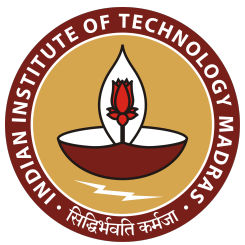
Ab-initio Synthesis of Amino acids

N Sowmya Manojna | Sahana Gangadharan
BE17B007 | BE17B038

Under the Guidance of
Prof. Karthik Raman
IBSE, RBC-DSAI, IIT Madras



Miller-Urey experiment

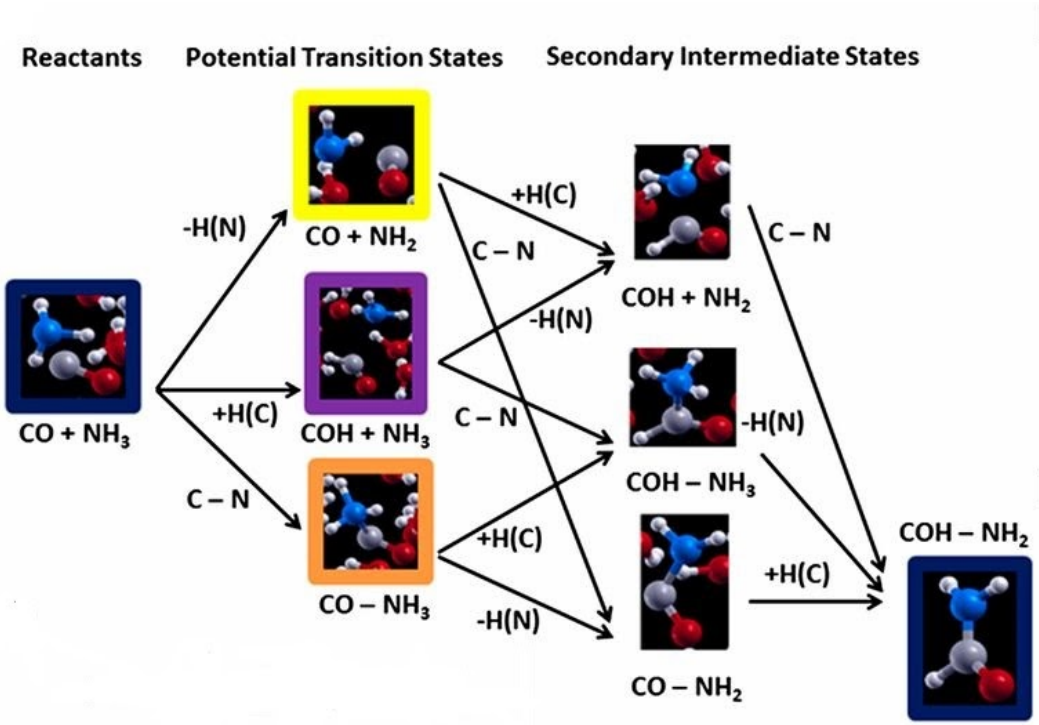


Figure 1: Metadynamic based analysis of the formation of formamide (Saitta and Saija (2014))

- Prebiotic earth - CH_4, NH_3, H_2O, and H_2
- Track formation of Amino acids
- Important intermediates!
- Strecker's amino acid synthesis

Miller-Urey experiment

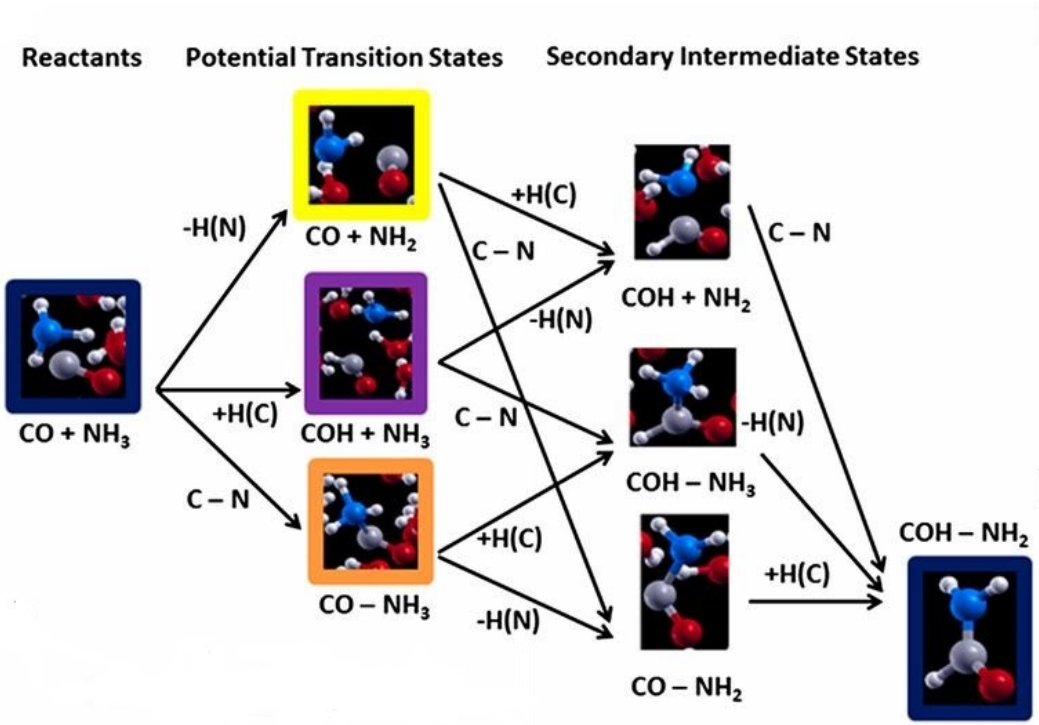


Figure 1: Metadynamic based analysis of the formation of formamide (Saitta and Saija (2014))

- Prebiotic earth - CH_4, NH_3, H_2O, and H_2
- Track formation of Amino acids
- Important intermediates!
- Strecker's amino acid synthesis

Miller-Urey experiment

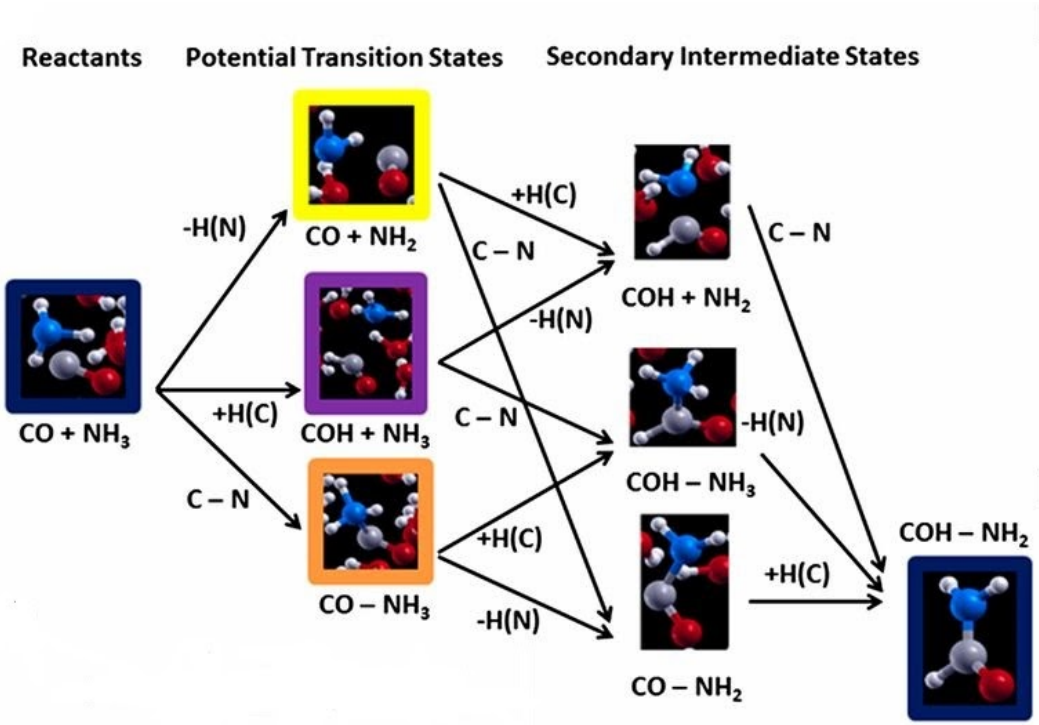


Figure 1: Metadynamic based analysis of the formation of formamide (Saitta and Saija (2014))

