

使用雙目標最佳化方法估算肌肉醣原分解之
代謝路徑的動力學參數

Estimation of kinetic parameters for modeling
metabolic pathways of muscle glycogenolysis
using two-objective optimization methods

研究生：張孝邦

指導教授：何信瑩

Student : Shiao-Bang Chang

Advisor : Shinn-Ying Ho

國立交通大學

生物資訊研究所



A Thesis Submitted to Institute of Bioinformatics
College of Biological Science and Technology
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of Master in
Bioinformatics

January 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年一月

使用雙目標最佳化方法估算肌肉醣原分解之 代謝路徑的動力學參數

學生：張孝邦

指導教授：何信瑩

國立交通大學生物資訊研究所碩士班

摘要

代謝路徑的模型重建是研究系統生物學的主要課題之一，建模的目的是為了研究蛋白質或整個代謝網路彼此之間的相互作用以及行為模式。例如肌肉之醣原分解是人體生理機能中不可或缺的一環，並且其代謝路徑的生化反應相當敏感。本研究的模型架構是以微分方程式為主軸的動力學模型，此模型能在反應過程中呈現出瞬間的濃度變化、酶所造成之通量及受酶影響之回應。

通常要重建一個動力學模型，例如肌肉醣原分解之代謝路徑模型，可使用被收集在特定資料庫中的數據，但可能有些參數值尚屬未知。這些動力學參數值通常是利用實驗方法去獲得，它們可能是從不同的實驗環境或是不同物種組成分批得來。由於代謝過程之中的動力學參數之間存在著熱力學限制以及反應代謝物的 Gibbs 生成自由能會決定反應平衡常數，使用這類參數值所建成的模型可能未遵守熱力學原理，導致不真實的生化反應。

本研究提出一套使用雙目標最佳化方法來估算代謝路徑模型的動力學參數並將之應用到肌肉醣原分解之代謝路徑的模型重建，其特色有(1)同時最小化系統總自由能來符合熱力學原理及最小化模型所估算反應物的濃度誤差及(2)使用智慧型基因演算法精確地解此雙目標最佳化問題的大量動力學參數。

從模擬實驗結果顯示最小化系統總自由能和最小化濃度誤差兩目標在提高模型精確度時有所衝突，使用雙目標最佳化方法有其必要。在模型精確度方面，我們以所提方法的估算參數值及現有方法的預設

參數值所建的模型做比較，結果顯示我們的方法能獲得較低的系統總自由能(-0.72566kJ vs. -0.06561kJ)和較小的濃度誤差(提升準確率9.20%)。對所建模型做反應物與產物的敏感度分析，可進一步了解整個代謝網路彼此之間的相互作用。本文所提方法亦適用於其它相似代謝路徑的模型重建。



Estimation of kinetic parameters for modeling metabolic pathways of muscle glycogenolysis using two-objective optimization methods

Student : Shiao-Bang Chang

Advisor : Dr. Shinn-Ying Ho

Institute of Bioinformatics
National Chiao Tung University

Abstract

The model reconstruction of metabolic pathways is one of the major researches of systems biology, whose goal is to study the interaction and behaviors among proteins or whole metabolic networks. Muscle glycogenolysis is essential to the human physiological functions and its biochemical reactions of metabolic pathways are sensitive. In this study, the structure of kinetic models bases on the principle of ordinary differential equations, which can express the temporal changing, fluxes and responses influenced by enzymes in the progress of reactions.

To construct a kinetic model, such as the metabolic pathways of muscle glycogenolysis, one can retrieve the values of kinetic parameters from specific databases. However, the parameter values for establishing the kinetic models were generally obtained from experimental methods, which may be derived separately from different experiments or combination of different species. Therefore, the constructed model may violate thermodynamics theory or be unknown, which results in unrealistic biochemical reactions.

This study proposes a two-objective optimization approach to estimation of kinetic parameters for modeling metabolic pathways and its application to muscle glycogenolysis. The merits of this approach is twofold: 1) simultaneously minimizing the total Gibbs free energy of the system according to thermodynamic theory and minimizing the estimated concentration errors; and 2) using intelligent genetic algorithms to solve accurately the two-objective optimization problem with a large number of ki-

netic parameters.

The simulation results reveal that the two objectives have conflicts in pursuit of accurate models. It means that the simultaneous optimization of the two objectives is necessary. In the aspect of model accuracy, the two models using the estimated and existing kinetic parameter values were compared. The results show that the proposed model has lower total Gibbs free energy of the system (-0.72566kJ vs. -0.06561kJ) and a smaller estimated concentration errors (improvement 9.20%). The sensitivity analysis on the reconstructed model can further understand the interactions of the metabolic networks. The proposed method is also effective to the model reconstruction of other similar metabolic pathways.



誌謝

此篇論文的研究完成需要特別感謝指導教授—何信瑩老師的巧思，當我在研究瓶頸中的循循善誘，讓我能逐步完成研究以及學習如何以邏輯思考地分析研究方向。另外還能從他的身教及言教中了解作人作事的道理所在，而這些都要感謝老師的不吝指導。

此研究的完成還須感謝清大化工系的汪上曉汪教授在化工熱力學和反應動力學上的指教以及本所尤禎祥老師於計算化學上的指教，感謝此兩位教授的幫忙，使得研究能逐步完成。

另外還要感謝實驗室義雄學長在我碩士生涯中於公、於私的教訓及關心，以及實驗室交接的各項幫忙。還有冠維、鐘錡與俊維學長在資訊能力上的教導及幫助，而佳達，凱迪及廖芹於提供實驗室生活中的趣味，當然還要感謝慧玲學姐和瓊慧學姐的大力鼓勵。感謝實驗室大家的關心以及和樂融融的相處能讓我不斷走下去。另外還要謝謝怡蒨在我身邊的關心與教訓也是能讓我完成研究的一大動力。

最後要謝謝我的家人，因為有他們支持及幫忙使我在求學階段不必擔心生活而終將完成學業上的研究！。



目錄

頁次

摘要	ii
Abstract	iv
誌謝	vi
目錄	vii
表目錄	x
圖目錄	xi
符號說明	xiii
一、 緒論	1
1.1 研究動機	1
1.2 相關研究	1
1.3 研究目的	3
二、 代謝反應路徑之動力學模型的結構	7
2.1 模型理論	7
2.1.1 動力學模型	7
2.1.2 擴展後的 Debye-Hückel 定律	10
2.1.3 熱力學上的限制	11
2.2 系統流程	13
2.2.1 數學架構	13
2.2.2 解題方法	16
2.3 方法應用	16
2.3.1 使用擴展 Debye Hückel 定律	18
2.3.2 動力學函數以及模型	19

三、 最佳化演算法之使用	22
3.1 智慧型基因演算法	22
3.1.1 直交表與因素分析	22
3.1.2 智慧型交配運算	23
3.2 方法流程	24
3.2.1 選擇運算	24
3.2.2 智慧型交配	24
3.2.3 突變運算及演化終止條件	25
3.3 智慧型多目標基因演算法	26
3.3.1 基於 Pareto 理論通適化且不因尺度影響之評估函數	26
3.3.2 演算法流程	28
四、 估算動力學參數之雙目標最佳化方法	29
4.1 估算動力學參數之模組化	29
4.2 評估函數的兩項目標函數	31
4.2.1 以加權指數法混合雙目標	32
4.2.2 以最佳解集合同時考量雙目標	33
4.3 染色體編碼方式	35
五、 在骨骼肌肉醣解代謝上的應用	36
5.1 反應速率方程式	36
5.2 計算範例說明	42
六、 實驗結果	45
6.1 利用智慧型基因演算法作加權指數法	45
6.1.1 實驗只使用 f_G 模擬無雜訊之數據	45
6.1.2 實驗只使用 f_E 模擬無雜訊之數據	46
6.1.3 混合使用 f_G 和 f_E 模擬無雜訊之數據	51
6.1.4 實驗混合使用 f_G 和 f_E 並使用加權指數	52
6.1.5 實驗只使用 f_E 模擬有雜訊之數據	53
6.2 用智慧型多目標基因演算法取得最佳解集合	58
6.2.1 實驗模擬無雜訊之數據	59

6.2.2 實驗模擬有雜訊之數據.....	60
6.2.3 實驗中使用真實實驗數據.....	62
6.3 與真實實驗數據之比較.....	63
6.3.1 驗證 f_G	63
6.3.2 濃度曲線圖.....	65
6.3.3 動力學參數.....	68
6.4 敏感度分析	71
6.5 模擬結果的其餘反應物之濃度.....	73
七、問題討論	80
7.1 討論.....	80
7.2 未來展望.....	81
參考文獻.....	82
附錄	89



表目錄

表 1 肌肉醣原分解中全部反應	17
表 2 反應物簡名及標號	19
表 3 反應物的起始濃度	20
表 4 參與反應的酵素活性	21
表 5 $L_8(2^7)$ 直交表	23
表 6 IGA 全實驗之統計表	58
表 7 動力學參數表	69
表 8 敏感度分析的實驗統計結果	72



圖目錄

圖 1 系統概念圖	5
圖 2 系統總體 Gibbs 自由能對反應程度	11
圖 3 肌肉醣原分解之反應途徑	18
圖 4 支配與被支配關係示意圖	27
圖 5 GPSIFF 之示意說明圖	27
圖 6 EKTOM 之說明及流程圖	30
圖 7 應用智慧型基因演算法求最佳解之流程	33
圖 8 應用智慧型多目標基因演算法求最佳解集合之流程	34
圖 9 染色體的編碼方式	35
圖 10 IGA 針對 f_G 做最佳化之結果	45
圖 11 IGA 針對 f_E 做最佳化之結果	46
圖 12 IGA 演化曲線圖-實驗組 1	47
圖 13 IGA 演化曲線圖-實驗組 2	48
圖 14 IGA 演化曲線圖-實驗組 3	48
圖 15 IGA 演化曲線圖-實驗組 4	49
圖 16 IGA 演化曲線圖-實驗組 5	49
圖 17 IGA 演化曲線圖-實驗組 6	50
圖 18 IGA 演化曲線圖-實驗組 7	50
圖 19 IGA 演化曲線圖-實驗組 8	51
圖 20 結合兩者結果	52
圖 21 混合兩目標函數之實驗結果	53
圖 22 IGA 針對 f_E 及 3% 雜訊	54
圖 23 IGA 具雜訊之演化曲線圖-實驗組 1	55
圖 24 IGA 具雜訊之演化曲線圖-實驗組 2	55
圖 25 IGA 具雜訊之演化曲線圖-實驗組 3	56
圖 26 IGA 具雜訊之演化曲線圖-實驗組 4	56
圖 27 IGA 具雜訊之演化曲線圖-實驗組 5	57
圖 28 雜訊存在與否之實驗示意圖	57
圖 29 IMOGA 對模擬實驗數據之結果	59
圖 30 IMOGA 與 IGA 比較圖	60
圖 31 IMOGA 對具雜訊之模擬數據結果-實驗組 1	61

圖 32 IMOGA 對具雜訊之模擬實驗數據圖-實驗組 2	61
圖 33 當雜訊存在，IGA 與 IMOGA 之比較圖	62
圖 34 IMOGA 對真實實驗數據圖	63
圖 35 三組 K 在模擬反應過程中系統總能量之變化	64
圖 36 ATP 及 ADP 之濃度變化	66
圖 37 LAC 及 PCr 之濃度變化	66
圖 38 HMP 之濃度變化	67
圖 39 G3P 及 PG 之濃度變化	68
圖 40 最終產物之產生速率的敏感度對反應物濃度 變化 1% 的結果	72
圖 41 最終產物之產生速率的敏感度對反應物濃度 變化 3% 的結果	73
圖 42 Cr 的濃度變化	74
圖 43 Pi 的濃度變化	75
圖 44 AMP 及 DHAP 之濃度變化	76
圖 45 G6P 及 FBP 之濃度變化	76
圖 46 GAP 及 1,3BPG 之濃度變化	77
圖 47 NADH 及 G1P 之濃度變化	77
圖 48 2PG 及 PEP 之濃度變化	78
圖 49 3PG 及 PYR 之濃度變化	78
圖 50 NAD 的濃度變化	79



符 號 說 明

$N_H(i)$: 該反應物 i 擁有的氫離子數

pH : 代表環境酸鹼度

Z_i : 該反應物 i 的電子價價數

I : 離子強度

R : 氣體常數

T : 絶對溫度

nV : 所有成分之總體積

P : 系統壓力

nS : 系統總熵

μ_i : 反應物 i 的化學勢

n_i : 反應物 i 的莫耳數

$K_{a,i}$: 反應物 i 的解離常數

K_m : 金屬離子 m 的平衡常數

K_{eq}^m : 反應 m 的反應平衡常數

K_j : 動力學模型中的參數。 j 為參數標號， $j=1 \dots 70$ 。

X_i : 反應物 i 的濃度

$Flux_m$: 反應 m 的反應通量

w_i^m : 在反應 m 中，反應物 i 的化學計量係數

ζ_m : 反應 m 的反應程度

ΔH : 反應焓

\bar{N}_H^i : 反應物 i 的平均氫離子數目

$\Delta_f G_i$: 反應物 i 的莫耳生成自由能

$\Delta_f G_i^o$: 修正後的反應物 i 之莫耳生成自由能

ΔG_m^o : 經轉換後，反應 m 的莫耳反應自由能

G : 系統總體 Gibbs 自由能

dG : 總體 Gibbs 自由能的改變量

$E_{am\ f}$: 第 m 個反應的正向酵素活性

$E_{am\ r}$: 第 m 個反應的逆向酵素活性

$E_{am\ app}$: 第 m 個反應，經修正後的酵素活性

K : 動力學參數組

一、緒論

1.1 研究動機

在系統生物學之中，要用電腦模擬細胞行為的研究是一項相當重要的課題[1]。而其中關於生物化學網路的研究裡有三個主要架構，分別是基因調節網路(genetic regulatory network)、生化網路(biochemical or metabolic network)和訊號傳遞網路 (signal transduction network)。這三個網路彼此交互作用而形成生物體內的運作[2-6]，經由一連串的刺激、反應跟訊息傳遞造成生物內部的行為。生化網路所進行的部分是細胞內的化學物質之間化學變化，而這個網路的結構是來自於數個連續的化學反應而串接起來成為一個有功能性的反應路徑，經由反應消耗在反應中的代謝物(metabolite)，生成下一個反應的受質(substrate)後接著繼續下一個反應直到整條反應路徑達到反應平衡而中止。代謝網路專注研究在反應之間彼此的行為，例如濃度、通量、激活或是抑制反應。系統生物學需要利用這些行為對反應途徑建構模型，藉此探討整條反應網路之行為變化所造成的影响。

為了能夠研究反應過程中的影響及變化，我們需要一套能作動力學分析的模型。動力學模型通常是使用微分方程組的數學架構，此架構存在著許多參數，參數數目與微分方程之大小成正相關。

本研究為提供一個能同時根據實驗數據以及考慮熱力學條件之下，將動力學模型之參數作系統最佳化設計，為了能更精確地計算參數值並正確地還原代謝反應之系統。這篇研究以肌肉醣解代謝反應為例，其共有 16 個微分方程式，含有 70 個動力學參數。

1.2 相關研究

目前已可從過去研究區分出代謝之網路結構[7-9]，為了能更深入生化反應的網路，之前大多嘗試使用 Flux Balance Analysis (FBA) [10-12]、Biochemical Systems Theory (BST) [13, 14]、Metabolic Control Analysis (MCA)[15]以上方法建構出一個能表示出反應流程的模型，並且藉由路徑或基本限制的方法[16-19]以應用反應通量的方式[10, 20，

21] 成功地描述其模型。

Flux Balance Analysis[11, 12]是根據研究系統大小而改變其化學計量矩陣並利用 Linear programming 找出最佳解。因為 FBA 能對應到大型代謝網路系統上，目前大多是應用在 *E. coli* 上，雖然這方法有陷入局部最佳解(local optimum)以及其解可能無效的問題存在，但依舊仍是廣受研究者喜愛的方法之一。

Metabolic Control Analysis[22]一開始是用來研究代謝路徑的調控關係，雖然後期被應用到信號傳遞網路及基因網路上。但此方法將變數分成具網路性質(control coefficients)和具本身特性(elasticity coefficients)兩類參數，能直接對照過去的控制理論是這方法的特性。

Biochemical Systems Theory[23]是根據微分方程組(Ordinary differential equations, ODEs)依 Power-Law 展開而成的數學模型，可根據模型變數之指數辨別其變數對於網路之影響能力。這套方法發展至今已有各種變形，如 S-system 模型可見一斑[24]。

還有一種方法是依據酵素動力學的反應方程式建造整個模型，我們選擇使用動力學模型來建構本篇研究的反應模型架構，對代謝系統建模的方法有很多已經在論文中依各種不同的動力學架構而被發表[25, 26]。這種方法因為具有基本的化學反應動力學理論存在，能呈現出在反應系統之中的短暫變化及酵素所造成之通量和對其影響之回應，使得模型具有生物意義。用系統組織的方式會有助於找出一些重要的系統特性，但這種特徵在當簡化問題時會被隱藏起來[27]，但我們並不能判斷這些資訊是不是我們所盼望發現的資訊。並且局部、分段的尋找解答也會限制解空間，無法更進一步的找出問題最適當的解答。所以當使用 ODEs 作建構模型的數學架構時，必須要使用系統化的方式計算方程組中的參數值。

在反應方程式中為了能精確的調適反應過程而需要許多動力學參數[28]，動力學參數的散佈[28-31]被用在探索反應彼此間之變化及潛在行為，但這些參數存在著不易定量的問題。雖然目前已經有大量的酵素動力學之參數被收集存放在一些特定資料庫之中[32, 33]，但是直接使用這些參數值卻不一定是最正確的方式[34, 35]。經由實驗取得的數值或者是在之前文獻裡所記載的實驗數據中，大多針對著整段反應路徑中的部分實驗所設計而經由計算而留下記錄。我們知道在實驗中要同時取得代謝網路所需要的數據是一件辛苦兼困難的工作，根據之

前研究要將部分已知的動力學參數組織並串連成為一個完整的連續反應中，是個廣為流傳使用的方式[25, 26]。但是在一個較複雜或是將數個生化路徑結合起來(例如：Glycolysis 結合 TCA cycle)的生化反應系統中，會從中發現在文獻取用的參數數值通常會是處在不同環境或不同物種下的實驗值[28]，這種問題會使得依據此數據所建造的模型可能失去一些它們本來的正確性。

在 2005 年 Liebermeister 和 Klipp 的研究[30]中提到，參數間的微小變化可能會影響定性上的行為模式，同時也會對系統行為有些許影響。基於熱力學的限制之下，動力學參數並不適宜被分別挑選得出；而統計上，這些參數同時跟其它因素相關的情況發生頻率不算小，因為參數彼此之間會因為反應中相關反應物的濃度、反應速度的快慢及反應熱力學的效應而互相影響其值，所以這些參數具有需要利用整體最佳化特性存在。此外為求得動力學參數還存在著幾個問題就是實驗測量數值之誤差、生物本身所存在的變異性以及大多數的動力學參數都是未知或是不明、不確定[30]。

要藉由電腦計算的方式求得模型裡之大量參數數據是件不容易的事，若要得到較為準確的化學動力學參數則還是需要依據相關的熱力學性質[36, 37]。用數值分析及模擬處理具有詳盡動力學機制的模型會有高維度空間(high-dimensional space)的問題[38]。要計算大量參數並最佳化是一個難以處理的問題，因此藉由其他方法將問題降階簡化以減少問題的複雜度和困難。這些方法目前已被接受且普遍使用[36, 38-40]。而使用演化式演算法尋找最佳解的方式是一個對使用動力學模型處理複雜反應系統的良好方法[41]，並且不會受到參數未知的限制是這個方法的優點。使用演化式演算法做這種問題的研究通常在評估函數上有兩種選擇：一個是對系統總 Gibbs 自由能最小化[42, 43]；另一個則是與實驗值之誤差做最小化[44, 45]。雖然陸續已有人提出為了使模型更接近現實情況需要再多考慮熱力學的能量[19]，但都是將其當作是一種判斷解答的限制，而不是同時將熱力學能量當成是一個最佳化的目標函數，至今仍尚未同時考量最小化實驗數據之誤差以及系統總能量之研究。

1.3 研究目的

從化學反應平衡的觀點，若是相當快速的化學反應或者是完全反

應，大多數的反應最後幾乎都會走向平衡狀態，在生物體內的化學反應亦是如此。而生化反應會為了需要達成某種目的而進行反應，當反應完成隨即又會回復到一個新的穩定狀態之下，等待著下一次反應的開始。這樣的情況稱為生理穩定(Physiological Steady State)，所以當代謝反應進行後會有一個穩定狀態的存在，但要注意的是該狀態並不會持續存在。

在我們的研究目標是希望藉由所使用的動力學模型計算反應過程中系統總能量的變化，找尋總能量和實驗誤差同時最小化之一組動力學參數最佳解，直接由智慧型基因演算法求出最適合這條反應的動力學參數。

代謝系統之中的動力學參數之間所存在的熱力學限制：反應代謝物的 Gibbs 生成自由能經由計算會決定反應平衡常數，它會產生出反應之間的動力學參數以及橫跨整個網路的限制。一般將來自於實驗的動力學參數值直接放入模型使用，但這樣的使用將有可能具有熱力學上的錯誤[34,35]。從酵素活性、質量守恆、能量守恆跟反應平衡限制之下，使用智慧型基因演算法搜尋具有反應過程中最低總自由能的整組反應之動力學參數。利用基因演算法的特性能有效地求解並且毫無簡化問題本身，在眾多基因演算法中被選擇使用的方法是智慧型基因演算法(Intelligent genetic algorithm, IGA)以及智慧型多目標基因演算法(Intelligent Multi-Objective Genetic Algorithm, IMOGA)，因為它能有效解決大量參數的問題以及本身搜尋最佳解的優秀能力[46]，而能採用整體最佳化考量的作法去避免局部解問題的發生。

我們所使用的目標函數有兩項：最小化系統 Gibbs 自由能之最小值以及最小化模型估算與實驗濃度數據之誤差總和。以前研究通常單獨使用其一當成它們的目標函數，但在我們實驗中發現當同時使用兩者會造成自相矛盾的結果：當實驗濃度之記錄無誤只用最小化誤差總和即可，但會導致違背熱力學原則；而當實驗濃度考量本身帶有誤差存在時，僅使用誤差總和最小化無法求得最佳解。為了處理這樣問題的發生，我們建議增加系統 Gibbs 自由能最小化當作另一項目標函數為參考依據。

雖然兩目標函數存在著衝突，但藉由我們所提出的演算法可適當處理，所以我們方法適合解決此類型的問題。而本篇研究目的就在於提出一個方法可以提供具有參數不明確問題的生化反應系統一組經過

整體系統最佳化、正確且穩定的動力學參數，並能夠同時符合熱力學上之限制及提高動力學模型與實驗數據的相似度。

圖 1. 是概略性地表示具有複雜關係之動力學參數需要經過整體最佳化的示意圖：

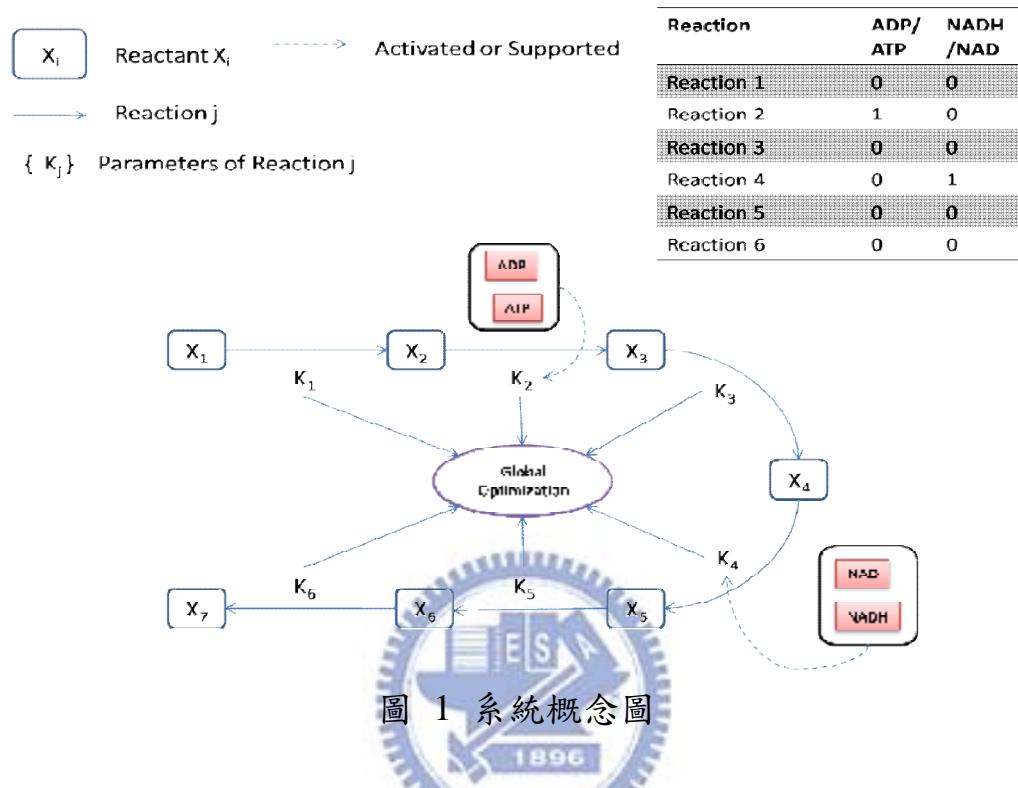


圖 1 系統概念圖

第二章會將我們所作的問題做個簡單的定義。敘述整個動力學模型的三大架構：(1) 骨架：根據生化反應式的動力學方程式組；(2) 連接關係：利用 Debye Hückel 定律 以及 Haldane relationship 計算反應間之通量；(3) 溶液環境：利用鍵結多項式、 H^+ 和 Mg^{2+} 的離子平衡式以及緩衝溶液的模擬，構成整個反應系統。在 2.3 應用於肌肉的肝醣分解之代謝機制上。

第三章介紹本篇研究使用之智慧型基因演算法機制以及所使用之評估函數。3.3 節更進一步利用智慧型多目標基因演算法(Intelligent Multi-Objective Genetic Algorithm, IMOGA)找出 Pareto Front，並從中提供一組最佳解供人以後參考使用。

第四章為介紹本研究所提出之方法 EKTOM(Estimation of Kinetic parameters using Two-objective Optimization Methods)之說明。並分別再對選擇不同最佳化演算法的搭配流程做詳細敘述。

第五章將針對本篇之應用簡單說明 EKTOM 其流程。

第六章為實驗結果及分析。首先證明在一個代謝網路中的系統總

能量中會有一個能量最低點的存在，說明使用 f_G 的意義並驗證。再利用 IGA 分別對 f_G , f_E 和 Hybrid(f_G+f_E) 作實驗，分析當實驗濃度有無誤差時的效果。由於 f_G 與 f_E 具有互相衝突的特性存在，故使用 IMOGA 作 Pareto front 從中提供一組最佳解以滿足雙目標之要求，將此最佳解與 Vinnakota[44]相比較其精確度及穩定性。



二、代謝反應路徑之動力學模型的結構

2.1 模型理論

2.1.1 中介紹所使用之動力學模型的相關方程式。2.1.2 是介紹 Debye-Hückel theory 以修正在不同環境下會發生改變的標準 Gibbs 生成自由能。2.1.3 說明考慮一化學反應系統時，熱力學的特性會使得系統中出現的限制條件及其原則。

2.1.1 動力學模型

這套模型有一些基本假設，就是跟陽離子鍵結的成分對酵素都有一樣的親和力，且在固定 pH 值之下我們維持 Mg^{2+} 和其他離子的濃度。另外我們參考基本的生理狀況下，先取 pH、溫度和離子強度起始值。在方程組中有關正向酵素活性值(E_{af})的計算我們引用 Scopes *et al.*[47] 所收集的實驗值，而逆向酵素活性值(E_{ar})值就使用 Haldane 關係式直接計算。模型中幾個反應的速率計算式之中，我們假設酵素活性會根據環境酸鹼度而改變，所以使用修正式改變所使用的酵素活性值。

根據熱力學公式計算出的反應平衡常數值(K_{eq}^m)當作反應指引的根據，想讓反應過後的濃度能在該反應之平衡常數的容許範圍內。當反應在不同環境之下， K_{eq}^m 也會根據反應環境不同而改變，在模型之中我們會重新計算做適當修正。當 pH 及 I 存在時，須考量根據它們所產生的影響而作應對的變化。當系統環境與實驗一致，對於同一生化反應所測得的實驗數據還是可能會不一樣，當然所測得之數據不可避免的還是須要考慮人為誤差。對於使用的動力學參數有不同環境的數據存在其中，對於使用這些數據建構模型還原的準確度上就會造成影響[34]。這份懷疑並不是指它不正確，因為之前的研究都是根據實驗數據所作，也可從文獻中看到模擬的結果。通常的確是有如同實驗值的傾向存在並呈現出來，但的確是有可再改善的空間。

使用建模型的反應參數想要在一致的條件跟環境之下能一次取得的全部數據資料，相信對於建造模型會有更精準的表現[48]。但是動力學模型架構的基本是一組微分方程組，在一組微分方程組之中最

難的問題就是該如何決定在方程式之中的大量參數[30]。我們從整個系統能量之觀點出發，而從化學能及熱力學能量的觀點發展相關代謝系統的研究已有先例[43, 49-52]。並且從熱力學觀點可以提供對問題的限制[53]，進而讓問題得到解決或者得到更接近真實的方法。而動力學參數會影響到反應平衡常數，改變反應的平衡狀態而影響著整體系統的運作。可知動力學參數與熱力學平衡關係非常密切。

目前已知道在 *E coli* 的細胞質液中大約會存在 0.15~0.20M 之間的離子強度[37]。而根據擴展的 Debye-Hückel 定律[54]可將反應物之生成自由能作計算修正，再依照各個成分的反應量計算該反應的反應自由能，再將該反應的反應自由能根據 van't Hoff 方程式¹計算該反應之反應平衡常數。迭代計算反應過程中的流通量、濃度、自由能、pH 值，以及反應後的反應平衡常數。

在使用動力學模型的好處是它較易於令人了解反應物所參與的反應，例如某化學反應：



該反應的一般速率表示式為：

$$Flux_m = \frac{\frac{E_{af} X_A X_B}{K_a K_b} - \frac{E_{ar} X_P X_Q}{K_p K_q}}{1 + \frac{X_A}{K_a} + \frac{X_B}{K_b} + \frac{X_A X_B}{K_a K_b} + \frac{X_P}{K_p} + \frac{X_Q}{K_q} + \frac{X_P X_Q}{K_p K_q}} \quad (2)$$

(2)式之中的 K_a 、 K_b 、 K_p 和 K_q 就是這篇研究所要求出的動力學參數值。 X_i 是反應物 i 的濃度， E_{af} 和 E_{ar} 是正逆向的酵素活性， $Flux_m$ 是的 m 個反應的反應通量。這篇研究所使用動力學模型的定義可參閱[55]。 E_{amf} 、 E_{amr} 兩者知其一即可藉由 Haldane 關係式，可從正向反應推算逆向反應，反之亦可。Haldane 關係式如下：

$$E_{amr} = \frac{E_{amf} \prod_{m=1}^M K_{mr}}{K_{eq}^m \prod_{m=1}^M K_{mf}} \quad (3)$$

以(3)式求出的值代入(2)式可得到該反應的反應通量。在(3)式中的 K_{eq}^m 為反應 m 的反應平衡常數， K_{mf} 是正向反應的動力學參數， K_{mr} 是

¹ van't Hoff 方程式： $\ln\left(\frac{K_2}{K_1}\right) = -\frac{\Delta H^\circ}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)$

逆向反應的動力學參數。

在開始計算反應之前需要先針對反應環境計算 $\Delta_f G_i^{o'}$ ，其值來自 expanded Debye-Hückel theory 的計算而得第 m 個反應的莫耳反應自由能 ($\Delta G_m^{o'}$)，

$$\Delta G_m^{o'} = \sum_{products} \Delta_f G_i^{o'} - \sum_{reactants} \Delta_f G_i^{o'} \quad (4)$$

並要考慮在環境中所存有之金屬離子的影響，我們使用 P 為陽離子鍵結多項式(binding polynomial)[56]：

$$P = 1 + \sum_{p=1}^N \frac{[H]^p}{\prod_{j=1}^p K_{a,j}} + \sum_{q=1}^Q \frac{[Metal]^q}{K_q} \quad (5)$$

並從(5)式的 P 進一步求出當下環境的反應平衡常數：

$$K_{eq}^m = \exp\left(\frac{\Delta_r G_m^{o'}}{RT}\right) * \frac{\prod_{products} P}{\prod_{reactants} P} \quad (6)$$

同時，也需計算在不同環境下的平均質子鍵結數 \bar{N}_H^1 [56]，用於計算反應中質子的變動。計算平均離子之鍵結數其詳細理論於 Alberty[54,56]之中。但因為所使用的熱力學參數是在標準狀態下的數據，所以還要使用 van't Hoff 方程式根據模型所假定的溫度對參數做修正。

