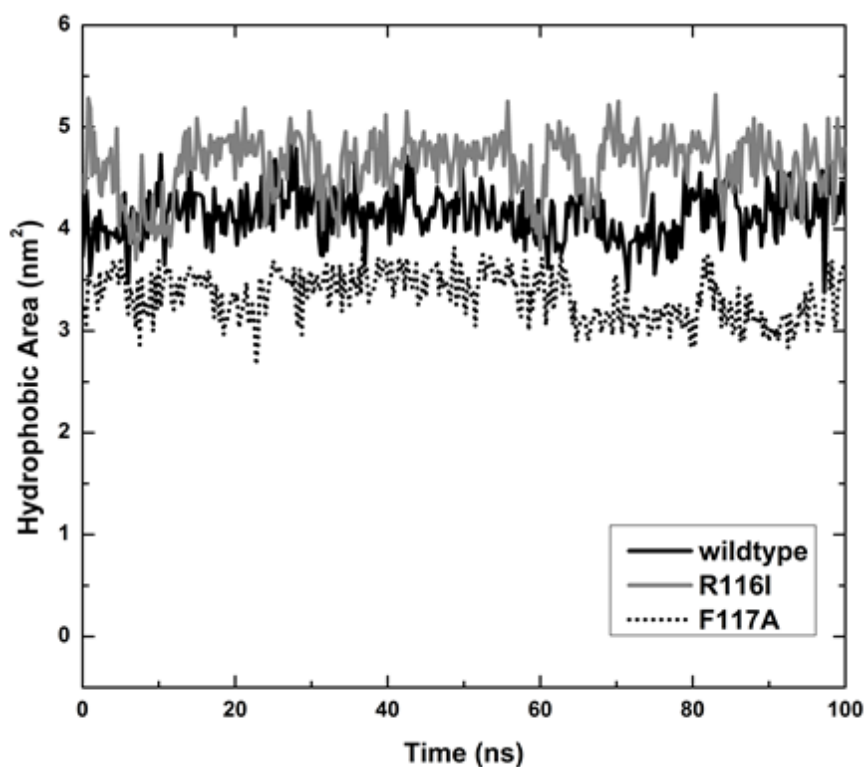


圖 4.2.13 HyDC 與其突變種之殘基 115-118 在 343K 下疏水面積隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.13.tif)

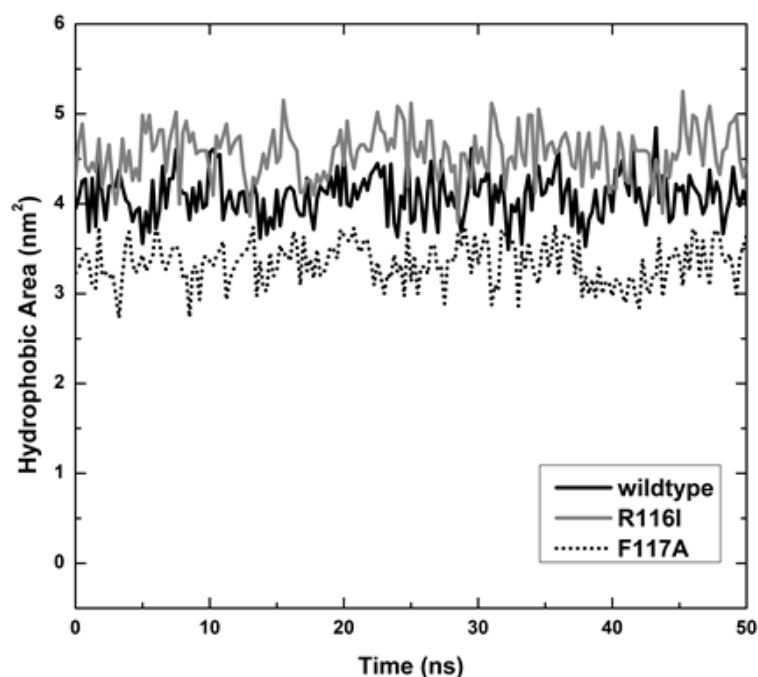


圖 4.2.14 HyDC 與其突變種之殘基 115-118 在 425K 下疏水面積隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.2.14.tif)

本節以分子動力學模擬八個突變種分別在 pH7、343K 與 pH7、425K 環境下結構的變化，在綜合比較結構穩定性與聚集程度後，分別找出最容易產生聚集的突變種 R116I 以及最不容易產生聚集的 F117A，而這兩個突變種的結構穩定性都與原態差異不大。進一步分析其結構的改變，發現雖然二級結構無顯著的改變，但在 RMSF 結果可以推測聚集傾向與不穩定區域 115-118 的擾動呈正相關，且疏水區域的裸露也與聚集傾向有關聯，過去研究亦有提及疏水區域的裸露為造成蛋白質聚集的原因之一[280]。

定性分析可能造成聚集的原因，在 R116I 的組別中，當突變成有 β -branched 的 isoleucine(I)時，因其立體障礙可能會使原本的結構瓦解，進而有更大的擾動以及疏水區域的裸露，進而產生聚集，此部分在 4.1.2 節已有提過。而在 F117A 的組別中，因殘基 115-118 在前人研究中就被判別為較不穩定的區域，當突變成無任何特殊官能基且穩定的 alanine(A)時，就有可能使得原本的聚集傾向下降。過去研究

即有以 alanine(A)作為突變目標，掃描對蛋白質功能扮演重要角色的殘基為何之技術[281]。

另外，就基因突變的觀點來看，在本研究中所挑選出之 R116I 與 F117A 之兩突變種，根據圖 4.2.15 所示，可以得知不論要進行 R→I 或是 F→A 的突變，至少都需要兩個密碼子(codon)的突變。然而在生物演化過程中，因連續兩個密碼子造成突變的機率是趨近於零的，因此在未來進行實驗來驗證本研究之模擬結果時，需要把問題納入考量才是。

		Second Letter				
		U	C	A	G	
1st letter	U	UUU Phe UUC UUA Leu UUG	UCU UCC Ser UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trp	U C A G
	C	CUU CUC Leu CUA CUG	CCU CCC Pro CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC Arg CGA CGG	U C A G
	A	AUU AUC Ile AUA AUG Met	ACU ACC Thr ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
	G	GUU GUC Val GUA GUG	GCU GCC Ala GCA GCG	GAU Asp GAC GAA Glu GAG	GGU GGC Gly GGA GGG	U C A G

圖 4.2.15 胺基酸與對應之基因密碼子(codon)一覽

<http://openwetware.org/wiki/Image:Codontable2.gif>

4.3 探討小分子抑制 HyDC 聚集之能力與機制

除了以單點突變探討 HyDC 不穩定區域(殘基 115-118)的結構穩定性與聚集傾向之外,本研究還利用分子對接模擬,嘗試在 16 種小分子中選出最具有抑制 HyDC 聚集能力的小分子,以作為治療因 HyDC 聚集引起之白內障的可能療方。本節先將 HyDC 與 16 種小分子進行對接,選出能量較穩定者,再以分子動力學於 343K、pH 7 的環境下模擬之,並且進一步分析兩者間的作用力。

4.3.1 分子對接結果

在本節分子對接模擬使用的 HyDC 結構,為過去本實驗室以酸性條件(pH 2)模擬在 196 ns 之結構,其結構在此時有四組新生成的 β -sheet,包括本研究的目標區域 115-118,如 2.2.3 節與圖 2.2.6 所述;而小分子的選用則參照表 2.3.4。在對接之前需先初步判斷對接的位置,HyDC 共有 6 個能進行對接的位置(site),其中僅有 site 5 和 site 6 和殘基 115-118 相鄰,前者屬於 HyDC 外側的位置,而後者則是處於 HyDC 內的位置,如圖 4.3.1 所示。

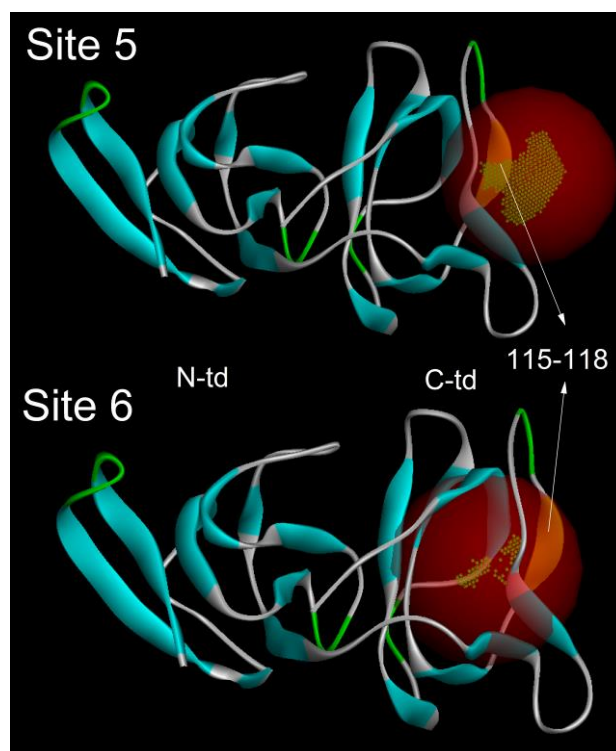


圖 4.3.1 HyDC 與小分子對接之位置示意圖
(檔案路徑名稱: Sean_Thesis_Figure/4.3.1.tif)

對接結果如表 4.3.1 所示，列出 HyDC 與每個小分子對接後的能量最穩定的前三名。比較 site 5 和 site 6，發現在 site 5 的位置與小分子都可以進行對接，而 site 6 卻僅有肉鹼(L-carnitine inner salt)、鄰香蘭素(Ortho-vanillin)以及 5-羥基-1,4-萘醌(5-hydroxy-1,4-naphthoquinone)能進行對接，而這三個小分子的分子量分別為 161.20、174.15 與 152.15，比起其他小分子都小很多(見表 2.3.4)，因此可以合理推測位於內側位置的 site 6，所具有的空間可能不夠大到足以容納其他較大的小分子，以致於無法進行對接；而處於在外側位置的 site 5 則是較無空間上的問題，所以都可以對接。另外，從能量值來分析，小分子對接在 site 5 的能量值大多都小於 0，而能順利對接在 site 6 的小分子，其能量值都非常大，分別約為 800、300、700 kJ/mol，表示 site 5 比起 site 6 應是較為穩定，因此接著針對 site 5 來分析。

表 4.3.1 HyDC 與小分子對接能量一覽

小分子名稱 (結構檔來源)	Cdocker Energy (kJ/mol) (Site 5)			Cdocker Energy (kJ/mol) (Site 6)		
	Pose1	Pose2	Pose3	Pose1	Pose2	Pose3
槲皮素 Quercetin (pdb code: 5AUW)[282]	-28.77	-27.10	-26.30	N/A	N/A	N/A
氯原酸 Chlorogenic acid (PubChem, CID=1794427)	-15.76	-15.41	-14.85	N/A	N/A	N/A
肉鹼 L-carnitine inner salt (PubChem, CID=10917)	-23.46	-19.10	-18.59	867.87	905.51	N/A
2,9-二甲基-4,10-菲羅 Neocuproine free base (PubChem, CID= 65237)	-0.17	1.34	1.42	N/A	N/A	N/A
酚紅 Phenol red (PubChem, CID=4766)	-3.83	-2.32	-1.58	N/A	N/A	N/A
鄰香蘭素 Ortho-vanillin (PubChem, CID=8991)	-12.66	-11.55	-11.40	308.02	323.74	362.75
2,2-二羥基二苯甲酮 2,2-dihydroxy-Benzophenone (PubChem, CID=69150)**	9.48	9.31	8.77	N/A	N/A	N/A
5-羥基-1,4-萘醌 5-hydroxy-1,4-naphthoquinone (pdb code: 3B7J)[283]	-12.64	-12.56	-11.95	677.84	694.56	737.34



天青藍 C Azure C (PubChem, CID=68277)	-8.89	-8.49	-7.96	N/A	N/A	N/A
精氨酸乙基醚 L-arginine ethyl ester dihydrochloride (PubChem, CID= 45073230)	-26.25	-24.93	-24.70	N/A	N/A	N/A
楊梅素 Myricetin (pdb code: 4GQR)**[284]	-32.98	-32.18	-29.26	N/A	N/A	N/A
迷迭香酸 Rosmarinic acid (pdb code: 4PWI)**[285]	-33.53	-29.55	-28.74	N/A	N/A	N/A
白藜蘆醇 Resveratrol (pdb code: 5CR1)**[286]	-12.58	-11.83	-11.22	N/A	N/A	N/A
薑黃素 Curcumin (pdb code: 4PME)**[287]	-31.12	-29.37	-24.71	N/A	N/A	N/A
亞甲基藍 Methylene blue (pdb code: 2YVE)[288]	-0.44	0.40	0.42	N/A	N/A	N/A
兒茶素 Epigallocatechin gallate (pdb code: 3OOB)[289]	-32.34	-31.41	-30.42	N/A	N/A	N/A

*N/A 表示對接能量過高或是有幾何障礙導致無法有效對接

**表示原結構有誤，吾人予以修正之

根據對接能量值來挑選，找出前五個最穩定的HyDC-小分子複合物(complex)，分別有薑黃素(Curcumin)、兒茶素(Epigallocatechin gallate)、楊梅素(Myricetin)、槲皮素(Quercetin)以及迷迭香酸(Rosmarinic acid)，並初步分析其對接後的交互作用力，如圖 4.3.2 所示。雖然這五種小分子與 HyDC 形成的作用力不盡相同，但仔細比較可以發現，這些小分子大多具有苯環以及 OH 基團，前者能與苯環類的殘基產生 pi-pi 作用力，後者則可能有靜電作用力或氫鍵的形成；且吾人發現在苯環上外接兩個 OH 基團時(即二酚)，比較容易產生 pi-pi 作用力與氫鍵作用力，但若苯環上外接三個 OH 基團時(即三酚)，則幾乎沒有上述的作用力。由此可知，雖然小分子都有類似的結構，但對於能產生作用力的要求仍是有一定的限制。另外，小分子除了與殘基 115-118 鄰近之外，也與殘基 102-106 接近且有作用力，亦可能對於穩定此複合物有重要的影響。

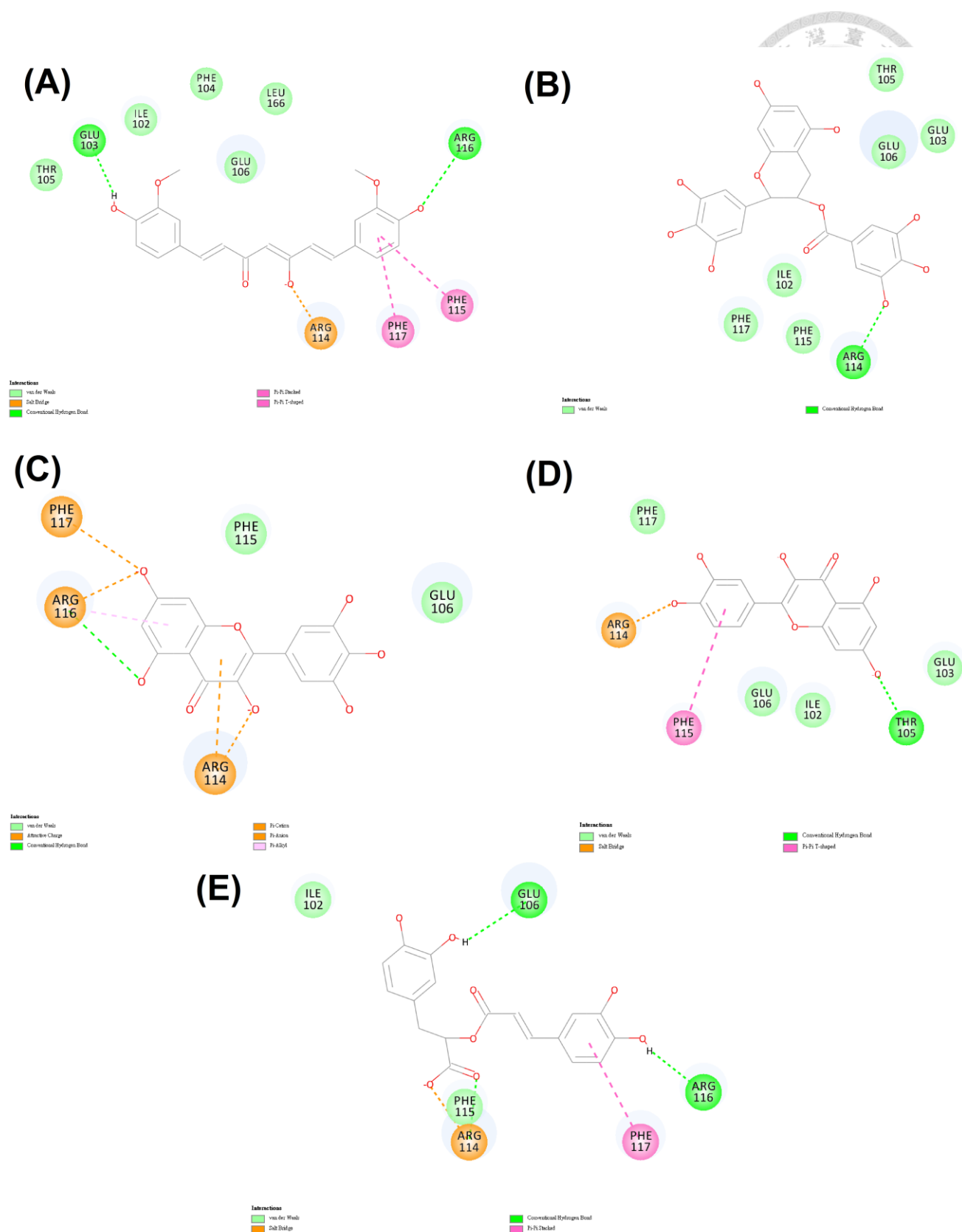


圖 4.3.2 HyDC 與小分子對接作用力分析結果(A)薑黃素(Curcumin) (B)兒茶素(Epigallocatechin gallate) (C)楊梅素(Myricetin) (D)槲皮素(Quercetin) (E)迷迭香酸(Rosmarinic acid)
(檔案路徑名稱：Sean_Thesis_Figure/4.3.2.tif)