- Prebiotic earth - CH_4, NH_3, H_2O, and H_2
- Track formation of Amino acids
- Important intermediates!
- Strecker's amino acid synthesis

Miller-Urey experiment

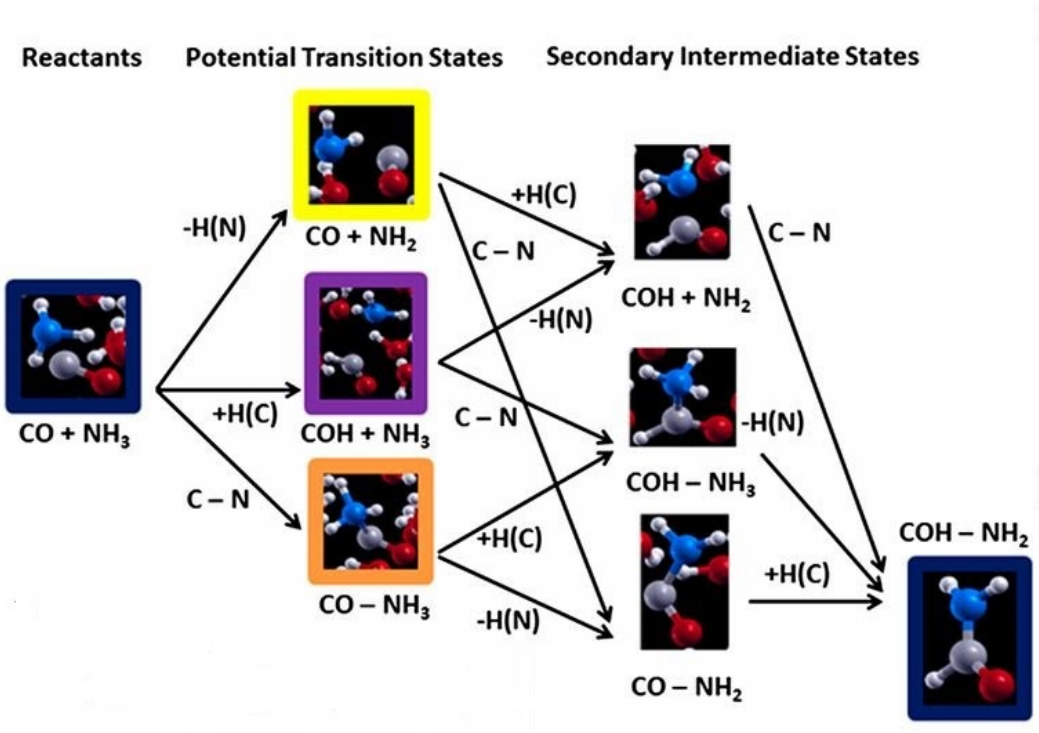


Figure 1: Metadynamic based analysis of the formation of formamide (Saitta and Saija (2014))

- Prebiotic earth - CH_4, NH_3, H_2O, and H_2
- Track formation of Amino acids
- Important intermediates!
- Strecker's amino acid synthesis

Thermodynamic data collection

- Reaction Mechanism Generator (RMG)
- Segregated in the form of libraries
- NASA and Group Additivity formats
- $C_p^o(T)$, $H^o(T)$ and $S^o(T)$, for two different temperature regimes were available.
- Highcharts of the same and Gibbs free energy

Thermodynamic data collection

- Reaction Mechanism Generator (RMG)
- Segregated in the form of libraries
- NASA and Group Additivity formats
- $C_p^o(T)$, $H^o(T)$ and $S^o(T)$, for two different temperature regimes were available.
- Highcharts of the same and Gibbs free energy

Thermodynamic data collection

- Reaction Mechanism Generator (RMG)
- Segregated in the form of libraries
- NASA and Group Additivity formats
- $C_p^o(T)$, $H^o(T)$ and $S^o(T)$, for two different temperature regimes were available.
- Highcharts of the same and Gibbs free energy

Thermodynamic data collection

- Reaction Mechanism Generator (RMG)
- Segregated in the form of libraries
- NASA and Group Additivity formats
- $C_p^o(T)$, $H^o(T)$ and $S^o(T)$, for two different temperature regimes were available.
- Highcharts of the same and Gibbs free energy

Thermodynamic data collection

- Reaction Mechanism Generator (RMG)
- Segregated in the form of libraries
- NASA and Group Additivity formats
- $C_p^o(T)$, $H^o(T)$ and $S^o(T)$, for two different temperature regimes were available.
- Highcharts of the same and Gibbs free energy

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup - Representation

class `Atom()`

Attributes:

- `_name`
- `_number`
- `_element`
- `_valency`

Member functions:

- `__init__()`
- `get_max_valency()`
- `current_valency()`
- `get_total_valency()`
- `repr()`

Initial Network Setup

- $H : C : O : N$ in ratio of 4:1:1:1, with a scaling factor of 8.
- Initial edges resulting in the formation of 2 CH_4, 2 NH_3 and 1 H_2O molecules

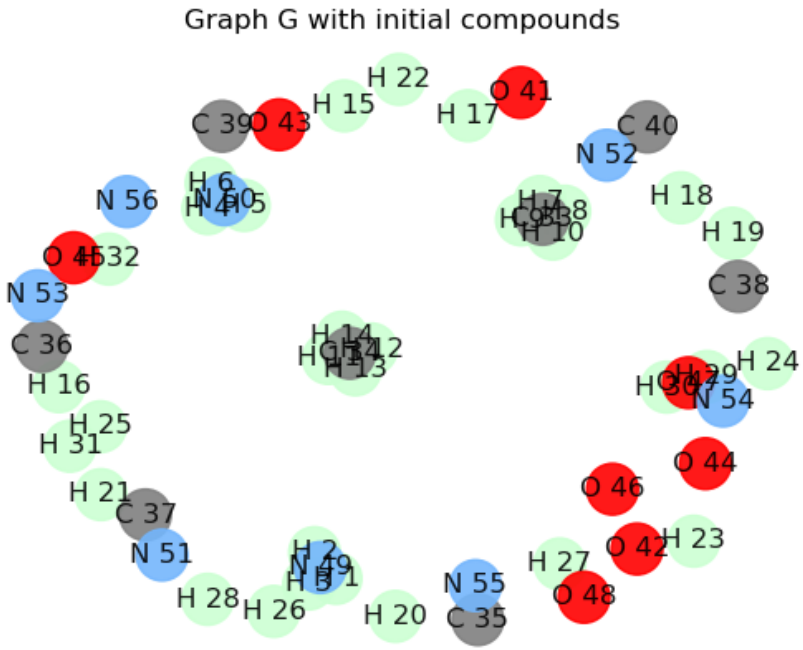


Figure 2: Initial network setup. Methane, Ammonia and Water are formed

Initial Network Setup

- $H : C : O : N$ in ratio of 4:1:1:1, with a scaling factor of 8.
- Initial edges resulting in the formation of 2 CH_4, 2 NH_3 and 1 H_2O molecules

