Electronic Supplementary Material (ESI) for Green Chemistry.

Artificial Intelligence System for Enhanced Automated 1,3-propanediol Green Biosynthesis

Jiacheng Huang[1#], Chade-Deng Li[1,2#], Haodong Zhao[1,2], Meng Yu[1], Aihui Zhang[1,3*], Baishan Fang[1,2,3*]

Affiliations:

1. Department of Chemical and Biochemical Engineering, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, 361005, China

2. Gulei Innovation Institute, Xiamen University, Zhangzhou, 363200, China

3. Shenzhen Research Institute of Xiamen University, Shenzhen, 518000, China

[#] These authors contributed equally to this work

[*]Corresponding authors: Tel: +86 0592 2185869

E-mail: zhangaihui@xmu.edu.cn (Aihui Zhang); fbs@xmu.edu.cn (Baishan Fang)

Table of Contents for Supporting Information

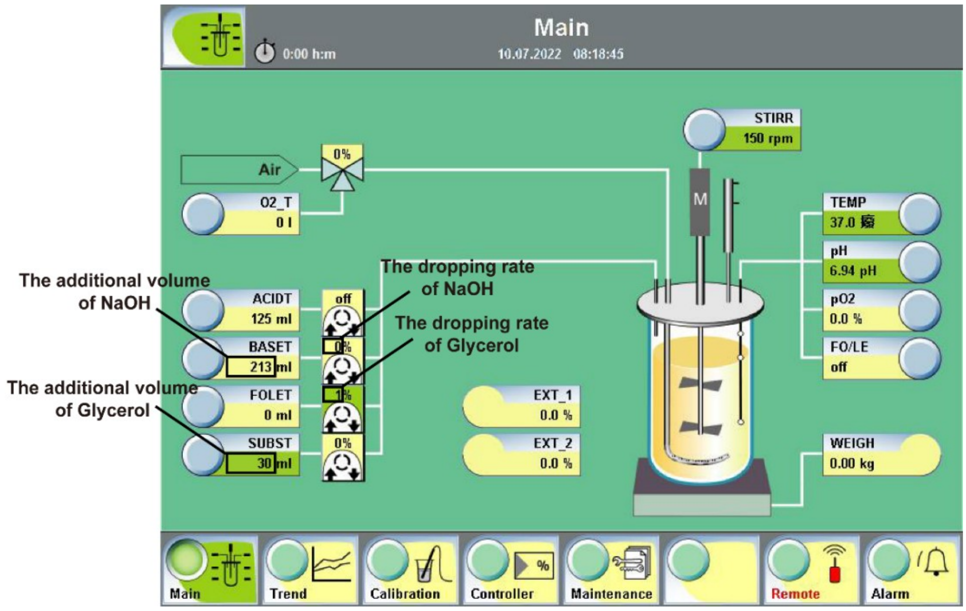

Figure S1. The real-time data of the flowmeter and speedometer for the base and feed are in the Sartorius bioreactor's supporting software.

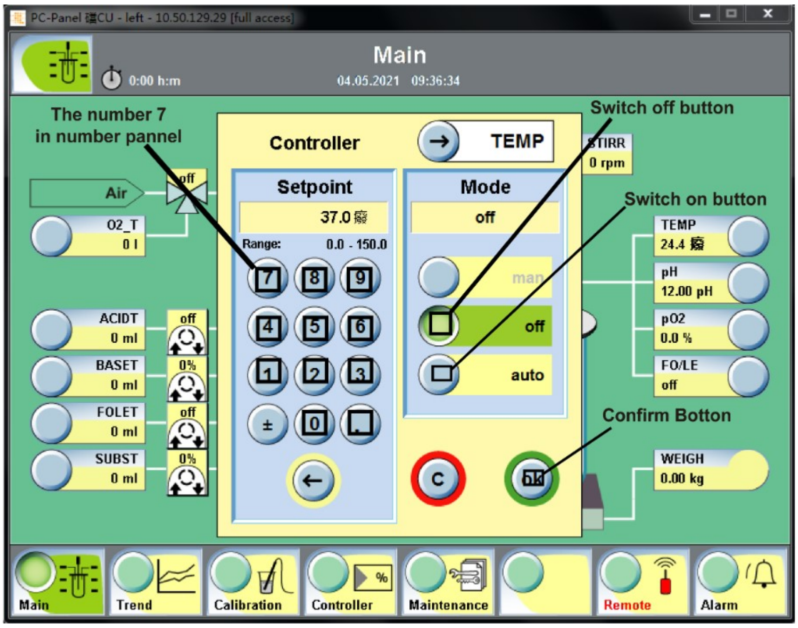

Figure S2. The number panel, on-off switch, and confirm button are in the Sartorius bioreactor's supporting software.

