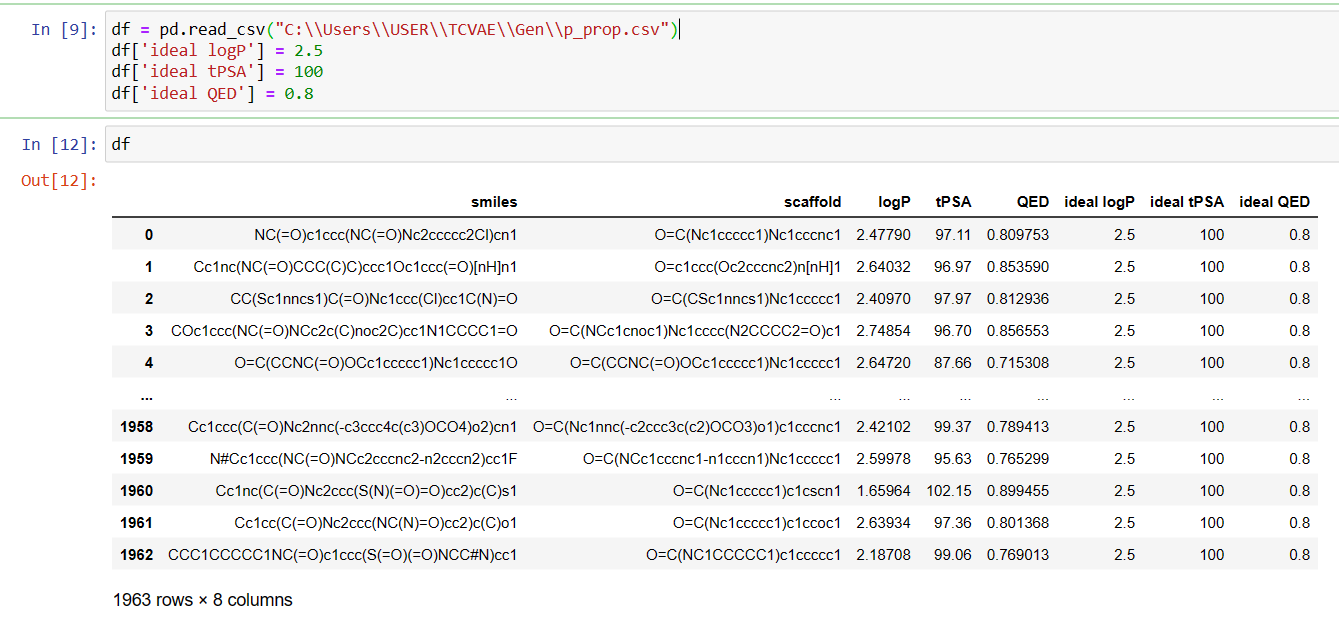
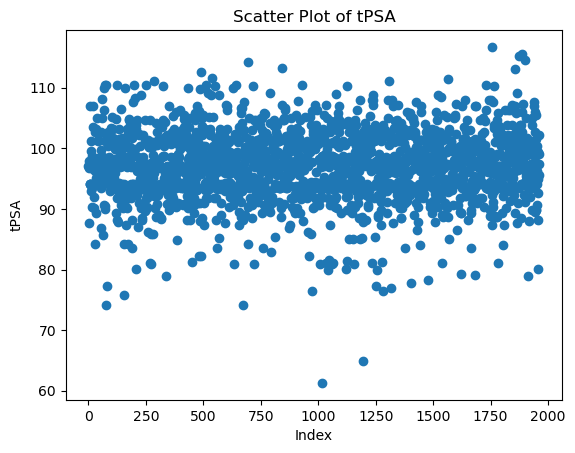
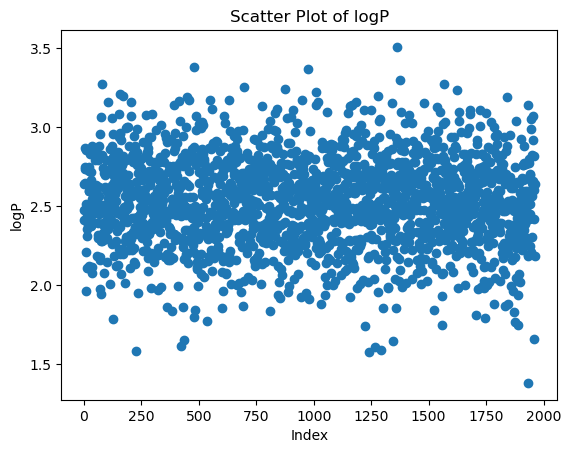
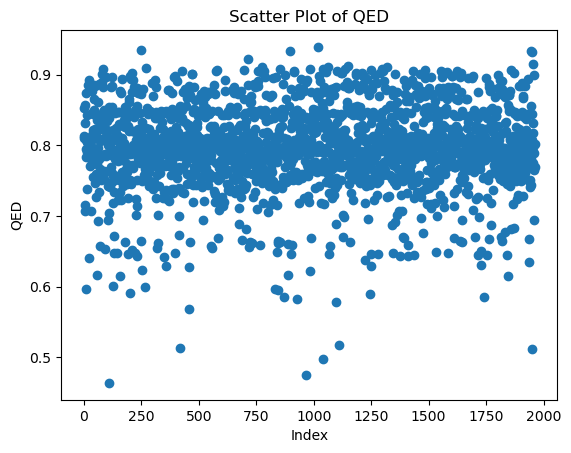
AI in ChE HW6 : molecule design

