

# Modelling Carbon Fixation and G-encoded Amino Acids

## Method

The Gillespie algorithm was employed to create stochastic simulations for metabolic pathways. This algorithm accounts for random deviations and thus does not result in deterministic outcomes. Rate constants were initially assigned based on the EC numbers of the reaction, where EC categories most commonly seen in the pathways were assigned higher values. The initial state of all the reactants and products were assigned arbitrarily, with a value of 5000 for the reactants and cofactors and 0 for products. The trajectory was then generated by starting from this initialized state and firing reactions in continuous time, updating the system state at every time step. To do this, two numbers were randomly sampled by the algorithm, one to determine when the next reaction fired, and another to determine which reaction fired next. Simulated trajectories for the production of key species were then plotted over time.

## Preliminary Results

### G-encoded Amino Acids

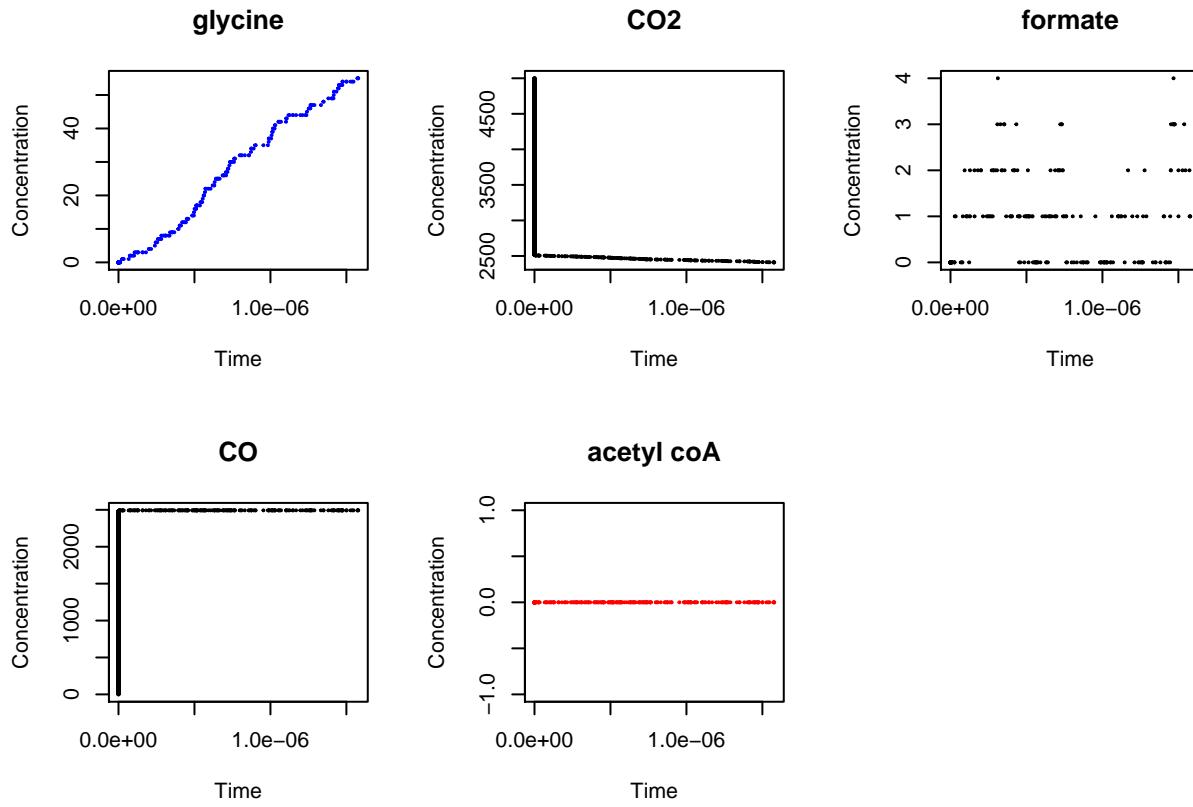
A model was initially created for the synthesis of all G-encoded amino acids, namely glycine, glutamate, alanine, valine, and aspartate, which are produced the earliest in metabolism. The preceding pathways, carbon fixation and incomplete reductive TCA, were also included.

While this initial model resulted in the synthesis of all amino acids, alanine was produced at a concentration two orders of magnitude less than the other four amino acids. This was initially thought to be due to competition for glutamate, which is consumed as a reactant for alanine, valine, and aspartate synthesis. Replacing the glutamate requirement with one for  $\text{NH}_4^+$ , NADP, and  $\text{H}^+$  resulted in increased synthesis of alanine, but synthesis of glycine and glutamate occurred one order of magnitude greater than the three other amino acids.

