Molecular Structure Design

For those primitive designs in which desired properties and performance are specified, it is often necessary to identify chemicals and chemical mixtures to meet the specifications

e.g., Thin polymer films to protect electronic devices
Refrigerants that do not react with ozone
Environmentally-friendly solvents
Low-viscosity lubricants
Proteins for pharmaceuticals with desired therapeutic effects
Polymer film to protect an electronic device

Device will operate at $T < 50^\circ C \rightarrow T_g \approx 100^\circ C$
A fairly dense layer is needed, $\rho = 1.5 \text{ g/cm}^3$
Small water absorption, $W = 0.005 \text{ gH}_2\text{O/gpolymer}$

Derringer and Markham (1985) create polymer repeat units involving 7 molecular groups:

-CH$_2$-, -CO-, -COO-, -O-, -CONH-
-CHOH-, -CHCl-

$$\min_{w.r.t. \ n} \left( \frac{\rho - \rho_{\text{spec}}}{\rho_{\text{spec}}} \right)^2 + \left( \frac{T_g - T_{g,\text{spec}}}{T_{g,\text{spec}}} \right)^2 + \left( \frac{W - W_{\text{spec}}}{W_{\text{spec}}} \right)^2$$

Yields - -((CH$_2$)$_3$(CHCl)$_6$)-

Semi-empirical, group contribution methods – provided by van Krevelen (1990)

Have the form:

$$p(n) = \frac{\sum_{i=1}^{N} A_i n_i}{\sum_{j=1}^{n} B_j n_j}$$

$N$ – no. of types of groups in repeat units
$N_i$ – no. of groups of type $i$

e.g., to estimate $T_g$ of

-((CH$_2$CH$_2$CHCl)$_n$)-
From van Krevelen

<table>
<thead>
<tr>
<th>Group</th>
<th>$A_i$</th>
<th>$B_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-CH$_2$-</td>
<td>2,700</td>
<td>14</td>
</tr>
<tr>
<td>-CHCl-</td>
<td>20,000</td>
<td>48.5</td>
</tr>
</tbody>
</table>

$$T_g = \frac{2,700 \times 2 + 20,000 \times 1}{14 \times 2 + 48.5 \times 1} = 332 \text{ K}$$

More fundamental estimation methods using microsimulation -
Molecular dynamics (MD), Monte-Carlo (MC) methods
Need pair potentials – e.g., Lennard-Jones
With added computer speed, methods gaining popularity

---

**Optimization to Locate Molecular Structure**

$$\min_n \sum_{j=1}^{P} (p_j \{n\} - p_j^{\text{spec}})^2$$

$$p_j^{\text{min}} \leq p_j \{n\} \leq p_j^{\text{max}}$$

$$n_i \in \{n_i^{\text{min}}, ..., n_i^{\text{max}}\}, \quad i = 1, ..., N$$

When adding a group to a repeating unit of a polymer,

the number of free attachments must be 2

$$f = \sum_{i=1}^{N} (v_i - 2)n_i + 2$$

$v_i$ is the valence, or number of free bonds, associated with group $i$
Example 2.2 (SSL, 2004) provides a GAMS program that locates 

\[-[(\text{CH}_2)_3(\text{CHCl})_6]-\]

at the minimum to give:

\[
\begin{align*}
T_g &= 384.7 \text{ K} \quad (T_g^{\text{spec}} = 383 \text{ K}) \\
\rho &= 1.4889 \text{ g/cm}^3 \quad (\rho^{\text{spec}} = 1.5 \text{ g/cm}^3) \\
W &= 0.0049 \text{ gH}_2\text{O/gpolymer} \quad (W^{\text{spec}} = 0.005 \text{ gH}_2\text{O/gpolymer})
\end{align*}
\]
For environmental and health reasons, chemical companies are increasingly challenged to find replacements - “designer” solvents that satisfy specifications for each application.

- e.g., lithographic printing
  - Ink is conveyed to “impression plant”
  - by a train of rubber rollers – “blankets”
  - Cleaned regularly
  - Solvent must -
    - dissolve dried ink
    - have small drying time – small $\Delta H^v$
    - have negligible swelling of blanket
    - be nonflammable

---

**Molecular Structure Design**

Sinha et al. (1999) create molecules involving 12 molecular groups:

- CH$_3$-, -CH$_2$-, Ar- (C$_6$H$_5$-), Ar= (C$_6$H$_4$=), -OH, CH$_3$CO-, -CH$_2$CO-, -COOH, CH$_3$COO-, -CH$_2$COO-, -CH$_3$O, and –CH$_2$O-

To avoid swelling, let $\delta_p > 6.3$ MPa$^{1/2}$

To give liquid at 1 atm, $T_b > 323$ K

$T_m < 223$ K

Keep bioconcentration factor low

Minimize $\Delta H^v$

Yields methyl-ethyl-ketone, diethyl-ketone, and ethylene glycol mono-methyl ether
Property Estimation Methods

Solubility paramater – accounting for polar interactions

\[
\delta_p = \sqrt{\frac{\sum_{i=1}^{N} n_i (1,000 F_i)}{V_0 + \sum_{i=1}^{N} n_i V_i}} \quad F_i \text{ from van Krevelen and Hoftyzer (1976)}
\]

\[V = \text{molar volume, cm}^3/\text{mol}\]

To assure solvent is liquid at 1 atm

\[T_b > 323 \text{ K}, \quad T_m < 223 \text{ K}\]

\[
T_b = T_{b0} \ln \left(\sum_{i=1}^{N} n_i T_{b_i}\right)
\]

\[
T_m = T_{m0} \ln \left(\sum_{i=1}^{N} n_i T_{m_i}\right)
\]

To reduce the drying time, minimize \(\Delta H^v\)

\[
\Delta H^v = H_0^v + \sum_{i=1}^{N} n_i H_i^v
\]

Also, must keep solvent accumulation in human tissue low

\[
\log\text{BCF} = 0.76 \log K_{ow} - 0.23
\]

where

\[
\log K_{ow} = \sum_{i=1}^{N} \chi_i^0 + 0.12 \sum_{i=1}^{N} \chi_i^1
\]

\(\chi_i^0\) and \(\chi_i^1\) are the fragment and factor of group i

Rule of thumb - \(\log_{10} K_{ow} < 4.0\)
Student Assignment

Rather than solve the constrained optimization problem, suggest several compounds and have students check whether specifications are achieved. If not, students suggest alternate compounds –

see Exercise 2.5 – SSL (2004)