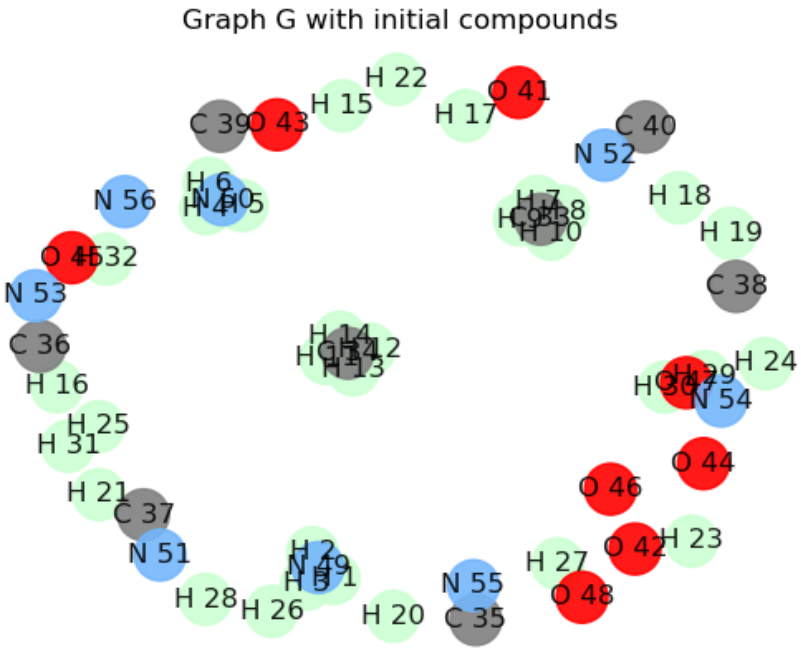


Figure 2: Initial network setup. Methane, Ammonia and Water are formed

What is the input for our model?

- Unknown - thermodynamic data.
- `Network H` - as a database that stores ΔG values for intermediates.
- `input_G()` - incorporates ΔG from `H` or takes input from user.

What is the input for our model?

- Unknown - thermodynamic data.
- `Network H` - as a database that stores ΔG values for intermediates.
- `input_G()` - incorporates ΔG from `H` or takes input from user.

What is the input for our model?

- Unknown - thermodynamic data.
- `Network H` - as a database that stores ΔG values for intermediates.
- `input_G()` - incorporates ΔG from `H` or takes input from user.

Random Graph Generation

- Generates connected components
- Maximum number of edges - `get_number_bonds()`
- Size of component formed - $(< 2 * \texttt{max\_size})$

Random Graph Generation

- Generates connected components
- Maximum number of edges - `get_number_bonds()`
- Size of component formed - $(< 2 * \texttt{max\_size})$

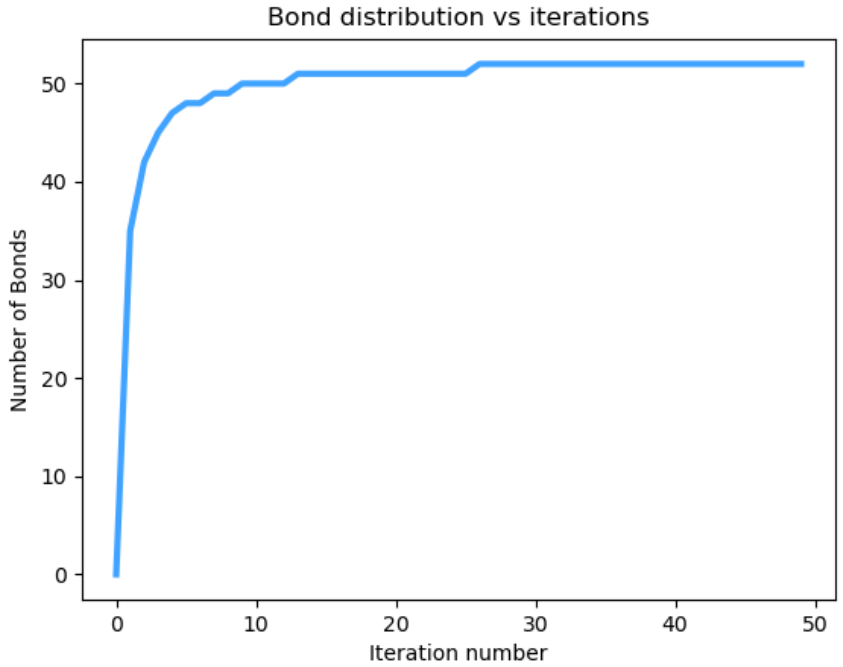


Figure 3: Maximum number of bonds that can be formed across iterations

Random Graph Generation

- Generates connected components
- Maximum number of edges - `get_number_bonds()`
- Size of component formed - $(< 2 * \texttt{max\_size})$

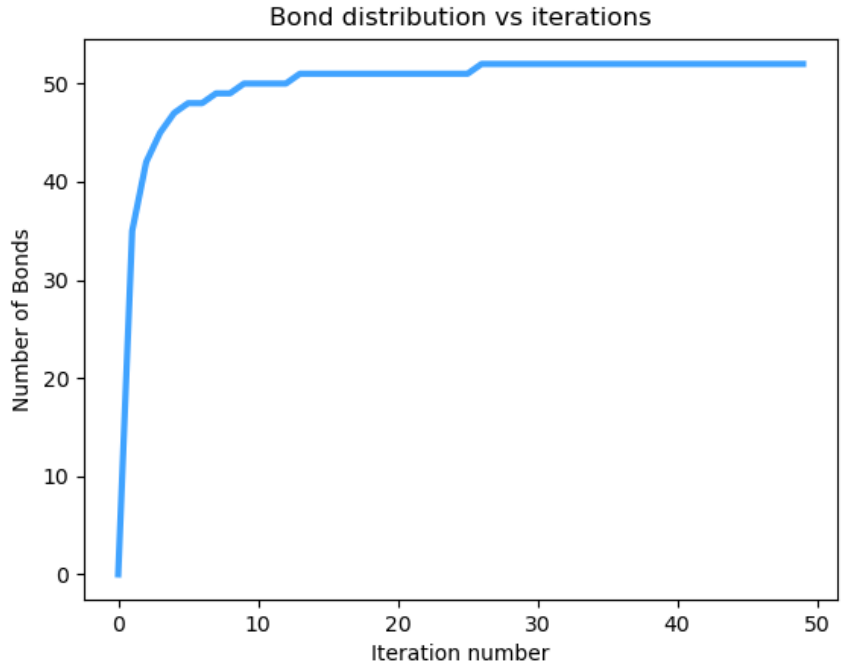


Figure 3: Maximum number of bonds that can be formed across iterations

Ensuring electrical neutrality

- Case 1: Atom and neighboring atoms don't have satisfied valency
- Case 2: All neighboring atoms have satisfied valency

Ensuring electrical neutrality

- Case 1: Atom and neighboring atoms don't have satisfied valency
- Case 2: All neighboring atoms have satisfied valency

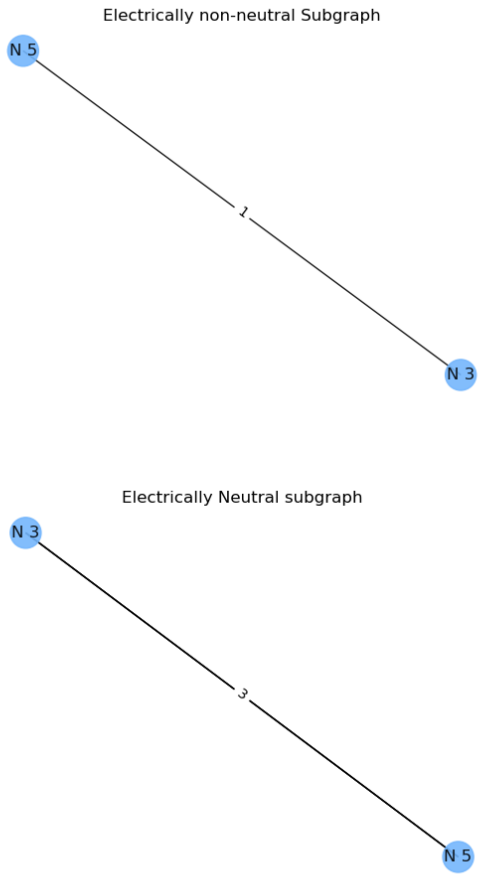


Figure 4: Neighboring atom doesn't have satisfied valency

Ensuring electrical neutrality

- Case 1: Atom and neighboring atoms don't have satisfied valency
- Case 2: All neighboring atoms have satisfied valency

