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.

SwissDrugDesign
SwissDock
SwissParam
SwissSidechain
SwissBioisostere
SwissTargetPrediction
SwissADME
SwissSimilarity
About us
Swiss Institute of Bioinformatics

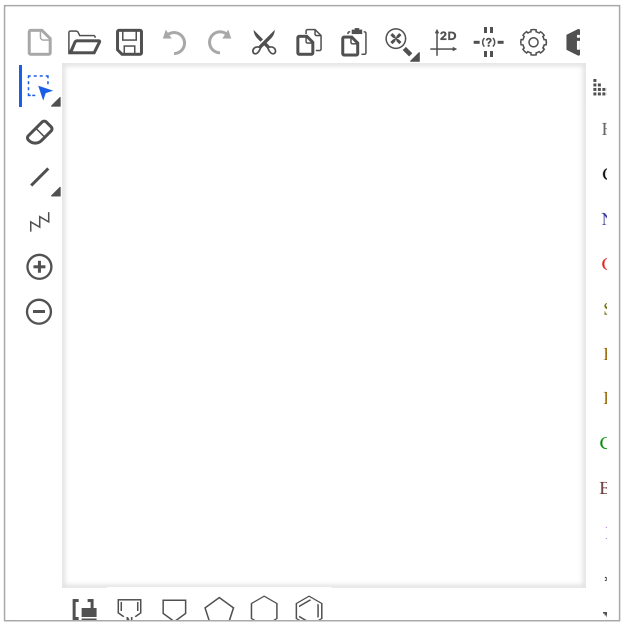
SwissADME

- Terms of Use
- Help
- FAQ
- Home

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:

```
NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@@H](N)[C@H](O[C@H]3O[C@
NC[C@@H]1CC[C@@H](N)[C@@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@H](
NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@H](O[C@H]3O[C@H
NC[C@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@H](O[C@H]3O[C@
NC[C@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@H](O[C@H]3O[C@
NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@@H](N)C(O[C@H]3O[C@H](CO
NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C@@H]3O[C@H
NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C@H]3O[C@H
NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C@@H]3O[C@H
NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C@@H]3O[C@H
NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@@H](O[C@@H]3O[C(
NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C@@H]3O[C@H
NC[C@@H]1O[C@H](O[C@H]2[C@H](N)C[C@H](N)[C@H](O[C@@H]3O[C@H]
NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@H](N)[C@H](O[C@H]3O[C@H
NC[C@@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@H](N)[C@H](O[C@H]3O[C@H
NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@@H](N)[C@H](O[C@H]3O[C@
NC[C@@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@@H](O[C@@H]3O[(
NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@H](O[C@H]3O[C@H
NC[C@@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@@H](O[C@@H]3O
NC[C@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@@H](N)[C@@H](O[C@@H]3O[(
```

Fill with an example | Clear | Run!

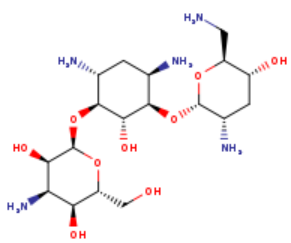
Show BOILED-Egg

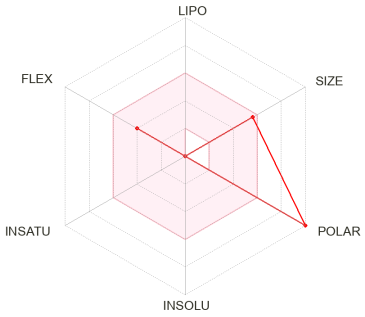
Retrieve data: CSV POWERED BY ChemAxon

Molecule 1

	Water Solubility
Log *S* (ESOL)	1.58

ESOL: Topological method implemented from





SMILES NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@H]([C@@H]([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@H]([C@H]2O)N)O)N)[C@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.57
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
Log $P_{o/w}$ (MLOGP)	-5.09
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	1.79e+04 mg/ml ; 3.82e+01 mol/l
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.	1.28
Solubility Class	8.95e+03 mg/ml ; 1.91e+01 mol/l
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	3.94
Solubility Class	4.11e+06 mg/ml ; 8.79e+03 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 | -6.58

Consensus Log $P_{o/w}$

Consensus Log $P_{o/w}$: Average of all five predictions | -4.53

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. | -13.58 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.	0.17
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

