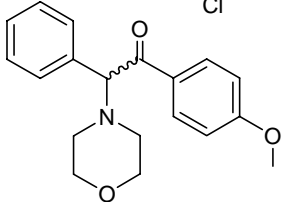


NCC SDF Specification

This document specifies the Structure-Data File (SDF) to accompany compounds shipped from BioFocus DPI to NCC customers.

SD File Field Name	Description	Data Type	Required Field	Example
[Structure Molfile]	Chemical structure of the compound in MDL Molfile format. Each molecular component of a mixture is represented once, regardless of the true stoichiometry of the mixture. For example, if a basic compound is a bis-HCl salt, then the HCl salt is represented only once. The correct proportional stoichiometry of a mixture is contained in the field NCC_STRUCTURE_REAL_MF.	Mol	Y	See example after table for Molfile corresponding to the structure: 
NCC_ISM	Daylight Isomeric SMILES representation of NCC_STRUCTURE.	Varchar	Y	<chem>O=C(C2=CC=C(OC)C=C2)C(C3=CC=CC=C3)N1CCOCC1.[Cl]</chem>
NCC_STRUCTURE_ID	Unique structure identifier assigned by MLSMR.	Varchar	Y	CPD000008708

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_STRUCTURE_SYNONYMS	Names or other identifiers for the compound. Multiple lines may be used.	Varchar	N	1-(4-Methoxy-phenyl)-2-morpholin-4-yl-2-phenyl-ethanone
NCC_SAMPLE_ID	Unique identifier assigned by MLSMR corresponding to a sample of NCC_STRUCTURE_ID.	Varchar	Y	SAM000077240
PUBCHEM_CID	Unique structure identifier assigned by PubChem.	Varchar	Y	9876543
PUBCHEM_SID	Substance identifier assigned by PubChem corresponding to NCC_SAMPLE_ID.	Varchar	Y	842125

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_STRUCTURE_REAL_MF	Molecular formula of the compound including correct proportional stoichiometry of a mixture in the format aMF ₁ .bMF ₂ .cMF ₃ , where a, b, and c are the stoichiometry factors for molecular components 1, 2, and 3, respectively, and MF ₁ , MF ₂ , and MF ₃ are the molecular formulae for molecular components 1, 2, and 3, respectively.	Varchar	Y	C19H21NO3.2HCl
NCC_STRUCTURE_REAL_AMW	Average molecular weight of the compound including correct proportional stoichiometry of a mixture.	Number	Y	347.8456
NCC_PARENT_AMW	Average molecular weight of parent compound free from salts and solvents.	Number	Y	311.3749
NCC_PARENT_EMW	Exact molecular weight of parent compound free from salts and solvents.	Number	Y	311.1521

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_SAMPLE_SUPPLIER	Source of compound sample.	Varchar	Y	TOCRIS
NCC_SUPPLIER_STRUCTURE_ID	Supplier's identifier corresponding to NCC_STRUCTURE.	Varchar	Y	BAS 00363530
NCC_SUPPLIER_SAMPLE_ID	Supplier's identifier corresponding to sample of NCC_STRUCTURE.	Varchar	N	12345
NCC_SUPPLIER_SAMPLE_COMMENTS	Additional information on sample provided by supplier.	Varchar	N	1:1 mixture of diastereoisomers
NCC_ANALYTICAL_PURITY_METHOD	Method used to determine the percent purity of the sample. Multiple lines will be included if multiple methods were used to determine purity.	Varchar	Y	LC-MS-UV214 LC-MS-ELS CofA

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_SAMPLE_PURITY	Percent purity of the sample as measured by NCC_ANALYTICAL_PURITY_METHOD. Multiple lines will be included if multiple methods were used to determine purity.	Number (one place decimal)	Y	93.7 100.0 80.5
NCC_PURITY_DATE	Date NCC_SAMPLE_PURITY was determined. Multiple lines will be included if multiple methods were used to determine purity.	Date	Y	01-APR-2005 02-APR-2005 25-MAY-2006
NCC_PLATE	Generic plate identifier containing aliquot of compound sample for NCC catalog number.	Varchar	Y	NCC-Plate 1
NCC_ALIQUOT_PLATE_BARCODE	Unique barcode of plate containing aliquot of compound sample.	Varchar	Y	NCC-123456
NCC_ALIQUOT_WELL_ID	Well location containing aliquot in plate.	Varchar	Y	A06