在初步分析 HyDC 與五種小分子的對接作用力後，接著以分子動力學模擬此複合物在動態環境下的穩定程度以及結構的變化。其中環境設定為 pH 7、343 K。



4.3.2 質心距離與最短距離分析結果

為了判斷 HyDC 與小分子在模擬過程中，是否仍保有作用力，吾人分別計算 HyDC 與小分子兩者的質心(center of mass)距離與最短距離，結果如圖 4.3.3 與圖 4.3.4。從結果可以發現，五種小分子在模擬過程中與 HyDC 的距離並沒有太大的變化，表示兩者間仍有一定程度作用力存在。然而，其複合物存在何種作用力以及作用力的強度則在 4.3.6 節討論之。

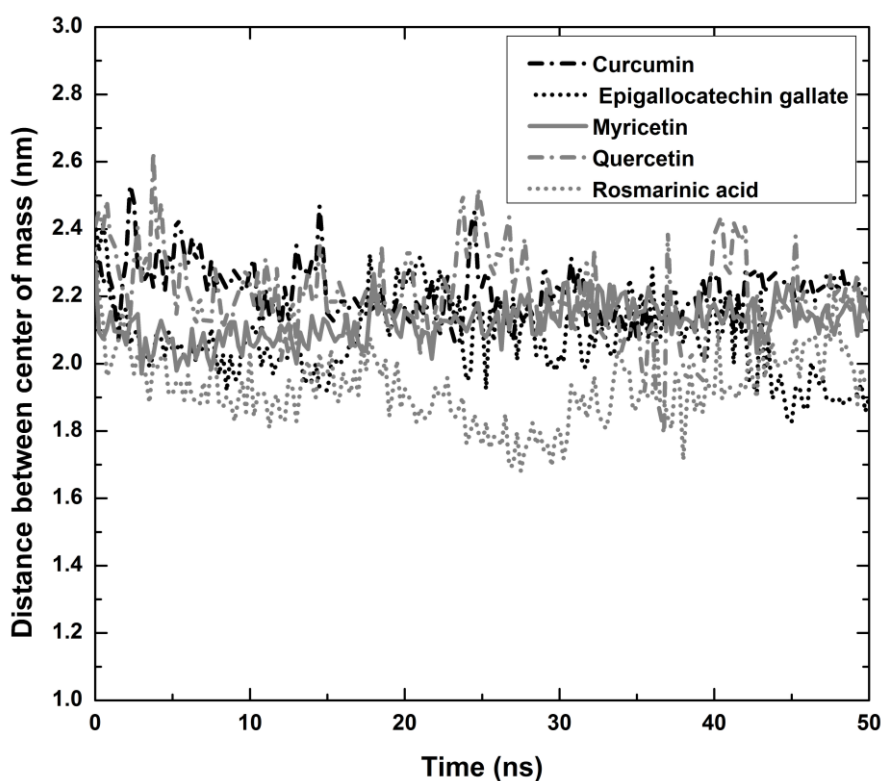


圖 4.3.3 HyDC 與小分子的質心距離隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.3.3.tif)

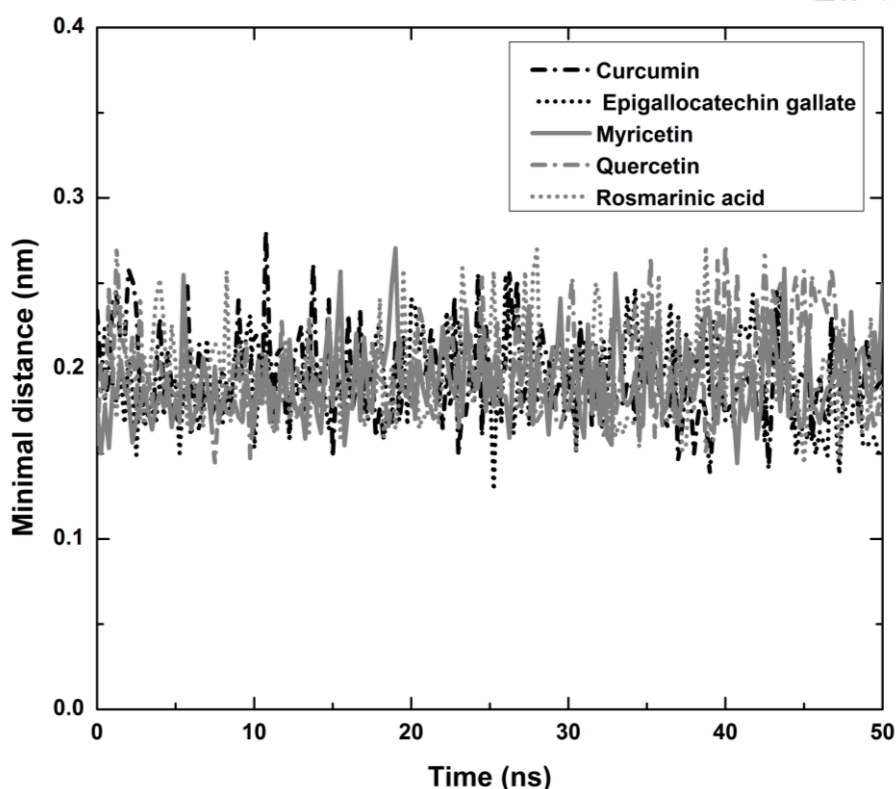


圖 4.3.4 H γ DC 與小分子的最短距離隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.3.4.tif)

4.3.3 RMSD 分析結果

為了瞭解 H γ DC 與小分子對接後的結構穩定性，以 RMSD 分析 H γ DC 結構的穩定性，結果如圖 4.3.5。在只有 H γ DC 的組別中，其 RMSD 值在 50 ns 時約上升至 0.34 左右；在有對接小分子的組別中，唯有對接槲皮素(Quercetin)的 RMSD 值明顯低於只有 H γ DC 的組別，約為 0.25~0.3 間，其他小分子對接之 RMSD 值都與只有 H γ DC 的組別沒有太大的差異，約為 0.35 左右。由此結果，吾人可以推測槲皮素(Quercetin)對於穩定 HGDC 結構的能力是最好的，但詳細情況仍以分析其對接作用力作為判斷的依據。

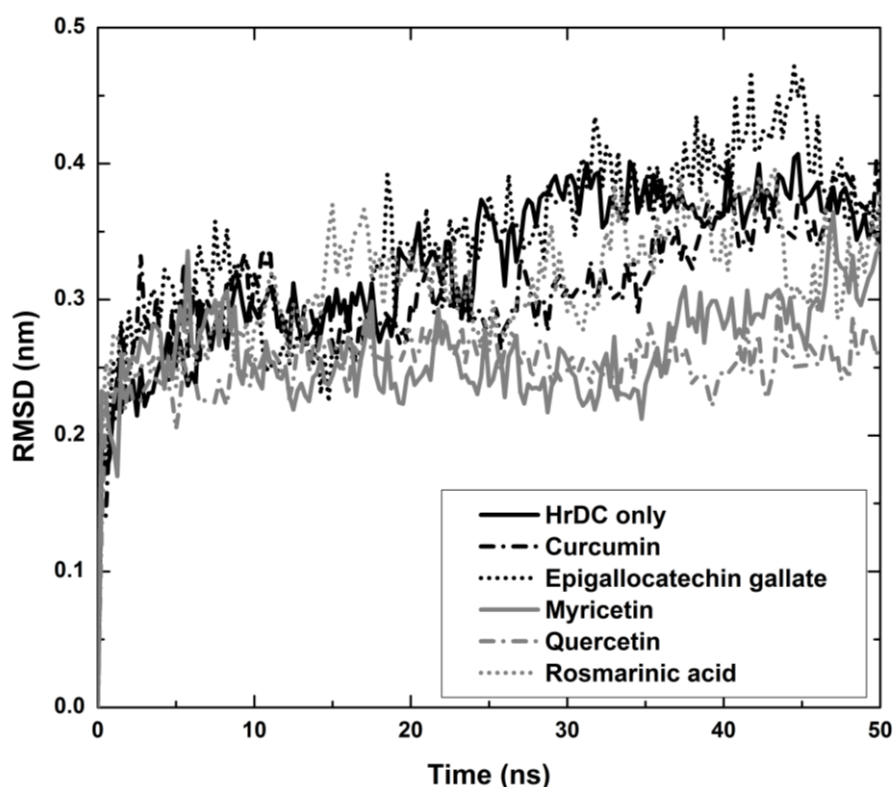


圖 4.3.5 H γ DC 與小分子對接之 RMSD 值隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.3.5.tif)

4.3.4 凡德瓦力與氫鍵作用力分析結果

蘭納-瓊斯勢(Lennard-Jones potential)能用來模擬兩個電中性的分子或原子間交互作用勢能，此處計算整個系統的交互勢能(potential energy)，結果如圖 4.3.6。只有 H γ DC 的組別於 0 ns 時，其能量明顯高於有對接小分子的組別，表示在有對接小分子的系統是較穩定的。隨著時間演進，只有 H γ DC 組別的能量雖有下降的趨勢，但在 50 ns 時仍略高於其他組別。除此之外，比較對接不同小分子的組別，其能量並沒有明顯的差異，可見就凡德瓦力分析而言，其穩定 H γ DC 的能力都是差不多的。

氫鍵作用力在分析在蛋白質與小分子對接時，也屬於非常重要的一種作用力。在此計算蛋白質與小分子在模擬過程中出現氫鍵的次數，並取平均值，結果如表

4.3.2 所示。根據結果可以得知楊梅素(Myricetin)與迷迭香酸(Rosmarinic acid)在模擬過程中具有的氫鍵較多，平均次數分別為 1.74 與 1.60。

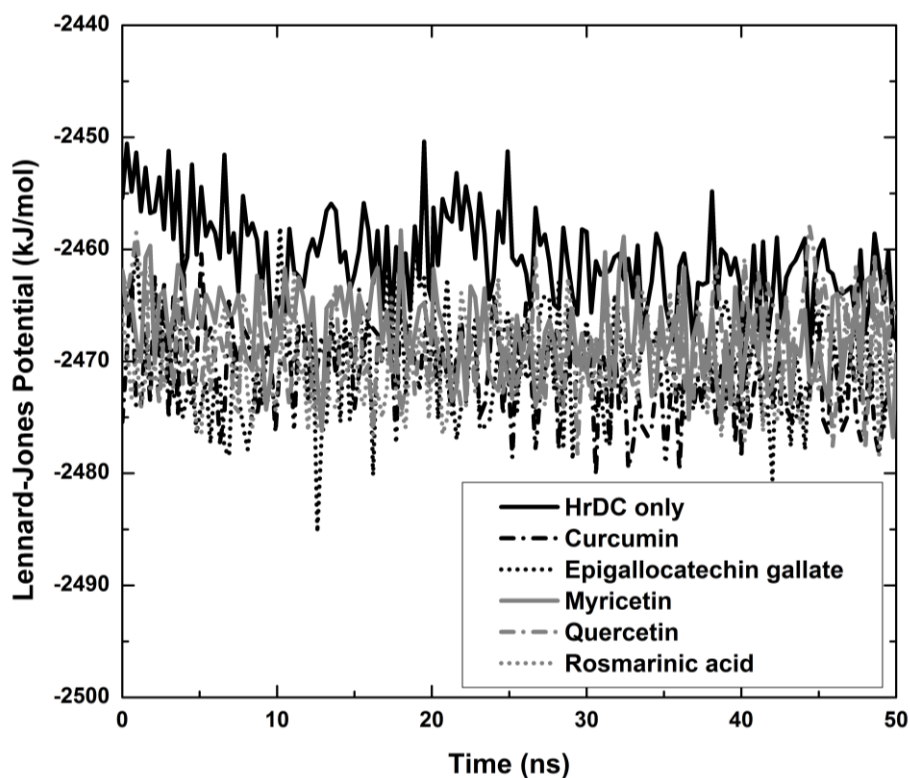


圖 4.3.6 H_γDC 與小分子對接之凡德瓦力隨時間變化圖
(檔案路徑名稱：Sean_Thesis_Figure/4.3.6.tif)

表 4.3.2 H_γDC 與小分子在模擬過程中出現氫鍵之平均次數

小分子	薑黃素 (Curcumin)	兒茶素 (Epigallocatechin gallate)	楊梅素 (Myricetin)	槲皮素 (Quercetin)	迷迭香酸 (Rosmarinic acid)
平均氫鍵 數目	1.12	1.36	1.74	1.32	1.60

然而，蛋白質與小分子的作用力並非只有凡德瓦力與氫鍵作用力，還包括疏水影響力(hydrophobic effect)以及庫倫靜電作用力等，吾人將在 4.3.6 節將深入探討之。

4.3.5 RMSF 分析結果

RMSF 能用來分析各殘基在模擬過程中各自的擾動程度，結果如圖 4.3.7，並分別探討主要對接的兩組殘基 102-106 以及 115-118。在殘基 102-106 中，其擾動值都約在 0.1~0.15 間，不僅較鄰近殘基的擾動小之外，不同小分子對接的組別也無明顯的差異。從 H γ DC 有無對接小分子的結果差異不大來推測，此區域可能本身即很穩定。在殘基 115-118 中，與薑黃素(Curcumin)對接組別 RMSF 值上升至 0.3 左右，而與兒茶素(Epigallocatechin gallate)對接組別 RMSF 值也上升至 0.2 左右，但在只有 H γ DC 的組別之 RMSF 值僅約在 0.1~0.15 間，其餘的對接組別則略小於只有 H γ DC 的組別。由上述結果可以推測，薑黃素(Curcumin)與兒茶素(Epigallocatechin gallate)雖然能與 H γ DC 進行對接，但可能因其產生的交互作用力不如其他小分子，使得結構上有較大的擾動。

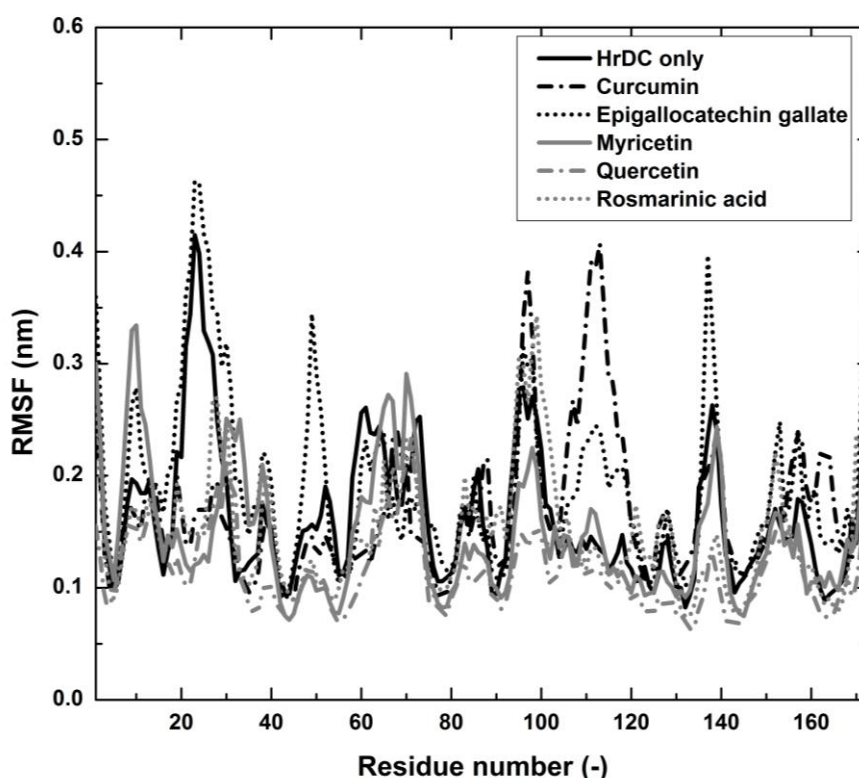


圖 4.3.7 H γ DC 與小分子對接之 RMSF 值結果圖
(檔案路徑名稱：Sean_Thesis_Figure/4.3.7.tif)

4.3.6 對接作用力分析結果

前幾節主要以概略性的方式探討 HyDC 與小分子間的相互關係，在本節將分析兩者間所有的作用力，並加以比較之。吾人將分析的作用力包括疏水影響力(hydrophobic effect)、氫鍵(hydrogen-bond)作用力、庫倫靜電力(coulomb electrostatic force)、苯環-苯環(aromatic-aromatic)作用力、苯環-硫(aromatic-sulphur)作用力以及陽離子-苯環(cation-aromatic)作用力，其相關的計算標準如表 4.3.3。

表 4.3.3 對接作用力計算標準一覽表

作用力	截斷半徑 (Å)
疏水(hydrophobic) [290]	5.0
氫鍵(hydrogen-bond)[291]	3.5 for O.N ; 4.0 for S
庫倫靜電力(coulomb electrostatic force)[290]	6.0
苯環-苯環(aromatic-aromatic)[292]	4.5 to 7.0
苯環-硫(aromatic-sulphur)[293]	5.3
陽離子-苯環(cation-aromatic)[294]	6.0

為了涵蓋上述所有的作用力，吾人將截斷半徑設定為 6.0 Å，計算 HyDC 所有重原子(即非氫原子)與小分子之所有原子的距離，當兩原子的距離小於 6.0 Å 即被判定為 contact。此法根據模擬的軌跡檔，計算 HyDC 每個重原子與小分子每個原子被判定為 contact 的次數，除以總計算次數，即可算出 contact 的分率。結果如圖 4.3.8 至圖 4.3.12，其中有產生作用力的殘基主要包括 91-93、99-108、114-117 等三個區間。另外，將 HyDC 重原子的編號與對照表列於表 4.3.4。

表 4.3.4 HyDC 重原子編號與殘基對照表

HyDC 重原子 編號	對應殘基	HyDC 重原子 編號	對應殘基	HyDC 重原子 編號	對應殘基
742-749	LEU91	843-850	ILE102	895-900	CYS108
750-761	TYR92	851-859	GLU103	938-948	ARG114
762-770	GLU93	860-870	PHE104	949-959	PHE115
822-825	GLY99	871-877	THR105	960-969	ARG116
826-834	GLN100	878-886	GLU106	971-981	PHE117
835-842	MET101	887-894	ASP107		