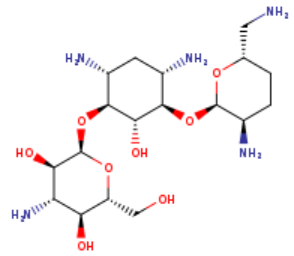
Leadlikeness

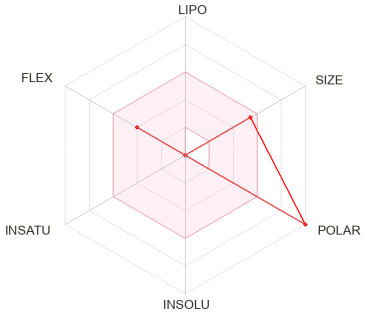
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

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)	6.42

Molecule 2





SMILES: NC[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]1O[C@H](CO)[C@H]([C@@H]([C@H]1O)N)O)N)N

Physicochemical Properties

Formula	C18H37N5O8
Molecular weight	451.52 g/mol
Num. heavy atoms	31
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	13
Num. H-bond donors	9
Molar Refractivity	104.82
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	247.94 Å²

Lipophilicity

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.41
Solubility	1.16e+04 mg/ml ; 2.57e+01 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.	1.26
Solubility	8.23e+03 mg/ml ; 1.82e+01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.71
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	-5.80
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-5.27
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.	-4.35
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	-5.70
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-3.88

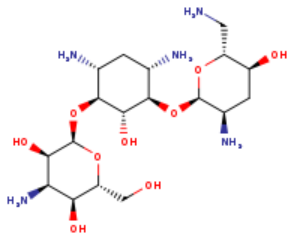
Log S (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	3.12
Solubility	5.98e+05 mg/ml ; 1.32e+03 mol/l
Class	
Solubility class: Log S scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 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

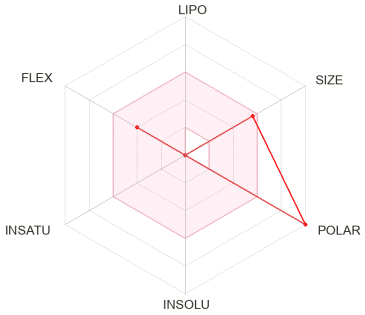
CYP2C9 inhibitor **Cytochrome P450 2C9 inhibitor:** SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor **Cytochrome P450 2D6 inhibitor:** SVM model built on 3664 molecules (training 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.	-13.17 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.	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	6.34
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 NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@H]([C@@H]([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@H]([C@H]2O)N)O)N)[C@@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.21
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.40

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.	-13.58 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.	0.17

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

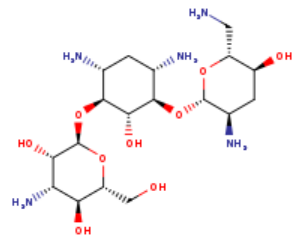
Leadlikeness

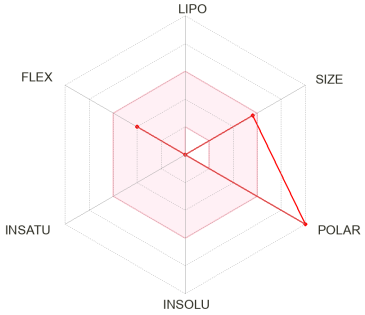
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.42

Molecule 4





SMILES	NC[C@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@H]([C@@H]([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@@H]([C@@H]2O)N)O)N)[C@@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.14
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	-4.61
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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-13.58 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

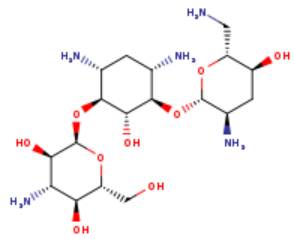
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 1 violation: WLOGP<-0.4
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140	No; 1 violation: TPSA>140
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.	0.17
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	

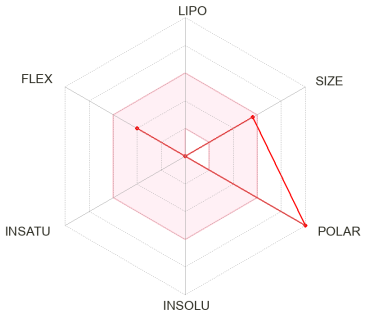
Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

