

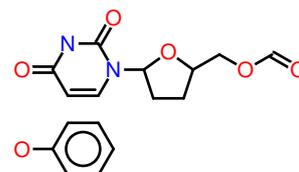
## 10. Übungsblatt

### Aufgabe 36 Molekülbeschreibungssprachen

a) Zeichnen Sie das durch die folgende SMILES-Zeichenkette beschriebene Molekül:  
S2c1c4c(ccc1N(c3c2cccc3)C(=O)C)cccc4

b) Zeichnen Sie das durch die folgende SLN-Zeichenkette beschriebene Molekül:  
NH2C[1]:N:C(:C(:N:C:@1C(=O)OCH3)C(=O)C[2]:CH:CH:CH:CH:CH:@2)NH2

c) Erzeugen Sie eine SMILES-Beschreibung dieses Moleküls:



d) Finden sie mindestens drei verschiedene SMILES-Beschreibungen von Phenol, d.h., dieses Moleküls:

### Aufgabe 37 Molekülbeschreibungssprachen

Zeichnen Sie das durch das folgende SDfile beschriebene Molekül:

```
4728
1016 111283D
4728
9 9 0 1 V2000
-0.0965 1.3884 0.0104 N 0 0 0 0 0 0 0 0 0 0
1.8297 -0.2821 -0.0166 S 0 0 0 0 0 0 0 0 0 0
1.0302 2.1371 0.0038 N 0 0 0 0 0 0 0 0 0 0
0.1607 0.0664 0.0009 C 0 0 0 0 0 0 0 0 0 0
2.1519 1.3924 -0.0108 C 0 0 0 0 0 0 0 0 0 0
3.3971 1.9238 -0.0196 N 0 0 0 0 0 0 0 0 0 0
-0.5873 -0.6698 0.0040 H 0 0 0 0 0 0 0 0 0 0
3.5072 2.8871 -0.0156 H 0 0 0 0 0 0 0 0 0 0
4.1746 1.3453 -0.0300 H 0 0 0 0 0 0 0 0 0 0
1 4 2 0 0 0 0
1 3 1 0 0 0 0
2 4 1 0 0 0 0
2 5 1 0 0 0 0
3 5 2 0 0 0 0
4 7 1 0 0 0 0
5 6 1 0 0 0 0
6 8 1 0 0 0 0
6 9 1 0 0 0 0
M END
$$$$
```

### Aufgabe 38 Teilgraphisomorphismen

- Wie oft tritt das Fragment C-C-C in Phenol auf (siehe Aufgabe 36d)?  
Mit anderen Worten: Wie viele Teilgraph-Isomorphismen gibt es?
- Besitzt Phenol (siehe Aufgabe 36d) einen von der Identität verschiedenen Automorphismus? Wie viele verschiedene Automorphismen besitzt Phenol?
- Erzeugen Sie alle zusammenhängenden Teilgraphen von Phenol (siehe 36d)!
- Finden Sie alle Brücken in Phenol (siehe Aufgabe 36d)!  
Sind irgendwelche dieser Brücken echte Brücken?
- Finden Sie alle Brücken in dem Graphen/Molekül aus Aufgabe 36c)!  
Welche dieser Brücken sind echte Brücken?

**Aufgabe 39** Suche nach häufigen Teilgraphen

- a) Finden Sie in der rechts gezeigten Datenbank (SMILES) alle häufigen (Teil-)Graphen, die Schwefel enthalten; minimaler Support  $s_{\min} = 3$ :
- |            |          |
|------------|----------|
| CCS(O)(O)N | CCS(=N)N |
| CCS(O)(C)N | CS(=N)N  |
| CS(O)(C)N  | CS(=N)O  |
- b) Warum ist es bei der Suche nach häufigen Teilgraphen schwieriger, redundante Suche zu vermeiden, als bei der Suche nach häufigen Itemmengen? Worin bestehen die Hauptprobleme?
- c) Was ist der Zweck der Konstruktion von kanonischen Kodewörtern von (Teil-)Graphen? Was ist ein kanonisches Kodewort? Was brauchen wir tatsächlich?
- d) Wie werden mit Hilfe eines kanonischen Kodewortes (Teil-)Graphen eindeutige Elterngraphen zugewiesen? Welche Information des Kodewortes wird benutzt?