# A Machine Learning Approach for Efficient Selection of Enzyme Concentrations and Its Application for Flux Optimization

**Anamya Ajjolli Nagaraja [1,2,3,4], Philippe Charton [2,3,4], Xavier F. Cadet [5], Nicolas Fontaine [5], Mathieu Delsaut [1], Birgit Wiltschi [6], Alena Voit [6], Bernard Offmann [7], Cedric Damour [1], Brigitte Grondin-Perez [1] and Frederic Cadet [2,3,4,*]**

Laboratory of Energy, Electronics and Processes (LE2P-EnergyLab), EA 4079, Faculty of Sciences and Technology, University of La Reunion, St Denis Cedex 97444, France; anamya.ajjolli-nagaraja@univ-reunion.fr (A.A.N.); mathieu.delsaut@univ-reunion.fr (M.D.); cedric.damour@univ-reunion.fr (C.D.); Brigitte.Grondin@univ-reunion.fr (B. G.-P.)

2 University of Paris, BIGR—Biologie Intégrée du Globule Rouge, Inserm, UMR_S1134, Paris F-75015, France; philippe.charton@univ-reunion.fr (P.C.); frederic.cadet.run@gmail.com (F.C.)

3 Laboratory of Excellence GR-Ex, Paris, France

4 DSIMB, UMR_S1134, BIGR, Inserm, Faculty of Sciences and Technology, University of La Reunion, Saint-Denis F-97715, France

5 PEACCEL—Protein Engineering Accelerator, 6 Square Albin Cachot, box 42, Paris 75013, France; xavier.cadet.fjf@gmail.com (X.F.C.); nicolas.fontaine@peaccel.com (N.F.)

6 ACIB—Austrian Centre of Industrial Biotechnology, Synthetic Biology Group, Petersgasse 14, Graz 8010, Austria; birgit.wiltschi@acib.at (B.W.); alenavoit@acib.at (A.V.)

7 Unité Fonctionnalité et Ingénierie des Protéines (UFIP), UFR Sciences et Techniques, Université de Nantes, UMR 6286 CNRS, 2, Chemin de la Houssinière, Nantes Cedex 03 44322, France; bernard.offmann@univ-nantes.fr

* Correspondence: frederic.cadet.run@gmail.com; Tel.: +33695918108

Supplementary Table:

Table S 1: Enzymes used in this study for the upper part of glycolysis. All enzymes were from Sigma.

| Enzyme | EC | Acronym | Source | Sigma Cat. - No. | Lot No. | Purity | MW | active enzyme | MW/subunit |
|---|---|---|---|---|---|---|---|---|---|
| Hexokinase | 2.7.1.1 | HK | baker's yeast | H4502-1KU | SLBT5451 | mix of isoenzymes HXK1 & HXK2 | 110.0 kDa | homodimer | 54 kDa |
| Phosphoglucoisomerase | 5.3.1.9 | PGI | baker's yeast | P5381-1KU | SLBW8689 | n.i. | 119.5 kDa | homodimer | 61.3 kDa |
| Phosphofructokinase | 2.7.1.11 | PFK | Bacillus stearothermophilus | F0137-100UN | SLBW6641 | n.i. | 136.5 kDa | homotetramer | 34 kDa |
| Fructose biphosphate aldolase | 4.1.2.13 | FBA | rabbit muscle | A2714-100UN | SLBR7752V SLBV7445 | >80% | 157.4 kDa | homotetramer | 39.3 kDa |
| Triose phosphate isomerase | 5.3.1.1 | TPI | Baker's yeast | T2507-5MG | 036H8025 | n.i. | 53.6 kDa | homodimer | 26.8 kDa |
| Glycerol-3-phosphate dehydrogenase | 1.1.1.8 | GDH | rabbit muscle | 10127752001 | 21866328 | traces of other enzymes | 75.2 kDa | homodimer | 37.6 kDa |
| Glyceraldehyde-3-phosphate dehydrogenase | 1.2.1.12 | GAPDH | rabbit muscle | G2267-1KU | SLBR0602V | >80% | 144.0 kDa | homotetramer | 36 kDa |
| Creatine kinase | 2.7.3.2 | CK | rabbit muscle | 10127566001 | 25998433 | traces of other enzymes | 86.2 kDa | homodimer | 43.1 kDa |
| Glucose-6-phosphate dehydrogenase | 1.1.1.49 | G6PDH | baker's yeast | G6378-250UN | SLBP6152V | -- | -- | -- | -- |