### Modeling carbon fixation and glycine synthesis

Given this, a model was created for carbon fixation, the earliest pathway in autotrophic metabolism, and glycine synthesis, which immediately follows it. This was done to inspect if earlier intermediates are being produced and consumed consistently with the reactions considered.

The Ljungdahl-Wood pathway was used to model carbon fixation. This pathway occurs in autotrophs, and is best understood in homoacetogenic bacteria, to result in the production of acetyl-coA from two molecules of  $\text{CO}_2$ . Understanding the carbon fixation pathway is key to modelling prebiotic metabolism, as carbon fixation allows for the establishment of autotrophic growth and development by reducing dependence on organic material for energy.



As seen in the figure above, this resulted in low levels of glycine synthesis, lower than in the earlier model which included all the amino acids. Acetyl-coA was not produced at the end of carbon fixation. Most earlier intermediates were also not produced, or produced only at low concentrations, except for carbon monoxide. In order to best explore the reasons behind this, the individual reactions in the carbon fixation pathway were further analyzed.

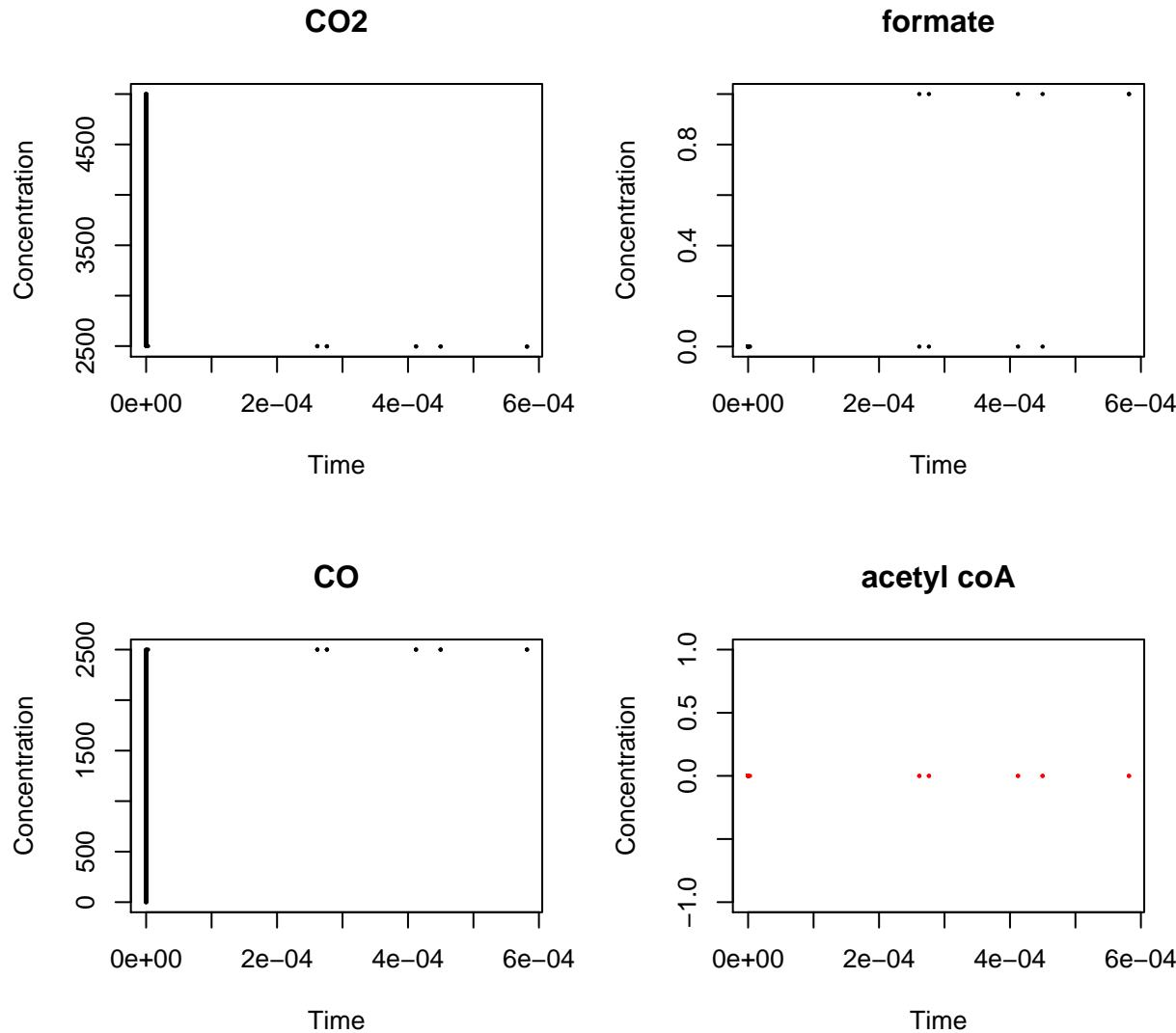
## Carbon fixation

CO<sub>2</sub> is used in two parallel pathways for carbon fixation. The first involves six steps from the conversion of CO<sub>2</sub> to methyl corrinoid (III) FeS clusters. The second involves a single step converting CO<sub>2</sub> to CO. Methyl corrinoid (III) FeS and CO are then used as inputs in the final step of CO<sub>2</sub> fixation, acetyl coA synthesis. The first model showed high levels of CO synthesis. However, methyl corrinoid (III) FeS was not produced. The tetrahydrofolates that are meant to be produced in preceding steps were also not synthesized, with the exception of N10-tetrahydrofolate, which is the earliest one produced in this pathway.