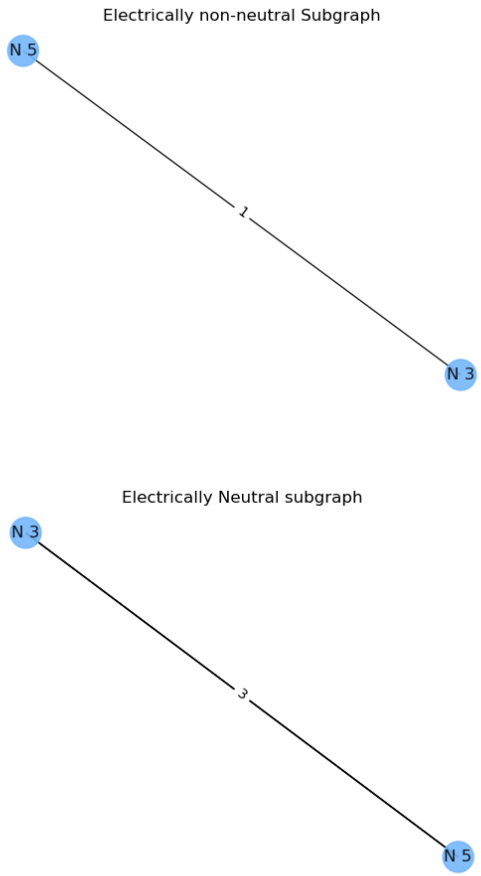
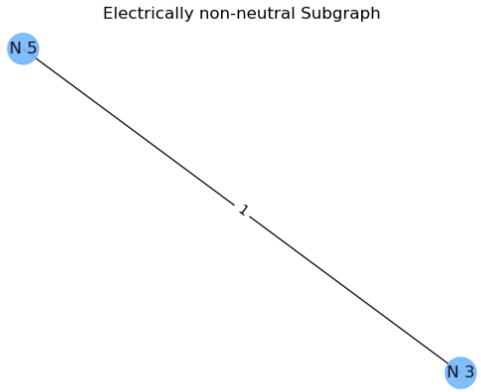


Figure 4: Neighboring atom doesn't have satisfied valency

Ensuring electrical neutrality

- Case 1: Atom and neighboring atoms don't have satisfied valency
- Case 2: All neighboring atoms have satisfied valency



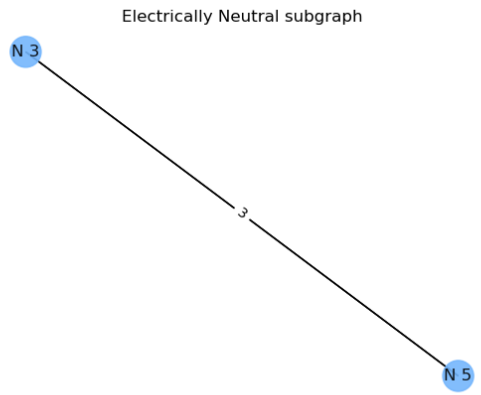
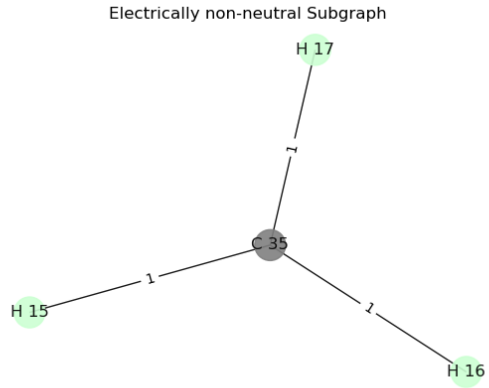


Figure 4: Neighboring atom doesn't have satisfied valency



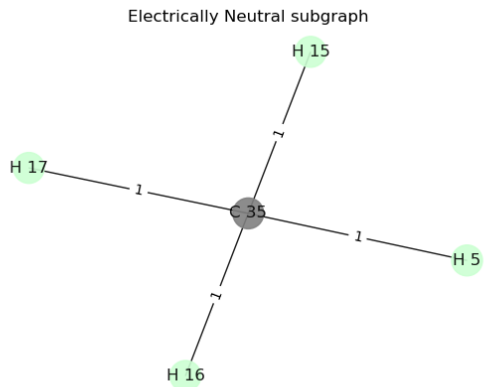


Figure 5: All neighboring atom have satisfied valency

Simulated Annealing

- Minimize sum total Gibbs free energy of network
- Probability p:

$$p = \exp\left(\frac{-(\Delta G^{\circ}_{tot,current} - \Delta G^{\circ}_{tot,best})}{T}\right)$$

$$T = (0.995)^{iteration} T_{intial}$$

- Enhanced sampling: `rearrange_connected_components`
  - `random.uniform(0.0, 1.0)` < 0.1: Complete reshuffling
  - 0.1 < `random.uniform(0.0, 1.0)` < 0.6: Two reshufflings
  - 0.6 < `random.uniform(0.0, 1.0)` < 0.8: One reshuffling
  - `random.uniform(0.0, 1.0)` > 0.8: No reshuffling

Simulated Annealing

- Minimize sum total Gibbs free energy of network
- Probability p:

$$p = \exp\left(\frac{-(\Delta G^{\circ}_{tot,current} - \Delta G^{\circ}_{tot,best})}{T}\right)$$

$$T = (0.995)^{iteration} T_{intial}$$

- Enhanced sampling: `rearrange_connected_components`
  - `random.uniform(0.0, 1.0)` < 0.1: Complete reshuffling
  - 0.1 < `random.uniform(0.0, 1.0)` < 0.6: Two reshufflings
  - 0.6 < `random.uniform(0.0, 1.0)` < 0.8: One reshuffling
  - `random.uniform(0.0, 1.0)` > 0.8: No reshuffling

Simulated Annealing

- Minimize sum total Gibbs free energy of network
- Probability p:

$$p = \exp\left(\frac{-(\Delta G^{\circ}_{tot,current} - \Delta G^{\circ}_{tot,best})}{T}\right)$$

$$T = (0.995)^{iteration} T_{intial}$$

- Enhanced sampling: `rearrange_connected_components`
  - `random.uniform(0.0, 1.0)` < 0.1: Complete reshuffling
  - 0.1 < `random.uniform(0.0, 1.0)` < 0.6: Two reshufflings
  - 0.6 < `random.uniform(0.0, 1.0)` < 0.8: One reshuffling
  - `random.uniform(0.0, 1.0)` > 0.8: No reshuffling

