



# REINVENT4 强化学习生成报告

登革病毒NS2B-NS3蛋白酶抑制剂从头设计

Run: run9\_t1200 | LibInvent架构 | 吡咯烷双芳环骨架

生成分子总数

**892,916**

有效率 99.3%

平均预测活性

**pIC50 7.56**

中位数 7.64

最高预测活性

**pIC50 8.25**

强效抑制剂范围

高活性分子

**871,165**

pIC50  $\geq$  7.5 (97.6%)

结构多样性

**884,161**

独特R基团组合

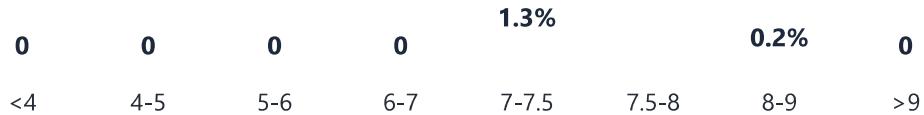
稳定性合格率

**99.5%**

4,491个触发警报

## 预测活性分布 (pIC50)

97.3%



活性高度集中在7.5-8.0区间，标准差0.83表明模型收敛良好。仅0.2%分子达到pIC<sub>50</sub>>8.0，提示该区域探索难度较大。

## 药物化学性质统计

**分子量 (MW)** **487.9 ± 71.6** [200-600]  
**97.2% 达标**

**亲脂性 (LogP)** **4.03 ± 1.31** [0-5]  
**77.7% 达标**

**氢键受体 (HBA)** **5.6 ± 1.5** [0-10]  
**99.9% 达标**

**氢键供体 (HBD)** **1.3 ± 0.6** [0-5]  
**100% 达标**

**极性表面积 (TPSA)** **88.2 ± 23.8** [30-140]  
**97.7% 达标**

**类药性 (QED)** **0.52 ± 0.15** [0-1]  
**100% 达标**

**合成难度 (SA)** **3.91 ± 0.52** [1-6]  
**98.8% 达标**

**Lipinski违规** **588, 933**  
**平均0.66次** 总次数

## 化学稳定性分析

仅0.5%的分子触发不稳定性警报，表明约束条件设置合理。最常见问题是8元环系统(0.2%)和偶氮结构(0.1%)，多卤代CF3仅占0.04%，已被有效控制。

[*;r8] 八元环	1,695次 (0.19%)	[NX3] [NX3] 偶氮	1,270次 (0.14%)
[SH] 硫醇	801次 (0.09%)	多卤代 CF3	320次 (0.04%)
[#8] [#8] 过氧	188次 (0.02%)	C#C 快烃	85次 (0.01%)

## 结构多样性与R基团偏好

884,161种独特R基团组合，平均每种仅出现1.0次，表明极高的结构多样性。模型偏好甲基( $C^*$ )、乙酰基( $CC(*)=O$ )和磺酰基( $S(=O)(=O)C$ )取代，这些基团在前20名中占据主导。

### Top 10 最常见R基团

$C^*$	157,741
$*C$	134,342
$CC\ (*)=O$	15,407
$C\ (*)\ (=O)\ C$	10,346
$*S\ (=O)\ (=O)\ C$	9,466

### Top 10 最高活性候选分子

#1

**pIC50 = 8.25**

CN1CCCCC1C(=O)N1C(c2cccnc2)CC(C(=O)NCC2CCNCC2)C1c1ccncc1



R-groups: N1(C)CCCCC1C(=O)\* | C1CNCCC1C\*

#2

**pIC50 = 8.23**

CN1CCCCC1C(=O)N1C(c2cccc2)CC(C(=O)NCC2CCNCC2)C1c1cccnc1



R-groups: \*C(=O)C1N(C)CCCC1 | C1NCCC(C\*)C1

#3

**pIC50 = 8.22**

CN1CCCC1C(=O)N1C(c2cccc2)CC(C(=O)NCC2CCNCC2)C1c1cccs1

R-groups: C(C1CCN1C)(=O)\* | C1CNCCC1C\*

#4

**pIC50 = 8.21**

O=C(NCC1CCCNCC1)C1CC(c2cccnc2)N(C(=O)c2cc3ncccc3[nH]2)C1



R-groups: [nH]1c2cccnc2cc1C(\*)=O | C1C(C\*)CCNCC1

#5

**pIC50 = 8.20**

CN1CCCCC1CN1C(c2cccnc2)CC(C(=O)NCC2CCNCC2)C1c1ccncc1

R-groups: C1C(C\*)N(C)CCC1 | C1NCCC(C\*)C1

所有Top 10分子均无稳定性警报，包含哌啶、吡啶和吡咯环系统，符合典型NS2B-NS3抑制剂结构特征

## 核心发现与结论

本次强化学习成功生成89万余个DENV NS2B-NS3蛋白酶抑制剂候选分子，平均预测活性pIC<sub>50</sub>达7.56，97.6%的分子达到临床前活性阈值( $pIC_{50} \geq 7.5$ )。结构多样性极高(88万种独特组合)，稳定性合格率99.5%，表明评分函数配置合理。所有Top 10分子均含氮杂环系统(吡啶/哌啶)，与文献报道的NS2B-NS3抑制剂药效团高度一致。分子量集中在 $488 \pm 72$  Da，LogP为 $4.0 \pm 1.3$ ，略高于理想范围但可通过后续优化改善。Lipinski违规率低(平均0.66次/分子)，提示良好的成药性。建议对Top 1000分子进行分子对接验证，并合成Top 20进行体外酶活测试。