This may be explained by competition for reactants either between the two parallel pathways, or within the pathway for methyl corrinoid (III) FeS synthesis. NADPH is consumed when CO<sub>2</sub> is used for formate synthesis, and again later on for the synthesis of 5,10-methylenetetrahydrofolate from 5,10-methyltetrahydrofolate. Formate is produced in oscillating concentrations and used in the synthesis of N10-tetrahydrofolate, which unlike 5,10-methylenetetrahydrofolate is successfully produced in this model. As such, NADPH may be consumed early in the pathway and depleted before 5,10-methyltetrahydrofolate can be produced. Alternatively, CO production from CO<sub>2</sub> involves only one step, which is a fifth order reaction and thus relatively rapid compared to formate synthesis that is only second order. CO<sub>2</sub> may be consumed more rapidly in CO synthesis than in the multi-step pathway to methyl corrinoid (III) FeS production, resulting in the high levels of CO produced in the model.

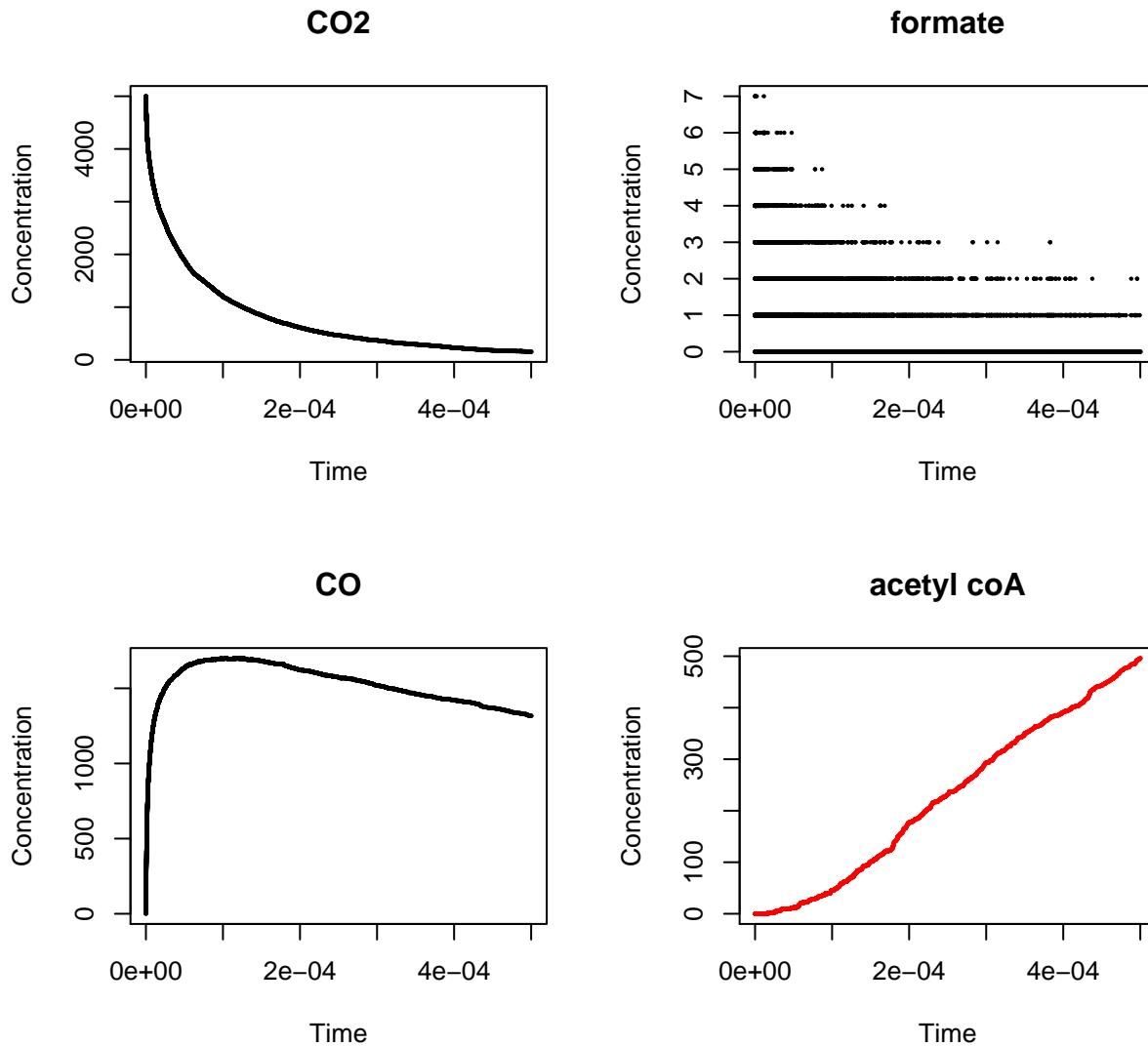
To test these hypotheses, a model was created including only the reactions for carbon fixation, introducing a few modifications.