B08504052 化工四 賴正鎧

1. 我的電腦跑的第一個模擬(“property conditioned sampling”)，有1963個分子成功，比例為1963/2000 = 98.15%。 讀進python的樣子如下:

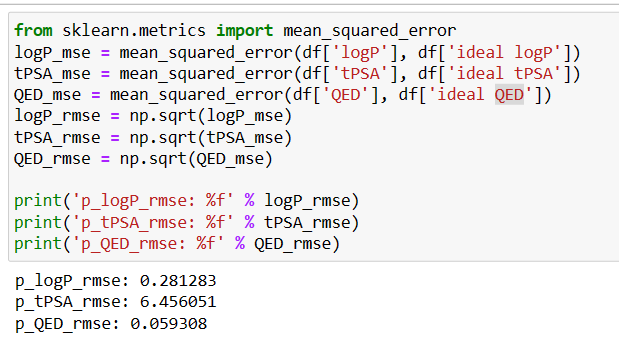


probability distribution of the logP, tPSA, and QED

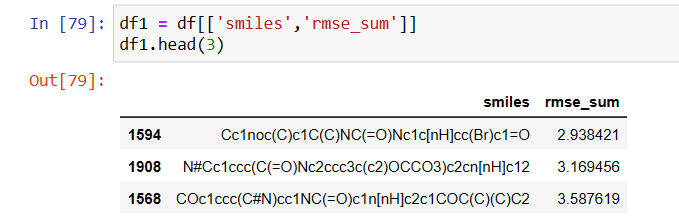
可以看到機率分布還蠻平均的，且確實有集中在我們要求的

