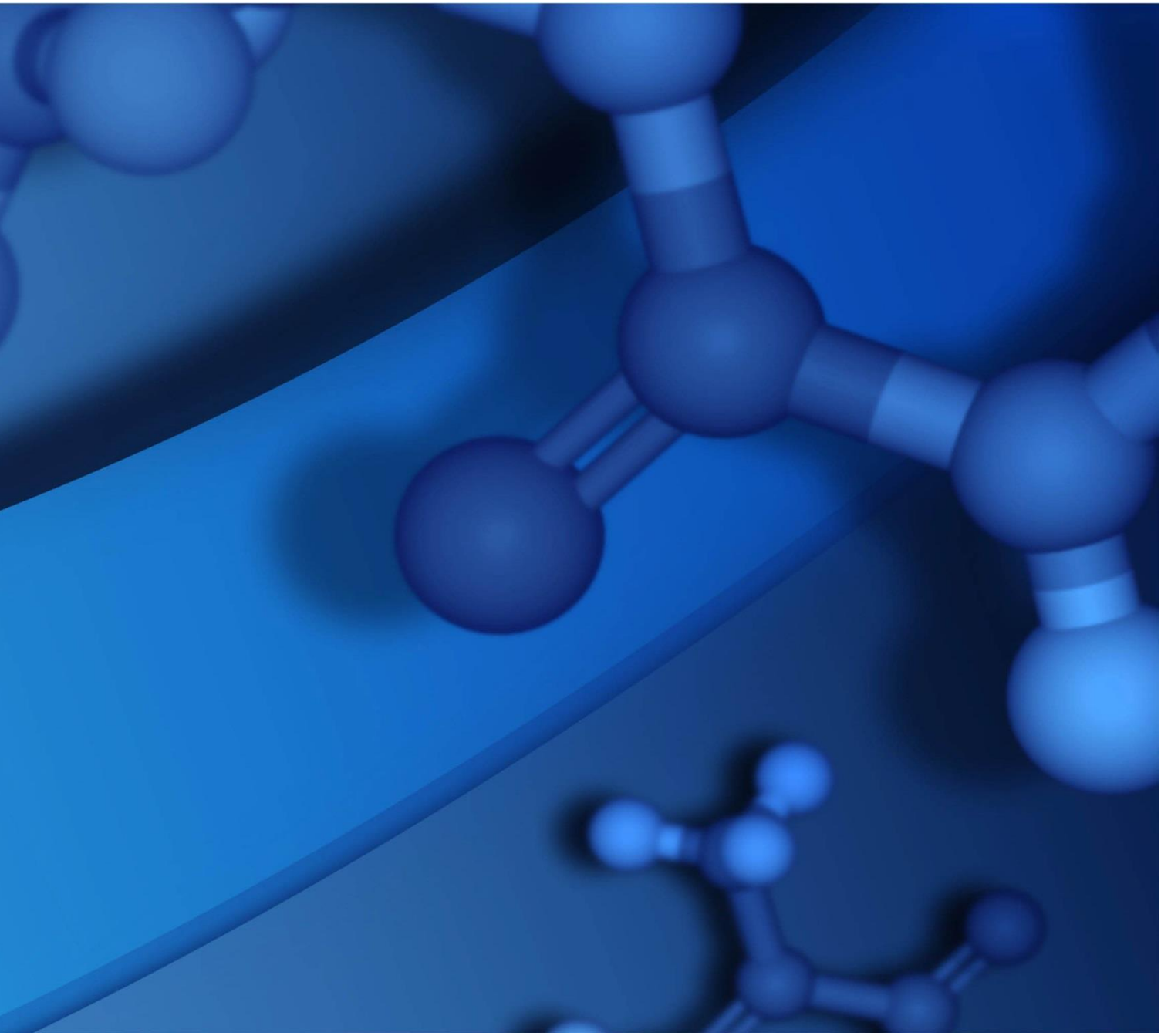
PRODUCT RELEASE DOCUMENT

BIOVIA Direct

2021

Minor



Third Party Notification

Your BIOVIA Direct 2021 installation contains third-party software components. For details, refer to the BIOVIA Program Directory at https://media.3ds.com/support/progdir/.

Acknowledgments and References

Dassault Systèmes may grant permission to republish or reprint its copyrighted materials. Requests should be submitted Dassault Systèmes Customer Support, either by visiting https://www.3ds.com/support and clicking **Call us** or **Submit a request**, or by writing to:

Dassault Systèmes Customer Support
10, Rue Marcel Dassault
78140 Vélizy-Villacoublay
FRANCE