#### Removing the requirement for NADPH in formate synthesis



In the figure above, the NADPH requirement was removed from formate synthesis, such that NADPH was only consumed for the production of 5,10-methylenetetrahydrofolate. However, while no NADPH depletion was seen in the model and acetyl-coA was not produced,  $\text{CO}_2$  was still rapidly consumed. This suggests that the pathway for methyl corrinoid (III) FeS synthesis was not occurring, and that the  $\text{CO}_2$  was funneled into  $\text{CO}$  synthesis. The internal competition for NADPH amongst tetrahydrofolate intermediates must not be a strong rate-limiting step.

## Decreasing the rate of reaction for CO synthesis

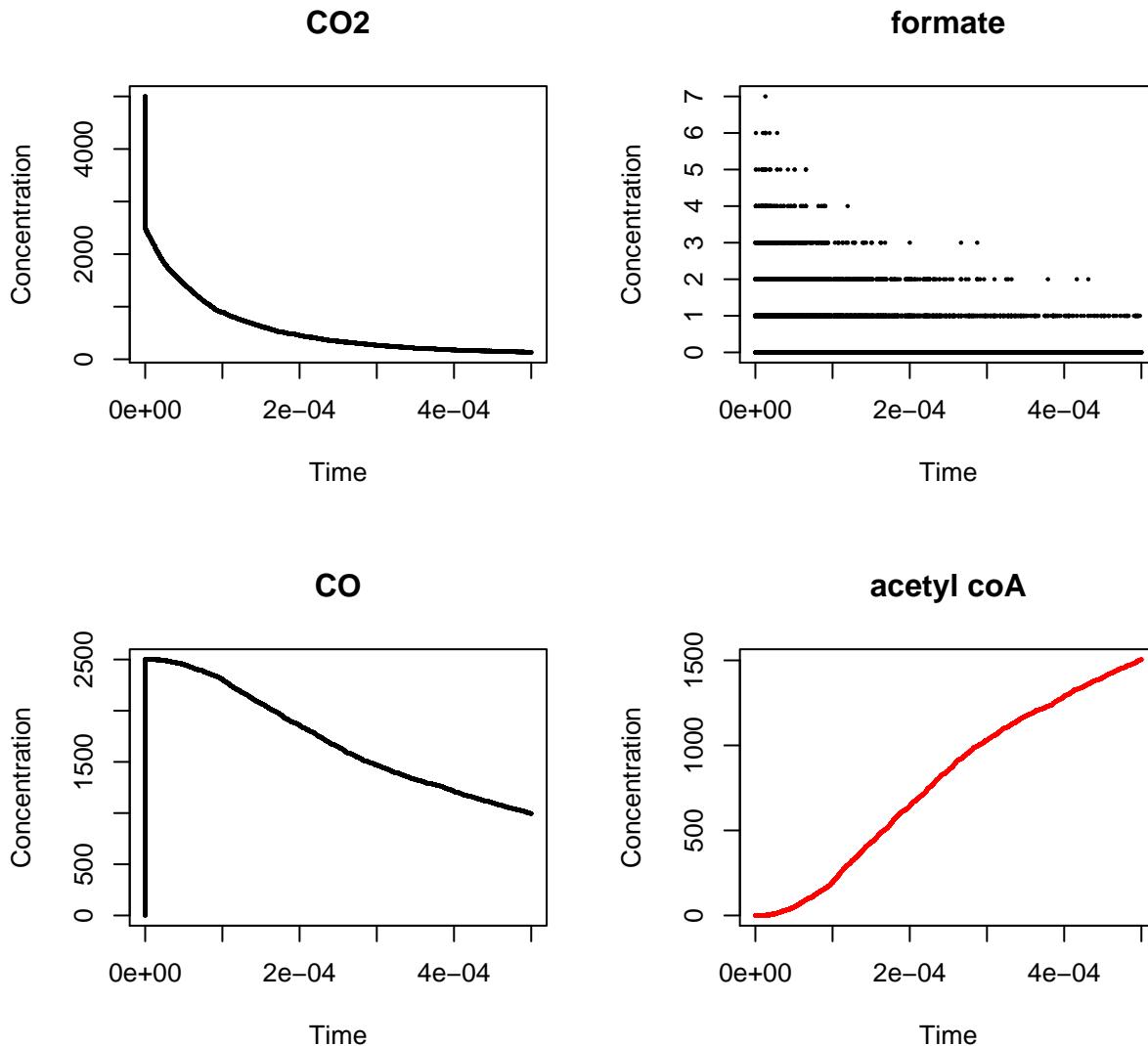


The figure above shows the results of a model where the rate of CO synthesis is decreased by a factor of  $10^{10}$ . Production of all the tetrahydrofolate intermediates occurred, although at a lower level for intermediates synthesized later in the pathway. CO and low levels of methyl corrinoid (III) FeS were both synthesized, which reacted to produce acetyl coA. This suggests that CO synthesis is likely the limiting step for this pathway. These results are seen when CO synthesis is decreased by at least a factor of  $10^3$ .

Making CO synthesis second order like formate synthesis, as opposed to the original fifth order, also generates the same results. This indicates that when CO and formate synthesis reactions compete for CO<sub>2</sub> at the same rate, carbon fixation proceeds to completion. Competition for CO<sub>2</sub> between the parallel pathways is thus rate-limiting.

In following models, the CO synthesis reaction is made first order with respect to H<sup>+</sup> and reduced FeS clusters, and thus third order overall. This is because the reaction is able to proceed with only one mol of each, and only requires the second mol to become more energetically stable. Making CO synthesis third order alone does not result in acetyl-coA production. This allows for the effectiveness of the subsequently introduced modifications to be considered.