Reaction simulation

Mediated by Hydrogen atoms

Reaction simulation

Mediated by Hydrogen atoms

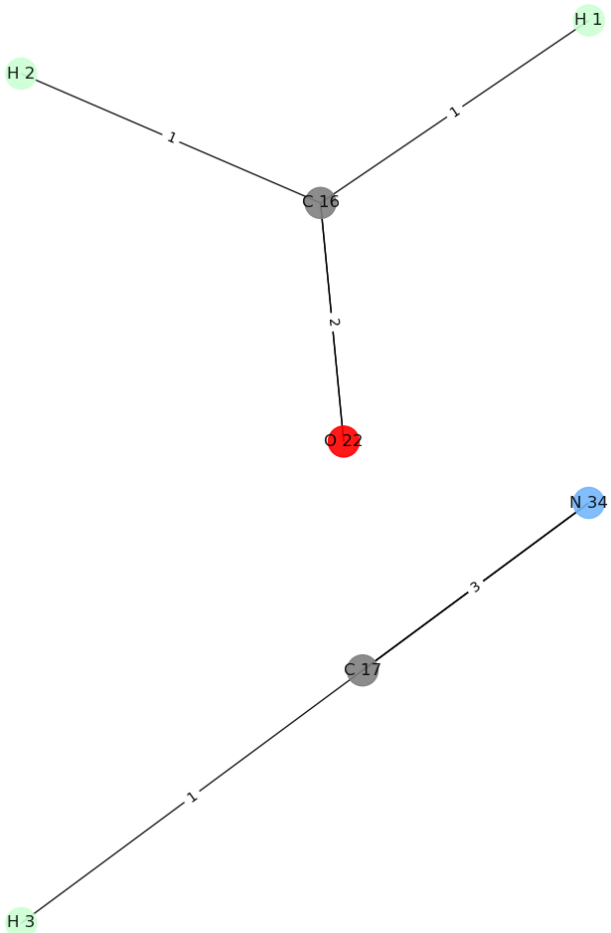


Figure 6: Reactants: Formaldehyde and Hydrogen Cyanide

Reaction simulation

Mediated by Hydrogen atoms

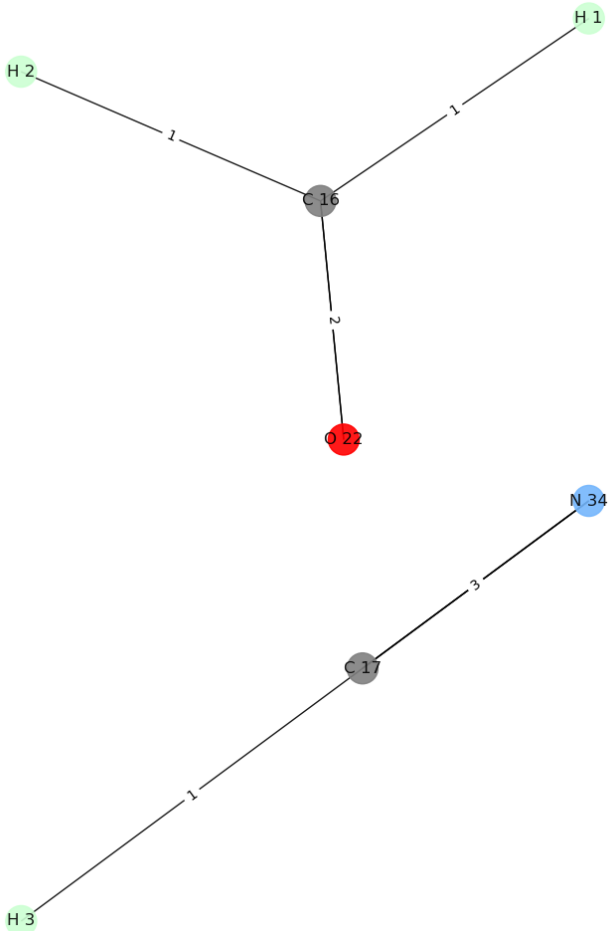


Figure 6: Reactants: Formaldehyde and Hydrogen Cyanide

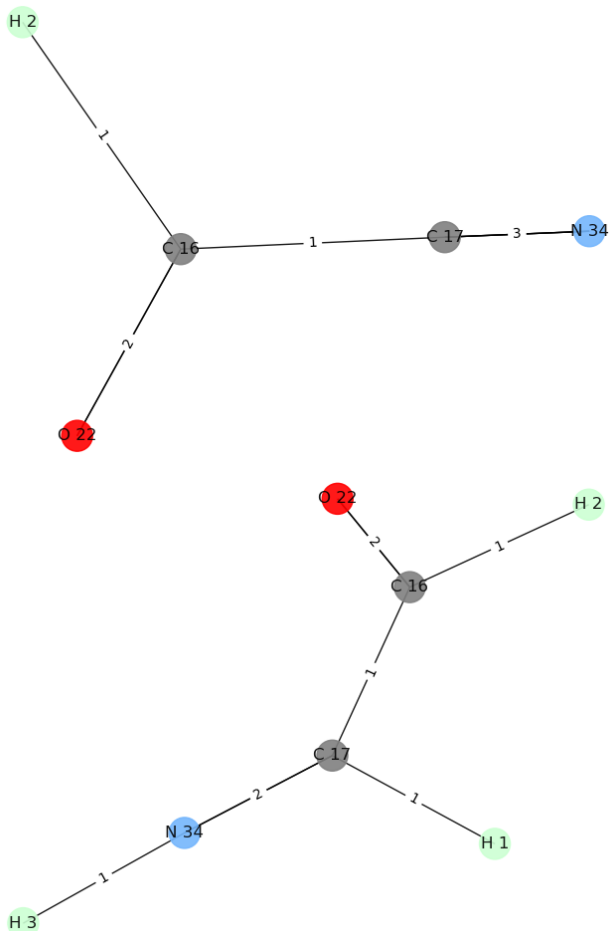


Figure 7: Products: Glyoxylonitrile and Iminoacetaldehyde

Strecker Reaction

Compounds were scanned for

- Aldehydes
- Ketones
- Amines

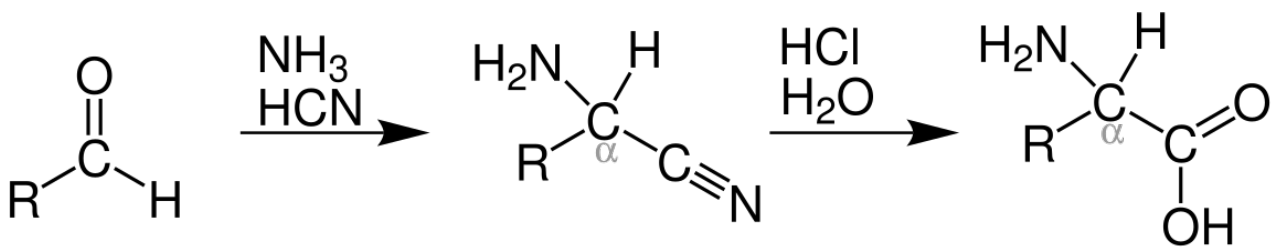


Figure 8: Strecker amino acid synthesis

Glycine synthesis

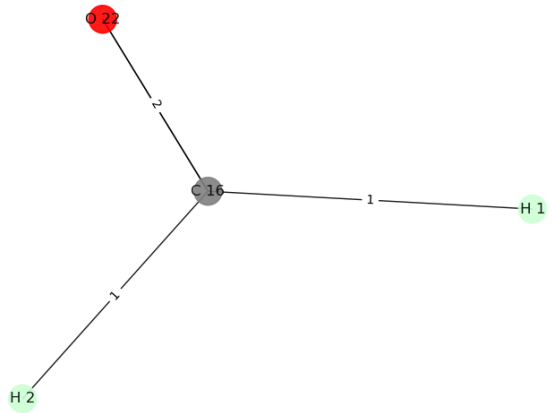


Figure 9: Initial reactant - Formaldehyde

Glycine synthesis

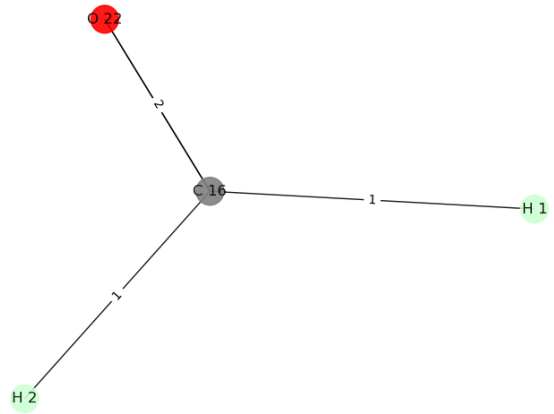


Figure 9: Initial reactant - Formaldehyde

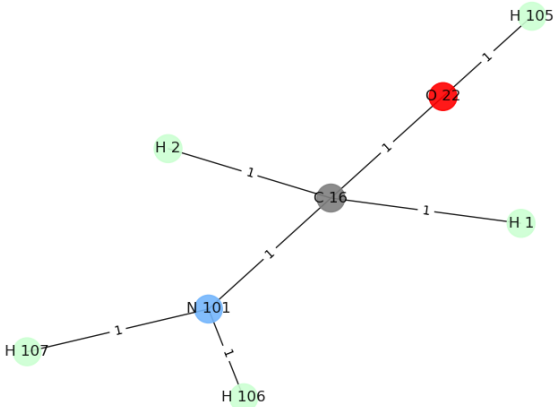


Figure 10: Addition of ammonia - Aminomethanol