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) 6.42

Molecule 5





SMILES NC[C@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@H]([C@@H]([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@@H]([C@H]2O)N)O)N)[C@@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.10

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.62

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.	-13.58 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.	0.17

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

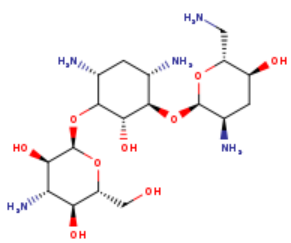
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350

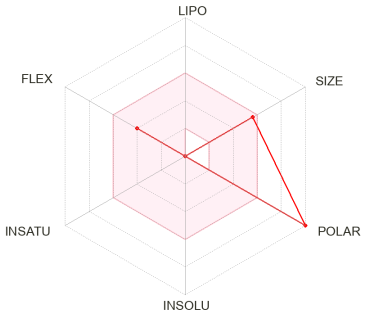
Synthetic accessibility

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)	6.42

Molecule 6

Water Solubility





SMILES NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@H](C([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@@H]([C@H]2O)N)O)N)[C@@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.46
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
Log $P_{o/w}$ (MLOGP)	-5.09
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.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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	

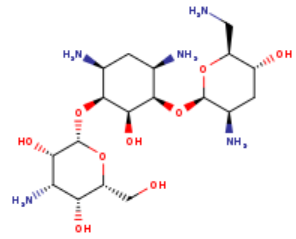
Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.55

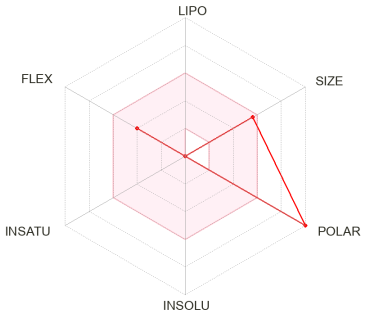
(training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / 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)	-13.58 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.	0.17
	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 penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.42

Molecule 7





SMILES	NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@@H]([C@@H]2O)O[C@@H]2O[C@H](CO)[C@@H]([C@@H]([C@@H]2O)N)O)N)[C@@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 mol/l
Class	Highly soluble
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4	

TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²
	Lipophilicity
Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	0.34
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.77

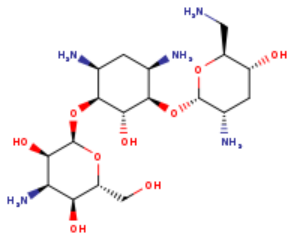
< Soluble < -2 Very < 0 < Highly	
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < 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

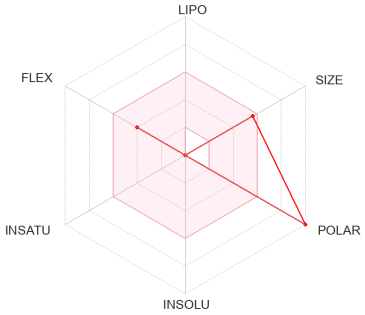
CYP2C9 inhibitor	
Cytochrome P450 2C9 inhibitor: SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor	
Cytochrome P450 2D6 inhibitor: SVM model built on 3664 molecules (training set) 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.	-13.58 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	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	6.42
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 8





SMILES NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@@H]([C@@H]([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@H]([C@H]2O)N)O)N)[C@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	0.75
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	-6.23
Log $P_{o/w}$ (WLOGP)	-6.30
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.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.69

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.	-13.58 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.	0.17

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

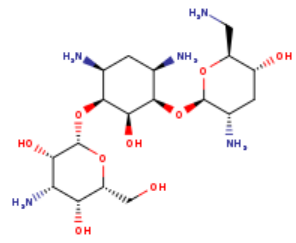
Leadlikeness

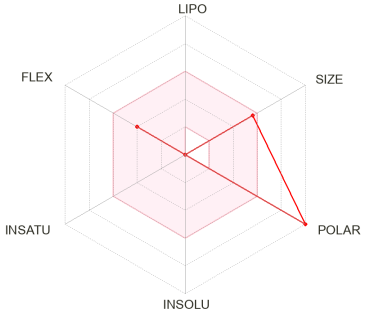
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

