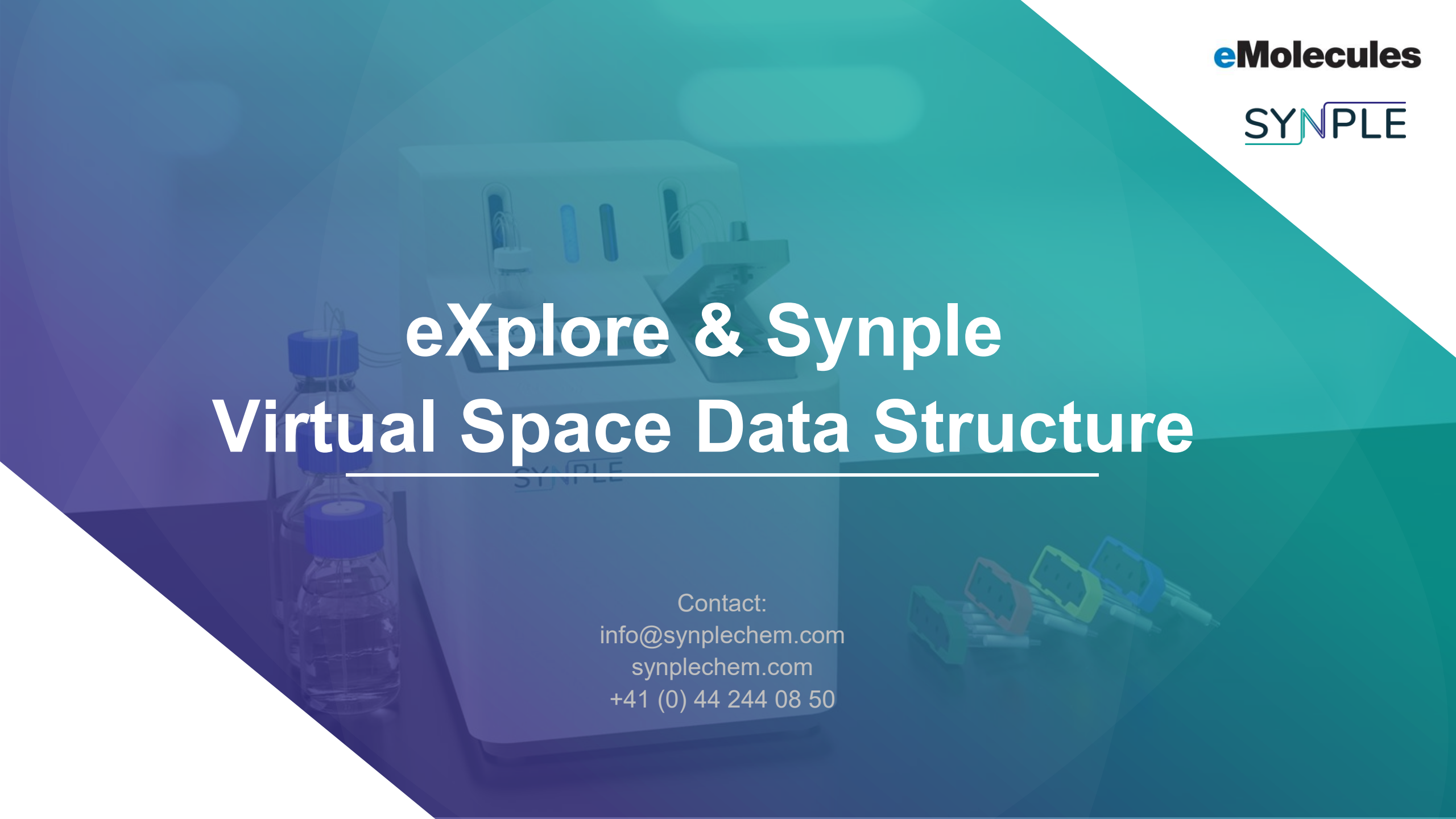
eMolecules

SYNPLE

eXplore & Synple
Virtual Space Data Structure



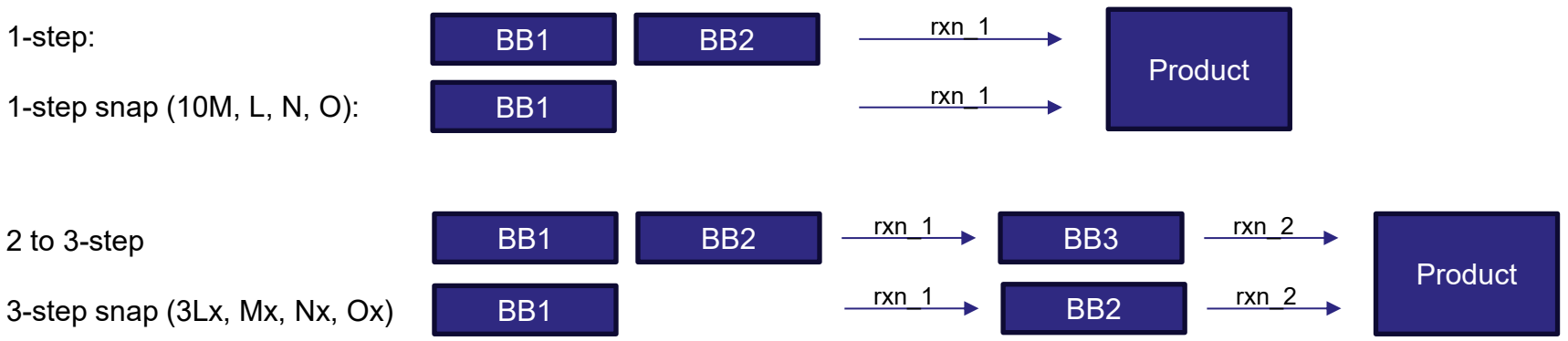
Contact:
info@synplechem.com
synplechem.com
+41 (0) 44 244 08 50

Overview

Several datasets are now available for the Synple and eXplore space and contain an array of metadata. This document aims to give an overview over the data structure of the library files.

Library sections:
Due to the library sizes, the data is split into chunks of ~15 GB (uncompressed) or ~5 (compressed). Additionally, each dataset is also downloadable as 10% random sample or 1% random sample. Other samples can be also provided on request

Reaction Enumeration:
Reaction were enumerated with the building block-logic below. Details on used reaction type and building block is entered for each compounds



On-demand synthesis price:
Each compound also show an indicative price. Upon order compounds price will be quoted based on current building block price and other requested services (e.g. additional analytics, reformatting, etc).

Custom data sets:
Filtered of modified datasets can be provided upon request. Please contact (benedikt.wanner@emolecules.com).

Data structure

Column	Data	Example
name	Unique molecule identifier	EMOL_rxn3AA_37173991_28304437_96308540
smiles	SMILES of enumerated library molecule	Cc1ncc(C)c(C(=O)N[C@H](CNC(=O)c2csc(-c3cccc(Br)c3)n2)Cc2ccccc2)n1
rxn_1	Reaction name of first step	abf
rxn_2	Reaction name of second step (empty in case of 1-step reactions)	ra
bb1_parent_id	eMolecules ID of first building block	37173991
bb1_smiles	SMILES of first building block	CC(C)(C)OC(=O)N[C@H](CN)Cc1ccccc1
bb2_parent_id	eMolecules ID of second building block	28304437
bb2_smiles	SMILES of second building block	O=C(O)c1csc(-c2cccc(Br)c2)n1
bb3_parent_id	eMolecules ID of third building block	96308540
bb3_smiles	SMILES of third building block	Cc1ncc(C)c(C(=O)O)n1
Molecule Properties: qed, MolWt, HeavyAtomMolWt, ExactMolWt, NumValenceElectrons, NumRadicalElectrons, TPSA, FractionCSP3, NumHAcceptors, NumHDonor, NumRotatableBonds, MolLogP	Various molecule properties of library compounds	
total_price	Indicator for price if compound is ordered from on-demand synthesis service (USD)	500