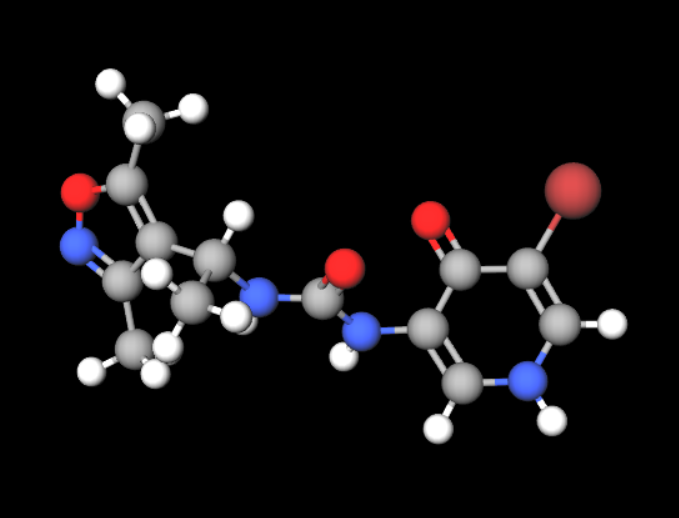
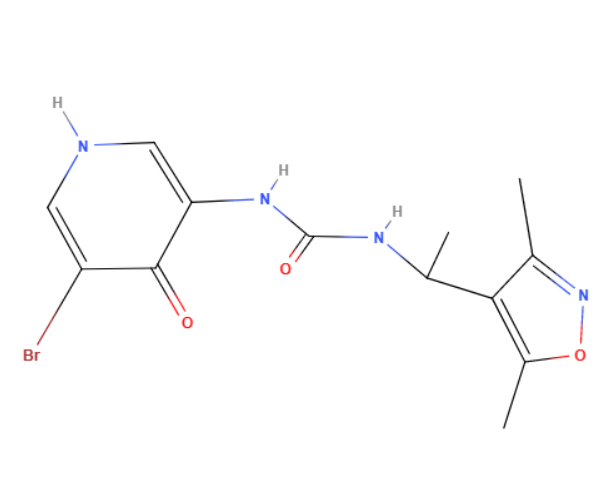
2.5 , 100 , 0.8 身上。

the RMSE in the three target properties

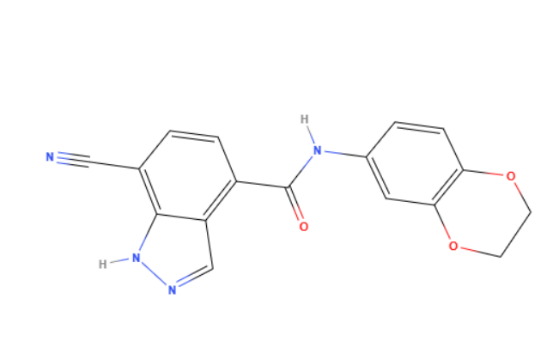
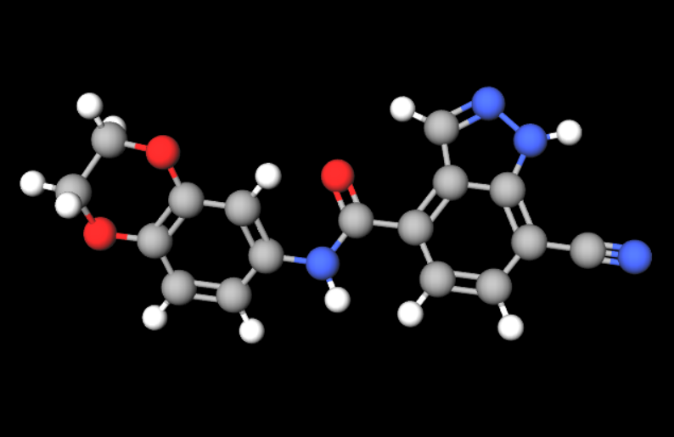


Best 3 molecule (rmse的加總最小者)