Table S 2: The measured enzyme activities for the enzymes involved in the upper part of glycolysis (see also Table 2 in the main text).

| **Enzyme** | **EC number** | **origin** | **specific activity (U/mg)** | **Comments** |
|---|---|---|---|---|
| Phosphoglucoisomerase | 5.3.1.9 | Baker's yeast | 556 | this study |
| Phosphofructokinase | 2.7.1.11 | *Bacillus stearothermophilus* | 73 | this study |
| Fructose biphosphate aldolase | 4.1.2.13 | Rabbit muscle | 10 | this study |
| Triose phosphate isomerase | 5.3.1.1 | Baker's yeast | 9500 | Manufacturer value |

Table S 3: The enzyme concentrations (mg/l) predicted from ANN and in-silico modelling to have higher flux values. For the experimental validation, we used relative concentrations of enzymes obtained as explained in Method S 1.

| Index | PGI mg/l | PFK mg/l | FBA mg/l | TPI mg/l | Index | PGI mg/l | PFK mg/l | FBA mg/l | TPI mg/l |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 70 | 5 | 5 | 21.9 | 22 | 4 | 11 | 85.24 | 1.66 |
| 2 | 33 | 1 | 66.23 | 1.66 | 23 | 4 | 16 | 80.24 | 1.66 |
| 3 | 55 | 7.5 | 22.5 | 16.9 | 24 | 4 | 16 | 79.24 | 2.66 |
| 4 | 4.23 | 2.62 | 76.42 | 18.62 | 25 | 5 | 15 | 80.24 | 1.66 |
| 5 | 7.36 | 3.21 | 86.61 | 4.72 | 26 | 5 | 16 | 79.24 | 1.66 |
| 6 | 5.38 | 3.01 | 86.61 | 6.9 | 27 | 5 | 16 | 78.24 | 2.66 |
| 7 | 8.62 | 3.47 | 86.61 | 3.19 | 28 | 5 | 16 | 77.24 | 3.66 |
| 8 | 28.1 | 8.42 | 50.95 | 14.43 | 29 | 6 | 15 | 79.24 | 1.66 |
| 9 | 22.56 | 7.84 | 50.95 | 20.55 | 30 | 6 | 15 | 78.24 | 2.66 |
| 10 | 19.19 | 17 | 56.04 | 9.67 | 31 | 6 | 15 | 77.24 | 3.66 |
| 11 | 2 | 10 | 88.24 | 1.66 | 32 | 6 | 16 | 78.24 | 1.66 |
| 12 | 2 | 10 | 86.24 | 3.66 | 33 | 6 | 16 | 77.24 | 2.66 |
| 13 | 2 | 11 | 82.24 | 6.66 | 34 | 7 | 12 | 78.24 | 4.66 |
| 14 | 2 | 12 | 80.24 | 7.66 | 35 | 7 | 15 | 78.24 | 1.66 |
| 15 | 2 | 13 | 85.24 | 1.66 | 36 | 7 | 15 | 77.24 | 2.66 |
| 16 | 2 | 14 | 84.24 | 1.66 | 37 | 7 | 16 | 77.24 | 1.66 |
| 17 | 2 | 15 | 83.24 | 1.66 | 38 | 8 | 13 | 79.24 | 1.66 |
| 18 | 2 | 16 | 78.24 | 5.66 | 39 | 8 | 15 | 77.24 | 1.66 |
| 19 | 3 | 10 | 85.24 | 3.66 | 40 | 9 | 12 | 78.24 | 2.66 |
| 20 | 3 | 12 | 85.24 | 1.66 | 41 | 10 | 12 | 78.24 | 1.66 |
| 21 | 3 | 16 | 80.24 | 2.66 | | | | | |

