Second Report on Nutrient Limitations in M9-medium for the E.coli iEC1364_W - P.putida iJN1463 consortium - uptake rates for limiting nutrients

Iván Martín Martín

02/04/2021

0. Preliminary Notes

The current analysis was performed to finally assess the effect of nutrient limitations (nitrogen, phosphate) in naringenin production for the E.coli iEC1364_W - P.putida iJN1463 consortium, using the M9 medium.

In this case, nutrient concentrations were constant for all executions (M9 layout), and the optimization strategy consisted on considering different NH_4 and P_i uptake rates as input parameters for SMAC. The **original parameters used in the SMAC optimization** were (`params.pcs` file):

- p1_sucr1 ordinal {-10,-8,-6,-4,-2} [-2]
- p2_biomassEc ordinal {0.05, 0.1, 0.15, 0.20, 0.25} [0.1]
- p3_frc2 ordinal {-20,-18,-16,-14,-12,-10,-8,-6,-4,-2} [-14]
- p4_biomassKT ordinal {0.05, 0.1, 0.15, 0.20, 0.25} [0.1]
- p5_nh4_Ec ordinal {-6, -4, -3, -2, -1} [-4]
- p6_nh4_KT ordinal {-12, -10, -8, -6, -4, -2} [-10]
- p7_pi_Ec ordinal {-0.5, -0.4, -0.3, -0.2, -0.1} [-0.3]
- p8_pi_KT ordinal {-1, -0.75, -0.50, -0.25, -0.1} [-0.75]

NH_4 and P_i uptake rates ranges might be changed in further FLYCOP executions, as part of the guided optimization described in this report.

Moreover, other **layout parameters** to note were:

- **Grid Size (inoculation point)**: 1, 1 (0, 0).
- **Number of cycles**: 240.
- `maxSpaceBiomass` **parameter**: 10.0 gL^{-1}.

The **fitness function** is represented here as **naringenin yield** (mM / gL^{-1}): total production of this product (mM), divided by final KT biomass (gL^{-1}). Note that all scenarios have been run with a **SD (standard deviation) cutoff value of (0.1) · (avgfitness)**.

It is worth noticing that the **main concerns of this analysis** were the following ones.

- Finding a range of optimal NH_4 and P_i uptake rates for the consortium to be optimized, in order to allow naringenin production under M9 medium conditions.
- Assessing the effect of M9 medium nutrient limitations in naringenin production (either positive, negative or neutral).
- Avoiding biomass loss (as happened in last report FLYCOP runs, "Report on Nutrient Limitations in M9-medium for the E.coli iEC1364_W - P.putida iJN1463 consortium") because of nutrient limitations.

1. Base Context FLYCOP run

A first FLYCOP run was executed with the parameters stated above, in order to find a preliminary subrange of optimal values for NH_4 and P_i uptake rates for subsequent FLYCOP runs. The general distribution of configurations were as follows.

- **10 cases of `ZeroDivisionError`**.

- **44 cases of a different error, `Non-optimal solution obtained`**. These cases were included in the final datatable of "configurationsResults_Scenario0.txt".

    - 7 of these cases did not experience biomass loss.
    - 37 of these cases did experience biomass loss.

- **46 acceptable cases**, without COBRA errors.

    - 5 of these cases did not experience biomass loss (none of which with excessive SD).
    - 41 of these cases did experience biomass loss.

Given these results, the performed analysis consisted on comparing a series of input and output parameters for each set of configurations (boxplot+scatter).

- **Acceptable cases**. Further subdivision, biomass loss *vs.* non-biomass loss configurations.
- **NonOptimalConfig cases**. Further subdivision, biomass loss *vs.* non-biomass loss configurations.

Note that an excessive SD was *not* considered as criteria for removing a given configuration from the analysis. However, it might be required to consider in further FLYCOP executions. In any case, there were just 10 cases of excessive SD: 5 of them, *NonOptimalConfig* cases with biomass loss; the remaining 5, *Acceptable* cases with biomass loss.

1.1. NH_4 uptakes rates

These values are negative under COBRA / SMAC optimization. Please, consider the absolute value while reading these interpretations.

1.1.1. For Escherichia coli

For those configurations with **NonOptimalConfig**, NH_4 uptake rates were:

- evenly distributed from -6 to -1 in the case of *biomass loss*.
- evenly distributed from -4 to -1 in the case of *non biomass loss*. They tended to lower values (near -1).

