Building models in RMG – PART 3

Shamel Merchant
• Reaction Library

Usage:

\[ A + B \rightleftharpoons C + D \quad (A) \quad (n) \quad (E_a) \]

1. If the reaction rate is measured experimentally

Pitfall:

Once reaction libraries has mixed high-P and Pdep rates, you cannot run Pdep module of RMG. This leads to double accounting for the rates.
– Literature rates are reporting the overall rate for this reaction not the disproportionation rate which is being represented above.

\[
\text{CH}_2 + \text{O}_2 \rightarrow \text{HO}_2
\]

– If we now run RMG with pressure dependence and use the reaction rate in Reaction Libraries then RMG uses it as a high P for disproportionation rate.

\[
\text{CH}_2 + \text{O}_2 \rightarrow \text{OO} + \text{HO}_2
\]
This is where our rate from reaction library gets used.
General tip for model construction

1. Start making your models initial at high pressure limit – the model is more transparent and easy to track.

Product-1 overpredicted
General tip for model construction

It's easier to inspect the pathway

High Pressure Mechanism

\[ k_{A=B} \] is too fast and reason for misprediction

Pressure – Dep Mechanism

This is not a direct pathway – but a well skipping reaction
\[ k_{A=B} \] Cannot be identified unless you go and run a separate populate reaction job and check all the rates.
2. Control the tolerance to get model which are adequate for your purpose

My general strategy is to start with as loose tolerance I can and check if my initial branching and major product distribution is okay.

In this particular case, I found out RMG had a too slow dehydration rate and underpredicted my H$_2$O speciation data. This is usually the first reaction and you do not need a very tight tolerance to get it in the model.
3. RMG cannot estimate thermochemistry for unsaturated cyclic molecules even with QM on. This error is enough to cause RMG to start exploring completely different pathways.

<table>
<thead>
<tr>
<th>H&lt;sub&gt;298&lt;/sub&gt;</th>
<th>CBS-QB3 (kcal/mol)</th>
<th>QM PM-7 (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50.95</td>
<td>40.75</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>H&lt;sub&gt;298&lt;/sub&gt;</th>
<th>CBS-QB3 (kcal/mol)</th>
<th>QM PM-7 (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>42.25</td>
<td>34.5</td>
<td></td>
</tr>
</tbody>
</table>
The problem with Seed libraries

- In some cases reactions might be missing out because of the interconnectivity between different seed libraries

- Example:

Seed 1:
\[
\text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (-0.5,0) -- (0.5,0);
\end{tikzpicture}}}
\rightarrow \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (-0.5,0) -- (0,0) -- (0.5,0);
\end{tikzpicture}}} + \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (0,0) circle (0.3);
\end{tikzpicture}}}
\]

Seed 2:
\[
\text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (-0.5,0) -- (0,0) -- (0.5,0);
\end{tikzpicture}}} + \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (0,0) circle (0.3);
\end{tikzpicture}}} \rightarrow \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (0,0) circle (0.3);
\end{tikzpicture}}} + \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (-0.5,0) -- (0,0) -- (0.5,0);
\end{tikzpicture}}}
\]

Seed 3: Combustion core
\[
\text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (-0.5,0) -- (0,0) -- (0.5,0);
\end{tikzpicture}}} \rightarrow \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (0,0) circle (0.3);
\end{tikzpicture}}} + \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (0,0) circle (0.3);
\end{tikzpicture}}} + \text{\rotatebox{-90}{\begin{tikzpicture}
\draw[thick,red] (-0.5,0) -- (0,0) -- (0.5,0);
\end{tikzpicture}}}
\]

All species are occurring in the seed, so this reaction will be missed out!
Detection of aromaticity
• Some species were wrongly recognized as being aromatic

• Benzene bonds were sort of unreactive (not included in R_Addition_Multiplebond)
  – Consequently all pathways that could form aromatic products, such as benzene, toluene… needed to be provided in libraries because RMG would not find them

\[
\text{ベンゼン} \rightarrow \text{フラン} + \text{H}^+ 
\]
New algorithm to detect aromaticity in PAH

