In order to avoid security-related warning messages when switching to secured connection, you may want either to:

- confirm the exception on the next page, or
- import our CA key in your web browser

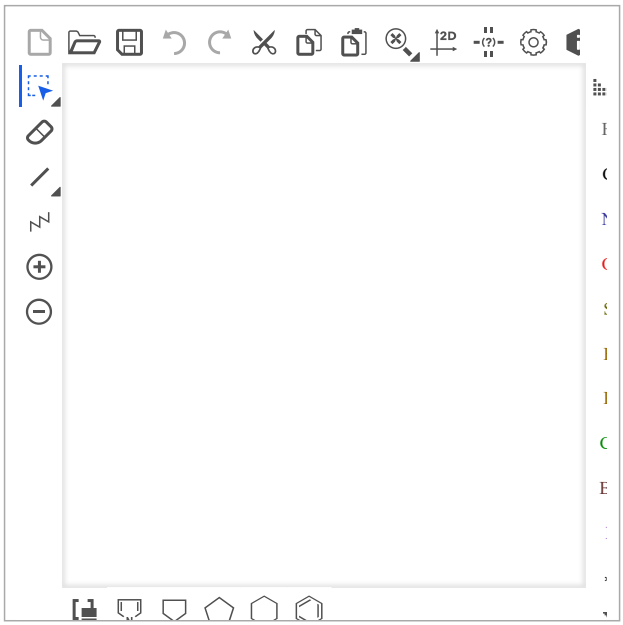
Click here to proceed.

SwissADME

This website allows you to compute physicochemical descriptors as well as to predict ADME parameters, pharmacokinetic properties, druglike nature and medicinal chemistry friendliness of one or multiple small molecules to support drug discovery.

The main article describing the web service and its underlying methodologies is SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* (2017) 7:42717.
For details about development and validation of iLOG, please refer to this article: iLOGP: a simple, robust, and efficient description of *n*-octanol/water partition coefficient for drug design using the GB/SA approach. *J. Chem. Inf. Model.* (2014) 54(12):3284-3301.
For details about development and validation of the BOILED-Egg, please refer to this article: A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules. *ChemMedChem* (2016) 11(11):1117-1121.

Developed and maintained by the Molecular Modeling Group of the SIB | Swiss Institute of Bioinformatics.



Enter a list of SMILES here:

```
CC[C@@H]1[C@@H](O)[C@@H](O[C@H]2[C@H](N)C[C@H](N)[C@@H](O[C@
CC[C@@H]1[C@@H](O)[C@@H](O[C@H]2[C@H](N)C[C@H](N)[C@H](O[C@@
CC[C@@H]1[C@@H](O)[C@@H](O[C@H]2[C@H](N)C[C@H](N)[C@@H](O[C@(
CC[C@@H]1[C@@H](O)[C@@H](O[C@H]2[C@H](N)C[C@H](N)[C@@H](O[C@(
CNC[C@@H]1CC[C@H](N)[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C(
CNC[C@@H]1CC[C@H](N)[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C(
CN[C@H](C)[C@H]1CC[C@@H](N)[C@@H](O[C@H]2[C@@H](N)C[C@@H](N)
CN[C@H](C)[C@@H]1CC[C@@H](N)[C@@H](O[C@H]2[C@@H](N)C[C@@H](N
CN[C@H](C)[C@@H]1CC[C@H](N)[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C(
CNC[C@@H]1CC[C@@H](N)[C@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@H]
CN[C@H](C)[C@@H]1CC[C@H](N)[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C(
CNC[C@@H]1CC[C@H](N)[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C(
CNC[C@@H]1CC[C@H](N)[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C(
CNC[C@H]1CC[C@@H](N)[C@@H](O[C@H]2[C@@H](N)C[C@@H](N)[C@@H](
CN[C@H](C)[C@@H]1CC[C@@H](N)[C@@H](O[C@H]2[C@@H](N)C[C@@H](
CO[C@H]1[C@H](O)[C@H](O)[C@H](O[C@H]2[C@H](O)[C@@H](CO)O[C@@
CNC[C@H]1CC[C@@H](N)[C@@H](O[C@H]2[C@@H](N)C[C@@H](N)[C@@H](
CO[C@H]1[C@H](O)[C@@H](O)[C@H](O[C@H]2[C@H](O)[C@@H](CO)O[C@(
CN[C@H](C)[C@H]1CC[C@@H](N)[C@@H](O[C@@H]2[C@@H](N)C[C@@H](N
CN[C@@H](C)[C@H]1CC[C@H](N)[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C(
```

Fill with an example | Clear | Run!

Show BOILED-Egg

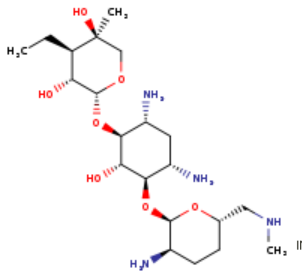
Retrieve data: CSV POWERED BY ChemAxon

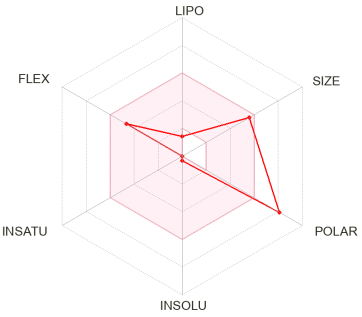
Molecule 1

Water Solubility

Log *S* (ESOL) -0.34

ESOL: Topological method implemented from





SMILES CNC[C@@H]1CC[C@H]([C@H](O1)O[C@@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@H]1OC[C@]([C@@H]([C@H]1O)CC)(C)O)N)N

Physicochemical Properties

Formula	C21H42N4O7
Molecular weight	462.58 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	11
Num. H-bond donors	7
Molar Refractivity	115.51
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	187.70 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.43
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-3.02
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-2.28
Log $P_{o/w}$ (MLOGP)	-2.16
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Delaney JS. 2004 J. Chem. Inf. Model.	
Solubility Class	2.10e+02 mg/ml ; 4.54e-01 mol/l
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble

Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	-0.36
Solubility Class	2.02e+02 mg/ml ; 4.37e-01 mol/l
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.18
Solubility Class	6.97e+02 mg/ml ; 1.51e+00 mol/l
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	Yes
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77	

Chem. Pharm. Bull.
Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.