For those **Acceptable** configurations, NH_4 uptake rates were:

- evenly distributed from -6 to -2 in the case of *biomass loss*.
- concentrated in -3 (3 cases), except for one case in -2 and another case in -6 (outliers); in the case of *non biomass loss*.

Therefore, a very low NH_4 uptake rate for *E.coli* (-2 or higher, negative value) might lead to non-optimal solutions. Moreover, a value under -4 (negative value) might also lead to a non-optimal solution and/ or to biomass loss. A potential better subrange for this parameter (`NH4_Ec`) would be (-2, -4). Note that this parameter might also be interrelated to other input parameter values: this individual effect of NH_4 uptake rate for *E.coli* would be just a partial cause for the outcome obtained.

1.1.2. For Pseudomonas putida

For those configurations with **NonOptimalConfig**, NH_4 uptake rates were:

- evenly distributed from -12 to -2 in the case of *biomass loss*.
- evenly distributed from -10 to -2 in the case of *non biomass loss*.

For those **Acceptable** configurations, NH_4 uptake rates were:

- evenly distributed from -12 to -2 in the case of *biomass loss.*
- concentrated in -12 to -10 (4 cases), except for one case in -4 (outlier); in the case of *non biomass loss.*

Thus a low NH_4 uptake rate for *P.putida* (above -10 or -8, negative value) might lead to non-optimal solutions, or acceptable solutions but with biomass loss. A potential better subrange for this parameter (NH4_KT) would be (-10, -12). Note that this parameter might also be interrelated to other input parameter values: this individual effect of NH_4 uptake rate for *P.putida* would be just a partial cause for the outcome obtained.

1.2. P_i uptakes rates

1.2.1. For Escherichia coli

For those configurations with **NonOptimalConfig**, P_i uptake rates were:

- distributed from -0.50 to -0.10 in the case of *biomass loss*, but more concentrated in the rank of -0.30 to -0.10.
- evenly distributed from -0.40 to -0.10 in the case of *non biomass loss.*

For those **Acceptable** configurations, P_i uptake rates were:

- evenly distributed from -0.50 to -0.30 in the case of *biomass loss.* Certain outliers around -0.20.
- dispersed to some extent, between -0.50 to -0.20, in the case of *non biomass loss.*

Therefore, P_i uptake for *E.coli* might not be specially influential in the outcome of the configuration, though a too low value (above -0.20, negative value) might not be a good choice (leading to non-optimal solutions). An acceptable subrange for this parameter (Pi_Ec) would be (-0.20, -0.50). Again, note the inter-relation between input parameter values for the unbiased interpretation of the final outcome obtained.

1.2.2. For Pseudomonas putida

For those configurations with **NonOptimalConfig**, P_i uptake rates were:

- evenly distributed from -1.0 to -0.1 in the case of *biomass loss.*
- evenly distributed from -1.0 to -0.1 in the case of *non biomass loss.* The value of -0.1 could be considered an outlier.

For those **Acceptable** configurations, NH_4 uptake rates were:

- evenly distributed from -1.0 to -0.1 in the case of *biomass loss.*
- dispersed to some extent, between -1.0 to -0.30, in the case of *non biomass loss.*

Thus P_i uptake for *P.putida* might not be specially influential in the outcome of the configuration. An acceptable subrange for this parameter (Pi_KT) would be (-0.30, -1.0), avoiding too low values (-0.2 or higher, negative value). Again, note the inter-relation between input parameter values for the unbiased interpretation of the final outcome obtained.

1.3. Carbon source uptake rates

1.3.1. For Escherichia coli

For those configurations with **NonOptimalConfig**, sucrose uptake rates were:

- evenly distributed from -10 to -2 in the case of *biomass loss.*
- taking values of -8 (2 cases) and -6 (5 cases), in the case of *non biomass loss.*

For those **Acceptable** configurations, sucrose uptake rates were:

- taking values of -2 and -4, in the case of *biomass loss.* Certain outliers around -8.0.

- esentially concentrated in -2, in the case of *non biomass loss*. One outlier in -4.