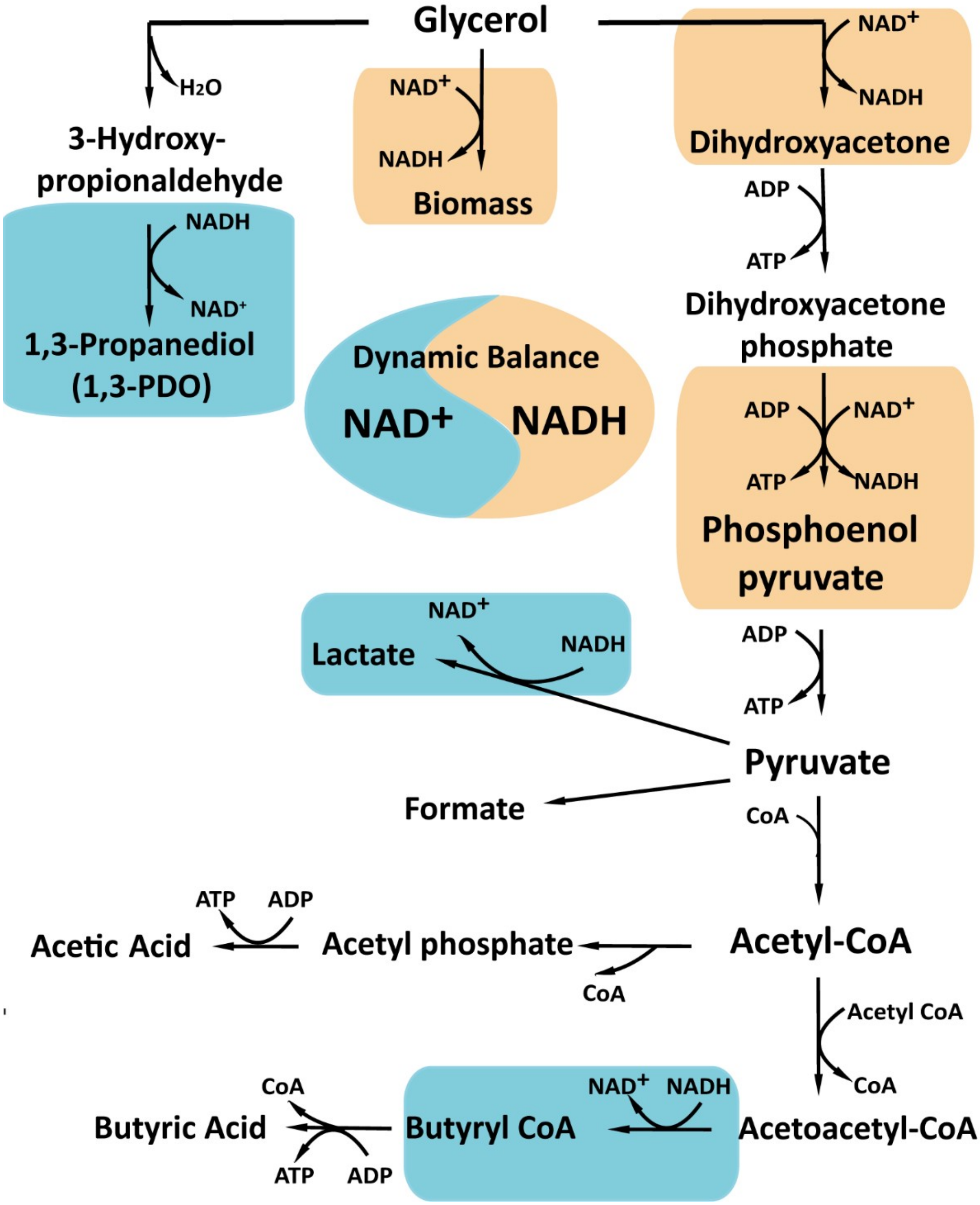

Figure S3. Schematic diagram of the main metabolic pathways closely associated with 1,3-PDO biosynthesis of *Clostridium butyricum* fed-batch fermentation process. The light blue part represents the pathway which would consume NADH, while the light red part represents the pathway which would generate NADH.

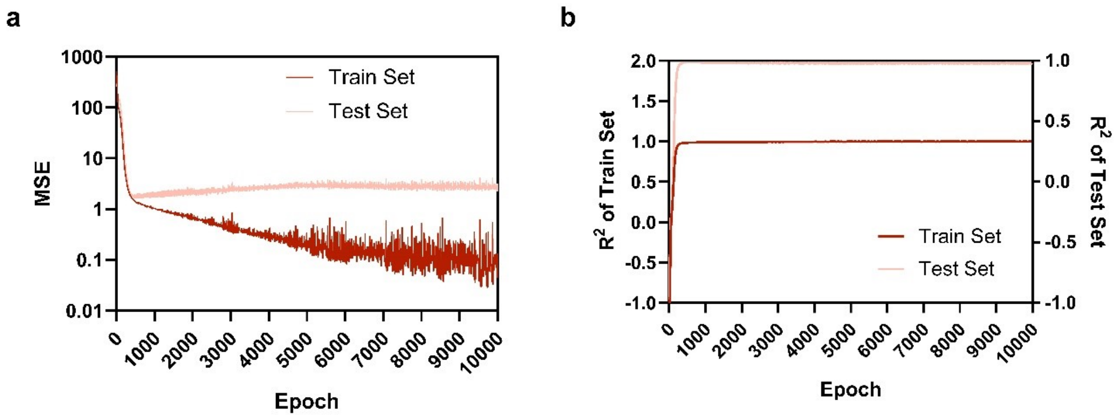

Figure S4. Training process of the soft sensor with best hyperparameters (4 layers with 80 nodes in each layer). The full process is monitored with two metrics, MSE (**a**) and R^2 (**b**).

