

Colin McNabb

Math 2270

Professor Gustafson

Biochemical Pathways as Linear Systems

Introduction:

The human body is an extremely complex system, especially when you look at it from a chemical standpoint. When you zoom in past the ordered structures of organs and organ systems you are confronted with what could be described as chaos. You have thousands of reactions happening constantly within each cell, of which there are trillions, in your body. Every time you make the slightest movement, smell something, see something, or even just simply sit still, your body is working at dizzying rates to keep you alive and able to function. One of the most basic reactions, and the one I will be examining, is cellular metabolism. Cellular metabolism is the process of taking glucose, the preferred source for fuel in the body, and converting it to carbon dioxide, water and, most importantly, Adenosine Triphosphate (ATP). Just about anyone who has taken an introductory course to biology knows the net reaction and overall process for this mechanism, but it is actually much more complex than many people realize. The amazing thing about biochemistry, however, is that it can be easily modeled in a linear system. Using systems of linear equations we can create stoichiometric matrices based on the individual chemical equations and use these matrices to determine the overall net reaction involved in this chemical pathway. The use of computing software, such as Maple or Matlab, has greatly reduced the tedium and time required to model this pathway.

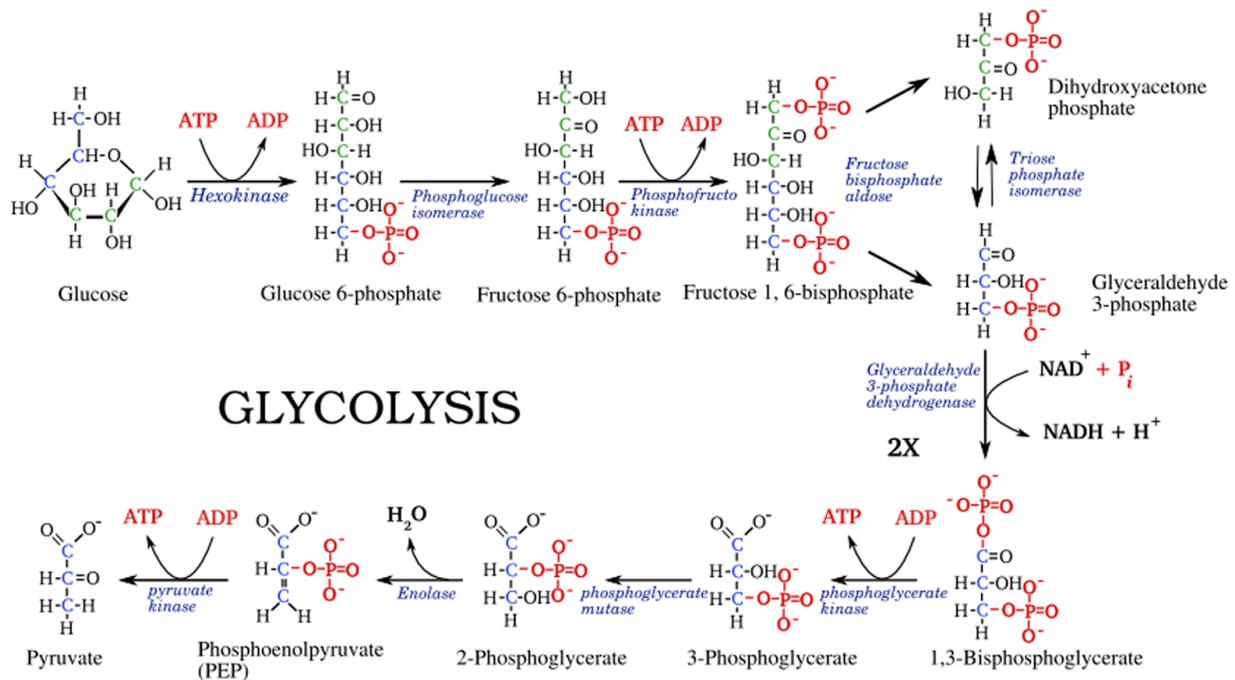
Chemical Reactions of Metabolism:

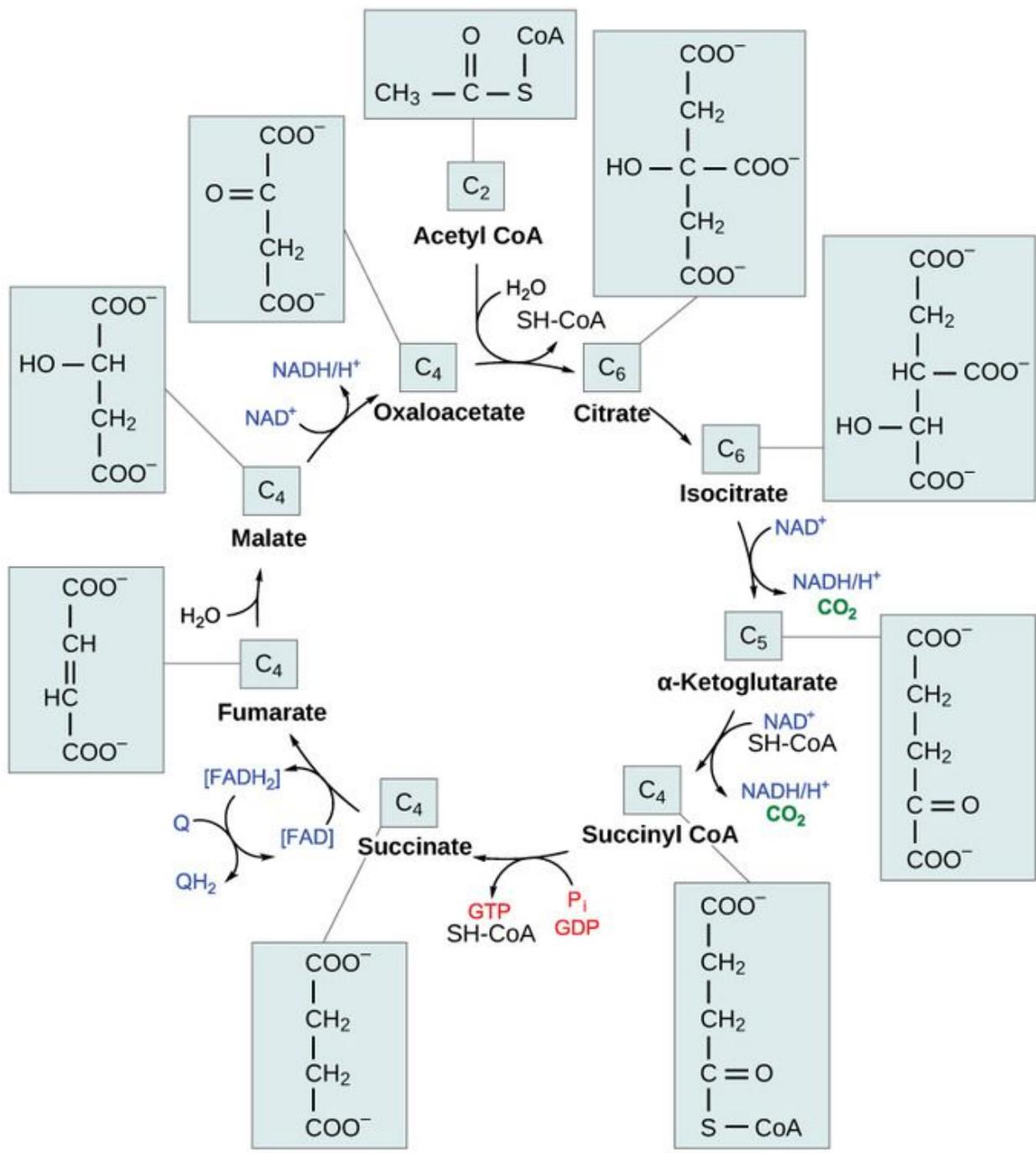
I will briefly explain the chemical pathway of the conversion of Glucose to ATP. There are approximately 21 reactions involved in this chemical pathway:

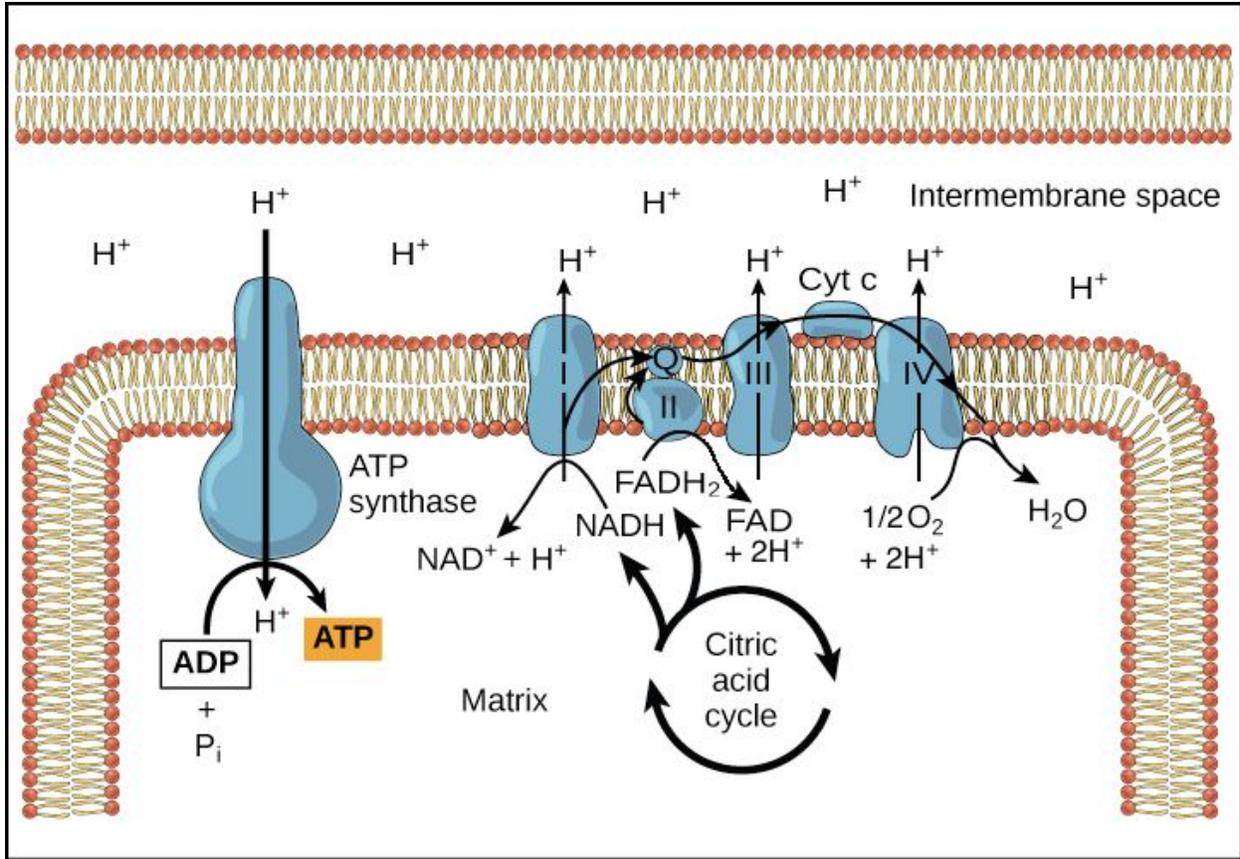
1. glucose + ATP = glucose 6-phosphate + ADP
2. glucose 6-phosphate = fructose 6-phosphate
3. fructose 6-phosphate + ATP = fructose 1,6-bisphosphate+ ADP
4. fructose 1,6-bisphosphate = glyceraldehyde phosphate + glyceraldehyde 3-phosphate
5. glyceraldehyde phosphate = glyceraldehyde 3-phosphate
6. glyceraldehyde 3-phosphate + Pi + NAD⁺ = 3-phospho-D-glyceroyl phosphate + NADH
7. 3-phospho-D-glyceroyl phosphate+ ADP= 3-phospho-D-glycerate + ATP
8. 3-phospho-D-glycerate = 2-phospho-D-glycerate
9. 2-phosphoglycerate = phosphoenolpyruvate +H₂O
10. phosphoenolpyruvate + ADP = pyruvate+ ATP
11. pyruvate + CoA + NAD⁺ = acetyl-CoA + CO₂ + NADH
12. acetyl-CoA + oxaloacetate + H₂O = citrate + CoA
13. citrate = cis-aconitate+ H₂O
14. cis-aconitate+ H₂O = isocitrate
15. isocitrate + NAD⁺ =2-oxoglutarate + CO₂ + NADH
16. 2-oxoglutarate + NAD⁺ + CoA = succinyl-CoA+ CO₂+ NADH
17. succinyl-CoA + Pi + ADP = succinate + ATP + CoA
18. succinate + NAD⁺ = fumarate + NADH
19. fumarate + H₂O = (S)-malate
20. (S)-malate + NAD⁺ = oxaloacetate + NADH
21. NADH + 1/2O₂ + 3Pi + 3ADP = NAD⁺ + 4H₂O + 3ATP