Log $P_{o/w}$ (SILICOS-IT)

SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com | -2.63

Consensus Log $P_{o/w}$

Consensus Log $P_{o/w}$: Average of all five predictions | -1.53

External: ACC=0.88 / AUC=0.94

CYP1A2 inhibitor

Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91 | No

CYP2C19 inhibitor

Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87 | No

CYP2C9 inhibitor

Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81 | No

CYP2D6 inhibitor

Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87 | No

CYP3A4 inhibitor

Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86 | No

Log K_p (skin permeation)

Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res. | -11.27 cm/s

Druglikeness

Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW ≤ 500 MLOGP ≤ 4.15 N or O ≤ 10 NH or OH ≤ 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA ≤ 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

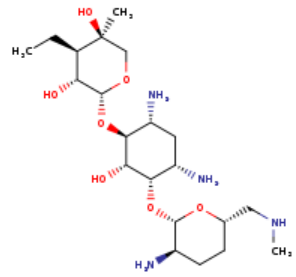
Leadlikeness

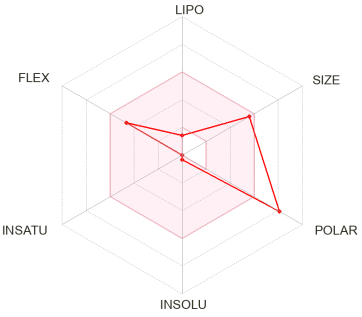
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.34

Molecule 2





SMILES: CNC[C@@H]1CC[C@H]([C@@H](O1)O[C@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@H]1OC[C@]([C@@H]([C@H]1O)CC)(C)O)N)N

Physicochemical Properties

Formula	C21H42N4O7
Molecular weight	462.58 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	11
Num. H-bond donors	7
Molar Refractivity	115.51
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	187.70 Å²

Lipophilicity

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	-0.34
Solubility	2.10e+02 mg/ml ; 4.54e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble

Log *S* (Ali)

Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	-0.36
Solubility	2.02e+02 mg/ml ; 4.37e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.19
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-3.02
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-2.28
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.16
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-2.63
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-1.58

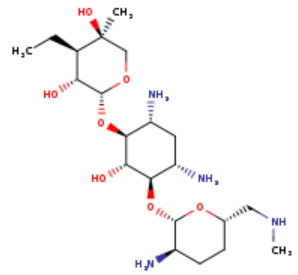
Log S (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.18
Solubility	6.97e+02 mg/ml ; 1.51e+00 mol/l
Class	
Solubility class: Log S scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble
Pharmacokinetics	
GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No

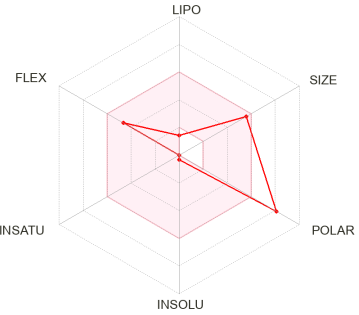
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-11.27 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW ≤ 500 MLOGP ≤ 4.15 N or O ≤ 10 NH or OH ≤ 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	No; 1 violation: TPSA>140
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem.	

Rotatable bonds ≤ 10 TPSA ≤ 140	
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA ≤ 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert
Brenk	
Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert
Leadlikeness	
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP ≤ 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350
Synthetic accessibility	6.34
Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties,	

trained on 12’782’590 molecules and tested on 40 external molecules (r^2 = 0.94)

Molecule 3





SMILES: CNC[C@@H]1CC[C@H]([C@@H](O1)O[C@@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@H]1OC[C@]([C@@H]([C@H]1O)CC)(C)O)N)N

Physicochemical Properties

Formula	C21H42N4O7
Molecular weight	462.58 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	11
Num. H-bond donors	7
Molar Refractivity	115.51
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	187.70 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.15
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-3.02
Log $P_{o/w}$ (WLOGP)	-2.28
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	-0.34
Solubility	2.10e+02 mg/ml ; 4.54e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	-0.36
Solubility	2.02e+02 mg/ml ; 4.37e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.18
Solubility	6.97e+02 mg/ml ; 1.51e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low

Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.16
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-2.63
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-1.59

BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	No
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set)	

and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-11.27 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

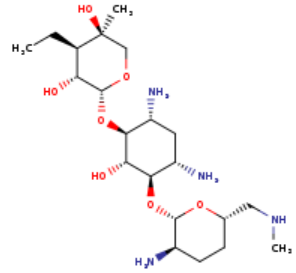
Leadlikeness

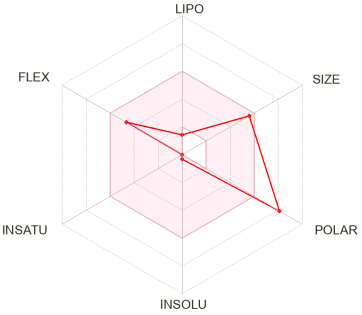
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.34

Molecule 4





SMILES: CNC[C@@H]1CC[C@H]([C@@H](O1)O[C@@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@H]1OC[C@]([C@@H]([C@H]1O)CC)(C)O)N)N

Physicochemical Properties

Formula: C21H42N4O7

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	-0.34
Solubility	2.10e+02 mg/ml ; 4.54e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble

Molecular weight	462.58 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	11
Num. H-bond donors	7
Molar Refractivity	115.51
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	187.70 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.15
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-3.02
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-2.28
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.16
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-2.63
Consensus Log $P_{o/w}$	-1.59
Consensus Log $P_{o/w}$: Average of all five	

Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	-0.36
Solubility	2.02e+02 mg/ml ; 4.37e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.18
Solubility	6.97e+02 mg/ml ; 1.51e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	No
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90	

predictions	External: ACC=0.84 / AUC=0.91	
	CYP2C19 inhibitor	
	Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
	CYP2C9 inhibitor	
	Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
	CYP2D6 inhibitor	
	Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
	CYP3A4 inhibitor	
	Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
	Log K_p (skin permeation)	
	Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-11.27 cm/s
		Druglikeness
	Lipinski	
	Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5