- Based on the initial algorithm that was there but recursive
- Smart checks: no saturated C atoms (forbid sp$^3$)
- Hückel rule: AROMATIC if $n_{\text{elec}} = 4n + 2$
- Code:

Old code:

For each ring

$n_{\text{elec}} = \text{EndoPiElectrons} + \text{ExoPiElectrons}$

6 / 6 / 6

$\text{6 / 6 / 6}$

$\text{6 / 4}$

$\text{6 / 4}$
New algorithm to detect aromaticity in PAH

New code:

For each unrecognized ring

\[ n_{\text{elec}} = \text{EndoPiElectrons} + (\text{ExoPiElectrons})_{\text{arom}} \]

4 / 6 / 6

6 / A / A

RMG group meeting, 05/09/2013
New algorithm to detect aromaticity in PAH

New code:

For each unrecognized ring

\[ n_{\text{elec}} = \text{EndoPiElectrons} + (\text{ExoPiElectrons})_{\text{arom}} \]

perylene
How to make B bonds active?

• Do not convert to B bonds (keep the original structure with the S/D bonds)

• For each ChemGraph representing a compound:

```python
thermo_graph = ChemGraph.copy(this)
thermo_graph.determineAromaticityAndWriteBBonds()
thermoData = gen.generateThermo(this);
    if(!this.fromprimarythermolibrary && this.isAromatic) {
        thermoData = gen.generateThermo(thermo_graph);
    }
```

• No changing the original ChemGraph
New resonance isomers: kekulized forms

- At a certain point we found out we had two indene structures

\begin{align*}
\text{\includegraphics[width=0.2\textwidth]{indene1}} & & \text{\includegraphics[width=0.2\textwidth]{indene2}}
\end{align*}

- `generateResonanceIsomersFromConjugatedRings()`
  - Check for each ring
    - Do I have alternating single and double bonds?
    - If YES then make corresponding resonance isomer

\begin{align*}
\text{\includegraphics[width=0.25\textwidth]{resonance1}} & & \text{\includegraphics[width=0.25\textwidth]{resonance2}}
\end{align*}
Initial attempt to get some 3D in
Intramolecular rate rules were mainly obtained from linear molecules. Rates are not suited for polycyclics, moreover most of them are simply impossible. Hard-coded check for feasibility.
Two checks

- **The unreachable H atom**
  - a) Check the distance: \( n \) (3)
  - b) How many atoms along minimum path are in the same ring as site 1? \( n_1 \) (3)
  - c) How many atoms along minimum path are in the same ring as site 2? \( n_2 \) (3)
  - d) If \( (n+1)/2 \leq n_x - 1 \) then this reaction won’t occur

- **Sterically hindered reactions**
  - a) Determine minimum path from site 1 to site 2
  - b) Number of atoms that are part of two rings in product? \( n_{biring} \)
  - c) If \( n_{biring} > 1 \) then reaction hampered
Examples

• The unreachable H atom

\[(7+1)/2 \leq 7-1\]

• Sterically hindered reactions

\[(7+1)/2 \leq 2-1\]
Branching the exo - endocyclic addition tree
Problem

- We only calculated the most likely reactions (fast)

- Doing so (and because of the averaging scheme going on in RMG) **ALL** less likely reactions were also fast

- Quick solution: give all rate rules priority 0 (= they are not used to average out the tree)

- But eventually people will want to have a working tree structure
• Old tree

\[
\begin{align*}
R_{\text{all}} & \rightarrow R3 \\
 & \quad \rightarrow R4 \\
 & \quad \quad \rightarrow R5_{\text{SS}} \\
 & \quad \quad \quad \rightarrow R5_{\text{SS}} \_D \\
 & \quad \quad \quad \rightarrow R5_{\text{SS}} \_T \\
R5 & \rightarrow R5_{\text{SD}} \\
 & \quad \rightarrow R5_{\text{ST}} \\
 & \quad \rightarrow R5_{\text{SB}} \\
R6 & \rightarrow R7 \\
\end{align*}
\]
• New tree

R_all

R3

R4

R5

R5_SS

R5_SM

R5_SD

R5_SD_D

R5_SD_T

R5_SD_CO

R5_ST

R5_SB

R6

R6_plus

R7

R8

Will help you to avoid slow reactions!
Questions