Synthetic accessibility

Synthetic accessibility score: from 1 (very easy) to 10 (very difficult) based on 1024 fragmental contributions (FP2) modulated by size and complexity penaties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.42

Molecule 9





SMILES	NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@@H]([C@@H]2O)O[C@@H]2O[C@H](CO)[C@@H]([C@@H]([C@@H]2O)N)O)N)[C@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.24
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	-4.59
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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-13.58 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

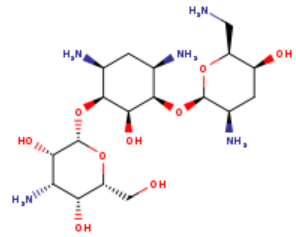
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 1 violation: WLOGP<-0.4
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140	No; 1 violation: TPSA>140
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.	0.17
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	

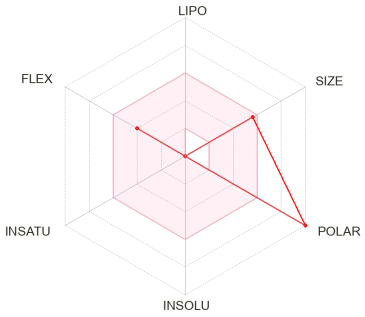
Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

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) 6.42

Molecule 10





SMILES NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@@H]([C@@H]2O)O[C@@H]2O[C@H](CO)[C@@H]([C@@H]([C@@H]2O)N)O)N)[C@@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.35

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.57

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.	-13.58 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.	0.17

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

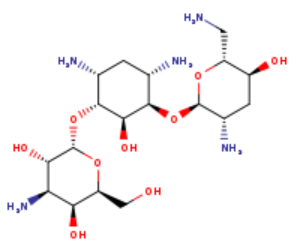
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350

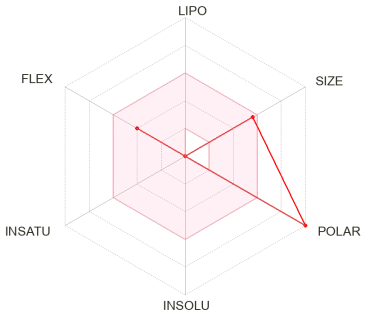
Synthetic accessibility

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)	6.42

Molecule 11

Water Solubility





SMILES NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@H]([C@H]([C@@H]2O)O[C@@H]2O[C@@H](CO)[C@H]([C@H]([C@@H]2O)N)O)N)[C@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.40
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
Log $P_{o/w}$ (MLOGP)	-5.09
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.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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

-6.58

Consensus Log $P_{o/w}$

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

-4.56

(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)

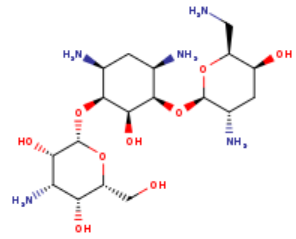
Skin permeation: QSPR model

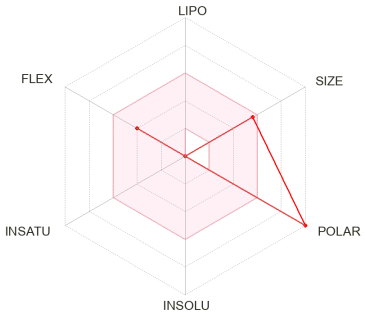
-13.58 cm/s

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.	0.17
	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)	6.42

Molecule 12





SMILES	NC[C@@H]1O[C@H](O[C@@H]2[C@H](N)C[C@@H]([C@@H]2O)O[C@@H]2O[C@H](CO)[C@@H]([C@@H]([C@@H]2O)N)O)N)[C@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 mol/l
Class	Highly soluble
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4	

TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²
	Lipophilicity
Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	-0.64
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.97

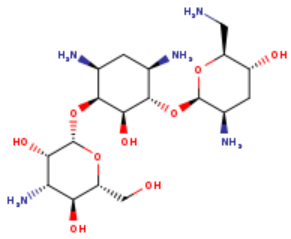
< Soluble < -2 Very < 0 < Highly	
Log *S* (SILICOS-IT)	
SILICOS-IT: Fragmental method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < 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