要能在一個環境多變的生物系統下計算精確自由能的理論值是相當困難的，因為有太多會對反應過程造成影響，例如濃度、溫度、電價數、離子強度或者是分子間碰撞次數都會影響到反應的能量。而這份研究從原子不滅之平衡式開始，將氫離子和鎂離子的濃度影響以用偏微分方程式展開[44]。

$$[H_{total}] = [H_{free}] + [H_{bound}] + [H_{reference}] \quad (7)$$

$$[Mg_{total}^{2+}] = [Mg_{free}^{2+}] + [Mg_{bound}^{2+}] \quad (8)$$

(7)式是氫離子的原子守恆式， $[H_{total}]$ 是氫離子總濃度， $[H_{free}]$ 是自由氫離子的濃度， $[H_{bound}]$ 是反應物鍵結之平均氫離子數， $[H_{reference}]$ 是在反應物上的氫離子濃度。(8)式是鎂離子的原子守恆式， $[Mg_{total}^{2+}]$ 是鎂離子總濃度， $[Mg_{free}^{2+}]$ 是自由鎂離子的濃度， $[Mg_{bound}^{2+}]$ 是跟反應物鍵結之平均鎂離子數。將兩式取微分並將氫、鎂的自由離子濃度移到

等式左側可得(9)、(10)式。

$$\frac{d[H_{\text{free}}^+]}{dt} = -\frac{d[H_{\text{bound}}^+]}{dt} - \frac{d[H_{\text{reference}}^+]}{dt} \quad (9)$$

$$\frac{d[Mg_{\text{free}}^{2+}]}{dt} = -\frac{d[Mg_{\text{bound}}^{2+}]}{dt} \quad (10)$$

下面兩式是計算在系統中，反應物鍵結之平均氫離子數和平均鎂離子數的變化。

$$\begin{aligned} \frac{d[H_{\text{bound}}^+]}{dt} &= \sum_{i=1}^I \left(\frac{d\bar{N}_H^i}{dt} C_i + \bar{N}_H^i \frac{dC_i}{dt} \right) \\ &= \sum_{i=1}^I C_i \frac{\partial \bar{N}_H^i}{\partial [H^+]} \frac{d[H^+]}{dt} + \sum_{i=1}^I \left(C_i \frac{\partial \bar{N}_H^i}{\partial [Mg^{2+}]} \frac{d[Mg^{2+}]}{dt} + \bar{N}_H^i \frac{dC_i}{dt} \right) \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{d[Mg_{\text{bound}}^{2+}]}{dt} &= \frac{\sum_{i=1}^I \frac{\partial \bar{N}_{Mg}^i}{\partial pH} \frac{dpH}{dt} C_i + \sum_{i=1}^I \bar{N}_{Mg}^i \frac{dC_i}{dt}}{-(1 + \sum_{i=1}^I \frac{\partial \bar{N}_H^i}{\partial Mg^{2+}} C_i)} \end{aligned} \quad (12)$$

由[44]的研究中可以得知在這個生化反應系統之下，整條反應途徑會受到 $[Mg^{2+}]$ 、 $[K^+]$ 和 $[H^+]$ 的影響。

根據[57]的研究顯示出生化反應中的酵素會隨著細胞質 pH 值的影響而敏感地反應在酵素活性上。所以在[44]的動力學模型中已將所使用的酵素其活性對 pH 加以修正，並放進動力學模型計算之中。在反應不斷進行時，pH、 Mg^{2+} 、酵素活性以及參與反應的反應物會不斷變化，而反應程度會隨著 $[Mg^{2+}]$ 、 $[K^+]$ 和 $[H^+]$ 做適度調整。反應結果我們會跟[47]中的 Exp.29 做驗證，我們將結果放在後面章節討論。

2.1.2 擴展後的 Debye-Hückel 定律

各反應物的反應消耗(或生成)單位莫耳生成熱($\Delta_f G_i^o$) 算出該成分在這個反應中所提供之能量，計算各成份的反應能量總和可得到反應熱。而生成熱會受到 pH、溫度(T)、離子強度(I)以及電價數(Z_i)的影響，根據擴展後的 Debye-Hückel 定律在不同溫度下的離子物質可用下式計算[54, 58]，且當離子強度為 0.05 ~ 0.25M 範圍內的計算結果相當具有公信力。其公式如下式：

$$\Delta_f G_i^{o'} = \Delta_f G_i^o (I = 0) + N_H(i)RT \ln(10) pH + \frac{2.91482(Z_i^2 - N_H(i))I^{1/2}}{1 + 1.6I^{1/2}} \quad (13)$$

根據(13)式將各個反應物的生成能 $\Delta_f G_i^{o'} (pH, I)$ 依固定的 pH 跟 I 逐一修正，計算出在定溫下各個代謝物的 Gibbs 生成自由能。

2.1.3 热力学上的限制

在熱力学平衡的觀點下，一連串的化學反應依序的發生，若反應為可逆反應，則反應終將達到反應平衡的狀態。根據勒沙特列原理 (Le Chatelier's principle) — 若一個已達平衡的系統被改變，該系統會隨之改變來抗衡該改變。在固定環境之下只會存在一個平衡點，當環境不變的情況下，平衡狀態就是唯一。而以圖 2 說明總體 Gibbs 自由能(G)與反應進行程度(ξ)微分值為零，是系統總體自由能的最低點。圖 2 是對單一反應的平衡說明，圖中之 $\Delta_r G$ 代表此反應的反應自由能在紅點上時的微小變化量，可等同於斜率理解。對單一化學反應而言，反應自由能與總體自由能相等。

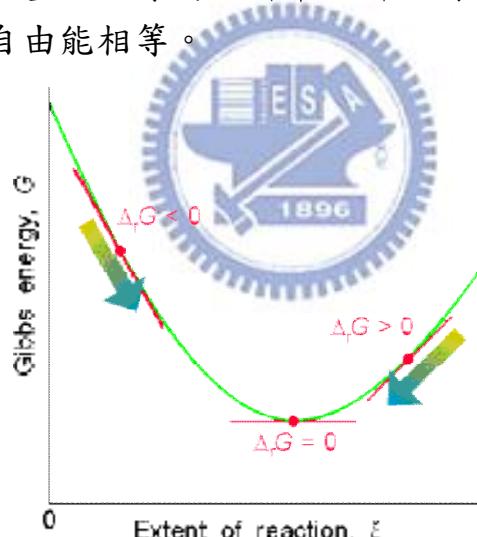


圖 2 系統總體 Gibbs 自由能對反應程度_[60]

在生物體內的代謝反應是被假定在屬於在定溫定壓之環境進行的。因為此類反應的反應量跟濃度成正比，不論是吸熱反應或是放熱反應都不足以改變系統環境的溫度。而在生物體內的壓力同樣不會有明顯的壓力變化存在，所以視為定壓狀態。由於生化反應路徑具有如此的特性，所以系統總體 Gibbs 自由能(G)照理應會有最小值的存在。

反應物經過反應後會因反應物濃度的生成消耗造成 G 的改變，因為 G 為 T 、 P 及系統各反應物莫耳數(n_i)的函數：

$$dG = (nV)dP - (nS)dT + \sum_{i=1}^I m_i dn_i, n_i = n_{io} + w_i^m x_m \quad (14)$$

在(14)式這個微分式當中 dG 是總體 Gibbs 自由能的改變量， nV 是總體積， nS 是系統總熵， m_i 是反應物的化學勢能。 w_i^m 是在第 m 個反應中反應物 i 的化學計量數， x_m 是第 m 個反應的反應程度。因為環境在定溫定壓下，上式變成

$$dG = \sum_{i=1}^I m_i dn_i \quad (15)$$

根據化學勢(chemical potential, μ_i)在此環境下的定義：

$$m_i \equiv \left[\frac{\partial(nG)}{\partial n_i} \right]_{T,P,n_j} \quad (16)$$

在[59]中由於化學勢是很難測的一個熱力學性質，但是在定溫、定壓的定義之下化學勢可等於 $\Delta_f G_i^{o'}$ ，所以在計算上就可以使用 $\Delta_f G_i^{o'}$

$$m_i^o \equiv \Delta_f G_i^{o'} \quad (17)$$

我們可將系統總自由能轉成下式：

$$G = G^\circ + \sum_{m=1}^M \sum_{i=1}^I w_i^m \Delta_f G_i^{o'} x_m = G^\circ + \sum_{m=1}^M \Delta G_m^{o'} x_m \quad (18)$$

每個反應物 i 所提供的能量變化是 $w_i^m \Delta_f G_i^{o'} x_m$ ，代表在反應 m 的反應自由能 $\Delta G_m^{o'}$ 。每次反應都會使 pH 改變造成 $\Delta_f G_i^{o'}$ 改變，故 G 也會跟著改變。

自然界的化學反應總是會伴隨著化學能的生成及消耗，而反應物本身會具有其物質本身的物理性質及化學性質。而反應物自己的標準生成熱就是其物理性質。反應物本身的標準生成熱會受到所處在的環境影響而改變，在這個研究之中會影響到標準生成熱的因素有溫度、離子強度和 pH。在實驗之中，我們給定溫度和離子強度，pH 和 $[Mg^{2+}]$ 則是隨著反應進行而改變，所以生成熱是不斷的受 pH 的影響而改變的。

由[60]的定義中：

“在封閉系統中的總體自由能，在固定 T、P 的不可逆程序中必須降低，且反應中系統總自由能 G 達到最小值時，即為平衡狀

態，在此平衡狀態時 $dG=0$ 。因此，若混合物中各成分不在化學平衡時，任何一個在固定 T、P 下的反應都會導致 G 的降低。”

所以在反應過程中，反應會朝著系統總體自由能降低的方向反應。當達到 $dG=0$ 時，為反應的平衡狀態。這是化學反應系統中的熱力學特性，可適用於研究的反應途徑的平衡狀態。由於生化反應會有達到反應平衡的特性，在化學反應限制之下濃度變化終究會趨於平衡，而平衡(equilibrium)是指當反應的正逆反應速率相等或者是濃度已不再增加或減少，即可被視為濃度達到平衡。

各個代謝物的標準莫耳生成熱 $\Delta_f G_i^o$ 都在表上查知，並經由(13)式計算在不同環境(例如 pH、T 及 I)所造成的影響偏差作修正，經由化學反應方程式計算在當此反應環境下該化學反應的反應自由能 ΔG_m^o ，所以全部反應自由能經加總起來變成整條代謝路徑的總自由能 G，而找尋反應過程中 G 的最低點就是我們設定的熱力學特性之目標函數。

2.2 系統流程



2.2.1 數學架構

首先假設整條反應路徑有 $m=1, 2, \dots, M$ ；共 M 條反應；反應物 i 的濃度(X_i)有 $i = 1, 2, \dots, N$ ，共 N 個反應物；對應反應的酵素活性 $E_{a_m f}$ 同為 M 個；反應速率式中使用的動力學參數 K_j ， $j = 1, 2, \dots, L$ ，共 L 個參數。

若反應 1 為：



分別將 X_A 至 X_Q 代入到(2)式成：

$$Flux_1 = \frac{\frac{E_{a1f} X_A X_B}{K_a K_b} - \frac{E_{a1r} X_P X_Q}{K_p K_q}}{1 + \frac{X_A}{K_a} + \frac{X_B}{K_b} + \frac{X_A X_B}{K_a K_b} + \frac{X_P}{K_p} + \frac{X_Q}{K_q} + \frac{X_P X_Q}{K_p K_q}} \quad (20)$$

$$E_{a1r} = \frac{E_{a1f} K_p K_q}{K_A K_B K_{eq}^1} \quad (21)$$

$$\ln\left(\frac{K_{eq}^1}{K_{eq}^1}\right) = \frac{-\Delta H^o}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \quad (22)$$

可得到(19)式的 $Flux_1$ 。整條反應路徑之中可能會有幾個反應會特別對pH敏感，需要特別透過下面通式加以對個別反應中的酵素通量修正以反應出pH對反應速度的影響[61, 62]。

$$E_{a1\ app} = E_{a1} \left(\frac{c}{1 + 10^{pKa1-pH} + 10^{pH-pKa2}} \right), \quad pKa1 < pH < pKa2 \quad (23)$$

上式中 $E_{a1\ app}$ 是修正後的酵素活性， $Ka1$ 及 $Ka2$ 是第一個氫離子的解離常數及第二個氫離子的解離常數， c 的值會因為反應酵素本身的性質而改變。

反應物的濃度變化則是根據相關反應的通量組合計算而成，假如一連續反應如下：



$X_E \rightarrow X_F$ ，為 $Flux_1$ ； $X_F \rightarrow X_G$ ，為 $Flux_2$ 。則反應後的濃度變化如下：

$$\frac{dX_E}{dt} = -Flux_1 \quad (25)$$

$$\frac{dX_F}{dt} = Flux_1 - Flux_2 \quad (26)$$

$$\frac{dX_G}{dt} = Flux_2 \quad (27)$$

而pH和 $[Mg^{2+}]$ 的變化速率為(11)、(12)式，若反應時間由 t_0 開始，而反應終止時間 t_f 由使用者決定：

$$X_{E1} = X_{Eo} + \frac{dX_E}{dt} \quad (28)$$

$$X_{F1} = X_{Fo} + \frac{dX_F}{dt} \quad (29)$$

$$X_{G1} = X_{Go} + \frac{dX_G}{dt} \quad (30)$$

$$pH_{t_f} = pH_{t_0} + \frac{dpH}{dt} \quad (31)$$

$$[Mg^{2+}]_{t_f} = [Mg^{2+}]_{t_0} + \frac{d[Mg_{\text{bound}}^{2+}]}{dt} \quad (32)$$

上列各式為在時間 t_0 至 t_1 的濃度變化量和 pH。

從上一步得到各個成分反應通量，我們再進一步將每次反應的反應通量用來計算在整個反應過程中，反應之間濃度的變化所造成能量之間的變化。在計算複雜的化學反應平衡中，常以反應程度(ζ)當做個別反應的時間軸。

$$n_i = n_{io} + W_i^m X_m \quad (33)$$

在一個封閉系統下，我們假設體積不隨反應的進行而改變，而反應物的濃度與莫耳數成正比，所以上式變成：

$$X_i = X_{io} + W_i^m X_m \quad (34)$$

經過反應後的個別反應的反應生成能：

$$\Delta G_m^{o'} = \sum_{m=1}^M w_i^m \Delta_f G_i^{o'} \quad (35)$$

(36)式所求得的總體自由能就是當整體反應達平衡時的總自由能，而此時 G 會是曲線的最低點。

$$G = G^\circ + \sum_{m=1}^M \sum_{i=1}^I w_i^m \Delta_f G_i^{o'} X_m = G^\circ + \sum_{m=1}^M \Delta G_m^{o'} X_m \quad (36)$$

$$dG = \sum_{m=1}^M \Delta G_m^{o'} dX_m \quad (37)$$

在我們的計算之中(5)式以及(13)式都會因為 pH 以及 $[Mg^{2+}]$ 的影響，(5)式還包括 $[K^+]$ 皆會隨反應進行而改變。在(13)式會影響該時間點的 $\Delta_f G_i^{o'}$ ；在(5)式會影響鍵結多項式，進而改變 K_{eq}^m 。接著會使反應通量受到影響，改變了反應物的反應速度， $\Delta_f G_i^{o'}$ 以及反應速率都會受到改變，系統的總自由能也是一樣。

而將 pH 及金屬陽離子的考量進模型之中，會使得模型的準確度上升，其效果已呈現在[44]。而本研究更進一步試著將會受到 pH 以及 $[Mg^{2+}]$ 的 G 當成評估函數之一，用以最小化系統總自由能。在整個反應過程之中不斷改變的 pH 以及 $[Mg^{2+}]$ 一定會造成 $\Delta_f G_i^{o'}$ 隨反應時間而改變。雖然變化幅度並不明顯，但這兩個因素的考量與否，會提高濃度曲線的準確度。所以在能量上面更不能忽略這兩個因素所造成的影响。

將參與反應的反應物在反應過程中所產生的能量加總起來得到整體反應的系統總自由能，每一次計算所得到的總系統自由能都在算反應當中該階段時的總能量，但計算並不是達到總系統自由能最低點即停止。而是在整個反應過程中能量最低點才是傳給演算法的函數分數。

2.2.2 解題方法

我們研究在於從熱力學以及化學平衡的觀點找出一組能通用於生化反應途徑的動力學模型之參數組，並且使之運用電腦模擬其化學反應其過程，而增進對系統生物學的了解與發展。為了找出一組最佳解，我們選擇使用能同時最佳化大量參數的智慧型基因演算法[46]。關於基因的編碼，我們使用實數編碼。將全部反應速率式中會使用到的動力學參數當成整條基因，反應速率式愈多，用到的動力學參數就愈多，則智慧型基因演算法找尋最佳解的能力顯得就愈強。

將每個反應物對時間的變化分別當成一個微分方程式，整條反應途徑就成了一組微分方程組。但因為有些生化反應本身的特性，讓反應物濃度會跟複數個反應同時相關，使得反應途徑的結構產生迴圈。具有這樣特性的微分方程組，被稱做「剛性方程式」。在[63]中使用基因演算法解這樣的微分方程組，計算動力學模型的參數。在[44]之中是直接使用 MATLAB(®Mathworks)中的解題器 — ode15s，可用在解常微分剛性方程式之中。在本篇研究之中，我們使用智慧型基因演算法搜尋最佳化的動力學參數，搭配 MATLAB 的 ode15s 解生化網路的常微分方程組，使用熱力學跟化學平衡當成我們的評估函數，藉此搜尋出最好的一組參數解。

2.3 方法應用

我們根據所參考論文中的反應方程式其中所列舉的反應速率常數(K_j)嘗試用智慧型基因演算法找尋整個反應系統中最佳的反應速率常數之組合。當染色體中的 K_j 值給起始值之後就讓反應速率方程式進行計算，模擬在單位時間下，計算連續反應後的濃度及能量變化，取出一組最符合我們所要求條件的解。

整條反應途徑中共 22 個反應物，在計算式裡我們以 X_{i0} 來表示初始濃度。 $i = 1, 2, \dots, 21, 22$ 。在 Ref.44 之中提到限制 $\text{PCr} + \text{Cr}$ 的總濃

度為 30mM。在[44]所討論的反應路徑裡，所引用的反應方程式模型[55]有包含以下特性：質量平衡和 Haldane 關係式所定義全部可逆之熱力學計算式。在 Vinnakota[44]中有更詳細的模型計算之解說。這個骨骼肌肉模型從 glycogen 到 lactate 共包括 16 個反應，我們將全部反應列在表 1，反應路徑放在圖 3。

表 1 肌肉醣原分解中全部反應

Reaction Name	Identifier	Chemical Reaction
<i>Glycogen Phosphorylase</i>	GP	$GLY_n + \Sigma Pi \leftrightarrow GLY_{n-1} + G1P$
<i>Phosphoglucomutase</i>	PGLM	$G1P \leftrightarrow G6P$
<i>Phosphoglucoisomerase</i>	PGI	$G6P \leftrightarrow F6P$
<i>Phosphofructokinase</i>	PFK	$F6P + \Sigma ATP \leftrightarrow F1,6P + \Sigma ADP$
<i>Aldolase</i>	ALD	$F1,6P \leftrightarrow GAP + DHAP$
<i>Triose Phosphate Isomerase</i>	TPI	$GAP \leftrightarrow DHAP$
<i>Glyceraldehyde-3-Phosphate Dehydrogenase</i>	GAPDH	$GAP + NAD + \Sigma Pi \leftrightarrow 1,3BPG + NADH$
<i>Glycerol-3-phosphate dehydrogenase</i>	G3PDH	$G3P + NAD \leftrightarrow DHAP + NADH$
<i>Phosphoglycerate Kinase</i>	PGK	$1,3BPG + \Sigma ADP \leftrightarrow 3PG + \Sigma ATP$
<i>Phosphoglyceromutase</i>	PGM	$3PG \leftrightarrow 2PG$
<i>Enolase</i>	ENOL	$2PG \leftrightarrow PEP$
<i>Pyruvate Kinase</i>	PK	$PEP + \Sigma ADP \leftrightarrow PYR + \Sigma ATP$
<i>Lactate Dehydrogenase</i>	LDH	$PYR + NADH \leftrightarrow LAC + NAD$
<i>Creatine Kinase</i>	CK	$PCr + \Sigma ADP \leftrightarrow Cr + \Sigma ATP$
<i>Adenylate Kinase</i>	ADK	$\Sigma ATP + \Sigma AMP \leftrightarrow 2\Sigma ADP$

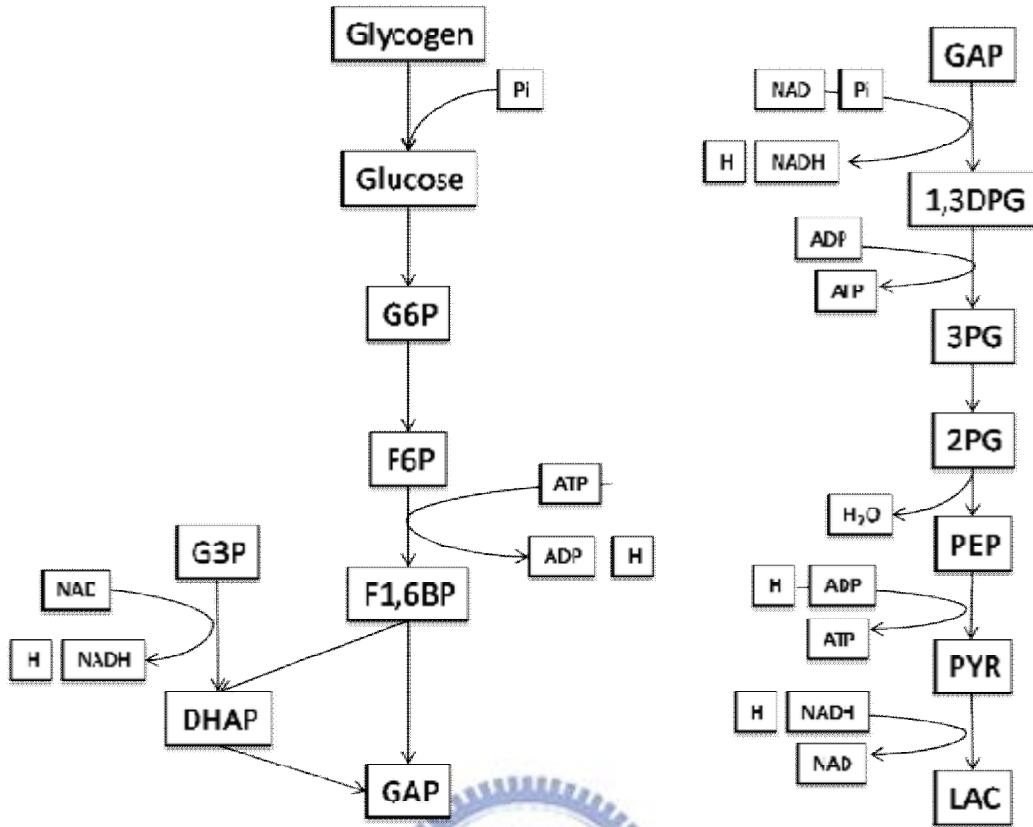


圖 3 肌肉糖原分解之反應途徑

2.3.1 使用擴展 Debye Hückel 定律

依照各反應物的反應消耗(或生成)單位莫耳生成熱($\Delta_f G_i^o$)算出該成分在這個反應狀況下中所提供之能量，計算各成份的反應能量總和可得到反應熱。而生成熱會受到 pH、溫度、離子強度以及電價數的影響，根據擴展的 Debye-Hückel 定律在不同溫度下的離子物質可用(13)式計算[54, 58]，且當離子強度為 0.05 ~ 0.25M 範圍內的計算結果相當值得信賴。

$$\Delta_f G_i^o = \Delta_f G_i^o(I=0) + N_H(i)RT \ln(10)pH + \frac{2.91482(Z_i^2 - N_H(i))I^{1/2}}{1+1.6I^{1/2}} \quad (13)$$

根據上面方程式將反應中各個反應物的生成能 $\Delta_f G_i^o(pH, I)$ 依特定的 pH 跟 I 逐一計算出在固定溫度下的各個代謝物的 $\Delta_f G_i^o$ 。經過計算得到 $\Delta_f G_i^o(pH, I)$ ，然後從 2.3.2 的動力學模型計算後，我們可以得到濃度變化值 ΔX_i ，但由於經反應後 pH 值會改變，所以我們需要再透過上式計算經過反應之後 $\Delta_f G_i^o(pH, I)$ 的改變，以提供下次反應計算所使用。

2.3.2 動力學函數以及模型

我們所使用的模型為具有連續性的標準微分方程之形式，而導數為該反應式的反應速率。從 Muscle glycogenolysis 的反應途徑中所導出的反應方程組(rate equations)中，我們要給三個種類的參數分別是在時間 t 的反應物濃度 $X_i(t)$ 、反應 m 的反應平衡常數 K_{eq}^m 和在染色體上所計算的動力學參數 K_j 。將上面參數依對應反應所需要的數值代入反應方程組裡之後，可得到該反應的反應通量(Flux_m)；然後將代謝物所參與反應的反應通量計算反應物的總反應量後，可得到該代謝物的反應速率。

從文獻[55]中所節錄出來的反應方程式中，令每個反應中的反應速率常數為 genes，將 genes 的值放進反應方程式裡，從 $X_i(t_0)$ 開始依序各個反應通量而改變每個反應物的濃度。所有反應物我們用表 2 表示。

表 2 反應物簡名及標號

Number	Abbreviation	Metabolite Name	Concentration.
1	GLY	Glycogen	X_1
2	G1P	glucose 1-phosphate	X_2
3	G6P	glucose 6-phosphate	X_3
4	F6P	fructose 6-phosphate	X_4
5	FBP	fructose 1,6-bisphosphate	X_5
6	DHAP	1,3-dihydroxyacetone phosphate	X_6
7	G3P	glycerol 3-phosphate	X_7
8	GAP	glyceraldehyde 3-phosphate	X_8
9	1,3BPG	glycerate 1,3-bisphosphate	X_9
10	3PG	glycerate 3-phosphate	X_{10}
11	2PG	glycerate 2-phosphate	X_{11}
12	PEP	phosphoenolpyruvate	X_{12}
13	PYR	Pyruvate	X_{13}
14	LAC	Lactate	X_{14}
15	Pi	inorganic phosphate	X_{15}
16	PCr	phosphocreatine	X_{16}

17	<i>Cr</i>	Creatine	X_{17}
18	<i>ATP</i>	adenosine-5'- triphosphate	X_{18}
19	<i>ADP</i>	adenosine diphosphate	X_{19}
20	<i>AMP</i>	adenosine monophosphate	X_{20}
21	<i>NAD</i>	Nicotinamide adenine di-nucleotide	X_{21}
22	<i>NADH</i>	Nicotinamide adenine di-nucleotide	X_{22}

表 3 反應物的起始濃度

Metabolite	X_{io} (mM)	Metabolite	X_{io} (mM)
GLY	0.04	PEP	10^{-9}
G1P	10^{-9}	PYR	10^{-9}
G6P	10^{-9}	LAC	10^{-9}
F6P	10^{-9}	Pi	0.03
FBP	10^{-9}	PCr	10^{-9}
DHAP	10^{-9}	Cr	0.03-PCr
G3P	10^{-9}	ATP	0.5
GAP	10^{-9}	ADP	0.005
13BPG	10^{-9}	AMP	10^{-9}
3PG	10^{-9}	NAD	0.0005
2PG	10^{-9}	NADH	10^{-9}

根據 Haldane relation 在全部反應之中計算逆反應速率值 E_{am_r} ，G3PDH、PGI 以及 PGK 是由逆反應速率值計算 E_{am_f} ，計算模型中所使用的值在表 4。 E_{am_f} 我們是採用[47]所記載的酵素活性，而起始條件來自之前的研究[64-67]。

當相對應反應式中的 E_{am_f} 、 K_{eq}^m 、 X_{io} 以及染色體上對應的 K_j 代入方程式計算，可以得到各個反應通量 $Flux_m$ 。在反應中的反應物中除了一開始的 glycogen 外，都同時有著生成通量和消耗通量。這些通量經過下列函式所得的值，就是反應物濃度的增減量。

表 4 參與反應的酵素活性

Enzyme	$E_{am\ f}$ (M/min)
GP	0.5
PGLM	0.48
PGI	0.88
PFK	0.056
ALD	0.104
TPI	12.0
GAPDH	1.265
PGK	1.12
PGM	1.12
EN	0.192
PK	1.44
LDH	1.92
ADK	0.88
CK	0.5

而反應物濃度變化速率如下：

$$\frac{dX_1}{dt} = -frac_a * Flux_a + frac_b * Flux_b \quad (38)$$

在[47]中第 29 號實驗中， $frac_a = 0.2\%$ ； $frac_b = 99.8\%$ ，其餘通量計算皆如同(25)式至(27)式之描述。從反應通量式可得 X_i 的濃度變化，依序將 X_i 的濃度變化求出，逐個將此計算值加到代入計算的濃度，然後成為下一個時間點計算的濃度值。總反應時間的長短是依據[47]的實驗而定，其反應總花費時間為 30 分鐘。

三、最佳化演算法之使用

3.1 智慧型基因演算法

為解決選擇動力學模型時的大量參數最佳化問題，必須藉由一強而有力的最佳化演算法。本文使用智慧型演化式演算法[41]做為最佳化工具。智慧型基因演算法結合了基因演算法與直交實驗兩種方法之特性，具有收斂速度快，精確度高的優點。

3.1.1 直交表與因素分析

IGA[46]和一般基因演算法最大的不同之處在於，IGA 在交配(crossover)的過程使用直交表(orthogonal arrays)挑選好的參數，使得IGA能克服參數過多染色體過長的問題。直交表(OA)與因素分析常用在品質控制的方法中[68]，也可運用於很有效率地改善交配操作。假設有兩水準的N個因素的直交表，那總共會有 2^N 種不同的組合。當四組(1,1), (1,2), (2,1)及(2,2)出現在全部實驗時，那麼兩個因素在直交表的行是彼此正交的。當任兩個因素在實驗集合是正交時，這個集合可被稱之為直交表(OA)。為了建立兩水準N個因素的直交表，我們可以得到一個整數 $n = \log_2^{[\log_2(N+1)]}$ ，來建立n列和(n-1)行的 $L_n(2^{n-1})$ 的直交表，然後選擇N行來使用。舉例來說表3.1則表示 $L_8(2^7)$ 直交表。

因素分析可以評估在評估函數中的因素效果，排名最有效果的因素以及決定每個因素的最好水準組合，進而促使評估函數有最佳的結果。直交實驗設計可以縮減因素分析的實驗次數。直交表的實驗次數在單一因素分析時只需n次。假設 y_t 是第t次實驗的函數評估值，我們定義 S_{jk} 是j因素在水準k的主效果：

$$S_{jk} = \sum_{t=1}^n y_t \times F_{k,t} \quad (39)$$

其中 F_k 為一個旗標值，，若第t次實驗中第j個因素選用水準為k，則 F_k 為1；若否，則 F_k 為0。若評估函數值為望大，則較大的主效果值表示對評估函數具有較佳的貢獻度；反之若評估函數望小，則主效果值小者貢獻度較佳。

主效果可以顯示因素中水準的個別影響。例如主效果 $S_{j1} > S_{j2}$ 則表示在參數最佳化的問題中，第j個因素水準值1對於整體最佳化函數的

貢獻大於水準值 2。如果相反的情形 $S_{j1} < S_{j2}$ ，則表示水準值 2 較佳。在各因素間無交互作用的前提下，主效果的分析可以用來推測出全實驗的最佳解。

直交因素實驗為一種部分因素實驗方式，可以有效減少參數設計時的實驗次數，並同時考慮實驗因素之間的交互作用。將直交因素實驗後的數據經過主效果分析，便可以將每個因素對於設計目標的貢獻優劣計算出來，推論出最佳解的實驗參數。

表 5 $L_8(2^7)$ 直交表

因 素 實 驗 編 號	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	1	1	1	2

3.1.2 智慧型交配運算

智慧型基因演算法與傳統基因演算法最大之不同，乃是以智慧型交配取代一般的單點交配或多點交配。傳統的交配方式無法評估染色體中參數個別的優劣，加上交配點是由隨機方式產生，得到的後代染色體品質不容易提昇。

事實上，我們可以將染色體交配過程視為一種因素實驗。將來自父代的兩個染色體中已切割好欲交配的片段，做為直交實驗的因素，並以染色體片段「互換」或「不換」做為兩種水準值。如此，以兩水準直交實驗產生出優良品質染色體的機率便可大幅提昇，進行步驟如下：

步驟一：令產生染色體中的油交配點所切割出的基因片段為實驗因素；假設因素數目為 n ，即有 n 個基因片段，選擇 $L_\beta(2^{\beta-1})$ 直交表的前 n 欄作為實驗之用，其中 $\beta = 2^{\lceil \log(n+1) \rceil}$ 。

