

# 使用 MolAICal 计算分子合成可行性预测、类药五原则和 PAINS 的教程

作者：Qifeng Bai (update 2020-08-03)

更多教程（含英文教程）请见如下：

MMolAICal 官方主页：<https://molaical.github.io>

MolAICal 官方主页中国镜像：<https://molaical.gitee.io>

MolAICal 中文博客：<https://molaical.gitee.io/cntutorial.html>

## 1. 简介

在本教程中，我们将介绍药物合成可行性预测，Lipinski's rule of five（里宾斯基五规则）和 PAINS 的计算方法。药物可合成性（SA）可用于预测化合物的合成难度。里宾斯基五规则也被称为类药五原则（RO5），是一种用来估计类药性或评估一种化合物是否具有药理或生物活性的法则。PAINS 全称是 Pan Assay Interference Compounds，指对大多数测试有干扰的化合物，PAINS 大致分为两类：对第一类来说，化合物在测试浓度下在溶液中形成胶体，蛋白被包裹在胶体中，底物因此无法接近酶的活性中心，因此这些化合物属于假阳性化合物；对于第二类来说，这些化合物的活性基团可以与蛋白受体形成共价作用，进而导致受体受到抑制，但这种抑制很难逆转，同时这些具有 PAINS 性质的配体可以与大多数靶点发生反应，缺乏特异性。如果药物相关文章要投稿到 Journal of Medicinal Chemistry 等杂志，它们可能要求 PAINS 的计算，此时，可以使用 MolAICal 来计算 PAINS。

## 2. 材料和工具

### 2.1. 软件需求

MolAICal: <https://molaical.github.io> 或 <https://molaical.gitee.io>

### 2.2. 示例文件

所有需要的教程文件可以从以下链接下载

[https://gitee.com/molaical/tutorials/tree/master/009-SA\\_Ro5\\_Pains](https://gitee.com/molaical/tutorials/tree/master/009-SA_Ro5_Pains)

## 3. 步骤

### 3.1. 可合成性（SA）预测计算

你可以通过分子的 SMILES 序列计算 SA，使用如下命令：

```
#> molaical.exe -tool sa -i "FC(F)(F)c1cc(ccc1)N5CCN(Cc2mnc3[C@H]4CCC[C@H]4Cn23)CC5"
```

或者你可以在一个包含许多 SMILES 序列的文件中计算配体的 SA

```
#> cd "009-SA_Ro5_Pains/SA"
```

```
#> molaical.exe -tool sa -i SmilTest.smi
```

**注意：**化合物 SA 的值越大，表明该化合物越易于合成。

### 3.2. 类药五原则计算

单个配体的 RO5 计算如下

```
#> cd "009-SA_Ro5_Pains/ro5"
```

```
#> molaical.exe -tool ruleoffive -f mol2 -n zinc_1879871.mol2
```

如果要计算多个配体的 **Lipinski's rule of five**，可以按以下步骤进行：

1. 在 Linux 的控制台中使用命令“ls > mol2List.dat”，或在 Windows 的 DOS 控制台中使用命令“dir /b > mol2List.dat”。打开生成的文件“mol2List.dat”并且删除所有无用的字符。确保文件“mol2List.dat”只包含配体的名称。

2. 执行如下命令：

```
#> molaical.exe -tool ruleoffive -f mol2list -i mol2List.dat -o result.dat
```

将会生成名为“result.dat”的文件，其中包含大量的 RO5 值。关于 RO5 的更多细节，请参阅 MolAICal 手册。

### 3.3. PAINS 计算

具有 SMILES 或 mol2 格式的单个配体可计算如下：

```
#> cd "009-SA_Ro5_Pains/pains"
```

```
#> molaical.exe -tool pains -f smi -n "c1ccccc1N=Nc1ccccc1"
```

```
#> molaical.exe -tool pains -f mol2 -n ZINC00154323.mol2
```

如果你想要计算多个配体的 PAINS，可以执行以下步骤：

对于 SMILES 格式：

```
#> molaical.exe -tool pains -f smilist -i painsTest.txt -o smiResult.txt
```

对于 Mol2 格式：

```
#> cd "009-SA_Ro5_Pains/pains/mol"
```

1. 在 Linux 控制窗口使用命令“ls > mol2List.dat”，或者在 Windows 的 DOS 窗口使用命令“dir /b > mol2List.dat”。打开生成的文件“mol2List.dat”并且删除所有无用的字符。确保文件“mol2List.dat”中只包含配体的名称。

2. 执行如下命令：

```
#> molaical.exe -tool pains -f mol2list -i mol2List.dat -o mol2Result.txt
```

将会生成名为“mol2Result.txt”的文件，其中包括所有配体的 PAINS 信息。

**注意：**建议使用 SMILES 格式进行 PAINS 的计算。