Table S 4: Specification of enzymes used for the calculation of cost for the preparatory stage of glycolysis from sigma. Specific activities are calculated by Fievet et al.

| Enzyme | Origin | Price (EUR) | Sold units (kU) | Specific activity (calc.) (U/mg) | Units (calc.) (kU) | MW (active enzyme) (kDa) |
|---|---|---|---|---|---|---|
| Phosphoglucoisomerase | Baker's yeast | 78.50 | 1.0 | 1370.0 | 1.0 | 119.5 |
| Phospho-fructokinase | *Bacillus stearothermophilus* | 178.00 | 0.1 | 70.0 | 0.1 | 136.5 |
| Fructose biphosphate aldolase | rabbit muscle | 48.75 | 0.1 | 42.0 | 0.1 | 157.4 |
| Triose phosphate isomerase | Baker's yeast | 146.00 | n.a. | 14690.0 | 50.0 | 53.6 |

Table S 5: Comparison of flux predicted between Fievet et al. selected concentration (JFievet) and new estimation during current work (Jobs).

| Index | PGI (mg/l) | PFK (mg/l) | FBA (mg/l) | TPI (mg/l) | $J_{obs}$ [μM $s^{-1}$] | $J_{Fievet}$ [μM $s^{-1}$] |
|---|---|---|---|---|---|---|
| 1 | 70.0 | 5.0 | 5.0 | 21.9 | 0.59 | 1.22 |
| 10 | 19.19 | 17.0 | 56.4 | 9.67 | 8.03 | 11.05 |

Table S 6: The calculated price for the μM of NADH consumed per second by the enzyme concentration selected for the experiment.

| Index | mg/l | | | | μM/s | | in EUR |
|---|---|---|---|---|---|---|---|
| | PGI | PFK | FBA | TPI | $J_{ANN}$ | $J_{EXP}$ | Price per μM |
| 11 | 2 | 10 | 88.24 | 1.66 | 12.24 | 15.7 | 0.213 |
| 12 | 2 | 10 | 86.24 | 3.66 | 12.06 | 16.3 | 0.208 |
| 13 | 2 | 11 | 82.24 | 6.66 | 12 | 12.1 | 0.294 |
| 14 | 2 | 12 | 80.24 | 7.66 | 12.03 | 16.6 | 0.222 |
| 15 | 2 | 13 | 85.24 | 1.66 | 12.7 | 13.9 | 0.263 |
| 16 | 2 | 14 | 84.24 | 1.66 | 12.74 | 18.3 | 0.205 |
| 17 | 2 | 15 | 83.24 | 1.66 | 12.72 | 17.1 | 0.226 |
| 18 | 2 | 16 | 78.24 | 5.66 | 12.16 | 20.1 | 0.202 |
| 19 | 3 | 10 | 85.24 | 3.66 | 12 | 14.4 | 0.241 |
| 20 | 3 | 12 | 85.24 | 1.66 | 12.53 | 15.8 | 0.230 |
| 21 | 3 | 16 | 80.24 | 2.66 | 12.44 | 20.6 | 0.198 |
| 22 | 4 | 11 | 85.24 | 1.66 | 12.32 | 15.4 | 0.235 |
| 23 | 4 | 16 | 80.24 | 1.66 | 12.49 | 16.1 | 0.257 |
| 24 | 4 | 16 | 79.24 | 2.66 | 12.36 | 19.3 | 0.216 |
| 25 | 5 | 15 | 80.24 | 1.66 | 12.48 | 18.5 | 0.223 |
| 26 | 5 | 16 | 79.24 | 1.66 | 12.41 | 17.8 | 0.237 |
| 27 | 5 | 16 | 78.24 | 2.66 | 12.29 | 16.3 | 0.261 |
| 28 | 5 | 16 | 77.24 | 3.66 | 12.18 | 19.7 | 0.217 |
| 29 | 6 | 15 | 79.24 | 1.66 | 12.41 | 17.8 | 0.237 |
| 30 | 6 | 15 | 78.24 | 2.66 | 12.29 | 19 | 0.223 |
| 31 | 6 | 15 | 77.24 | 3.66 | 12.19 | 21 | 0.203 |
| 32 | 6 | 16 | 78.24 | 1.66 | 12.34 | 15.6 | 0.277 |
| 33 | 6 | 16 | 77.24 | 2.66 | 12.23 | 17.8 | 0.244 |
| 34 | 7 | 12 | 78.24 | 4.66 | 12 | 17.1 | 0.237 |
| 35 | 7 | 15 | 78.24 | 1.66 | 12.33 | 17.7 | 0.243 |
| 36 | 7 | 15 | 77.24 | 2.66 | 12.22 | 18.8 | 0.230 |
| 37 | 7 | 16 | 77.24 | 1.66 | 12.27 | 20.4 | 0.216 |
| 38 | 8 | 13 | 79.24 | 1.66 | 12.26 | 15.9 | 0.263 |
| 39 | 8 | 15 | 77.24 | 1.66 | 12.26 | 17.9 | 0.245 |
| 40 | 9 | 12 | 78.24 | 2.66 | 12.04 | 15.8 | 0.265 |
| 41 | 10 | 12 | 78.24 | 1.66 | 12.05 | 13.6 | 0.312 |

