QSAR Prediction Reporting Format (QPRF)

The QPRF is a harmonised template for summarising and reporting substance-specific predictions generated by (Q)SAR models. The present QPRF is generated by LMC software, University "Prof. As. Zlatarov", Bourgas, Bulgaria

QSAR prediction for 1,1,3-trimethyl-3-phenylindan

1. SUBSTANCE

1.1 CAS number (entered by the author of prediction):

3910-35-8

1.2 EC number:

not reported

1.3 Chemical name:

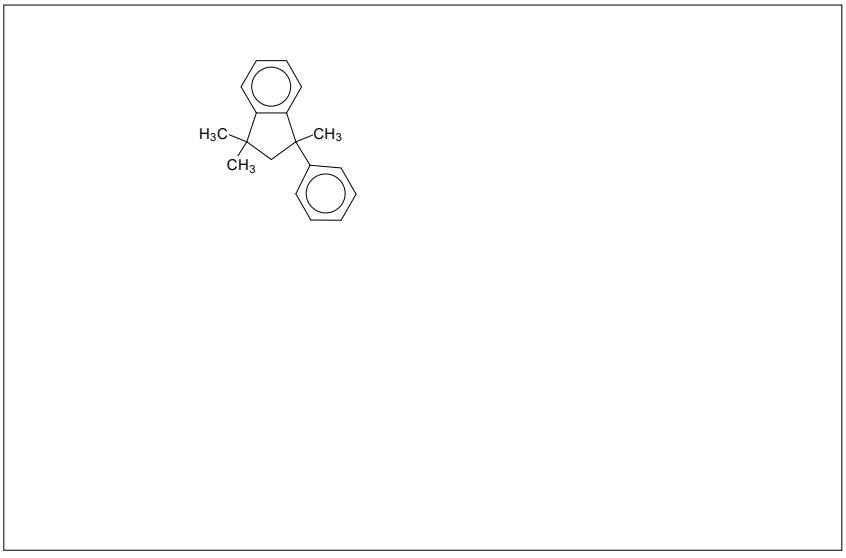
1,1,3-trimethyl-3-phenylindan

1.4 Structural formula:

a. Molecular formula:

C18H20

b. 2D structure:



1.5 Structure codes:

a. SMILES (used for prediction):

CC1(C)CC(C)(c2ccccc2)c2ccccc12

b. Other structural representation:

No other structural representations are used to generate the prediction

c. Stereochemical features:

Stereochemical features are not used to generate the prediction

2. GENERAL INFORMATION

2.1 Report date:

13.07.2023

2.2 Report authors:

Contact details are not provided by the author of prediction

2.3 Data sponsors:

Data sponsors are not provided by the author of prediction

3. PREDICTION

3.1 Endpoint (OECD Principle 1):

a. Endpoint:

Ready biodegradability, OECD 301C
Biological Oxygen Demand (BOD) according to Modified MITI Test (I) (TG 301 C) conditions (http://www.oecd.org/chemicalsafety/risk-assessment/1948209.pdf).
BOD is the oxygen used by aerobic microorganisms to mineralize the test substance after a given time (28 days), corrected for oxygen uptake by the blank inoculum control after the same time, and related to the theoretical oxygen demand needed for full mineralization of the substance.

Concomitant predictions:
Primary half-life
Ultimate half-life
Metabolites and their quantitative distribution

b. Dependent variable:

BOD, %

3.2 Algorithm (OECD Principle 2):

a. Model or submodel name:

CATALOGIC 301C

b. Model version:

v.12.17 - October 2021

c. Reference to QMRF:

available in OASIS Catalogic v.5.15.2.14

d. Predicted value (model result):

0.01 ± 5.07E-3

Concomitant predictions :
Not ready degradable
Primary Half Life = 3m 10d
Ultimate Half Life = 6y 5m 7d

e. Predicted value (comments):

No additional comments are provided by the author of prediction

f. Input for prediction:

