Discovering a Green Pesticide Candidate for Controlling Bac-terial Plant Disease: 1,2,3,4-Tetrahydro-β-carboline as a Poten-tial Biofilm Inhibitor

Puying Qi [1, 4]*, Hongwu Liu [2], Yue Li [1], Daiyu Tang [3], Lei Shao [1], Junjie Wang [1, 4], Qifei Zhu [1], Taibai Jiang [1], Lixia Li [1], Shilong Jiang [1], Fan Wu [1], Yanhong Guo [1], Yong Liu[1], Lihong Shi [1,]*, Yin Wang [1,]* and Jun Sun [5,]*.

[1] Guizhou Province Engineering Research Center of Medical Resourceful Healthcare Products, College of Pharmacy, Guiyang Healthcare Vocational University, Guiyang, 550081, China.

[2] National Key Laboratory of Green Pesticide, Key Laboratory of Green Pesticide and Agricultural Bioengineering, Ministry of Education, Center for R&D of Fine Chemicals of Guizhou University, Guiyang, 550025, China.

[3] Guizhou Provincial Center for Disease Control and Prevention，Guiyang, 550004, China.

[4] Guizhou Provincial Doctor Innovation Station, Southeastern Guizhou Province, 556400, China.

[5] College of Chemistry and Chemical Engineering, Guizhou University of Engineering Science, Xueyuan Road, Qixingguan District, Bijie 551700, China. and Analytical and Testing Center, Guizhou University of Engi-neering Science, Xueyuan Road, Qixingguan District, Bijie 551700, China

* Correspondence: qpuying000@163.com or qipuying@gyhvu.edu.cn (P.Q.); shilihong@zju.edu.cn (L.S.); ada_wy@126.com (Y.W.); sun-jun@gues.edu.cn (J.S.)

Table S1. ADMETlab assessment.

1. Physicochemical Property

Property	Value	Comment
Molecular Weight	174.12	Contain hydrogen atoms. Optimal:100~600
Volume	187.227	Van der Waals volume
Density	0.93	Density = MW / Volume
nHA	2	Number of hydrogen bond acceptors. Optimal:0~12
nHD	2	Number of hydrogen bond donors. Optimal:0~7
nRot	0	Number of rotatable bonds. Optimal:0~11

nRing	3	Number of rings. Optimal:0~6
MaxRing	13	Number of atoms in the biggest ring. Optimal:0~18
nHet	2	Number of heteroatoms. Optimal:1~15
fChar	0	Formal charge. Optimal:-4 ~4
nRig	15	Number of rigid bonds. Optimal:0~30
Flexibility	0.0	Flexibility = nRot /nRig
Stereo Centers	2	Optimal: □ 2
TPSA	24.06	Topological Polar Surface Area. Optimal:0~140
logS	-1.673	Log of the aqueous solubility. Optimal: -4~0.5 log mol/L
logP	1.226	Log of the octanol/water partition coefficient. Optimal: 0~3
logD	1.414	logP at physiological pH 7.4. Optimal: 1~3

2. Medicinal Chemistry

Property	Value	Decision	Comment
QED	0.623	●	■ A measure of drug-likeness based on the concept of desirability; ■ Attractive: > 0.67; unattractive: 0.49~0.67; too complex: < 0.34
SAscore	3.196	●	■ Synthetic accessibility score is designed to estimate ease of synthesis of drug-like molecules. ■ SAscore □ 6, difficult to synthesize; SAscore <6, easy to synthesize
Fsp3	0.455	●	■ The number of sp3 hybridized carbons / total carbon count, correlating with melting point and solubility. ■ Fsp^3 □0.42 is considered a suitable value.
MCE-18	47.125	●	■ MCE-18 stands for medicinal chemistry evolution. ■ MCE-18□45 is considered a suitable value.
NPscore	0.653	-	■ Natural product-likeness score. ■ This score is typically in the range from □5 to 5. The higher the score is, the higher the probability is that the molecule is a NP.
Lipinski Rule	Accepted	●	■ MW □ 500; logP □ 5; Hacc □ 10; Hdon □ 5 ■ If two properties are out of range, a poor absorption or permeability is possible, one is acceptable.
Pfizer Rule	Accepted	●	logP > 3; TPSA < 75 Compounds with a high log P (>3) and low TPSA (<75) are likely to be toxic.
GSK Rule	Accepted	●	■ MW □ 400; logP □ 4 ■ Compounds satisfying the GSK rule may have a more favorable ADMET profile
Golden Triangle	Rejected	●	■ 200 □ MW □ 50; -2 □ logD □ 5 ■ Compounds satisfying the Golden Triangle rule may have a more favorable ADMET profile.