步驟二：令因素 j 的水準 1 與水準 2 分別表示來自父代染色體 P_1

與 P_2 第 j 個基因片段。

步驟三：根據直交表，計算各因素組合實驗的評估值 y_t ， $t=1, 2, \dots, \beta$ 。

步驟四：計算主效果 S_{jk} ，其中 $j = 1, 2, \dots, n$ ， $k = 1, 2$ 。

步驟五：決定各因素的最佳水準。在評估函數望大時，則選擇主效果值較大之水準；在評估函數望小時，則各因素的最佳水準為主效果值較小之水準。如評估函數望大且 $S_{j1} > S_{j2}$ ，則因素 j 的最佳水準為 1；反之則最佳水準為 2。

步驟六：根據各因素的最佳水準，選擇對應父代染色體中的基因片段，組合成第一個子代染色體。

步驟七：將各因素的主效果差值 ($|S_{j1} - S_{j2}|$) 排名，差值越大者排名越高。

步驟八：以類似第一個子代染色體的方式來組合因素，將差值排名最差的因素，選擇與第一個子代相反的水準，則可產生第二個子代染色體。



3.2 方法流程

本論文以智慧型基因演算法做為挑選動力學參數的核心演算法，因此在本章節便以基因演算法的各步驟來介紹我們所提出的方法。

3.2.1 選擇運算

我們選用二元競爭法作為基因演算法的選擇運算，每次進行選擇時由族群中隨機挑選出兩條染色體，評估值較佳的染色體則可以進入基因重組運算(即交配與突變運算)。傳統的輪盤法選擇運算，在基因運算法演化至接近最佳解時，由於各染色體評估值變化不大時，造成每個染色體被選擇至下一代的機率都差不多，無法有效選出優秀的染色體。

而二元競爭選擇法的優點是，即使在染色體評估值變化不大時，仍可透過競爭方式取出表現比較優秀的染色體，避免基因演算法演化過早收斂。

3.2.2 智慧型交配

智慧型交配運算在本論文所提方法中扮演一個相當重要的角色。

智慧型交配運算結合了兩水準直交實驗與交配運算，能夠有效率的產生具有優秀評估值的新染色體。假設我們使用 $L_{N+1}(2^N)$ 直交表來做智慧型交配運算，本染色體交配運算的詳細進行步驟如下：

步驟一：假設即將進行交配運算的兩條染色體為 P_1 、 P_2 ，比對 P_1 、 P_2 內的參數基因，並將重複出現在兩條染色體內的參數基因移動至染色體末端，其相對應的控制基因。

步驟二：隨機將參數基因切割成 $[N/2]$ 個基因片段。每個基因片段及代表直交表的一個因素。 P_1 中第 j 個基因片段，及代表因素 j 的第一個水準值； P_2 則代表第二個水準值。

步驟三：計算直交表中每個染色體排列組合方式的適應值 f_t ， $t = 1, 2, 3, \dots, N+1$ 。

步驟四：計算各因素之主效果 S_{jk} ， $j = 1, 2, 3, \dots, N$ ， $k = 1, 2$ 。

步驟五：決定各因素的最佳水準。在評估函數望大時，則選擇主效果值較大之水準；在評估函數望小時，則各因素的最佳水準為主效果值較小之水準。如評估函數望大且 $S_{j1} > S_{j2}$ ，則因素 j 的最佳水準為 1；反之則最佳水準為 2。

步驟六：根據各參數的最佳水準，選擇對應父代染色體中的參數，組合出第一個子代染色體。

步驟七：將各參數的主效果差值排名，差值越大者排名越高。

步驟八：以類似第一個子代染色體的方式組合參數，除了盤明最差的參數選擇的水準與第一個子代染色體相反，可產生第二個子代染色體。

3.2.3 突變運算及演化終止條件

突變運算使用任意合理之亂數運算。假設一條染色體編碼之總長為 N 個位元，突變率為 P_m ，則每次的突變運算，隨機由染色體中選出 $'N * P_m'$ 個位元，然後以亂數變化所選中的動力學參數。

基因演算法所需的演化時間必須視問題的複雜度而定，較一般化的作法是設定演算法的評估次數與問題中的參數數目成正比。一旦終止條件達到設定值，演化即停止，並輸出所搜尋過的最佳解、用最佳解當模型參數所計算出的各個代謝物濃度值和反應計算過程中整體

自由能的值。

3.3 智慧型多目標基因演算法

IMOGA 在演化過程中基本上還是依循基因演算法的過程，但是在評估函數方式不像 IGA 是直接將目標函數當成評估函數而作演化判斷之根據，而是使用基於 Pareto 理論，使用通適化且不受尺度因素影響的評估函數(Generalized Pareto-based Scale-Independent Fitness Function , GPSIFF)作為評估函數。

3.3.1 基於 Pareto 理論通適化且不因尺度影響之評估函數

為了分辨出各染色體之間的優劣，本篇應用了基於 Pareto 理論為基礎的記分方式以避免受到尺度因素的影響，並且對於被支配解和未被支配解給於具有區分能力的適應函數值，用以取代傳統有失準確性的排名法和距離方式，稱之 GPSIFF。

GPSIFF 使用類競爭式(Tournament-Like)的記分方式來評估 Pareto 解集中染色體個體 x 的適應值，GPSIFF 的數學式如下：

$$Score(x) = p - q + c \quad (40)$$

其中 p 表示在目前欲評估的解集中 x 所支配的個體數目， q 表示在目前欲評估的解集中能夠把 x 支配的個體數目， c 是一個較大的正整數，以保證求出的適應值為一正整數。通常以目前參與評估運算的所有個體的數目作為正整數 c 的值。

GPSIFF 的優點如下：

(a)不需調整權重值：基於 Pareto 理論來評估解的好壞，沒有權重加總法需決定權重值的困難，也不會受到人為主觀判斷的影響。(b)不需考量尺度因素：由於各目標函數值的尺度適應值不盡相同，在權重加總法中需考慮到尺度因素，以免使得權重設定失之準確。(c)以積分方式有效辨識不同解的優劣程度：取代傳統排名法可能將不同的解給予相同的排名，以及距離法有尺度因素影響的缺點，以精確的記分評估解的優劣程度。

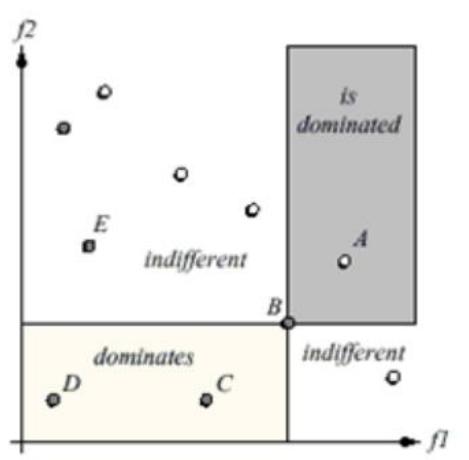


圖 4 支配與被支配關係示意圖

我們用圖 4 表示出在兩個目標中同時做最小化問題之說明，在 Pareto 解集中所有個體在雙目標軸的關係。我們以點 B 為例，點 B 之 f_1 和 f_2 同時都比點 A 小，所以點 A 被點 B 所支配；同理，點 C 與點 D 位於點 B 之左下角，點 B 就被點 C 與點 D 所支配，所以點 B 將不會被收在 Pareto 解集中。

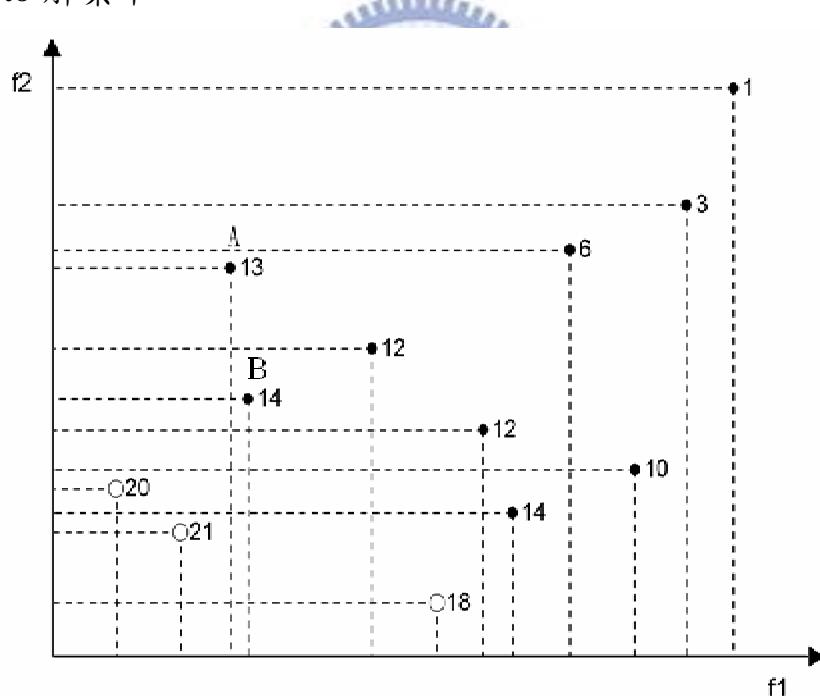


圖 5 GPSIFF 之示意說明圖^[46]

同樣的最小化問題，在圖 5 中點上的數字為適應值 c 。以 A 為例，其 $c=12$ ，TM 是未被支配解 ~ 是被支配解， $p=3$ ， $q=2$ ，所以適應值為 13。

3.3.2 演算法流程

智慧型多目標演算法之流程，其步驟詳細敘述如下：

步驟一：(初始化) 亂數產生起初使用的族群數量 N_{pop} 個染色體以及兩個空的優秀基因集合，一個是 E ；一個是 E' 。

步驟二：(評估初始值) 計算族群裡全部染色體的兩項目標函數值，並藉由 GPSIFF 分配每條染色體一個評估值。

步驟三：(更新優秀基因集合) 將未被支配(non-dominated)的染色體同時丟入 E 和 E' ，然後清空 E' 。考量 E 中的所有染色體，將被支配(dominated)的染色體移除。若 N_E 的數量大於原本所設定數量則將從中亂數去除超過的部分。

步驟四：(挑選) 從族群裡用 binary tournament selection 挑選出 $N_{\text{pop}}-N_{\text{ps}}$ 個染色體，並從 E 中亂數挑選出 N_{ps} 個染色體形成一組新的族群。其中 $N_{\text{ps}}=N_{\text{pop}} \times p_s$ ，若 $N_{\text{ps}} > N_E$ ，則令 $N_{\text{ps}}=N_E$ 。

步驟五：(重組) 藉由 Intelligent Gene Collector (IGC) 運作從 $N_{\text{pop}} \times p_c$ 選擇親代。每次 IGC 皆是由 OA 重組因子(副產物)找出未被支配的染色體以及兩個子代加入至 E' 。

步驟六：(突變) 根據 p_m 對整個族群進行突變機制

步驟七：(終止條件) 假設已滿足停止條件即可停止演算，反之回到步驟二。

四、 估算動力學參數之雙目標最佳化方法

4.1 估算動力學參數之模組化

本篇所提出的估算動力學參數的雙目標最佳化方法(Estimation of Kinetic parameters using Two-objective Optimization Methods, EKTOM)是為了要估算動力學模型中所必要的參數，藉由同時考慮將與真實數據之誤差以及系統熱力學之特性，利用智慧型基因演算法得出最佳化過的所求參數。

由於不一樣的起始狀態，一樣的環境條件下最終還是會達到一樣的平衡常數(因為溫度與壓力不變，化學反應不變則平衡狀態不變)。所以反應速率就會因應起始濃度的不同而加速反應或減速，但確定的是反應速率常數必定會產生改變其值使該反應能達到一樣的平衡狀態。

我們藉由計算總體反應自由能的方法求解最小化整體反應自由能在這反應方程組中的最佳反應速率常數之組合，基於能量的熱力學限制加上系統最佳化在整個系統一舉得出整組速率常數值。在整個流程計算中，我們要求的是反應速率常數、反應物濃度、反應 Gibbs 自由能，而這些數值循序漸進求出來才有最後的總反應自由能。整體計算過程我們將用圖來表示使過程能更清楚及容易了解。

反應總耗時設為 30 分鐘做為一次反應結束。改變基因演算法中上的 K_j 值(也就是染色體上的基因)重作計算以得到新的評估值。從不斷改變 K_j 值的組合來計算整條反應系統中最小的總體自由能，將具有這樣條件的一組 K_j 值當成是在我們已知之特定條件(T、I)下算出的最佳解。接著其他條件不變，用[44]所使用的 K_j 值放入我們的模型計算在這樣情況下反應經過後的 G 值比較之間的差距。

要將一個代謝路徑建成動力學模型，需要的兩方面分別是定義良好的速率方程式及精準的動力學參數。由於化學反應式都已經確定，但根據不同模型詮釋即會產生不同的速率方程式，使用不同的速率方程式就會需要不同動力學參數。但通常都很難將所需要的動力學參數

完整收集使用。所以我們提出用決定論的演化式演算法將動力學參數估算並利用其建構代謝路徑模型。

圖 6 為利用 EKTOM 估算之流程，其步驟詳細敘述如下：

步驟一：依照 4.3 將基因編碼成所要最佳化的動力學參數集合 K 。

步驟二：確定所要建立之代謝路徑，並將其化學反應式依方程式 (2) 展開速率方程式。

步驟三：根據本篇研究 2.1.1 之數學架構建立該動力學模型之微分式以及相互關係。

步驟四：利用 MATLAB 內建之解題器 ode15s 或自行建立能適用於剛性方程式解題器解速率之微分式，得全程反應物濃度。

步驟五：利用步驟四之結果計算本篇所使用之雙目標。

步驟六：依照使用者需求之目標函數，加權指數法的 IGA 或是根據 Pareto 理論之 IMOOGA，兩者擇一使用。

步驟七：IGA 之結果為一組 K ；IMOOGA 之結果為最佳解 K 之集合。

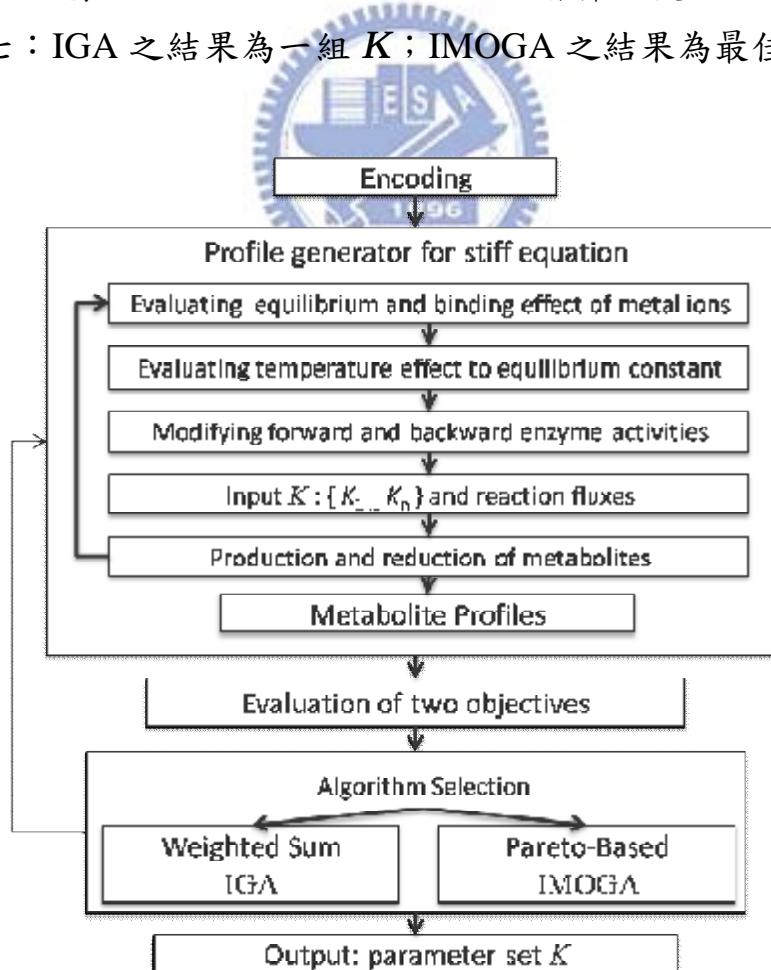


圖 6 EKTOM 之說明及流程圖

4.2 評估函數的兩項目標函數

我們所使用的目標函數有兩項，一項是根據動力學模型所計算之反應物濃度變化，估算其模型之濃度(X_i^{est})與評估反應物濃度(X_i^{exp})的差別，最小化與實驗數據之誤差作為第一項目標函數。模型濃度與實驗所記錄之濃度誤差總和(f_E)之計算如下式：

$$\min f_E = \sum_{i=1}^I |X_i^{est} - X_i^{exp}| \quad (41)$$

其中實驗記錄之濃度有 ADP、ATP、PCr、LAC、HMP、G3P 及 PG 共七個，而 HMP 是 Hexose monophosphate 的總濃度，PG 是 P3G + P2G+PYR 的總濃度。

另一項是有關系統總體自由能的函數。根據化學反應所設計的反應動力學方程式計算出反應程度(X_m)以及濃度變化量($\frac{dX_i}{dt}$)，在一個封閉系統下，我們應假設體積不隨化學反應的進行而改變，而反應物的濃度與莫耳數成正比，所以方程式(33)變成：

$$X_i = X_{io} + W_i^m X_m \quad (34)$$

經過反應後個別反應之反應生成能是從下式經計算而得：

$$\Delta G_m^{o'} = \sum_{m=1}^M w_i^m \Delta_f G_i^{o'} \quad (35)$$

然後進入能量計算的部分獲得單位時間內的在反應平衡前重覆計算至所設定的反應時間秒數(t_f)，然後記錄每次迭代計算出的濃度值與總反應自由能直到停止條件成立。如此可得到反應過程中系統總自由能的變化過程。能量的變化可被視為反應進行的方向，當反應持續進行時，系統總自由能是呈現減少的。從圖 2 的變化看到在達到平衡之曲線前段是負的，顯示反應是朝向系統總自由能減少的方向進行。而系統總自由能最小值愈低可代表反應的驅動力愈大，這是因為自然界的一項原則：能量值愈低其系統愈安定。由於真實的反應系統應具有較低之系統總自由能的特性，所以我們演算法中的第二個目標函數定為將反應過程之系統總體自由能的最低能量(f_G)最小化。

$$f_G = \min \left(G^\circ + \sum_{m=1}^M \Delta G_m^{o'} x_m \right) \quad (42)$$

目前 Gibbs 自由能最小化已被應用在一般的網路理論之中，而生化過程裡自由能熵和焓的變化是平衡熱力學中最基本的問題[43]。雖然這個目標函數通常是為了決定系統在平衡狀態時的各成分組態[41]，但是它依舊是一個建立生化模型的好方法[42]。

4.2.1 以加權指數法混合雙目標

由於在智慧型基因演算法(IGA)中所使用的評估函數在此篇研究中認定為單目標的演化式演算法。所以使用的目標函數利用加權指數法將兩項目標函數結合應用在 IGA 的評估函數中。

圖 6 是本篇研究使用 IGA 的方法流程圖，其演化過程與一般基因演算法的差異除了在 3.2 節已詳述過程外，不同的地方在於評估函數之計算。在圖 7 之 Fitness evaluation 的部分我們把演化過程中詳細進行步驟如下：

步驟一：將演化過程中產生之染色體送進 MATLAB。

步驟二：利用 MATLAB 解動力學模型中的剛性方程式之後，得到全反應物之濃度矩陣回傳給演算法主體。

步驟三：利用濃度矩陣分別計算 f_G 和 f_E ，依據不同的評估函數計算適應值，為該染色體之得分。

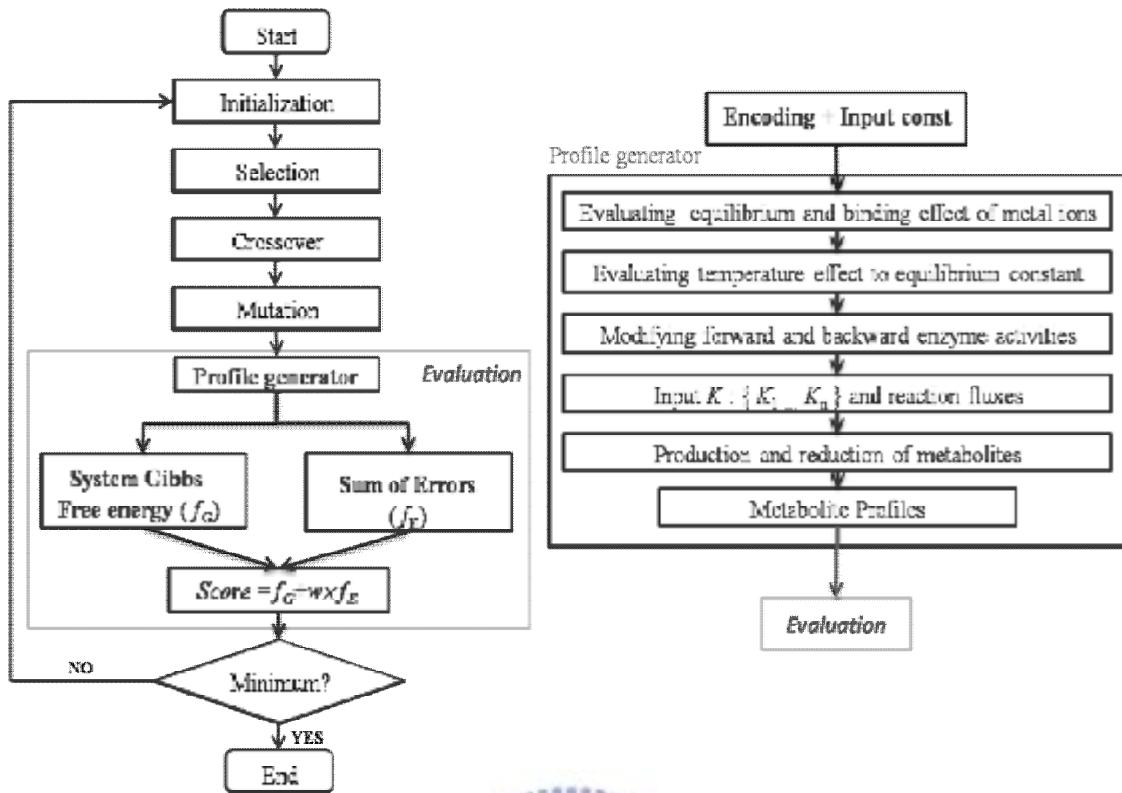


圖 7 應用智慧型基因演算法求最佳解之流程

在智慧型基因演算法中所設定的終止條件是「每代最小適應值之差異 $< 10^{-5}$ 的次數為 30 次」，在演算法終止後，可得最佳化過的一組動力學參數。

4.2.2 以最佳解集合同時考量雙目標

為了建出一套符合真實的動力學模型，當然不可缺少考慮減少模型估計值與真實實驗數據的誤差。但同樣為了符合真實反應之熱力學要求，我們認為多考慮系統總自由能為另一個最佳化之目標，也應視為模型須符合真實反應進行之條件。故我們使用智慧型多目標基因演算法，追求並建立出同時滿足兩項目標且更為精確之模型為此篇的研究目的。

在設定目標函數時， f_G 和 f_E 本身即存在著決定性的差異—函數值尺度不同且兩者差異極大，使得在 IGA 中運用加權指數法也無法得到良好的結果。並且在 IGA 的實驗中確定兩目標函數在最佳化中會互相衝突，所以最佳解集合將會出現 Pareto-Front 的分佈現象。

在智慧型多目標基因演算法中使用的評估函數為 GPSIFF，其優點

已在 3.3.1 節所提到，而其三大優點分別對應並解決在 IGA 實驗中所暴露出的缺點。

在圖 8 中我們把演化過程中詳細進行步驟如下：

步驟一：將演化過程中產生之染色體送進 MATLAB。

步驟二：利用 MATLAB 解動力學模型中的剛性方程式之後，將得到全反應物之濃度矩陣回傳給演算法主體。

步驟三：利用濃度矩陣分別計算 f_G 和 f_E ，並此染色體投影至以 f_G 和 f_E 為軸之解空間。

步驟四：藉由 GPSIFF 的計分方式將所有參與評估之個體計算其適應值，依去蕪存菁法保留較好之個體(non-dominated)。

步驟五：將未被支配(non-dominated)的染色體同時丟入 E 和 E' ，然後清空 E' 。考量 E 中的所有染色體，將被支配(dominated)的染色體移除。

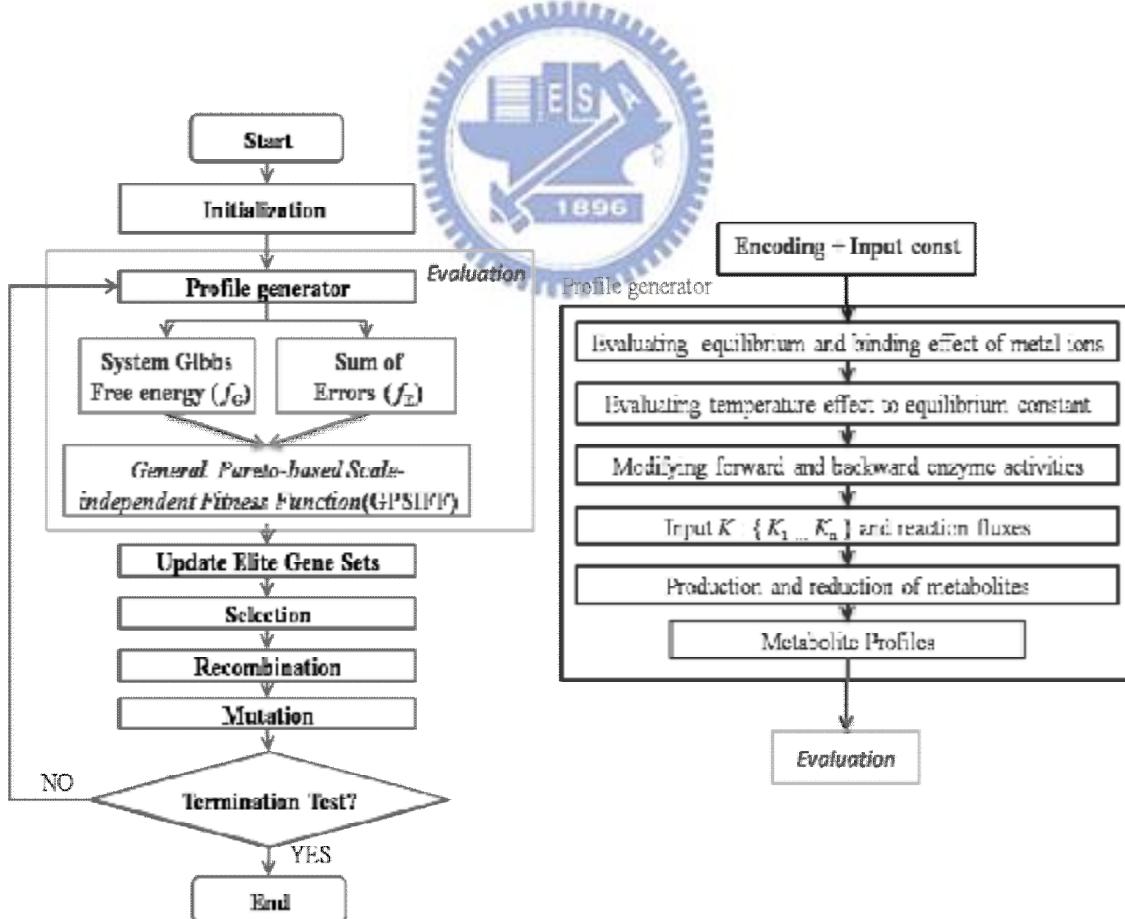


圖 8 應用智慧型多目標基因演算法求最佳解集合之流程
在智慧型多目標基因演算法中所設定的終止條件為個體計算次數

大於或等於 50000 次，達成此終止條件即結束演算法。當程式結束會將最外層的 Pareto-Front 解集合以及其實際參數值組合輸出。

4.3 染色體編碼方式

本計畫使用整條染色體的基因皆為實數編碼，用來對應各個反應中的動力學參數值。編碼方式如圖 9 所示：

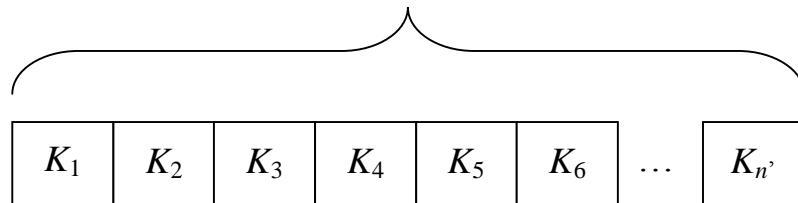


圖 9 染色體的編碼方式

由於為求動力學參數的精確數值，故在編碼時選用實數數值， K 代表動力學參數，總數 $n = 70$ 。因為實驗的反應計算量龐大，用傳統基因演算法會需要大量的計算時間。所以使用智慧型基因演算法能有效降低計算時間，同時也能得到較好的效能。



五、在骨骼肌肉醣解代謝上的應用

5.1 反應速率方程式

本小節收錄 Vinnakota[44,55]中骨骼肌肉醣解反應之所有詳細速率反應式(如式 42-a)、Haldane 關係式(如式 42-b)和酵素活性修正式(如式 42-c)。

從式 42 開始至式 57 為止，共有 16 個參與反應式的酵素名稱，名稱後面之 4 組號碼為 EC 號碼(Enzyme Commission)，每個編號各有及代表意義。式 58 到式 79 是全部反應物濃度對時間的微分式[55]。

Glycogen Phosphorylase A (2.4.1.1)

$$Flux_{GPa} = \frac{E_{a1f} \left(\frac{X_1 X_{15}}{K_{iGLY_f} K_{Pi}} \right) - E_{a1r} \left(\frac{X_1 X_2}{K_{GLYb} K_{iG1P}} \right)}{\left(1 + \frac{X_1}{K_{iGLY_f}} + \frac{X_{15}}{K_{iPi}} + \frac{X_1}{K_{iGLYb}} + \frac{X_2}{K_{iG1P}} + \frac{X_1 X_{15}}{K_{GLY_f} K_{Pi}} + \frac{X_1 X_2}{K_{GLYb} K_{iG1P}} \right)} \quad (42-a)$$

$$E_{a1r} = \frac{E_{a1f} K_{GLYb} K_{iG1P}}{K_{iGLY_f} K_{Pi} K_{eq}^1} \quad (42-b)$$

$$E_{a1app} = E_{a1f} \left(\frac{1.404}{1 + 10^{5.94-pH} + 10^{pH-7.29}} \right) , 5.75 \leq pH \leq 7.3 \quad (42-c)$$

Glycogen Phosphorylase B (2.4.1.1)

$$Flux_{GPb} = \frac{E_{a1f} \left(\frac{X_1 X_{15}}{K_{iGLY_f} K_{Pi}} \right) - E_{a1r} \left(\frac{X_1 X_2}{K_{iGLYb} K_{G1P}} \right)}{\left(1 + \frac{X_1}{K_{iGLY_f}} + \frac{X_{15}}{K_{iPi}} + \frac{X_1}{K_{iGLYb}} + \frac{X_2}{K_{iG1P}} + \frac{X_1 X_{15}}{K_{iGLY_f} K_{Pi}} + \frac{X_1 X_2}{K_{iGLYb} K_{G1P}} \right)} \left(\frac{\frac{AMP^{pH}}{K_{AMP}}}{1 + \frac{AMP^{pH}}{K_{AMP}}} \right) \quad (43-a)$$

$$E_{a1r} = \frac{E_{a1f} K_{iGLYb} K_{G1P}}{K_{iGLY_f} K_{iPi} K_{eq}^1} \quad (43-b)$$

$$E_{a1app} = E_{a1f} \left(\frac{1.75}{1 + 10^{6.12-pH} + 10^{pH-7.03}} \right) , 5.9 \leq pH \leq 7.2 \quad (43-c)$$

Phosphoglucomutase (2.7.1.41)

$$Flux_{PGLM} = \frac{E_{a2f} \left(\frac{X_2}{K_{G1P}} \right) - E_{a2r} \left(\frac{X_3}{K_{G6P}} \right)}{1 + \frac{X_2}{K_{G1P}} + \frac{X_3}{K_{G6P}}} \quad (44\text{-a})$$

$$E_{a2r} = \frac{E_{a2f} K_{G6P}}{K_{G1P} K_{eq}^2} \quad (44\text{-b})$$

$$E_{a2app} = E_{a2f} \left(\frac{1.329}{1 + 10^{6.64-pH} + 10^{pH-8.36}} \right) , 5.5 \leq pH \leq 7.9 \quad (44\text{-c})$$

Phosphoglucoisomerase (5.3.1.9)

$$Flux_{PGI} = \frac{E_{a3f} \left(\frac{X_3}{K_{G6P}} \right) - E_{a3r} \left(\frac{X_4}{K_{F6P}} \right)}{1 + \frac{X_3}{K_{G16P}} + \frac{X_4}{K_{F6P}}} \quad (45\text{-a})$$

$$E_{a3f} = \frac{E_{a3r} K_{G6P} K_{eq}^3}{K_{F6P}} \quad (45\text{-b})$$

$$E_{a3app} = E_{a3f} \left(\frac{1.0}{1 + 10^{6.36-pH} + 10^{pH-9.91}} \right) , 6.0 \leq pH \leq 10.0 \quad (45\text{-c})$$