Reactions 1-10 are the steps in of glycolysis, which happens in the cytosol of a cell. And reactions 11-21 are known as the Citric Acid Cycle or the Krebs' Cycle. Glycolysis, in actuality, produces 2 pyruvate per glucose for entry into the citric acid cycle. This means that every step of the Citric Acid Cycle happens twice, yielding 6 CO₂ as byproducts, fully accounting for every carbon in glucose. This can also be seen in the final step, where ½ of an O₂ molecule would be produced. Because you cannot create half a molecule, this reaction happens twice, consuming one full molecule of O₂. For future clarity, Pi is inorganic phosphate which, when bound to another molecule, is a source of a high energy bond, which, when broken, can be a source of

the energy required for the formation of new bonds between substrates; ADP is Adenosine Diphosphate, which is ATP with one phosphate group removed as Pi. The matrices we will produce will be absent of any H⁺ or metal ion quantities because they are often at pre-specified concentrations at points in the body and their overall quantities will not change. Finally, below you will see a graphical representation of glycolysis, the citric acid cycle and the oxidative phosphorylation. Below that we can finally move past the chemistry portion and explore the mathematics behind metabolism!







You will notice the presence of $FADH_2$ which did not appear in the above reactions. This is because $FADH_2$ acts as an electron carrier for $NADH$ generated in glycolysis, which happens in the cytosol of a cell. Because $NADH$ cannot pass through membranes, a process known as the glycerol phosphate shuttle utilizes $FADH_2$ to help move electrons from $NADH$ outside the mitochondrial matrix to $NADH$ inside the mitochondrial matrix.

Mathematics Behind Metabolism:

The calculation of a biochemical pathway through linear algebra is done with a stoichiometric number matrix. This is a matrix which gives the stoichiometric number for the various molecules encountered in an independent reaction. In these matrices, the number of rows is determined by the number of substrates encountered in a pathway, and the number of columns is determined by the number of reactions involved in an overall pathway. In each individual reaction, the reactants will be given a (-) in front of their stoichiometric number because they are being consumed, while the products will have a (+) stoichiometric number. This matrix can then be multiplied by a column vector, which is called the pathway vector. The pathway vector will have a row for each reaction (or column) of the number matrix. For a given set of reactions a solution will not always exist. The solution will only exist when the final column of the number matrix is a linear combination of the columns of the number matrix.

$$\begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} & \dots & \mathbf{V}_{1R} \\ \mathbf{V}_{21} & \mathbf{V}_{22} & \dots & \mathbf{V}_{2R} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \mathbf{V}_{N1} & \mathbf{V}_{N2} & \dots & \mathbf{V}_{NR} \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \\ \dots \\ \mathbf{S}_R \end{bmatrix} = \begin{bmatrix} \mathbf{V}_{net1} \\ \mathbf{V}_{net2} \\ \dots \\ \mathbf{V}_{netN} \end{bmatrix}$$

$$S_1 \begin{bmatrix} V_{11} \\ V_{21} \\ \dots \\ V_{N1} \end{bmatrix} + S_2 \begin{bmatrix} V_{12} \\ V_{22} \\ \dots \\ V_{N2} \end{bmatrix} + \dots + S_R \begin{bmatrix} V_{1R} \\ V_{2R} \\ \dots \\ V_{NR} \end{bmatrix} = \begin{bmatrix} V_{net1} \\ V_{net2} \\ \dots \\ V_{netN} \end{bmatrix}$$

The trick with this system is that it is easy to calculate the net energy production when you can multiply a number matrix by a pathway vector, but what if you don't know the pathway vector? This seems like the kind of roadblock that could stop an aspiring mathematician right in their tracks. Luckily, there is a workaround. We can create a stoichiometric number matrix without any of the reactions involving energetic molecules, namely Pi, ADP, ATP, NADH, NAD⁺. We can then simplify these reactions down to their lowest stoichiometric number, which can provide a pathway vector for the abbreviated matrix. This will not balance any energetic molecules, but it will allow us to balance the carbon, oxygen and sulfur used. This will generate pathway vectors which we can then use to calculate a net energy output for the metabolism depending on starting substrates.