Glycine synthesis

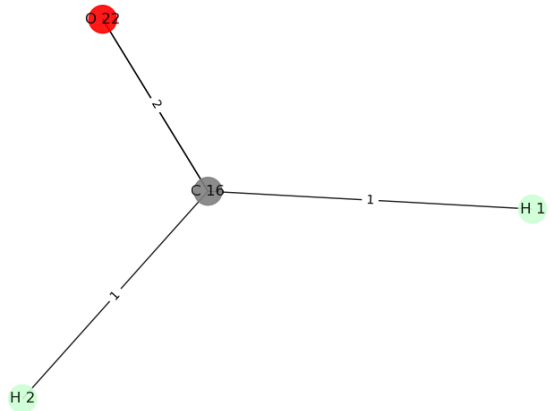


Figure 9: Initial reactant - Formaldehyde

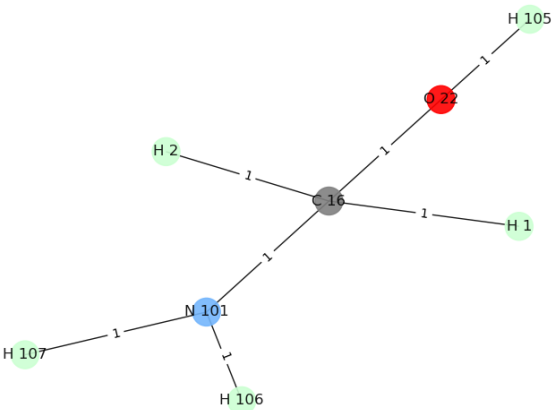


Figure 10: Addition of ammonia - Aminomethanol

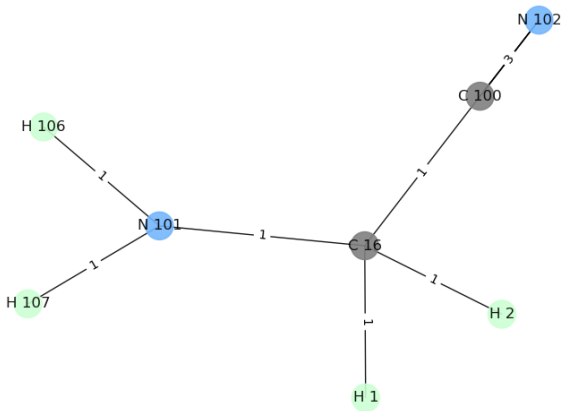


Figure 11: Addition of HCN - Aminoacetonitrile

Glycine synthesis

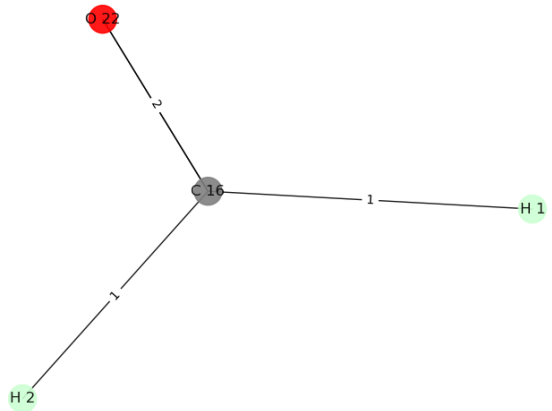


Figure 9: Initial reactant - Formaldehyde

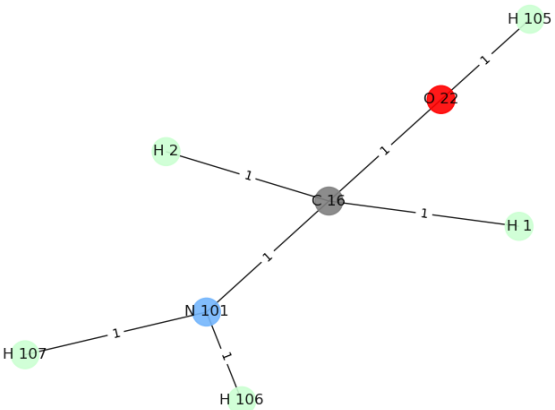


Figure 10: Addition of ammonia - Aminomethanol

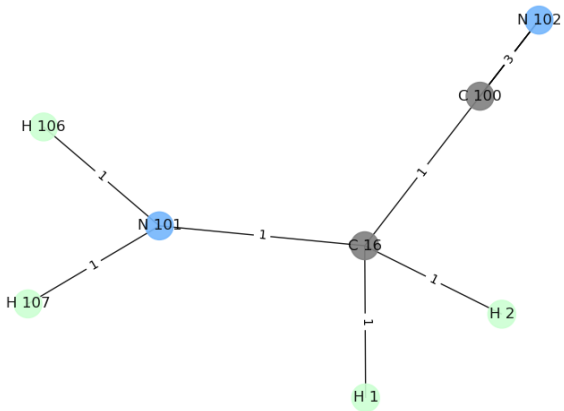


Figure 11: Addition of HCN - Aminoacetonitrile

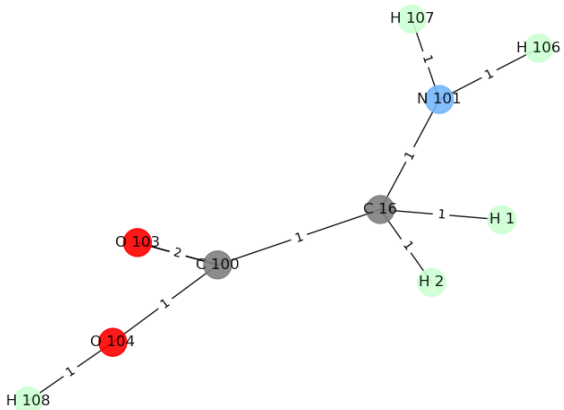


Figure 12: Acid hydrolysis - Glycine

Alanine synthesis

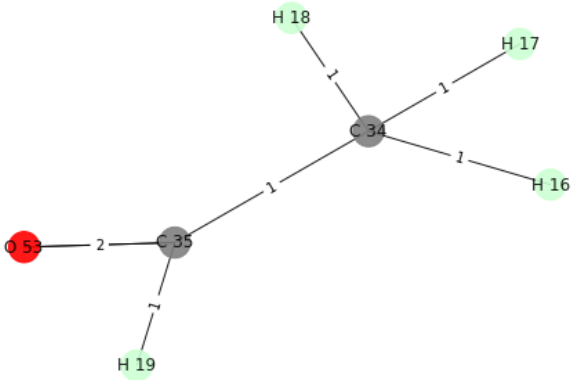


Figure 13: Initial reactant - Acetaldehyde

Alanine synthesis

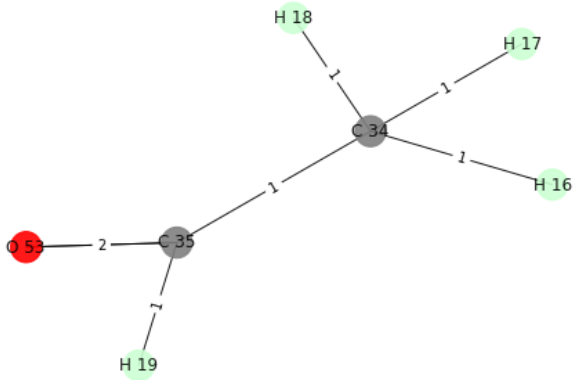


Figure 13: Initial reactant - Acetaldehyde

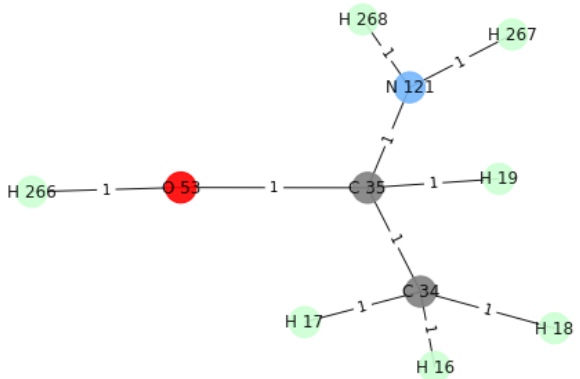