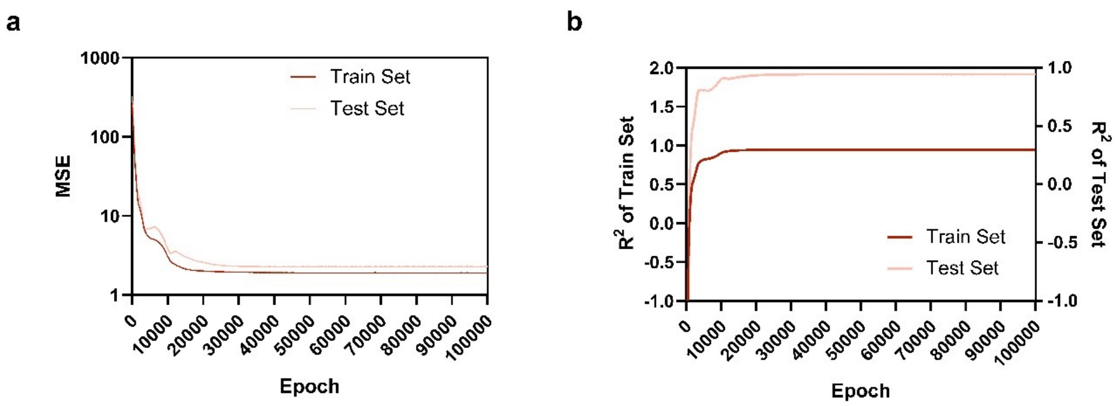

Figure S5. Training process of the predictor with best hyperparameters (8 hidden layers with 60 nodes in each layer). The full process is monitored with two metrics, MSE (**a**) and R^2 (**b**).

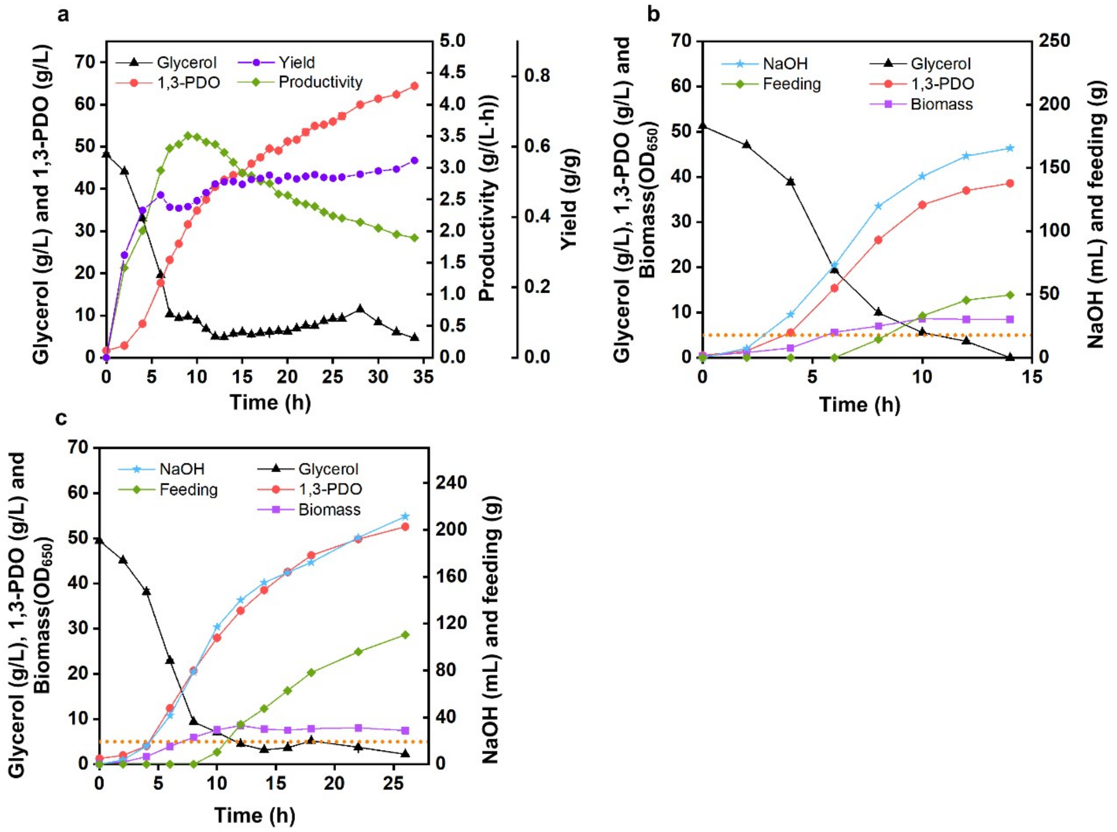

Figure S6. Time course of 1,3-PDO AI fed-batch fermentation. The relationship among 1,3-PDO concentrations, yield, and productivity (**a**). The concentrations of glycerol, 1,3-PDO (g/L), biomass (OD650), feeding (g), and NaOH (mL) in the AI fed-batch fermentations in which the threshold of NaOH dropping rate was set to 5 (**b**), 10 (**c**) mL/h.