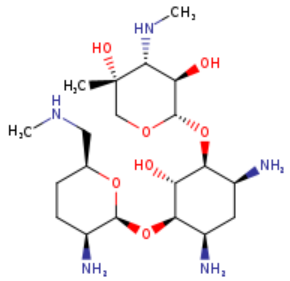
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA ≤ 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert
Brenk	
Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert
Leadlikeness	No; 1 violation: MW>350
Leadlikeness: implemented from	

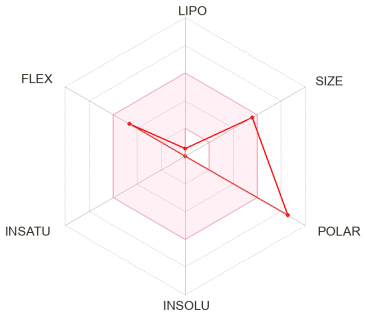
Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94) 6.34

Molecule 5





SMILES: CNC[C@@H]1CC[C@@H]([C@@H](O1)O[C@@H]1[C@H](N)C[C@@H]([C@@H]([C@H]1O)O[C@@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N

Physicochemical Properties

Formula	C20H41N5O7
Molecular weight	463.57 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	113.50
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)

iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.65

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.60
Solubility	1.85e+03 mg/ml ; 4.00e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (Ali)

Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.96
Solubility	4.18e+03 mg/ml ; 9.02e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (SILICOS-IT)

SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.53
Solubility	1.58e+03 mg/ml ; 3.40e+00 mol/l

Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.53
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.72
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-3.14
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.79
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.51

Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No

CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.34 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6

Muegge

Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 ≤ MW ≤ 600 -2 ≤ XLOGP ≤ 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

Leadlikeness

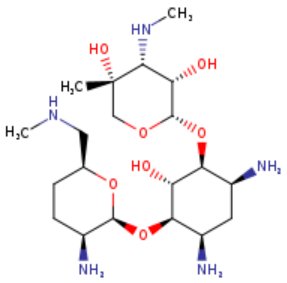
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350

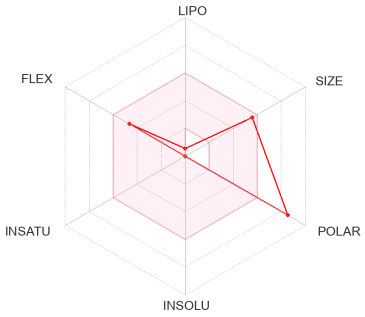
Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.33

Molecule 6

Water Solubility





SMILES: CNC[C@@H]1CC[C@@H]([C@@H](O1)O[C@@H]1[C@H](N)C[C@@H]([C@@H]([C@H]1O)O[C@@H]1OC[C@]([C@@H]([C@@H]1O)NC)(C)O)N)N

Physicochemical Properties

Formula	C20H41N5O7
Molecular weight	463.57 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	113.50
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.36
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.53
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.72
Log $P_{o/w}$ (MLOGP)	-3.14
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994	

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.60
Solubility	1.85e+03 mg/ml ; 4.00e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.96
Solubility	4.18e+03 mg/ml ; 9.02e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.53
Solubility	1.58e+03 mg/ml ; 3.40e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	Yes
P-glycoprotein substrate: SVM model built on 1033 molecules	

Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.79
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.56

(training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation)	-12.34 cm/s
Skin permeation: QSPR model	

implemented from Potts RO and Guy RH. 1992 Pharm. Res.	
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
	Medicinal Chemistry
PAINS	0 alert
Pan Assay Interference Structures:	

implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem | 0 alert

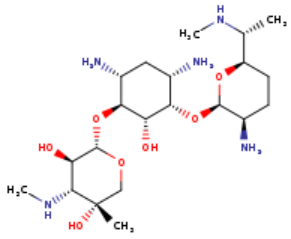
Leadlikeness

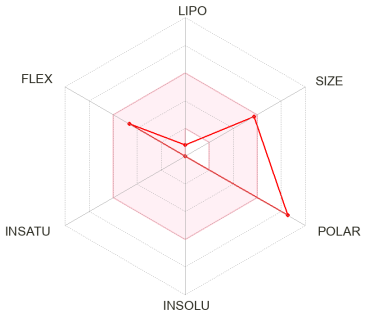
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7 | No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult)
based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94) | 6.33

Molecule 7





SMILES: CN[C@@H]([C@H]1CC[C@H]([C@H](O1)O[C@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N)C

Physicochemical Properties

Formula	C21H43N5O7
Molecular weight	477.60 g/mol
Num. heavy atoms	33
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	118.31
TPSA	199.73 Å²

Topological Polar Surface Area:

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model. | 0.24

Solubility | 8.37e+02 mg/ml ; 1.75e+00 mol/l

Class

Solubility class: Log *S* scale
Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly | Highly soluble

Log *S* (Ali)

Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model. | 0.51

Solubility | 1.54e+03 mg/ml ; 3.23e+00 mol/l

Class | Highly soluble

Solubility class: Log *S* scale
Insoluble < -10 < Poorly < -6 < Moderately < -4

Calculated from Ertl P. et al. 2000 J. Med. Chem.	