Attached will be the Maple file with the matrices used in the following calculations. The dimensions of the unabbreviated matrix will be 30x21. There are 30 total molecular species and 21 reactions used in the pathway. The abbreviated number matrix will be 25x21. The energetic molecules listed above have been removed from the reactions with some waters added to help balance charges. To save on space, and because this is a lesson mostly on math and not the biochemistry possibilities, I will only include 3 of the possible pathways. One starting from glucose, one starting from a phosphorylated version of glucose, and one coming from an

already partially oxidized species stemming from glucose. The columns of the unabbreviated matrix will go in the following order of molecular species, while the abbreviated matrix will simply follow in the same order with the energetic molecules removed:

- | | |
|-------------------------------|---------------------------------------|
| 1. Glucose | 17. Glyceraldehyde 3-phosphate |
| 2. Pi | 18. 3-phospho-D-glycerol
phosphate |
| 3. ADP | 19. 3-phospho-D-glycerate |
| 4. NAD ⁺ | 20. 2-phospho-D-glycerate |
| 5. Pyruvate | 21. Phosphoenolpyruvate |
| 6. ATP | 22. Oxaloacetate |
| 7. NADH | 23. Citrate |
| 8. H ₂ O | 24. Cis-aconitate |
| 9. CoA | 25. Isocitrate |
| 10. Acetyl-CoA | 26. 2-oxoglutarate |
| 11. CO ₂ | 27. Succinyl-CoA |
| 12. O ₂ | 28. Succinate |
| 13. Glucose 6-phosphate | 29. Fumarate |
| 14. Fructose 6-phosphate | 30. Malate |
| 15. Fructose 1,6-bisphosphate | |
| 16. Glycerone phosphate | |

generates the ATP. So NADH's energy comes not from itself, but from its potential as a reducing agent for the electron transport chain.

The results on the energy output make sense as well. The results in column 1 are the results when the starting material is glucose. This is the compound the body uses so we can use this as a base line. You will notice the results from column 2 have a higher energy output than glucose. This is because the starting compound used in this pathway was Fructose 1,6-bisphosphate, which is a glucose derivative that has been phosphorylated. Because the starting material has already been phosphorylated, the energy input to phosphorylate it has been negated, increasing the net energy output. From the other side, the energy output from column 3 is lower than that of glucose. This is because that starting material, Phosphoenolpyruvate, is a derivative of glucose which has already been partially oxidized so it cannot be oxidized to the same extent of glucose.

Conclusion:

Using the example of metabolism, which is fairly complex but by no means the most complex chemical reaction the body undergoes, we can see that as long as certain aspects of a chemical pathway are known that we can model it in a linear system. Once we have a linear representation of it, we can then use simple matrix properties to solve for any unknown or desired quantities.

References:

Alberty, R.a. "Calculation of Biochemical Net Reactions and Pathways by Using Matrix Operations."

Biophysical Journal 71.1 (1996): 507-15. Web.

Akinola, R. O., S. Y. Kutchin, I. A. Nyam, and O. Adeyanju. "Using Row Reduced Echelon Form in Balancing Chemical Equations." *Advances in Linear Algebra & Matrix Theory* 06.04 (2016): 146-57.

Web.

Rhodes, Carl Douglas, Kirsten Fertuck, P. David Josephy, Roger E. Koeppe, and Jeremy M. Berg. *Biochemistry, Eighth Edition*. New York: W.H. Freeman and, a Macmillan Education, 2015. Print.

Lay, David C., Steven R. Lay, and Judi J. McDonald. *Linear Algebra and Its Applications*. Boston: Pearson, 2016. Print.

"Oxidative Phosphorylation." *Khan Academy*. N.p., n.d. Web. 18 Apr. 2018.

"Glycolysis." *Khan Academy*. N.p., n.d. Web. 18 Apr. 2018.

"The Citric Acid Cycle." *Khan Academy*. N.p., n.d. Web. 18 Apr. 2018.