PAINS	0 alerts	-	Pan Assay Interference Compounds, frequent hitters, Alpha-screen artifacts and reactive compound.
ALAR M NMR	0 alerts	-	Thiol reactive compounds.
BMS	0 alerts	-	Undesirable, reactive compounds.
Chelator Rule	0 alerts	-	Chelating compounds.

3. Absorption

Property	Value	Decision	Comment
Caco-2 Permeability	-4.987	●	Optimal: higher than -5.15 Log unit
MDCK Permeability	6e-06	●	■ low permeability: < 2 × 10^{-6} cm/s ■ medium permeability: 2–20 × 10^{-6} cm/s ■ high passive permeability: > 20 × 10^{-6} cm/s
Pgp-inhibitor	0.0	●	■ Category 1: Inhibitor; Category 0: Non-inhibitor; ■ The output value is the probability of being Pgp-inhibitor
Pgp-substrate	0.293	●	■ Category 1: substrate; Category 0: Non-substrate; ■ The output value is the probability of being Pgp-substrate
HIA	0.012	●	■ Human Intestinal Absorption ■ Category 1: HIA+(HIA < 30%); Category 0: HIA-(HIA < 30%); The output value is the probability of being HIA+
$F_{20\%}$	0.027	●	■ 20% Bioavailability ■ Category 1: $F_{20\%}$+ (bioavailability < 20%); Category 0: $F_{20\%}$- (bioavailability □ 20%); The output value is the probability of being $F_{20\%}$+
$F_{30\%}$	0.024	●	■ 30% Bioavailability ■ Category 1: $F_{30\%}$+ (bioavailability < 30%); Category 0: $F_{30\%}$- (bioavailability □ 30%); The output value is the probability of being $F_{30\%}$+

4. Distribution

Property	Value	Decision	Comment
PPB	21.79%	●	■ Plasma Protein Binding ■ Optimal: < 90%. Drugs with high protein-bound may have a low therapeutic index.
VD	3.381	●	■ Volume Distribution ■ Optimal: 0.04-20L/kg
BBB Penetration	0.868	●	■ Blood-Brain Barrier Penetration ■ Category 1: BBB+; Category 0: BBB-; The output value is the probability of being BBB+
Fu	66.67%	●	■ The fraction unbound in plasms ■ Low: <5%; Middle: 5~20%; High: > 20%

5. Metabolism

Property	Value	Comment
CYP1A2 inhibitor	0.103	■ Category 1: Inhibitor; Category 0: Non-inhibitor; ■ The output value is the probability of being inhibitor.
CYP1A2 substrate	0.219	■ Category 1: Substrate; Category 0: Non-substrate; ■ The output value is the probability of being substrate.
CYP2C19 inhibitor	0.065	■ Category 1: Inhibitor; Category 0: Non-inhibitor; ■ The output value is the probability of being inhibitor.
CYP2C19 substrate	0.884	■ Category 1: Substrate; Category 0: Non-substrate; ■ The output value is the probability of being substrate.
CYP2C9 inhibitor	0.011	■ Category 1: Inhibitor; Category 0: Non-inhibitor; ■ The output value is the probability of being inhibitor.
CYP2C9 substrate	0.192	■ Category 1: Substrate; Category 0: Non-substrate; ■ The output value is the probability of being substrate.
CYP2D6 inhibitor	0.422	■ Category 1: Inhibitor; Category 0: Non-inhibitor; ■ The output value is the probability of being inhibitor.
CYP2D6 substrate	0.871	■ Category 1: Substrate; Category 0: Non-substrate; ■ The output value is the probability of being substrate.
CYP3A4 inhibitor	0.068	■ Category 1: Inhibitor; Category 0: Non-inhibitor; ■ The output value is the probability of being inhibitor.
CYP3A4 substrate	0.367	■ Category 1: Substrate; Category 0: Non-substrate; ■ The output value is the probability of being substrate.

6. Excretion

Property	Value	Decision	Comment
CL	9.98	●	■ Clearance ■ High: >15 mL/min/kg; moderate: 5-15 mL/min/kg; low: <5 mL/min/kg
$T_{1/2}$	0.386	-	■ Category 1: long half-life ; Category 0: short half-life; ■ long half-life: >3h; short half-life: <3h ■ The output value is the probability of having long half-life.