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.38
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.10
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.33
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.92
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.56
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.31

< Soluble < -2 Very < 0 < Highly	
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.52
Solubility	1.56e+03 mg/ml ; 3.27e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

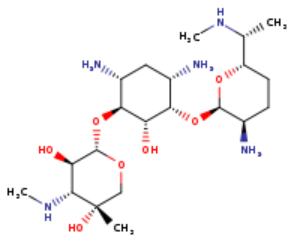
GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No

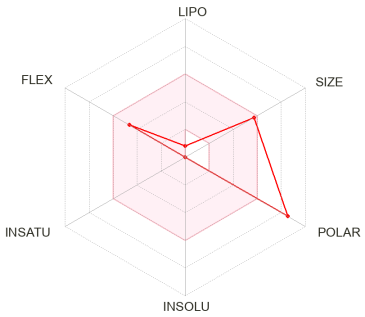
CYP2C9 inhibitor **Cytochrome P450 2C9 inhibitor:** SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor **Cytochrome P450 2D6 inhibitor:** SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor **Cytochrome P450 3A4 inhibitor:** SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation) **Skin permeation:** QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.12 cm/s
	Druglikeness
Lipinski **Lipinski (Pfizer) filter:** implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW ≤ 500 MLOGP ≤ 4.15 N or O ≤ 10 NH or OH ≤ 5	No; 2 violations: NorO>10, NHorOH>5
Ghose **Ghose filter:** implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber **Veber (GSK) filter:** implemented from Veber DF. et al. 2002 J. Med. Chem.	No; 1 violation: TPSA>140

Rotatable bonds ≤ 10 TPSA ≤ 140	
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA ≤ 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert
Brenk	
Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert
Leadlikeness	
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP ≤ 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350
Synthetic accessibility	6.51
Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties,	

trained on 12’782’590 molecules and tested on 40 external molecules (r^2 = 0.94)

Molecule 8





SMILES: CN[C@@H]([C@@H]1CC[C@H]([C@H](O1)O[C@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N)C

Physicochemical Properties

Formula	C21H43N5O7
Molecular weight	477.60 g/mol
Num. heavy atoms	33
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	118.31
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	3.19
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.10
Log $P_{o/w}$ (WLOGP)	-3.33
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.24
Solubility	8.37e+02 mg/ml ; 1.75e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.51
Solubility	1.54e+03 mg/ml ; 3.23e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.52
Solubility	1.56e+03 mg/ml ; 3.27e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low

Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.92
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.56
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.14

BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	No
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set)	

and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.12 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

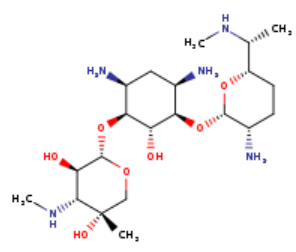
Leadlikeness

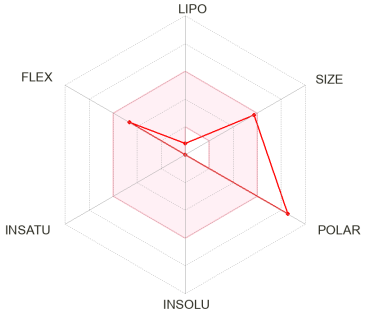
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.51

Molecule 9





SMILES	CN[C@@H]([C@@H]1CC[C@@H]([C@@H](O1)O[C@@H]1[C@H](N)C[C@@H]([C@@H]([C@H]1O)O[C@@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N)C

Physicochemical Properties

Formula	C21H43N5O7

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.24
Solubility	8.37e+02 mg/ml ; 1.75e+00 mol/l
Class	Highly soluble

Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly

Molecular weight	477.60 g/mol
Num. heavy atoms	33
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	118.31
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	3.26
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.10
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.33
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.92
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.56
Consensus Log $P_{o/w}$	-2.13
Consensus Log $P_{o/w}$: Average of all five	

Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.51
Solubility	1.54e+03 mg/ml ; 3.23e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.52
Solubility	1.56e+03 mg/ml ; 3.27e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	No
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90	

predictions	External: ACC=0.84 / AUC=0.91	
	CYP2C19 inhibitor	
	Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
	CYP2C9 inhibitor	
	Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
	CYP2D6 inhibitor	
	Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
	CYP3A4 inhibitor	
	Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
	Log K_p (skin permeation)	
	Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.12 cm/s
		Druglikeness
	Lipinski	
	Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5

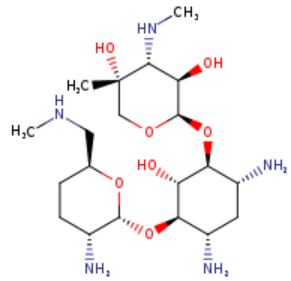
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA ≤ 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 ≤ MW ≤ 600 -2 ≤ XLOGP ≤ 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert
Brenk	
Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert
Leadlikeness	No; 1 violation: MW>350
Leadlikeness: implemented from	

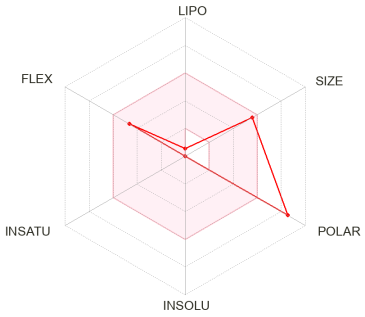
Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94) 6.51

Molecule 10





SMILES: CNC[C@@H]1CC[C@H]([C@H](O1)O[C@@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N

Physicochemical Properties

Formula	C20H41N5O7
Molecular weight	463.57 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	113.50
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)

iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.61

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.60
Solubility	1.85e+03 mg/ml ; 4.00e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (Ali)

Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.96
Solubility	4.18e+03 mg/ml ; 9.02e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (SILICOS-IT)

SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.53
Solubility	1.58e+03 mg/ml ; 3.40e+00 mol/l

Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.53
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.72
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-3.14
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.79
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.51

Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No

CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.34 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6

Muegge

Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 ≤ MW ≤ 600 -2 ≤ XLOGP ≤ 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

Leadlikeness

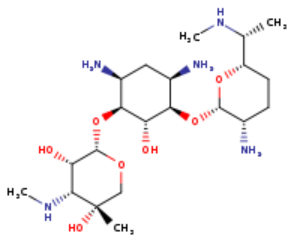
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350

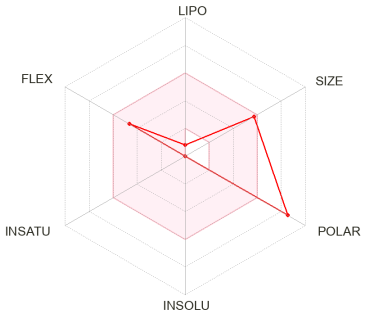
Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.33

Molecule 11

Water Solubility





SMILES CN[C@@H]([C@@H]1CC[C@@H]([C@@H](O1)O[C@@H]1[C@H](N)C[C@@H]([C@@H]([C@H]1O)O[C@@H]1OC[C@]([C@@H]([C@@H]1O)NC)(C)O)N)N)C

Physicochemical Properties

Formula	C21H43N5O7
Molecular weight	477.60 g/mol
Num. heavy atoms	33
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	118.31
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.68
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.10
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.33
Log $P_{o/w}$ (MLOGP)	-2.92
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994	

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.24
Solubility	8.37e+02 mg/ml ; 1.75e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.51
Solubility	1.54e+03 mg/ml ; 3.23e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.52
Solubility	1.56e+03 mg/ml ; 3.27e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	Yes
P-glycoprotein substrate: SVM model built on 1033 molecules	

Chem. Pharm. Bull.
Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.