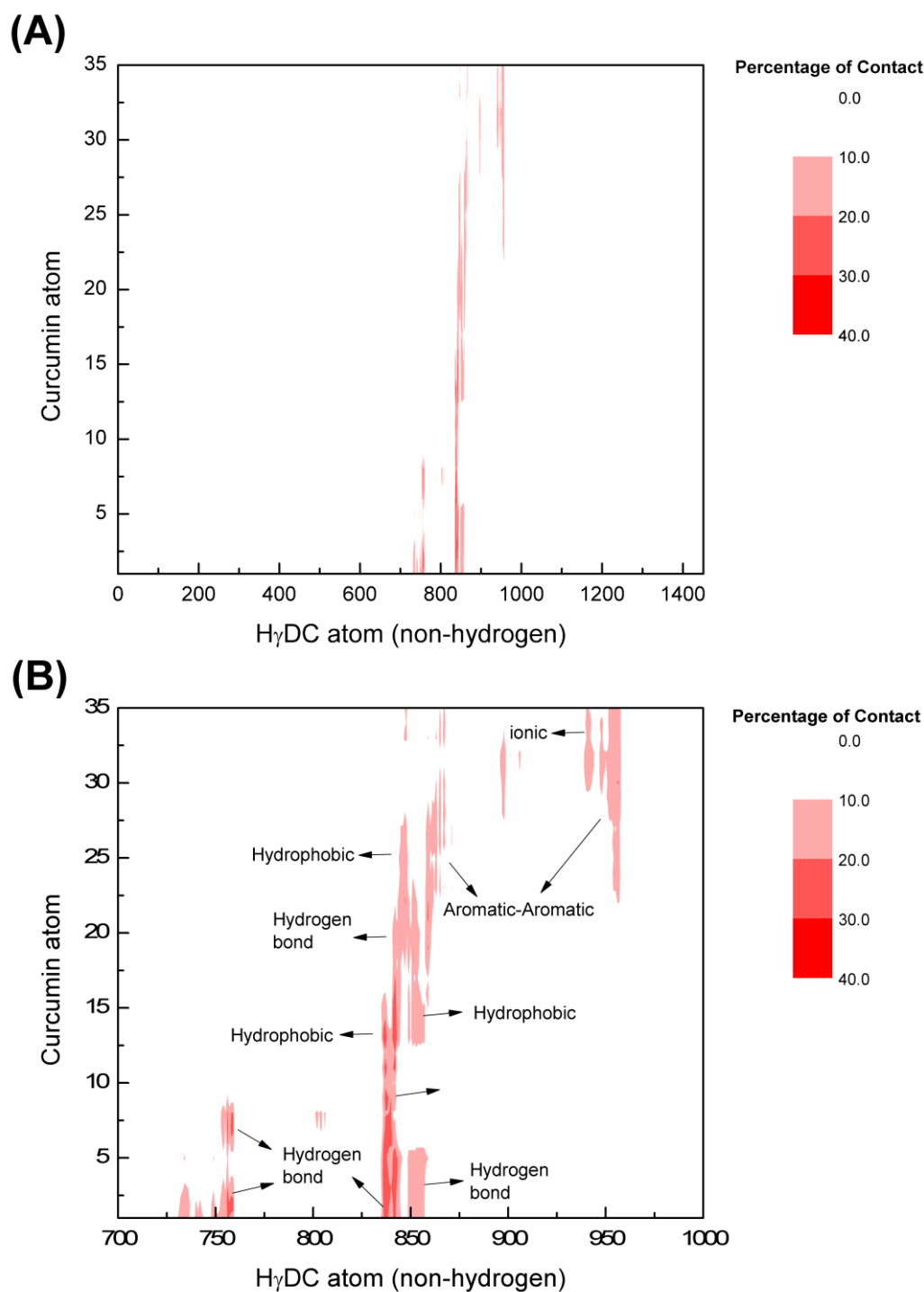


圖 4.3.8 H γ DC 與薑黃素(Curcumin)交互作用力之分析圖(A)全圖 (B)局部放大圖
作用殘基: LEU91(742-749)、TYR92(750-761)、MET101(835-842)、ILE102(843-850)、
GLU103(851-859)、PHE104(860-870)、ARG114(938-948)、PHE115(949-959)
(檔案路徑名稱: Sean_Thesis_Figure/4.3.8.tif)

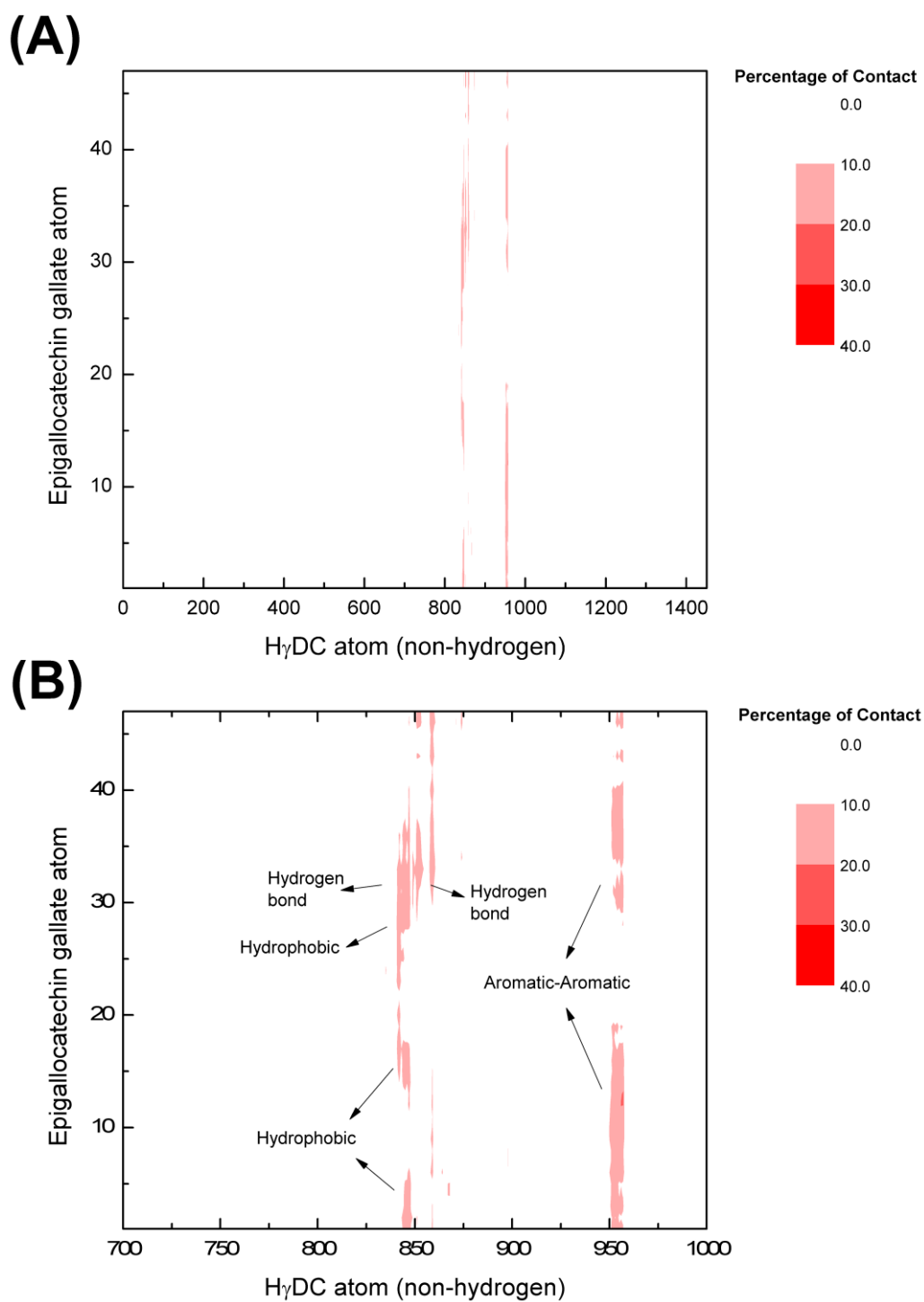


圖 4.3.9 H γ DC 與兒茶素(Epigallocatechin gallate)交互作用力之分析圖

(A)全圖 (B)局部放大圖

作用殘基：MET101(835-842)、ILE102(843-850)、GLU103(851-859)、
PHE115(949-959)

(檔案路徑名稱：Sean_Thesis_Figure/4.3.9.tif)

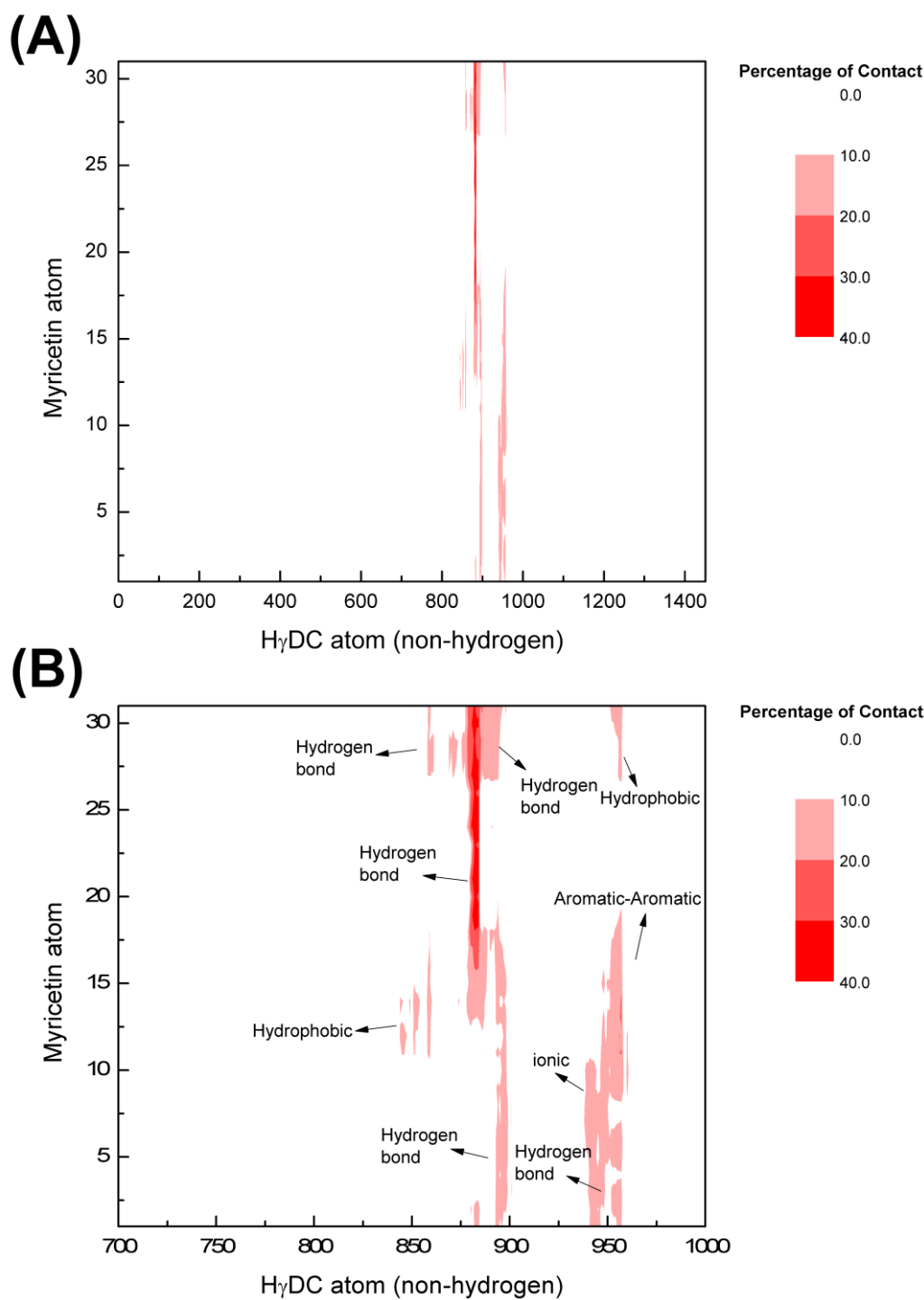


圖 4.3.10 H₇DC 與楊梅素(Myricetin)交互作用力之分析圖(A)全圖 (B)局部放大圖
 作用殘基：MET101(835-842)、ILE102(843-850)、GLU103(851-859)、
 PHE104(860-870)、GLU106(878-886)、ASP107(887-894)、CYS108(895-900)、
 ARG114(938-948)、PHE115(949-959)
 (檔案路徑名稱：Sean_Thesis_Figure/4.3.10.tif)

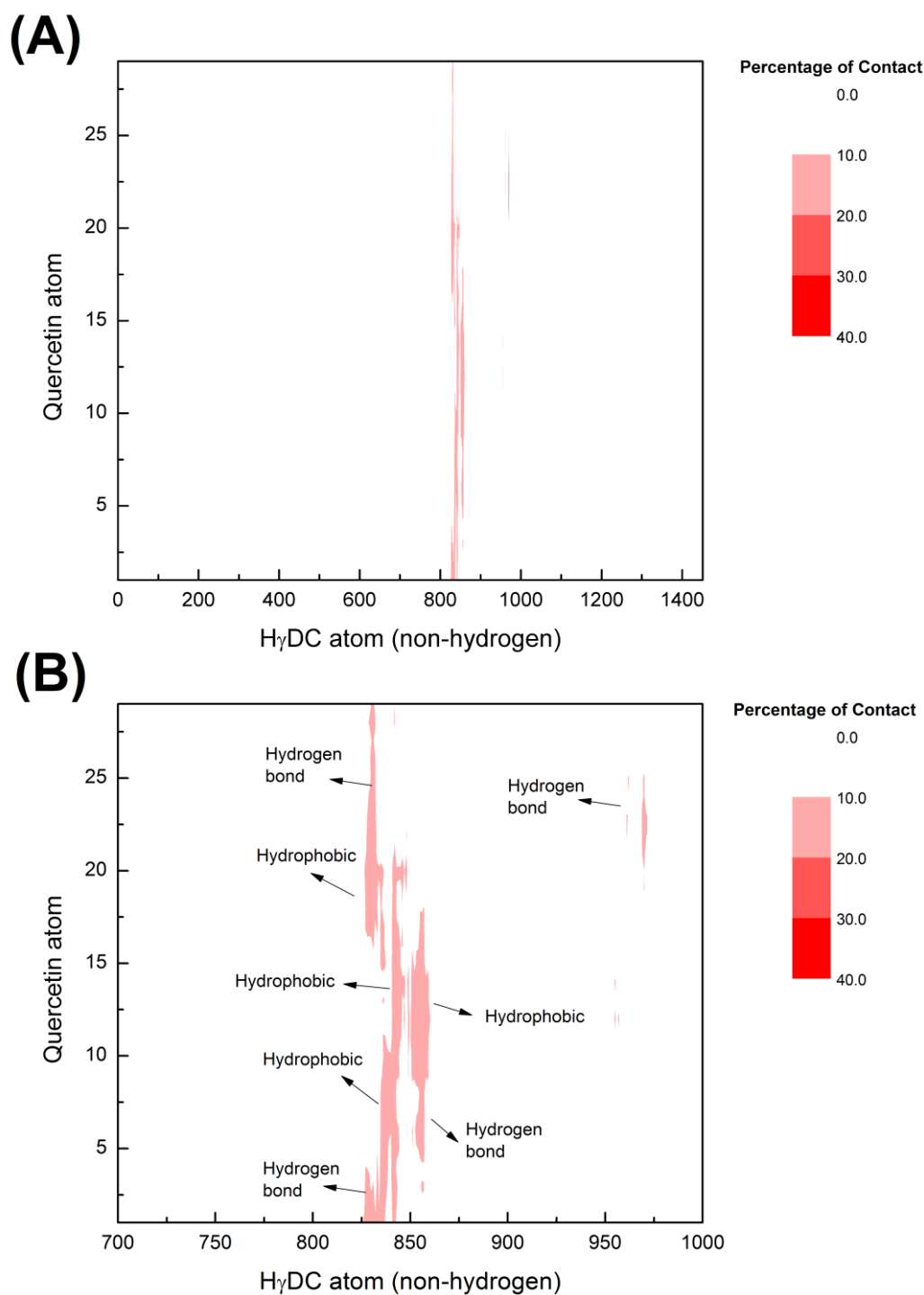


圖 4.3.11 H γ DC 與槲皮素(Quercetin)交互作用力之分析圖(A)全圖 (B)局部放大圖
 作用殘基：GLY99(822-825)、GLN100(826-834)、MET101(835-842)、
 ILE102(843-850)、GLU103(851-859)、ARG116(960-969)
 (檔案路徑名稱：Sean_Thesis_Figure/4.3.11.tif)

