**1. Glass Transition Temperature（Tg，玻璃转变温度）**

* **定义**：聚合物从“硬脆的玻璃态”变成“柔韧的橡胶态”时的温度。
* **物理意义**：低于 Tg，链段运动受限，材料像硬塑料；高于 Tg，链段可以局部运动，材料表现为柔软、弹性。
* **生活例子**：
  + **聚苯乙烯（PS，凡士林杯、一次性餐具）**：Tg ≈ 100 °C，室温 (~20 °C) 明显硬脆。
  + **天然橡胶**：Tg ≈ –70 °C，室温下非常柔韧。
  + **聚甲基丙烯酸甲酯（PMMA，亚克力）**：Tg ≈ 105 °C，用于透明面板，耐高温。

**2. Thermal Conductivity（Tc，热导率）**

* **定义**：单位厚度、单位温差下的热流量，常用单位 W/(m·K)。
* **物理意义**：热导率越高，材料越“导热”；越低，则越“保温隔热”。
* **生活例子**：
  + **泡沫塑料（保温材料，如聚苯乙烯泡沫）**：Tc ≈ 0.03 W/(m·K)，极佳保温。
  + **普通聚乙烯（PE）**：Tc ≈ 0.4 W/(m·K)，常见塑料袋、薄膜绝缘性能一般。
  + **金属（对比）**：铝约 200 W/(m·K)，导热非常快。

**3. Density（密度）**

* **定义**：单位体积内的质量，常用 g/cm³ 或 kg/m³。
* **物理意义**：同体积下质量的多少，影响材料强度、浮沉、包装重量。
* **生活例子**：
  + **聚丙烯（PP，可漂浮在水面）**：≈ 0.90 g/cm³，轻、可浮于水。
  + **聚对苯二甲酸乙二酯（PET，塑料瓶）**：≈ 1.38 g/cm³，比水重，会下沉。
  + **聚氯乙烯（PVC，给水管）**：≈ 1.4 g/cm³，坚固但略重。

**4. Radius of Gyration（Rg，回转半径）**

* **定义**：从分子质心到所有原子位置的均方根距离，度量高分子链在空间中的“伸展”尺寸。
* **物理意义**：Rg 越大，链越“张开”，溶液粘度和分子间缠结越强；
* **生活例子**：
  + 在**配制润发露或凝胶**时，高 Rg 的聚合物（如高分子聚乙烯醇）会显著增稠、形成丝滑手感。
  + **低分子量聚乙二醇（PEG）** Rg 只有几个纳米；而用于粘度改性的高分子 Rg 可达几十、上百纳米。

**5. Fractional Free Volume（FFV，分数自由体积）**

* **定义**：材料总体积中“没有被高分子链占据”的那部分体积分数。
* **物理意义**：FFV 越大，分子链之间空隙越多，气体/水分子透过更易——影响阻隔性和扩散性。
* **生活例子**：

| **聚合物类型** | **FFV** | **主要影响** |
| --- | --- | --- |
| **低 FFV（≈0.1）** | PVDC 真空包装 | 极低气体透过率（优秀阻隔） |
| **中等 FFV（≈0.2–0.3）** | PET 瓶 | 中等透过率（普通饮料瓶） |
| **高 FFV（≈0.4–0.5）** | PDMS（硅橡胶） | 非常高的气体透过性（用于气体分离膜） |

* **PVDC（聚偏二氯乙烯）**：FFV ≈ 0.10，几乎是小分子透不过去的“密实”材料，用于食品包装。
* **PET（聚对苯二甲酸乙二醇酯）**：FFV ≈ 0.20–0.25，透氧透水都在可控范围，所以常做饮料瓶。
* **PDMS（聚二甲基硅氧烷）**：FFV ≈ 0.45–0.50，因为分子链极其松散，多孔隙，非常透气，常用作气体分离和透气膜。

**小结表**

| **指标** | **含义** | **典型聚合物例子** | **取值范围（约）** |
| --- | --- | --- | --- |
| Tg | 玻璃态 ↔ 橡胶态 转变温度 | PS、PMMA、橡胶 | –100 ℃ ~ +200 ℃ |
| Tc | 导热能力 | 泡沫、PE、金属（对比） | 0.02 ~ 0.5 W/mK |
| 密度 | 单位体积质量 | PP (0.9)、PET (1.38) | 0.8 ~ 1.6 g/cm³ |
| Rg | 链在溶液/熔体中的“尺寸” | PEG、PVA… | 几 nm ~ 几百 nm |
| FFV | 链间空隙体积分数 | PE 薄膜(25%)、PVDC(10%) | 0.1 ~ 0.5 |