Log $P_{o/w}$ (SILICOS-IT)

SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com

-3.56

Consensus Log $P_{o/w}$

Consensus Log $P_{o/w}$: Average of all five predictions

-2.25

(training set)
and tested on 415 molecules (test set)
10-fold CV: ACC=0.72 / AUC=0.77
External: ACC=0.88 / AUC=0.94

CYP1A2 inhibitor

Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set)
and tested on 3000 molecules (test set)
10-fold CV: ACC=0.83 / AUC=0.90
External: ACC=0.84 / AUC=0.91

No

CYP2C19 inhibitor

Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set)
and tested on 3000 molecules (test set)
10-fold CV: ACC=0.80 / AUC=0.86
External: ACC=0.80 / AUC=0.87

No

CYP2C9 inhibitor

Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set)
and tested on 2075 molecules (test set)
10-fold CV: ACC=0.78 / AUC=0.85
External: ACC=0.71 / AUC=0.81

No

CYP2D6 inhibitor

Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set)
and tested on 1068 molecules (test set)
10-fold CV: ACC=0.79 / AUC=0.85
External: ACC=0.81 / AUC=0.87

No

CYP3A4 inhibitor

Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set)
and tested on 2579 molecules (test set)
10-fold CV: ACC=0.77 / AUC=0.85
External: ACC=0.78 / AUC=0.86

No

Log K_p (skin permeation)

-12.12 cm/s

Skin permeation: QSPR model

implemented from Potts RO and Guy RH. 1992 Pharm. Res.	
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
	Medicinal Chemistry
PAINS	0 alert
Pan Assay Interference Structures:	

implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem — 0 alert

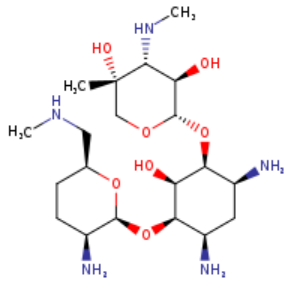
Leadlikeness

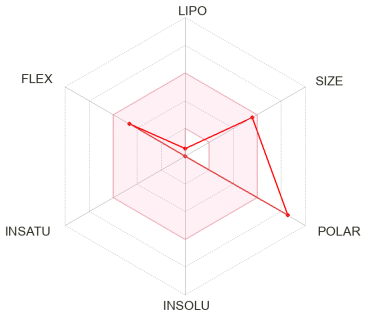
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7 — No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult)
based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94) — 6.51

Molecule 12





SMILES: CNC[C@@H]1CC[C@@H]([C@@H](O1)O[C@@H]1[C@H](N)C[C@@H]([C@@H]([C@@H]1O)O[C@@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N

Physicochemical Properties

Formula	C20H41N5O7
Molecular weight	463.57 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	113.50
TPSA	199.73 Å²

Topological Polar Surface Area:

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model. — 0.60

Solubility — 1.85e+03 mg/ml ; 4.00e+00 mol/l

Class

Solubility class: Log *S* scale
Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly — Highly soluble

Log *S* (Ali)

Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model. — 0.96

Solubility — 4.18e+03 mg/ml ; 9.02e+00 mol/l

Class — Highly soluble

Solubility class: Log *S* scale
Insoluble < -10 < Poorly < -6 < Moderately < -4

Calculated from Ertl P. et al. 2000 J. Med. Chem.	

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.01
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.53
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.72
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-3.14
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.79
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.63

< Soluble < -2 Very < 0 < Highly	
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.53
Solubility	1.58e+03 mg/ml ; 3.40e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

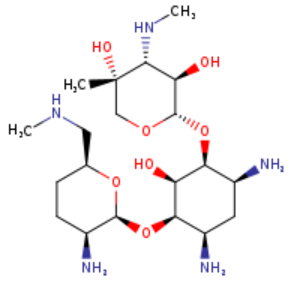
GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No

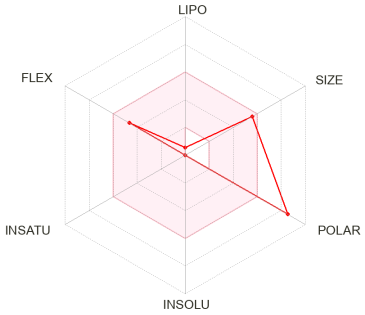
CYP2C9 inhibitor **Cytochrome P450 2C9 inhibitor:** SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor **Cytochrome P450 2D6 inhibitor:** SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor **Cytochrome P450 3A4 inhibitor:** SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation) **Skin permeation:** QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.34 cm/s
	Druglikeness
Lipinski **Lipinski (Pfizer) filter:** implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW ≤ 500 MLOGP ≤ 4.15 N or O ≤ 10 NH or OH ≤ 5	No; 2 violations: NorO>10, NHorOH>5
Ghose **Ghose filter:** implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber **Veber (GSK) filter:** implemented from Veber DF. et al. 2002 J. Med. Chem.	No; 1 violation: TPSA>140

Rotatable bonds ≤ 10 TPSA < 140	
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert
Brenk	
Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert
Leadlikeness	
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP ≤ 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350
Synthetic accessibility	6.33
Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties,	

trained on 12’782’590 molecules and tested on 40 external molecules (r^2 = 0.94)

Molecule 13





SMILES: CNC[C@@H]1CC[C@@H]([C@@H](O1)O[C@@H]1[C@H](N)C[C@@H]([C@@H]([C@@H]1O)O[C@@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N