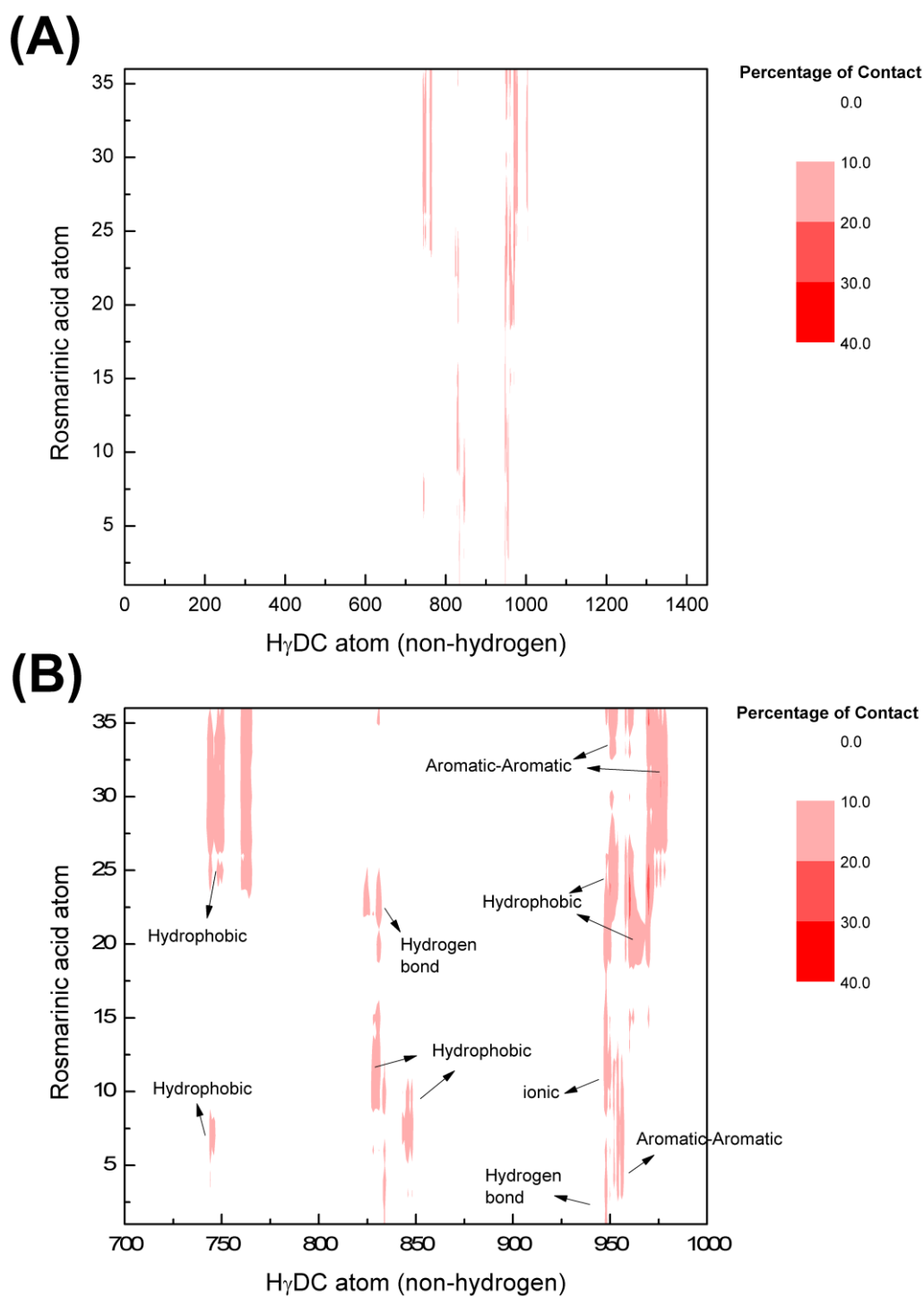


圖 4.3.12 H₇DC 與迷迭香酸(Rosmarinic acid)交互作用力之分析圖

(A)全圖 (B)局部放大圖

作用殘基：LEU91(742-749)、GLU93(762-770)、GLY99(822-825)、GLN100(826-834)、
ILE102(843-850)、ARG114(938-948)、PHE115(949-959)、ARG116(960-969)、
PHE117(971-981)

(檔案路徑名稱：Sean_Thesis_Figure/4.3.12.tif)

比較 HyDC 與不同小分子的對接作用力圖，發現不同的小分子會與不同的殘基接觸，並產生不同的作用力。吾人透過三個面向來比較結果：



1. 接觸 HyDC 殘基的數目：若小分子能與越多的殘基相互作用，表示此小分子的大小與位向都能與蛋白質殘基的形狀匹配(match)，即可能會有較強的作用力。根據此面向，薑黃素(Curcumin)與迷迭香酸(Rosmarinic acid)是可能具有較多作用力的小分子。
2. 小分子的使用程度：若小分子有較多的原子(官能基)能與 HyDC 作用，越能將此小分子的功能都發揮到淋漓盡致，即表示 HyDC 與小分子間可能會有較強的作用力。根據此面向，薑黃素(Curcumin)、楊梅素(Myricetin)與槲皮素(Quercetin)幾乎所有的原子都與 HyDC 產生作用力，可能為作用力較多的小分子；其餘兩個小分子則有少部分區域未與 HyDC 產生作用力。
3. 與不穩定殘基(115-118)的鍵結情形：由於 HyDC 之殘基 115-118 為過去研究中不穩定的區域，所以若能與此區域產生較多的作用力，則較能有效地使蛋白質變得穩定，進而抑制其聚集。根據此面向，除槲皮素(Quercetin)外，其餘四種小分子都會與殘基 F115 有作用力，且迷迭香酸(Rosmarinic acid)還與 R116 以及 F117 有作用力，合理推測其應為較有可能使 HyDC 變得更穩定的小分子。

由於上述第二點描述之小分子使用程度在相互比較下的差異並沒有很大，吾人不予採用。因此，綜合上述第一點與第三點，吾人推測迷迭香酸(Rosmarinic acid)為本研究的 16 種小分子中，最具有穩定 HyDC 結構的能力。

第五章 結論與建議



5.1 結論

本論文使用電腦模擬的方式探討 HyDC 結構之不穩定區域 115-118，吾人先以生物資訊工具預測單點突變後的結構穩定性與聚集傾向，接著利用分子動力學模擬 HyDC 與其突變種之結構隨時間的變化情形，藉以瞭解此不穩定區域在 HyDC 中扮演的角色；最後則是使用分子對接的方式，透過分析 HyDC 與小分子對接的交互作用力，找出最有可能抑制 HyDC 聚集的小分子。吾人將結果分為下列三個部分：

1. 以生物資訊工具預測 HyDC 及其突變種之穩定性與聚集傾向

吾人以地毯式搜索的方式，將不穩定區域 115-118 的 76 個突變種，以多種生物資訊工具分別預測其結構穩定性與聚集傾向。在結構穩定性方面，由於不同生物資訊工具的預測結果不盡相同，在進行數據整合後，分別得出每個殘基之最穩定突變種(F115W、R116I、F117Y、N118I)與最不穩定之突變種(F115A、R116S、F117A、N118H)；吾人推測當把特定殘基進行單點突變時，造成結構不穩定的因素主要是由於側鏈性質的改變，包括立體障礙從大變成小，極性變成非極性，或是帶電變成不帶電等。在聚集傾向方面，吾人同樣地整合多種生物資訊工具，找出一致具有高聚集傾向的突變種 R116I 與 N118I；吾人推測當突變成具有 β -branched 的胺基酸時，容易因為其較大的立體結構使得其區間的結構瓦解，使得結構較為開展。

2. 以分子動力學探討 HyDC 及其突變種之結構的改變

吾人進一步以分子動力學模擬 HyDC 與上述 8 個突變種的結構變化，同時分析 RMSD、原態接觸以及聚集傾向等結果。綜合比較發現，F115A 的穩定性、開展程

度以及聚集傾向都與原態沒有太大的區別；F115W 與 F117Y 為結構較穩定且較無開展的突變種，然而前者的聚集傾向與原態沒有明顯的差異，後者的聚集傾向則是減少的；R116I、R116S 與 N118I 結構穩定性和開展程度都與原態差不多，但卻有較大的聚集傾向，其中 R116I 的聚集傾向極為顯著地大；而 F117A 結構穩定性和原態也差異不大，但其聚集傾向卻是減少的；最後是 N118H 的突變種，其結構會變得比較不穩定，而聚集傾向是略為減少的。

接著，吾人挑選聚集傾向最大的 R116I 以及最小的 F117A 為目標突變種，進一步分析其對於不穩定區域 115-118 的影響。在 RMSF 結果中，不穩定區域 115-118 的擾動情形依大小排列分別為 R116I、原態、F117A；在 DSSP 二級結構分析中，三者無明顯的差異；在疏水程度方面，依大小排列分別為 R116I、原態、F117A。因此吾人推論聚集傾向的程度，可能與不穩定區域的擾動情形以及疏水程度有所關連，而其擾動可能還沒有達到能影響二級結構的程度。

分析可能造成聚集的原因，在 R116I 的組別中，當突變成有 β -branched 的 isoleucine 時，因其立體障礙可能使原本的結構瓦解，此部分在第 1 點已提及。而在 F117A 的組別中，因殘基 115-118 在前人研究中就被判別為較不穩定的區域，當突變成無任何特殊官能基且穩定的 alanine 時，就有可能使得原本的聚集傾向下降。

3. 探討小分子抑制 HyDC 聚集之能力與機制

吾人以分子對接的方式，將本研究選用的 16 種小分子與 HyDC 進行對接，透過能量值的排序找出最穩定的 5 種複合物，分別有薑黃素(Curcumin)、兒茶素(Epigallocatechin gallate)、楊梅素(Myricetin)、槲皮素(Quercetin)以及迷迭香酸(Rosmarinic acid)。接著，以分子動力學模擬整體結構的改變以及分析 HyDC 與上述 5 種小分子間可能產生的作用力。吾人發現雖這 5 種小分子都能與 HyDC 有不同的作用力，且都能使 HyDC 結構變得較穩定，但其中迷迭香酸(Rosmarinic acid)

與本研究探討之不穩定殘基 115-118 產生的作用力最多，推測其為本研究的 16 種小分子中，最具有穩定 H γ DC 結構的能力。

因此，根據以上結果，吾人希望藉由分子層面的信息，以更了解 H γ DC 不穩定區域的行為，並嘗試找出較可能抑制 H γ DC 聚集的藥物，提供未來預防或治療白內障等相關研究的可能方向。

5.2 建議

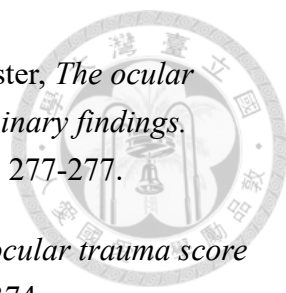
本研究主要著重於電腦模擬的部分，雖然能得出 H γ DC 在分子層面的訊息，但最終仍須搭配實驗結果相互印證才是。期許未來能以實驗的方式，對 H γ DC 進行突變並測試其結構的改變，以及將本論文提及的多種小分子進行抑制 H γ DC 聚集行為的測試，並相互檢視模擬與實驗的相同與不相同處，藉以更了解 H γ DC 的行為，希望能對於未來預防白內障形成相關的研究有更進一步的幫助。

Reference

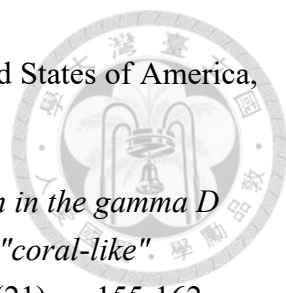


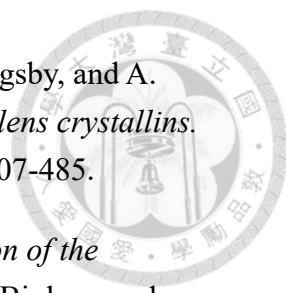
1. Voleti, V.B. and J.P. Hubschman, *Age-related eye disease*. Maturitas, 2013. **75**(1): p. 29-33.
2. Allen, D. and A. Vasavada, *Cataract and surgery for cataract*. British Medical Journal, 2006. **333**(7559): p. 128-132.
3. Gimbel, H.V. and A.A. Dardzhikova, *Consequences of waiting for cataract surgery*. Current Opinion in Ophthalmology, 2011. **22**(1): p. 28-30.
4. Pascolini, D. and S.P. Mariotti, *Global estimates of visual impairment: 2010*. British Journal of Ophthalmology, 2012. **96**(5): p. 614-618.
5. John Forrester, A.D., Paul McMenamin, William Lee *The Eye: Basic Sciences in Practice*. 1996, London: W.B. Saunders Company
6. Kardong, K., *Vertebrates: Comparative anatomy, function, evolution*. 5th ed. 2008, Boston: McGraw-Hill. pp.676-677.
7. Hecht, E., *Optics*. 2nd ed. 1987, Addison Wesley.
8. Shirk J. S., S.M., Scribner D., Fleet E., Stroman R., Baer E., Hilter A., *NRL Review*. 2006. p. 53-61.
9. Flaugh, S.L., M.S. Kosinski-Collins, and J. King, *Interdomain side-chain interactions in human gamma D crystallin influencing folding and stability*. Protein Science, 2005. **14**(8): p. 2030-2043.
10. Mahendiran, K., C. Elie, J.C. Nebel, A. Ryan, and B.K. Pierscionek, *Primary sequence contribution to the optical function of the eye lens*. Scientific Reports, 2014. **4**.
11. Moreau, K.L. and J.A. King, *Protein misfolding and aggregation in cataract disease and prospects for prevention*. Trends in Molecular Medicine, 2012. **18**(5): p. 273-282.
12. Asbell, P.A., L. Dualan, J. Mindel, D. Brocks, M. Ahmad, and S. Epstein, *Age-related cataract*. Lancet, 2005. **365**(9459): p. 599-609.
13. Santana, A. and M. Waiswol, *The genetic and molecular basis of congenital cataract*. Arquivos Brasileiros De Oftalmologia, 2011. **74**(2): p. 136-142.

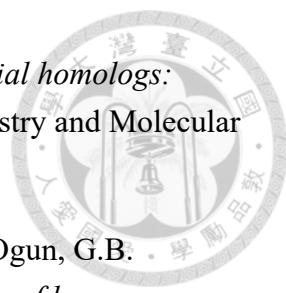
- 
14. Robinson N.E. and R. A.B., *Molecular Clocks: Deamidation of Asparaginy and Glutaminyl Residues in Peptides and Proteins*. 2004, Althouse Press: Cave Junction, Ore. .
 15. Lindner, H. and W. Helliger, *Age-dependent deamidation of asparagine residues in proteins*. Experimental Gerontology, 2001. **36**(9): p. 1551-1563.
 16. Ray, N.J., D. Hall, and J.A. Carver, *Deamidation of N76 in human gamma S-crystallin promotes dimer formation*. Biochimica Et Biophysica Acta-General Subjects, 2016. **1860**(1): p. 315-324.
 17. Taylor, H.R., *Epidemiology of age-related cataract*. Eye, 1999. **13**: p. 445-448.
 18. Stevens, V.J., C.A. Rouzer, V.M. Monnier, and A. Cerami, *Diabetic Cataract Formation - Potential Role of Glycosylation of Lens Crystallins*. Proceedings of the National Academy of Sciences of the United States of America, 1978. **75**(6): p. 2918-2922.
 19. Cordes, F.C., *Galactosemia Cataract - a Review - Report of a Case*. American Journal of Ophthalmology, 1960. **50**(6): p. 1151-1158.
 20. Falls, H.F., G.H. Lowrey, and W.C. Anderson, *Galactosemia Cataract - Report of a Case*. American Journal of Ophthalmology, 1951. **34**(9): p. 1271-1274.
 21. Schutte, E. and M. Reim, *Cataract by Galactokinase Deficiency*. Klinische Monatsblatter Fur Augenheilkunde, 1975. **166**(4): p. 553-556.
 22. Endres, W. and Y.S. Shin, *Cataract and Metabolic Disease*. Journal of Inherited Metabolic Disease, 1990. **13**(4): p. 509-516.
 23. Park, S. and E.H. Lee, *Association between metabolic syndrome and age-related cataract*. International Journal of Ophthalmology, 2015. **8**(4): p. 804-811.
 24. Sabanayagam, C., J.J. Wang, P. Mitchell, A.G. Tan, E.S. Tai, T. Aung, S.M. Saw, and T.Y. Wong, *Metabolic Syndrome Components and Age-Related Cataract: The Singapore Malay Eye Study*. Investigative Ophthalmology & Visual Science, 2011. **52**(5): p. 2397-2404.
 25. Tasman W. and J. E.A., *Traumatic cataract*. Duane's Clinical Ophthalmology, 1997. **1**: p. 13-4.

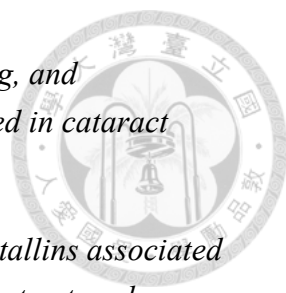
- 
26. Kuhn, F., R. Maisiak, R. Morris, C.D. Witherspoon, and V. Mester, *The ocular trauma score (OTS): Developmental considerations and preliminary findings*. Investigative Ophthalmology & Visual Science, 1997. **38**(4): p. 277-277.
 27. Acar, U., O.Y. Tok, D.E. Acar, A. Burcu, and F. Ornek, *A new ocular trauma score in pediatric penetrating eye injuries*. Eye, 2011. **25**(3): p. 370-374.
 28. Zhu, L.L., Z.C. Wu, F. Dong, J. Feng, D.H. Lou, C.X. Du, and P.F. Ren, *Two kinds of ocular trauma score for paediatric traumatic cataract in penetrating eye injuries*. Injury-International Journal of the Care of the Injured, 2015. **46**(9): p. 1828-1833.
 29. Klein, B.E.K., R. Klein, and K.E. Lee, *Incidence of age-related cataract over a 10-year interval - The Beaver Dam Eye Study*. Ophthalmology, 2002. **109**(11): p. 2052-2057.
 30. Truscott, R., *Age-related human nuclear cataract. Blindness due to inexorable protein deterioration*. Acta Ophthalmologica, 2013. **91**.
 31. Burch, J.W., D.C. Mair, G.M. Meny, G. Moroff, S.S.T. Ching, M.A. Naidoff, E. Steuer, S. Loftus, J.R. Armstrong, and B.E.K. Klein, *Prevalence of posterior subcapsular cataract in granulocyte donors*. Transfusion, 2004. **44**(9): p. 8A-8A.
 32. Yu, X.N., D.N. Lyu, X.R. Dong, J.L. He, and K. Yao, *Hypertension and Risk of Cataract: A Meta-Analysis*. Plos One, 2014. **9**(12).
 33. Hanson, S.R.A., D.L. Smith, and J.B. Smith, *Deamidation and disulfide bonding in human lens gamma-crystallins*. Experimental Eye Research, 1998. **67**(3): p. 301-312.
 34. Lapko, V.N., D.L. Smith, and J.B. Smith, *Methylation and carbamylation of human gamma-crystallins*. Protein Science, 2003. **12**(8): p. 1762-1774.
 35. Li, D.Y., R.F. Borkman, R.H. Wang, and J. Dillon, *Mechanisms of Photochemically Produced Turbidity in Lens Protein Solutions*. Experimental Eye Research, 1990. **51**(6): p. 663-669.
 36. Stephan, D.A., E. Gillanders, D. Vanderveen, D. Freas-Lutz, G. Wistow, A.D. Baxevanis, C.M. Robbins, A. VanAuken, M.I. Quesenberry, J. Bailey-Wilson, S.H.H. Juo, J.M. Trent, L. Smith, and M.J. Brownstein, *Progressive juvenile-onset punctate cataracts caused by mutation of the gamma D-crystallin gene*.