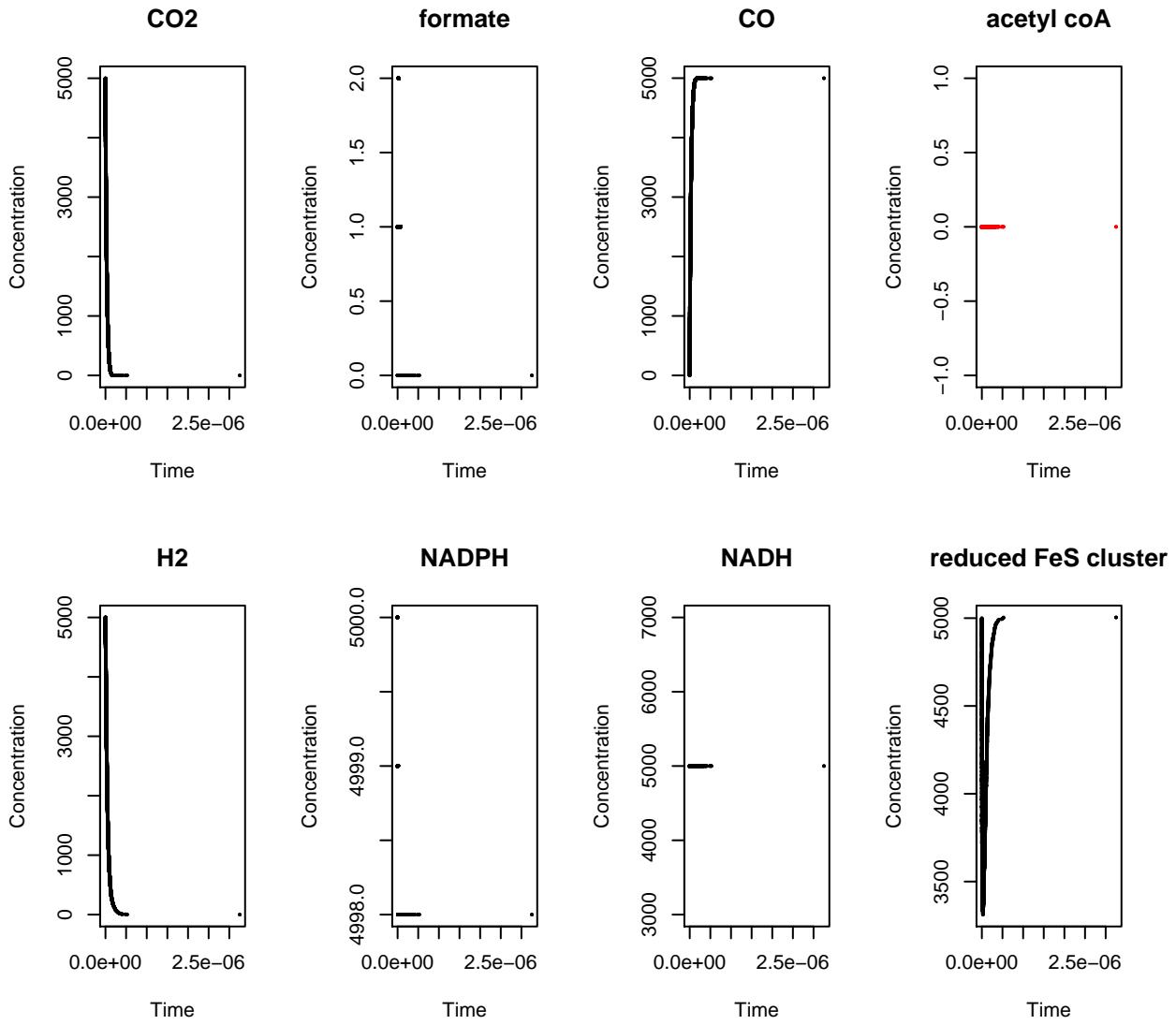
### Increasing the initial H<sup>+</sup> concentration



In the figure above, the initial H<sup>+</sup> concentration was increased by a factor of 3. This similarly results in the synthesis of CO, all tetrahydrofolate intermediates, and acetyl coA. This may be because both CO synthesis and many of the steps to methyl corrinoid (III) FeS production involve redox reactions, but most of the available protons are consumed for the faster reaction, CO synthesis. Increasing H<sup>+</sup> concentration prevents inhibition of the pathway by H<sup>+</sup> depletion, allowing methyl corrinoid (III) FeS synthesis to proceed.

