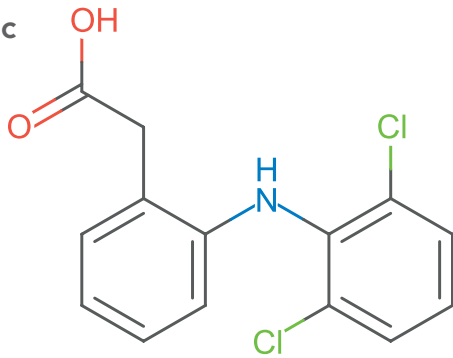# Which fragments are in these molecules?

Diclofenac
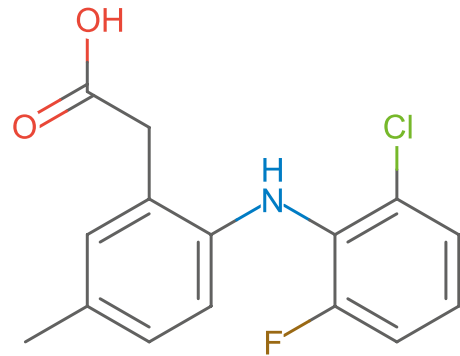
→ Spot the fragments present in each molecule. Write '1' below the corresponding boxes or write '0' if not found.
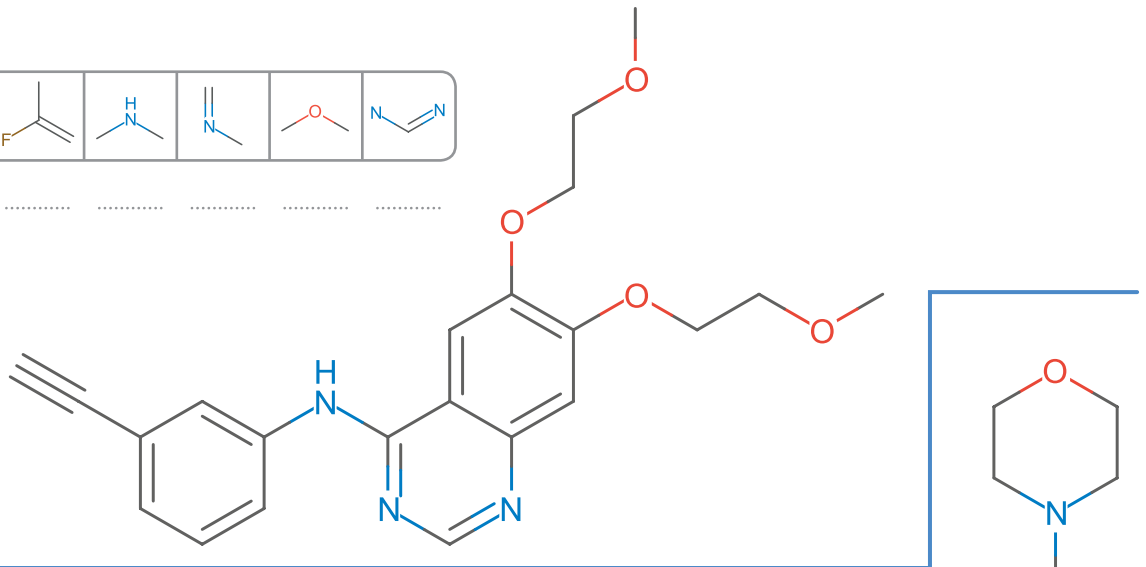
Example:

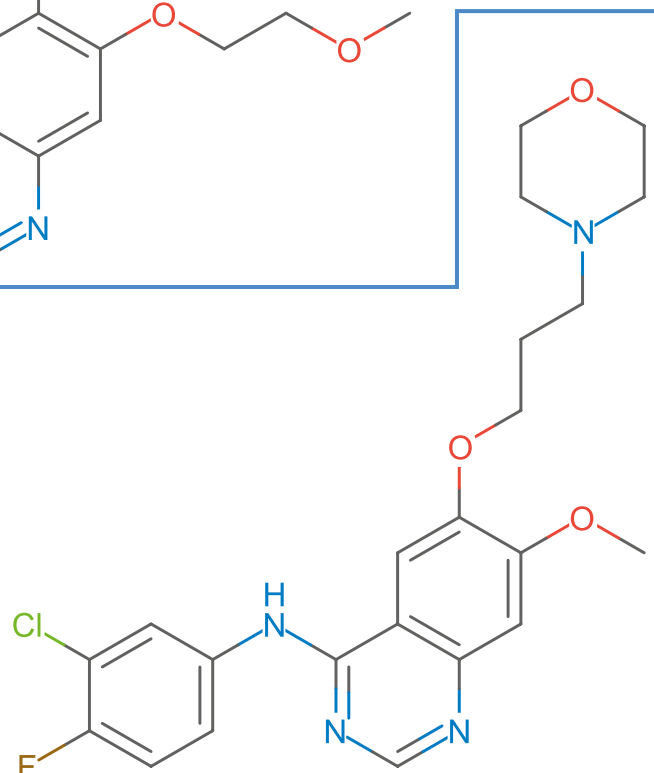| OH/O		Cl	F	H/N	N	O	N N
1	1	1	0	1	0	0	0

Lumiracoxib

Erlotinib

Gefitinib

Similarity calculation
The Tanimoto coefficient

SIB — Swiss Institute of Bioinformatics

Formula

$$T = \frac{M_{11}}{M_{11} + M_{10} + M_{01}}$$

M_{11} → number of times '1' is found in A and B

M_{01} → number of times '0' is found in A and '1' in B

M_{10} → number of times '1' is found in A and '0' in B

$$0 \leq T \leq 1$$

$T = 0$: molecules are completely different

$T = 1$: molecules are identical

Example

A

| 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |

B

| 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |

$M_{11} = 3$, $M_{10} = 1$, $M_{01} = 2$

$$T = \frac{3}{3 + 1 + 2} = 0.5$$

→ Compare all 4 molecules by completing the table below with the Tanimoto coefficients you have calculated:

	Diclofenac	Lumiracoxib	Erlotinib	Gefitinib
Diclofenac	1			
Lumiracoxib		1		
Erlotinib			1	
Gefitinib				1

→ Which molecule pairs are most similar?

....

....