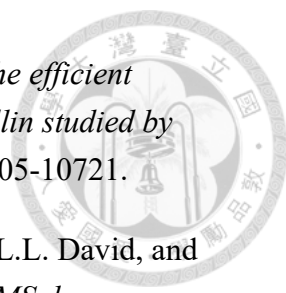
Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(3): p. 1008-1012.

- 
37. Mackay, D.W., U.P. Andley, and A. Shiels, *A missense mutation in the gamma D crystallin gene (CRYGD) associated with autosomal dominant "coral-like" cataract linked to chromosome 2q*. Molecular Vision, 2004. **10**(21): p. 155-162.
38. Kmoch, S., J. Brynda, B. Asfaw, K. Bezouska, P. Novak, P. Rezacova, L. Ondrova, M. Filipec, J. Sedlacek, and M. Elleder, *Link between a novel human gamma D-crystallin allele and a unique cataract phenotype explained by protein crystallography*. Human Molecular Genetics, 2000. **9**(12): p. 1779-1786.
39. Heon, E., M. Priston, D.F. Schorderet, G.D. Billingsley, P.O. Girard, N. Lubsen, and F.L. Munier, *The gamma-crystallins and human cataracts: A puzzle made clearer*. American Journal of Human Genetics, 1999. **65**(5): p. 1261-1267.
40. Santhiya, S.T., M.S. Manohar, D. Rawley, P. Vijayalakshmi, P. Namperumalsamy, P.M. Gopinath, J. Loster, and J. Graw, *Novel mutations in the gamma-crystallin genes cause autosomal dominant congenital cataracts*. Journal of Medical Genetics, 2002. **39**(5): p. 352-358.
41. Schafheimer, N., Z. Wang, K. Schey, and J. King, *Tyrosine/Cysteine Cluster Sensitizing Human gamma D-Crystallin to Ultraviolet Radiation-Induced Photoaggregation in Vitro*. Biochemistry, 2014. **53**(6): p. 979-990.
42. Bollinger, K.E. and R.H.S. Langston, *What can patients expect from cataract surgery?* Cleveland Clinic Journal of Medicine, 2008. **75**(3): p. 193-+.
43. Ashwin, P.T., S. Shah, and J.S. Wolffsohn, *Advances in cataract surgery*. Clinical and Experimental Optometry, 2009. **92**(4): p. 333-342.
44. Hoehenwarter, W., J. Klose, and P.R. Jungblut, *Eye lens proteomics*. Amino Acids, 2006. **30**(4): p. 369-389.
45. Jester, J.V., *Corneal crystallins and the development of cellular transparency*. Seminars in Cell & Developmental Biology, 2008. **19**(2): p. 82-93.
46. Andley, U.P., *Crystallins in the eye: Function and pathology*. Progress in Retinal and Eye Research, 2007. **26**(1): p. 78-98.

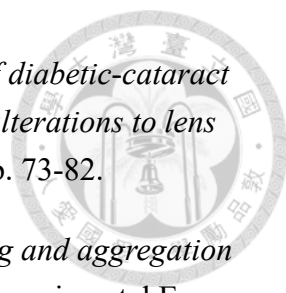
- 
47. Bloemendal, H., W. de Jong, R. Jaenicke, N.H. Lubsen, C. Slingsby, and A. Tardieu, *Ageing and vision: structure, stability and function of lens crystallins*. Progress in Biophysics & Molecular Biology, 2004. **86**(3): p. 407-485.
 48. Dejong, W.W., J.A.M. Leunissen, and C.E.M. Voorter, *Evolution of the Alpha-Crystallin Small Heat-Shock Protein Family*. Molecular Biology and Evolution, 1993. **10**(1): p. 103-126.
 49. Horwitz, J., *Alpha-Crystallin Can Function as a Molecular Chaperone*. Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(21): p. 10449-10453.
 50. Bloemendal, H. and W.W. Dejong, *Lens Proteins and Their Genes*. Progress in Nucleic Acid Research and Molecular Biology, 1991. **41**: p. 259-281.
 51. de Jong, W.W., G.J. Caspers, and J.A.M. Leunissen, *Genealogy of the alpha-crystallin - small heat-shock protein superfamily*. International Journal of Biological Macromolecules, 1998. **22**(3-4): p. 151-162.
 52. Braun, N., M. Zacharias, J. Peschek, A. Kastenmuller, J. Zou, M. Hanzlik, M. Haslbeck, J. Rappsilber, J. Buchner, and S. Weinkauf, *Multiple molecular architectures of the eye lens chaperone alpha B-crystallin elucidated by a triple hybrid approach*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(51): p. 20491-20496.
 53. Groenen, P.J.T.A., K.B. Merck, W.W. Dejong, and H. Bloemendal, *Structure and Modifications of the Junior Chaperone Alpha-Crystallin - from Lens Transparency to Molecular Pathology*. European Journal of Biochemistry, 1994. **225**(1): p. 1-19.
 54. Augusteyn, R.C., *alpha-crystallin: a review of its structure and function*. Clinical and Experimental Optometry, 2004. **87**(6): p. 356-366.
 55. Maulucci, G., M. Papi, G. Arcovito, and M. De Spirito, *The Thermal Structural Transition of alpha-Crystallin Inhibits the Heat Induced Self-Aggregation*. Plos One, 2011. **6**(5).
 56. Nagaraj, R.H., R.B. Nahomi, N.H. Mueller, C.T. Raghavan, D.A. Ammar, and J.M. Petrash, *Therapeutic potential of alpha-crystallin*. Biochimica Et Biophysica Acta-General Subjects, 2016. **1860**(1): p. 252-257.

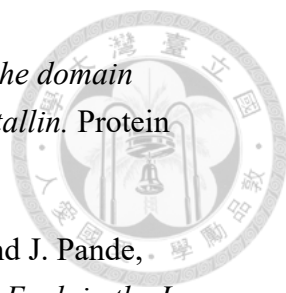
- 
57. Jaenicke, R. and C. Slingsby, *Lens crystallins and their microbial homologs: Structure, stability, and function*. Critical Reviews in Biochemistry and Molecular Biology, 2001. **36**(5): p. 435-499.
 58. Basak, A., O. Bateman, C. Slingsby, A. Pande, N. Asherie, O. Ogun, G.B. Benedek, and J. Pande, *High-resolution X-ray crystal structures of human gamma D crystallin (1.25 angstrom) and the R58H mutant (1.15 angstrom) associated with aculeiform cataract*. Journal of Molecular Biology, 2003. **328**(5): p. 1137-1147.
 59. Norledge, B.V., E.M. Mayr, R. Glockshuber, O.A. Bateman, C. Slingsby, R. Jaenicke, and H.P.C. Driessen, *The X-ray structures of two mutant crystallin domains shed light on the evolution of multi-domain proteins*. Nature Structural Biology, 1996. **3**(3): p. 267-274.
 60. Lampi, K.J., J.T. Oxford, H.P. Bachinger, T.R. Shearer, L.L. David, and D.M. Kapfer, *Deamidation of human beta B1 alters the elongated structure of the dimer*. Experimental Eye Research, 2001. **72**(3): p. 279-288.
 61. Berbers, G.A.M., W.A. Hoekman, H. Bloemendal, W.W. Dejong, T. Kleinschmidt, and G. Braunitzer, *Homology between the Primary Structures of the Major Bovine Beta-Crystallin Chains*. European Journal of Biochemistry, 1984. **139**(3): p. 467-479.
 62. Zigler, J.S., J. Horwitz, and J.H. Kinoshita, *Human Beta-Crystallin .1. Comparative Studies on the Beta-1-Crystallins Beta-2-Crystallins and Beta-3-Crystallins*. Experimental Eye Research, 1980. **31**(1): p. 41-55.
 63. Zigler, J.S., M. Valapala, P. Shang, S. Hose, M.F. Goldberg, and D. Sinha, *beta A3/A1-crystallin and persistent fetal vasculature (PW) disease of the eye*. Biochimica Et Biophysica Acta-General Subjects, 2016. **1860**(1): p. 287-298.
 64. Slingsby, C., B. Norledge, A. Simpson, O.A. Bateman, G. Wright, H.P.C. Driessen, P.F. Lindley, D.S. Moss, and B. Bax, *X-ray diffraction and structure of crystallins*. Progress in Retinal and Eye Research, 1997. **16**(1): p. 3-29.
 65. Brakenhoff, R.H., H.J.M. Aarts, F.H. Reek, N.H. Lubsen, and J.G.G. Schoenmakers, *Human Gamma-Crystallin Genes - a Gene Family on Its Way to Extinction*. Journal of Molecular Biology, 1990. **216**(3): p. 519-532.

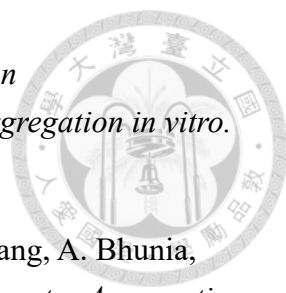
- 
66. Kosinski-Collins, M.S. and J. King, *In vitro* unfolding, refolding, and polymerization of human gamma D crystallin, a protein involved in cataract formation. *Protein Science*, 2003. **12**(3): p. 480-490.
67. Das, P., J.A. King, and R.H. Zhou, *Aggregation of gamma-crystallins associated with human cataracts via domain swapping at the C-terminal beta-strands*. *Proceedings of the National Academy of Sciences of the United States of America*, 2011. **108**(26): p. 10514-10519.
68. Ji, F.L., L.M.I. Koharudin, J. Jung, and A.M. Gronenborn, *Crystal structure of the cataract-causing P23T D-crystallin mutant*. *Proteins-Structure Function and Bioinformatics*, 2013. **81**(9): p. 1493-1498.
69. Kong, F.R. and J. King, *Contributions of aromatic pairs to the folding and stability of long-lived human gamma D-crystallin*. *Protein Science*, 2011. **20**(3): p. 513-528.
70. Slingsby, C. and G.J. Wistow, *Functions of crystallins in and out of lens: Roles in elongated and post-mitotic cells*. *Progress in Biophysics & Molecular Biology*, 2014. **115**(1): p. 52-67.
71. Hemmingsen, J.M., K.M. Gernert, J.S. Richardson, and D.C. Richardson, *The Tyrosine Corner - a Feature of Most Greek Key Beta-Barrel Proteins*. *Protein Science*, 1994. **3**(11): p. 1927-1937.
72. Hamill, S.J., E. Cota, C. Chothia, and J. Clarke, *Conservation of folding and stability within a protein family: The tyrosine corner as an evolutionary cul-de-sac*. *Journal of Molecular Biology*, 2000. **295**(3): p. 641-649.
73. Spector, A. and D. Roy, *Disulfide-Linked High Molecular-Weight Protein Associated with Human Cataract*. *Proceedings of the National Academy of Sciences of the United States of America*, 1978. **75**(7): p. 3244-3248.
74. Xia, Z., Z.X. Yang, T. Huynh, J.A. King, and R.H. Zhou, *UV-radiation Induced Disruption of Dry-Cavities in Human gamma D-crystallin Results in Decreased Stability and Faster Unfolding*. *Scientific Reports*, 2013. **3**.
75. Chen, J.J., S.L. Flaugh, P.R. Callis, and J. King, *Mechanism of the highly efficient quenching of tryptophan fluorescence in human gamma D-crystallin*. *Biochemistry*, 2006. **45**(38): p. 11552-11563.

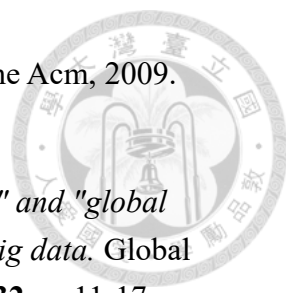
- 
76. Chen, J.J., D. Toptygin, L. Brand, and J. King, *Mechanism of the efficient tryptophan fluorescence quenching in human gamma D-crystallin studied by time-resolved fluorescence*. Biochemistry, 2008. **47**(40): p. 10705-10721.
 77. Searle, B.C., S. Dasari, P.A. Wilmarth, M. Turner, A.P. Reddy, L.L. David, and S.R. Nagalla, *Identification of protein modifications using MS/MS de novo sequencing and the OpenSea alignment algorithm*. Journal of Proteome Research, 2005. **4**(2): p. 546-554.
 78. Wilmarth, P.A., S. Tanner, S. Dasari, S.R. Nagalla, M.A. Riviere, V. Bafna, P.A. Pevzner, and L.L. David, *Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: Does deamidation contribute to crystallin insolubility?* Journal of Proteome Research, 2006. **5**(10): p. 2554-2566.
 79. MacCoss, M.J., W.H. McDonald, A. Saraf, R. Sadygov, J.M. Clark, J.J. Tasto, K.L. Gould, D. Wolters, M. Washburn, A. Weiss, J.I. Clark, and J.R. Yates, *Shotgun identification of protein modifications from protein complexes and lens tissue*. Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(12): p. 7900-7905.
 80. Hanson, S.R.A., A. Hasan, D.L. Smith, and J.B. Smith, *The major in vivo modifications of the human water-insoluble lens crystallins are disulfide bonds, deamidation, methionine oxidation and backbone cleavage*. Experimental Eye Research, 2000. **71**(2): p. 195-207.
 81. Kosinski-Collins, M.S., S.L. Flaugh, and J. King, *Probing folding and fluorescence quenching in human gamma D crystallin Greek key domains using triple tryptophan mutant proteins*. Protein Science, 2004. **13**(8): p. 2223-2235.
 82. Olivares-Quiroz, L., Guzmán-López, Orlando, Jardón-Valadez, *Physical Biology of Proteins and Peptides*. 2015.
 83. Pande, A., O. Annunziata, N. Asherie, O. Ogun, G.B. Benedek, and J. Pande, *Decrease in protein solubility and cataract formation caused by the Pro23 to Thr mutation in human gamma D-crystallin*. Biochemistry, 2005. **44**(7): p. 2491-2500.
 84. Evans, P., K. Wyatt, G.J. Wistow, O.A. Bateman, B.A. Wallace, and C. Slingsby, *The P23T cataract mutation causes loss of solubility of folded gamma D-crystallin*. Journal of Molecular Biology, 2004. **343**(2): p. 435-444.

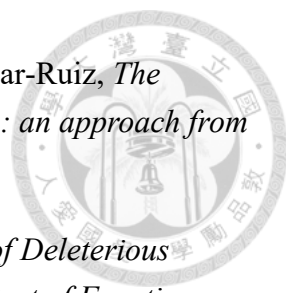
- 
85. Hejtmancik, J.F., *The genetics of cataract: Our vision becomes clearer*. American Journal of Human Genetics, 1998. **62**(3): p. 520-525.
 86. Flaugh, S.L., I.A. Mills, and J. King, *Glutamine deamidation destabilizes human gamma D-crystallin and lowers the kinetic barrier to unfolding*. Journal of Biological Chemistry, 2006. **281**(41): p. 30782-30793.
 87. Talla, V., N. Srinivasan, and D. Balasubramanian, *Visualization of in situ intracellular aggregation of two cataract-associated human gamma-crystallin mutants: Lose a tail, lose transparency*. Investigative Ophthalmology & Visual Science, 2008. **49**(8): p. 3483-3490.
 88. Schafheimer, N. and J. King, *Tryptophan Cluster Protects Human D-Crystallin from Ultraviolet Radiation-Induced Photoaggregation In Vitro*. Photochemistry and Photobiology, 2013. **89**(5): p. 1106-1115.
 89. Moran, S.D., T.Q.O. Zhang, S.M. Decatur, and M.T. Zanni, *Amyloid Fiber Formation in Human gamma D-Crystallin Induced by UV-B Photodamage*. Biochemistry, 2013. **52**(36): p. 6169-6181.
 90. Wu, J.W., M.E. Chen, W.S. Wen, W.A. Chen, C.T. Li, C.K. Chang, C.H. Lo, H.S. Liu, and S.S.S. Wang, *Comparative Analysis of Human gamma D-Crystallin Aggregation under Physiological and Low pH Conditions*. Plos One, 2014. **9**(11).
 91. Varma, S.D., K.R. Hegde, and S. Kovtun, *Inhibition of selenite-induced cataract by caffeine*. Acta Ophthalmologica, 2010. **88**(7): p. e245-e249.
 92. Shin, K.H., S.S. Lim, and D.K. Kim, *Studies on the inhibitory effects of medicinal plants on cataract formation part 5 - Effect of byakangelicin, an aldose reductase inhibitor, on galactosemic cataracts, the polyol contents and Na⁺,K(+)ATPase activity in sciatic nerves of streptozotocin-induced diabetic rats*. Phytomedicine, 1998. **5**(2): p. 121-127.
 93. Nagai, R., M. Nagai, S. Shimasaki, J.W. Baynes, and Y. Fujiwara, *Citric acid inhibits development of cataracts, proteinuria and ketosis in streptozotocin (type 1) diabetic rats*. Biochemical and Biophysical Research Communications, 2010. **393**(1): p. 118-122.
 94. Shi, Q., H. Yan, M.Y. Li, and J.J. Harding, *Effect of a combination of carnosine and aspirin eye drops on streptozotocin-induced diabetic cataract in rats*. Molecular Vision, 2009. **15**(228-29): p. 2129-2138.