### Regenerating reducing agents

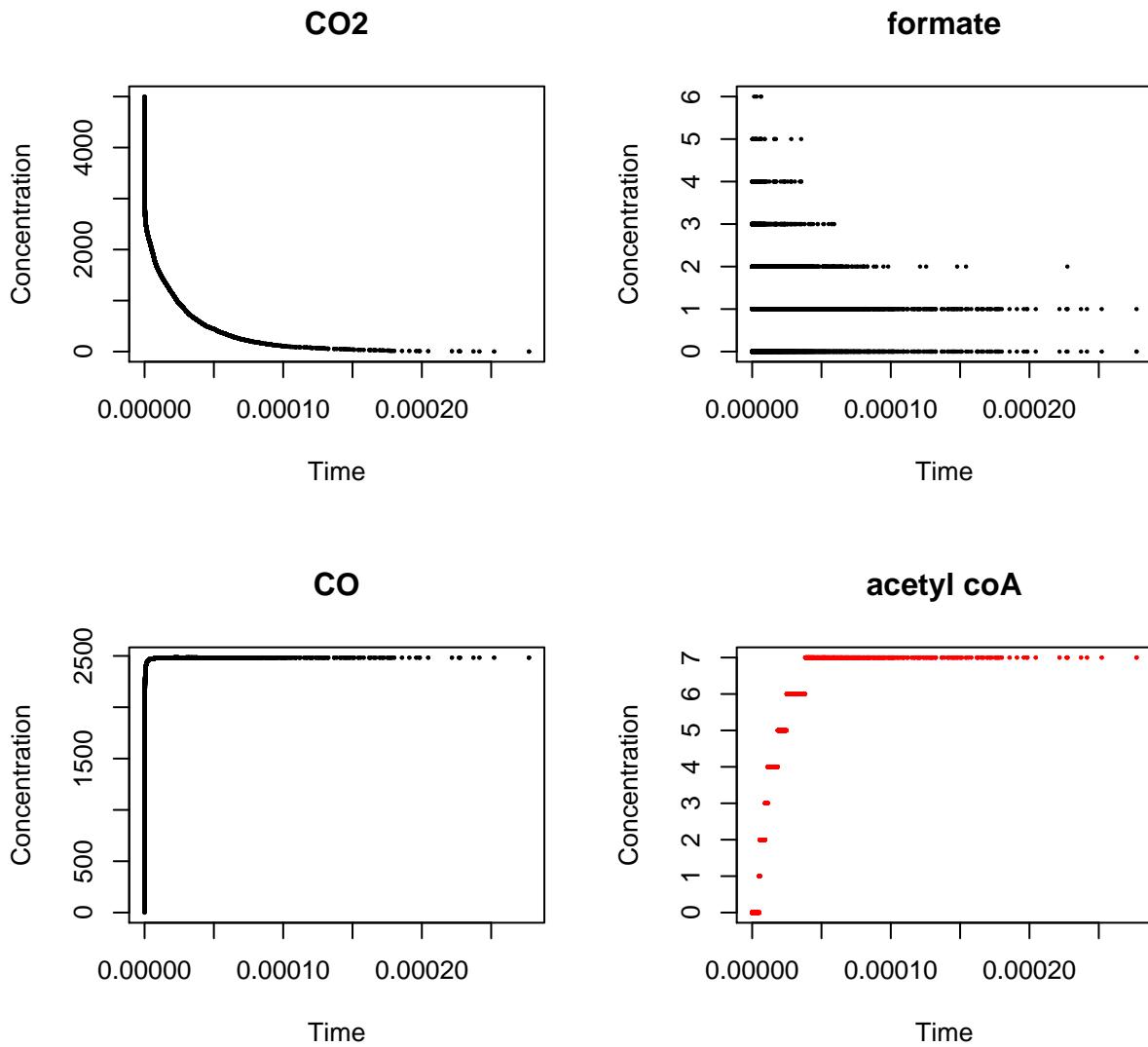
The competition for reagents and reducing agents suggests that a model with only forward reactions, unable to regenerate the reactants consumed, will inhibit itself over time. Reducing agents will be regenerated after they are oxidized if they are able to react with hydrogen in the environment. The introduction of reactions for NADPH, NADP, and reduced FeS cluster synthesis is modeled below.



Reduced FeS clusters were generated, although acetyl coA, NADPH, and NADP were not. This may be because the CO synthesis pathway still occurs at a greater rate, such that most of the available hydrogen is being used to regenerate only the reduced FeS clusters that are consumed in CO synthesis. Reduced FeS cluster synthesis is also third order, while NADPH and NADH are only second order overall, which means it is occurring at a faster rate and able to sequester more hydrogen than the other two regenerative steps.