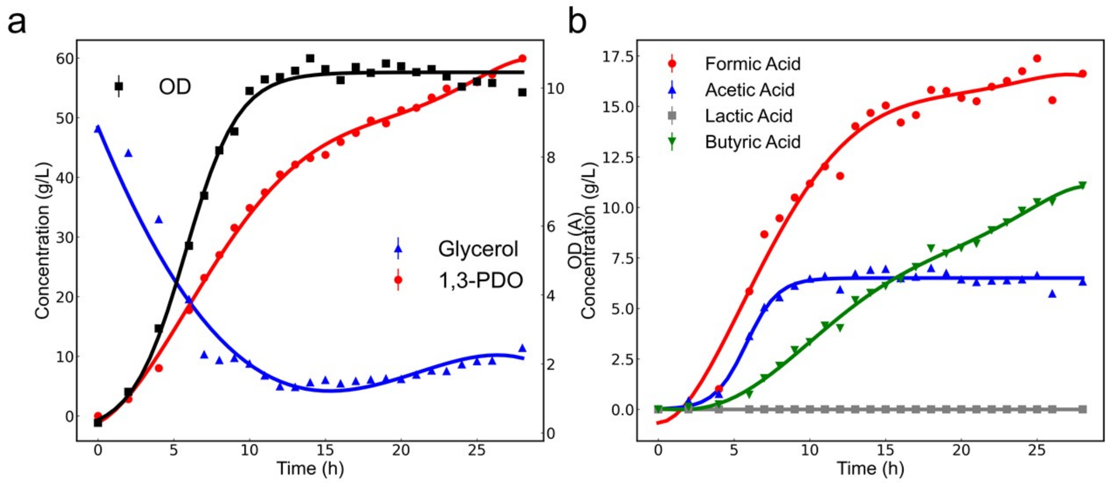

Figure S7. Curve fitting of the time course in 1,3-PDO fed-batch fermentation of G1. (**a**): the scale in the left vertical ordinate for the concentrations (g/L) of glycerol and 1,3-PDO, while the scale in the right vertical ordinate for the OD (A) of *C. butyricum*; (**b**): the scale in the left vertical ordinate for the concentrations (g/L) of formic acid, acetic acid, lactic acid, and butyric acid.

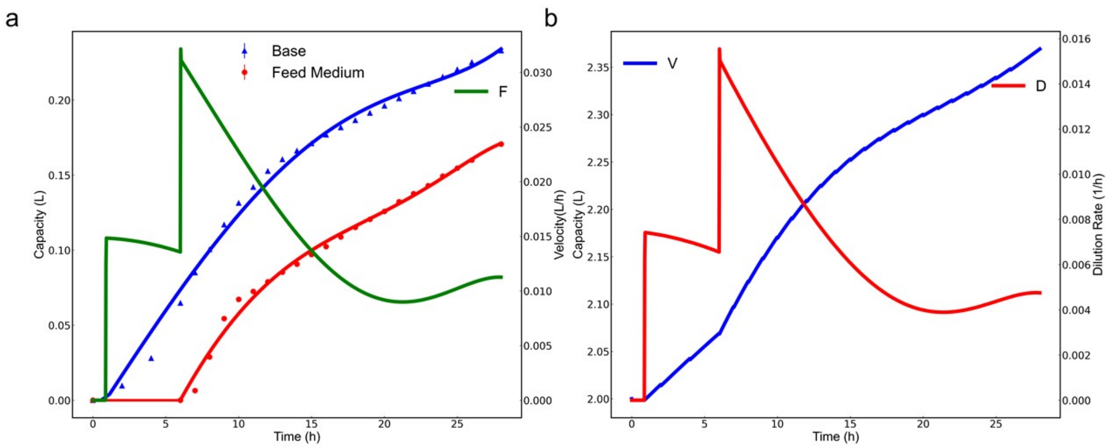

Figure S8. Curve fitting of the time course in 1,3-PDO fed-batch fermentation of G1. (**a**): the scale in the left vertical ordinate for the volume (L) of adding base and feed medium, while the scale in the right vertical ordinate for their total flow rate (L/h); (**b**): the scale in the left vertical ordinate for the volume (L) of the bioreactor, while the scale in the right vertical ordinate for dilution rate (1/h).

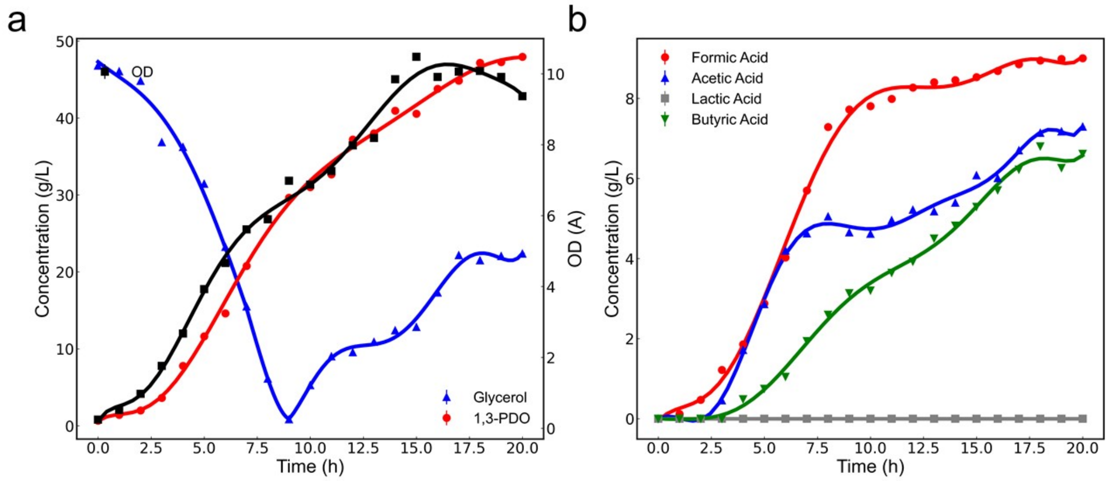

Figure S9. Curve fitting of the time course in 1,3-PDO fed-batch fermentation of G4. (**a**): the scale in the left vertical ordinate for the concentrations (g/L) of glycerol and 1,3-PDO, while the scale in the right vertical ordinate for the OD (A) of *C. butyricum*; (**b**): the scale in the left vertical ordinate for the concentrations (g/L) of formic acid, acetic acid, lactic acid, and butyric acid.