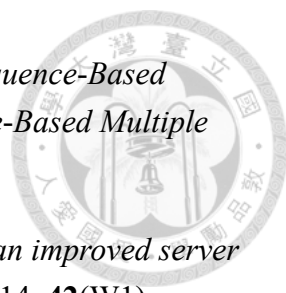
- 
95. Varsha, M.K.N.S., T. Raman, and R. Manikandan, *Inhibition of diabetic-cataract by vitamin K1 involves modulation of hyperglycemia-induced alterations to lens calcium homeostasis*. Experimental Eye Research, 2014. **128**: p. 73-82.
 96. Goulet, D.R., K.M. Knee, and J.A. King, *Inhibition of unfolding and aggregation of lens protein human gamma D crystallin by sodium citrate*. Experimental Eye Research, 2011. **93**(4): p. 371-381.
 97. Gong, B., L.Y. Zhang, D.S.C. Lam, C.P. Pang, and G.H.F. Yam, *Sodium 4-phenylbutyrate ameliorates the effects of cataract-causing mutant gammaD-crystallin in cultured cells*. Molecular Vision, 2010. **16**(107-10): p. 997-1003.
 98. Wu, J.W.R., C.Y. Kao, L.T.W. Lin, W.S. Wen, J.T. Lai, and S.S.S. Wang, *Human gamma D-crystallin aggregation induced by ultraviolet C irradiation is suppressed by resveratrol*. Biochemical Engineering Journal, 2013. **78**: p. 189-197.
 99. Srivastava, S.K. and N.H. Ansari, *Prevention of Sugar-Induced Cataractogenesis in Rats by Butylated Hydroxytoluene*. Diabetes, 1988. **37**(11): p. 1505-1508.
 100. Linklater, H.A., T. Dzialoszynski, H.L. Mcleod, S.E. Sanford, and J.R. Trevithick, *Modeling Cortical Cataractogenesis .8. Effects of Butylated Hydroxytoluene (Bht) in Reducing Protein Leakage from Lenses in Diabetic Rats*. Experimental Eye Research, 1986. **43**(3): p. 305-314.
 101. Ansari, N.H., A. Bhatnagar, E. Fulep, P. Khanna, and S.K. Srivastava, *Trolox Protects Hyperglycemia-Induced Cataractogenesis in Cultured Rat Lens*. Research Communications in Chemical Pathology and Pharmacology, 1994. **84**(1): p. 93-&.
 102. Creighton, M.O. and J.R. Trevithick, *Cortical Cataract Formation Prevented by Vitamin-E and Glutathione*. Experimental Eye Research, 1979. **29**(6): p. 689-693.
 103. Chen, J., P.R. Callis, and J. King, *Mechanism of the very efficient quenching of tryptophan fluorescence in human gamma D- and gamma S-crystallins: the gamma-crystallin fold may have evolved to protect tryptophan residues from ultraviolet photodamage*. Biochemistry, 2009. **48**(17): p. 3708-16.
 104. Moreau, K.L. and J. King, *Hydrophobic core mutations associated with cataract development in mice destabilize human gammaD-crystallin*. J Biol Chem, 2009. **284**(48): p. 33285-95.

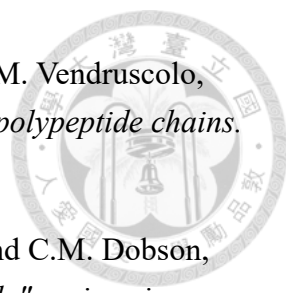
- 
105. Das, P., J.A. King, and R.H. Zhou, *beta-strand interactions at the domain interface critical for the stability of human lens gamma D-crystallin*. Protein Science, 2010. **19**(1): p. 131-140.
 106. Banerjee, P.R., S.S. Puttamadappa, A. Pande, A. Shekhtman, and J. Pande, *Increased Hydrophobicity and Decreased Backbone Flexibility Explain the Lower Solubility of a Cataract-Linked Mutant of YD-Crystallin*. Journal of Molecular Biology, 2011. **412**(4): p. 647-659.
 107. Goulet, D.R., K.M. Knee, and J.A. King, *Inhibition of unfolding and aggregation of lens protein human gamma D crystallin by sodium citrate*. Exp Eye Res, 2011. **93**(4): p. 371-81.
 108. Sahin, E., J.L. Jordan, M.L. Spatara, A. Naranjo, J.A. Costanzo, W.F. Weiss, A.S. Robinson, E.J. Fernandez, and C.J. Roberts, *Computational design and biophysical characterization of aggregation-resistant point mutations for gammaD crystallin illustrate a balance of conformational stability and intrinsic aggregation propensity*. Biochemistry, 2011. **50**(5): p. 628-39.
 109. Wang, L., X. Chen, Y. Lu, J. Wu, B. Yang, and X. Sun, *A novel mutation in gammaD-crystallin associated with autosomal dominant congenital cataract in a Chinese family*. Mol Vis, 2011. **17**: p. 804-9.
 110. Ji, F.L., J.W. Jung, and A.M. Gronenborn, *Structural and Biochemical Characterization of the Childhood Cataract-Associated R76S Mutant of Human gamma D-Crystallin*. Biochemistry, 2012. **51**(12): p. 2588-2596.
 111. Mishra, S., R.A. Stein, and H.S. Mchaourab, *Cataract-linked gamma D-crystallin mutants have weak affinity to lens chaperones alpha-crystallins*. Febs Letters, 2012. **586**(4): p. 330-336.
 112. Moreau, K.L. and J.A. King, *Cataract-causing defect of a mutant gamma-crystallin proceeds through an aggregation pathway which bypasses recognition by the alpha-crystallin chaperone*. PLoS One, 2012. **7**(5): p. e37256.
 113. Ji, F.L., J. Jung, L.M.I. Koharudin, and A.M. Gronenborn, *The Human W42R gamma D-Crystallin Mutant Structure Provides a Link between Congenital and Age-related Cataracts*. Journal of Biological Chemistry, 2013. **288**(1): p. 99-109.
 114. Moran, S.D., T.O. Zhang, S.M. Decatur, and M.T. Zanni, *Amyloid fiber formation in human gammaD-Crystallin induced by UV-B photodamage*. Biochemistry, 2013. **52**(36): p. 6169-81.

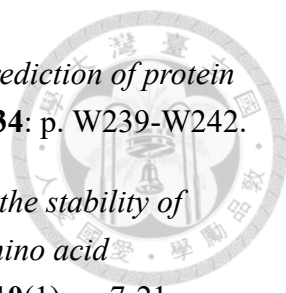
- 
115. Schafheimer, N. and J. King, *Tryptophan cluster protects human gammaD-crystallin from ultraviolet radiation-induced photoaggregation in vitro*. Photochem Photobiol, 2013. **89**(5): p. 1106-15.
116. DiMauro, M.A., S.K. Nandi, C.T. Raghavan, R.K. Kar, B.L. Wang, A. Bhunia, R.H. Nagaraj, and A. Biswas, *Acetylation of Gly1 and Lys2 Promotes Aggregation of Human gamma D-Crystallin*. Biochemistry, 2014. **53**(46): p. 7269-7282.
117. Schafheimer, N., Z. Wang, K. Schey, and J. King, *Tyrosine/cysteine cluster sensitizing human gammaD-crystallin to ultraviolet radiation-induced photoaggregation in vitro*. Biochemistry, 2014. **53**(6): p. 979-90.
118. Yang, Z., Z. Xia, T. Huynh, J.A. King, and R. Zhou, *Dissecting the contributions of beta-hairpin tyrosine pairs to the folding and stability of long-lived human gammaD-crystallins*. Nanoscale, 2014. **6**(3): p. 1797-807.
119. Serebryany, E. and J.A. King, *Wild-type Human gamma D-crystallin Promotes Aggregation of Its Oxidation-mimicking, Misfolding-prone W42Q Mutant*. Journal of Biological Chemistry, 2015. **290**(18): p. 11491-11503.
120. Eugene Serebryany, Takumi Takata, Erika Erickson, Nathaniel Schafheimer, Yongting Wang, and J.A. King, *Aggregation of Trp > Glu point mutants of human gamma-D crystallin provides a model for hereditary or UV-induced cataract*. PROTEIN SCIENCE 2016. **25**: p. pp. 1115-1128.
121. Garcia-Manyes, S., D. Giganti, C.L. Badilla, A. Lezamiz, J. Perales-Calvo, A.E.M. Beedle, and J.M. Fernandez, *Single-molecule Force Spectroscopy Predicts a Misfolded, Domain-swapped Conformation in human D-Crystallin Protein*. Journal of Biological Chemistry, 2016. **291**(8): p. 4226-4235.
122. Quintanar, L., J.A. Dominguez-Calva, E. Serebryany, L. Rivillas-Acevedo, C. Haase-Pettingell, C. Amero, and J.A. King, *Copper and Zinc Ions Specifically Promote Nonamyloid Aggregation of the Highly Stable Human gamma-D Crystallin*. Acs Chemical Biology, 2016. **11**(1): p. 263-272.
123. Chang C.K., W.S.S., Lo C.H., Hsiao H.C., Wu J.W., *Investigation of the early stages of human γ D-crystallin aggregation process*. Journal of Biomolecular Structure and Dynamics 2016: p. 1-13.
124. Boland, M.V., *Big Data, Big Challenges*. Ophthalmology, 2016. **123**(1): p. 7-8.

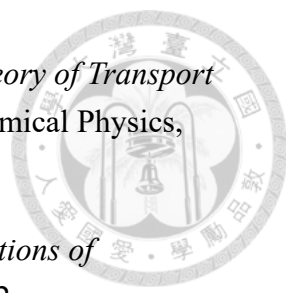
- 
125. Jacobs, A., *The Pathologies of Big Data*. Communications of the Acm, 2009. **52**(8): p. 36-44.
 126. Jang, S.M. and P.S. Hart, *Polarized frames on "climate change" and "global warming" across countries and states: Evidence from Twitter big data*. Global Environmental Change-Human and Policy Dimensions, 2015. **32**: p. 11-17.
 127. *Community cleverness required*. Nature, 2008. **455**(7209): p. 1-1.
 128. Pijanowski, B.C., A. Tayyebi, J. Doucette, B.K. Pekin, D. Braun, and J. Plourde, *A big data urban growth simulation at a national scale: Configuring the GIS and neural network based Land Transformation Model to run in a High Performance Computing (HPC) environment*. Environmental Modelling & Software, 2014. **51**: p. 250-268.
 129. Reichman, O.J., M.B. Jones, and M.P. Schildhauer, *Challenges and Opportunities of Open Data in Ecology*. Science, 2011. **331**(6018): p. 703-705.
 130. Salsbury, F.R., *Molecular dynamics simulations of protein dynamics and their relevance to drug discovery*. Current Opinion in Pharmacology, 2010. **10**(6): p. 738-744.
 131. Durrant, J.D. and J.A. McCammon, *Molecular dynamics simulations and drug discovery*. BMC Biology, 2011. **9**.
 132. Orobitg, M., F. Guirado, F. Cores, J. Lladós, and C. Notredame, *High Performance computing improvements on bioinformatics consistency-based multiple sequence alignment tools*. Parallel Computing, 2015. **42**: p. 18-34.
 133. Reblova, M. and K. Reblova, *RNA secondary structure, an important bioinformatics tool to enhance multiple sequence alignment: a case study (Sordariomycetes, Fungi)*. Mycological Progress, 2013. **12**(2): p. 305-319.
 134. Jones, J., J.K. Field, and J.M. Risk, *A comparative guide to gene prediction tools for the bioinformatics amateur*. International Journal of Oncology, 2002. **20**(4): p. 697-705.
 135. Birdsell, J.A., *Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution*. Molecular Biology and Evolution, 2002. **19**(7): p. 1181-1197.

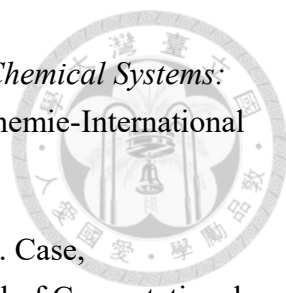
- 
136. Santiesteban-Toca, C.E., G.M. Casanola-Martin, and J.S. Aguilar-Ruiz, *The machine learning techniques in the protein structure prediction: an approach from bioinformatics*. Afinidad, 2014. **71**(567): p. 219-227.
137. Rajith, G.P.D.C.B. and C. Chakraborty, *Predicting the Impact of Deleterious Mutations in the Protein Kinase Domain of FGFR2 in the Context of Function, Structure, and Pathogenesis-a Bioinformatics Approach*. Applied Biochemistry and Biotechnology, 2013. **170**(8): p. 1853-1870.
138. Cao, C.H., J.H. Chen, C.Q. Lyu, J. Yu, W. Zhao, Y. Wang, and D.R. Zou, *Bioinformatics Analysis of the Effects of Tobacco Smoke on Gene Expression*. Plos One, 2015. **10**(12).
139. Shen, E., X.H. Diao, C. Wei, Z.H. Wu, L. Zhang, and B. Hu, *MicroRNAs target gene and signaling pathway by bioinformatics analysis in the cardiac hypertrophy*. Biochemical and Biophysical Research Communications, 2010. **397**(3): p. 380-385.
140. Doolittle, R.F., *The Roots of Bioinformatics in Protein Evolution*. Plos Computational Biology, 2010. **6**(7).
141. Chen, C.W., J. Lin, and Y.W. Chu, *iStable: off-the-shelf predictor integration for predicting protein stability changes*. BMC Bioinformatics, 2013. **14**.
142. Pires, D.E.V., D.B. Ascher, and T.L. Blundell, *DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach*. Nucleic Acids Research, 2014. **42**(W1): p. W314-W319.
143. Yin, S.Y., F. Ding, and N.V. Dokholyan, *Eris: an automated estimator of protein stability*. Nature Methods, 2007. **4**(6): p. 466-467.
144. Yin, S., F. Ding, and N.V. Dokholyan, *Modeling backbone flexibility improves protein stability estimation*. Structure, 2007. **15**(12): p. 1567-1576.
145. Ding, F. and N.V. Dokholyan, *Emergence of protein fold families through rational design (vol 2, pg 725, 2006)*. Plos Computational Biology, 2006. **2**(10): p. 1325-1325.
146. Folkman, L., B. Stantic, and A. Sattar, *Feature-based multiple models improve classification of mutation-induced stability changes*. BMC Genomics, 2014. **15**.

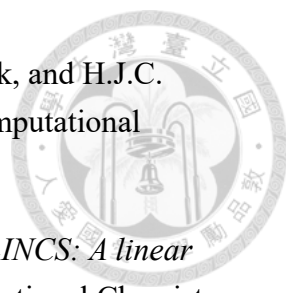
- 
147. Folkman, L., Stantic, B., Sattar, A., Zhou, Y., *EASE-MM: Sequence-Based Prediction of Mutation-Induced Stability Changes with Feature-Based Multiple Models*. Journal of Molecular Biology, 2016.
 148. Walsh, I., F. Seno, S.C.E. Tosatto, and A. Trovato, *PASTA 2.0: an improved server for protein aggregation prediction*. Nucleic Acids Research, 2014. **42**(W1): p. W301-W307.
 149. Fernandez-Escamilla, A.M., F. Rousseau, J. Schymkowitz, and L. Serrano, *Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins*. Nature Biotechnology, 2004. **22**(10): p. 1302-1306.
 150. Rousseau, F., J. Schymkowitz, and L. Serrano, *Protein aggregation and amyloidosis: confusion of the kinds?* Current Opinion in Structural Biology, 2006. **16**(1): p. 118-126.
 151. Linding, R., J. Schymkowitz, F. Rousseau, F. Diella, and L. Serrano, *A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins*. Journal of Molecular Biology, 2004. **342**(1): p. 345-353.
 152. Conchillo-Sole, O., N.S. de Groot, F.X. Aviles, J. Vendrell, X. Daura, and S. Ventura, *AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides*. BMC Bioinformatics, 2007. **8**.
 153. Oliveberg, M., *Waltz, an exciting new move in amyloid prediction*. Nature Methods, 2010. **7**(3): p. 187-188.
 154. Maurer-Stroh, S., M. Debulpaep, N. Kuemmerer, M.L. de la Paz, I.C. Martins, J. Reumers, K.L. Morris, A. Copland, L. Serpell, L. Serrano, J.W.H. Schymkowitz, and F. Rousseau, *Exploring the sequence determinants of amyloid structure using position-specific scoring matrices (vol 7, pg 237, 2010)*. Nature Methods, 2010. **7**(10): p. 855-855.
 155. Garbuzynskiy, S.O., M.Y. Lobanov, and O.V. Galzitskaya, *FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence*. Bioinformatics, 2010. **26**(3): p. 326-332.
 156. Tartaglia, G.G. and M. Vendruscolo, *The Zygggregator method for predicting protein aggregation propensities*. Chemical Society Reviews, 2008. **37**(7): p. 1395-1401.

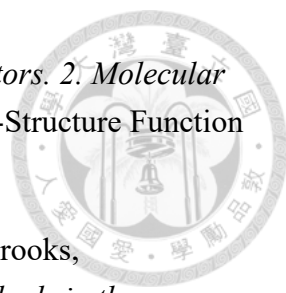
- 
157. Dubay, K.F., A.P. Pawar, F. Chiti, J. Zurdo, C.M. Dobson, and M. Vendruscolo, *Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains*. Journal of Molecular Biology, 2004. **341**(5): p. 1317-1326.
 158. Pawar, A.P., K.F. DuBay, J. Zurdo, F. Chiti, M. Vendruscolo, and C.M. Dobson, *Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases*. Journal of Molecular Biology, 2005. **350**(2): p. 379-392.
 159. Tartaglia, G.G., A.P. Pawar, S. Campioni, C.M. Dobson, F. Chiti, and M. Vendruscolo, *Prediction of aggregation-prone regions in structured proteins*. Journal of Molecular Biology, 2008. **380**(2): p. 425-436.
 160. Zambrano, R., M. Jamroz, A. Szczasiuk, J. Pujols, S. Kmiecik, and S. Ventura, *AGGRESKAN3D (A3D): server for prediction of aggregation properties of protein structures*. Nucleic Acids Research, 2015. **43**(W1): p. W306-W313.
 161. Kumar, M.D.S., K.A. Bava, M.M. Gromiha, P. Prabakaran, K. Kitajima, H. Uedaira, and A. Sarai, *ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions*. Nucleic Acids Research, 2006. **34**: p. D204-D206.
 162. Bava, K.A., M.M. Gromiha, H. Uedaira, K. Kitajima, and A. Sarai, *ProTherm, version 4.0: thermodynamic database for proteins and mutants*. Nucleic Acids Research, 2004. **32**: p. D120-D121.
 163. Capriotti, E., P. Fariselli, and R. Casadio, *I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure*. Nucleic Acids Research, 2005. **33**: p. W306-W310.
 164. Masso, M. and I.I. Vaisman, *Accurate prediction of stability changes in protein mutants by combining machine learning with structure based computational mutagenesis*. Bioinformatics, 2008. **24**(18): p. 2002-2009.
 165. Cheng, J.L., A. Randall, and P. Baldi, *Prediction of protein stability changes for single-site mutations using support vector machines*. Proteins-Structure Function and Bioinformatics, 2006. **62**(4): p. 1125-1132.
 166. Dehouck, Y., A. Grosfils, B. Folch, D. Gilis, P. Bogaerts, and M. Rومان, *Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0*. Bioinformatics, 2009. **25**(19): p. 2537-2543.