SMILES

g. Descriptor values:

not reported

h. Observed Meta Data:

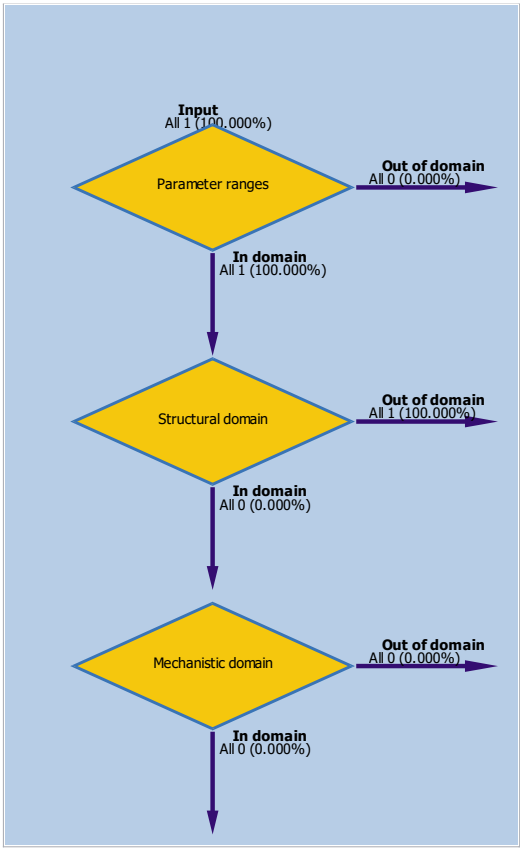
not reported

3.3 Applicability domain (OECD Principle 3):

a. Domain:

The LMC stepwise approach was used to define the applicability domain. It contains three layers:
• General properties requirements (log KOW, MW, WS)
• Structural domain (Atom Centered Fragments (ACFs))
• Mechanistic domain

Details could be found in:
• LMC QMRF CATALOGIC 301C v.12.17
• Dimitrov S, Dimitrova G., Pavlov T., Dimitrova D., Patlevisz G., Niemela J., Mekenyan O., A stepwise approach for defining the applicability domain of SAR and QSAR models, J. Chem. Inf. Model., 45, 839-849 (2005)



i. Parameter domain:

Log(Kow):::
range = [-10.7 .. 29.2]
calculated: 5.91 (In domain)

MOL._WEIGHT
range = [6.93 .. 1460]Da
calculated: 236Da (In domain)

WaterSolubility_FR
range = [0 .. 1000000]mg/L
calculated: 0.25mg/L (In domain)

CONCLUSION:
The chemical fulfils the general properties requirements

ii. Structural fragment domain:

The following ACF are identified:
Fragments in correctly predicted training chemicals – 62.50%
Fragments in non-correctly predicted training chemicals – 0.00%
Fragments not present in the training chemicals – 37.50%

CONCLUSION:
The chemical is out of the interpolation structural space

iii. Mechanistic domain:

The simulator of metabolism successfully simulated the mineralization

CONCLUSION:
The chemical is in the mechanistic domain.

b. Structural analogues in the training set:

not reported

3.4 Appropriate measures of goodness-of–fit, robustness and predictivity (OECD principle 4):

External Validation:
CATALOGIC 301C model was subject of external validation by different institutions and companies. Most of these validations were based on biodegradation results found in miscellaneous ready tests and practically are examination of the extrapolation potency of the model:
• BASF, 2004 – 488 chemicals tested in OECD 301A-F and ISO 14593 tests. External validation set is proprietary and is not available,
• Syracuse Research Corporation (SRC) & USA EPA – 336 chemicals tested in miscellaneous (not specified) ready tests. External validation set is not available,
• NITE Japan, 2005 – 1321 chemicals tested in OECD 301 C modified MITI I test. The validation set of 200 chemicals is available in NITE Report No. 1, CMC/QSAR/VR1002 (2005), the remaining part of the validation set is proprietary and is not available,
• RIVM, 2007 – 83 chemicals tested in OECD 301 B (79 chemicals) and 301C (4 chemicals) tests. External validation set is proprietary and is not available.