**Supplementary Figures:**

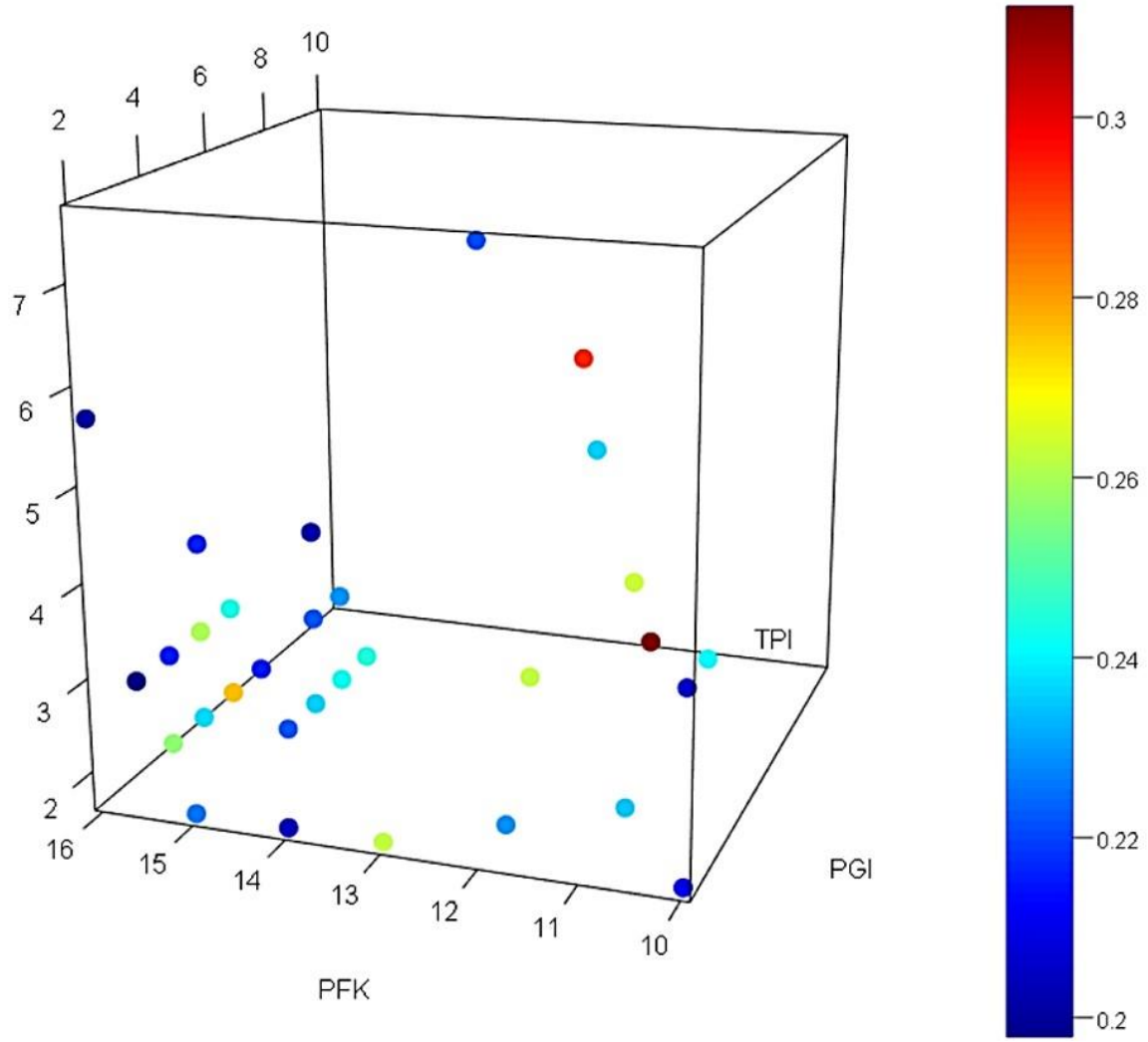


Figure S 1: The cost predicted (in EUR) for the four-enzyme concentration (PGI, PFK, FBA and TPI) selected for experimental validation. The blue is lowest, to highest in red.