Physicochemical Properties

Formula	C20H41N5O7
Molecular weight	463.57 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	113.50
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.01
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.53
Log $P_{o/w}$ (WLOGP)	-3.72
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.60
Solubility	1.85e+03 mg/ml ; 4.00e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.96
Solubility	4.18e+03 mg/ml ; 9.02e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.53
Solubility	1.58e+03 mg/ml ; 3.40e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low

Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-3.14
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.79
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.63

BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	No
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set)	

and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.34 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

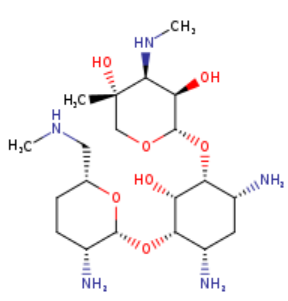
Leadlikeness

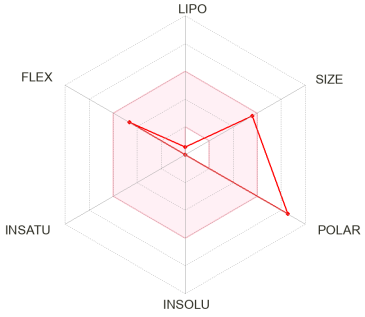
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.33

Molecule 14





SMILES	CNC[C@H]1CC[C@H]([C@H](O1)O[C@H]1[C@@H](N)C[C@H]([C@H]([C@H]1O)O[C@@H]1OC[C@]([C@H]([C@H]1O)NC)(C)O)N)N

Physicochemical Properties

Formula	C20H41N5O7

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.60
Solubility	1.85e+03 mg/ml ; 4.00e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Molecular weight	463.57 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	113.50
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.02
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.53
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.72
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-3.14
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.79
Consensus Log $P_{o/w}$	-2.63
Consensus Log $P_{o/w}$: Average of all five	

Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.96
Solubility	4.18e+03 mg/ml ; 9.02e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.53
Solubility	1.58e+03 mg/ml ; 3.40e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	No
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90	

predictions	External: ACC=0.84 / AUC=0.91	
	CYP2C19 inhibitor	
	Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
	CYP2C9 inhibitor	
	Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
	CYP2D6 inhibitor	
	Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
	CYP3A4 inhibitor	
	Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
	Log K_p (skin permeation)	
	Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.34 cm/s
		Druglikeness
	Lipinski	
	Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5

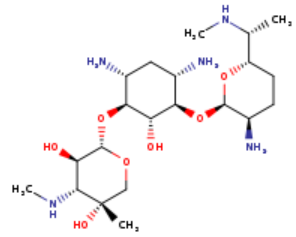
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert
Brenk	
Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert
Leadlikeness	No; 1 violation: MW>350
Leadlikeness: implemented from	

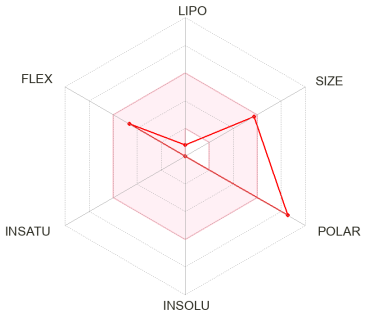
Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94) 6.33

Molecule 15





SMILES: CN[C@@H]([C@@H]1CC[C@H]([C@H](O1)O[C@@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N)C

Physicochemical Properties

Formula	C21H43N5O7
Molecular weight	477.60 g/mol
Num. heavy atoms	33
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	118.31
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.82

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.24
Solubility	8.37e+02 mg/ml ; 1.75e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.51
Solubility	1.54e+03 mg/ml ; 3.23e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.52
Solubility	1.56e+03 mg/ml ; 3.27e+00 mol/l

Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.10
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.33
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.92
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.56
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.22

Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No

CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.12 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6

Muegge

Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 ≤ MW ≤ 600 -2 ≤ XLOGP ≤ 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

Leadlikeness

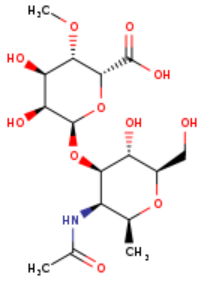
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350

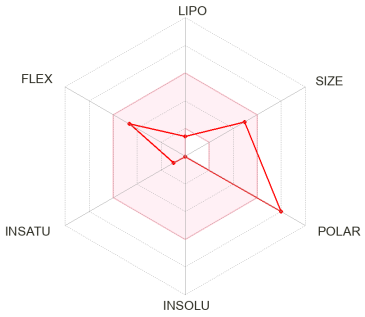
Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.51

Molecule 16

Water Solubility





SMILES OC[C@H]1O[C@@H](C)[C@H]([C@H]([C@@H]1O)O[C@@H]1O[C@@H](C(=O)O)[C@H]([C@@H]([C@@H]1O)O)OC)NC(=O)C

Physicochemical Properties

Formula	C16H27NO11
Molecular weight	409.39 g/mol
Num. heavy atoms	28
Num. arom. heavy atoms	0
Fraction Csp3	0.88
Num. rotatable bonds	7
Num. H-bond acceptors	11
Num. H-bond donors	6
Molar Refractivity	88.56
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	184.24 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	-0.12
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-3.00
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.44
Log $P_{o/w}$ (MLOGP)	-3.52
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull.	

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	-0.03
Solubility	3.85e+02 mg/ml ; 9.41e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	-0.31
Solubility	2.02e+02 mg/ml ; 4.93e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	2.13
Solubility	5.54e+04 mg/ml ; 1.35e+02 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	Yes
P-glycoprotein substrate: SVM model built on 1033 molecules	

Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.