Statistics obtained by external validation:
• BASF, 2004:
- Overall accuracy – 72%,
- Specificity (ready biodegradable) = 50%,
- Sensitivity (not ready biodegradable) = 82%,
- Applicability domain was not accounted for.

• SRC & USA EPA, 2004:
- Overall accuracy – 81%,
- Specificity (ready biodegradable) = 57%,
- Sensitivity (not ready biodegradable) = 92%,
- Applicability domain was not accounted for.

• NITE Japan, 2005:
- Overall accuracy – 80%,
- Specificity (ready biodegradable) = 55%,
- Sensitivity (not ready biodegradable) = 88%,
- Applicability domain was analyzed and it was found better predictability for in the domain chemicals.

• RIVM, 2007:
- Overall accuracy – 86%,
- Specificity (ready biodegradable) = 65%,
- Sensitivity (not ready biodegradable) = 90%,
- Applicability domain was analyzed and domain chemicals were correctly predicted.

Statistics for goodness-of-fit:
• Residual Sum of Squares, RSS = 98.91
• Coefficient of correlation, R = 0.87
• Root mean square error, SR = 0.197

Correctness of classification:
• Sensitivity (not ready degradable): 93% for chemicals with BOD <= 50%
• Specificity (ready degradable): 77% for chemicals with BOD >= 70%
• Accuracy of predictions for chemicals within the border range (accounting for the tolerated test error) 50% < BOD < 70% is 31%.

3.5 The chemical and biological mechanisms according to the model underpinning the predicted result (OECD principle 5):

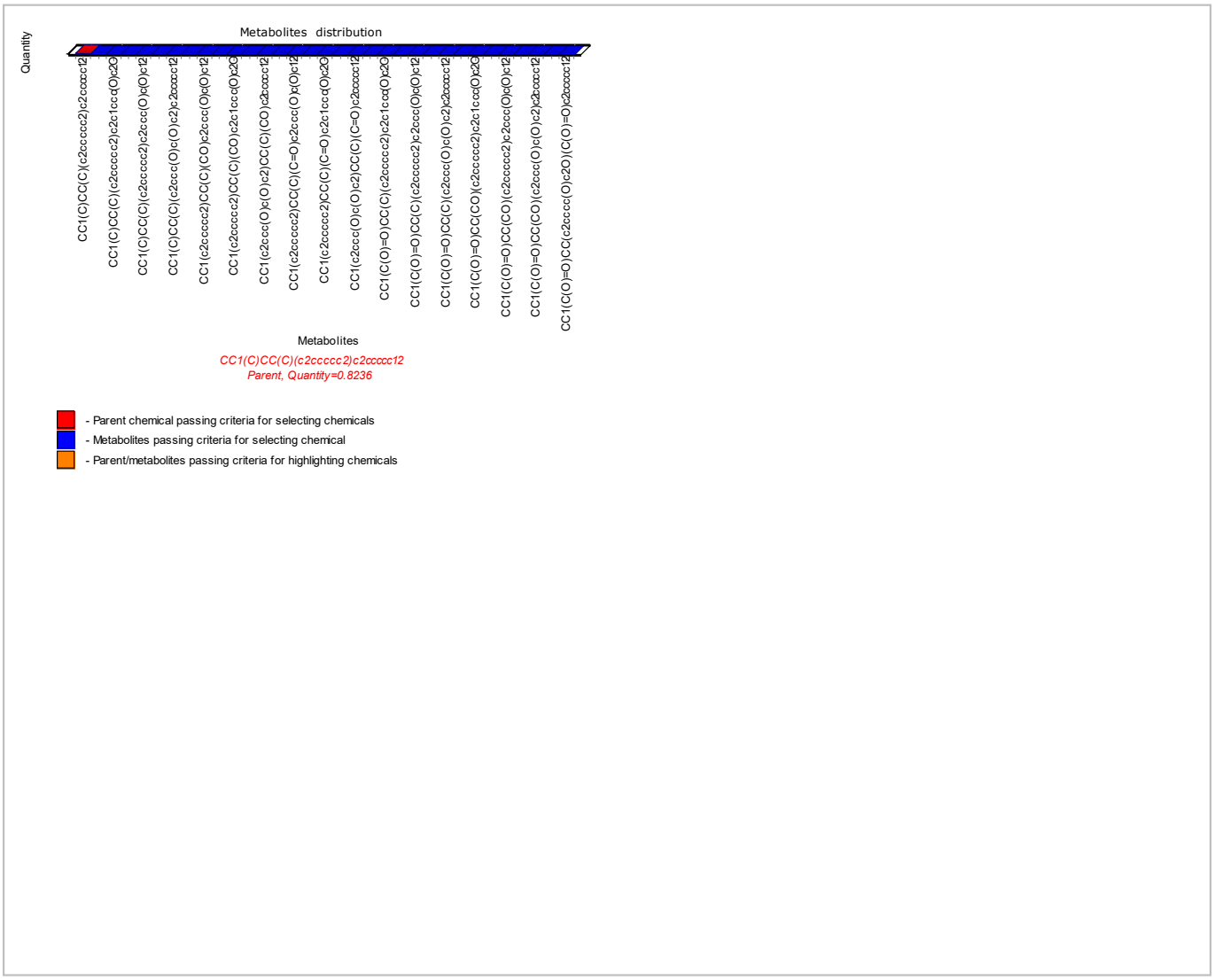
Detailed information about mechanistic basis of the model underpinning the prediction is available in model QMRF, section 8.