7. Toxicity

Property	Value	Decision	Comment
hERG Blockers	0.177	●	■ Category 1: active; Category 0: inactive; ■ The output value is the probability of being active.
H-HT	0.652	●	■ Human Hepatotoxicity ■ Category 1: H-HT positive(+); Category 0: H-HT negative(-); ■ The output value is the probability of being toxic.
DILI	0.044	●	■ Drug Induced Liver Injury. ■ Category 1: drugs with a high risk of DILI; Category 0: drugs with no risk of DILI. The output value is the probability of being toxic.
AMES Toxicity	0.918	●	■ Category 1: Ames positive(+); Category 0: Ames negative(-); ■ The output value is the probability of being toxic.
Rat Oral Acute Toxicity	0.825	●	■ Category 0: low-toxicity; Category 1: high-toxicity; ■ The output value is the probability of being highly toxic.
FDAMDD	0.854	●	■ Maximum Recommended Daily Dose ■ Category 1: FDAMDD (+); Category 0: FDAMDD (-) ■ The output value is the probability of being positive.
Skin Sensitization	0.618	●	■ Category 1: Sensitizer; Category 0: Non-sensitizer; ■ The output value is the probability of being sensitizer.
Carcinogenicity	0.09	●	■ Category 1: carcinogens; Category 0: non-carcinogens; ■ The output value is the probability of being toxic.
Eye Corrosion	0.005	●	■ Category 1: corrosives ; Category 0: noncorrosives ■ The output value is the probability of being corrosives.
Eye Irritation	0.037	●	■ Category 1: irritants ; Category 0: nonirritants ■ The output value is the probability of being irritants.
Respiratory Toxicity	0.975	●	■ Category 1: respiratory toxicants; Category 0: respiratory nontoxicants ■ The output value is the probability of being toxic.

8. Environmental toxicity

Property	Value	Comment
Bioconcentration Factors	0.511	■ Bioconcentration factors are used for considering secondary poisoning potential and assessing risks to human health via the food chain. ■ The unit is □log10[(mg/L)/(1000*MW)]
IGC_{50}	2.987	■ Tetrahymena pyriformis 50 percent growth inhibition concentration ■ The unit is □log10[(mg/L)/(1000*MW)]
$LC_{50}FM$	3.056	■ 96-hour fathead minnow 50 percent lethal concentration ■ The unit is □log10[(mg/L)/(1000*MW)]
$LC_{50}DM$	4.875	■ 48-hour daphnia magna 50 percent lethal concentration ■ The unit is □log10[(mg/L)/(1000*MW)]

9. Tox21 pathway

Property	Value	Decision	Comment
NR-AR	0.004	●	■ Androgen receptor ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
NR-AR-LBD	0.002	●	■ Androgen receptor ligand-binding domain ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
NR-AhR	0.224	●	■ Aryl hydrocarbon receptor ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
NR-Aromatase	0.004	●	■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
NR-ER	0.081	●	■ Estrogen receptor ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
NR-ER-LBD	0.007	●	■ Estrogen receptor ligand-binding domain ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
NR-PPAR-gamma	0.002	●	■ Peroxisome proliferator-activated receptor gamma ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
SR-ARE	0.067	●	■ Antioxidant response element ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
SR-ATAD5	0.017	●	■ ATPase family AAA domain-containing protein 5 ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
SR-HSE	0.241	●	■ Heat shock factor response element ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.
SR-MMP	0.02	●	■ Mitochondrial membrane potential ■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.

SR-p53	0.012	●	■ Category 1: actives ; Category 0: inactives; ■ The output value is the probability of being active.

10. Toxicophore Rules

Property	Value	Comment
Acute Toxicity Rule	0 alerts	■ 20 substructures ■ acute toxicity during oral administration
Genotoxic Carcinogenicity Rule	1 alerts	■ 117 substructures ■ carcinogenicity or mutagenicity
NonGenotoxic Carcinogenicity Rule	0 alerts	■ 23 substructures ■ carcinogenicity through nongenotoxic mechanisms
Skin Sensitization Rule	3 alerts	■ 155 substructures ■ skin irritation
Aquatic Toxicity Rule	1 alerts	■ 99 substructures ■ toxicity to liquid(water)
NonBiodegradable Rule	0 alerts	■ 19 substructures ■ non-biodegradable
SureChEMBL Rule	0 alerts	■ 164 substructures ■ MedChem unfriendly status

Figure S1. The antibacterial activity of four THC analogues and thiodiazole-copper against Xoo.