### Changing the rate constants

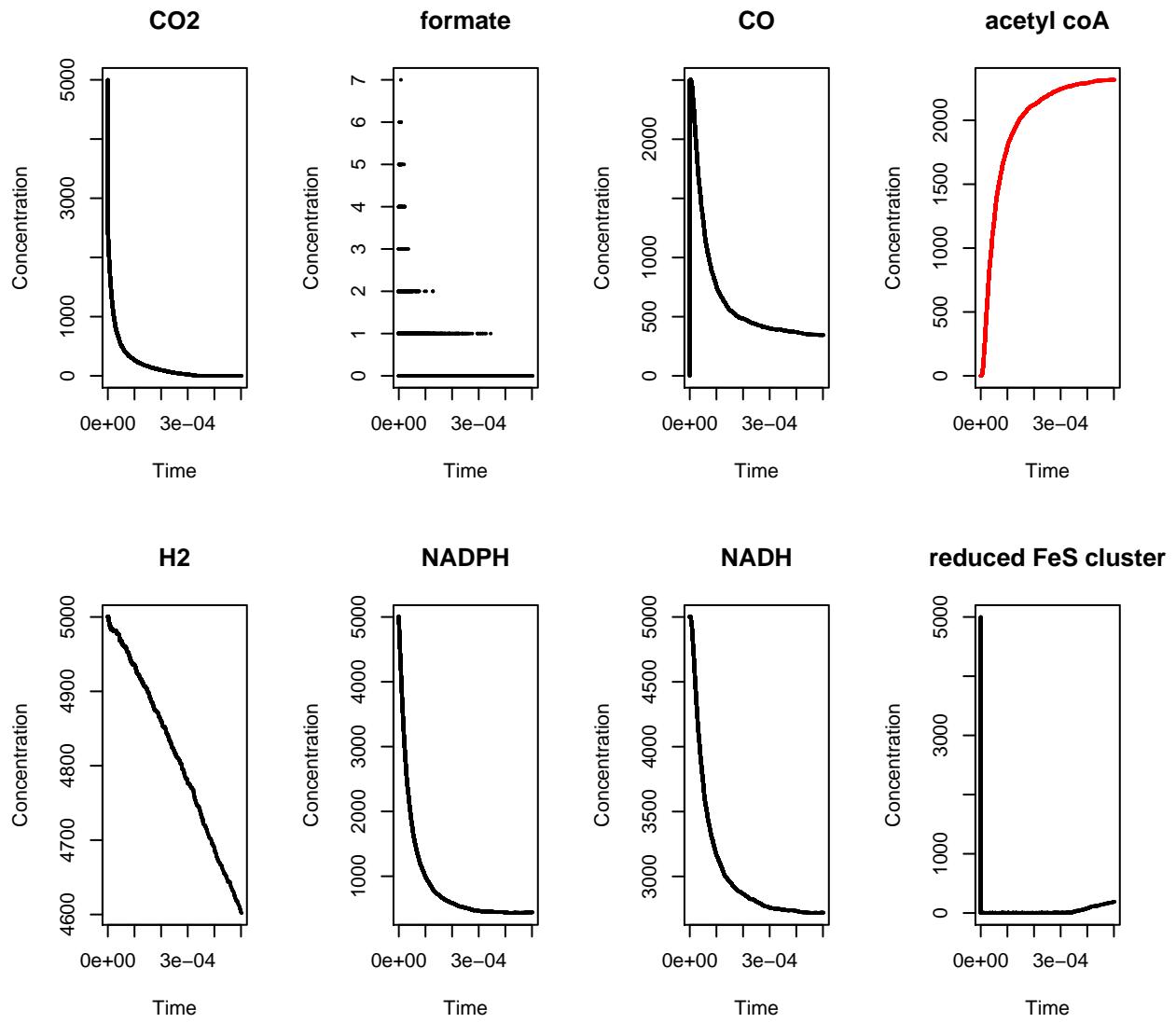
Rate constants were obtained for every reaction in the pathway based on the reported  $k_{cat}$  of each, although values were only available for enzyme-catalyzed reactions based on experimental models in physiological conditions.  $k_{cat}$  values were chosen from enzymes found in bacteria or archaea, for reactions going in the direction that they occur in metabolism. These values were then supplied as initial parameters in the model for carbon fixation, and the results of this are shown below.



Low levels of acetyl-coA synthesis occurred, which was not detected in the unmodified model. CO was still successfully being produced at high concentrations. Formate and proceeding tetrahydrofolate intermediates were synthesized successfully, although only in very low or oscillating concentrations. Formate synthesis has a  $k_{cat}$  two orders of magnitude greater than that of CO synthesis in physiological conditions, thus allowing for the methyl corrinoid (III) FeS synthesis pathway to proceed despite the problems described earlier.

### Combined model

A model was then created with these rate constants from literature, an excess of  $H^+$  to simulate a reducing environment, and reactions for the regeneration of reducing agents. The regenerative reactions were given a reaction rate one order of magnitude less than the slowest forward reaction, CO synthesis. The reaction for reduced FeS synthesis was also made first order with respect to oxidized FeS and thus second order overall. The results of this are shown below.



Acetyl-coA, formate, and CO were all produced. The tetrahydrofolate intermediates were also synthesized. Interestingly, reduced FeS clusters were regenerated at very low levels towards the end of the simulation, but NADPH and NADH were not.