Phosphofructokinase (2.7.1.11)

$$Flux_{PFK} = \left(\frac{E_{a4f} \left(\frac{X_{18} X_4}{K_{ATP} K_{F6P}} \right) - E_{a4r} \left(\frac{X_{19} X_5}{K_{ADP} K_{FBP}} \right)}{\Delta} \right) \times \left(\frac{1 + aL \left(\frac{\Delta'}{\Delta} \right)^3}{1 + L \left(\frac{\Delta'}{\Delta} \right)^4} \right) \quad (46\text{-a})$$

$$\Delta = \left(1 + \frac{X_4}{K_{F6P}} \right) \times \left(1 + \frac{X_{18}}{K_{ATP}} \right) + \left(1 + \frac{X_{19}}{K_{ADP}} \right) \times \frac{X_5}{K_{FBP}} + \frac{X_{19}}{K_{ADP}} \quad (46\text{-b})$$

$$\Delta' = \left(1 + \frac{X_4}{K_{F6P}} \right) \times \left(1 + \frac{X_{18}}{K_{ATP}} \right) + \left(1 + \frac{X_{19}}{K_{ADP}} \right) \times \frac{X_5}{K_{FBP}} + \frac{X_{19}}{K_{ADP}} \quad (46\text{-c})$$

$$L = L_o \left(\frac{\left(1 + \frac{X_{18}}{K_{iATP}} \right) \times \left(1 + e \frac{X_{20}}{K_{aAMP}} \right)^4}{\left(1 + d \frac{X_{18}}{K_{iATP}} \right) \times \left(1 + \frac{X_{20}}{K_{aAMP}} \right)} \right) \quad (46\text{-d})$$

$$E_{a4r} = \frac{E_{a4f} K_{FBP} K_{ADP}}{K_{F6P} K_{ATP}} \quad (46-e)$$

$$E_{a4app} = E_{a4f} \left(\frac{1.0}{1 + \left(\frac{pH}{6.8} \right)^{-30}} \right) \quad , 6.4 \leq pH \leq 8.5 \quad (46-f)$$

Aldolase (4.1.2.13)

$$Flux_{ALD} = \frac{E_{a5f} \left(\frac{X_5}{K_{FBP}} \right) - E_{a5r} \left(\frac{X_8 X_6}{K_{GAP} K_{DHAP}} \right)}{1 + \frac{X_5}{K_{FBP}} + \frac{X_6}{K_{DHAP}} + \frac{X_8}{K_{GAP}}} \quad (47-a)$$

$$E_{a5r} = \frac{E_{a5f} K_{DHAP} K_{GAP}}{K_{FBP} K_{eq}^5} \quad (47-b)$$

$$E_{a5app} = E_{a5f} \left(\frac{1.013}{1 + 10^{5.32-pH} + 10^{pH-9.15}} \right) \quad , 6.0 \leq pH \leq 9.0 \quad (47-c)$$

Triose Phosphate Isomerase (5.3.1.1)

$$Flux_{TPI} = \frac{E_{a6f} \left(\frac{X_8}{K_{GAP}} \right) - E_{a6r} \left(\frac{X_6}{K_{DHAP}} \right)}{1 + \frac{X_6}{K_{DHAP}} + \frac{X_8}{K_{GAP}}} \quad (48-a)$$

$$E_{a6r} = \frac{E_{a6f} K_{DHAP}}{K_{GAP} K_{eq}^6} \quad (48-b)$$

Glyceraldehyde-3-Phosphate Dehydrogenase (1.2.1.12)

$$\mathcal{U}X_{GAPDH} = \frac{E_{a7f} \left(\frac{X_8 X_{21} X_{15}}{K_{GAP} K_{NAD} K_{Pi}} \right) - E_{a7r} \left(\frac{X_9 X_{22}}{K_{13BPG} K_{NADH}} \right)}{1 + \frac{X_8}{K_{GAP}} + \frac{X_{21}}{K_{NAD}} + \frac{X_{15}}{K_{Pi}} + \frac{X_8 X_{21}}{K_{GAP} K_{NAD}} + \frac{X_8 X_{21} X_{15}}{K_{GAP} K_{NAD} K_{Pi}} + \frac{X_9}{K_{13BPG}} + \frac{X_{22}}{K_{NADH}} + \frac{X_9 X_{22}}{K_{13BPG} K_{NADH}}} \quad (49-a)$$

$$E_{a7r} = \frac{E_{a7f} K_{13BPG} K_{NADH}}{K_{GAP} K_{NAD} K_{eq}^7} \quad (49-b)$$

$$E_{a7app} = E_{a7f} \times 0.007 \times e^{0.8979 \times pH} \quad , 7.2 \leq pH \leq 8.7 \quad (49-c)$$

Phosphoglycerate Kinase (2.7.2.3)

$$Flux_{PGK} = \frac{E_{a8f} \left(\frac{X_9 X_{19}}{K_{13BPG} K_{ADP}} \right) - E_{a8r} \left(\frac{X_{10} X_{18}}{K_{3PG} K_{ATP}} \right)}{\left(1 + \frac{X_9}{K_{13BPG}} + \frac{X_{19}}{K_{ADP}} + \frac{X_9 X_{19}}{K_{13BPG} K_{ADP}} + \frac{X_{10}}{K_{3PG}} + \frac{X_{18}}{K_{ATP}} + \frac{X_{10} X_{18}}{K_{3PG} K_{ATP}} \right)} \quad (50\text{-a})$$

$$E_{a8f} = \frac{E_{a8r} K_{13BPG} K_{ADP} K_{eq}^8}{K_{3PG} K_{ATP}} \quad (50\text{-b})$$

Phosphoglycerate mutase (5.4.2.1)

$$Flux_{PGM} = \frac{E_{a9f} \left(\frac{X_{10}}{K_{3PG}} \right) - E_{a9r} \left(\frac{X_{11}}{K_{2PG}} \right)}{1 + \frac{X_{10}}{K_{3PG}} + \frac{X_{11}}{K_{2PG}}} \quad (51\text{-a})$$

$$E_{a9r} = \frac{E_{a9f} K_{2PG}}{K_{3PG} K_{eq}^9} \quad (51\text{-b})$$

$$E_{a9app} = E_{a9f} \left(\frac{0.989}{1 + 10^{5.62-pH} + 10^{pH-8.74}} \right) , \quad 4.9 \leq pH \leq 8.7 \quad (51\text{-c})$$

Enolase (4.2.1.11)

$$Flux_{ENOL} = \frac{E_{a10f} \left(\frac{X_{11}}{K_{2PG}} \right) - E_{a10r} \left(\frac{X_{12}}{K_{PEP}} \right)}{1 + \frac{X_{11}}{K_{2PG}} + \frac{X_{12}}{K_{PEP}}} \quad (52\text{-a})$$

$$E_{a10r} = \frac{E_{a10f} K_{PEP}}{K_{2PG} K_{eq}^{10}} \quad (52\text{-b})$$

Pyruvate Kinase (2.7.1.40)

$$Flux_{PK} = \frac{E_{a11f} \left(\frac{X_{12} X_{19}}{K_{PEP} K_{ADP}} \right) - E_{a11r} \left(\frac{X_{13} X_{18}}{K_{PYR} K_{ATP}} \right)}{\left(1 + \frac{X_{12}}{K_{PEP}} + \frac{X_{19}}{K_{ADP}} + \frac{X_{12} X_{19}}{K_{PEP} K_{ADP}} + \frac{X_{13}}{K_{PYR}} + \frac{X_{18}}{K_{ATP}} + \frac{X_{13} X_{18}}{K_{PYR} K_{ATP}} \right)} \quad (53\text{-a})$$

$$E_{a11r} = \frac{E_{a11f} K_{ATP} K_{PYR}}{K_{PEP} K_{ADP} K_{eq}^{11}} \quad (53\text{-b})$$

$$E_{a11app} = E_{a11f} \left(\frac{1.05}{1 + 10^{5.58-pH} + 10^{pH-8.79}} \right) , 5.0 \leq pH \leq 9.0 \quad (53-c)$$

Lactate Dehydrogenase (1.1.1.27)

$$Flux_{LDH} = \frac{E_{a12f} \left(\frac{X_{13}X_{22}}{K_{PYR}K_{NADH}} \right) - E_{a11r} \left(\frac{X_{14}X_{21}}{K_{LAC}K_{NAD}} \right)}{\left(1 + \frac{X_{13}}{K_{PYR}} + \frac{X_{22}}{K_{NADH}} + \frac{X_{13}X_{22}}{K_{PYR}K_{NADH}} + \frac{X_{14}}{K_{LAC}} + \frac{X_{21}}{K_{NAD}} + \frac{X_{14}X_{21}}{K_{LAC}K_{NAD}} \right)} \quad (54-a)$$

$$E_{a12r} = \frac{E_{a12f} K_{LAC} K_{NAD}}{K_{PYR} K_{NADH} K_{eq}^{12}} \quad (54-b)$$

$$E_{a12app} = E_{a12f} (-0.1134 \times pH + 1.6069) , 5.48 \leq pH \leq 10.49 \quad (54-c)$$

Creatine Kinase (2.7.3.2)

$$Flux_{CK} = \frac{E_{a13f} \left(\frac{X_{18}X_{Cr}}{K_{ATP}K_{Cr}} \right) - E_{a13r} \left(\frac{X_{19}X_{PCr}}{K_{ADP}K_{PCr}} \right)}{\left(1 + \frac{X_{19}}{K_{iADP}} + \frac{X_{PCr}}{K_{iPCr}} + \frac{X_{19}X_{PCr}}{K_{iADP}K_{PCr}} + \frac{X_{18}}{K_{iATP}} + \frac{X_{18}X_{Cr}}{K_{iATP}K_{Cr}} \right)} \quad (55-a)$$

$$E_{a13f} = \frac{E_{a13r} K_{iATP} K_{Cr} K_{eq}^{13}}{K_{iADP} K_{PCr}} \quad (55-b)$$

Adenylate Kinase (2.7.4.3)

$$Flux_{ADK} = \frac{E_{a14f} \left(\frac{X_{18}X_{20}}{K_{ATP}K_{AMP}} \right) - E_{a14r} \left(\frac{X_{19}}{K_{ADP}} \right)^2}{\left(1 + \frac{X_{18}}{K_{ATP}} + \frac{X_{20}}{K_{AMP}} + \frac{X_{18}X_{20}}{K_{ATP}K_{AMP}} + \frac{2X_{19}}{K_{ADP}} + \frac{X_{19}^2}{K_{ADP}^2} \right)} \quad (56-a)$$

$$E_{a14r} = \frac{E_{a14f} K_{ADP}^2}{K_{ATP} K_{AMP} K_{eq}^{14}} \quad (56-b)$$

Glycerol-3-phosphate dehydrogenase (1.1.1.8)

$$Flux_{G3PDH} = \frac{E_{a15f} \left(\frac{X_7 X_{21}}{K_{G3P} K_{NAD}} \right) - E_{a15r} \left(\frac{X_6 X_{22}}{K_{DHAP} K_{NADH}} \right)}{\left(1 + \frac{X_7}{K_{G3P}} + \frac{X_{21}}{K_{NAD}} \right) \left(1 + \frac{X_6}{K_{DHAP}} + \frac{X_{22}}{K_{NADH}} \right)} \quad (57\text{-a})$$

$$E_{a15f} = \frac{E_{a15r} K_{G3P} K_{NAD} K_{eq}^{15}}{K_{DHAP} K_{NADH}} \quad (57\text{-b})$$

$$\frac{dGLY}{dt} = -(0.002 \times Flux_{GPa} + 0.998 \times Flux_{GPb}) \quad (58)$$

$$\frac{dG1P}{dt} = (0.002 \times Flux_{GPa} + 0.998 \times Flux_{GPb}) - Flux_{PGLM} \quad (59)$$

$$\frac{dG6P}{dt} = Flux_{PGLM} - Flux_{PGI} \quad (60)$$

$$\frac{dF6P}{dt} = Flux_{PGI} - Flux_{PFK} \quad (61)$$

$$\frac{dFBP}{dt} = Flux_{PFK} - Flux_{ALD} \quad (62)$$

$$\frac{dDHAP}{dt} = Flux_{ALD} + Flux_{TPI} + Flux_{G3PDH} \quad (63)$$

$$\frac{dGAP}{dt} = Flux_{ALD} - Flux_{TPI} - Flux_{GAPDH} \quad (64)$$

$$\frac{d13BPG}{dt} = Flux_{GAPDH} - Flux_{PGK} \quad (65)$$

$$\frac{d3PG}{dt} = Flux_{PGK} - Flux_{PGM} \quad (66)$$

$$\frac{d2PG}{dt} = Flux_{PGM} - Flux_{ENOL} \quad (67)$$

$$\frac{dPEP}{dt} = Flux_{ENOL} - Flux_{PK} \quad (68)$$

$$\frac{dPYR}{dt} = Flux_{PK} - Flux_{LDH} \quad (69)$$

$$\frac{dLAC}{dt} = Flux_{LDH} \quad (70)$$

$$\frac{dPi}{dt} = -(0.002 \times Flux_{GPa} + 0.998 \times Flux_{GPb}) - Flux_{GAPDH} \quad (71)$$

$$\frac{dADP}{dt} = Flux_{PFK} - Flux_{PGK} - Flux_{PK} + 2Flux_{ADK} + Flux_{CK} \quad (72)$$

$$\frac{dATP}{dt} = -Flux_{PFK} + Flux_{PGK} + Flux_{PK} - Flux_{ADK} - Flux_{CK} \quad (73)$$

$$\frac{dAMP}{dt} = -Flux_{ADK} \quad (74)$$

$$\frac{dPCR}{dt} = Flux_{CK} \quad (75)$$

$$\frac{dCr}{dt} = -Flux_{CK} \quad (76)$$

$$\frac{dNADH}{dt} = Flux_{GAPDH} - Flux_{LDH} \quad (77)$$

$$\frac{dNAD}{dt} = -Flux_{GAPDH} + Flux_{LDH} \quad (78)$$

$$\frac{dG3P}{dt} = -Flux_{G3PDH} \quad (79)$$

5.2 計算範例說明

本篇研究以骨骼肌肉的醣解代謝路徑作為應用方法之說明。全反應共有 16 個反應、22 個反應物和 70 個動力學參數。

為了計算標準反應平衡常數，我們需要每個反應中的反應物之 $\Delta_f G_i^o$ ，經由(4)式算 $\Delta_f G_m^o$ 。根據(5)式計算 \mathbf{P} ，最後藉由(6)式計算該 K_{eq}^m 。因為反應並不在標準狀態(25°C , 1atm)下進行，所以須使用 van't Hoff 方程式修正為在 37°C 下的反應平衡常數。而反應的進行將會使得反應環境的 pH 值不斷發生改變，造成經過 Debye-Hückel 計算的 $\Delta_f G_i^o$ 一直改變計算結果，同時也使得反應平衡常數也隨著變化。但根據使用公式定義，在當時環境下的平衡常數應為公式之計算結果，只是因為反應的進行會使得反應環境發生變化，所以反應的平衡須重新計算，計算後的平衡常數是當時反應狀態的平衡常數。

E_{af} 會受 pH 值的影響，所以我們利用(23)式修正表 4 所提供的 E_{af} ，經(3)式以及套用修正後反應平衡常數計算另一方向的酵素活性。然後將反應之雙向酵素活性代入(2)式求出該反應之反應通量。

在此將更詳細的使用反應速率方程式，該反應之 Haldane 關係

式和活性修正式。

以 *Phosphoglycerate Kinase* (2.7.2.3) 舉例：

參與化學反應的反應物有 1,3BPG、3PG、ATP 及 ADP。首先，我們需要計算每個反應物的 P ：

$$P_{1,3BPG} = 1 + 10^{-pH + pK_{a1-1,3BPG}}$$

$$P_{3PG} = 1 + 10^{-pH + pK_{a1-3PG}}$$

$$P_{ATP} = 1 + 10^{-pH + pK_{a1-ATP}} + X_{Mg} * 10^{pKa_{Mg^{2+}-ATP}} + X_{K^+} * 10^{pK_{aK^+-ATP}}$$

$$P_{ADP} = 1 + 10^{-pH + pK_{a1-ADP}} + X_{Mg} * 10^{pKa_{Mg^{2+}-ADP}} + X_{K^+} * 10^{pK_{aK^+-ADP}}$$

然後利用 Debye Hückel 定律計算每個反應物修正後的 $\Delta_f G_i^o$ ：

$$\Delta_f G_{1,3BPG}^o = \Delta_f G_{1,3BPG}^o (I = 0) + N_H(1,3BPG)RT_1 \ln(10)pH + \frac{2.91482(Z_{1,3BPG}^2 - N_H(1,3BPG))I^{1/2}}{1 + 1.6I^{1/2}}$$

$$\Delta_f G_{3PG}^o = \Delta_f G_{3PG}^o (I = 0) + N_H(3PG)RT_1 \ln(10)pH + \frac{2.91482(Z_{3PG}^2 - N_H(3PG))I^{1/2}}{1 + 1.6I^{1/2}}$$

$$\Delta_f G_{ATP}^o = \Delta_f G_{ATP}^o (I = 0) + N_H(ATP)RT_1 \ln(10)pH + \frac{2.91482(Z_{ATP}^2 - N_H(ATP))I^{1/2}}{1 + 1.6I^{1/2}}$$

$$\Delta_f G_{ADP}^o = \Delta_f G_{ADP}^o (I = 0) + N_H(ADP)RT_1 \ln(10)pH + \frac{2.91482(Z_{ADP}^2 - N_H(ADP))I^{1/2}}{1 + 1.6I^{1/2}}$$

根據上列數式計算 $\Delta_r G_8^o$ ：

$$\Delta_r G_8^o = \Delta_f G_{3PG}^o + \Delta_f G_{ATP}^o - \Delta_f G_{1,3BPG}^o + \Delta_f G_{ADP}^o$$

最後計算反應平衡常數 $K_{eq}^{8\circ}$ ：

$$K_{eq}^{8\circ} = \exp\left(\frac{\Delta_r G_8^o}{RT_1}\right) * \frac{P_{3PG} * P_{ATP}}{P_{1,3BPG} * P_{ADP}}$$

因為反應並不在標準狀態(25°C, 1atm)下進行，所以須使用 van't Hoff 方程式修正為在 37°C 下的反應平衡常數：

$$\ln\left(\frac{K_{eq}^{8\circ}}{K_{eq}^{8\circ}}\right) = \frac{-\Delta H^o}{R} \left(\frac{1}{T'} - \frac{1}{T_1}\right) \text{。}$$

並藉由 Haldane 關係式計算逆向酵素活性：

$$E_{a8f} = \frac{E_{a8r} K_{13BPG} K_{ADP} K_{eq}^8}{K_{3PG} K_{ATP}}$$

然後將動力學參數、正逆酵素活性和 37°C 下的反應平衡常數代入速率反應方程式：

$$Flux_{PGK} = \frac{E_{a8f} \left(\frac{X_{1,3BPG} X_{ADP}}{K_{13BPG} K_{ADP}} \right) - E_{a8r} \left(\frac{X_{3PG} X_{ATP}}{K_{3PG} K_{ATP}} \right)}{\left(1 + \frac{X_{1,3BPG}}{K_{13BPG}} + \frac{X_{ADP}}{K_{ADP}} + \frac{X_{1,3BPG} X_{ADP}}{K_{13BPG} K_{ADP}} + \frac{X_{3PG}}{K_{3PG}} + \frac{X_{ATP}}{K_{ATP}} + \frac{X_{3PG} X_{ATP}}{K_{3PG} K_{ATP}} \right)}$$

得到 PGK 的反應通量。

最後根據反應物所參予反應的反應通量計算反應物的生成及消耗速率：

$$\begin{aligned} \frac{dX_{1,3BPG}}{dt} &= Flux_{GAPDH} - Flux_{PGK} \\ \frac{dX_{3PG}}{dt} &= Flux_{PGK} - Flux_{PGM} \\ \frac{dX_{2PG}}{dt} &= Flux_{PGM} - Flux_{ENOL} \end{aligned}$$

當 16 個反應的反應通量全計算過一遍後，將與氫離子及鎂離子有參予反應的反應通量經由(11)式、(12)式計算這兩個離子的生成與消耗量計算 H^+ 與 Mg^{2+} 的濃度變化。

在染色體的編碼上，直接使用實數型編碼，動力學參數組 K 一共 70 個參數。在計算動力學模型時，須提供解題器個別反應物的 K_a 、該反應之 ΔH 、22 個反應物之起始濃度和 16 個酵素活性以計算(2)至(6)式，而在演算法中須提供所要估算之反應物濃度數據計算目標函數 f_E 。

六、 實驗結果

6.1 利用智慧型基因演算法作加權指數法

首先我們先使用單目標的智慧型基因演算法分別利用 f_G 及 f_E 對動力學參數作最佳化。實驗內容為當使用實驗所觀測的反應物濃度變化作誤差最小化，並假設所測得之濃度數據有無雜訊存在的問題作實驗驗證。在下列實驗所使用之模擬濃度數據是我們起初在驗證 f_G 值時的最佳組合解之濃度變化。

6.1.1 實驗只使用 f_G 模擬無雜訊之數據

圖 10 是 IGA 的評估函數只使用 f_G 值的實驗結果。因為在群落起始化放入 Vinnakota[44]的參數組當作種子，故我們在結果中標示出來。

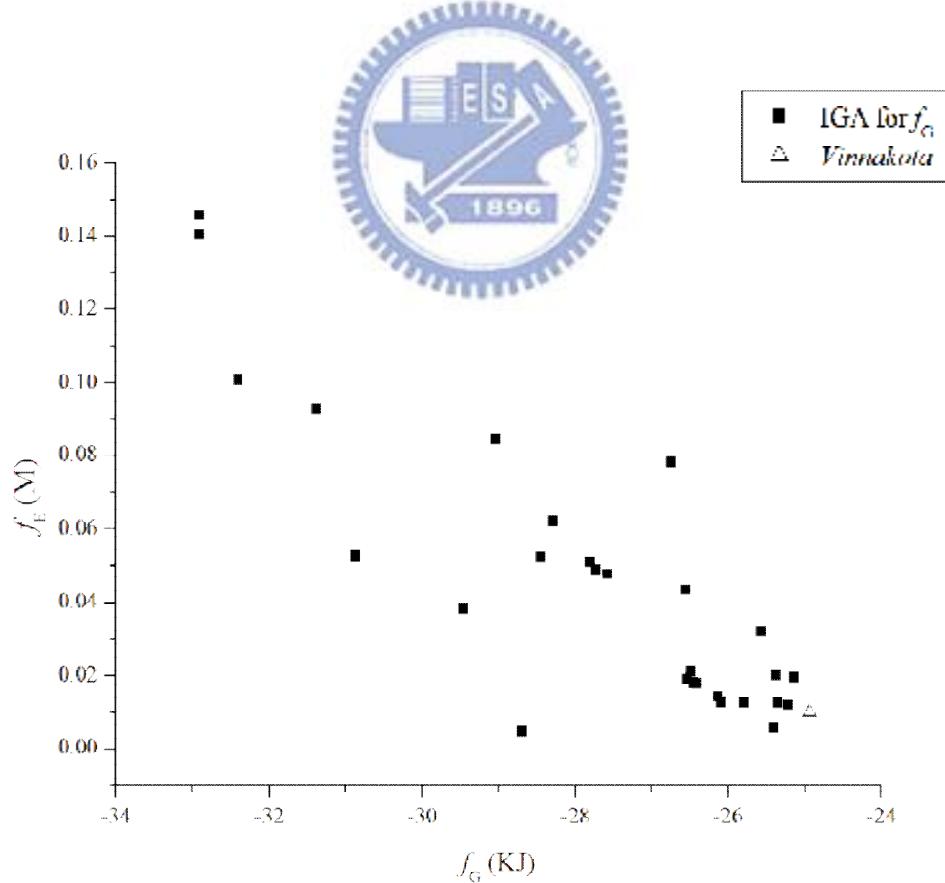


圖 10 IGA 針對 f_G 做最佳化之結果

在圖中 f_G 值的確是可以遠小於 Vinnakota[44]，但是當 $f_G < -26$ 之後，與濃度曲線之誤差就會快速上升。因為能量的計算會跟濃度有極大相

關，所以動力學參數的組合會讓反應物濃度反應過量使 f_G 過小而不符合模擬實驗數據。另外，反應物濃度變化曲線與系統總自由能的最低點之間的對應關係並無對等，一組濃度變化曲線只會對應得到一個系統總自由能的曲線，但反之則不然。一個系統總自由能曲線可能對應無限多組的濃度變化曲線。

我們從上圖可發現單就 f_G 此一目標最佳化，分佈結果相當分散，很難確實地將反應過程具體重現出其正確的濃度變化曲線。但還是可利用此單一目標搜尋到左下角能支配 Vinnakota[44]的 K ，故此目標雖可利用效果卻不佳。

6.1.2 實驗只使用 f_E 模擬無雜訊之數據

圖 11 是使用 f_E 作單目標的實驗結果，我們並將所模擬之實際答案以及 Vinnakota[44]的位置表示在圖 11 上。我們共做了 30 組實驗，而 IGA 最終輸出的結果大多數確實集中在精確解附近。從圖上結果的表現可以呈現出當實驗數據無誤差存在的情形時，IGA 確實能針對模型準確性找尋最佳解的優秀能力。

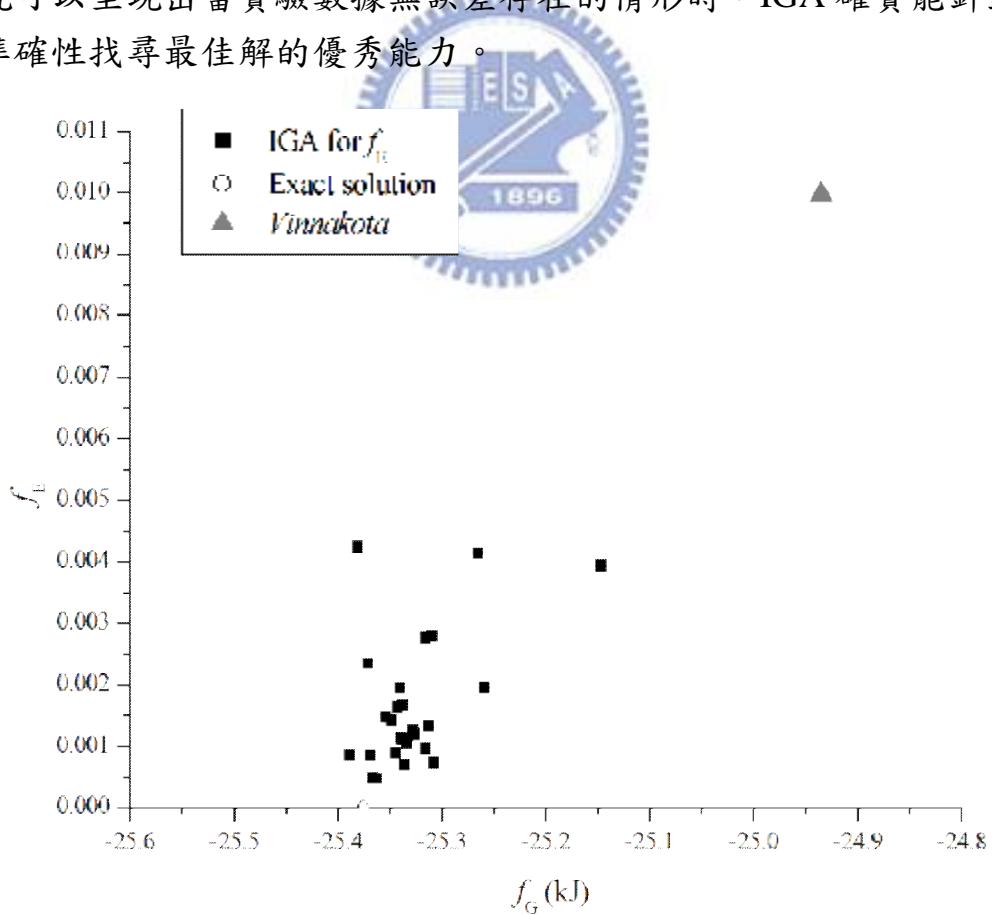


圖 11 IGA 針對 f_E 做最佳化之結果

但是在 IGA 演化曲線的最末端就出現演化之震盪現象，我們從 30

組實驗中挑選 8 組而將演化過程曲線之結果表示在圖 12 至圖 19。在圖 12 到圖 19 中 IGA 的演化軌跡全部都由右上方至左下方，顯示出 K 演化過程中 f_G 與 f_E 兩個評估函數是同時下降，而在 f_E 值在 0.001 ~ 0.002 範圍內開始出現震盪的現象。

震盪現象的出現則代表著兩項目標函數開始有互相衝突的性質存在，但是由於此實驗是對 f_E 值作最佳化，所以當震盪現象出現後不再是往左下搜尋而變成向正下方搜尋。

但圖 17 却也呈現震盪現象的軌跡，所以即使 f_E 值還不到 0.002 之前也是有可能發生互相衝突。但也有可能是此組不夠優異，參數的組合搭配不好才產生震盪現象。