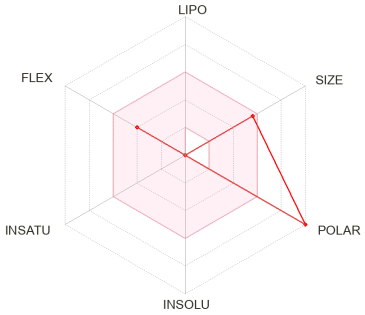
CYP2C9 inhibitor **Cytochrome P450 2C9 inhibitor:** SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor **Cytochrome P450 2D6 inhibitor:** SVM model built on 3664 molecules (training 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.	-13.58 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.	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	6.42
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 13





SMILES NC[C@@H]1O[C@H](O[C@H]2[C@H](N)C[C@@H]([C@@H]([C@@H]2O)O[C@@H]2O[C@H](CO)[C@H]([C@@H]([C@@H]2O)N)O)N)[C@@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	0.68
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	-6.23
Log $P_{o/w}$ (WLOGP)	-6.30
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.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.70

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.	-13.58 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.	0.17

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

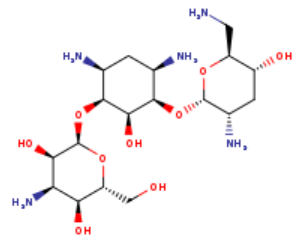
Leadlikeness

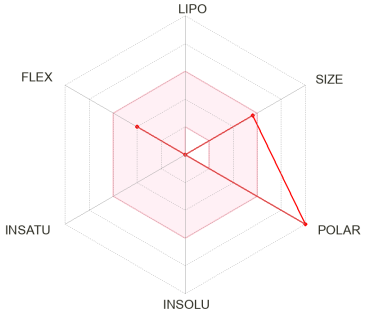
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

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)	6.42

Molecule 14





SMILES	NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@@H]([C@@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@H]([C@H]2O)N)O)N)[C@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.46
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	-4.55
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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-13.58 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

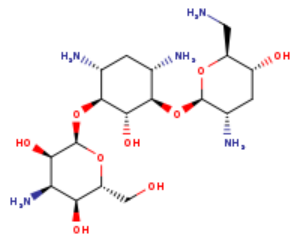
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 1 violation: WLOGP<-0.4
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140	No; 1 violation: TPSA>140
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.	0.17
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	

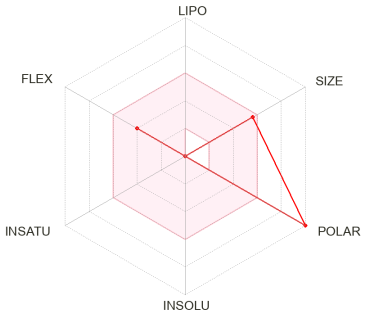
Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

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) 6.42

Molecule 15





SMILES NC[C@@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@H]([C@@H]([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@H]([C@H]2O)N)O)N)[C@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	0.97

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.65

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.	-13.58 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.	0.17

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

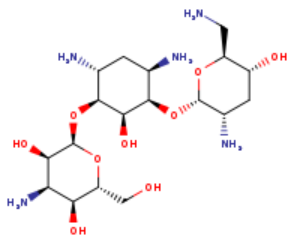
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds ≤ 7	No; 1 violation: MW>350

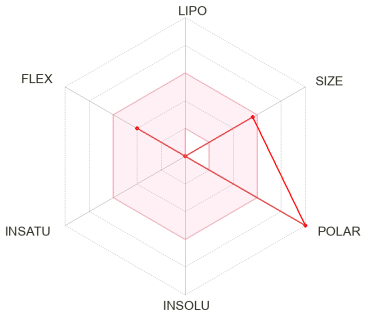
Synthetic accessibility

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)	6.42

Molecule 16

Water Solubility





SMILES NC[C@@H]1O[C@@H](O[C@@H]2[C@H](N)C[C@H]([C@@H]([C@@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@H]([C@H]2O)N)O)N)[C@H](C[C@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	2.22
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
Log $P_{o/w}$ (MLOGP)	-5.09
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.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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	