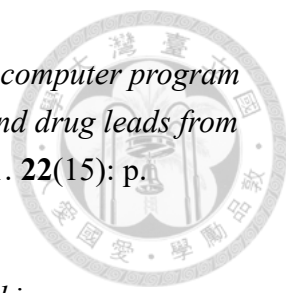
- 
167. Parthiban, V., M.M. Gromiha, and D. Schomburg, *CUPSAT: prediction of protein stability upon point mutations*. Nucleic Acids Research, 2006. **34**: p. W239-W242.
 168. Topham, C.M., N. Srinivasan, and T.L. Blundell, *Prediction of the stability of protein mutants based on structural environment-dependent amino acid substitution and propensity tables*. Protein Engineering, 1997. **10**(1): p. 7-21.
 169. Worth, C.L., R. Preissner, and T.L. Blundell, *SDM-a server for predicting effects of mutations on protein stability and malfunction*. Nucleic Acids Research, 2011. **39**: p. W215-W222.
 170. Pires, D.E.V., D.B. Ascher, and T.L. Blundell, *mCSM: predicting the effects of mutations in proteins using graph-based signatures*. Bioinformatics, 2014. **30**(3): p. 335-342.
 171. da Silva, C.H., D.E.V. Pires, R.C. Minardi, C. Ribeiro, C.J.M. Veloso, J.C.D. Lopes, W. Meira, G. Neshich, C.H.I. Ramos, R. Habesch, and M.M. Santoro, *Protein cutoff scanning: A comparative analysis of cutoff dependent and cutoff free methods for prospecting contacts in proteins*. Proteins-Structure Function and Bioinformatics, 2009. **74**(3): p. 727-743.
 172. Yin, S., L. Biedermannova, J. Vondrasek, and N.V. Dokholyan, *MedusaScore: An accurate force field-based scoring function for virtual drug screening*. Journal of Chemical Information and Modeling, 2008. **48**(8): p. 1656-1662.
 173. Saraboji, K., M.M. Gromiha, and M.N. Ponnuswamy, *Average assignment method for predicting the stability of protein mutants*. Biopolymers, 2006. **82**(1): p. 80-92.
 174. Bordner, A.J. and R.A. Abagyan, *Large-scale prediction of protein geometry and stability changes for arbitrary single point mutations*. Proteins-Structure Function and Bioinformatics, 2004. **57**(2): p. 400-413.
 175. Trovato, A., F. Chiti, A. Maritan, and F. Seno, *Insight into the structure of amyloid fibrils from the analysis of globular proteins*. Plos Computational Biology, 2006. **2**(12): p. 1608-1618.
 176. Kabsch, W. and C. Sander, *Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features*. Biopolymers, 1983. **22**(12): p. 2577-2637.
 177. Sirocco, F. and S.C.E. Tosatto, *TESE: generating specific protein structure test set ensembles*. Bioinformatics, 2008. **24**(22): p. 2632-2633.

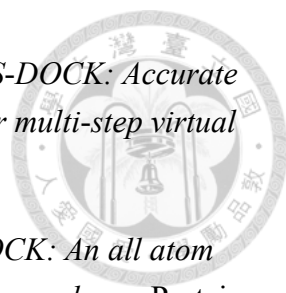
- 
178. Irving, J.H. and J.G. Kirkwood, *The Statistical Mechanical Theory of Transport Processes .4. The Equations of Hydrodynamics*. Journal of Chemical Physics, 1950. **18**(6): p. 817-829.
179. Karplus, M. and J.A. McCammon, *Molecular dynamics simulations of biomolecules*. Nature Structural Biology, 2002. **9**(9): p. 646-652.
180. Zhang, J., W.F. Li, J. Wang, M. Qin, L. Wu, Z.Q. Yan, W.X. Xu, G.H. Zuo, and W. Wang, *Protein Folding Simulations: From Coarse-Grained Model to All-Atom Model*. Iubmb Life, 2009. **61**(6): p. 627-643.
181. Lindorff-Larsen, K., S. Piana, R.O. Dror, and D.E. Shaw, *How Fast-Folding Proteins Fold*. Science, 2011. **334**(6055): p. 517-520.
182. Sun, X.T., Z.W. Feng, T.J. Hou, and Y.Y. Li, *Mechanism of Graphene Oxide as an Enzyme Inhibitor from Molecular Dynamics Simulations*. Acs Applied Materials & Interfaces, 2014. **6**(10): p. 7153-7163.
183. Cole, D.J., J. Tirado-Rives, and W.L. Jorgensen, *Molecular dynamics and Monte Carlo simulations for protein-ligand binding and inhibitor design*. Biochimica Et Biophysica Acta-General Subjects, 2015. **1850**(5): p. 966-971.
184. Ishitobi, M., *Introduction to molecular dynamics simulation*. Journal of Japanese Society of Tribologists, 2000. **45**(9): p. 649-654.
185. Novoselov, K.S., A.K. Geim, S.V. Morozov, D. Jiang, M.I. Katsnelson, I.V. Grigorieva, S.V. Dubonos, and A.A. Firsov, *Two-dimensional gas of massless Dirac fermions in graphene*. Nature, 2005. **438**(7065): p. 197-200.
186. Liu, H.Y., M. Elstner, E. Kaxiras, T. Frauenheim, J. Hermans, and W.T. Yang, *Quantum mechanics simulation of protein dynamics on long timescale*. Proteins-Structure Function and Genetics, 2001. **44**(4): p. 484-489.
187. Kowalewski, M. and S. Mukamel, *Stimulated Raman signals at conical intersections: Ab initio surface hopping simulation protocol with direct propagation of the nuclear wave function*. Journal of Chemical Physics, 2015. **143**(4).
188. Zhao, L.B., D.H. Xiao, and I.I. Fabrikant, *Simulation of wave-function microscopy images of Stark resonances*. Physical Review A, 2015. **91**(4).

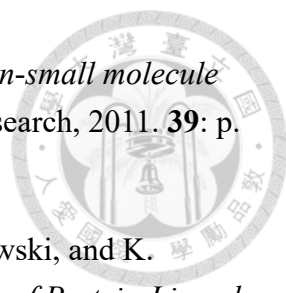
- 
189. Karplus, M., *Development of Multiscale Models for Complex Chemical Systems: From H+H-2 to Biomolecules (Nobel Lecture)*. Angewandte Chemie-International Edition, 2014. **53**(38): p. 9992-10005.
190. Wang, J.M., R.M. Wolf, J.W. Caldwell, P.A. Kollman, and D.A. Case, *Development and testing of a general amber force field*. Journal of Computational Chemistry, 2004. **25**(9): p. 1157-1174.
191. Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A.D. MacKerell, *CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields*. Journal of Computational Chemistry, 2010. **31**(4): p. 671-690.
192. Ritschl, F., M. Fait, K. Fiedler, J.E.H. Kohler, B. Kubias, and M. Meisel, *An extension of the consistent valence force field (CVFF) with the aim to simulate the structures of vanadium phosphorus oxides and the adsorption of n-butane and of 1-butene on their crystal planes*. Zeitschrift Fur Anorganische Und Allgemeine Chemie, 2002. **628**(6): p. 1385-1396.
193. Gaedt, K. and H.D. Holtje, *Consistent valence force-field parameterization of bond lengths and angles with quantum chemical ab initio methods applied to some heterocyclic dopamine D-3-receptor agonists*. Journal of Computational Chemistry, 1998. **19**(8): p. 935-946.
194. Witter, R., L. Seyfart, G. Greiner, S. Reissmann, J. Weston, E. Anders, and U. Sternberg, *Structure determination of a pseudotripeptide zinc complex with the COSMOS-NMR force field and DFT methods*. Journal of Biomolecular Nmr, 2002. **24**(4): p. 277-289.
195. Christen, M., P.H. Hunenberger, D. Bakowies, R. Baron, R. Burgi, D.P. Geerke, T.N. Heinz, M.A. Kastenholtz, V. Krautler, C. Oostenbrink, C. Peter, D. Trzesniak, and W.F. Van Gunsteren, *The GROMOS software for biomolecular simulation: GROMOS05*. Journal of Computational Chemistry, 2005. **26**(16): p. 1719-1751.
196. Wang, J., C.H. Tan, E. Chanco, and R. Luo, *Quantitative analysis of Poisson-Boltzmann implicit solvent in molecular dynamics*. Physical Chemistry Chemical Physics, 2010. **12**(5): p. 1194-1202.

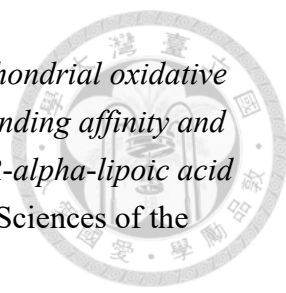
- 
197. Van der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J.C. Berendsen, *GROMACS: Fast, flexible, and free*. Journal of Computational Chemistry, 2005. **26**(16): p. 1701-1718.
198. Hess, B., H. Bekker, H.J.C. Berendsen, and J.G.E.M. Fraaije, *LINCS: A linear constraint solver for molecular simulations*. Journal of Computational Chemistry, 1997. **18**(12): p. 1463-1472.
199. Essmann, U., L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen, *A Smooth Particle Mesh Ewald Method*. Journal of Chemical Physics, 1995. **103**(19): p. 8577-8593.
200. Le Roux, S. and V. Petkov, *ISAACS - interactive structure analysis of amorphous and crystalline systems*. Journal of Applied Crystallography, 2010. **43**: p. 181-185.
201. Lengauer, T. and M. Rarey, *Computational methods for biomolecular docking*. Current Opinion in Structural Biology, 1996. **6**(3): p. 402-406.
202. Jain, A.N., *Scoring functions for protein-ligand docking*. Current Protein & Peptide Science, 2006. **7**(5): p. 407-420.
203. Lensink, M.F., R. Mendez, and S.J. Wodak, *Docking and scoring protein complexes: CAPRI 3rd edition*. Proteins-Structure Function and Bioinformatics, 2007. **69**(4): p. 704-718.
204. Sircar, A. and J.J. Gray, *Antibody-antigen docking with simultaneous CDR-H3 modeling*. Biophysical Journal, 2007: p. 318a-318a.
205. Jia, J.Q., Q.Y. Wu, H. Yan, and Z.Z. Gui, *Purification and molecular docking study of a novel angiotensin-I converting enzyme (ACE) inhibitory peptide from alcalase hydrolysate of ultrasonic-pretreated silkworm pupa (Bombyx mori) protein*. Process Biochemistry, 2015. **50**(5): p. 876-883.
206. DeMarco, M.L. and V. Daggett, *From conversion to aggregation: Protofibril formation of the prion protein*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(8): p. 2293-2298.
207. Roberts, V.A., M.E. Pique, L.F. Ten Eyck, and S. Li, *Predicting protein-DNA interactions by full search computational docking*. Proteins-Structure Function and Bioinformatics, 2013. **81**(12): p. 2106-2118.

- 
208. Goldman, B.B. and W.T. Wipke, *QSD quadratic shape descriptors. 2. Molecular docking using quadratic shape descriptors (QSDock)*. Proteins-Structure Function and Genetics, 2000. **38**(1): p. 79-94.
209. Feig, M., A. Onufriev, M.S. Lee, W. Im, D.A. Case, and C.L. Brooks, *Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures*. Journal of Computational Chemistry, 2004. **25**(2): p. 265-284.
210. Singh, T., D. Biswas, and B. Jayaram, *AADS - An Automated Active Site Identification, Docking, and Scoring Protocol for Protein Targets Based on Physicochemical Descriptors*. Journal of Chemical Information and Modeling, 2011. **51**(10): p. 2515-2527.
211. Goodsell, D.S., G.M. Morris, and A.J. Olson, *Automated docking of flexible ligands: Applications of AutoDock*. Journal of Molecular Recognition, 1996. **9**(1): p. 1-5.
212. Trott, O. and A.J. Olson, *Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading*. Journal of Computational Chemistry, 2010. **31**(2): p. 455-461.
213. Kim, D.S., C.M. Kim, C.I. Won, J.K. Kim, J. Ryu, Y. Cho, C. Lee, and J. Bhak, *BetaDock: Shape-Priority Docking Method Based on Beta-Complex*. Journal of Biomolecular Structure & Dynamics, 2011. **29**(1): p. 219-242.
214. Hottmann, K., *Detmold, 7 to 9. November 2008: "rock blaster, bridge builders, pioneer: the English composer Ethel Smyth (1858-1944)" by Katharina Hottmann, Hamburg*. Musikforschung, 2009. **62**(2): p. 157-158.
215. Lee, H.S. and Y. Zhang, *BSP-SLIM: A blind low-resolution ligand-protein docking approach using predicted protein structures*. Proteins-Structure Function and Bioinformatics, 2012. **80**(1): p. 93-110.
216. Taylor, J.S. and R.M. Burnett, *DARWIN: A program for docking flexible molecules*. Proteins-Structure Function and Genetics, 2000. **41**(2): p. 173-191.
217. Grosdidier, A., V. Zoete, and O. Michielin, *EADock: Docking of small molecules into protein active sites with a multiobjective evolutionary optimization*. Proteins-Structure Function and Bioinformatics, 2007. **67**(4): p. 1010-1025.

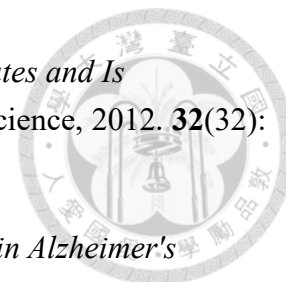
- 
218. Pang, Y.P., E. Perola, K. Xu, and F.G. Prendergast, *EUDOC: A computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases*. Journal of Computational Chemistry, 2001. **22**(15): p. 1750-1771.
219. Gaudreault, F. and R.J. Najmanovich, *FlexAID: Revisiting Docking on Non-Native-Complex Structures*. Journal of Chemical Information and Modeling, 2015. **55**(7): p. 1323-1336.
220. Zhao, Y. and M.F. Sanner, *FLIPDock: Docking flexible ligands into flexible receptors*. Abstracts of Papers of the American Chemical Society, 2007. **233**: p. 152-152.
221. Yang, J.M. and C.C. Chen, *GEMDOCK: A generic evolutionary method for molecular docking*. Proteins-Structure Function and Bioinformatics, 2004. **55**(2): p. 288-304.
222. Dominguez, C., R. Boelens, and A.M.J.J. Bonvin, *HADDOCK: A protein-protein docking approach based on biochemical or biophysical information*. Journal of the American Chemical Society, 2003. **125**(7): p. 1731-1737.
223. Venkatachalam, C.M., X. Jiang, T. Oldfield, and M. Waldman, *LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites*. Journal of Molecular Graphics & Modelling, 2003. **21**(4): p. 289-307.
224. Shin, W.H., L. Heo, J. Lee, J. Ko, C. Seok, and J. Lee, *LigDockCSA: Protein-Ligand Docking Using Conformational Space Annealing*. Journal of Computational Chemistry, 2011. **32**(15): p. 3226-3232.
225. Liu, M. and S.M. Wang, *MCDOCK: A Monte Carlo simulation approach to the molecular docking problem*. Journal of Computer-Aided Molecular Design, 1999. **13**(5): p. 435-451.
226. Chang, D.T.H., Y.J. Oyang, and J.H. Lin, *MEDock: a web server for efficient prediction of ligand binding sites based on a novel optimization algorithm*. Nucleic Acids Research, 2005. **33**: p. W233-W238.
227. Thomsen, R. and M.H. Christensen, *MolDock: A new technique for high-accuracy molecular docking*. Journal of Medicinal Chemistry, 2006. **49**(11): p. 3315-3321.

- 
228. Sauton, N., D. Lagorce, B.O. Villoutreix, and M.A. Miteva, *MS-DOCK: Accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening*. BMC Bioinformatics, 2008. **9**.
229. Gupta, A., A. Gandhimathi, P. Sharma, and B. Jayaram, *ParDOCK: An all atom energy based Monte Carlo docking protocol for protein-ligand complexes*. Protein and Peptide Letters, 2007. **14**(7): p. 632-646.
230. Schneidman-Duhovny, D., Y. Inbar, R. Nussinov, and H.J. Wolfson, *PatchDock and SymmDock: servers for rigid and symmetric docking*. Nucleic Acids Research, 2005. **33**: p. W363-W367.
231. Trosset, J.Y. and H.A. Scheraga, *PRODOCK: Software package for protein modeling and docking*. Journal of Computational Chemistry, 1999. **20**(4): p. 412-427.
232. Pei, J.F., Q. Wang, Z.M. Liu, Q.L. Li, K. Yang, and L.H. Lai, *PSI-DOCK: Towards highly efficient and accurate flexible ligand docking*. Proteins-Structure Function and Bioinformatics, 2006. **62**(4): p. 934-946.
233. Chung, J.Y., S.J. Cho, and J.M. Hah, *A Python-based Docking Program Utilizing a Receptor Bound Ligand Shape: PythDock*. Archives of Pharmacal Research, 2011. **34**(9): p. 1451-1458.
234. Brylinski, M. and J. Skolnick, *Q-Dock: Low-resolution flexible ligand docking with pocket-specific threading restraints*. Journal of Computational Chemistry, 2008. **29**(10): p. 1574-1588.
235. Ruiz-Carmona, S., D. Alvarez-Garcia, N. Foloppe, A.B. Garmendia-Doval, S. Juhos, P. Schmidtke, X. Barril, R.E. Hubbard, and S.D. Morley, *rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids*. Plos Computational Biology, 2014. **10**(4).
236. Chen, H.M., B.F. Liu, H.L. Huang, S.F. Hwang, and S.Y. Ho, *SODOCK: Swarm optimization for highly flexible protein-ligand docking*. Journal of Computational Chemistry, 2007. **28**(2): p. 612-623.
237. Jain, A.N., *Surflex-Dock 2.1: Robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search*. Journal of Computer-Aided Molecular Design, 2007. **21**(5): p. 281-306.