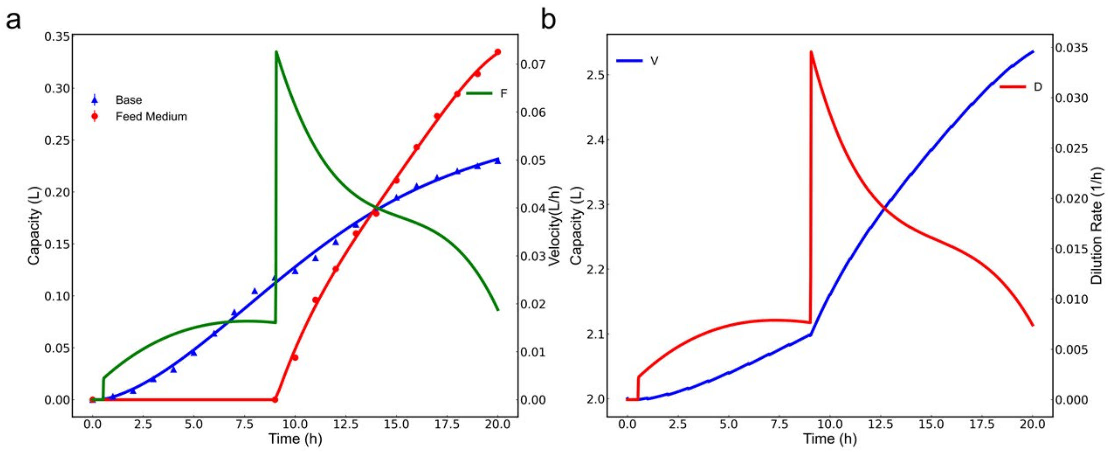

Figure S10. Curve fitting of the time course in 1,3-PDO fed-batch fermentation of G4. (**a**): the scale in the left vertical ordinate for the volume (L) of adding base and feed medium, while the scale in the right vertical ordinate for therir total flow rate (L/h); (**b**): the scale in the left vertical ordinate for the volume (L) of the bioreactor, while the scale in the right vertical ordinate for dilution rate (1/h).