了解了这五个指标的物理意义和现实例子，就能更好地把“SMILES → 分子结构特征”映射到它们在日常材料中的表现。

**1. 元素与官能团组成（Composition & Functional Groups）**

* **元素含量**：C/H/O/N/Cl… 的原子百分比（at %），特别是氧、氮等极性原子的比例，会显著影响 Tg 和相互作用强度。
* **官能团计数**：羰基 (C=O)、羟基 (–OH)、酯 (–COO–)、胺 (–NH₂/–NH–) 等关键基团的个数或密度（per repeat unit）。
* **极性/氢键能力**：H-bond donor/acceptor 数量，影响链间作用力，进而影响 Tg、密度和热导率。

**示例**

bash

CopyEdit

–C=O count

–OH count

H-bond donors (#NH, #OH)

H-bond acceptors (#O, #N)

**2. 骨架刚性与链柔性（Backbone Rigidity & Flexibility）**

* **可旋转键数（Rotatable Bonds）**：越多链段越柔软，Tg 通常越低；
* **芳香环/刚性环系比例**：苯环、杂环比例越高，链段刚性↑ → Tg↑；
* **分支度（Branching Index）**：侧链越多，链段间缠结和 packing 复杂度↑，影响 Rg 和 FFV。

**示例**

java

CopyEdit

NumRotatableBonds

AromaticRingCount

BranchCount (从 SMILES 括号“(”和“)”次数估算)

**3. 体积与形状描述符（Size & Shape）**

* **Molecular Weight of repeat unit**：分子量越大，链段“尺寸”越大，Rg ↑；
* **分子体积估算**：van der Waals 体积或 Connolly surface area（可用 RDKit 或 Mordred 计算）；
* **拓扑极矩（Wiener Index, Radius of Gyration of monomer）**：粗略反映 monomer 的伸展性。

**4. 拓扑/图论特征（Topological Descriptors）**

* **连接指数（Zagreb indices, Balaban’s J）**：反映分子图的分支和环路复杂度；
* **环系指标**：环闭合数、平均最短路径长度（Average Shortest Path）；
* **聚合物专用**：从 SMILES 中提取重复单元，用连通分量性质（component complexity）。

**5. 极性与疏/亲水性（Polarity & Hydrophobicity）**

* **LogP 或 ClogP**：反映分子疏水/亲水平衡，影响聚合物链间相互作用和 free volume；
* **TPSA（Topological Polar Surface Area）**：极性表面积越大，链间吸附越强，密度↑、FFV↓。

**6. 群贡献／经验参数（Group‐Contribution Features）**

Van Krevelen、Hoy 等聚合物性质的群贡献方法早已给出针对 Tg、密度、FFV 的各官能团贡献系数：

* 对每种基团（–CH₂–、–C(CH₃)₃、–COO–…）计数后，按文献系数线性组合
* 既可做可解释的“基团贡献”特征，也能在样本较少时稳定建模

**7. 指纹与嵌入（Fingerprints & Embeddings）**

* **Morgan／ECFP 指纹**（radius=2–3）：自动捕捉局部环境和支链模式；
* **Graph‐Neural‐Net 嵌入**：用 DGL、PyTorch‐Geometric 对 SMILES-graph 做端到端学习。

**8. 文本／序列级特征（SMILES-String Stats）**

* **SMILES 长度**、**括号深度**、**数字（环标记）频次**——反映分支和环路复杂度；
* **n-gram 统计**：比如 CC=O、C(C)(C) 这样的常见子串出现频率。

**实战建议**

1. **工具包**：RDKit、Mordred、ChemDes 均可自动批量算到上面大部分描述符。
2. **特征筛选**：先做单变量相关性（Tg vs. each descriptor），然后用 LASSO/树模型挑最重要的特征。
3. **聚合物特殊性**：先把“重复单元” SMILES 切出来（两端的 \*），再针对这个片段做分子量、体积、拓扑的计算。
4. **组合法**：物理化学＋图论＋指纹，混合建模通常比单一特征类别更稳健。