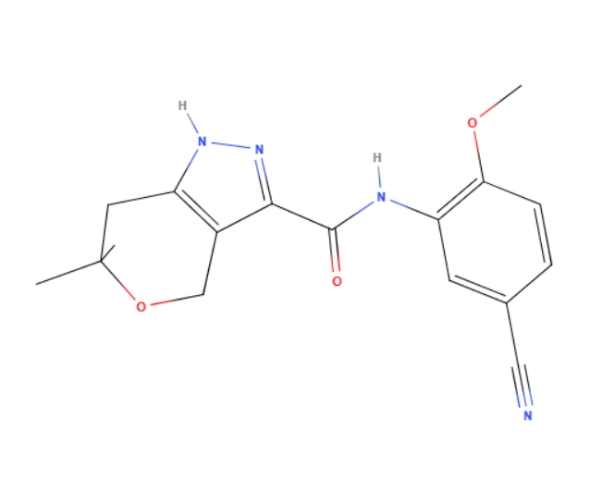


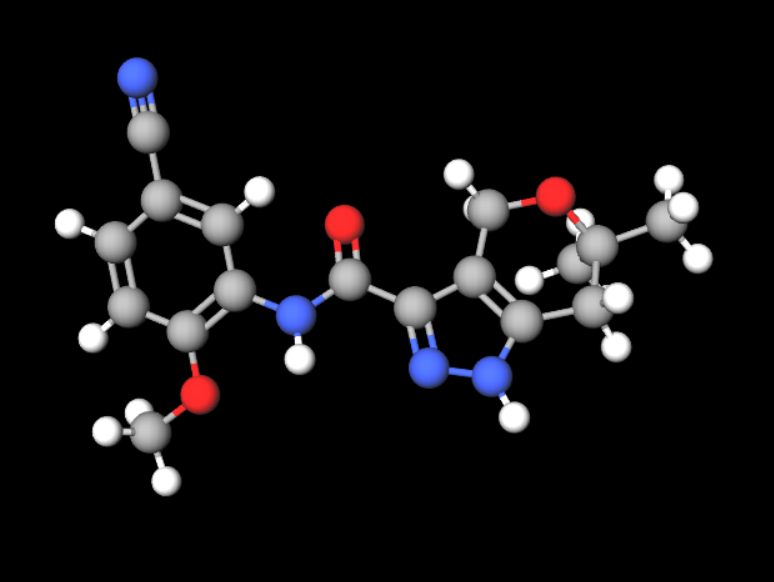
1st: Cc1noc(C)c1C(C)NC(=O)Nc1c[nH]cc(Br)c1=O

