Small Molecules

Importance 1: Required, 2: Required if available, 3: Optional
Batch Information Only available from DSGC
Canonical Information Required (2 out of 4) if DSGC does not provide batch information

LINCS Field Name	Related to	Description	Comments	Importance	Batch Information	Canonical Information
SM_Name	canonical	The primary name for the (parent) compound (in a standardized representation) as chosen by LINCS	common, recognizable, name	1	-	YES
SM_LINCS_ID	canonical	The global LINCS ID (parent) compound (in a standardized representaiton)	LINCS internal ID; this is a batch independent ID	1	-	-
SM_Alternative_Name	canonical	List of synonymous compound names, drug name (if applicable), and other alternative names	drug names can come from common drug name available from DrugBank, TTD, NPC or other resources	2	-	-
SM_PubChem_CID	canonical	CID that corresponds to the standardized parent compound in NCBIs PubChem database	This is the PubChem compound ID, and not to be confused with the substance ID. CID will be checked against structure by the DCIC	1	-	YES
SM_ChEBI_ID	canonical	ChEBI ID that corresponds to the standardized parent compound in NCBIs PubChem database; after applying the same business rules	ChEBI ID; ChEBI is curated and we may submit important structure to ChEBI	2	-	YES
SM_InChI_Parent	canonical	InChi representation of parent (standardized) chemical structure	standardized small molecule representation; documented standardization protocol; standardization will be performed by DCIC; data will be cross checked by DCIC if possible	2	-	-
SM_InChI_Key_Parent	canonical	InChi key of parent (standardized) chemical structure	standardized small molecule representation; documented standardization protocol; standardization will be performed by DCIC; data will be cross checked by DCIC if possible	2	-	-
SM_SMILES_Parent	canonical	Canonical isomeric SMILES representation of parent (standardized) chemical structure	standardized small molecule representation; documented standardization protocol; standardization will be performed by DCIC; SMILES will be checked against other data (if availble) to verify business rules	1	-	YES
SM_Center_Name	batch	Name of the LINCS Center that is using the small molecule		1	YES	-
SM_Center_Compound_ID	canonical	Center-specific compound ID; typically based on parent structure after compound standardization	LINCS DSGC-specific canonical ID. This will be assigned by a given LINCS DSGC according to its compound registration scheme.	1	YES	-
SM_Center_Sample_ID	batch	Center-specific ID of the tested sample	LINCS DSGC-specific compound batch ID. This will be assigned by a given LINCS DSGC according to its compound registration scheme.	1	YES	-
SM_Provider	batch	Vendor or lab that supplied the compound	list of vendor names and other providers	1	YES	-
SM_Provider_Catalog_ID	batch	ID or catalogue number assigned to the specific supplied sample by the vendor or provider	-	1	YES	-
SM_Provider_Batch_ID	batch	Batch or lot number assigned to the specific supplied sample by vendor or provider	-	1	YES	-
SM_Salt	batch	Reference to counter-ions and other addends present in the compounds formulation	HMS LINCS has developed a three digit "salt code" used as part of the facility ID for its compounds	2	-	-
SM_SMILES_Batch	batch	SMILES representation only as provided from the vendor; full structure with all information, i.e. isomeric SMILES	this includes salts and addends and no structure canonicalization; full structure with all stereo and isomer information; i.e. isomeric SMILES	2	YES	-
SM_InChI_Batch	batch	InChi representation of the actual sample (batch) structure	this includes salts and addends and no structure canonicalization; full structure	2	-	-
SM_InChI_Key_Batch	batch	InChi key of the actual sample (batch) structure	-	2	-	-
SM_Molecular_Mass	batch	Molecular mass of one molecule (including addends) of the compound in Daltons (unified atomic mass units)	this includes salts and addends and no structure canonicalization; full structure	2	-	-
SM_Purity	batch	Purity of the compound in percent	suggested by MIABE; if available	3	YES	-
SM_Purity_Method	batch	Method for determining the purity of the compound	-	3	YES	-
SM_Aqueous_Solubility	batch	The actual (measured) aqueous solubility of the compound in mg / mL or g / L	suggested by MIABE; if available	3	-	-
SM_LogP	batch	The logarithm of the actual (measured) water/octanol partition coefficient (logP) or hydrophobicity score	suggested by MIABE; if available	3	-	-