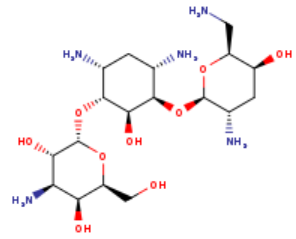
Chem. Pharm. Bull. Lipinski PA. et al. 2001 Adv. Drug. Deliv. Rev.	
Log $P_{o/w}$ (SILICOS-IT)	
SILICOS-IT: Hybrid fragmental/topological method calculated by FILTER-IT program, version 1.0.2, courtesy of SILICOS-IT, http://www.silicos-it.com	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.40

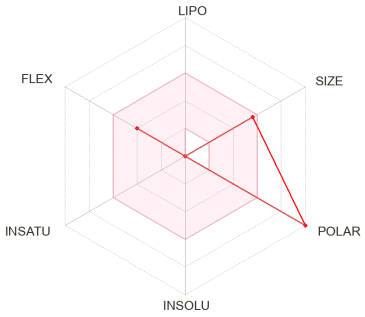
(training set) and tested on 415 molecules (test set) 10-fold CV: ACC=0.72 / AUC=0.77 External: ACC=0.88 / AUC=0.94	
CYP1A2 inhibitor	
Cytochrome P450 1A2 inhibitor: SVM model built on 9145 molecules (training set) and tested on 3000 molecules (test set) 10-fold CV: ACC=0.83 / AUC=0.90 External: ACC=0.84 / 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)	-13.58 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.	0.17
	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 penalties, trained on 12'782'590 molecules and tested on 40 external molecules (r^2 = 0.94)	6.42

Molecule 17





SMILES: NC[C@@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@H]([C@H]([C@@H]2O)O[C@@H]2O[C@@H](CO)[C@H]([C@H]([C@@H]2O)N)O)N)[C@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	268.17 Å²

Topological Polar Surface Area:

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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.	1.12
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.61

< 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Soluble

Pharmacokinetics

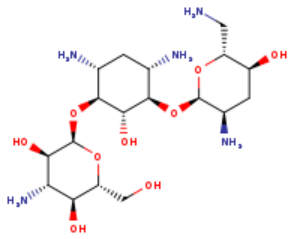
GI absorption	
Gatrointestinal absorption: according to the white of the BOILED-Egg	Low
BBB permeant	
BBB permeation: according to the yolk of the BOILED-Egg	No
P-gp substrate	
P-glycoprotein substrate: SVM model built on 1033 molecules (training set) and tested on 415 molecules (test set) 10-fold CV: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

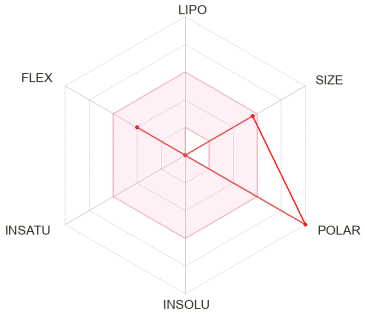
CYP2C9 inhibitor **Cytochrome P450 2C9 inhibitor:** SVM model built on 5940 molecules (training set) and tested on 2075 molecules (test set) 10-fold CV: ACC=0.78 / AUC=0.85 External: ACC=0.71 / AUC=0.81	No
CYP2D6 inhibitor **Cytochrome P450 2D6 inhibitor:** SVM model built on 3664 molecules (training 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.	-13.58 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.	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	6.42
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 18





SMILES NC[C@H]1O[C@H](O[C@@H]2[C@@H](N)C[C@H]([C@@H]([C@H]2O)O[C@H]2O[C@H](CO)[C@H]([C@@H]([C@H]2O)N)O)N)[C@@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.46
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	-6.23
Log $P_{o/w}$ (WLOGP)	-6.30
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.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.55

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.	-13.58 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.	0.17

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

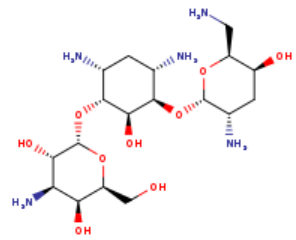
Leadlikeness

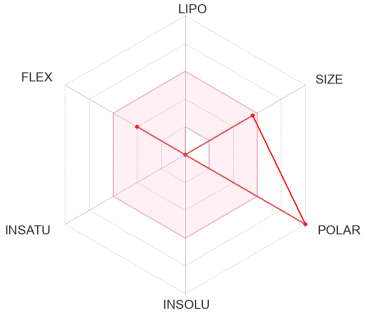
Leadlikeness: implemented from Teague SJ. 1999 Angew. Chem. Int. Ed. 250 < MW < 350 XLOGP < 3.5 Num. rotatable bonds < 7	No; 1 violation: MW>350