2nd: N#Cc1ccc(C(=O)Nc2ccc3c(c2)OCCO3)c2cn[nH]c12

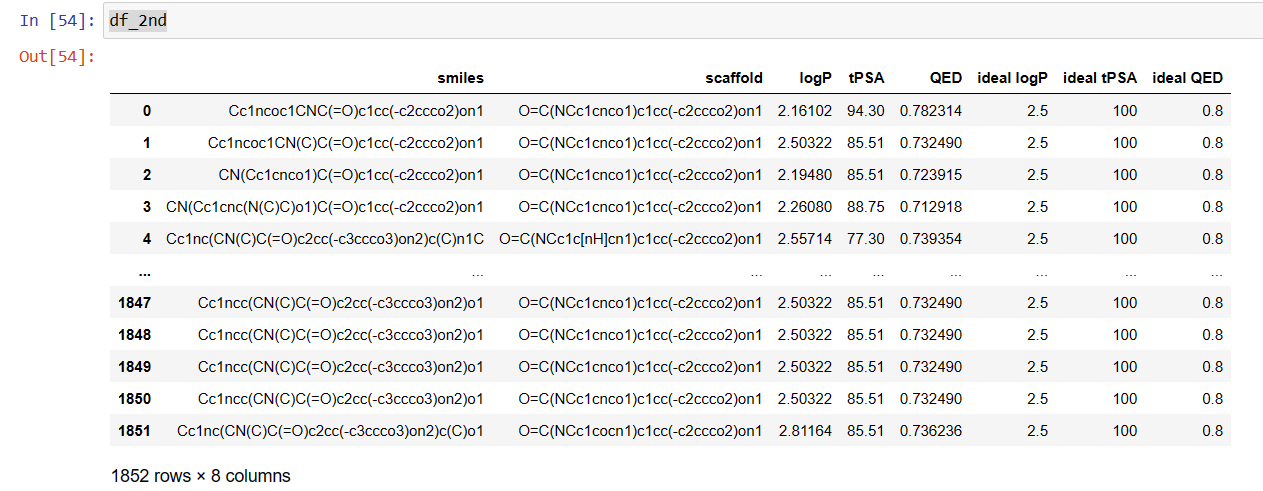


3rd: COc1ccc(C#N)cc1NC(=O)c1n[nH]c2c1COC(C)(C)C2

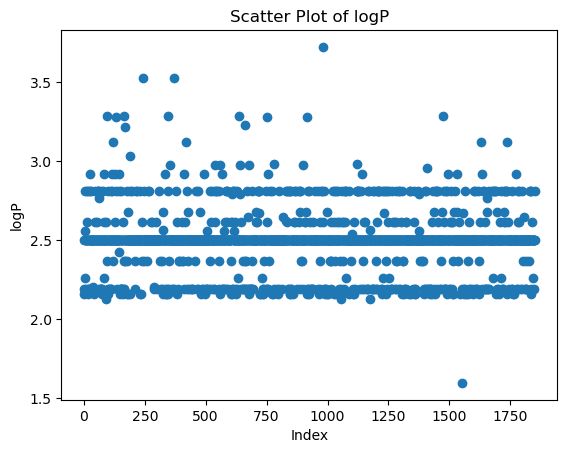
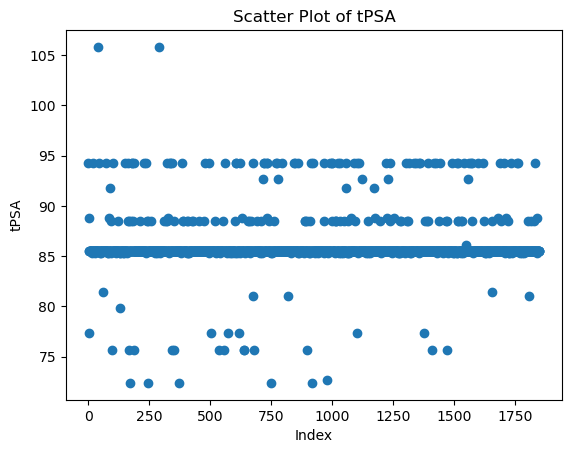


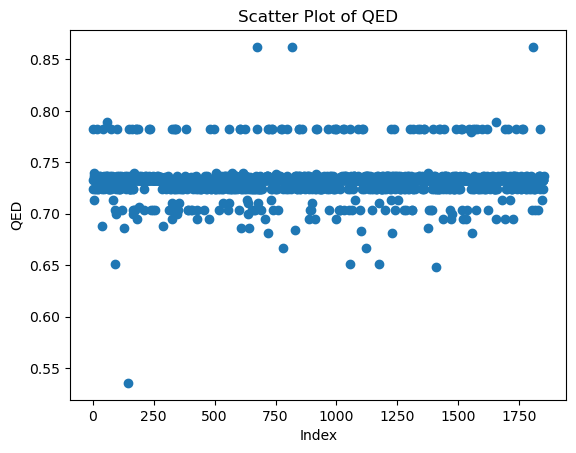


1. 我的電腦跑的第二個模擬(“property and scaffold conditioned sampling”)，有1852個分子成功，比例為1852/2000 = 92.6 %。 讀進python的樣子如下:

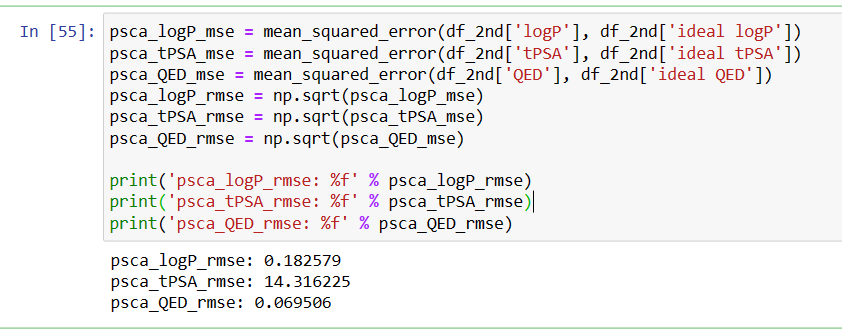


probability distribution of the logP, tPSA, and QED

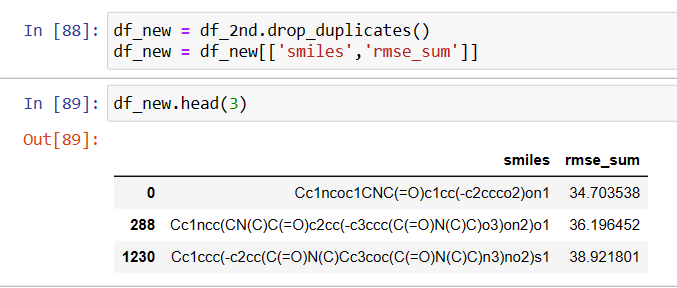
 這裡我發現一個問題，那就是我電腦模擬出來的分子重複了非常多，因此在畫圖的時候可以看得非常明顯，最後發現只有56個分子是獨特的。



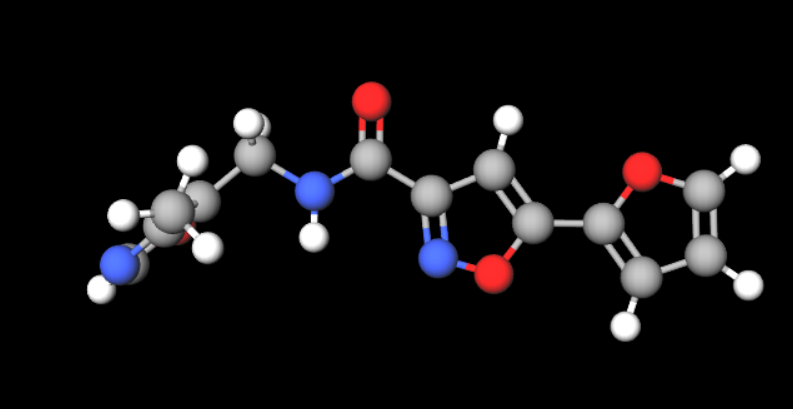
the RMSE in the three target properties

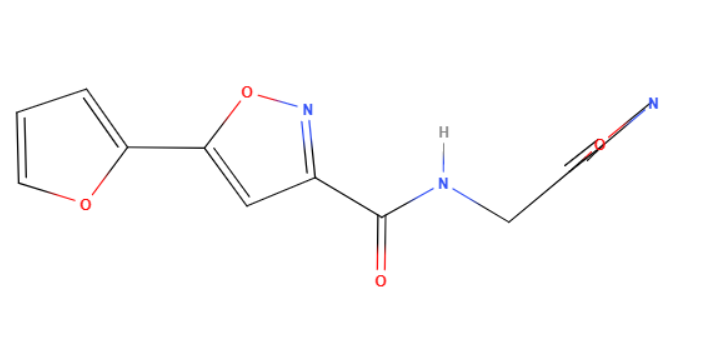


Best 3 molecule (rmse的加總最小者)