Log $P_{o/w}$ (SILICOS-IT)

SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com -3.18

Consensus Log $P_{o/w}$

Consensus Log $P_{o/w}$: Average of all five predictions -2.65

(training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94

CYP1A2 inhibitor

Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91 No

CYP2C19 inhibitor

Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87 No

CYP2C9 inhibitor

Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81 No

CYP2D6 inhibitor

Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87 No

CYP3A4 inhibitor

Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86 No

Log K_p (skin permeation) -10.93 cm/s

Skin permeation: QSPR model

implemented from Potts RO and Guy RH. 1992 Pharm. Res.	
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 1 violation: WLOGP<-0.4
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.11
	Medicinal Chemistry
PAINS	0 alert
Pan Assay Interference Structures:	

implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem — 0 alert

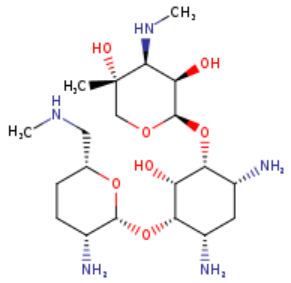
Leadlikeness

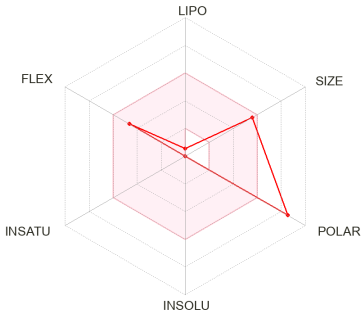
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7 — No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94) — 5.38

Molecule 17





SMILES: CNC[C@H]1CC[C@H]([C@H](O1)O[C@H]1[C@@H](N)C[C@H]([C@H]([C@H]1O)O[C@H]1OC[C@]([C@H]([C@H]1O)NC)(C)O)N)N

Physicochemical Properties

Formula	C20H41N5O7
Molecular weight	463.57 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	113.50
TPSA	199.73 Å²

Topological Polar Surface Area:

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model. — 0.60

Solubility — 1.85e+03 mg/ml ; 4.00e+00 mol/l

Class

Solubility class: Log *S* scale
Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly — Highly soluble

Log *S* (Ali)

Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model. — 0.96

Solubility — 4.18e+03 mg/ml ; 9.02e+00 mol/l

Class — Highly soluble

Solubility class: Log *S* scale
Insoluble < -10 < Poorly < -6 < Moderately < -4

Calculated from Ertl P. et al. 2000 J. Med. Chem.	

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.22
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.53
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.72
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-3.14
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.79
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.59

< Soluble < -2 Very < 0 < Highly	
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.53
Solubility	1.58e+03 mg/ml ; 3.40e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

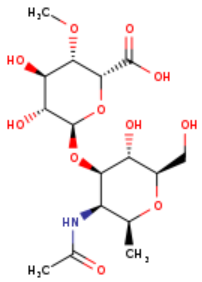
GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No

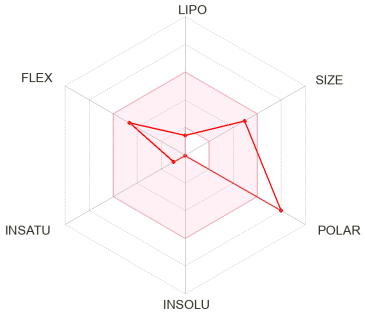
CYP2C9 inhibitor **Cytochrome P450 2C9 inhibitor:** SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor **Cytochrome P450 2D6 inhibitor:** SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor **Cytochrome P450 3A4 inhibitor:** SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation) **Skin permeation:** QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.34 cm/s
	Druglikeness
Lipinski **Lipinski (Pfizer) filter:** implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW ≤ 500 MLOGP ≤ 4.15 N or O ≤ 10 NH or OH ≤ 5	No; 2 violations: NorO>10, NHorOH>5
Ghose **Ghose filter:** implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber **Veber (GSK) filter:** implemented from Veber DF. et al. 2002 J. Med. Chem.	No; 1 violation: TPSA>140

Rotatable bonds ≤ 10 TPSA < 140	
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP ≤ 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5
Bioavailability Score	
Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17
Medicinal Chemistry	
PAINS	
Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert
Brenk	
Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert
Leadlikeness	
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP ≤ 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350
Synthetic accessibility	6.33
Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties,	

trained on 12’782’590 molecules and tested on 40 external molecules (r^2 = 0.94)

Molecule 18





SMILES OC[C@H]1O[C@@H](C)[C@H]([C@H]([C@@H]1O)O[C@@H]1O[C@@H](C(=O)O)[C@H]([C@@H]([C@H]1O)O)OC)NC(=O)C

Physicochemical Properties

Formula	C16H27NO11
Molecular weight	409.39 g/mol
Num. heavy atoms	28
Num. arom. heavy atoms	0
Fraction Csp3	0.88
Num. rotatable bonds	7
Num. H-bond acceptors	11
Num. H-bond donors	6
Molar Refractivity	88.56
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	184.24 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	0.71
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-3.00
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.44

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	-0.03
Solubility	3.85e+02 mg/ml ; 9.41e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	-0.31
Solubility	2.02e+02 mg/ml ; 4.93e-01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Very soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	2.13
Solubility	5.54e+04 mg/ml ; 1.35e+02 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low

Log $P_{o/w}$ (MLOGP)

MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev. | -3.52

Log $P_{o/w}$ (SILICOS-IT)

SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com | -3.18

Consensus Log $P_{o/w}$

Consensus Log $P_{o/w}$: Average of all five predictions | -2.49

BBB permeant

BBB permeation: according to the yolk of the BOILED-Egg | No

P-gp substrate

P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94 | Yes

CYP1A2 inhibitor

Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91 | No

CYP2C19 inhibitor

Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87 | No

CYP2C9 inhibitor

Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81 | No

CYP2D6 inhibitor

Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87 | No

CYP3A4 inhibitor | No

Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set)

and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-10.93 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 1 violation: WLOGP<-0.4
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6
Muegge	
Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA < 150 Num. rings < 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds < 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.11

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

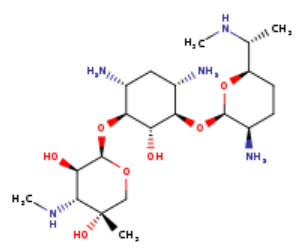
Leadlikeness

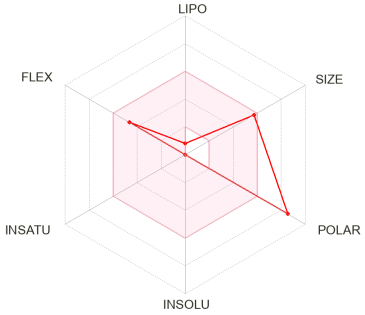
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	5.38