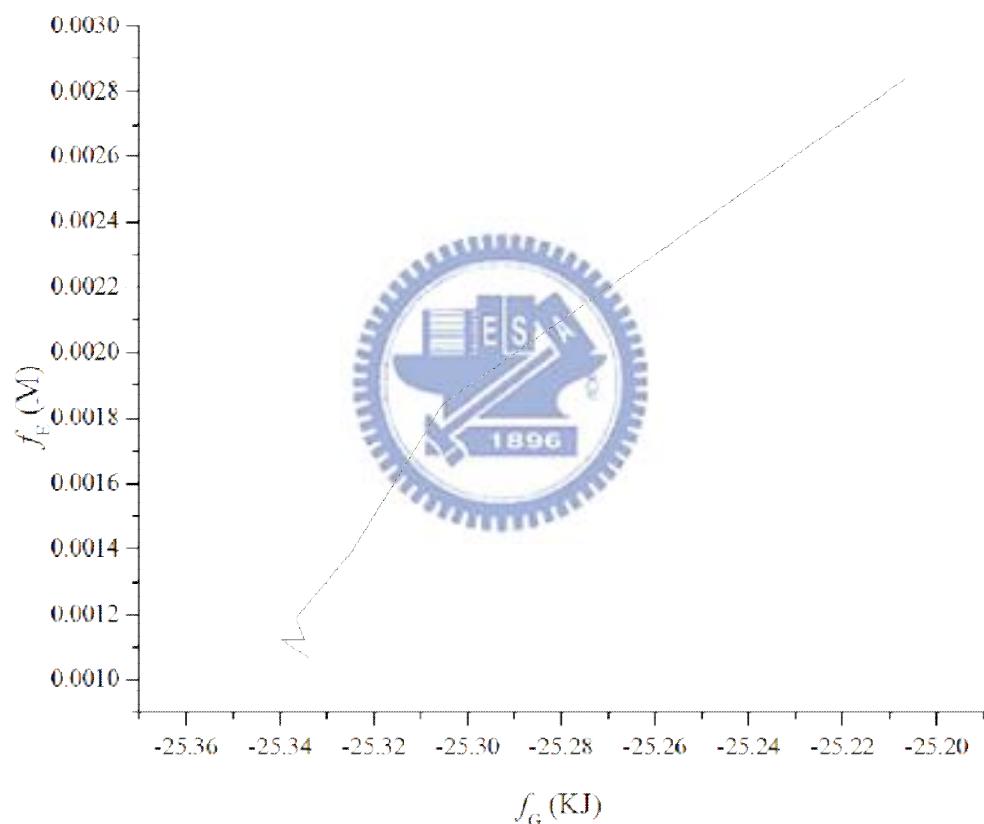


圖 12 IGA 演化曲線圖-實驗組 1

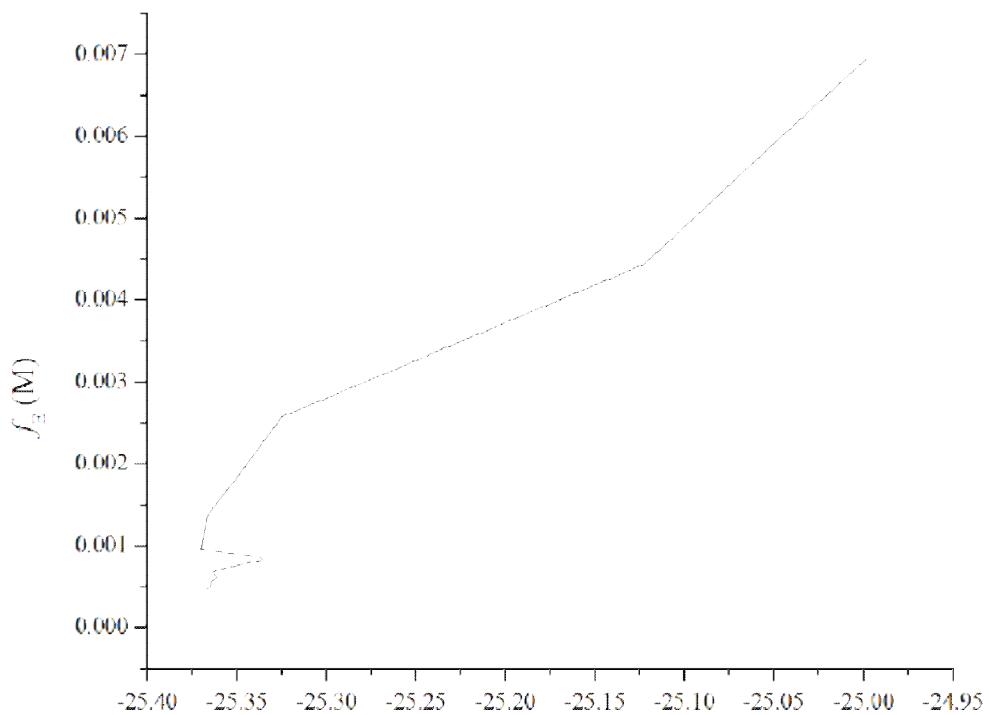


圖 13 IGA 演化曲線圖-實驗組 2

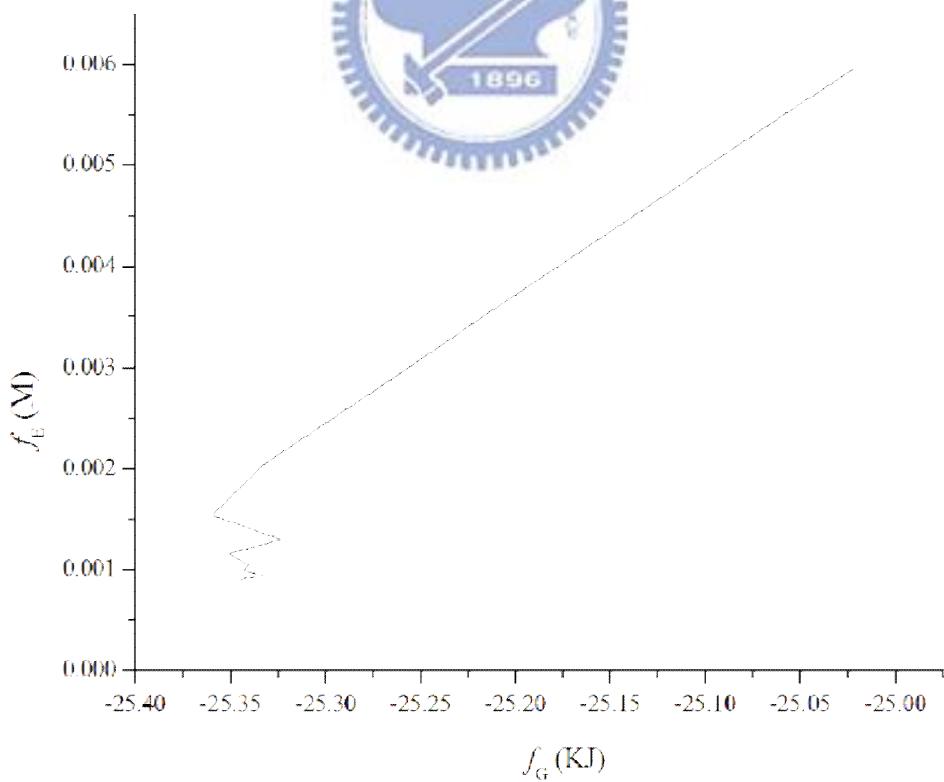


圖 14 IGA 演化曲線圖-實驗組 3

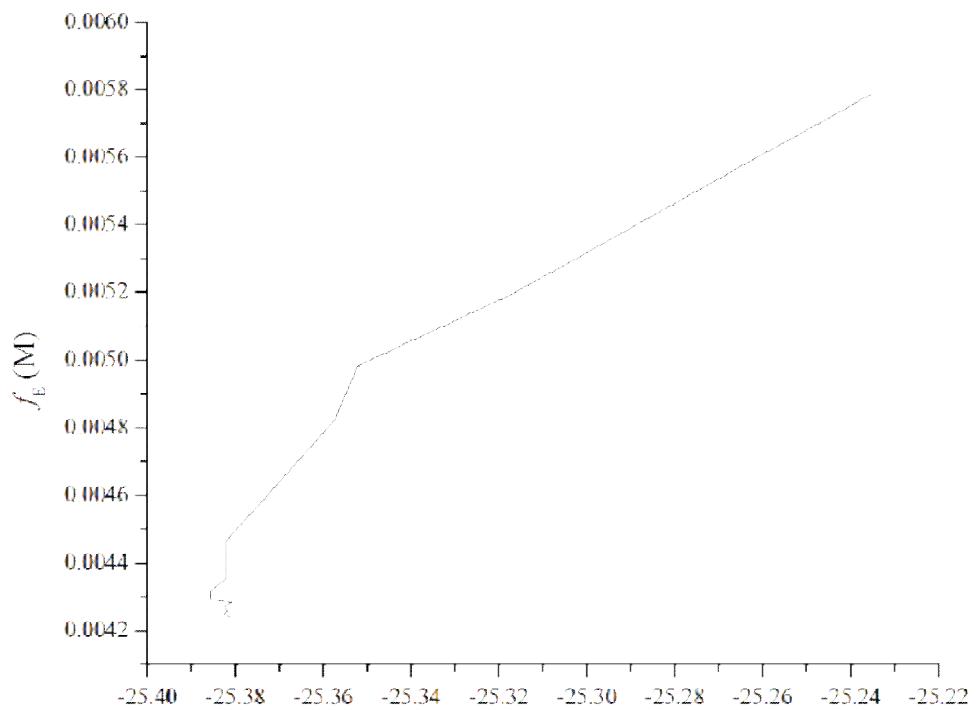


圖 15 IGA 演化曲線圖-實驗組 4

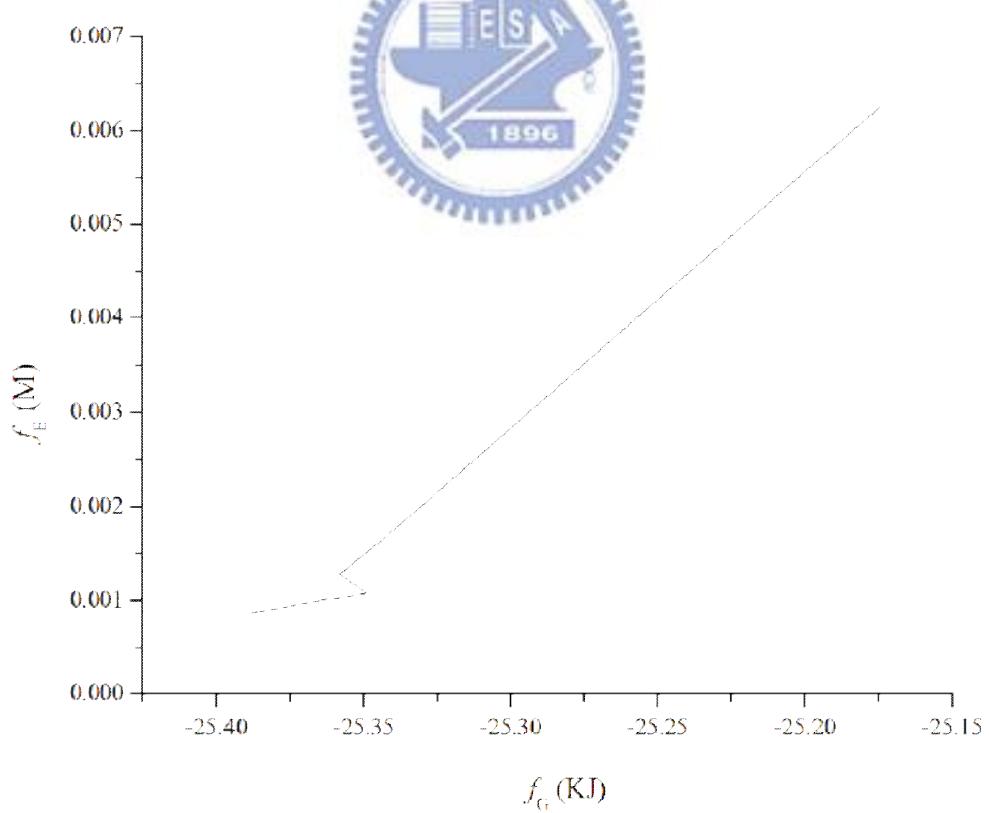


圖 16 IGA 演化曲線圖-實驗組 5

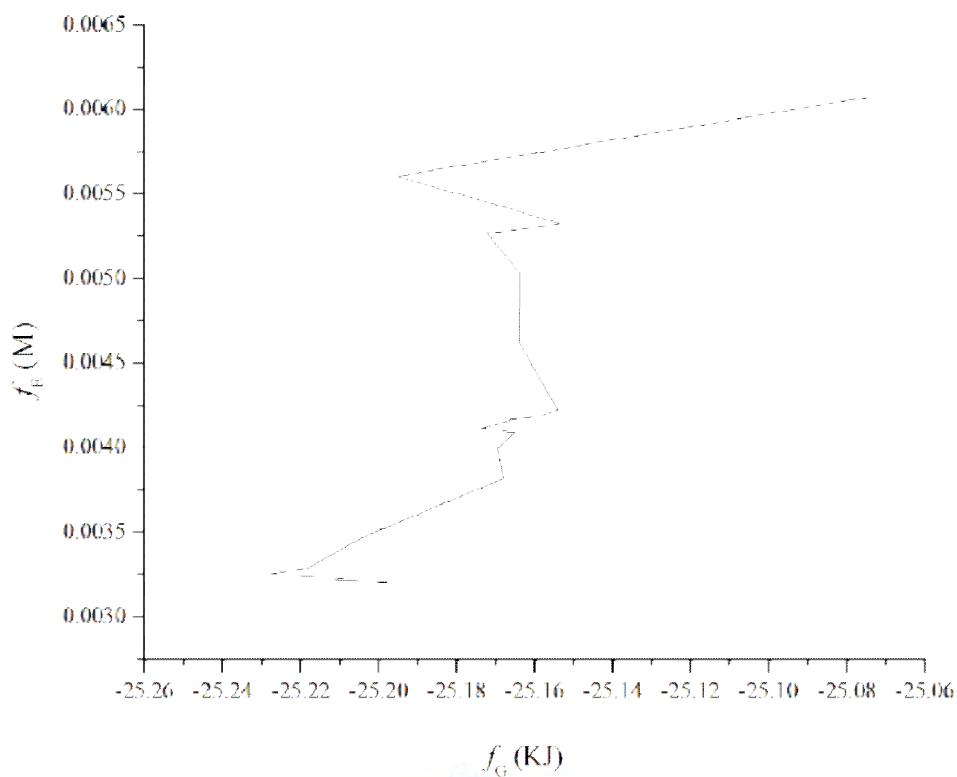


圖 17 IGA 演化曲線圖-實驗組 6

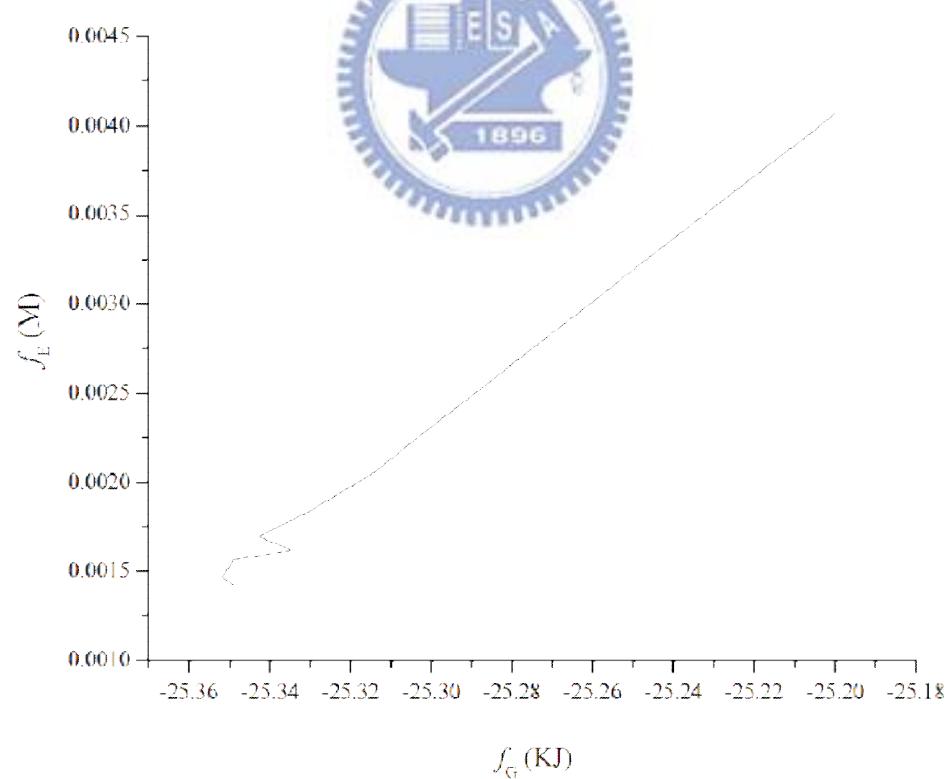


圖 18 IGA 演化曲線圖-實驗組 7

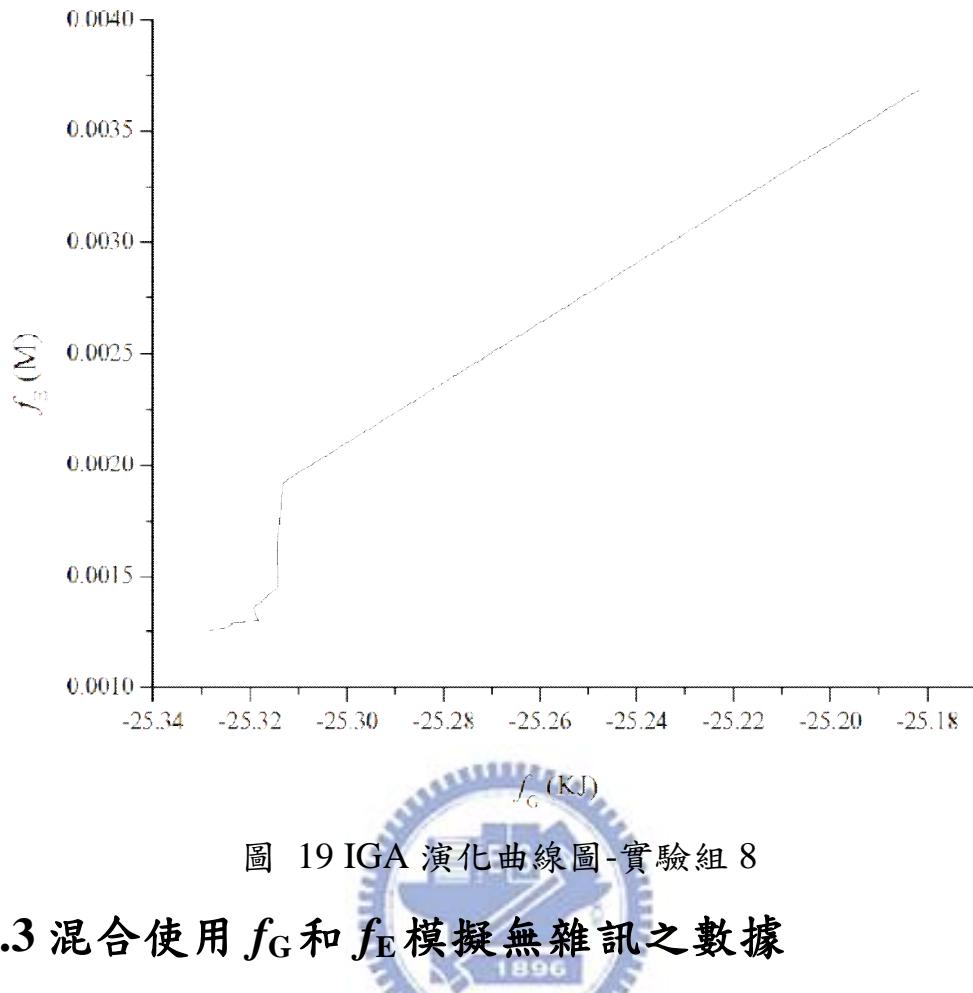


圖 19 IGA 演化曲線圖-實驗組 8

6.1.3 混合使用 f_G 和 f_E 模擬無雜訊之數據

我們將圖 10 和圖 11 合併在圖 20，顯示兩項單獨使用而分別得到的實驗結果，而單獨使用 f_E 的 30 組實驗比單獨使用 f_G 集中。會造成兩組解差異甚大的原因在於 f_G 值最小化所得的數值分布太廣， f_G 從 -25 ~ -34 都會存在最佳解，而 f_E 只會存在於 0 ~ 0.16 之間，所以兩個目標函數之間具有尺度差距的問題存在。

另一點可從圖 20 觀察出來的性質是所有最佳解分佈的結果具有一下凹曲線的邊界。雖然實驗組別不多而不夠明顯能看出，但是根據此圖解集合之分佈以及 6.1.2 演化曲線的震盪現象顯示出存在有兩目標函數結合之需要。

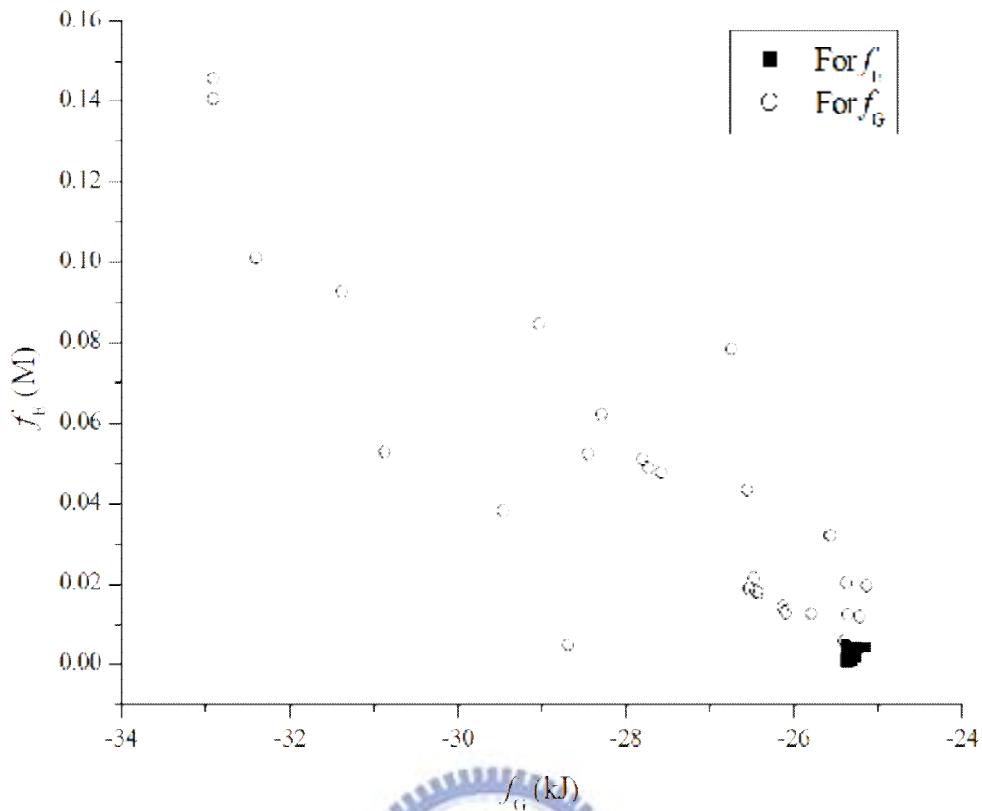


圖 20 結合兩者結果

6.1.4 實驗混合使用 f_G 和 f_E 並使用加權指數

有鑑於 6.1.2 以及 6.1.3 的結果以及發生原因，在此小節我們將使用加權指數 w 將兩項目標函數混合成為單目標 IGA 的新目標函數，如下式所示：

$$\text{Fitness Function} = f_G + w \times f_E \quad (80)$$

我們將實驗結果表示在圖 21， w 的範圍為 1.0 至 9.0。實驗結果顯示一旦當 $w > 2.0$ 後，使用 IGA 的結果會得到類似於單只使用 f_E 的結果趨勢，集中在 $(f_G, f_E) = (-25.35, 0.001)$ 的周圍。

但這樣混合兩項目標函數並沒有解決尺度大小不同的問題，而其搜尋出的最佳解跟只使用 f_E 相似並未進一步改善實驗結果，所以我們將在 6.2 引入智慧型多目標基因演算法(IMOGA)解決尺度大小不同的問題。

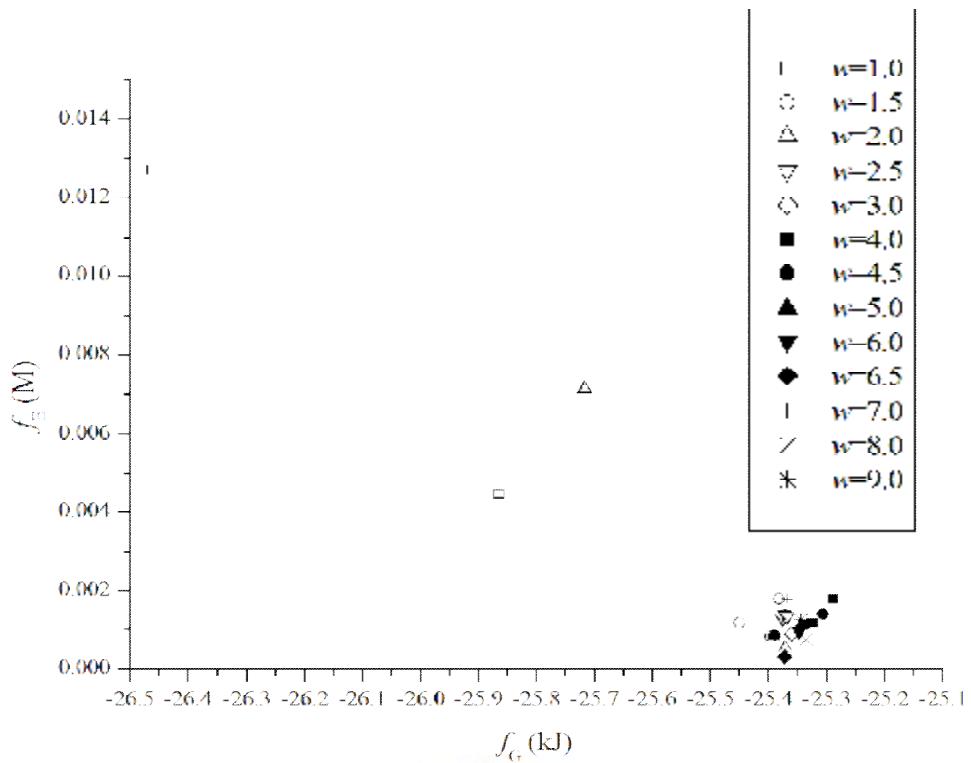


圖 21 混合兩目標函數之實驗結果

6.1.5 實驗只使用 f_E 模擬有雜訊之數據

從圖 22 中的實驗結果顯示 IGA 只使用 f_E 對無法有效的將模型參數最佳化，並在實驗中絕大部分都落在 $f_E = 0.065$ 附近而無法超過 $f_E = 0.005$ 的界線，我們估計 0.005 很有可能是一個界限值。但是實驗結果之群落分散成兩堆，使得結果不甚穩定集中。

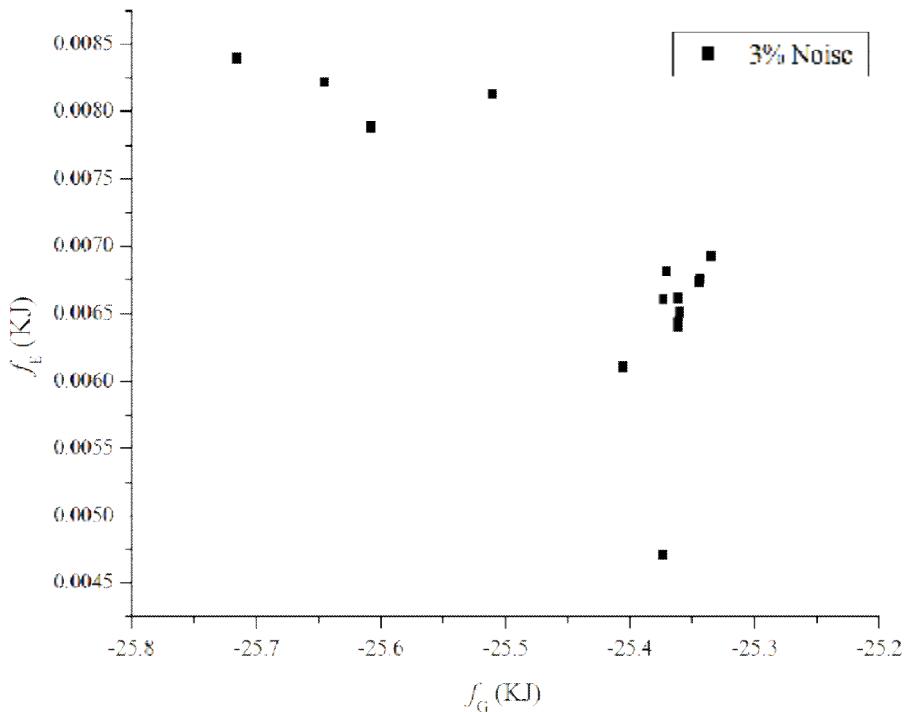


圖 22 IGA 針對 f_E 及 3% 雜訊

在圖 23 至圖 27 的結果顯示出 IGA 最小化的過程中小到 $f_E = 0.0065$ 附近就會出現震盪的現象，如同實驗 6.1.2 的演化曲線末端所出現情況極為相似。證明不論雜訊是否存在於，兩者都會產生震盪的情形，所以兩項目標衝突是一個這問題本身就存在的特性之一。