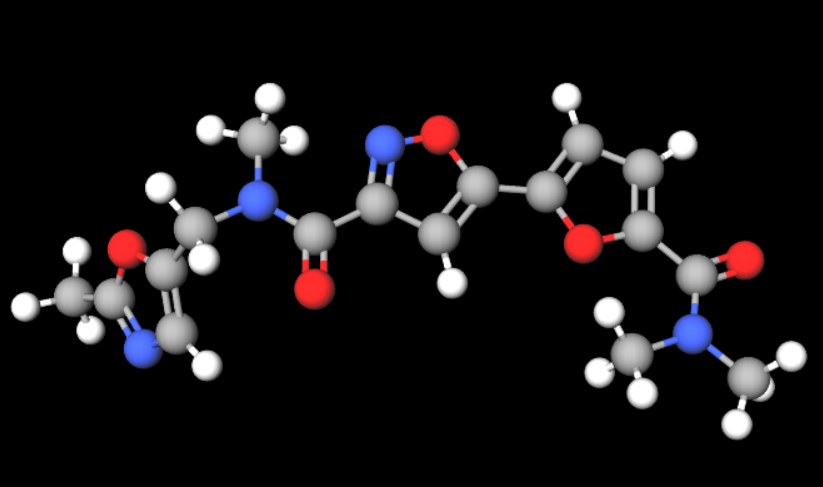


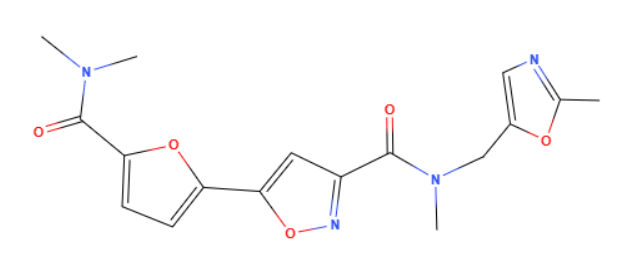
1st : Cc1ncoc1CNC(=O)c1cc(-c2ccco2)on1



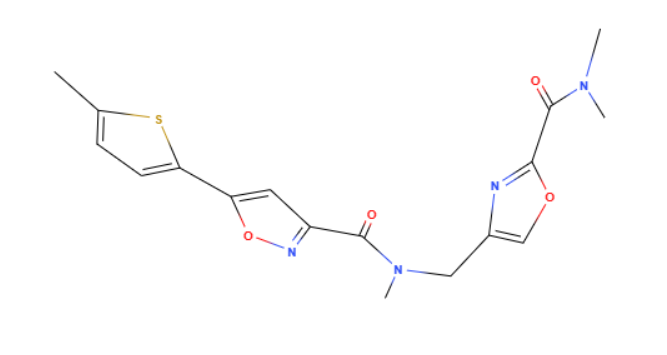
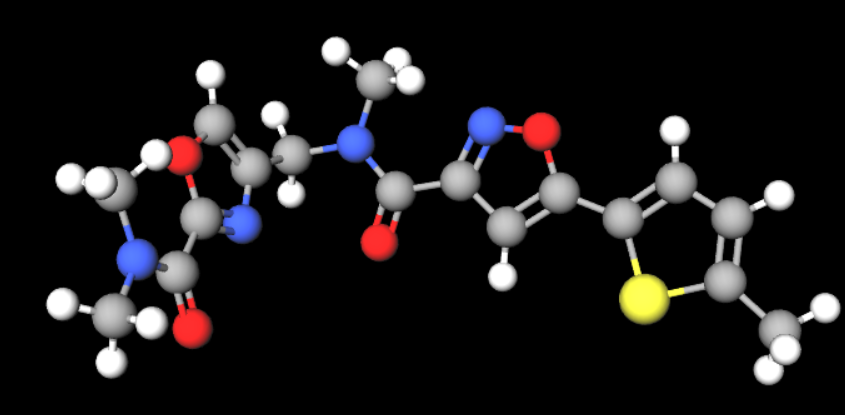


2nd : Cc1ncc(CN(C)C(=O)c2cc(-c3ccc(C(=O)N(C)C)o3)on2)o1





3rd : Cc1ccc(-c2cc(C(=O)N(C)Cc3coc(C(=O)N(C)C)n3)no2)s1



Comparison :

1.加了Murko Scaffold 會讓模擬變困難，在一般中階筆電上，也就是我的電腦上有點難跑，最後跑出來的結果只有56/2000 = 2.8%

2.在Rmse上，兩種模擬方法都不錯，沒有太大的差異，不過第一個相較之下表現更好一些。