Therefore, high or very high values of sucrose uptake for *E.coli* (under -4, negative values) might lead to non-optimal solutions. A potential better subrange for this parameter (`sucr1`) would be (-2, -4) - lower values might also avoid biomass loss in acceptable configurations. Again, note the inter-relation between input parameter values for the unbiased interpretation of the final outcome obtained.

1.3.2. For Pseudomonas putida

For those configurations with **NonOptimalConfig**, fructose uptake rates were:

- evenly distributed from -20 to -2.5 in the case of *biomass loss*.
- evenly distributed from -7.5 to -2.5 in the case of *non biomass loss*.

For those **Acceptable** configurations, fructose uptake rates were:

- evenly distributed from -20 to -2.5 in the case of *biomass loss*.
- dispersed to some extent, between -16 to -4, in the case of *non biomass loss*.

Thus a less error-prone subrange of fructose uptake rates might be (-4, -16), avoiding the values in the extremes of the interval considered so far. However, this value rate might be inter-related to other input parameters in the determination of the final outcome of a given configuration.

1.4. Final sucrose concentration (consumed by E.coli)

For those configurations with **NonOptimalConfig**, final sucrose concentrations were:

- evenly distributed from 100 to 0 (mM) in the case of *biomass loss*.
- evenly distributed from 90 to 60 (mM) in the case of *non biomass loss*. One outlier around 30 mM.

For those **Acceptable** configurations, final sucrose concentrations were:

- evenly distributed from 90 to 0 (mM) in the case of *biomass loss*.
- dispersed to some extent, between 85 to 10 (mM), in the case of *non biomass loss*.

Thus in those cases with no biomass loss, final sucrose concentrations were never exhausted, but there are some configurations with sucrose exhaustion within the sets of biomass loss for both acceptable and NonOptimalConfig groups.

1.5. Initial Biomass values

1.5.1. For Escherichia coli

For those configurations with **NonOptimalConfig**, initial *E.coli* concentrations were:

- evenly distributed from 0.05 to 0.25 (gL^{-1}) in the case of *biomass loss*.
- evenly distributed from 0.05 to 0.15 (gL^{-1}) in the case of *non biomass loss*. One outlier around 0.25 gL^{-1}.

For those **Acceptable** configurations, initial *E.coli* concentrations were:

- evenly distributed from 0.05 to 0.25 (gL^{-1}) in the case of *biomass loss*.
- 2 points taking a value of 0.05 gL^{-1}, 2 points in 0.10 gL^{-1}, 1 outlier (0.25 gL^{-1}); in the case of *non biomass loss*.

Thus these configurations with nitrogen limitation and restricted NH_4 uptake rates benefit from lower values of initial *E.coli* biomass (0.05 or lower, 0.1) (gL^{-1}).

1.5.2. For Pseudomonas putida

For those configurations with **NonOptimalConfig**, initial *P.putida* concentrations were:

- evenly distributed from 0.05 to 0.25 (gL^{-1}) in the case of *biomass loss*.
- evenly distributed from 0.05 to 0.20 (gL^{-1}) in the case of *non biomass loss*. One outlier around 0.25 gL^{-1}.

For those **Acceptable** configurations, initial *P.putida* concentrations were:

- evenly distributed from 0.05 to 0.25 (gL^{-1}) in the case of *biomass loss*.
- 4 points in 0.05 gL^{-1}, 1 point in 0.1 gL^{-1} (outlier); in the case of *non biomass loss*.

Thus these configurations with nitrogen limitation and restricted NH_4 uptake rates also benefit from lower values of initial *P.putida* biomass (0.01 - 0.075) (gL^{-1}).

1.6. Final Biomass values

1.6.1. For Escherichia coli

For those configurations with **NonOptimalConfig**, final *E.coli* concentrations were:

- evenly distributed from 0.00 to 1.00 (gL^{-1}) in the case of *biomass loss*. There was a certain concentration of points near 0.0 gL^{-1}.
- evenly distributed from 0.10 to 0.30-0.40 (gL^{-1}) in the case of *non biomass loss*. One outlier around 1.60 gL^{-1}.

For those **Acceptable** configurations, final *E.coli* concentrations were:

- evenly distributed from 0.00 to 2.00 (gL^{-1}) in the case of *biomass loss*.
- 3 points around 1.0 gL^{-1}, 2 points around 2.0 gL^{-1}; in the case of *non biomass loss*.