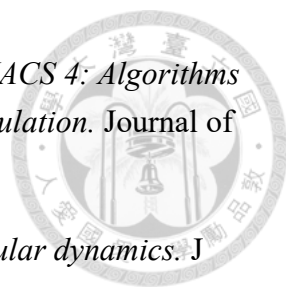
- 
238. Grosdidier, A., V. Zoete, and O. Michielin, *SwissDock, a protein-small molecule docking web service based on EADock DSS*. Nucleic Acids Research, 2011. **39**: p. W270-W277.
239. Plewczynski, D., M. Lazniewski, M. Von Grotthuss, L. Rychlewski, and K. Ginalski, *VoteDock: Consensus Docking Method for Prediction of Protein-Ligand Interactions*. Journal of Computational Chemistry, 2011. **32**(4): p. 568-581.
240. Wu, G.S., D.H. Robertson, C.L. Brooks, and M. Vieth, *Detailed analysis of grid-based molecular docking: A case study of CDOCKER - A CHARMM-based MD docking algorithm*. Journal of Computational Chemistry, 2003. **24**(13): p. 1549-1562.
241. Ramana, B.V., T.N. Raju, V.V. Kumar, and P.U.M. Reddy, *Defensive role of quercetin against imbalances of calcium, sodium, and potassium in galactosemic cataract*. Biological Trace Element Research, 2007. **119**(1): p. 35-41.
242. Cornish, K.M., G. Williamson, and J. Sanderson, *Quercetin metabolism in the lens: Role in inhibition of hydrogen peroxide induced cataract*. Free Radical Biology and Medicine, 2002. **33**(1): p. 63-70.
243. Kim, Y.S., N.H. Kim, Y.M. Lee, and J.S. Kim, *Preventive Effect of Chlorogenic Acid on Lens Opacity and Cytotoxicity in Human Lens Epithelial Cells*. Biological & Pharmaceutical Bulletin, 2011. **34**(6): p. 925-928.
244. Kim, C.S., J. Kim, Y.M. Lee, E. Sohn, K. Jo, and J.S. Kim, *Inhibitory Effects of Chlorogenic Acid on Aldose Reductase Activity In Vitro and Cataractogenesis in Galactose-fed Rats*. Archives of Pharmacal Research, 2011. **34**(5): p. 847-852.
245. Rathi, S.S., J.K. Grover, V. Vikrant, and N.R. Biswas, *Prevention of experimental diabetic cataract by Indian Ayurvedic plant extracts*. Phytotherapy Research, 2002. **16**(8): p. 774-777.
246. Necula, M., R. Kaye, S. Milton, and C.G. Glabe, *Small molecule inhibitors of aggregation indicate that amyloid beta oligomerization and fibrillization pathways are independent and distinct*. Journal of Biological Chemistry, 2007. **282**(14): p. 10311-10324.
247. Emerit, J., A. Edeas, and F. Bricaire, *Neurodegenerative diseases and oxidative stress*. Biomedicine & Pharmacotherapy, 2004. **58**(1): p. 39-46.

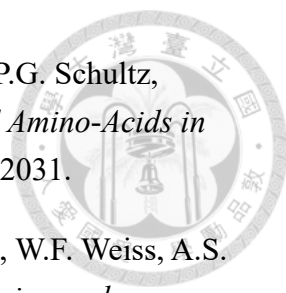
- 
248. Liu, J.K., D.W. Killilea, and B.N. Ames, *Age-associated mitochondrial oxidative decay: Improvement of carnitine acetyltransferase substrate-binding affinity and activity in brain by feeding old rats acetyl-L-carnitine and/or R-alpha-lipoic acid (vol 99, 1876, 2002)*. Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(10): p. 7184-7184.
249. Villar-Pique, A., A. Espargaro, R. Sabate, N.S. de Groot, and S. Ventura, *Using bacterial inclusion bodies to screen for amyloid aggregation inhibitors*. Febs Journal, 2012. **279**: p. 447-447.
250. Taniguchi, S., N. Suzuki, M. Masuda, S. Hisanaga, T. Iwatsubo, M. Goedert, and M. Hasegawa, *Inhibition of heparin-induced tau filament formation by phenothiazines, polyphenols, and porphyrins*. Journal of Biological Chemistry, 2005. **280**(9): p. 7614-7623.
251. Duhaime, A.S., *Characterization of zeta-crystallin inhibition by juglone*. Biochemical and Biophysical Research Communications, 1996. **218**(3): p. 648-652.
252. Shiraki, K., M. Kudou, S. Nishikori, H. Kitagawa, T. Imanaka, and M. Takagi, *Arginine ethylester prevents thermal inactivation and aggregation of lysozyme*. European Journal of Biochemistry, 2004. **271**(15): p. 3242-3247.
253. Hamada, H. and K. Shiraki, *L-Argininamide improves the refolding more effectively than L-arginine*. Journal of Biotechnology, 2007. **130**(2): p. 153-160.
254. Park, S.U., M.R. Uddin, H. Xu, Y.K. Kim, and S.Y. Lee, *Biotechnological applications for rosmarinic acid production in plant*. African Journal of Biotechnology, 2008. **7**(25): p. 4959-4965.
255. Ono, K., K. Hasegawa, H. Naiki, and M. Yamada, *Curcumin has potent anti-amyloidogenic effects for Alzheimer's beta-amyloid fibrils in vitro*. Journal of Neuroscience Research, 2004. **75**(6): p. 742-750.
256. Yang, F.S., G.P. Lim, A.N. Begum, O.J. Ubeda, M.R. Simmons, S.S. Ambegaokar, P.P. Chen, R. Kaye, C.G. Glabe, S.A. Frautschy, and G.M. Cole, *Curcumin inhibits formation of amyloid beta oligomers and fibrils, binds plaques, and reduces amyloid in vivo*. Journal of Biological Chemistry, 2005. **280**(7): p. 5892-5901.
257. Sontag, E.M., G.P. Lotz, N. Agrawal, A. Tran, R. Aron, G. Yang, M. Necula, A. Lau, S. Finkbeiner, C. Glabe, J.L. Marsh, P.J. Muchowski, and L.M. Thompson,

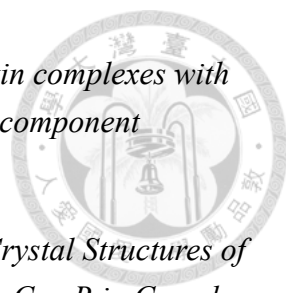
Methylene Blue Modulates Huntingtin Aggregation Intermediates and Is Protective in Huntington's Disease Models. Journal of Neuroscience, 2012. **32**(32): p. 11109-11119.



258. Atamna, H. and R. Kumar, *Protective Role of Methylene Blue in Alzheimer's Disease via Mitochondria and Cytochrome c Oxidase.* Journal of Alzheimers Disease, 2010. **20**: p. S439-S452.
259. Necula, M., L. Breydo, S. Milton, R. Kaye, W.E. van der Veer, P. Tone, and C.G. Glabe, *Methylene blue inhibits amyloid A beta oligomerization by promoting fibrillization.* Biochemistry, 2007. **46**(30): p. 8850-8860.
260. Porat, Y., A. Abramowitz, and E. Gazit, *Inhibition of amyloid fibril formation by polyphenols: Structural similarity and aromatic interactions as a common inhibition mechanism.* Chemical Biology & Drug Design, 2006. **67**(1): p. 27-37.
261. Ferreira, N., M.J. Saraiva, and M.R. Almeida, *Epigallocatechin-3-Gallate as a Potential Therapeutic Drug for TTR-Related Amyloidosis: "In Vivo" Evidence from FAP Mice Models.* Plos One, 2012. **7**(1): p. 232-241.
262. Ehrnhoefer, D.E., M. Duennwald, P. Markovic, J.L. Wacker, S. Engemann, M. Roark, J. Legleiter, J.L. Marsh, L.M. Thompson, S. Lindquist, P.J. Muchowski, and E.E. Wanker, *Green tea (-)-epigallocatechin-gallate modulates early events in huntingtin misfolding and reduces toxicity in Huntington's disease models.* Human Molecular Genetics, 2006. **15**(18): p. 2743-2751.
263. Beckstein, O., E.J. Denning, J.R. Perilla, and T.B. Woolf, *Zippering and Unzippering of Adenylate Kinase: Atomistic Insights into the Ensemble of Open <-> Closed Transitions.* Journal of Molecular Biology, 2009. **394**(1): p. 160-176.
264. Berendsen, H.J.C., D. Vanderspoel, and R. Vandrunen, *Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation.* Computer Physics Communications, 1995. **91**(1-3): p. 43-56.
265. Lindahl, E., B. Hess, and D. Van Der Spoel, *GROMACS 3.0: a package for molecular simulation and trajectory analysis.* Journal of Molecular Modeling, 2001. **7**(8): p. 306-317.
266. Van Der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J. Berendsen, *GROMACS: fast, flexible, and free.* J Comput Chem, 2005. **26**(16): p. 1701-18.

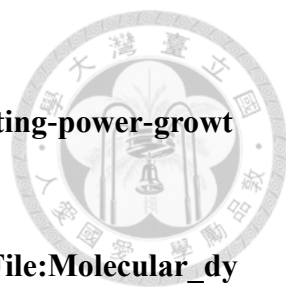
- 
267. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl, *GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation*. Journal of Chemical Theory and Computation, 2008. **4**(3): p. 435-447.
268. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics*. J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
269. Michaud-Agrawal, N., E.J. Denning, T.B. Woolf, and O. Beckstein, *MDAnalysis: A toolkit for the analysis of molecular dynamics simulations*. J Comput Chem, 2011.
270. Chandrasekhar, I., M. Kastenholtz, R.D. Lins, C. Oostenbrink, L.D. Schuler, D.P. Tieleman, and W.F. van Gunsteren, *A consistent potential energy parameter set for lipids: dipalmitoylphosphatidylcholine as a benchmark of the GROMOS96 45A3 force field*. European Biophysics Journal with Biophysics Letters, 2003. **32**(1): p. 67-77.
271. Michaud-Agrawal, N., E.J. Denning, T.B. Woolf, and O. Beckstein, *Software News and Updates MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations*. Journal of Computational Chemistry, 2011. **32**(10): p. 2319-2327.
272. Schuttelkopf, A.W. and D.M.F. van Aalten, *PRODRG: a tool for high-throughput crystallography of protein-ligand complexes*. Acta Crystallographica Section D-Biological Crystallography, 2004. **60**: p. 1355-1363.
273. Stadtman, T.C., *Selenocysteine*. Annual Review of Biochemistry, 1996. **65**: p. 83-100.
274. Zhang, Y., P.V. Baranov, J.F. Atkins, and V.N. Gladyshev, *Pyrrolysine and selenocysteine use dissimilar decoding strategies*. Journal of Biological Chemistry, 2005. **280**(21): p. 20740-20751.
275. Holley, L.H. and M. Karplus, *Protein Secondary Structure Prediction with a Neural Network*. Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(1): p. 152-156.
276. Eddy, S.R., *Where did the BLOSUM62 alignment score matrix come from?* Nature Biotechnology, 2004. **22**(8): p. 1035-1036.
277. Hartl, F.U., A. Bracher, and M. Hayer-Hartl, *Molecular chaperones in protein folding and proteostasis*. Nature, 2011. **475**(7356): p. 324-332.

- 
278. Cornish, V.W., M.I. Kaplan, D.L. Veenstra, P.A. Kollman, and P.G. Schultz, *Stabilizing and Destabilizing Effects of Placing Beta-Branched Amino-Acids in Protein Alpha-Helices*. *Biochemistry*, 1994. **33**(40): p. 12022-12031.
279. Sahin, E., J.L. Jordan, M.L. Spatara, A. Naranjo, J.A. Costanzo, W.F. Weiss, A.S. Robinson, E.J. Fernandez, and C.J. Roberts, *Computational Design and Biophysical Characterization of Aggregation-Resistant Point Mutations for gamma D Crystallin Illustrate a Balance of Conformational Stability and Intrinsic Aggregation Propensity*. *Biochemistry*, 2011. **50**(5): p. 628-639.
280. Rajan, R.S., M.E. Illing, N.F. Bence, and R.R. Kopito, *Specificity in intracellular protein aggregation and inclusion body formation*. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. **98**(23): p. 13060-13065.
281. Morrison, K.L. and G.A. Weiss, *Combinatorial alanine-scanning*. *Current Opinion in Chemical Biology*, 2001. **5**(3): p. 302-307.
282. Yokoyama, T., Y. Kosaka, and M. Mizuguchi, *Structural Insight into the Interactions between Death-Associated Protein Kinase 1 and Natural Flavonoids*. *Journal of Medicinal Chemistry*, 2015. **58**(18): p. 7400-7408.
283. Kong, Y.H., L. Zhang, Z.Y. Yang, C. Han, L.H. Hu, H.L. Jiang, and X. Shen, *Natural product juglone targets three key enzymes from Helicobacter pylori: inhibition assay with crystal structure characterization*. *Acta Pharmacologica Sinica*, 2008. **29**(7): p. 870-876.
284. Williams, L.K., C.M. Li, S.G. Withers, and G.D. Brayer, *Order and Disorder: Differential Structural Impacts of Myricetin and Ethyl Caffeate on Human Amylase, an Antidiabetic Target*. *Journal of Medicinal Chemistry*, 2012. **55**(22): p. 10177-10186.
285. Yokoyama, T., Y. Kosaka, and M. Mizuguchi, *Inhibitory Activities of Propolis and Its Promising Component, Caffeic Acid Phenethyl Ester, against Amyloidogenesis of Human Transthyretin*. *Journal of Medicinal Chemistry*, 2014. **57**(21): p. 8928-8935.
286. Florio, P., C. Folli, M. Cianci, D. Del Rio, G. Zanotti, and R. Berni, *Transthyretin Binding Heterogeneity and Anti-amyloidogenic Activity of Natural Polyphenols and Their Metabolites*. *Journal of Biological Chemistry*, 2015. **290**(50): p. 29769-29780.

- 
287. Ciccone, L., L. Tepshi, S. Nencetti, and E.A. Stura, *Transthyretin complexes with curcumin and bromo-estradiol: evaluation of solubilizing multicomponent mixtures*. New Biotechnology, 2015. **32**(1): p. 54-64.
288. Itou, H., N. Watanabe, M. Yao, Y. Shirakihara, and I. Tanaka, *Crystal Structures of the Multidrug Binding Repressor Corynebacterium glutamicum CgmR in Complex with Inducers and with an Operator*. Journal of Molecular Biology, 2010. **403**(2): p. 174-184.
289. Urusova, D.V., J.H. Shim, D.J. Kim, S.K. Jung, T.A. Zykova, A. Carper, A.M. Bode, and Z.G. Dong, *Epigallocatechin-gallate Suppresses Tumorigenesis by Directly Targeting Pin1*. Cancer Prevention Research, 2011. **4**(9): p. 1366-1377.
290. Tina, K.G., R. Bhadra, and N. Srinivasan, *PIC: Protein Interactions Calculator*. Nucleic Acids Research, 2007. **35**: p. W473-W476.
291. Overington, J., M.S. Johnson, A. Sali, and T.L. Blundell, *Tertiary Structural Constraints on Protein Evolutionary Diversity - Templates, Key Residues and Structure Prediction*. Proceedings of the Royal Society B-Biological Sciences, 1990. **241**(1301): p. 132-145.
292. Burley S.K., P.G., *Aromatic-aromatic interaction: a mechanism of protein structure stabilization*. Science and Technology of Nuclear Installations, 1985. **229**: p. 23-29.
293. Reid, K.S.C., P.F. Lindley, and J.M. Thornton, *Sulfur-Aromatic Interactions in Proteins*. Febs Letters, 1985. **190**(2): p. 209-213.
294. Sathyapriya, R. and S. Vishveshwara, *Interaction of DNA with clusters of amino acids in proteins*. Nucleic Acids Research, 2004. **32**(14): p. 4109-4118.

網路資料網站

- 圖 2.1.1 正常水晶體(左)與混濁水晶體(右)折射入射光之示意圖
<http://www.improveeyesighthq.com/cataracts.html>
- 圖 2.1.4 白內障手術更換人工水晶體之示意圖
<http://www.ncascade.com/services/cataract-surgery/>

- 
- 圖 2.3.1 電腦運算能力進展的趨勢圖
<https://www.quora.com/Where-can-I-see-a-graph-of-computing-power-growth-for-the-next-38-years-2050>
 - 圖 2.3.2 分子動力學模擬演算流程圖
https://en.wikipedia.org/wiki/Molecular_dynamics#/media/File:Molecular_dynamics_algorithm.png
 - 圖 2.3.3 量子/分子模擬示意圖 <http://www.ch.ntu.edu.tw/nobel/2013.html>
 - 圖 2.3.4 Gromacs 模擬流程圖
<http://2013.igem.org/Team:Manchester/FabProteinModel>
 - 圖 2.3.5 Gromacs 模擬流程示意圖(A)建立 $2*2*2\text{ nm}^3$ 的盒子(B)在盒中加入溶劑模型(C)在盒中加入離子(藍:鈉離子,綠:氯離子)
<http://2013.igem.org/Team:Manchester/FabProteinModel>
 - 圖 2.3.7 分子對接模型示意圖
[https://en.wikipedia.org/wiki/Docking_\(molecular\)#/media/File:Docking_representation_2.png](https://en.wikipedia.org/wiki/Docking_(molecular)#/media/File:Docking_representation_2.png)
 - 圖 4.2.15 胺基酸與對應之基因密碼子(codon)一覽
<http://openwetware.org/wiki/Image:Codontable2.gif>