

pChem 使用说明

Zhengcong Fei
Institute of Computing Technology
Chinese Academy of Sciences



中国科学院计算技术研究所
Institute of Computing Technology, Chinese Academy of Sciences



pChem快速使用方法

■ 一. quick start

- 1. 写pChem.cfg文件
- 2. 运行命令 `python quick_start.py` 进行一键修饰搜索

要求 python 3.x 版本



■ 二. Professional Usage

□ 1. 写pChem.cfg文件

□ 2. 运行命令 `python common_modification_search.py` 进行开放式搜索

查看 `common_modification_list.txt` 选择常见修饰（不选择则直接删除）

□ 3. 运行命令 `python blind_search.py` 进行盲搜

查看 `mass_diff_list.txt` 选择未知修饰（不选择直接删除）

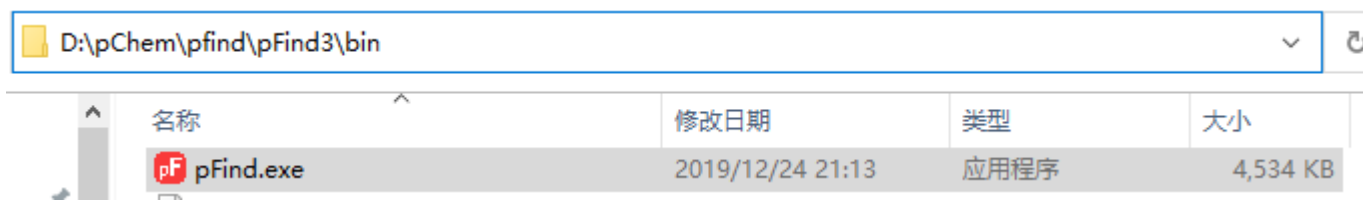
□ 4. 运行命令 `python close_identify.py`



参数文件填写

■ 1. pFind安装路径

- 因为需要调用pFind，所以需要填写路径。
- 操作：右键桌面pfind图标 -> 打开文件夹所在位置



选择bin文件夹的上一层位置，即 D:\pChem\pfind\pFind3



■ 2. fasta文件和pf2文件路径

- 如果是raw文件，则需要调用pParse做格式转换。

```
fasta_path=D:\pChem\pfindworkpFindTask1\fasta\uniprot-reviewed-20200326_con.fasta
```

```
msmsnum=1
```

```
msmspath1=D:\pChem\IPM\raw\QE_Plus_YJ_HJX_IPM_PH8_50per_20190730_F1_R1_HCDFT.pf2
```

如果有多个raw文件，则可以写为：

```
msmsnum=2
```

```
msmspath1=...
```

```
msmspath2=...
```



■ 2. fasta文件和pf2文件路径

- 因为目前只支持pf2文件，如果是raw文件等其他格式，可以使用文件目录下的pParse进行格式转化。

- 使用方法，执行以下命令，当前目录下会生成PF2文件

`pParse.exe -D 需要转化文件地址`

例如：

`pParse.exe -D D:\human.raw`



■ 3. 输出结果保存路径

```
output_path=D:\pChem\pChem\results\
```



■ 4. 开放式搜索模块

- 是否使用开放式刻画数据 True/ False

`open_flag=False`

- 如果使用开放式，则需要指定自动选择常见修饰的个数

`common_modification_number=3`

- 如果跳过开放式搜索，则需要指定常见修饰的类型

`common_modification_list=Carbamidomethyl[C];Oxidation[M];`



■ 5. 盲搜模块

- 是否引入同位素标签，没有则填 -1.0

如有6个C同位素标记，则为 6×1.00335

`mass diff diff=6.0201`

- 质量数必须大于设定阈值 单位：Da

`min_mass_modification=200.0`



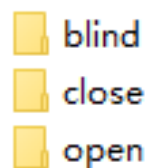
- 5. 限定式搜索模块
 - 进入限定式搜索的未知修饰个数

`close_mass_diff_number=2`



pChem结果输出

- 1. 三次（二次） pFind鉴定结果， 分别对应每次搜索， 存在于结果输出路径：



- 2. pChem.summary 盲搜结果统计信息：

pChem.summary - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

rank	mass modification	Total number	position: occurrence times
1	PFIND_DELTA_258.14	1207	{'C': 1170, 'N-SIDE': 105, 'A': 21, 'E': 3, 'N': 2, 'G': 2, 'P': 2, 'Y': 1, 'D': 1, 'V': 1, 'S': 1, 'H': 1, 'C_SIDE': 1, 'T': 1, 'L': 1}
2	PFIND_DELTA_258.15	996	{'C': 895, 'N-SIDE': 136, 'A': 62, 'P': 7, 'E': 6, 'V': 5, 'M': 4, 'S': 4, 'Q': 3, 'G': 2, 'T': 2, 'I': 2, 'L': 1, 'N': 1, 'R': 1, 'F': 1}
3	PFIND_DELTA_252.12	883	{'C': 851, 'N-SIDE': 79, 'A': 25, 'G': 3, 'P': 2, 'T': 1, 'N': 1}
4	PFIND_DELTA_252.13	602	{'C': 558, 'N-SIDE': 112, 'A': 18, 'P': 6, 'I': 5, 'E': 4, 'G': 3, 'N': 2, 'Y': 1, 'D': 1, 'V': 1, 'L': 1, 'F': 1, 'Q': 1}
5	PFIND_DELTA_274.14	160	{'C': 78, 'M': 44, 'N-SIDE': 36, 'A': 14, 'L': 5, 'F': 4, 'G': 3, 'N': 2, 'I': 2, 'P': 2, 'S': 1, 'T': 1, 'E': 1, 'D': 1, 'Q': 1, 'V': 1}
6	PFIND_DELTA_268.12	135	{'M': 59, 'C': 39, 'N-SIDE': 25, 'A': 16, 'I': 5, 'G': 4, 'P': 3, 'L': 2, 'F': 2, 'S': 1, 'E': 1, 'T': 1, 'D': 1, 'K': 1}
7	PFIND_DELTA_258.16	66	{'C': 39, 'A': 14, 'N-SIDE': 13, 'S': 4, 'M': 2, 'P': 1, 'Q': 1, 'D': 1, 'F': 1, 'I': 1, 'V': 1, 'W': 1}
8	PFIND_DELTA_252.14	38	{'C': 29, 'A': 5, 'N-SIDE': 5, 'E': 2, 'D': 1, 'Q': 1}
9	PFIND_DELTA_274.15	34	{'C': 12, 'M': 11, 'A': 7, 'N-SIDE': 7, 'P': 2, 'T': 1, 'S': 1}
10	PFIND_DELTA_315.17	37	{'C': 20, 'A': 14, 'N-SIDE': 5, 'S': 2, 'L': 1}
11	PFIND_DELTA_309.15	33	{'C': 28, 'N-SIDE': 5, 'A': 2, 'E': 1, 'N': 1, 'M': 1}
12	PFIND_DELTA_315.16	25	{'C': 16, 'A': 6, 'N-SIDE': 2, 'S': 1, 'F': 1, 'M': 1}



分子式推断工具

□ 输入 分子量, 元素列表, 误差范围 (单位ppm) , 输出可能的分子式

□ 例: 分子量256.0, 元素列表CHNOS, 误差范围20ppm, 则输入命令:

```
python composition_inference.py -mass 256.0 -element_list
```

```
CHNOS -error_range 20.0
```