Thus the higher final *E.coli* concentrations were obtained for acceptable configurations without biomass loss. In this category there were no values of final *E.coli* near 0.

1.6.2. For Pseudomonas putida

For those configurations with **NonOptimalConfig**, final *P.putida* concentrations were:

- distributed from 0.00 to 1.75 (gL^{-1}) in the case of *biomass loss*. There was a certain concentration of points near 0.0 gL^{-1}.
- evenly distributed from 0.10 to 1.0 (gL^{-1}) in the case of *non biomass loss*.

For those **Acceptable** configurations, final *P.putida* concentrations were:

- evenly distributed from 0.00 to 1.10 (gL^{-1}) in the case of *biomass loss*. There was a certain concentration of points near 0.0 gL^{-1}. -2 points around 0.50 gL^{-1}, 2 points around 0.75 gL^{-1}, 1 point around 1.0 gL^{-1}; in the case of *non biomass loss*.

Thus the higher final *P.putida* concentrations were obtained for acceptable configurations without biomass loss. In this category there were no values of final *P.putida* near 0.

1.7. Final production (naringenin and p-coumarate)

1.7.1. p-Coumarate by Escherichia coli

For those configurations with **NonOptimalConfig**, final pCA concentrations were:

- distributed from 0 to 50 (mM) in the case of *biomass loss*, with a higher concentration of points under 15 mM (values above 35 mM might be considered outliers).

- 5 points 0 or near 0 mM, 1 value around 5 mM, 1 value around 15 mM, in those cases of *non biomass loss.*

For those **Acceptable** configurations, final pCA concentrations were:

- evenly distributed from 0 to 30 mM in the case of *biomass loss.* Higher concentration of points under 10 mM.
- All values in 0 or near 0 mM, in the case of *non biomass loss.*

Thus the best category in terms of final pCA (the lower, the better) would be the one with acceptable configurations without biomass loss.

1.7.2. Naringenin by Pseudomonas putida

For those configurations with **NonOptimalConfig**, final naringenin concentrations were:

- distributed from 0 to 30 (mM) in the case of *biomass loss*, with a higher concentration of points under 10 mM (values above 25 mM can be considered outliers).
- even distribution of points between 10 to 20 mM, with one outlier around 25 mM; in those cases of *non biomass loss.*

For those **Acceptable** configurations, final naringenin concentrations were:

- distributed from 0 to 25 mM in the case of *biomass loss*, with a higher concentration of points under 10 mM. Also, some ouliers from 40 to 100 mM.
- 2 points around 10 mM, 1 point around 25 mM, 1 point around 35 mM, 1 point (outlier) around 80 mM; in the case of *non biomass loss.*

Thus the best category in terms of final naringenin would be the one with acceptable configurations without biomass loss: intermediate concentration, no values near 0 neither too disproportionate.

1.8. Nitrogen final concentrations

This nutrient (NH_4) would be consumed by both microorganisms. For those configurations with **NonOptimalConfig**, final NH_4 concentrations were:

- distributed from 18 to 0 (mM) in the case of *biomass loss.* Higher concentration of points between 10 to 18 mM.
- evenly distributed from 18 to 0 (mM) in the case of *non biomass loss.*

For those **Acceptable** configurations, final NH_4 concentrations were:

- evenly distributed from 18 to 0 (mM) in the case of *biomass loss.* Intermediate cluster of points in 0 mM.
- all values in 0 mM, in the case of *non biomass loss.*

Interestingly, NH_4 was always exhausted in those cases of acceptable configurations *without* biomass loss, as opposed to the rest of categories here considered.

1.9. Phosphate final concentrations

This nutrient (P_i) would be consumed by both microorganisms. For those configurations with **NonOptimalConfig**, final P_i concentrations were:

- distributed from 70 to 66 (mM) in the case of *biomass loss*, with certain outliers between 66 and 50 mM. Some configurations with very subtle production of P_i.
- distributed between 69 to 58 (mM) in the case of *non biomass loss.*

For those **Acceptable** configurations, final P_i concentrations were:

- evenly distributed from 68 to 62 (mM) in the case of *biomass loss*. Certain outliers between 62 and 50 mM.
- 2 points in 67.5 mM, 1 point around 64 mM, 1 point around 60 mM, 1 point around 57.5 mM; in the case of *non biomass loss*.