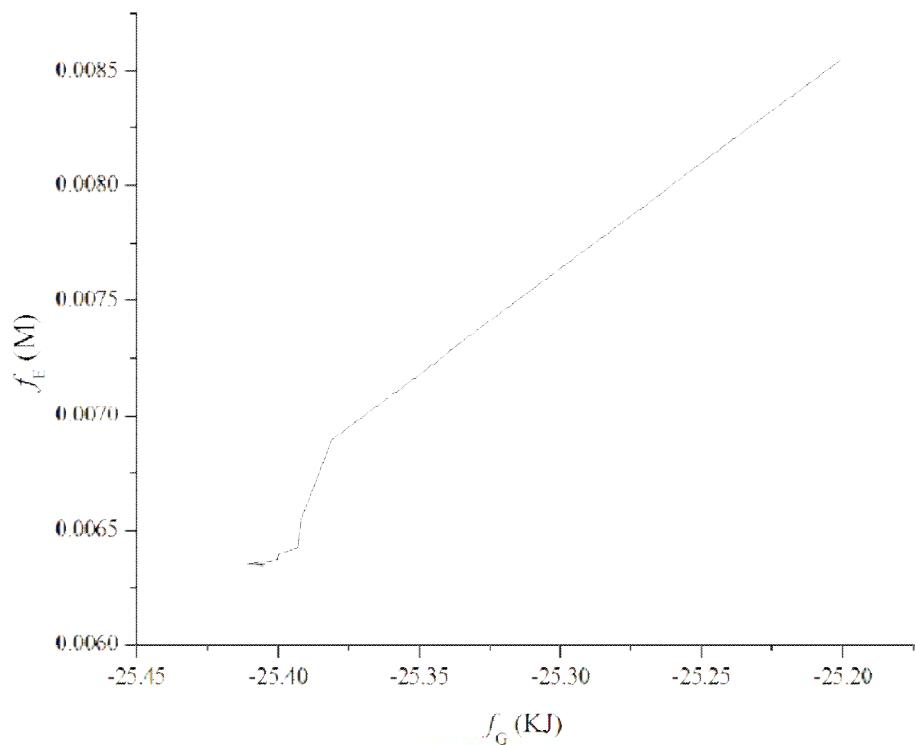


圖 23 IGA 具雜訊之演化曲線圖-實驗組 1

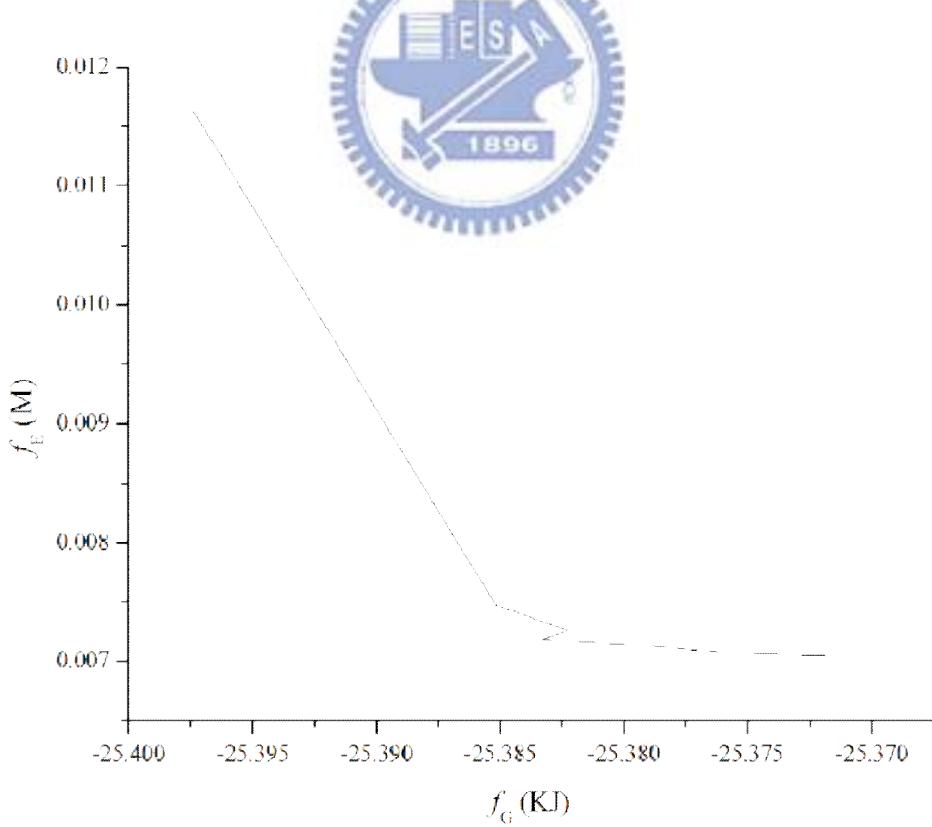


圖 24 IGA 具雜訊之演化曲線圖-實驗組 2

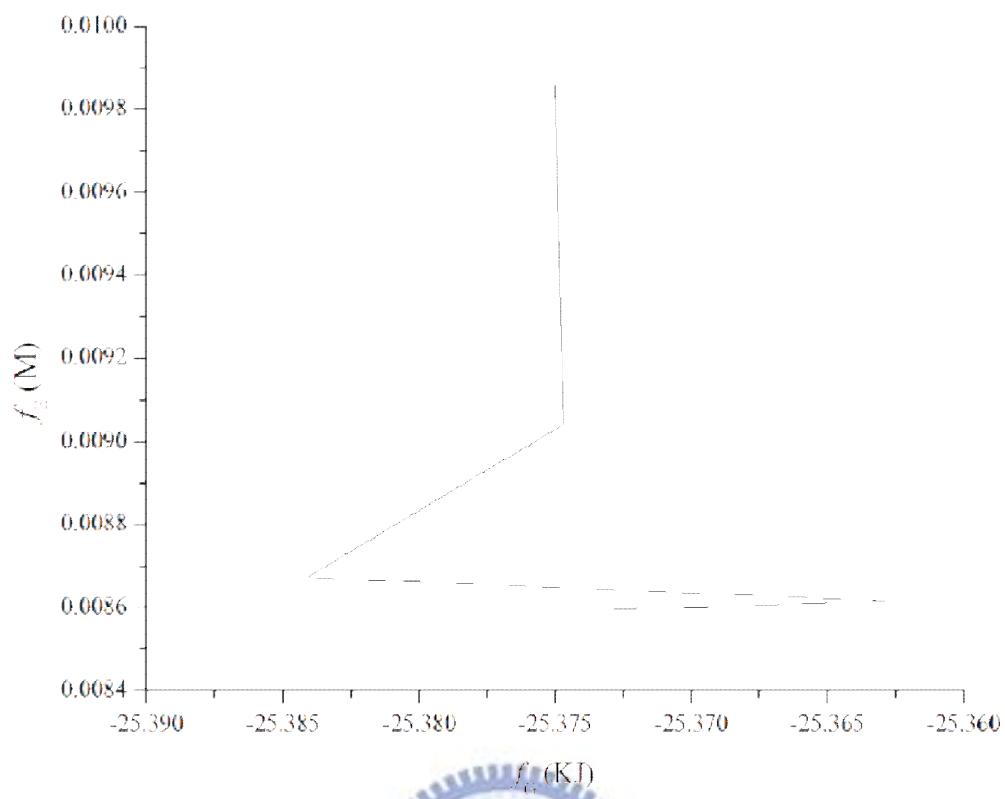


圖 25 IGA 具雜訊之演化曲線圖-實驗組 3

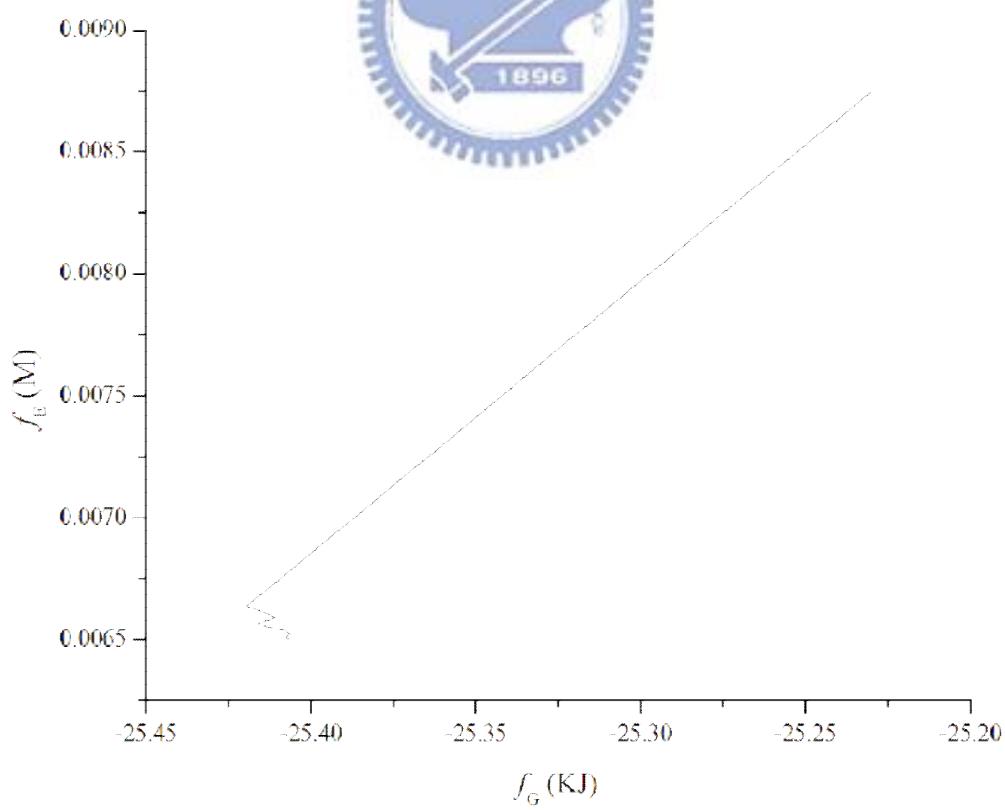


圖 26 IGA 具雜訊之演化曲線圖-實驗組 4

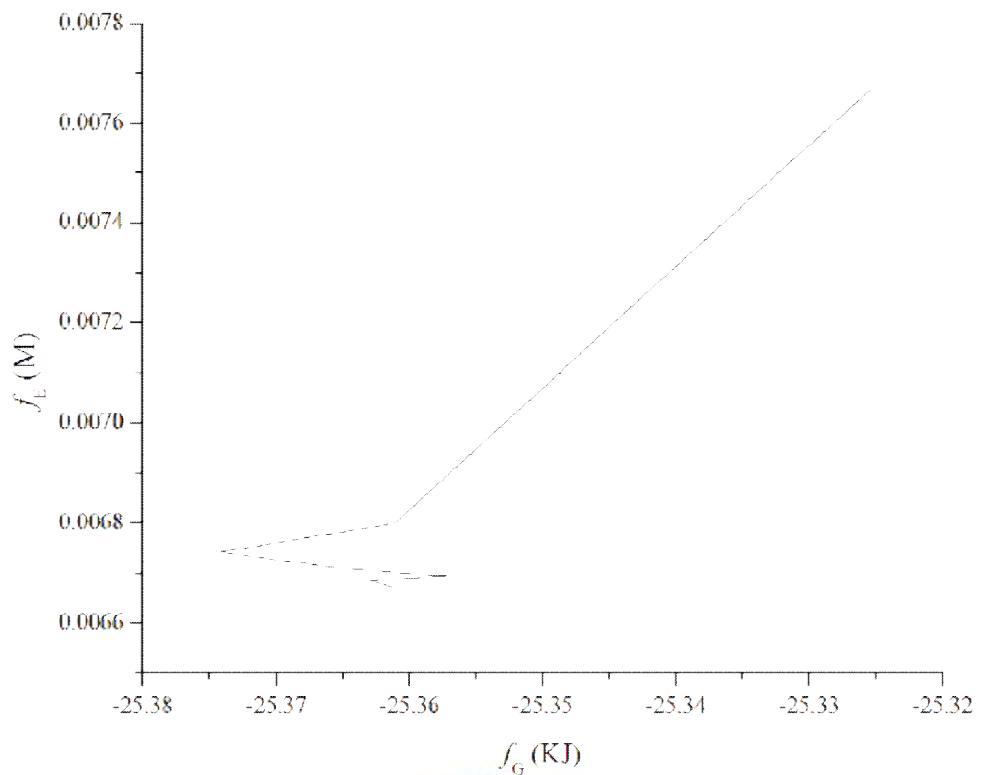


圖 27 IGA 具雜訊之演化曲線圖-實驗組 5

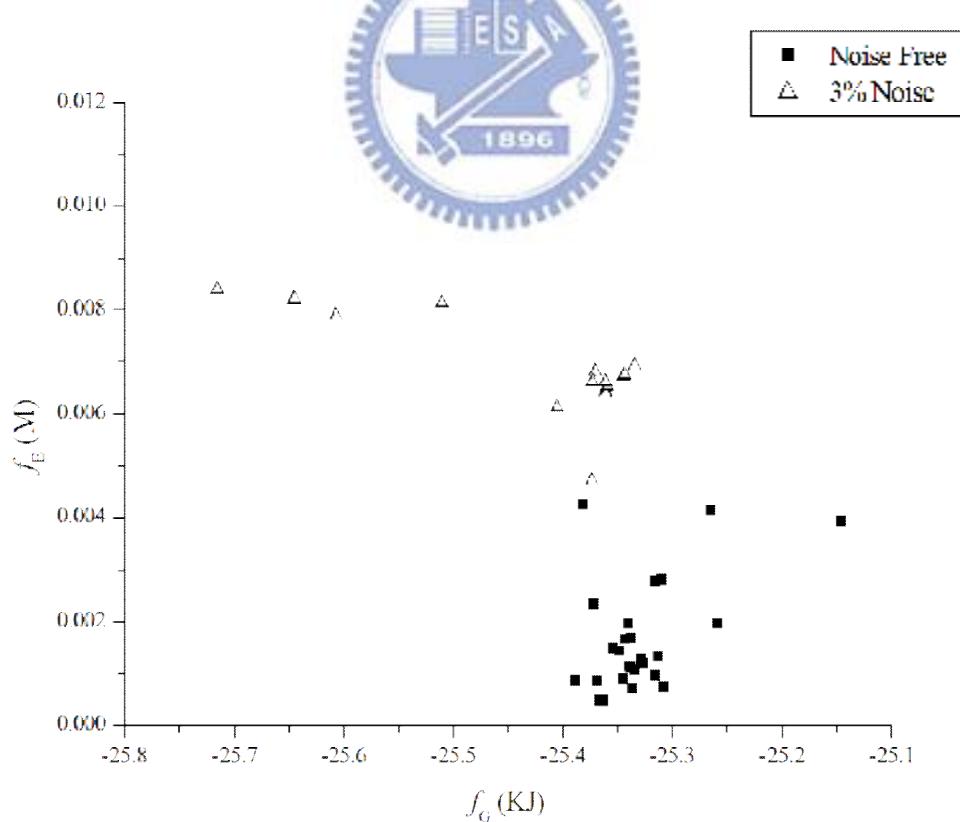


圖 28 雜訊存在與否之實驗示意圖

圖 28 中將雜訊有無的實驗結果放在一起，可從圖中結果之分佈得

到雜訊的存在確實可以幫助分辨 K 的好壞。當沒雜訊時，只追求縮小與實驗濃度數值的誤差就可使得模型的雙目標皆可達成。但當實驗濃度數據存在誤差時追求最小化的效果將會有較大的機率落在同一水平的 f_E 上，但是其 f_G 的差異性很大將不便於選擇 K 。所以經由圖 28 的顯示，我們認為在選擇最佳之動力學參數值的集合中是需要考慮使用雙目標的觀念去強化判定 K 的好壞。

我們將使用 IGA 所作的實驗經統計後把結果作成表 6。從表 6 中觀察實驗 2 與實驗 5，在實驗 2 的 f_E 之變異數雖然比實驗 5 大，但 f_G 之變異數卻遠小於實驗 5。這結果代表著當實驗值具有雜訊時，除了建模的準確性會受到影響外，同時也會讓 f_G 的偏差增大。兩者變異數有極大差異，讓我們發現雜訊的存在對於 f_G 的影響更大。

進一步的從 Variance 單行觀察，實驗 1 > 實驗 4 > 實驗 2。雖然實驗 4 之數值仍然遠大於實驗 2，但在 Best 中卻能更優於實驗 2 的結果。所以我們基於此結果所呈現的數據資料，相信需要考量同時將 f_E 和 f_G 最小化才能解決並使兩項目標一起滿足。

表 6 IGA 全實驗之統計表

		Mean		Variance		Best	
		f_G	f_E	f_G	f_E	f_G	f_E
Noise Free	only f_G	-27.7006	0.044302	5.472078	0.001417	-32.908	0.140509
	only f_E	-25.3285	0.001668	0.002318	1.19E-06	-25.3627	0.00047
	Hybrid	-25.4364	0.001956	0.058782	6.66E-06	-25.3717	0.000302
3% Noise	only f_E	-25.4311	0.006882	0.015707	9.07E-07	-25.3733	0.004709

6.2 用智慧型多目標基因演算法取得最佳解集合

6.2 的實驗是將最佳化演算法由 IGA 置換成 Intelligent Multi-Objective Genetic Algorithm (IMOGA) 同時對 f_G 與 f_E 作最佳化，由於兩者存在互相衝突的特性勢必會呈現 Pareto Front 的曲線存在，而 IMOGA 的目的即是將這曲線上的最佳解集合求出並呈現 Pareto-front。

根據 IMOGA 中 GPSIFF 的特性是不受到多項目標函數尺度差異而求出 Pareto Front，可用來解決上述所提出之問題：目標函數尺度不同之問題。

在我們的問題中兩項目標函數皆是求最小化，所以我們所得之曲線皆呈現向下凹陷的趨勢。由於 IMOOGA 作出之 Pareto Front 範圍太廣超出我們所需要的部份，故有關 Pareto Front 之實驗圖我們只取 $f_E \leq 0.02$ 的範圍作圖。

6.2.1 實驗模擬無雜訊之數據

圖 29 為此實驗之結果，共有 96 個點。因為 IMOOGA 在搜尋能力上並沒有像 IGA 針對單一方向之最佳解作深度搜索，所以相較之下所形成的 Pareto Front 就有可能被 IGA 所突破，但理論上若演化代數目設定夠長的時間，則 IMOOGA 就可能搜尋出 IGA 的解。圖 30 是 IMOOGA 與 IGA 兩相比較以證實結果，IGA 比之 IMOOGA 在單一方向搜尋具有非常優異的能力。

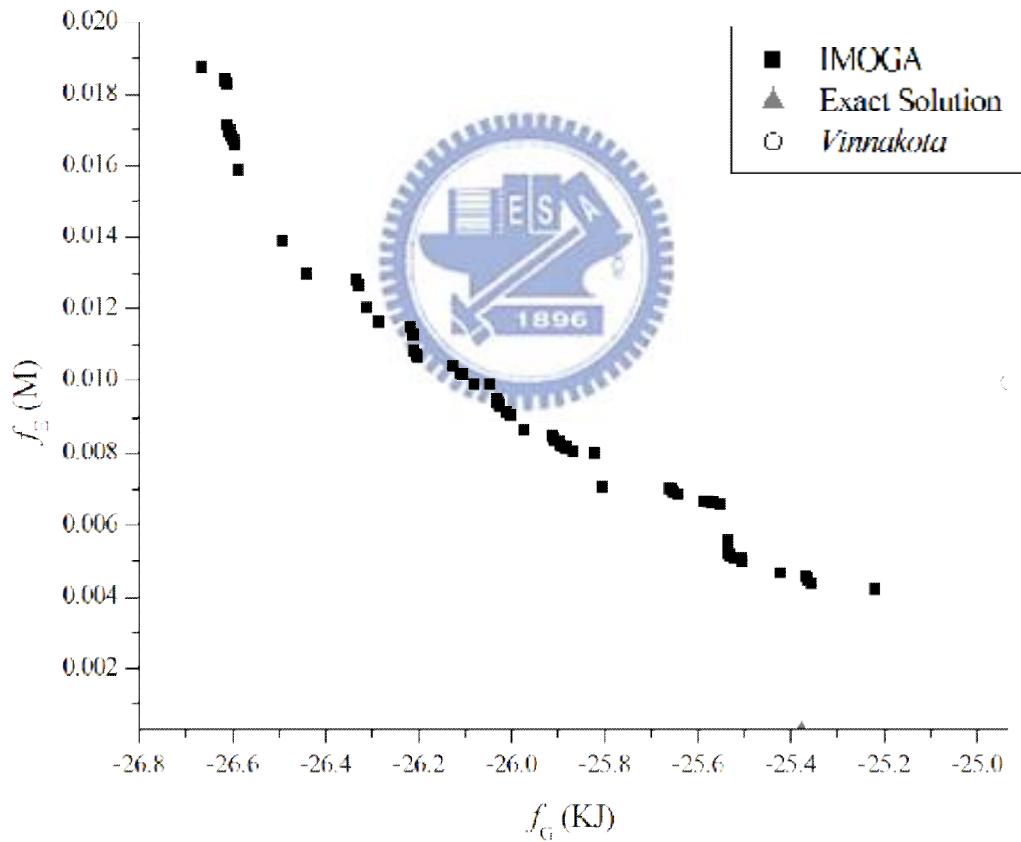


圖 29 IMOOGA 對模擬實驗數據之結果

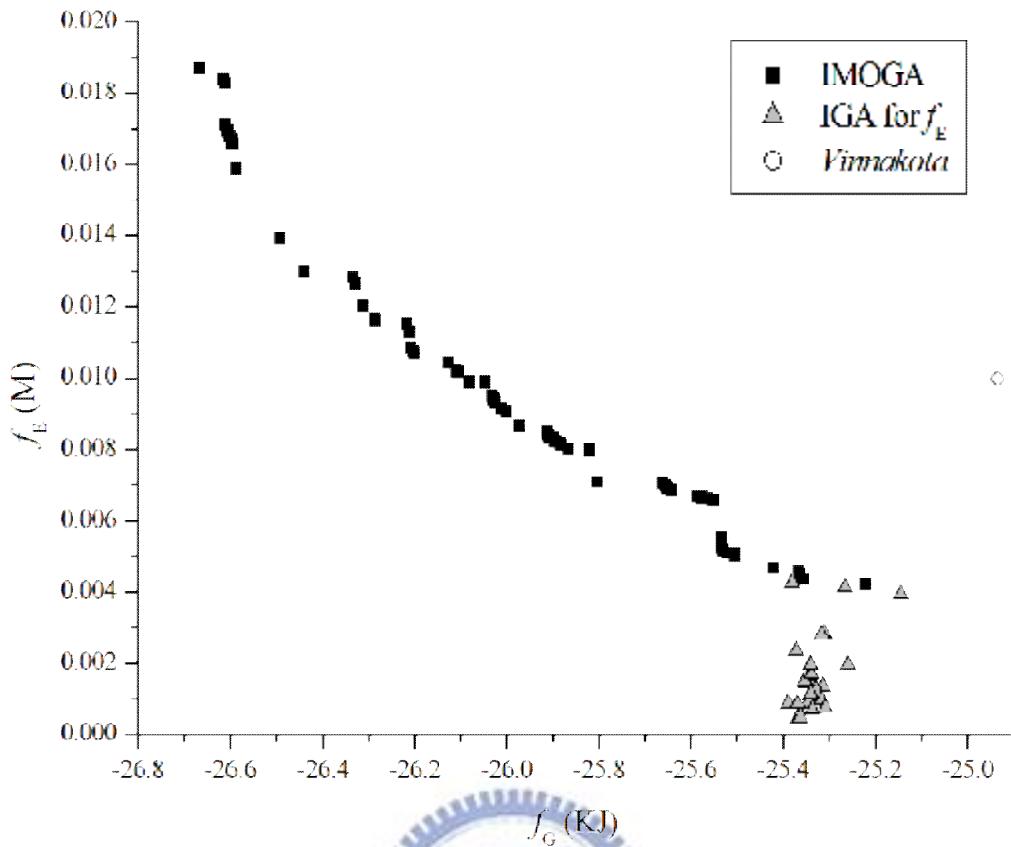


圖 30 IMOGA 與 IGA 比較圖

由於 IGA 只對單一目標作搜尋而將準確度提升，我們可從實驗結果證實。但是對於需要同時考量兩項目標函數而使用權重值的方法，是不能適宜地解決兩者尺度不同的問題，我們已在實驗 6.1.4 提出證明。而 IMOGA 所搜尋出的解不像 IGA 能在深度上能作更有效搜尋，但卻能提供同時考量兩項目標所組合出的最佳解。

相對於 Vinnakota[44]解，IMOGA 證實能有效的提升準確率並同時考量熱力學的要求，在建模所用之參數 Vinnakota[44]解不能有效建出貼近真實之實驗結果。

6.2.2 實驗模擬有雜訊之數據

在圖 31 和圖 32 的結果顯示在有雜訊的狀況下，IMOGA 在 f_E 方面所能下降的程度有限，情況如同有雜訊之 IGA 一樣。圖 33 中我們將 IMOGA 與 IGA 之結果結合，圖上顯示 IGA 的大多數實驗會聚集在 IMOGA 的 Pareto Front 的延伸線之上。但從實驗 6.1.5 中及其統計結果，只有考量準確度為目標將使得另一目標的變異數變大，而改用 IMOGA 就可從 Pareto 解集之中提供一組雙目標兼具考量之參數。

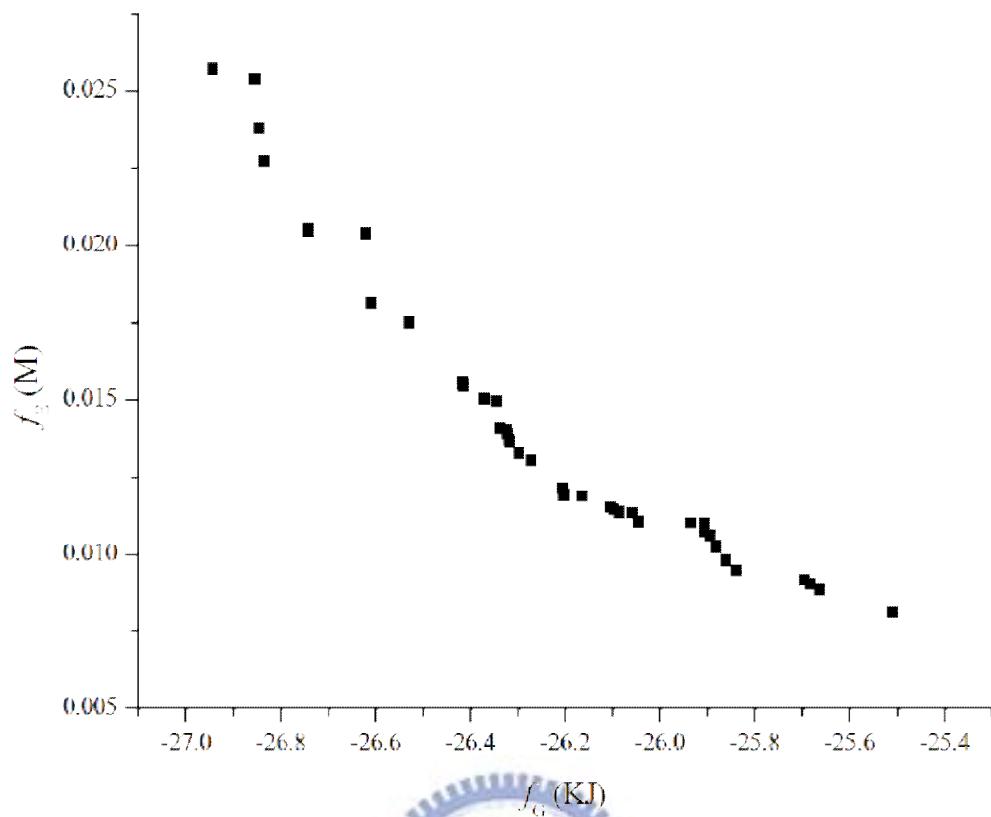


圖 31 IMOGA 對具雜訊之模擬數據結果-實驗組 1

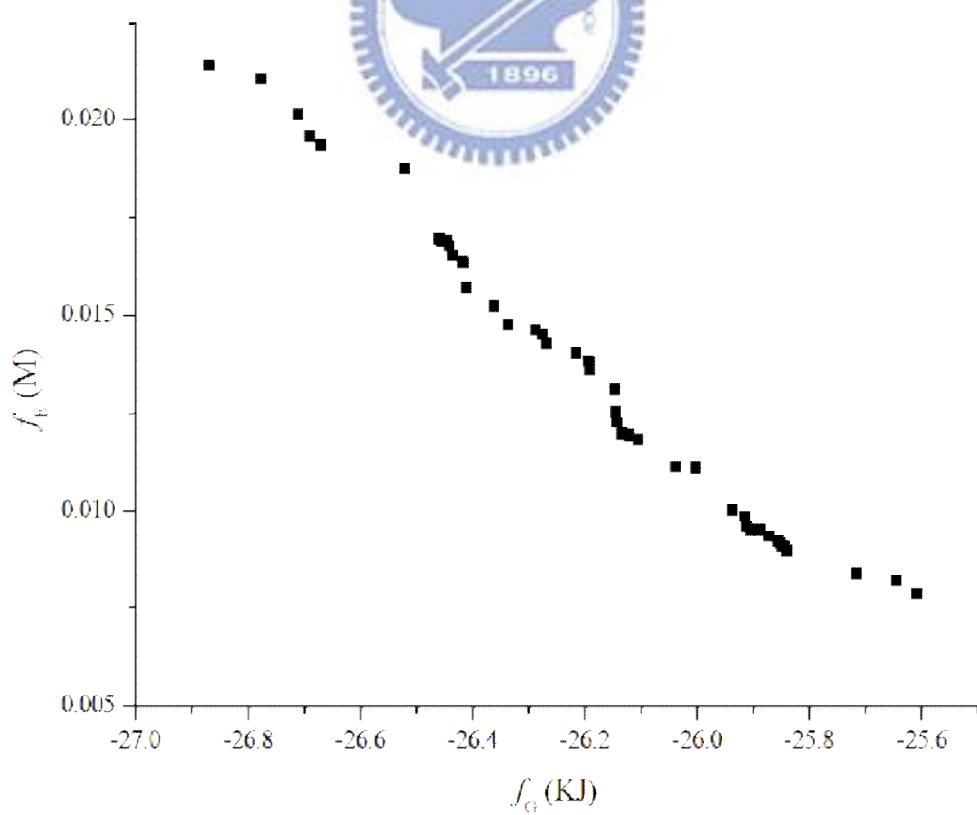


圖 32 IMOGA 對具雜訊之模擬實驗數據圖-實驗組 2

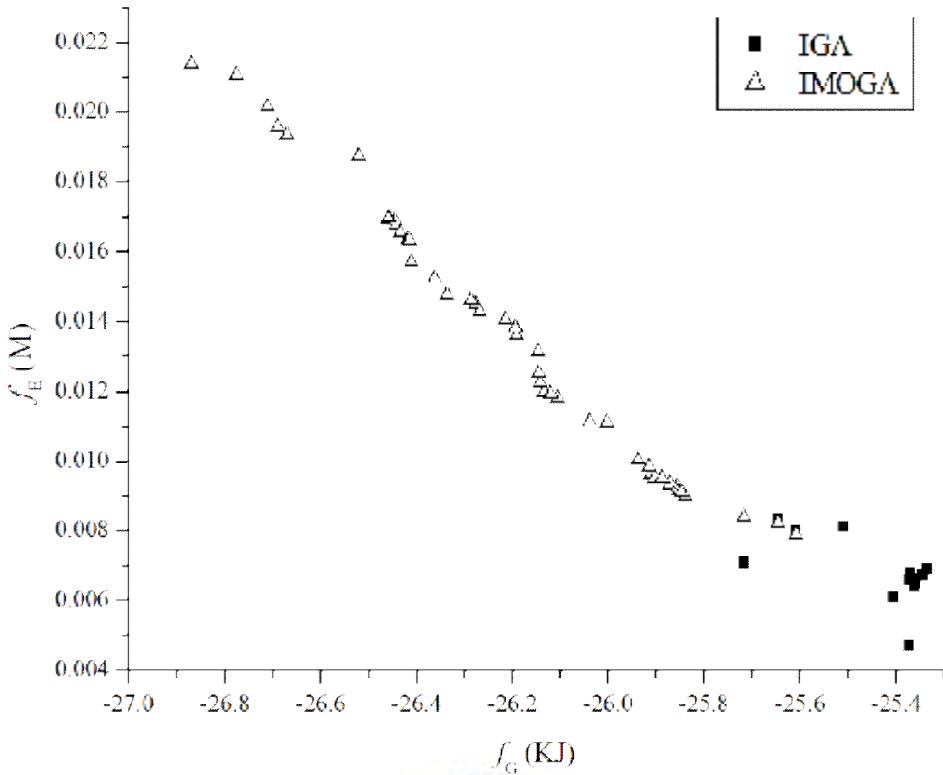


圖 33 當雜訊存在，IGA 與 IMOGA 之比較圖

6.2.3 實驗中使用真實實驗數據

最後我們 IMOGA 針對真實實驗數據作最佳化，結果即為圖 34。所比較之 K 為 Vinnakota[44]所提供之參數的結果確實顯示著被 IMOGA 的解所支配，效能雖不像之前實驗結果具有非常明顯之距離，但實際上依然能比 Vinnakota 解提升 9.20% 的準確率，而系統總能量所下降的幅度更是明顯差距有 10 倍之多，系統總能量的下降幅度可看圖 35。

另外我們還可能圖 34 上看到當將 Pareto Front 自 $f_E = 0.0275$ 分成兩段時， $f_E > 0.0275$ 的部分斜率遠比 $f_E < 0.0275$ 之部分陡峭。顯示出一但 $f_G < -26.4$ 之後，對於模型之誤差就會急速上升，而若取這部分的解之建模時 G 會出現與 Vinnakota[44]完全不同的圖形，其系統總能量之最低點會隨著時間不斷下降，而且濃度變化圖將會完全不同。因為系統總能量是跟量有關的一個系統上之特性(extensive)，當直接搜尋能量最低點所存在之參數解的結果即是反應物之消耗量大量減少，系統總自由能自然會呈現大幅下降的現象，但這樣的結果是忽略生物系統體內生理平衡機制的結果，所以在考慮能量最低點時可能會有一個可信賴的判斷值存在，而本實驗裡 $f_G = -26.4\text{KJ}$ 應是代表此一物理量的數值。

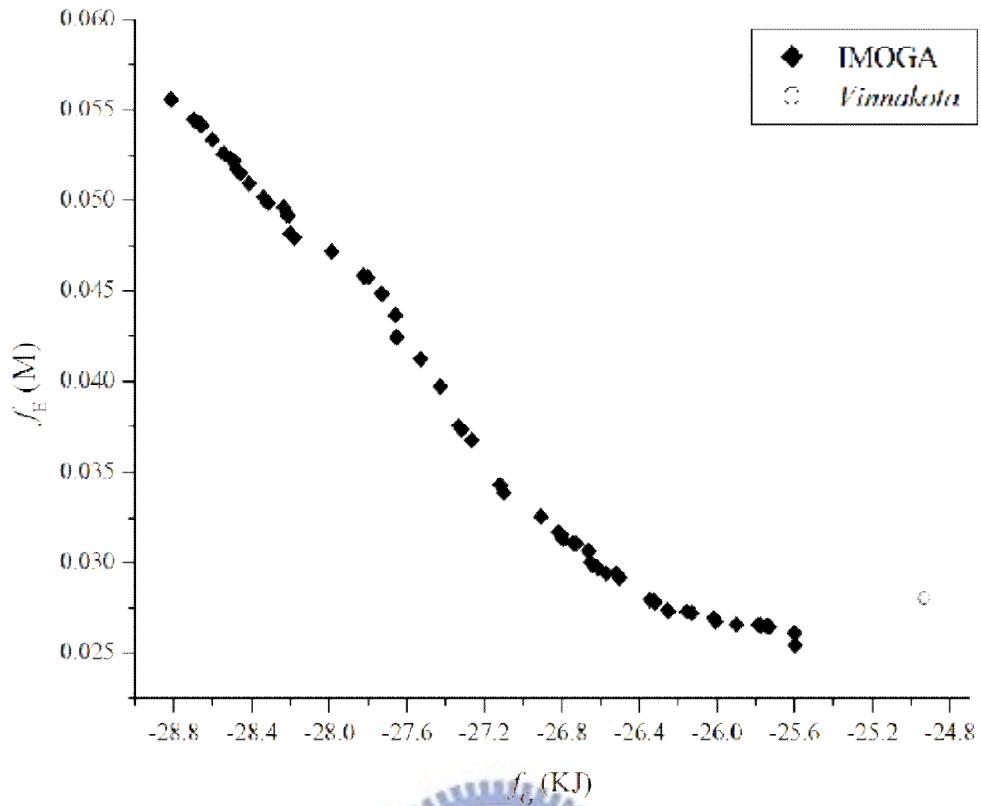


圖 34 IMOGA 對真實實驗數據圖

6.3 與真實實驗數據之比較

我們將實驗中所用的兩個演算法(IGA，IMOGA)分別取出其最佳解之代表作濃度、反應自由能來比較。

6.3.1 驗證 f_G

G 是根據系統中反應不斷進行所估計算出的一種存在於整體系統的熱力學性質，根據反應過程的進行使得計算結果產生改變。由於在計算 G 的定義是使用 μ_i 作為每個反應物的能量，但實際上的 μ_i 值難以得到，所以借用定義以 $\Delta_f G_i^o$ 代換 μ_i 計算 G 。而在 profile generator 中所計算的 1800 秒裡，每一點代謝物濃度跟 G 都在當下是屬於一個平衡狀態下所計算得到的結果。這是因為隨著反應的進行，每個時間點下的環境參數，例如 pH、 Mg^{2+} 的濃度都會改變。其中尤其以氫離子濃度的變化會不斷影響每個反應的平衡常數值，這個論點可由 Vinnakota[44]的實驗結果得到證明。故可說整個反應的平衡點不斷地在變化，但在每個時間點上卻又是一個暫時的平衡狀態。

再者，在計算式(42)中使用反應程度 ζ_m 來運算，由於 ζ_m 是表達

第 m 個反應進行程度的變數，它也可表示系統的組成及恆溫下之 G [60]。在反應開始時 ζ_m 等於零值其值將不斷增加到反應結束。在圖 35 中的 G 值曲線是將(42)式之模擬計算的結果直接呈現出來。

反應初期系統總自由能急速下降是符合熱力學的描述：反應會朝向減少總能量的方向前進，並且在熱力學原則上系統總能量會自然地走向總能量較低的系統狀態。系統總能量會自然地走向總能量較低的系統狀態是反應進行的驅動力，所以 f_G 設定的意義就是認為當 G 越低，就會越接近真實結果。

在圖 35 中，因為剛開始的反應商數皆遠小於反應平衡常數，故反應非常迅速；當反應商數逼近反應常數後，反應速度開始減緩，而反應速度之轉折點就是在 f_G 的位置。通過 f_G 點後 G 值就開始上升，原因在於當反應減緩後 ATP 又開始產生不再繼續消耗，所以能量又開始儲存。但 GLY 經過反應消耗後並未再生成，且生成物隨著反應進行增加也是令能量上升的原因之一。

根據熱力學的定義，系統的熱力學性質可以表示出反應進行的傾向。所以若是 f_G 值愈低，則代表著指反應愈容易朝此路徑進行，而認為是實際反應所遵循之方向，在圖 35 裡呈現實驗過程中確實有著上述現象發生。

