

高分子を扱うとき

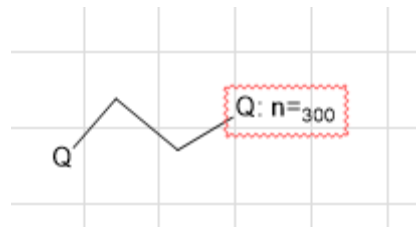
Polysmilesを使います

概要

- SMILESは高分子に非対応ですが、高分子の構造もそこそこ記述できるように、書式を拡張しました(= polysmiles)
- 基本ルール: 繰り返し構造を「Q」として表示
 - nやpdi、結合構造などを補足したいときは、
 - Q: n=123, pdi=1.3, connect=block
という感じにする
 - 低分子の混合物も、一応記述可能(次ページ)

Polyethylene (n=300)

chemdraw

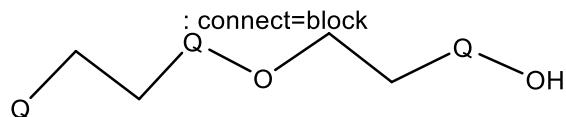


Smiles

(chemdrawから出力可能)

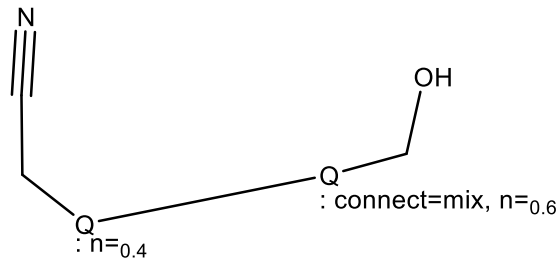
[Q]CC[Q: n=300]

Polyethylene-block-PEO (OH末端)



[Q]CC[Q: connect=block]OCC[Q]O

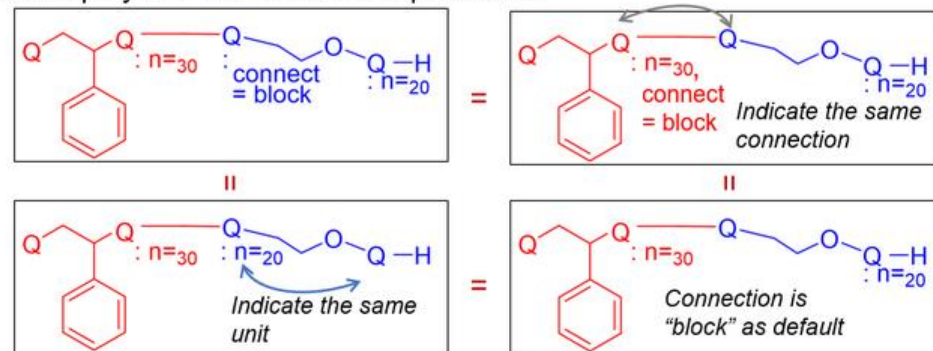
アセニト / メタノール
= 4/6



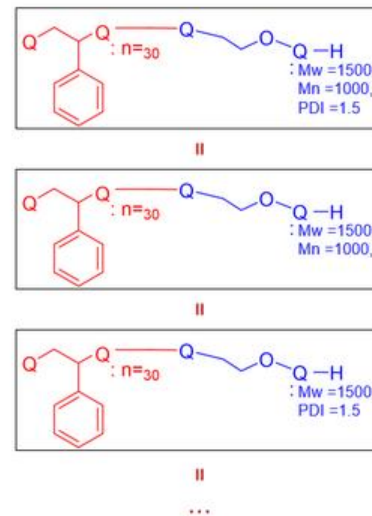
N#CC[Q: n=0.4][Q: connect=mix, n=0.6]CO

細かな表記ルール

Same polymer with different expressions



PolySMILES calculate molecular weights automatically from given information (n, Mw, Mn, PDI)



使い方: Smilesのlistを渡すだけです

```
4  
5 psm=PolySMILES()  
6 desc_df=psm.auto(smiles_list)  
7 desc_df
```

	SMILES	total MW	JR_BoilingPoint	JR_MeltingPoint	JR_CriticalTemp	JR_
0	[Q]CC[Q: n=300]	8416.2	245.36	112.3	401.954	
1	[Q]CC[Q: connect=block]OCC[Q]O	3622.36	301.191	149.331	460.919	
2	N#CC[Q: n=0.4][Q: connect=mix, n=0.6]CO	71.079	333.128	170.549	512.914	
3	C1(CCC1)C	70.135	325.01	160.53	511.698	
4	[O-][N+]#N	44.013	262.3	0	422.504	
...	
111	Cc1cccc(c1)C(C)C	134.222	459.62	226.4	670.516	
112	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	450.88	931.76	450.4	1158.75	
113	CC(C)C(=O)OC	102.133	371.78	173.44	550.251	
114	CCCC(O)CCCC	130.231	474.38	225.74	637.364	
115	CC1(C)CCCCC1	112.216	402.43	211.2	612.782	

その他

- Descriptorは、各成分の加重平均として計算
- 分子量が不明の場合は $n=50$ と仮定
- Block, random, mixなどの繋がり方に関する情報は今のところ、プログラム上は無視