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_ALIQUOT_SOLVENT	Solvent used to dissolve the sample to create plated aliquot.	Varchar	Y	DMSO
NCC_ALIQUOT_CONC	Concentration of the plated aliquot.	Number	Y	10.0
NCC_ALIQUOT_CONC_UNIT	Unit of NCC_ALIQUOT_CONC. mM = millimolar.	Varchar	Y	mM
NCC_ALIQUOT_VOLUME	Volume of the plated aliquot.	Number	Y	50
NCC_ALIQUOT_VOLUME_UNIT	Unit of NCC_ALIQUOT_VOLUME. uL = microliter.	Varchar	Y	uL

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_ALIQUOT_DATE	Date the aliquot was plated.	Date	Y	01-DEC-2007
NCC_SAMPLE_HAZARD S	Known hazards of the sample. May contain multiple lines.	Varchar	N	Carcinogen Mutagen
NCC_CONTROLLED_SU BSTANCE	Classification of DEA controlled substances.	Varchar	N	DEA SCHEDULE 2
NCC_STABILITY	Conditions under which the compound is known to be unstable. May contain multiple lines.	Varchar	N	UV Air Heat
NCC_NUM_HDONORS	Count of O-H or N-H bonds (NH ₂ counts as two) in parent compound.	Number	Y	2

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_NUM_HACCEPTORS	Count of O or N atoms in parent compound.	Number	Y	6
NCC_NUM_ROTATABLE BONDS	Count of non-terminal single bonds that are not in a ring where the two atoms in the single bond are each bonded to a total of at least two non-hydrogen atoms.	Number	Y	10
NCC_SSSR	Smallest Set of Smallest Rings	Number	Y	2
NCC_TPSA	Topological Polar Surface Area (Angstroms squared) calculated in Daylight using Ertl method.	Number	Y	78.09
NCC ALOGP	Hydrophobicity LogP calculated from atomic contributions (octanol/ water partition coefficient) via ALogPS (Tetko) method.	Number	N	2.21

NCC SD File Specification

SD File Field Name	Description	Data Type	Required Field	Example
NCC_ALOGS	Hydrophobicity LogS calculated from atomic contributions (octanol/ water partition coefficient) via ALogPS (Tetko) method.	Number	N	-3.85
NCC_WATER_SOLUBILITY	Water solubility of parent compound in ug / mL calculated at pH 7.0 by ACD/Labs Solubility_Batch.	Number	N	11.0

Example entry for Structure Molfile:

```

24 25 0 0 0 0 0 0 0 0999 V2000
 1.5114  2.0052  0.0000 Cl 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.0023  0.8273  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.7150  0.4125  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.7150 -0.4125  0.0000 N  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0.7150  0.4148  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.0023  1.4873  0.0000 O  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.4323  0.8250  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.7081 -2.0625  0.0000 O  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0.7150 -0.4102  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 1.4300  0.8273  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 2.1427 -0.4102  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 1.4300 -0.8227  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 2.1427  0.4148  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.4277 -0.8273  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0.0023 -0.8227  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 2.8577 -0.8227  0.0000 O  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.4254 -1.6523  0.0000 C  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

NCC SD File Specification

```
0.0046 -1.6477 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-2.1427 0.4125 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.4323 1.6500 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2.8554 -1.4827 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-2.1427 2.0625 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-2.8577 0.8250 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-2.8577 1.6500 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3 2 1 4
4 3 1 0
5 2 1 0
6 2 2 0
7 3 1 0
8 18 1 0
9 5 1 0
10 5 2 0
11 12 1 0
12 9 2 0
13 10 1 0
14 4 1 0
15 4 1 0
16 11 1 0
17 14 1 0
18 15 1 0
19 7 2 0
20 7 1 0
21 16 1 0
22 20 2 0
23 19 1 0
24 22 1 0
13 11 2 0
8 17 1 0
24 23 2 0
M END
```