

# QNPR

---

## QNPR descriptors

These descriptors are used for QNPR (Quantitative Name Property Relationship) thus giving their name. The descriptors are derived directly from the compounds name or SMILES strings.

For each molecule either canonical SMILES or IUPAC name are split into fragments of a specified length, which is determined by the configuration. All numbers 0-9 are substituted with § symbol.

Thus we will get for

- CCC: C CC and CCC as descriptors
- c1ccccc1: c § c§ c§c §cc ccc and cc\$ as descriptors

when using fragments of length 1-3

## Parameters

Parameter	Effect	
Fragments from to	Create string fragments with length from x to y	
Minimum fragment count threshold	If there are not at least # occurrences of the pattern in the whole dataset, filter the descriptor out	
Type of fragments	Naming scheme (SMILES, IUPAC, ...)	

## Literature

(1) Thormann M, Vidal D, Almstetter M, Pons M; Nomen Est Omen: Quantitative Prediction of Molecular Properties Directly from IUPAC Names; The Open Applied Informatics Journal; 1 (1), 28-32, 2007.