Therefore, phosphate consumption would be moderate in all configurations here analyzed, being between 1-2 mM to 10 mM during the whole simulation. However, it seemed to be more pronounced in those acceptable configurations without biomass loss, where NH_4 was always finally exhausted.

1.10. Oxygen final concentrations

This nutrient (oxygen) would be consumed by both microorganisms. For those configurations with **NonOptimalConfig**, final oxygen concentrations were:

- distributed from 1000 to 850 (mM) in the case of *biomass loss*, with one outlier in 775 mM.
- evenly distributed between 975 to 925 (mM) in the case of *non biomass loss*.

For those **Acceptable** configurations, final oxygen concentrations were:

- distributed from 1000 to 750 (mM) in the case of *biomass loss*. Higher concentration of points between 1000 and 850 mM.
- 2 points in 950 mM, 1 point around 900 mM, 1 point around 875 mM, 1 point around 825 mM; in the case of *non biomass loss*.

Thus oxygen consumption did not seem to be an important output difference between the categories here considered.

1.11. Fitness (naringenin yield) (mM / gL⁻¹)

Recall that fitness is defined as naringenin yield (mM / gL⁻¹): total production of this product (mM), divided by final KT biomass (gL⁻¹). For those configurations with **NonOptimalConfig**, fitness values were:

- distributed from near 0 to a disproportionate value (not checked) in the case of *biomass loss*. Disproportionate values were the result of intermediate naringenin production with very low *P.putida* final biomass.
- evenly distributed between 10 to 50, with one outlier around 100; in those cases of *non biomass loss*.

For those **Acceptable** configurations, final oxygen concentrations were:

- distributed from near 0 to a disproportionate value (not checked) in the case of *biomass loss*. Disproportionate values were the result of intermediate naringenin production with very low *P.putida* final biomass.
- 2 points in 10, 1 point around 50, 2 points around 75; in the case of *non biomass loss*.

Fitness might not be a good reference for configurations comparison in this FLYCOP run, since it would be moderately or severely influenced by biomass loss and final values of *P.putida* biomass.

1.12. Quick Conclusions and setting a new (optimized) FLYCOP run

The **main conclusions** obtained from the last FLYCOP run were the following ones.

- There were an important number of **non-optimal configurations**, possibly due to limitations related to nitrogen availability and NH_4 uptake rates for both microbes (generally excessive for a restricted NH_4 initial concentration).

- Despite the last statement, NH_4 **was finally exhausted in all 5 acceptable configurations without biomass loss** (pay attention to this detail in FLYCOP runs to come).

- **Phosphate consumption** was moderate in general, though slightly more pronounced in those 5 acceptable configurations without biomass loss.

- *E.coli* development seem to benefit from a **lower sucrose uptake rate**. On the other hand, *P.putida* also experience a more even growth under **a more compact range of fructose uptake rates** (avoid last extreme values set).

- **Initial biomass values** should be between (approx.) 0.05 and 0.1 gL^{-1} for a better evolution of the consortium in time.

- In general, the **main output parameters** (final biomass values, naringenin and pCA final concentrations, final sucrose) had **better results in the group of acceptable configurations without biomass loss**.

- **Fitness values were not very indicative**, since they were somewhat biased by the *non-optimal configuration* effect and biomass loss. Ideally, the significance of this parameter should be regained in further FLYCOP executions.

Given these facts, the **guided optimization** for the FLYCOP parameters in the current report lead to the execution of a new FLYCOP run with the next parameters (SMAC input parameters).

- p1_sucr1 ordinal {-3.5,-3,-2.5,-2,-1.5, -1} [-2]
- p2_biomassEc ordinal {0.01,0.05,0.075,0.1} [0.05]
- p3_frc2 ordinal {-16,-14,-12,-10,-8,-6,-4} [-14]
- p4_biomassKT ordinal {0.01,0.05,0.075,0.1} [0.05]
- p5_nh4_Ec ordinal {-3,-2.5,-2,-1.5} [-3]
- p6_nh4_KT ordinal {-12,-11.5,-11,-10.5,-10} [-12]
- p7_pi_Ec ordinal {-0.5, -0.4, -0.3, -0.2} [-0.3]
- p8_pi_KT ordinal {-1, -0.75, -0.50, -0.3} [-0.75]

The rest of parameters took the same value as in the recently commented FLYCOP run.