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)	6.42

Molecule 19





SMILES: NC[C@@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@H]([C@H]([C@@H]2O)O[C@@H]2O[C@@H](CO)[C@H]([C@H]([C@@H]2O)N)O)N)[C@H](C[C@@H]1O)N

Physicochemical Properties

Formula: C18H37N5O9

Water Solubility

Log *S* (ESOL)

ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 mol/l
Class	
Solubility class: Log *S* scale Insoluble < -10 < Poorly < -6 < Moderately < -4 < Soluble < -2 Very < 0 < Highly	Highly soluble

Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.00
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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	-4.64
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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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.	-13.58 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

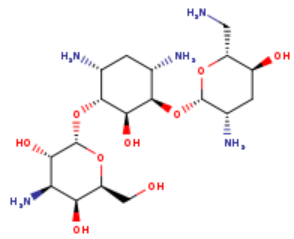
Ghose	
Ghose filter: implemented from Ghose AK. et al. 1999 J. Comb. Chem. 160 ≤ MW ≤ 480 -0.4 ≤ WLOGP ≤ 5.6 40 ≤ MR ≤ 130 20 ≤ atoms ≤ 70	No; 1 violation: WLOGP<-0.4
Veber	
Veber (GSK) filter: implemented from Veber DF. et al. 2002 J. Med. Chem. Rotatable bonds ≤ 10 TPSA ≤ 140	No; 1 violation: TPSA>140
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.	0.17
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	

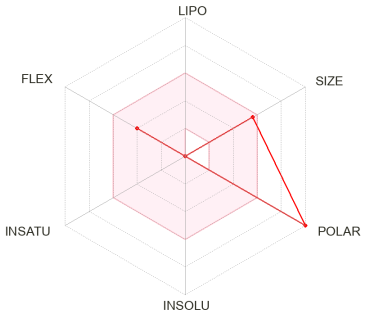
Teague SJ. 1999 Angew. Chem. Int. Ed.
250 < MW < 350
XLOGP < 3.5
Num. rotatable bonds < 7

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) 6.42

Molecule 20





SMILES NC[C@H]1O[C@@H](O[C@@H]2[C@@H](N)C[C@H]([C@H]([C@@H]2O)O[C@@H]2O[C@@H](CO)[C@H]([C@H]([C@@H]2O)N)O)N)[C@H](C[C@@H]1O)N

Physicochemical Properties

Formula	C18H37N5O9
Molecular weight	467.51 g/mol
Num. heavy atoms	32
Num. arom. heavy atoms	0
Fraction Csp3	1.00
Num. rotatable bonds	6
Num. H-bond acceptors	14
Num. H-bond donors	10
Molar Refractivity	105.98
TPSA	
Topological Polar Surface Area: Calculated from Ertl P. et al. 2000 J. Med. Chem.	268.17 Å²

Lipophilicity

Log $P_{o/w}$ (iLOGP)	
iLOGP: in-house physics-based method implemented from Daina A et al. 2014 J. Chem. Inf. Model.	1.43

Water Solubility

Log *S* (ESOL)	
ESOL: Topological method implemented from Delaney JS. 2004 J. Chem. Inf. Model.	1.58
Solubility	1.79e+04 mg/ml ; 3.82e+01 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.	1.28
Solubility	8.95e+03 mg/ml ; 1.91e+01 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	3.94
Solubility	4.11e+06 mg/ml ; 8.79e+03 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	-6.23
Log $P_{o/w}$ (WLOGP)	
WLOGP: Atomistic method implemented from Wildman SA and Crippen GM. 1999 J. Chem. Inf. Model.	-6.30
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.	-5.09
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	-6.58
Consensus Log $P_{o/w}$	
Consensus Log $P_{o/w}$: Average of all five predictions	-4.55

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.	-13.58 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.	0.17

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 ($r^2 = 0.94$)	6.42

Swiss Institute of Bioinformatics - © 2022