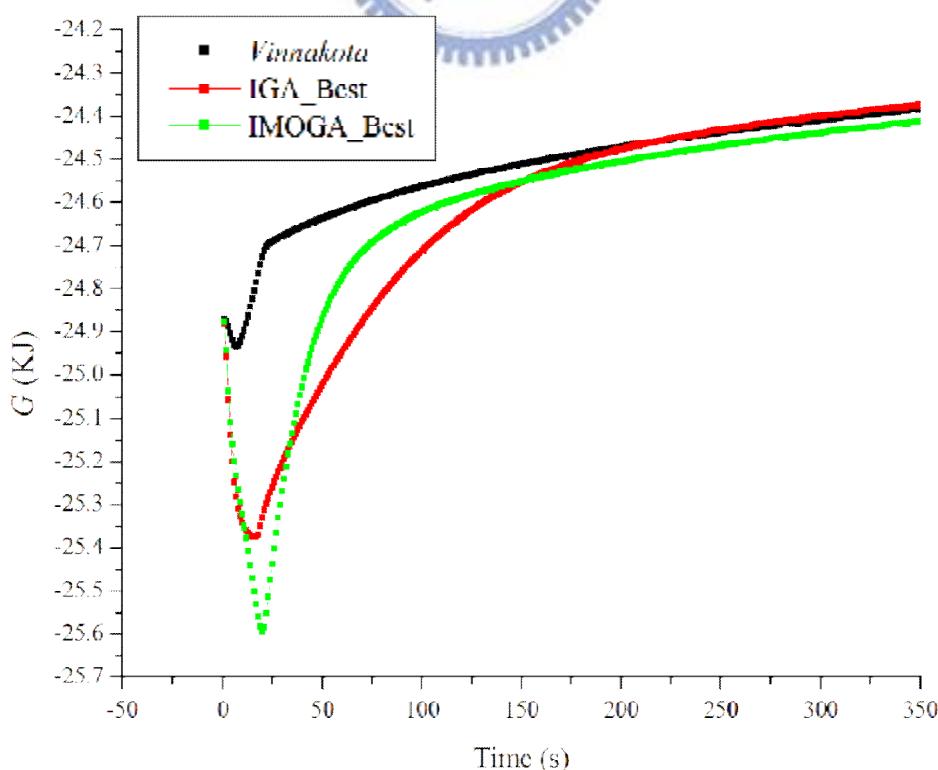


圖 35 三組 K 在模擬反應過程中系統總能量之變化

上圖是使用在我們最佳化過程中，分別從 IGA 和 IMOOGA 的實驗裡挑出最好的一組 K ，從三組 K 所建出的模型在整個反應過程裡，證實了當模擬之代謝物濃度與真實實驗之濃度數據兩者之間的誤差下降，同時反應過程中的 f_G 值也會具有較低的數值。

圖 35 所使用的 *Vinnakota*[44]的動力學參數確實不如使用最佳化過的 K 所估算出的 f_G 值低，並且與另兩組所造成的差距相當明顯，其效果比我們當時所預期的更好。根據所顯示的實驗結果我們判斷若要建立一套使用動力學模型的生化反應網路系統，在最佳化之過程中可以考慮使用 f_G 值當作是一個有效的評估函數。接下去的實驗結果會呈現同時利用 f_G 值及 f_E 值作完最佳化後的反應物濃度變化的過程。

6.3.2 濃度曲線圖

我們將 *Vinnakota*[44]的濃度變化分別從圖 36 至圖 39 將 IMOOGA 及 IGA 之結果表示出來，有 ATP、ADP、LAC、PCr、Hexose monophosphate(HMP)、G3P 和 Phosphoglycerates(PG)，共七種反應物作代表。雖然前篇研究有 fructose diphosphate(FDP)，但其數據資料只有前 20 秒故捨去其表示。經過最佳化過的參數我們全部放在表 7，下面模型所作出的圖皆是根據所列出之參數所建。

在圖 36 中，雖然兩個反應物在三組參數不同的情況下所建的曲線極相似，即使我們可從表 7 看到 IGA 所使用之參數有著極大的差異卻不足以影響 ATP 與 ADP 之變化。

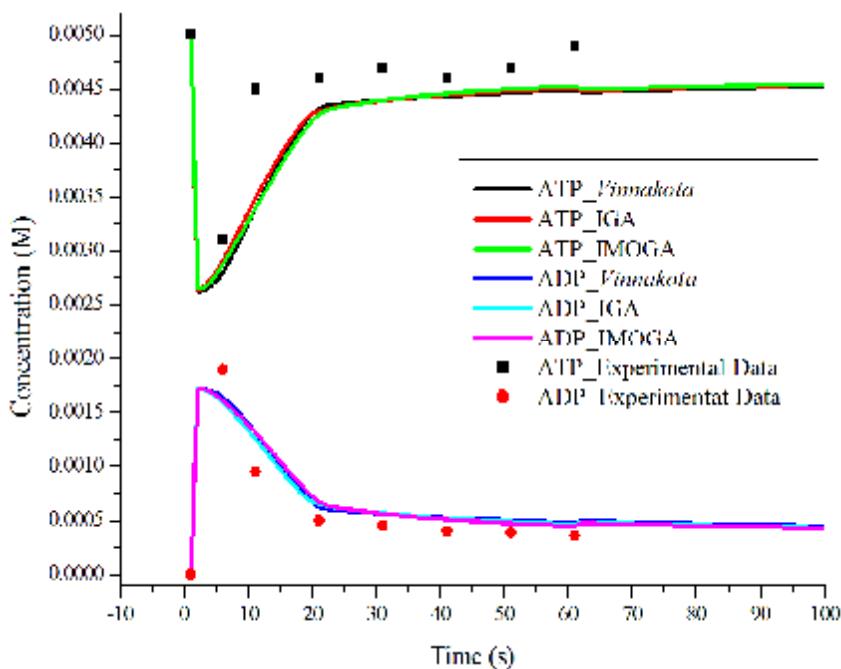


圖 36 ATP 及 ADP 之濃度變化

但在圖 37 即開始有些許的差異，可以看到濃度曲線明顯的區分成兩條，而 IGA 與 IMOGA 的趨勢一致，並且在反應初期的變化上比 Vinnakota[44]更接近真實實驗數據。

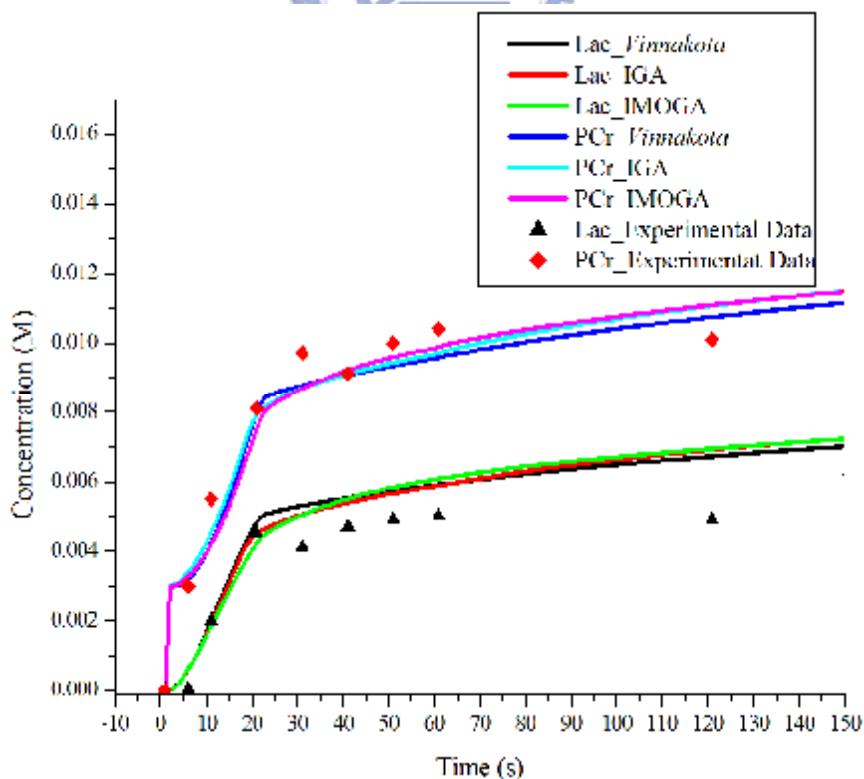


圖 37 LAC 及 PCr 之濃度變化

圖 38 的 HMP 是反應路徑中前段產物之總和，起初由肝醣大量分解成 G1P、磷酸化而後被快速消耗掉。圖中 Vinnakota 解不能適當地反應初期迅速的反應變化，而 IGA 所提供之參數解卻能將其變化顯現出來，但在 100 秒附近的反應趨勢卻不像其他兩者能有些許程度表示。不過整體看起來 IGA 已是最能呈現出 HMP 的反應過程。

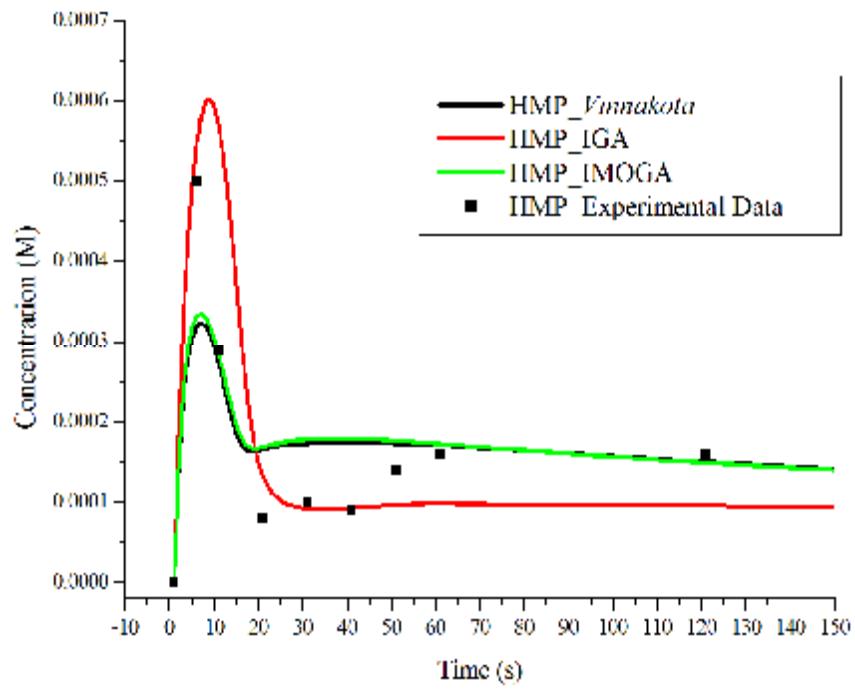


圖 38 HMP 之濃度變化

圖 39 是最能看出參數經過最佳化過後的結果。三組參數之中表現最好的是 IMOGA，而最差是 Vinnakota 解，從圖 38 及圖 39 可以發現 Vinnakota 解對於反應中迅速變化的部分無法有相對應的變化。IMOGA 的變化比 IGA 呈現出的曲線更接近真實結果，而反應過程劇烈的變化必定會造成系統總能量更明顯的改變並反應出來。根據以上結果我們如此推論，當反應過程之中含有反應物濃度劇烈變化者其總能量變化就更明顯，而自發反應必定會造成總能量下降，所以 f_G 想必會隨著 f_E 的下降而下降。這項推論即是實驗 6.3.1 的想法，我們從結果反推也是如此。故 f_G 即是一項有效的目標函數，能使我們演算法有效率地建構更準確之代謝反應網路模型。

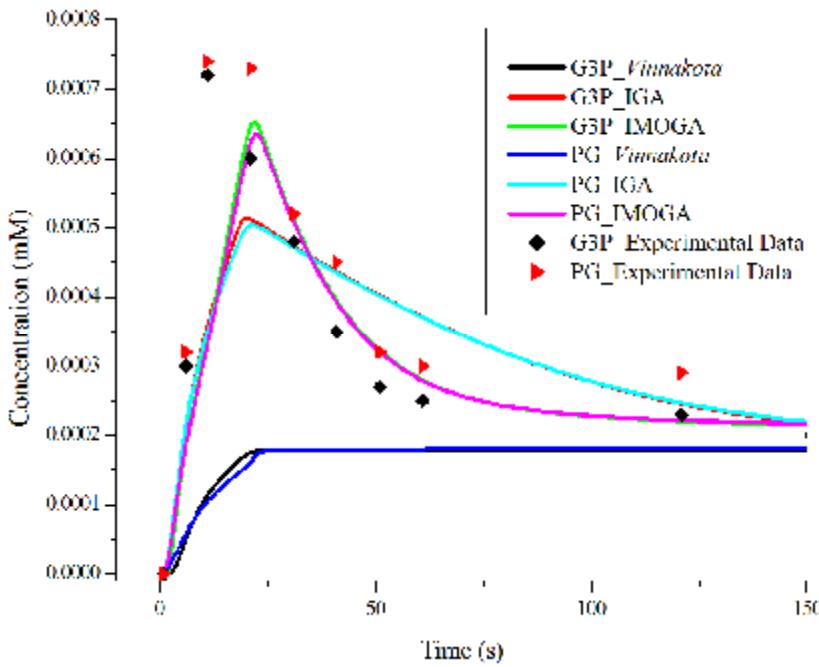


圖 39 G3P 及 PG 之濃度變化

6.3.3 動力學參數

本小節將 5.3.1 所使用的參數全部列表在表 7，而第六和七行是與 K_{iD} 之差異百分比。 K_{iIMOGA} 跟 K_{iD} 的參數在 GP、PGLM、PGM、PFK 和 ALD 中差異不大($< 10^{-4}$)，但從 TPI 開始變動幅度即產生極大的變動。我們認為在醣原分解的反應路徑中，由於 ALD、TPI 和 GAPDH 造成一個反應上的迴圈，同時也因為這個迴圈使得模型的數學問題成為”剛性”方程式而加強問題的難解程度。

因為上述三個所使用的動力學參數經最佳化後就會改變接下去的反應進行，不是單純只有影響改變下一個反應。所以整體最佳化的效果就能在此看出端詳，我們藉由整體最佳化來調整參數使反應過程能接近真實數據。從實驗 6.3.1 的結果得證，當模型更接近真實實驗時，其動力學參數會與文獻值之間的差異甚大，動輒倍增。可知單從文獻上所引用之參數實際上並不適用，更何況若環境不同時所造成的差異。所以當環境不同，起始條件的差異都會影響代謝網路之模型，而文獻所提供之 Vinnakota 解[44]並不一定通用在不同條件下。我們認為只要環境、起始條件不同，建模所需之參數就要重新尋找。而本篇提出一個同時考慮準確度與熱力學性質的演算法以提供他人使用我們結果，或套用我們方法而應用在其他代謝路徑上。

表 7 動力學參數表

Enzyme	Para-parameter	K_{iD} (mM)	K_{iIMOGA} (M)	K_{iIGA} (M)	$1 - \frac{K_{iIMOGA}}{K_{iD}}$	$1 - \frac{K_{iIGA}}{K_{iD}}$
Glycogen Phosphorylase A	K_{GLYf}	1.7	0.001700	0.00170659	-0.007%	0.388%
	K_{Pi}	4	0.004000	0.0109351	-0.007%	173.378%
	K_{iGLY}	2	0.002000	0.00151098	-0.007%	-24.451%
	K_{iPi}	4.7	0.004700	0.00557467	-0.007%	18.610%
	K_{GLYb}	0.15	0.000150	0.000151929	-0.007%	1.286%
	K_{G1P}	2.7	0.002700	0.00151353	-0.007%	-43.943%
	K_{iG1P}	10.1	0.010099	0.0114701	-0.007%	13.565%
Glycogen Phosphorylase B	K_{Pi}	0.2	0.000002	7.54E-05	-98.992%	-62.294%
	K_{iPi}	4.6	0.004600	0.0135604	-0.007%	194.791%
	K_{iGLYf}	15	0.014999	0.00285541	-0.007%	-80.964%
	K_{G1P}	1.5	0.001500	0.00244186	-0.007%	62.791%
	K_{iG1P}	7.4	0.007399	0.0178094	-0.007%	140.668%
	K_{iGLYb}	4.4	0.004400	0.0122687	-0.007%	178.834%
	K'_{AMP}	2.66E-03	0.002660	0.00255501	-0.007%	-3.967%
Phosphoglucomutase	nH	1.7505	1.750380	1.75016	-0.007%	-0.019%
	K_{G1P}	0.063	0.000063	7.30E-05	-0.007%	15.911%
Phosphoglucoisomerase	K_{G6P}	0.03	0.000030	2.56E-05	-0.007%	-14.724%
	K_{G6P}	0.48	0.000480	5.99E-05	-0.007%	-87.526%
Phosphofructokinase	K_{F6P}	0.18	0.000180	0.000275898	-0.007%	53.277%
	K'_{F6P}	20	0.019999	0.00912798	-0.007%	-54.360%
	K_{ATP}	0.08	0.000080	7.73E-05	-0.007%	-3.365%
	K'_{ATP}	0.25	0.000250	0.000243457	-0.007%	-2.617%
	K_{FBP}	4.02	0.004020	0.00456604	-0.007%	13.583%
	K'_{FBP}	4.02	0.004020	0.00159457	-0.007%	-60.334%
	K_{ADP}	2.7	0.002700	0.00295898	-0.007%	9.592%
	K'_{ADP}	2.7	0.002700	0.00297427	-0.007%	10.158%

	K_{iATP}	0.87	0.000870	0.000867673	-0.007%	-0.267%
	K_{aAMP}	0.06	0.000060	1.69E-05	-0.007%	-71.804%
	d	0.01	0.009999	0.0100237	-0.007%	0.237%
	e	0.01	0.000101	0.00150358	-0.007%	-84.964%
	Lo	13	12.999100	35.6249	-98.992%	174.038%
Aldolase	K_{FBP}	0.05	0.000050	5.78E-06	-0.007%	-88.447%
	K_{DHAP}	2	0.002000	0.00466114	-0.007%	133.057%
	K_{GAP}	1	0.001000	0.000646723	-0.007%	-35.328%
Triose Phosphate Isomerase	K_{GAP}	0.32	0.000003	0.000158169	-99.000%	-50.572%
	K_{DHAP}	0.61	0.000006	0.000195534	-98.992%	-67.945%
Glyceral-3-phosphate Dehydrogenase	K_{G3P}	0.18	0.000180	0.000527933	-0.007%	193.296%
	K_{NAD}	0.012	0.000012	1.38E-05	-0.007%	15.203%
	K_{DHAP}	0.22	0.000002	2.27E-05	-98.992%	-89.664%
	K_{NADH}	0.008	0.000001	1.41E-05	-85.736%	76.199%
Glyceraldehyde-3-Phosphate Dehydrogenase	K_{GAP}	0.025	0.000012	7.35E-06	-52.548%	-70.614%
	K_{NAD}	0.09	0.000001	2.66E-05	-98.992%	-70.463%
	K_{Pi}	0.29	0.000003	0.000835339	-98.992%	188.048%
	K_{13BPG}	0.0008	0.000004	7.75E-07	345.840%	-3.091%
	K_{NADH}	0.0033	0.000011	8.36E-06	243.809%	153.205%
Phosphoglycerate Kinase	K_{13BPG}	0.002	0.004044	0.00123141	202114%	61470.50%
	K_{ADP}	0.008	0.000035	1.31E-05	334.023%	63.490%
	K_{3PG}	1.2	0.003327	0.00061294	177.245%	-48.922%
	K_{ATP}	0.35	0.000004	0.000261185	-98.992%	-25.376%
Phosphoglycerate Mutase	K_{3PG}	0.2	0.000428	0.000581441	113.758%	190.721%
	K_{2PG}	0.014	0.000037	7.83E-06	164.026%	-44.081%

Enolase	K_{2PG}	0.1	0.000278	0.00028364	178.311%	183.640%
	K_{PEP}	0.37	0.001028	0.000844988	177.870%	128.375%
Pyruvate Kinase	K_{PEP}	0.08	0.000228	0.000138452	185.194%	73.065%
	K_{ADP}	0.3	0.000285	7.82E-05	-5.025%	-73.928%
	K_{PYR}	7.05	0.026351	0.0163444	273.779%	131.835%
	K_{ATP}	1.13	0.000654	0.00285131	-42.122%	152.328%
Lactate Dehydrogenase	K_{PYR}	0.335	0.000870	0.000996787	159.626%	197.548%
	K_{NADH}	0.002	0.000004	5.80E-06	94.174%	190.013%
	K_{LAC}	17	0.000171	0.027375	-98.992%	61.029%
	K_{NAD}	0.849	0.002798	0.000113903	229.539%	-86.584%
Creatine Kinase	K_{PCr}	1.11	0.004647	0.000194698	318.626%	-82.460%
	K_{iATP}	3.5	0.007122	0.0084113	103.493%	140.323%
	K_{iADP}	0.135	0.000001	0.000252576	-98.992%	87.093%
	K_{iPCr}	3.9	0.006561	0.00616382	68.239%	58.047%
	K_{Cr}	3.8	0.018852	0.00943007	396.108%	148.160%
Adenylate Kinase	K_{AMP}	0.32	0.000003	5.00E-05	-98.992%	-84.371%
	K_{ATP}	0.27	0.001159	0.000345612	329.422%	28.004%
	K_{ADP}	0.35	0.000832	0.000966257	137.773%	176.073%

6.4 敏感度分析

我們藉由改變反應物的起始濃度來對乳酸鹽的反應生成量測試模型的敏感度，計算式如下[55]：

$$Sensitivity_i^y = \frac{\partial \ln y(X_i, t)}{\partial \ln X_{io}} \quad . \quad (81)$$

乳酸鹽的流出量我們以 y 作表示， X_{io} 為反應物起始濃度。上式之物理意義為起始濃度影響產物流出量之回應。各別使用五個反應物作 30 次實驗。實驗結果用表 8 表示實驗後的統計數據，圖 40 及圖 41 是根據表 8 所做的圖。

表 8 敏感度分析的實驗統計結果

	Model		GLY	Pi	Cr	ATP	NAD
Perturbed by 1%	Vinnakota	Average	0.004238	0.082649	0.802315	0.213333	0.002285
		SD	0.000526	0.007135	0.011551	0.003378	0.003706
	IGA	Average	0.001200	0.046935	0.836579	0.214850	0.001411
		SD	0.001363	0.017387	0.007178	0.002298	0.002571
	IMOGA	Average	0.001421	0.071635	0.812530	0.246450	0.001211
		SD	0.001120	0.032187	0.006551	0.002128	0.002466
Perturbed by 3%	Vinnakota	Average	0.003964	0.084460	0.801409	0.210983	0.000441
		SD	0.003965	0.084503	0.801432	0.210898	0.000449
	IGA	Average	0.000551	0.040854	0.837407	0.218783	0.000421
		SD	0.000871	0.006988	0.003591	0.015397	0.000234
	IMOGA	Average	0.006510	0.071653	0.822070	0.228300	0.000844
		SD	0.001071	0.006988	0.003769	0.014597	0.000122

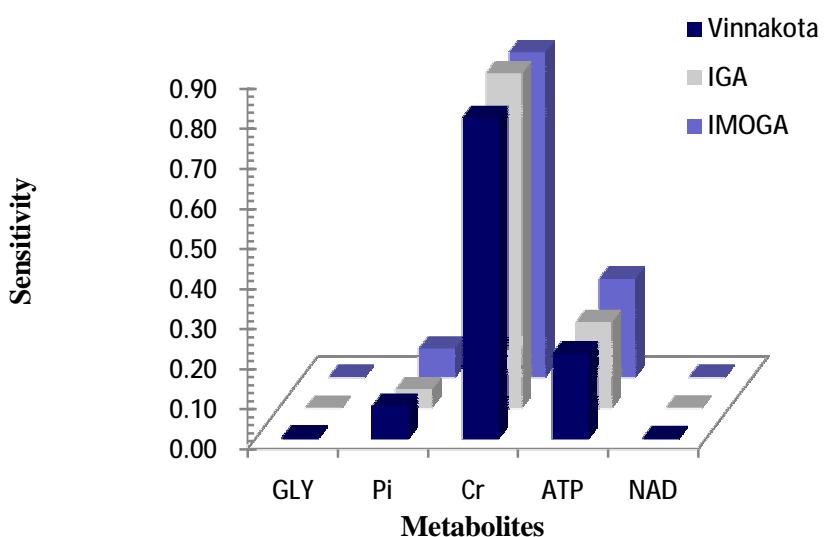


圖 40 最終產物之產生速率的敏感度對反應物濃度
變化 1%的結果

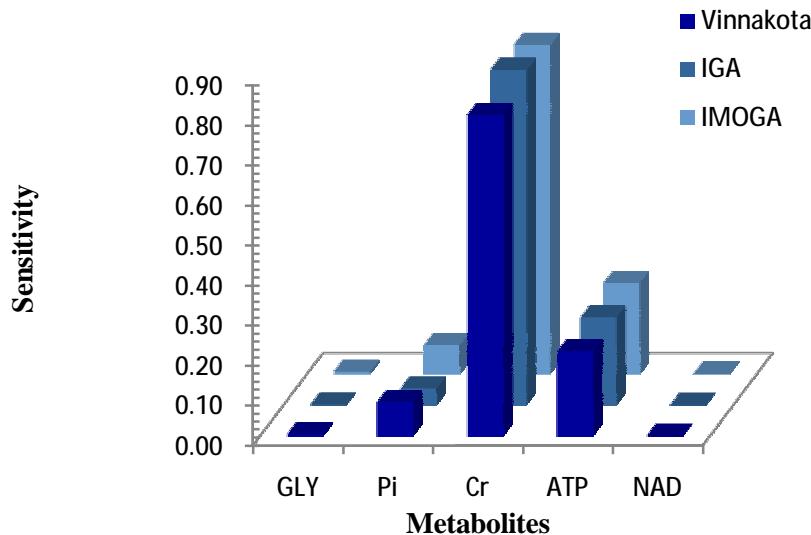


圖 41 最終產物之產生速率的敏感度對反應物濃度變化 3% 的結果

從圖 40 跟圖 41 的結果來看，Cr、ATP 及 Pi 明顯會使得乳酸鹽的生成量發生改變，特別是 Cr 最為明顯。但是 Cr 與 ATP 的敏感度在 IGA 或 IMOGA 却相距不大，而 Pi 是兩邊差別最顯著的。由於 IMOGA 能比 IGA 多考量熱力學的影響，故磷酸化對反應過程造成的影響能比 IGA 明顯，圖示結果與 Lambeth[55]相似，並且 Pi 的影響同時也符合 Richard *et al.* [69]的觀點而得到證實。

反應物 Cr 的濃度高低影響著 CK 之反應速率並正比於 ATP 的消耗速率，而使得乳酸鹽的產量增加。同樣地，ATP 本身之濃度高低也影響 ATP 的消耗速率，所以 Cr 及 ATP 是與 ATP 的消耗有直接影響，故得到高敏感度的結果。GLY 及 NAD 的濃度變化對產物乳酸鹽的反應量不具足夠影響力，無法改變反應之結果。

高敏感度在循環結構中是本來就會存在的特性，但現在若是 K 是由任意值所組成，其結果可能就會劣於不需輔酶的酵素。但是，特別高的敏感度卻可能是有數個狀況所構成，就像 Cr 的結果就可能是酵素活性，本身及 ATP 濃度高低、溫度或酸鹼度所共同造成的結果。

6.5 模擬結果的其餘反應物之濃度

最後我們將其他反應物的濃度變化結果呈現出來。在大部分的反應物中都有一個現象，就是反應物濃度最後會回到本來濃度。這樣

的結果符合所謂生理穩定的狀態，還原至原本濃度等待下一次反應的發生。另外大多數屬於被 IGA 或是 IMOGA 最佳化過的這些反應物的濃度變化過程皆呈現比較劇烈的濃度變化之情形。從 6.4 的結果可以得到最容易影響到最終產物的反應物依序是 Cr、ATP 和 Pi。

圖 42 是 Cr 濃度變化。Cr 是反應過程中濃度變化最大的少數反應物，並且又會跟 ATP 產生反應，所以對於最終產物的影響會是最敏感的結果相當合理。

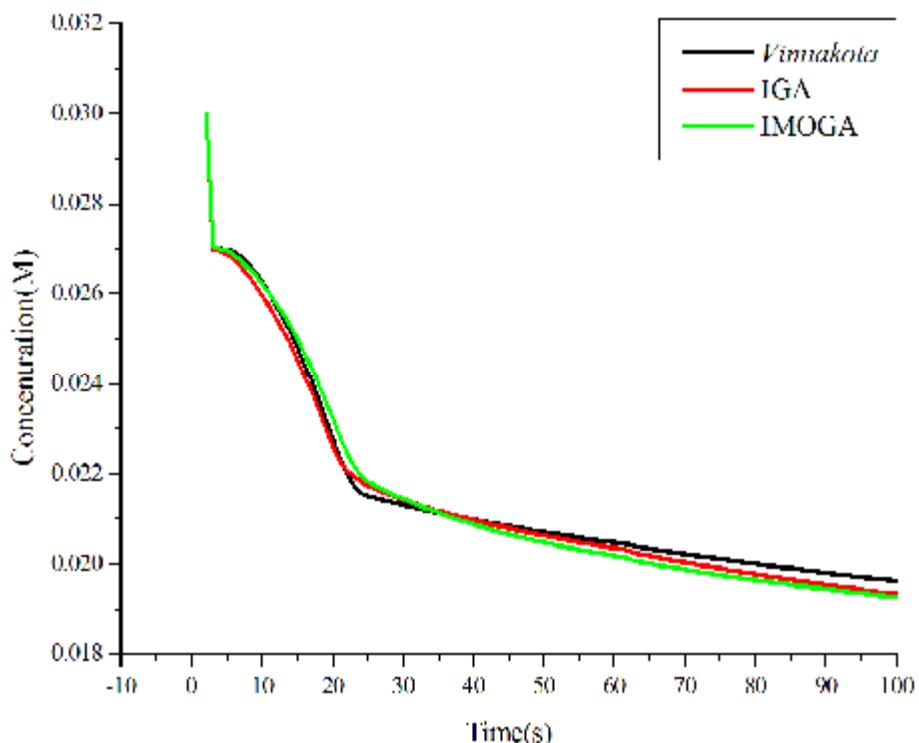


圖 42 Cr 的濃度變化

相較於圖 43 的 Pi 其反應量跟 Cr 差不多，但是對於最終產物的影響卻差異很大。將濃度變化與 ATP 及 Cr 相對照，可看到 Pi 對於反應劇烈的變化速度相較之下就較為和緩。但不論是 Pi 或是 Cr，IMOGA 都是變化最大的。

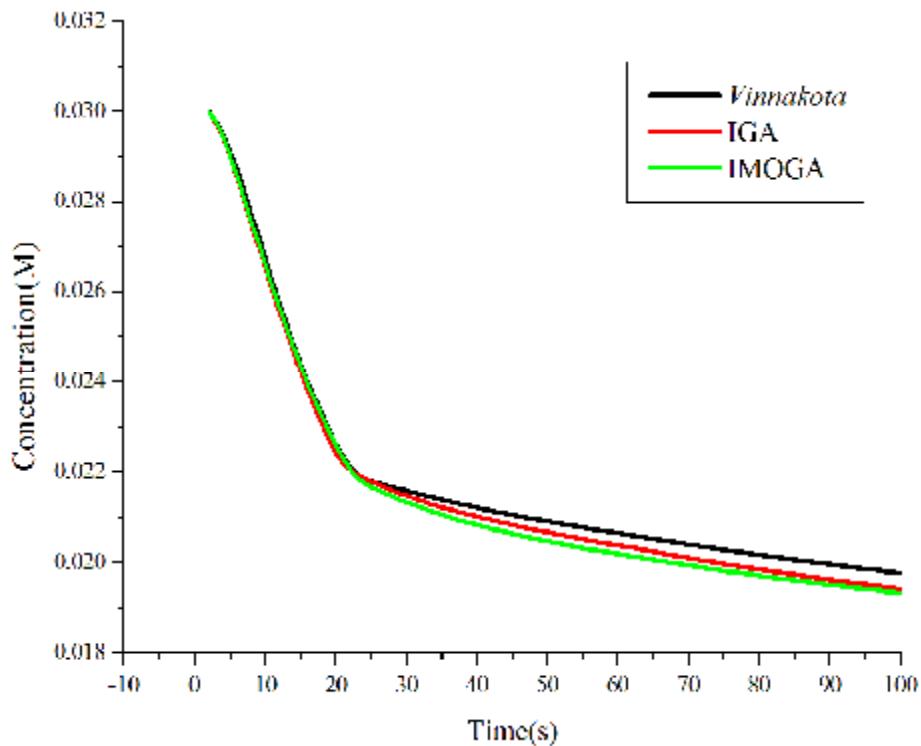


圖 43 Pi 的濃度變化

其餘 13 個反應物的濃度陸續陳列在圖 44 至圖 50，大致上 IGA 的效果會顯示在反應前端的反應物；Vinnakota [44]會在反應中段，也就是在造成剛性方程式的片段；而 IMO GA 的部分在反應途徑末端才顯得特別靈敏。此外還有一個有趣的現象，IMO GA 跟 IGA 所造成的反應物濃度變化曲線大部分相似，但反應量會有所差異。