APPENDIX 1

Distribution of metabolites of the parent chemical

Criteria for selecting chemicals:
- Parents
- Take the parent no matter thresholds

Criteria for highlighting chemicals:
- no requirements



APPENDIX 2

Map of metabolic transformations

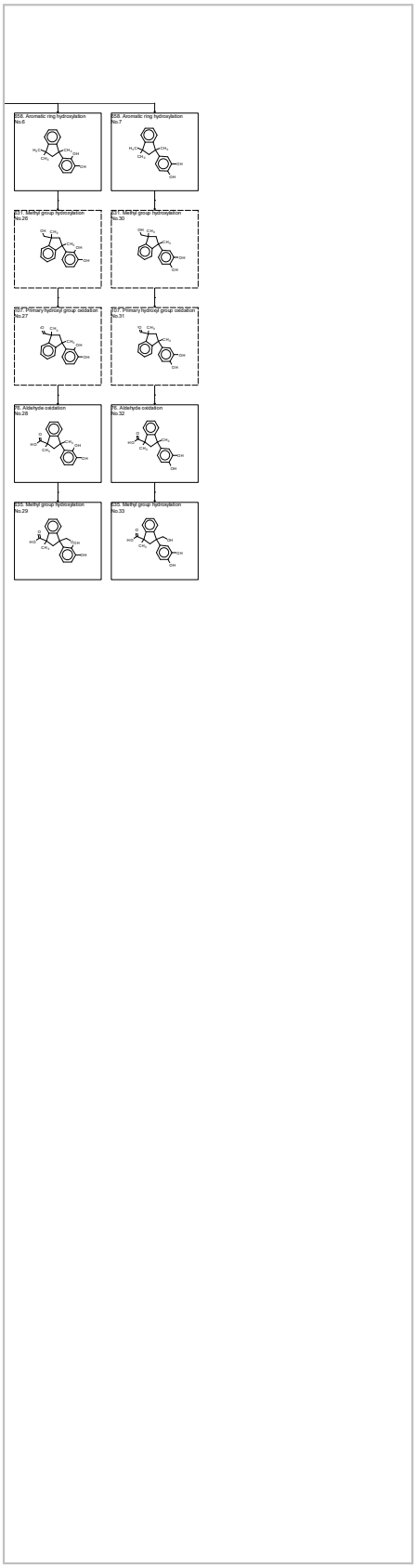
Criteria for highlighting chemicals:
· no requirements