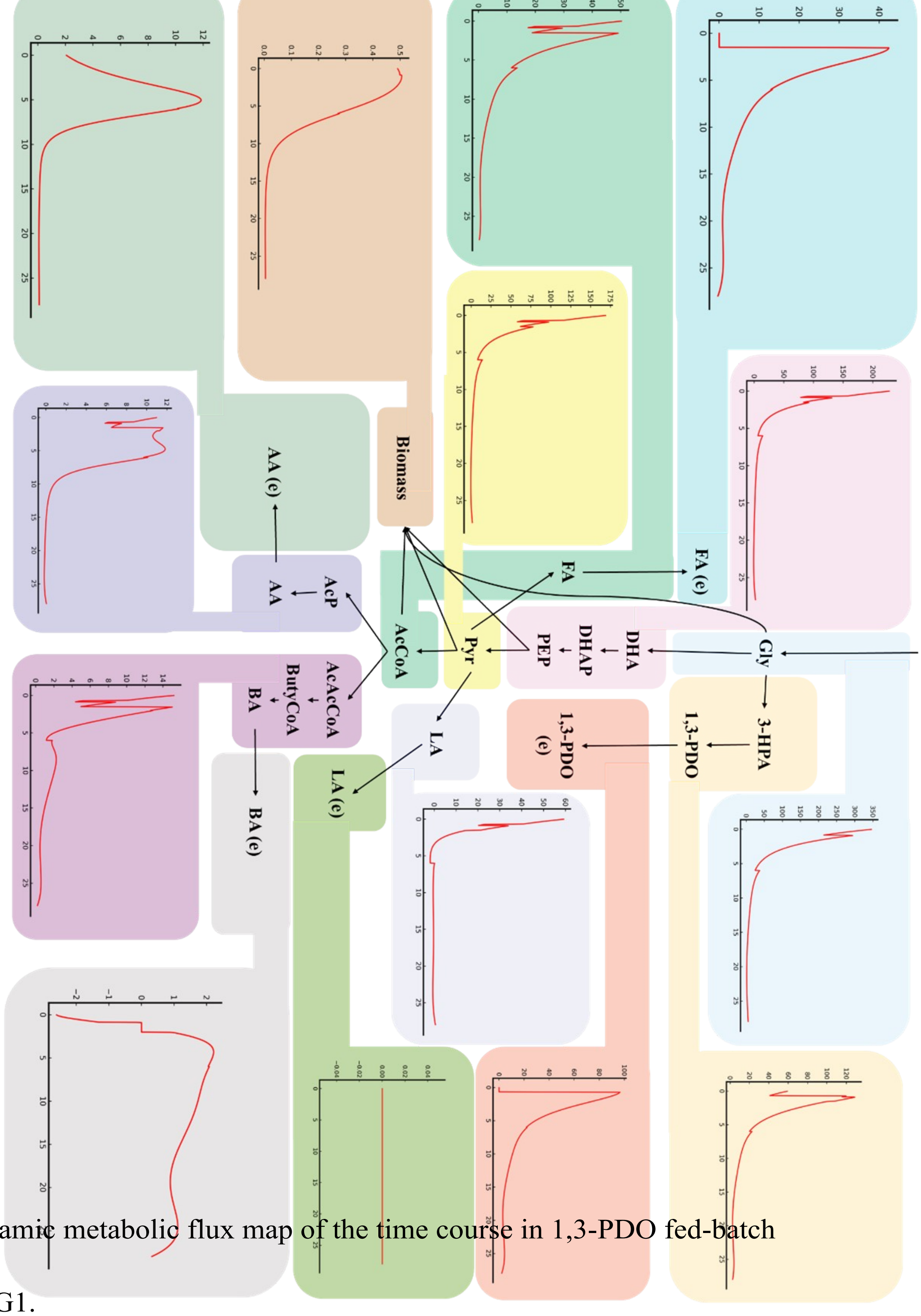

Figure S11. Dynamic metabolic flux map of the time course in 1,3-PDO fed-batch fermentation of G1.

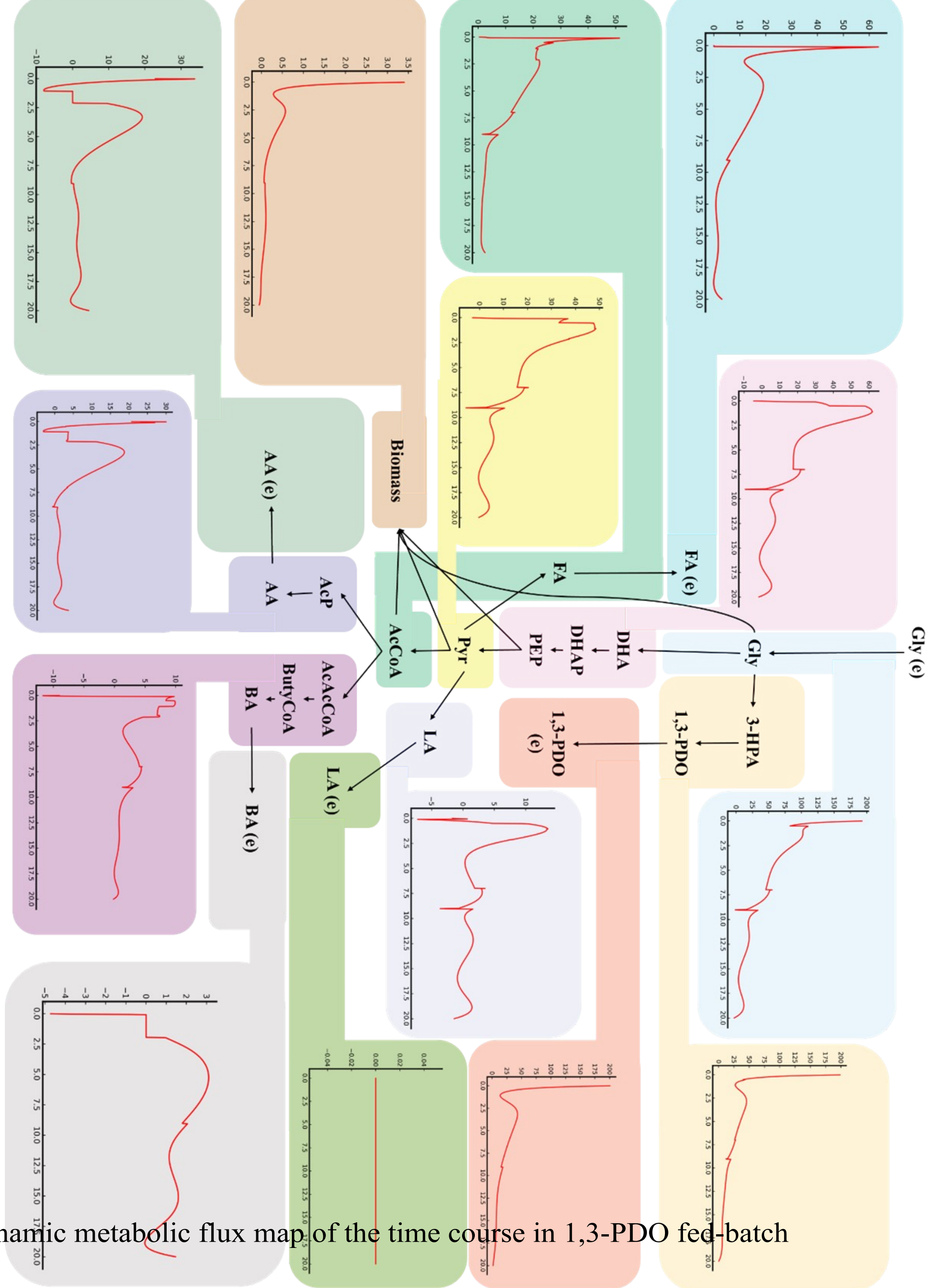

Figure S12. Dynamic metabolic flux map of the time course in 1,3-PDO fed-batch fermentation of G4.