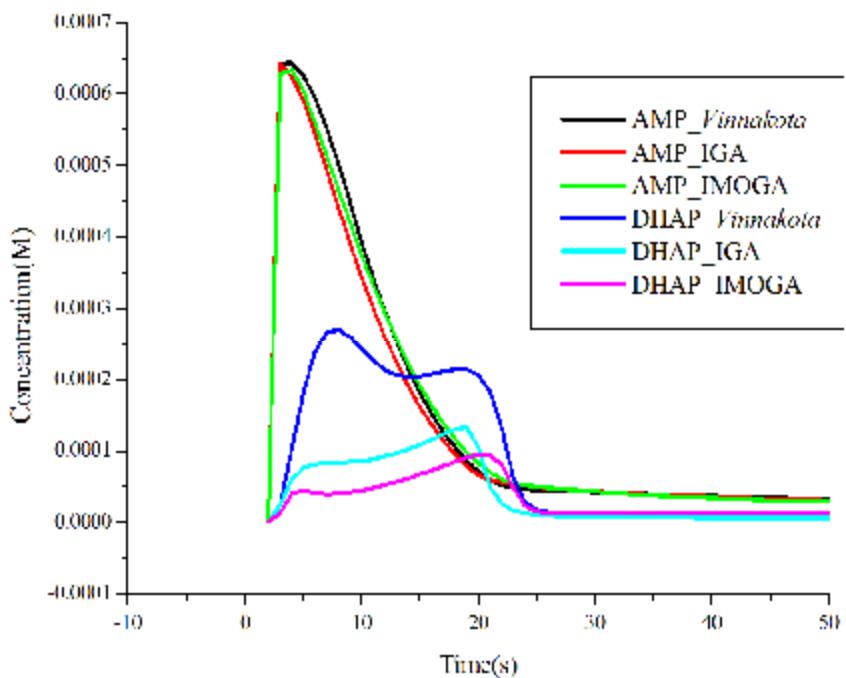


圖 44 AMP 及 DHAP 之濃度變化

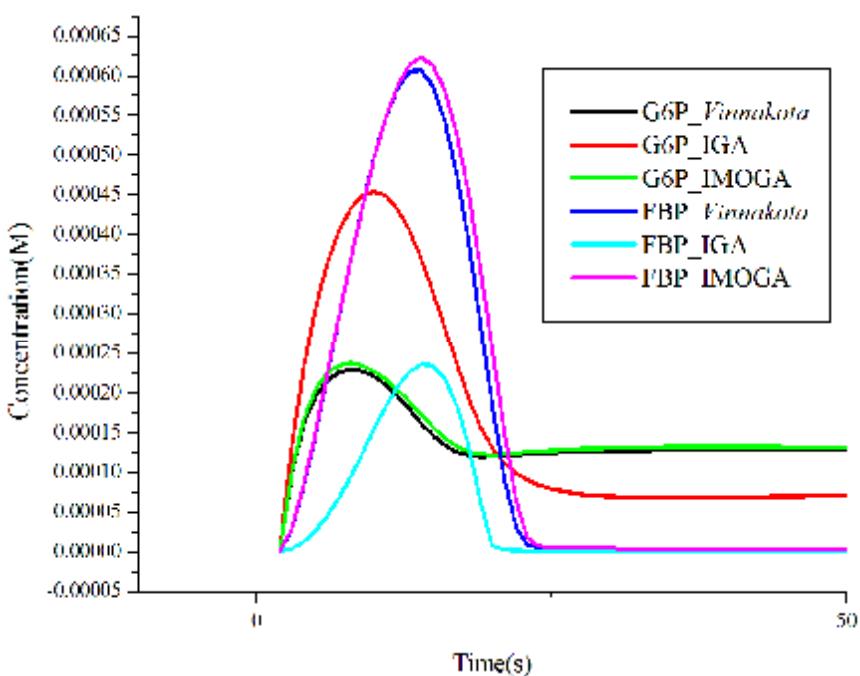


圖 45 G6P 及 FBP 之濃度變化

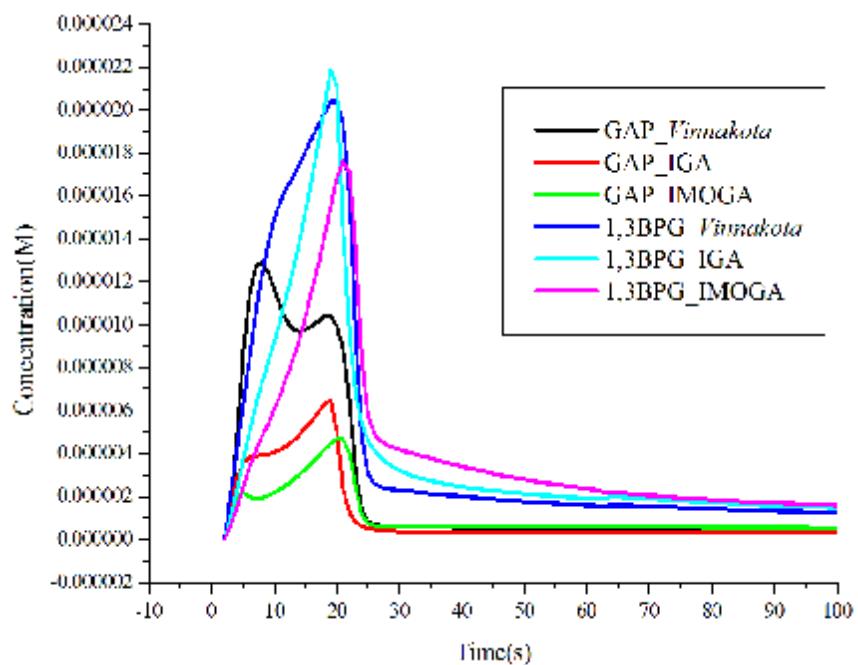


圖 46 GAP 及 1,3BPG 之濃度變化

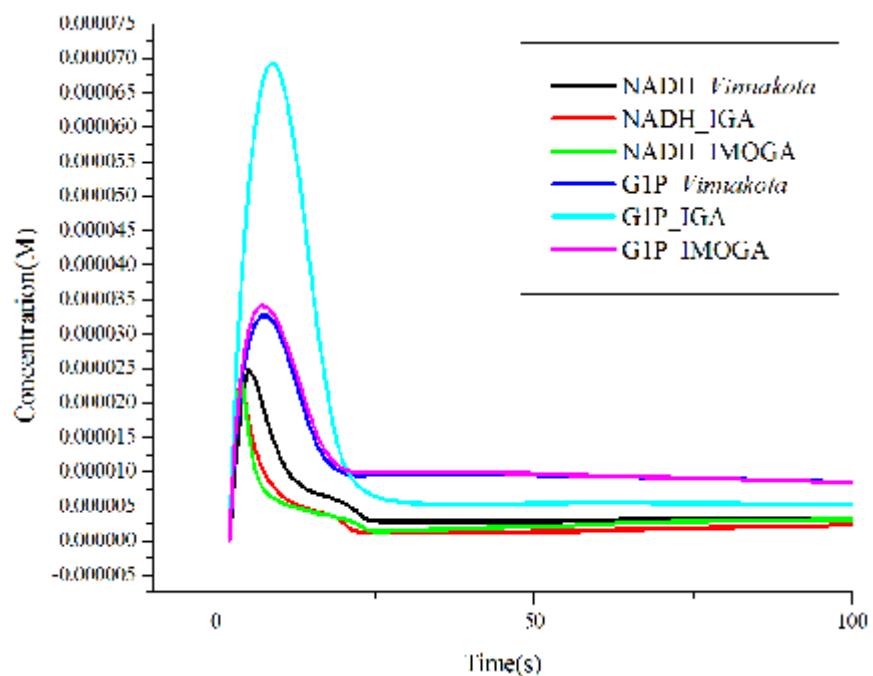


圖 47 NADH 及 G1P 之濃度變化

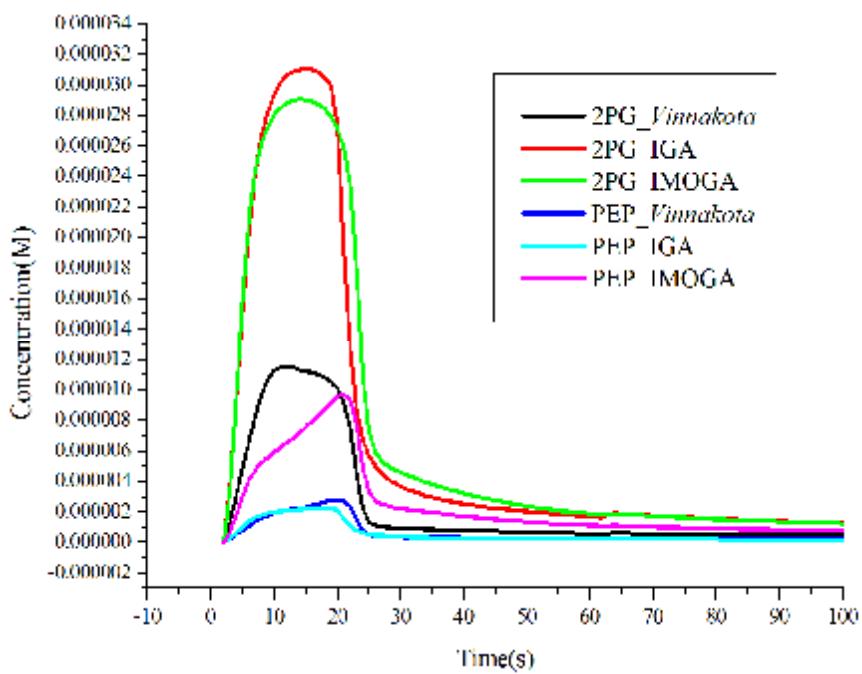


圖 48 2PG 及 PEP 之濃度變化

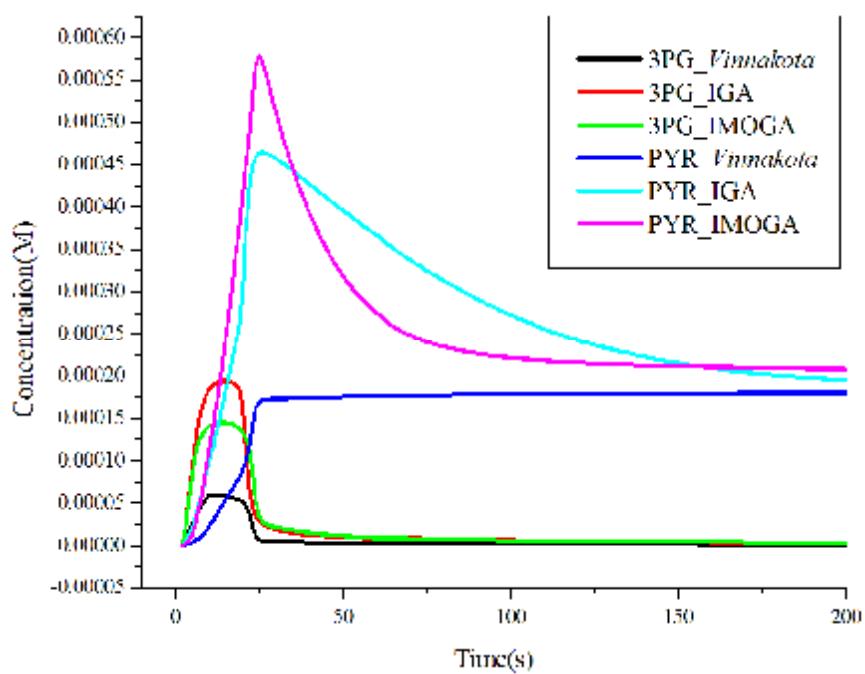


圖 49 3PG 及 PYR 之濃度變化

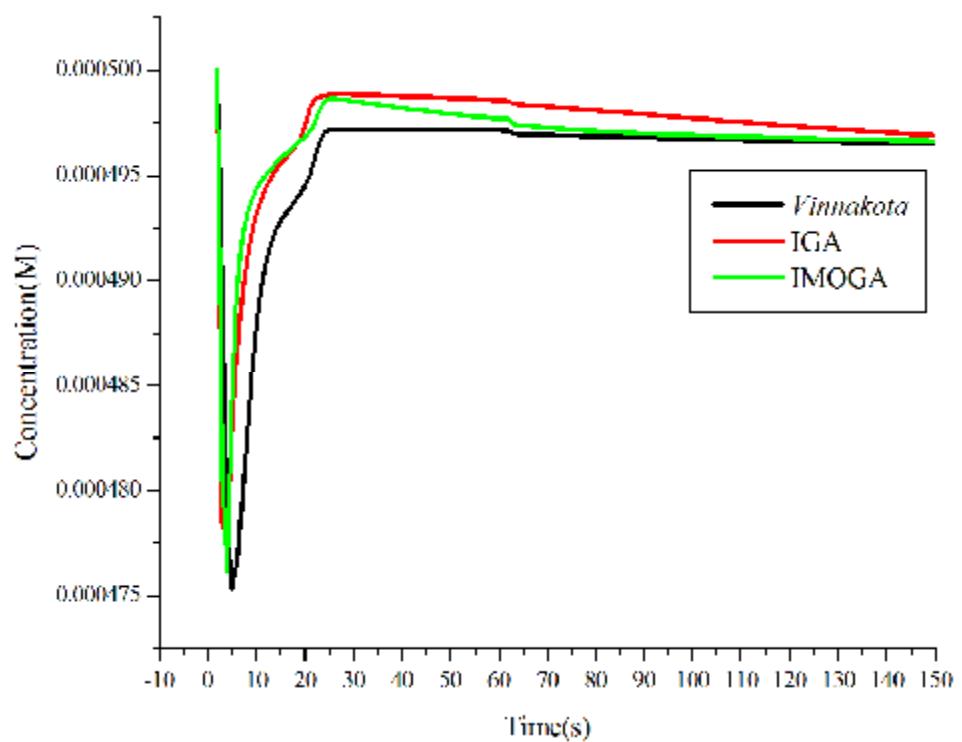


圖 50 NAD 的濃度變化



七、問題討論

7.1 討論

在本研究裡我們共使用兩種整體最佳化的演化式演算法 IGA 及 IMOGA 去估計生化網路模型的動力學參數值，主要是為了想從反應過程中的濃度數據將代謝網路重建。使用 IMOGA 是除了利用反應物濃度之外，再把系統上的熱力學性質一併考慮整合而成的最佳化之方法。由於取系統總自由能最低點及濃度誤差最小化都是屬於一般代謝網路上共有的性質，故 IMOGA 之應用能更加寬廣而不被特定代謝網路所侷限。本文中只應用在肌肉之醣解代謝上，主因是 *Lambeth*[55] 及 *Kushmerick* [44] 的模型已將相關要素整合進來，並且善加考慮 pH 對反應進行的影響。由於多考量 pH 的效應就會使得系統更複雜，估算參數的困難性也會相對提高[70, 71]，但是 pH 效應不可避免地存在於大多數的生化反應上。同樣地，其他代謝網路可藉由第二章的相關理論逐步建出其模型架構，並且越是考慮詳盡的代謝路徑之動力學模型就越會需要更多的動力學參數。遺憾的是大多數參數皆屬未知數，但建模型又需要這些未知之參數，所以利用我們所提出的 EKTOM 將系統參數可估計出一組最佳化之參數組合以提供使用。

我們運用 IGA 將模型分別將兩目標函數最佳化，兩組實驗並沒有集中。而在 f_E 的末端會發現兩者衝突的現象；而有雜訊存在時 f_E 既降不下去， f_G 也會發散。我們也用改變權重值的方式，但依舊無法有兩者並行的方法，最終才使用 IMOGA 一次滿足兩項目標。使用 IMOGA 所得到之最佳解由於多考量 f_G ，會比只考量 f_E 之最佳解更能反映出過程中反應物濃度之劇烈變化。IMOGA 的 f_E 雖然並沒有比 IGA 的最佳解低，但是在敏感度測試上兩者大部份的影響皆相似，而在 P_i 的敏感度卻更明顯。此外，兩者分別對反應物起始濃度加入 1% 和 3% 的擾亂的實驗結果相同顯示所建之模型的穩定且能呈現出對初始濃度之改變應有的回應，對於此系統具有相當優秀之模擬能力。

經由多考慮熱力學造成的模型能反應出瞬間的濃度變化最是能顯示出這模型最佳化後的優秀，在 *Smith* 和 *Missen*[59] 中認為反應速率改變(不論是反應或是質量傳遞)的迅速對應就可能使得模型有用。

並且每當反應速率的改變相對緩慢時，動力學模型的效用一般來說都會減少。所以思考如何加入能讓動力學模型迅速對應反應速度之改變的條件就會更顯得重要，而 EKTOM 的實驗結果就顯示出確實能解決這項擔憂。

然而利用決定論的演算法計算化學平衡將可能對產物的預測造成錯誤，因為關於封閉系統的化學平衡模型起初就必須規定系統質量須守恆[41]，但是生物系統是不能被獨立出來的。生物系統會與環境進行能量及成分的交換，並搭配熱量的消散[43]。所以反應後期過程的平衡狀態將會改變，由於反應速率與化學平衡息息相關，因此將會使得反應速率產生不可預期的變化。在我們數學模型中就代表動力學參數值將可能不會是反應初期所定下的值，而須做些修正。因為上述之原因，即可能使得對產物的預測發生誤差。

7.2 未來展望

在已知系統結構及實驗資料下估計模型使用的參數是為了將校準模型而讓實驗結果能再現於模型上[72]，所以參數之最佳化需要對整個系統做。但是卻沒有演算法能在有限時間內準確地解一般性的問題。建立代謝網路的要點在能表示出反應過程的變化，所以模型中反應物濃度的變化是觀察模型好壞的條件之一[59]。存在於生物網路間交錯的相互關係中，有些調節關係的觸發條件是跟會濃度梯度有關，所以將代謝網路模型盡可能地還原出實驗結果將會有助於研究系統生物學。

往後能將代謝網路與基因調節網路或是與信號傳遞網路整合是未來系統生物學的趨勢，實際上也已有如此研究逐漸被發表出來[2-6]，所以我們建出的網路模型在未來將可經整合而成更複雜的系統而讓生物系統的模擬更為完善。

參考文獻

1. Kitano, H., *Systems biology: a brief overview*. Science, 2002. **295**(5560): p. 1662-4.
2. Griggs, D.W. and M. Johnston, *Regulated expression of the GAL4 activator gene in yeast provides a sensitive genetic switch for glucose repression*. Proc Natl Acad Sci U S A, 1991. **88**(19): p. 8597-601.
3. Natarajan, K., et al., *Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast*. Mol Cell Biol, 2001. **21**(13): p. 4347-68.
4. Oh, M.K. and J.C. Liao, *Gene expression profiling by DNA micro-arrays and metabolic fluxes in Escherichia coli*. Biotechnol Prog, 2000. **16**(2): p. 278-86.
5. Quan, J.A., et al., *Regulation of carbon utilization by sulfur availability in Escherichia coli and Salmonella typhimurium*. Microbiology, 2002. **148**(Pt 1): p. 123-31.
6. Yeang, C.H. and M. Vingron, *A joint model of regulatory and metabolic networks*. BMC Bioinformatics, 2006. **7**: p. 332.
7. Michal, G., *Biochemical pathways : an atlas of biochemistry and molecular biology*. 1998, New York: John Wiley. xi.
8. Kanehisa, M., et al., *The KEGG databases at GenomeNet*. Nucleic Acids Res, 2002. **30**(1): p. 42-6.
9. Joshi-Tope, G., et al., *Reactome: a knowledgebase of biological pathways*. Nucleic Acids Res, 2005. **33**(Database issue): p. D428-32.
10. Forster, J., et al., *Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network*. Genome Res, 2003. **13**(2): p. 244-53.
11. Nolan, R.P., A.P. Fenley, and K. Lee, *Identification of distributed metabolic objectives in the hypermetabolic liver by flux and energy balance analysis*. Metab Eng, 2006. **8**(1): p. 30-45.
12. Herrgard, M.J., et al., *Integrated analysis of regulatory and meta-*

- bolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*.* Genome Res, 2006. **16**(5): p. 627-35.
- 13. Savageau, M.A., *Biochemical systems analysis. 3. Dynamic solutions using a power-law approximation.* J Theor Biol, 1970. **26**(2): p. 215-26.
 - 14. Alvarez-Vasquez, F., et al., *Simulation and validation of modelled sphingolipid metabolism in *Saccharomyces cerevisiae*.* Nature, 2005. **433**(7024): p. 425-30.
 - 15. Salter, M., R.G. Knowles, and C.I. Pogson, *Metabolic control.* Essays Biochem, 1994. **28**: p. 1-12.
 - 16. Varma, A. and B.O. Palsson, *Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110.* Appl Environ Microbiol, 1994. **60**(10): p. 3724-31.
 - 17. Schuster, S., T. Dandekar, and D.A. Fell, *Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering.* Trends Biotechnol, 1999. **17**(2): p. 53-60.
 - 18. Schilling, C.H., D. Letscher, and B.O. Palsson, *Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective.* J Theor Biol, 2000. **203**(3): p. 229-48.
 - 19. Beard, D.A., S.D. Liang, and H. Qian, *Energy balance for analysis of complex metabolic networks.* Biophys J, 2002. **83**(1): p. 79-86.
 - 20. Reed, J.L., et al., *An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR).* Genome Biol, 2003. **4**(9): p. R54.
 - 21. Kuepfer, L., U. Sauer, and L.M. Blank, *Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*.* Genome Res, 2005. **15**(10): p. 1421-30.
 - 22. Westerhoff, H.V. and B.O. Palsson, *The evolution of molecular biology into systems biology.* Nat Biotechnol, 2004. **22**(10): p. 1249-52.
 - 23. Voit, E.O., *Utility of Biochemical Systems Theory for the analysis of metabolic effects from low-dose chemical exposure.* Risk Anal, 2000. **20**(3): p. 393-402.

24. Chou, I.C., H. Martens, and E.O. Voit, *Parameter estimation in biochemical systems models with alternating regression*. Theor Biol Med Model, 2006. **3**: p. 25.
25. Heinrich, R.a.S., S., *The regulation of cellular systems*. 1996: Chapman & Hall.
26. Klipp, E., Herwig, R, Kowald, A, Wierling, C, and Lehrach, H, *Systems biology in practice. Concepts, implementation, and application*. 2005, Weinheim: Wiley-VCH.
27. Alfieri, R., et al., *A data integration approach for cell cycle analysis oriented to model simulation in systems biology*. BMC Syst Biol, 2007. **1**: p. 35.
28. Steuer, R., et al., *Structural kinetic modeling of metabolic networks*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(32): p. 11868-11873.
29. Brown, K.S. and J.P. Sethna, *Statistical mechanical approaches to models with many poorly known parameters*. Phys Rev E Stat Nonlin Soft Matter Phys, 2003. **68**(2 Pt 1): p. 021904.
30. Liebermeister, W. and E. Klipp, *Biochemical networks with uncertain parameters*. Syst Biol (Stevenage), 2005. **152**(3): p. 97-107.
31. Schwartz, J.M. and M. Kanehisa, *Quantitative elementary mode analysis of metabolic pathways: the example of yeast glycolysis*. BMC Bioinformatics, 2006. **7**: p. 186.
32. Goldberg, R.N., Y.B. Tewari, and T.N. Bhat, *Thermodynamics of enzyme-catalyzed reactions--a database for quantitative biochemistry*. Bioinformatics, 2004. **20**(16): p. 2874-7.
33. Schomburg, I., et al., *BRENDA, the enzyme database: updates and major new developments*. Nucleic Acids Res, 2004. **32**(Database issue): p. D431-3.
34. Liebermeister, W. and E. Klipp, *Bringing metabolic networks to life: integration of kinetic, metabolic, and proteomic data*. Theor Biol Med Model, 2006. **3**: p. 42.
35. Liebermeister, W. and E. Klipp, *Bringing metabolic networks to life: convenience rate law and thermodynamic constraints*. Theor Biol Med Model, 2006. **3**: p. 41.
36. Ren, Z. and S.B. Pope, *Reduced description of complex dynamics in*

- reactive systems.* J Phys Chem A, 2007. **111**(34): p. 8464-74.
- 37. Maskow, T. and U. von Stockar, *How reliable are thermodynamic feasibility statements of biochemical pathways?* Biotechnol Bioeng, 2005. **92**(2): p. 223-30.
 - 38. Shaik, O.S., et al., *Derivation of a quantitative minimal model from a detailed elementary-step mechanism supported by mathematical coupling analysis.* J Chem Phys, 2005. **123**(23): p. 234103.
 - 39. Gorban, A.N., I.V. Karlin, and A.Y. Zinovyev, *Invariant grids for reaction kinetics.* Physica a-Statistical Mechanics and Its Applications, 2004. **333**: p. 106-154.
 - 40. Okino, M.S. and M.L. Mavrovouniotis, *Simplification of Mathematical Models of Chemical Reaction Systems.* Chem Rev, 1998. **98**(2): p. 391-408.
 - 41. Zayer, A., U. Riedel, and J. Warnatz, *A hybrid genetic algorithm approach to calculating chemical equilibrium and detonation parameters in condensed energetic materials.* Combustion Theory and Modelling, 2006. **10**(5): p. 799-813.
 - 42. Jarungthammachote, S. and A. Dutta, *Equilibrium modeling of gasification: Gibbs free energy minimization approach and its application to spouted bed and spout-fluid bed gasifiers.* Energy Conversion and Management, 2008. **49**(6): p. 1345-1356.
 - 43. Qian, H. and D.A. Beard, *Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium.* Biophys Chem, 2005. **114**(2-3): p. 213-20.
 - 44. Vinnakota, K., M.L. Kemp, and M.J. Kushmerick, *Dynamics of muscle glycogenolysis modeled with pH time course computation and pH-dependent reaction equilibria and enzyme kinetics.* Biophys J, 2006. **91**(4): p. 1264-87.
 - 45. Srividhya, J., et al., *Reconstructing biochemical pathways from time course data.* Proteomics, 2007. **7**(6): p. 828-838.
 - 46. Ho, S.Y., L.S. Shu, and J.H. Chen, *Intelligent evolutionary algorithms for large parameter optimization problems.* Ieee Transactions on Evolutionary Computation, 2004. **8**(6): p. 522-541.
 - 47. Scopes, R.K., *Studies with a reconstituted muscle glycolytic system. The rate and extent of creatine phosphorylation by anaerobic gly-*

- colysis. *Biochem J*, 1973. **134**(1): p. 197-208.
48. Stephanopoulos, G., H. Alper, and J. Moxley, *Exploiting biological complexity for strain improvement through systems biology*. *Nature Biotechnology*, 2004. **22**(10): p. 1261-1267.
49. Beard, D.A., et al., *Thermodynamic constraints for biochemical networks*. *J Theor Biol*, 2004. **228**(3): p. 327-33.
50. Beard, D.A. and H. Qian, *Thermodynamic-based computational profiling of cellular regulatory control in hepatocyte metabolism*. *Am J Physiol Endocrinol Metab*, 2005. **288**(3): p. E633-44.
51. Ederer, M. and E.D. Gilles, *Thermodynamically feasible kinetic models of reaction networks*. *Biophys J*, 2007. **92**(6): p. 1846-57.
52. Henry, C.S., L.J. Broadbelt, and V. Hatzimanikatis, *Thermodynamics-based metabolic flux analysis*. *Biophys J*, 2007. **92**(5): p. 1792-805.
53. Qian, H., D.A. Beard, and S.D. Liang, *Stoichiometric network theory for nonequilibrium biochemical systems*. *Eur J Biochem*, 2003. **270**(3): p. 415-21.
54. Alberty, R.A., *Standard transformed Gibbs energies of coenzyme A derivatives as functions of pH and ionic strength*. *Biophys Chem*, 2003. **104**(1): p. 327-34.
55. Lambeth, M.J. and M.J. Kushmerick, *A computational model for glycogenolysis in skeletal muscle*. *Ann Biomed Eng*, 2002. **30**(6): p. 808-27.
56. Alberty, R.A., *Thermodynamics of Biochemical Reactions*. 2003, Hoboken: John Wiley and Sons.
57. Cornish-Bowden, A., *Fundamentals of enzyme kinetics*. 3rd ed. 2004, London: Portland Press. xvi.
58. Goldberg, R.N. and Y.B. Tewari, *Thermodynamics of the disproportionation of adenosine 5'-diphosphate to adenosine 5'-triphosphate and adenosine 5'-monophosphate I. Equilibrium model*. *Biophys Chem*, 1991. **40**(3): p. 241-61.
59. Smith, W.R. and R.W. Missen, *Chemical reaction equilibrium analysis : theory and algorithms*. 1982, New York: Wiley. xvi.
60. Smith, J.M., H.C. Van Ness, and M.M. Abbott, *Introduction to chemical engineering thermodynamics*. 7th ed. McGraw-Hill chem-

- ical engineering series. 2005, Boston: McGraw-Hill. xviii.
61. Segel, I.H., *Enzyme kinetics : behavior and analysis of rapid equilibrium and steady-state enzyme systems*. Wiley classics library ed. 1993, New York: John Wiley & Sons. xxii.
 62. Cornish-Bowden, A., *Fundamentals of enzyme kinetics*. Rev. ed. 1995, London: Portland Press. xiii.
 63. Matsubara, Y., et al., *Parameter estimation for stiff equations of biosystems using radial basis function networks*. BMC Bioinformatics, 2006. **7**: p. 230.
 64. Arnold, H. and D. Pette, *Binding of glycolytic enzymes to structure proteins of the muscle*. Eur J Biochem, 1968. **6**(2): p. 163-71.
 65. Cheetham, M.E., et al., *Human muscle metabolism during sprint running*. J Appl Physiol, 1986. **61**(1): p. 54-60.
 66. Harris, R.C., E. Hultman, and L.O. Nordesjo, *Glycogen, glycolytic intermediates and high-energy phosphates determined in biopsy samples of musculus quadriceps femoris of man at rest. Methods and variance of values*. Scand J Clin Lab Invest, 1974. **33**(2): p. 109-20.
 67. Sahlin, K., *NADH in human skeletal muscle during short-term intense exercise*. Pflugers Arch, 1985. **403**(2): p. 193-6.
 68. Taguchi, G.i., S. Konishi, and American Supplier Institute., *Orthogonal arrays and linear graphs : tools for quality engineering*. 1987, Dearborn, Mich.: American Supplier Institute. vii.
 69. Richard, P., et al., *Sustained oscillations in free-energy state and hexose phosphates in yeast*. Yeast, 1996. **12**(8): p. 731-40.
 70. Ishii, N., et al., *Dynamic simulation of an in vitro multi-enzyme system*. FEBS Lett, 2007. **581**(3): p. 413-20.
 71. Dash, R.K., et al., *Modeling cellular metabolism and energetics in skeletal muscle: large-scale parameter estimation and sensitivity analysis*. IEEE Trans Biomed Eng, 2008. **55**(4): p. 1298-318.
 72. Rodriguez-Fernandez, M., P. Mendes, and J.R. Banga, *A hybrid approach for efficient and robust parameter estimation in biochemical pathways*. Biosystems, 2006. **83**(2-3): p. 248-65.

73. Yang, J., et al., *Kinetic Monte Carlo method for rule-based modeling of biochemical networks*. Phys Rev E Stat Nonlin Soft Matter Phys, 2008. **78**(3 Pt 1): p. 031910.



附錄

全資料來自 [44]

Reference species	ΔG°	$N_{H(i)}$	pK_{a_1}	$\Delta H(I=0)$	pK_{a_2}	$pK_{a_{Mg}}$	$\Delta H(I=0)$	pK_{a_K}	$\Delta H(I=0)$
G1P ²⁻	-1756.87	11	6.09	-1.7		2.48	-12		
G6P ²⁻	-1763.94	11	6.11						
F6P ²⁻	-1760.8	11	5.89						
FDP ⁴⁻	-2601.4	10	6.4		5.92	2.7			
G3P ²⁻	-1339.25	5	6.22	-3.1		1.63			
DHAP ²⁻	-1296.26	5	5.9			1.57			
GAP ²⁻	-1288.6	5	6.45						
13DPG ⁴⁻	-2356.14	4	7.5						
3PG ³⁻	-1502.54	4	6.21						
2PG ³⁻	-1496.38	4	7		2.45		1.18		
PEP ³⁻	-1263.65	2	6.35		2.26		1.08		
PYR ⁻	-472.27	3	2.49						
LAC ⁻	-516.72	5	3.67	-0.33		0.98			
HPi ²⁻	-1096.1	1	6.75	3		1.65	-2.9	0.5	
IMP ²⁻	<i>Not avail.</i>	11	6.34	-2		1.67			
AMP ²⁻	-1040.45	12	6.29	-3		1.92	-7.5		
ADP ³⁻	-1906.13	12	6.38	-3		3.25	-15	1	
ATP ⁴⁻	-2768.1	12	6.48	-5		4.19	-18	1.17	-1
HPCr ²⁻	<i>Not avail.</i>	9	4.5	2.66		1.6	8.19	0.31	
HCr ⁰	Not avail.	8	2.63						
Glycogen(n)	0	$0n+1$	-						
Glycogen($n-1$)	679.1	$0(n-1)+1$	-						
NAD ⁻	0	26	-						
NADH ²⁻	22.65	27	-						
H ⁺	0	1	-						