Due to high number of generated metabolites the map depiction is restricted to levels <= 10

The map of metabolic transformations split into 2 parts is available below:

(1:1)

(1:2)


#28. Aromatic ring hydroxylation
No.6
#28. Aromatic ring hydroxylation
No.7
#1. Methyl group hydroxylation
No.26
#1. Methyl group hydroxylation
No.30
#2. Primary hydroxyl group oxidation
No.27
#2. Primary hydroxyl group oxidation
No.31
#3. Aldehyde oxidation
No.28
#3. Aldehyde oxidation
No.32
#35. Methyl group hydroxylation
No.29
#35. Methyl group hydroxylation
No.33

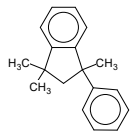
APPENDIX 3

List of metabolites of the parent chemical

Criteria for selecting chemicals:
· Parents
· Take the parent no matter thresholds

Parent No.1:

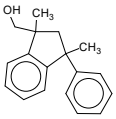
ID 1



#:	**2.1**
Cas#:	**3910-35-8**
Chem. Name:	**1,1,3-trimethyl-3-phenylindan**
Smiles:	**CC1(C)CC(C)(c2ccccc2)c2ccccc12**
Quantity [mol/mol parent]:	**0.8236**
log Kow:	**5.9100**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**0.01**

Metabolite No.2:

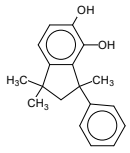
ID 2



#:	**2.2**
Smiles:	**CC1(c2ccccc2)CC(C)(CO)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**4.4400**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.3:

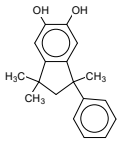
ID 3



#:	**2.3**
Smiles:	**CC1(C)CC(C)(c2ccccc2)c2c1ccc(O)c2O**
Quantity [mol/mol parent]:	**0.02741**
log Kow:	**4.9500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.4:

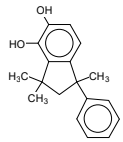
ID 4



#:	**2.4**
Smiles:	**CC1(C)CC(C)(c2ccccc2)c2cc(O)c(O)cc12**
Quantity [mol/mol parent]:	**0.02741**
log Kow:	**4.9500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.5:

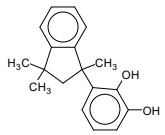
ID 5



#:	**2.5**
Smiles:	**CC1(C)CC(C)(c2ccccc2)c2ccc(O)c(O)c12**
Quantity [mol/mol parent]:	**0.02741**
log Kow:	**4.9500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.6:

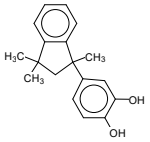
ID 6



#:	**2.6**
Smiles:	**CC1(C)CC(C)(c2cccc(O)c2O)c2ccccc12**
Quantity [mol/mol parent]:	**0.02741**
log Kow:	**4.9500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.7:

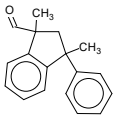
ID 7



#:	**2.7**
Smiles:	**CC1(C)CC(C)(c2ccc(O)c(O)c2)c2ccccc12**
Quantity [mol/mol parent]:	**0.02741**
log Kow:	**4.9500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.8:

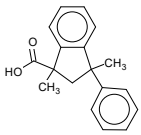
ID 8



#:	**2.8**
Smiles:	**CC1(c2ccccc2)CC(C)(C=O)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**4.4200**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.9:

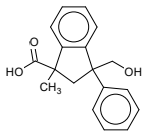
ID 9



#:	**2.9**
Smiles:	**CC1(C(O)=O)CC(C)(c2ccccc2)c2ccccc12**
Quantity [mol/mol parent]:	**0.03163**
log Kow:	**4.3100**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.10:

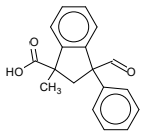
ID 10



#:	**2.10**
Smiles:	**CC1(C(O)=O)CC(CO)(c2ccccc2)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**2.8400**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.11:

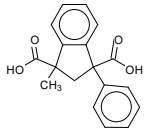
ID 11



#:	**2.11**
Smiles:	**CC1(C(O)=O)CC(C=O)(c2ccccc2)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.3300**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.12:

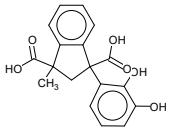
ID 12



#:	**2.12**
Smiles:	**CC1(C(O)=O)CC(c2ccccc2)(C(O)=O)c2ccccc12**
Quantity [mol/mol parent]:	**0.002395**
log Kow:	**2.6300**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.13:

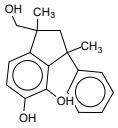
ID 13



#:	**2.13**
Smiles:	**CC1(C(O)=O)CC(c2cccc(O)c2O)(C(O)=O)c2ccccc12**
Quantity [mol/mol parent]:	**7.613E-5**
log Kow:	**1.6700**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.14:

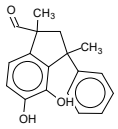
ID 14



#:	**2.14**
Smiles:	**CC1(c2ccccc2)CC(C)(CO)c2ccc(O)c(O)c12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.15:

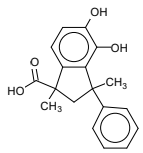
ID 15



#:	**2.15**
Smiles:	**CC1(c2ccccc2)CC(C)(C=O)c2ccc(O)c(O)c12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4600**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.16:

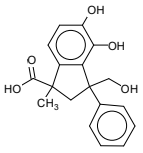
ID 16



#:	**2.16**
Smiles:	**CC1(C(O)=O)CC(C)(c2ccccc2)c2c1ccc(O)c2O**
Quantity [mol/mol parent]:	**0.001012**
log Kow:	**3.3500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.17:

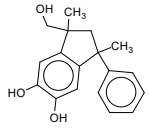
ID 17



#:	**2.17**
Smiles:	**CC1(C(O)=O)CC(CO)(c2ccccc2)c2c1ccc(O)c2O**
Quantity [mol/mol parent]:	**3.878E-5**
log Kow:	**1.8800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.18:

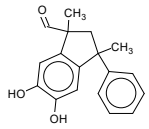
ID 18



#:	**2.18**
Smiles:	**CC1(c2ccccc2)CC(C)(CO)c2cc(O)c(O)cc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.19:

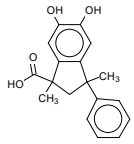
ID 19



#:	**2.19**
Smiles:	**CC1(c2ccccc2)CC(C)(C=O)c2cc(O)c(O)cc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4600**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.20:

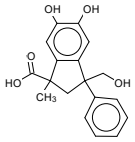
ID 20



#:	**2.20**
Smiles:	**CC1(C(O)=O)CC(C)(c2ccccc2)c2cc(O)c(O)cc12**
Quantity [mol/mol parent]:	**0.001012**
log Kow:	**3.3500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.21:

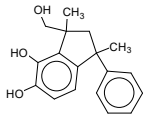
ID 21



#:	**2.21**
Smiles:	**CC1(C(O)=O)CC(CO)(c2ccccc2)c2cc(O)c(O)cc12**
Quantity [mol/mol parent]:	**3.878E-5**
log Kow:	**1.8800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.22:

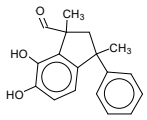
ID 22



#:	**2.22**
Smiles:	**CC1(c2ccccc2)CC(C)(CO)c2c1ccc(O)c2O**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.23:

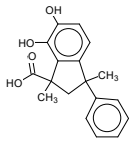
ID 23



#:	**2.23**
Smiles:	**CC1(c2ccccc2)CC(C)(C=O)c2c1ccc(O)c2O**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4600**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.24:

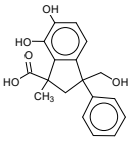
ID 24



#:	**2.24**
Smiles:	**CC1(C(O)=O)CC(C)(c2ccccc2)c2ccc(O)c(O)c12**
Quantity [mol/mol parent]:	**0.001012**
log Kow:	**3.3500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.25:

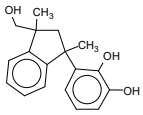
ID 25



#:	**2.25**
Smiles:	**CC1(C(O)=O)CC(CO)(c2ccccc2)c2ccc(O)c(O)c12**
Quantity [mol/mol parent]:	**3.878E-5**
log Kow:	**1.8800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.26:

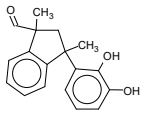
ID 26



#:	**2.26**
Smiles:	**CC1(c2cccc(O)c2O)CC(C)(CO)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.27:

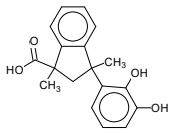
ID 27



#:	**2.27**
Smiles:	**CC1(c2cccc(O)c2O)CC(C)(C=O)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4600**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.28:

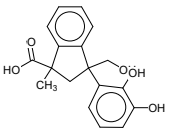
ID 28



#:	**2.28**
Smiles:	**CC1(C(O)=O)CC(C)(c2cccc(O)c2O)c2ccccc12**
Quantity [mol/mol parent]:	**0.001012**
log Kow:	**3.3500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.29:

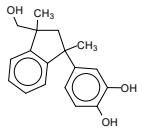
ID 29



#:	**2.29**
Smiles:	**CC1(C(O)=O)CC(CO)(c2cccc(O)c2O)c2ccccc12**
Quantity [mol/mol parent]:	**3.878E-5**
log Kow:	**1.8800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.30:

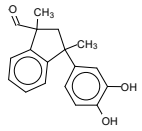
ID 30



#:	**2.30**
Smiles:	**CC1(c2ccc(O)c(O)c2)CC(C)(CO)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.31:

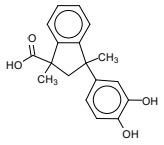
ID 31



#:	**2.31**
Smiles:	**CC1(c2ccc(O)c(O)c2)CC(C)(C=O)c2ccccc12**
Quantity [mol/mol parent]:	**0**
log Kow:	**3.4600**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.32:

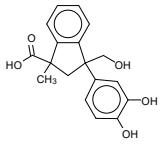
ID 32



#:	**2.32**
Smiles:	**CC1(C(O)=O)CC(C)(c2ccc(O)c(O)c2)c2ccccc12**
Quantity [mol/mol parent]:	**0.001012**
log Kow:	**3.3500**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

Metabolite No.33:

ID 33



#:	**2.33**
Smiles:	**CC1(C(O)=O)CC(CO)(c2ccc(O)c(O)c2)c2ccccc12**
Quantity [mol/mol parent]:	**3.878E-5**
log Kow:	**1.8800**
BOD_Observed [28.00 days] :	**No data**
BOD [28.00 days] :	**Disabled**

APPENDIX 4

Glossary of abbreviations

Abbreviation	Explanation
ACFs	Atom Centered Fragments
BOD	Biological Oxygen Demand, in % or decimal fraction
CO2	Carbon dioxide released for 28 days, in % or decimal fraction from theoretical amount of carbon dioxide
log Kow	Octanol/Water partition coefficient
MW	Molecular weight, Da
R	Coefficient of correlation
RSS	Residual Sum of Squares
SR	Root mean square error
ThCO2	Theoretical amount of carbon dioxide
WS	Water solubility, mg/l