Figure 14: Addition of ammonia - Ethylamine

Alanine synthesis

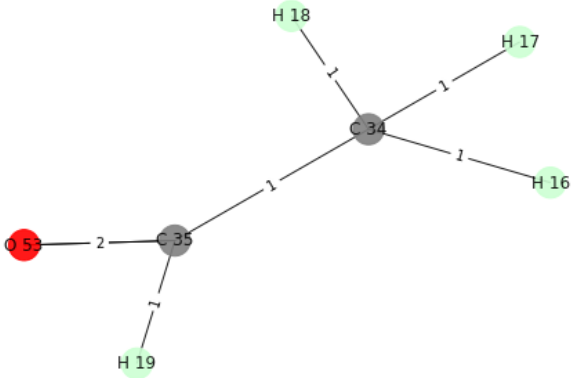


Figure 13: Initial reactant - Acetaldehyde

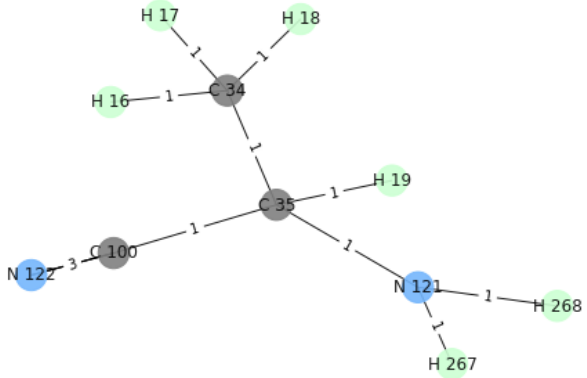


Figure 15: Addition of HCN: 2-Aminopropanenitrile

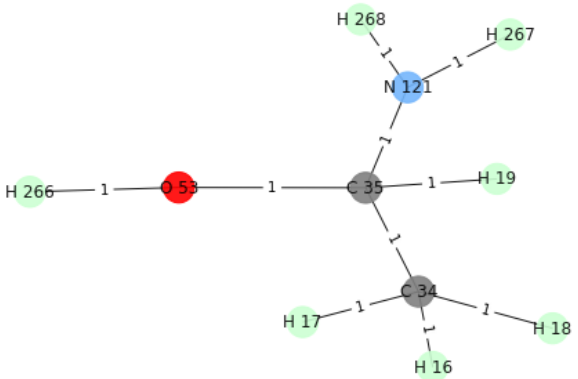


Figure 14: Addition of ammonia - Ethylamine

Alanine synthesis

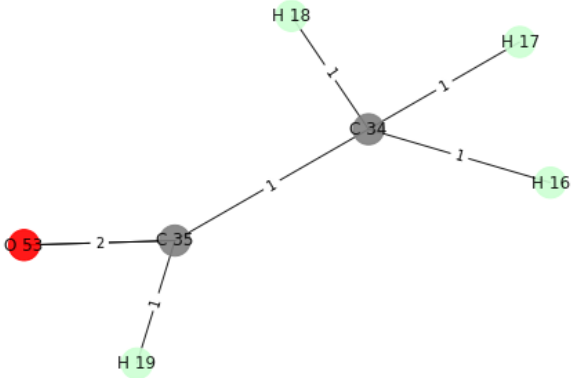


Figure 13: Initial reactant - Acetaldehyde

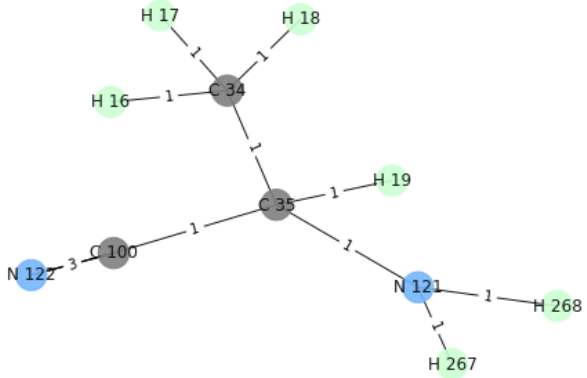


Figure 15: Addition of HCN: 2-Aminopropanenitrile

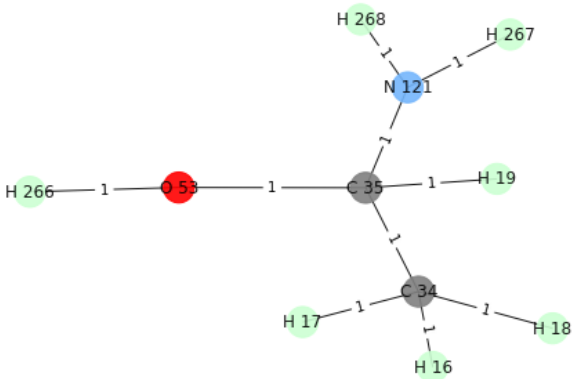


Figure 14: Addition of ammonia - Ethylamine

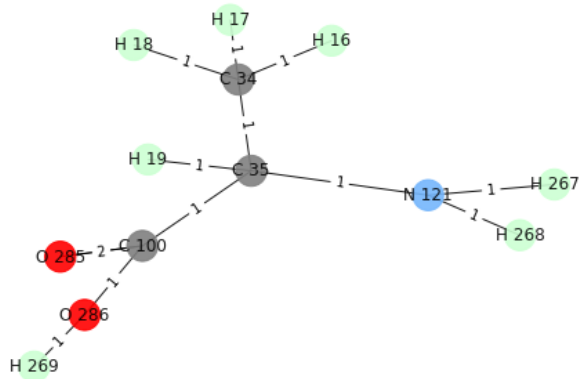


Figure 16: Acid hydrolysis - Alanine

Results

- Key intermediates:
  - Glycine: Formaldehyde, Aminoacetonitrile and Hydrogen Cyanide
  - Alanine: Acetaldehyde, Ethylamine and Hydrogen Cyanide
- Gibbs free energy across reaction coordinates