Molecule 19





SMILES	CN[C@@H]([C@H]1CC[C@H]([C@H](O1)O[C@@H]1[C@@H](N)C[C@H]([C@@H]([C@H]1O)O[C@H]1OC[C@]([C@@H]([C@H]1O)NC)(C)O)N)N)C

Physicochemical Properties

Formula	C21H43N5O7

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.24
Solubility	8.37e+02 mg/ml ; 1.75e+00 mol/l

Class

Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Molecular weight	477.60 g/mol
Num. heavy atoms	33
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	118.31
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.05
Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.10
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.33
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.92
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.56
Consensus Log $P_{o/w}$	-2.37
Consensus Log $P_{o/w}$: Average of all five	

Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.51
Solubility	1.54e+03 mg/ml ; 3.23e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.52
Solubility	1.56e+03 mg/ml ; 3.27e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	No
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90	

predictions	External: ACC=0.84 / AUC=0.91	
	CYP2C19 inhibitor	
	Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
	CYP2C9 inhibitor	
	Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
	CYP2D6 inhibitor	
	Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
	CYP3A4 inhibitor	
	Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
	Log K_p (skin permeation)	
	Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.12 cm/s
		Druglikeness
	Lipinski	
	Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5

Ghose

Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 2 violations: WLOGP<-0.4, #atoms>70

Veber

Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140

Egan

Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6

Muegge

Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 < MW < 600 -2 < XLOGP < 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. < 10 H-bond don. < 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

Leadlikeness — No; 1 violation: MW>350

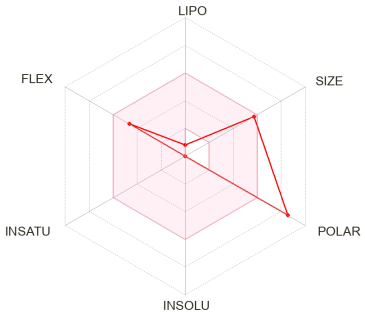
Leadlikeness: implemented from

Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94) 6.51

Molecule 20



SMILES: CN[C@H]([C@H]1CC[C@@H]([C@@H](O1)O[C@@H]1[C@H](N)C[C@@H]([C@@H]([C@@H]1O)O[C@H]1OC[C@]([C@H]([C@H]1O)NC)(C)O)N)N)C

Physicochemical Properties

Formula	C21H43N5O7
Molecular weight	477.60 g/mol
Num. heavy atoms	33
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	7
Num. H-bond acceptors	12
Num. H-bond donors	8
Molar Refractivity	118.31
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	199.73 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.56

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	0.24
Solubility	8.37e+02 mg/ml ; 1.75e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (Ali)	
Ali: Topological method implemented from Ali J. et al. 2012 J. Chem. Inf. Model.	0.51
Solubility	1.54e+03 mg/ml ; 3.23e+00 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	0.52
Solubility	1.56e+03 mg/ml ; 3.27e+00 mol/l

Log $P_{o/w}$ (XLOGP3)	
XLOGP3: Atomistic and knowledge-based method calculated by XLOGP program, version 3.2.2, courtesy of CCBG, Shanghai Institute of Organic Chemistry	-4.10
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-3.33
Log $P_{o/w}$ (MLOGP)	
MLOGP: Topological method implemented from Moriguchi I. et al. 1992 Chem. Pharm. Bull. Moriguchi I. et al. 1994 Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	-2.92
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-3.56
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-2.27

Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	Yes
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / AUC=0.91	No
CYP2C19 inhibitor	
Cytochrome P450 2C19 inhibitor: SVM model built on 9272 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.80 / AUC=0.86 External: ACC=0.80 / AUC=0.87	No
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No

CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) and tested on 1068 molecules (test set) 10-fold CV: ACC=0.79 / AUC=0.85 External: ACC=0.81 / AUC=0.87	No
CYP3A4 inhibitor	
Cytochrome P450 3A4 inhibitor: SVM model built on 7518 molecules (training set) and tested on 2579 molecules (test set) 10-fold CV: ACC=0.77 / AUC=0.85 External: ACC=0.78 / AUC=0.86	No
Log K_p (skin permeation)	
Skin permeation: QSPR model implemented from Potts RO and Guy RH. 1992 Pharm. Res.	-12.12 cm/s
	Druglikeness
Lipinski	
Lipinski (Pfizer) filter: implemented from Lipinski CA. et al. 2001 Adv. Drug Deliv. Rev. MW < 500 MLOGP < 4.15 N or O < 10 NH or OH < 5	No; 2 violations: NorO>10, NHorOH>5
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 < MW < 480 -0.4 < WLOGP < 5.6 40 < MR < 130 20 < atoms < 70	No; 2 violations: WLOGP<-0.4, #atoms>70
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds < 10 TPSA < 140	No; 1 violation: TPSA>140
Egan	
Egan (Pharmacia) filter: implemented from Egan WJ. et al. 2000 J. Med. Chem. WLOGP < 5.88 TPSA < 131.6	No; 1 violation: TPSA>131.6

Muegge

Muegge (Bayer) filter: implemented from Muegge I. et al. 2001 J. Med. Chem. 200 ≤ MW ≤ 600 -2 ≤ XLOGP ≤ 5 TPSA ≤ 150 Num. rings ≤ 7 Num. carbon > 4 Num. heteroatoms > 1 Num. rotatable bonds ≤ 15 H-bond acc. ≤ 10 H-bond don. ≤ 5	No; 4 violations: XLOGP3<-2, TPSA>150, H-acc>10, H-don>5

Bioavailability Score

Abbott Bioavailability Score: Probability of F > 10% in rat implemented from Martin YC. 2005 J. Med. Chem.	0.17

Medicinal Chemistry

PAINS

Pan Assay Interference Structures: implemented from Baell JB. & Holloway GA. 2010 J. Med. Chem.	0 alert

Brenk

Structural Alert: implemented from Brenk R. et al. 2008 ChemMedChem	0 alert

Leadlikeness

Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penalties, trained on 12'782'590 molecules and tested on 40 external molecules ($r^2 = 0.94$)	6.51

Swiss Institute of Bioinformatics - © 2022