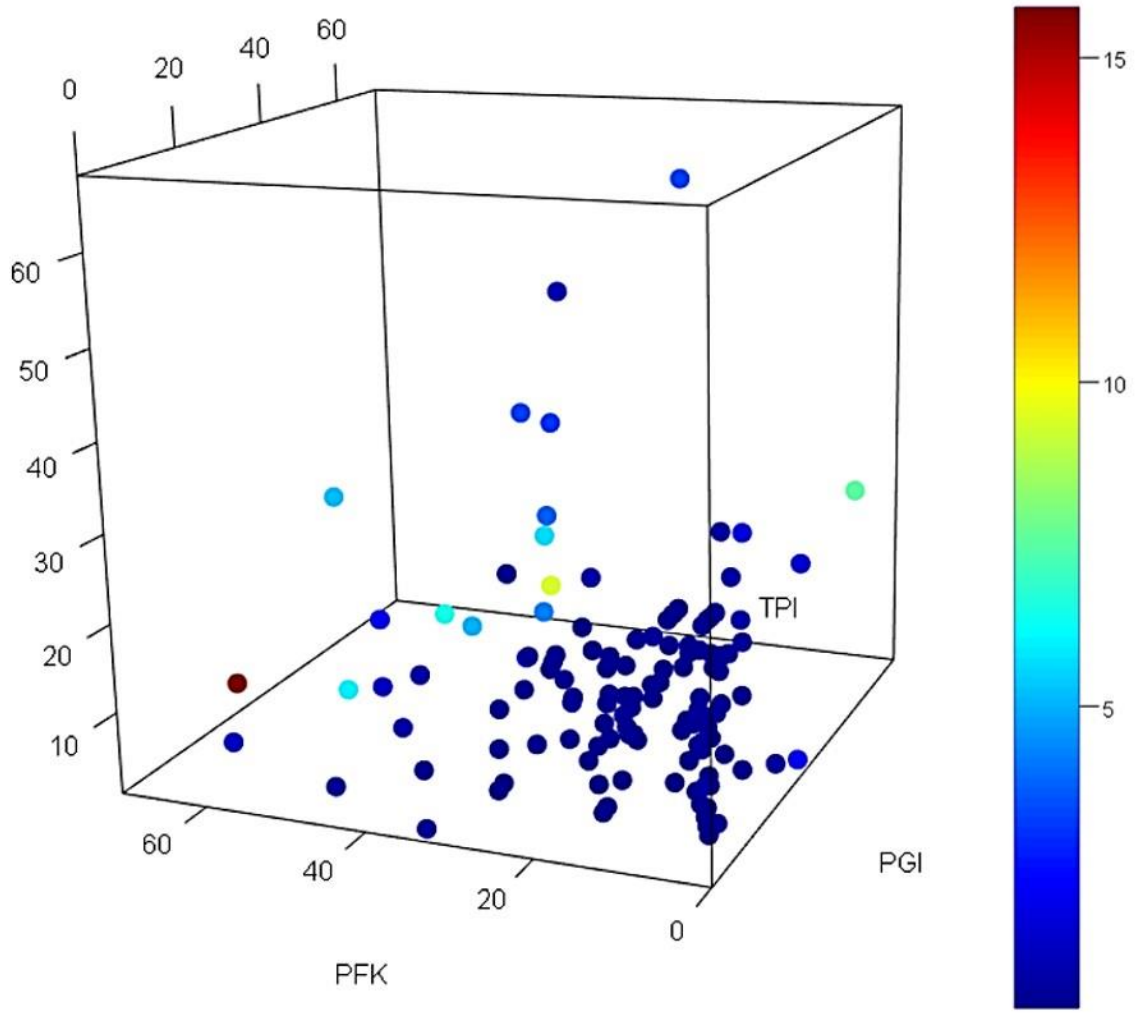


Figure S 2: The cost predicted (in EUR) for the four-enzyme concentration (PGI, PFK, FBA and TPI) selected by Fievet et al. (2006). The blue is lowest, to highest in red.

**Supplementary methods:**

**Methods S1: concentration based on relative activity:**

Our new methodology predicts the flux through the upper part of glycolysis based on the concentrations of the four enzymes, PGI, PFK, FBA and TPI. To make our prediction comparable to that of Fiévet et al, it was necessary to employ relative enzyme activities rather than enzyme concentrations. Depending on the specific activity of the enzyme preparation, the concentration of the enzyme represents a particular activity, which can vary from batch to batch. To account for this, we took the enzyme concentrations listed in Fiévet et al. (see Table S 3, indices 1-10) and transformed them into enzyme activities (see Equation 1) by employing the specific activities that were indicated in the paper (see also Table S 4). Then, we used the specific activities that we had assessed for the enzymes (see Table S 2 and Table 2 of the main text) and transformed the enzyme activities back into enzyme concentrations. These concentrations are indicated in Table 2 of the main text. The enzyme concentrations in 2 were used for the prediction (ANN predicted flux, $J_{ANN}$ and simulated flux $J_{copasi}$) while the