Table S1 The carbon-conserving metabolic model of *C. butyricum* [1-3]. The abbreviations are listed in Table S2

Rate	Reaction
r_1	Gly + NADH → 1,3-PDO + NAD^+
r_2	Gly + $2NAD^+$ → PEP + 2NADH

r_3	PEP → Pyr
r_4	Pyr → FA + AcCoA
r_5	AcCoA → AA
r_6	Pyr + NADH → LA + NAD^+
r_7	AcCoA + AcCoA + 2NADH → BA + 2NAD^+
r_8	Gly(e) → Gly
r_9	1,3-PDO → 1,3-PDO(e)
r_{10}	FA → FA(e)
r_{11}	AA → AA(e)
r_{12}	LA → LA(e)
r_{13}	BA → BA(e)
r_{14}	NAD → NADH
r_{15}	3.07 Gly + 0.49 PEP + 3.52 Pyr + 2.10 AcCoA + 8.89 NADH → Biomass + 2.53 NAD^+

Table S2 The abbreviations of Table S1.

Abbreviations	**Full Name**
Gly	Glycerol
NADH	Reduced-nicotinamide Adenine Dinucleotide

1,3-PDO	1,3-Propanediol
NAD	Nicotinamide Adenine Dinucleotide
PEP	Phosphoenolpyruvate
Pyr	Pyruvic Acid
FA	Formic Acid
AcCoA	Acetyl-CoA
AA	Acetic Acid
LA	Lactic Acid
BA	Butyric Acid
Gly(e)	Glycerol (extracellular)
1,3-PDO(e)	1,3-Propanediol (extracellular)
FA(e)	Formic Acid (extracellular)
AA(e)	Acetic Acid (extracellular)
LA(e)	Lactic Acid (extracellular)
BA(e)	Butyric Acid (extracellular)
Biomass	Biomass

Table S3 Curve fitting results in Fig S7.

Goal	**Equation**	R^2
Gly (e)	$y = -6.16t + 2.06\times10^{-1}t^2 + 2.62\times10^{-3}t^3 - 1.39\times10^{-4}t^4$	0.951

1,3-PDO (e)	$y = 1.32t + 5.21\times10^{-1}t^2 - 3.40\times10^{-2}t^3 + 4.49\times10^{-5}t^5 - 8.53\times10^{-7}t^6$	0.997
OD	$y = \frac{10.45}{1+e^{-0.57\cdot(t-5.76)}}$	0.993
FA (e)	$y = 1.21\times10^{-2}t + 3.22\times10^{-1}t^2 - 2.93\times10^{-2}t^3 + 1.01\times10^{-3}t^4 - 1.23\times10^{-5}t^5$	0.980
AA (e)	$y = \frac{6.50}{1+e^{-0.95\cdot(t-5.81)}}$	0.981
LA (e)	$y = 0$	1.000
BA (e)	$y = -4.42\times10^{-2}t - 9.31\times10^{-3}t^2 + 1.32\times10^{-2}t^3 - 1.23\times10^{-3}t^4 + 4.33\times10^{-5}t^5 - 5.35\times10^{-7}t^6$	0.996

Table S4 Curve fitting result in Figure S8.

Goal	**Equation**	R^2
Base	$y = 1.49\times10^{-2}t - 1.21\times10^{-5}t^3 + 2.03\times10^{-10}t^6$	0.992
Feed	$y = \begin{cases} 0 & t \le 6 \\ 2.81\times10^{-2}t - 8.80\times10^{-4}t^2 + 1.91\times10^{-8}t^5 - 3.12\times10^{-13}t^8 & t > 6 \end{cases}$	0.993

Table S5 Curve fitting result in Figure S9.

Goal	**Equation**	R^2

Gly (e)	$y=\begin{cases} -2.20t+0.40t^2-0.19t^3+0.01t^4 & 0<t\le 7 \\ 4.87t-6.67t^2+2.30t^3-3.73\times10^{-1}t^4+ & \\ \quad 2.62\times10^{-2}t^5-6.09\times10^{-4}t^6 & 7<t\le 9 \\ -2.86\times10^{3}t-5.47\times10^{2}t^2-5.43\times10^{1}t^3+ & \\ \quad 2.97t^4-8.50\times10^{-2}t^5+9.94\times10^{-4}t^6 & 7<t\le 9 \end{cases}$	0.993
1,3-PDO (e)	$y=1.92t-1.74t^2+7.86\times10^{-1}t^3-1.26\times10^{-1}t^4$ $+1.01\times10^{-2}t^5-4.32\times10^{-4}t^6+9.46\times10^{-6}t^7-8.26\times10^{-8}t^8$	0.998
OD	$y=9.33\times10^{-1}t-8.70\times10^{-1}t^2+4.40\times10^{-1}t^3-9.08\times10^{-2}t^4$ $+9.54\times10^{-3}t^5-5.40\times10^{-4}t^6+1.57\times10^{-5}t^7-1.84\times10^{-7}t^8$	0.995
FA (e)	$y=5.44\times10^{-1}t-3.59\times10^{-1}t^2+1.36\times10^{-1}t^3-1.35\times10^{-2}t^4$ $+7.92\times10^{-7}t^5+7.04\times10^{-5}t^6-3.84\times10^{-6}t^7+6.39\times10^{-8}t^8$	0.997
AA (e)	$y=2.87\times10^{-1}t-3.34\times10^{-1}t^2-1.31\times10^{-2}t^3+1.04\times10^{-1}t^4-3.45\times10^{-2}t^5$ $+5.18\times10^{-3}t^6-4.27\times10^{-4}t^7+2.01\times10^{-5}t^8-5.06\times10^{-7}t^9+5.31\times10^{-9}t^{10}$	0.997
LA (e)	$y=0$	1.000
BA (e)	$y=-5.28\times10^{-2}t+6.69\times10^{-2}t^2-4.58\times10^{-2}t^3+1.72\times10^{-2}t^4$ $+2.65\times10^{-3}t^5+1.97\times10^{-4}t^6-7.03\times10^{-6}t^7-9.76\times10^{-8}t^8$	0.997

Table S6 Curve fitting result in Figure S10.