Results

- Key intermediates:
  - Glycine: Formaldehyde, Aminoacetonitrile and Hydrogen Cyanide
  - Alanine: Acetaldehyde, Ethylamine and Hydrogen Cyanide
- Gibbs free energy across reaction coordinates

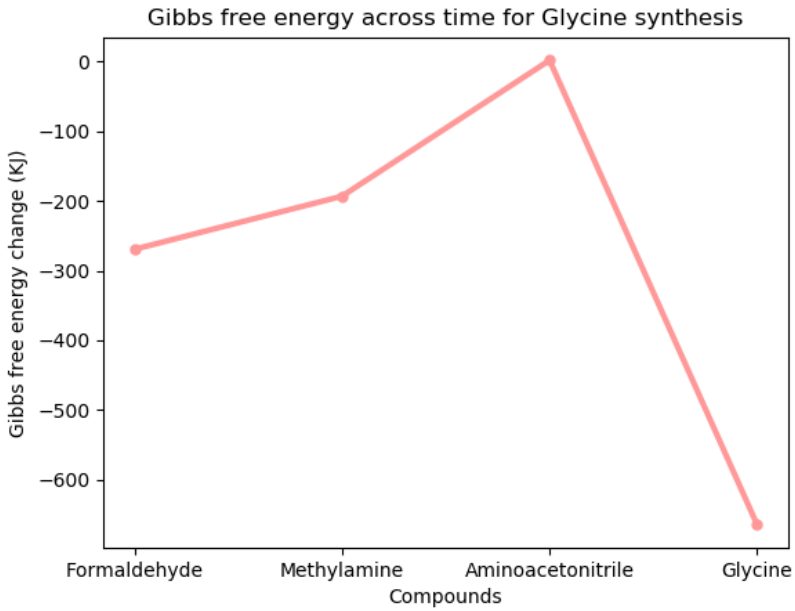


Figure 17: Gibbs free energy across reaction coordinates for Glycine synthesis

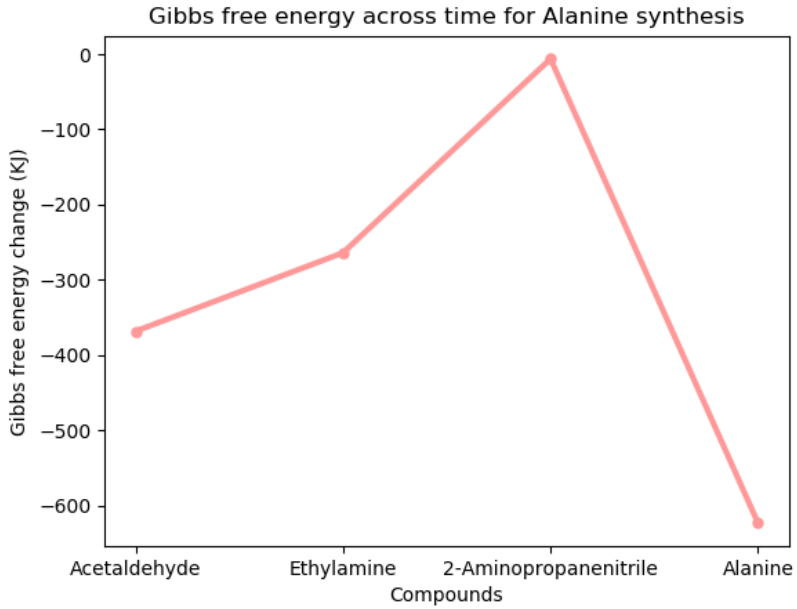


Figure 18: Gibbs free energy across reaction coordinates for Alanine synthesis

Challenges

- Electrically charged compounds
  - Unavailability of Gibbs free energy data
  - We overcame by introducing function - `get_neutral_compound()`
- Computationally expensive
  - Reduced number of simulations of Lewis Acid-Base reaction.

Challenges

- Electrically charged compounds
  - Unavailability of Gibbs free energy data
  - We overcame by introducing function - `get_neutral_compound()`
- Computationally expensive
  - Reduced number of simulations of Lewis Acid-Base reaction.

Thank You!

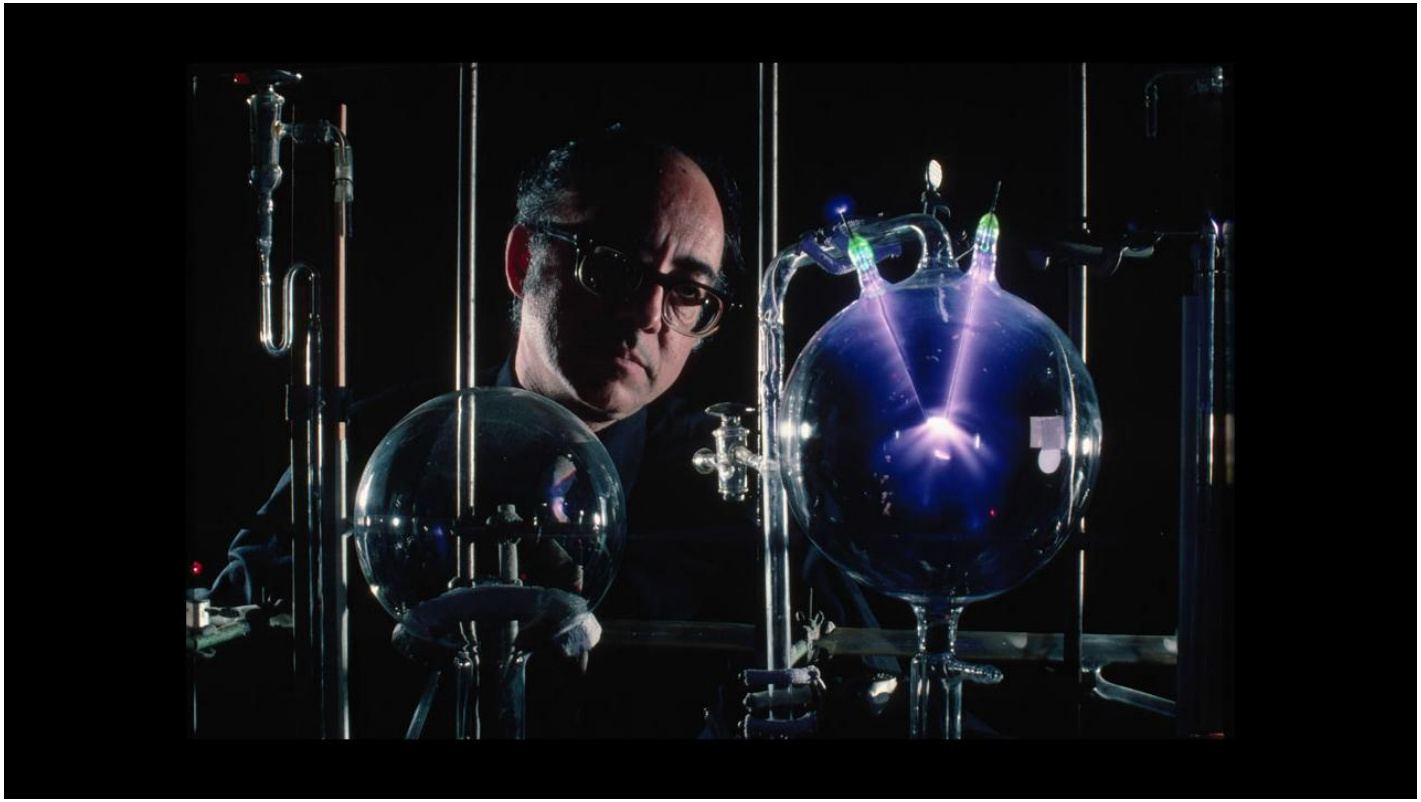


Figure 19: Stanley L. Miller overlooking a spark discharge apparatus 1994. (©Roger Ressmeyer/CORBIS)