concentrations indicated in Table 2, index 11-41 were used for the experimental assessment of the flux.

$$\text{enzyme concentration } \left(\frac{\text{mg}}{\text{L}}\right) = 1000 * \frac{enzyme\ activity\ (\frac{U}{mL})}{specific\ acitvity\ (\frac{U}{mg})} \text{ ---------- Equation 1}$$

$$\text{C } \left(\frac{\text{mg}}{\text{L}}\right) = 1000 * \frac{U_v\ (\frac{U}{mL})}{U_s\ (\frac{U}{mg})} \text{---------- Equation 1}$$

C, enzyme concentration (mg/l); $U_v$, enzyme activity per volume (U/mL); $U_s$, specific enzyme activity (U/mg)

**Method S2: Cost Calculation**

We predict the cost for μM/s of flux through the pathway as follows:

**Cost per 1U of enzyme:** For each enzyme (PGI, PFK, FBA, TPI) the cost was calculated as below using theTable S 4.

$$P_U\ (\frac{\text{EUR}}{\text{U}}) = \frac{Price\ (EUR)}{Units\ sold\ (U)}$$

$P_U$, price per unit

**Cost per reaction through the whole pathway:** the cost per 1 ml of reaction is calculated as follows:

$$C_R\ (EUR) = \sum(U_v * P_U) \text{ ----- Equation 3}$$

$C_R$, cost per reaction; $U_v$, enzyme activity per reaction volume (U/mL)

**Cost per one μM/s flux:** Cost for the conversion of 1 μM NADH in 1 second is calculated using Equation 4:

$$C_{flux}\ (EUR/\frac{\mu M}{s}) = \frac{C_R\ (EUR)}{f\ (\frac{\mu M}{s})}$$

$C_{flux}$, cost per flux of 1 μM/s; f, estimated flux (μM/s)