Goal	**Equation**	R^2
Base	$y=2.32\times10^{-3}t+2.09\times10^{-3}t^2-1.24\times10^{-4}t^3+2.15\times10^{-6}t^4$	0.998
Feed	$y=\begin{cases} 0 & t\le 9 \\ 4.42\times10^{-1}t-4.07\times10^{-2}t^2+1.78\times10^{-3}t^3-2.94\times10^{-5}t^4 & t>9 \end{cases}$	0.998

Table S7 By-products in the different models for fed-batch fermentation

Model	Glycerol Control	Formic acid (g/L)	Acetic acid (g/L)	Butyric acid (g/L)	Lactic acid (g/L)
Constant-speed fed-batch	22 mL/h	12.53±0.44	4.91±0.06	7.06±0.00	0.39±0.00
	33 mL/h	11.32±0.15	4.36±0.10	5.50±0.21	1.00±0.00
	44 mL/h	10.11±0.06	3.69±0.01	4.55±0.02	1.12±0.12
	55 mL/h	8.26±0.46	3.79±0.11	3.77±0.08	1.50±0.00
	66 mL/h	7.23±0.01	2.95±0.00	3.72±0.13	1.73±0.17
Pulsed fed-batch	/	7.89±0.00	3.20±0.01	3.79±0.01	0.43±0.00
Variable-speed fed-batch	5 g/L	10.24±0.13	6.16±0.14	6.37±0.24	0.29±0.02
	10 g/L	12.78±0.01	6.34±0.03	8.03±0.04	0.35±0.02
	15 g/L	13.84±0.12	7.74±0.15	8.66±0.30	0.21±0.00
Artificial intelligence fed-batch	5 g/L	15.08±0.35	7.56±0.34	11.90±0.01	0.31±0.00
	10 g/L	14.31±0.17	7.01±0.01	8.97±0.01	0.54±0.01
	15 g/L	13.27±0.44	6.30±0.20	6.31±0.06	0.35±0.01

Table S8 Parameter comparisons of the different models for fed-batch fermentation

Model	Glycerol Control	ATP[1] (mol)	NADH[2] (mol)	Material Cost[3] (¥/$kg_{1,3\text{-}PDO}$)	Substrate utilization rate[4] (%)

Constant-speed fed-batch	22 mL/h	0.99	0.74	16.21	92.2 %
	33 mL/h	0.87	0.65	19.66	61.1 %
	44 mL/h	0.80	0.61	20.43	62.5 %
	55 mL/h	0.69	0.53	20.66	71.4 %
	66 mL/h	0.72	0.55	26.64	46.7 %
Pulsed fed-batch	/	0.56	0.42	15.73	88.7 %
Variable-speed fed-batch	5 g/L	1.08	0.82	14.20	80.0 %
	10 g/L	1.22	0.92	13.34	98.2 %
	15 g/L	1.45	1.10	17.77	92.8 %
Artificial intelligence fed-batch	5 g/L	1.68	1.25	12.43	96.6 %
	10 g/L	1.49	1.13	14.35	89.8 %
	15 g/L	1.03	0.77	18.67	78.0 %

ATP[1](mol): ATP produced from glycerol oxidative pathway, =Lactic acid (mol) + Acetic acid (mol) × 2 + Butyric acid (mol) × 3

NADH[2] (mol): NADH produced from glycerol oxidative pathway, =Lactic acid (mol) + Acetic acid (mol) × 2 + Butyric acid (mol)× 1.5

Medium Cost[3] was calculated based on the unit price for medium component (Table S9)

Substrate utilization rate (%)[4]= (1 - Glycerol remnant (g) / Glycerol total (g)) × 100%

Table S9 Unit price for medium component

Medium component	Unit price (¥/kg)

Crude glycerol	3.50
Yeast extract	80.00
$(NH_4)_2SO_4$	1.80
$MgSO_4 \cdot 7H_2O$	1.60
KH_2PO_4	3.30
$K_2HPO_4 \cdot 3H_2O$	8.50
$CaCl_2$	0.85
NaOH	0.80

References

1. Papoutsakis E T. Equations and calculations for fermentations of butyric acid bacteria. Biotechnology and Bioengineering, 2000, 67(6): 813-826.
2. Celińska E. Debottlenecking the 1, 3-propanediol pathway by metabolic engineering. Biotechnology Advances, 2010, 28(4): 519-530.
3. Cai G, Jin B, Saint C, et al. Metabolic flux analysis of hydrogen production network by *Clostridium butyricum* W5: effect of pH and glucose concentrations. International Journal of Hydrogen Energy, 2010, 35(13): 6681-6690.