Contents

Customer Release Details

Release of BIOVIA Direct 2021

Release Type: Minor

28 Oct 2020

Key Technical Points

- BIOVIA Direct 2021 supports Oracle 19c Release Update 7 (19.7.0.0.0) and Oracle 12c R2 (12.2.0.1). For details on Oracle 19c support, see Technote T61-2020.
- BIOVIA Direct 2021 adds extended options for similarity searches and calculations. In addition to the standard MDL keys, you can now add Accord or Pipeline Pilot fingerprints to the Direct domain index. The similarity operators provide a parameter to use the alternative fingerprints rather than the standard MDL keys.
- BIOVIA Direct 2021 adds new operators to calculate HELM2 and (R)InChI AuxInfo data.
- BIOVIA Direct 2021 supports the enhanced reaction format that adds representations for reagents, solvents, and catalysts.
- Due to a change that fixes the perception of stereo chemistry at double-bonds, you might need to re-register particular existing structures after creating a Direct 2021 domain index. For details see the Release Notes on DIR-4107 below or the *Upgrading Indexes* chapter in the *Administration Guide for BIOVIA Direct 2021*.
- As of version 2018, BIOVIA Direct requires a license file to pass a pre-installation license validation that provides access to the Direct installers.

Support Dates

For information regarding currently supported releases and the dates on which mainstream support ends, see https://www.3ds.com/support/policies-offers/lifecycle-policy/.

For complete details about support policies, see https://www.3ds.com/support/policies-offers/.

Getting Help

If you have questions, visit Dassault Systèmes Customer Support at https://www.3ds.com/support and click either **Call us** or **Submit a request**.

System Requirements and Installation

System Requirements

The system requirements for BIOVIA Direct 2021 are identified in the document *BIOVIA Direct System Requirements* for BIOVIA Direct 2021. This document is included in the BIOVIA Direct 2021 documentation zip file.

Installation

The installation of BIOVIA Direct 2021 is detailed in the documents *Installation and Configuration Guide (Windows)* and *Installation and Configuration Guide (Linux)* for BIOVIA Direct 2021, respectively. These files are included in the BIOVIA Direct 2021 documentation zip file.

Post Installation

Post installation steps for BIOVIA Direct 2021 are detailed in the documents *Installation and Configuration Guide (Windows)* and *Installation and Configuration Guide (Linux)* for BIOVIA Direct 2021, respectively. These files are included in the BIOVIA Direct 2021 documentation zip file.

Release Notes

This section provides information about the new features, enhancements, fixed defects, and known issues in BIOVIA Direct 2021. For information about previous interim releases of BIOVIA Direct 2017, see the Product Release Documents for the interim releases (Direct 2017 R2/2017 R2 SP1, Direct 2018, Direct 2018 SP1, Direct 2018 SP2, Direct 2019, Direct 2020).

Change Advice

Per the GAMP 5 and ICH Q9 Guideline, BIOVIA understands the requirement to apply Quality Risk Management to computerized systems. The Release Notes include Change Advice information to aid in the risk assessment of changes to a validated environment and determines the required effort to validate the release of BIOVIA Direct. The Change Advice:

- Provides a measure of the potential harm of the issue in the environment
- Identifies the complexity of change

Risk Definition

- **High:** Update affects several distinct areas of the code or is in a complex area of the code
- **Medium:** Update affects a few distinct areas of the code or is in a stable area of the code
- **Low:** Update affects an isolated area of the code or is in a stable area of the code

New Features

- BIOVIA Direct 2021 adds options to use Accord and Pipeline Pilot fingerprints for similarity calculation and searching
- BIOVIA Direct 2021 adds new operators to calculate HELM2 and (R)InChI AuxInfo data.
- BIOVIA Direct 2021 supports the new enhanced reaction format that adds representations for reagents, solvents and catalysts.

Enhancements

This release of BIOVIA Direct includes the following enhancements.

Description	Areas Affected	Risk	User Story

Since since Direct 9, a molecule or reaction containing a symmetrically tetra-substituted double bond, one that cannot be cis or trans, might not match itself with a FLEXMATCH search. The same symmetrically tetra-substituted structure with an "either" double bond might not match the structure with the normal double bond. This issue corrected in Direct 2021.
Newly inserted molecules or reactions will search correctly—structures in existing databases must be updated to correct the problem. Locate these structures using mdlaux.scanindex with the 'CTAB' argument. Any structures that do not find themselves, but do find themselves when ignoring stereochemistry, should be updated. For example:
`SQL> execute mdlaux.scanindex('mdctestix', 'ctab');`
 	
`BEGIN mdlaux.scanindex('mdctestix', 'ctab'); END;`
`*`
`ERROR at line 1:`
`MDL-1152: Molecule CTAB scan has errors, use MDLAUX.LOGTABLE to view errors`
`MDL-1112: MDLAUX.SCANINDEX has warnings or errors, use MDLAUX.LOGTABLE to view them`
`ORA-06512: at "C$DIRECTS2021.MDLAUXOP", line 4895`
`ORA-06512: at "C$DIRECTS2021.MDLAUXOP", line 4885`
`ORA-06512: at line 1`
 	
`SQL> select msg from table(mdlaux.logtable('mdctestix', 'asc'));`
 	
`MSG`
`----------------------------------------------`
`Scanning molecule index MDCTESTIX on table MDCTEST.MDCTEST`
	Molecule Index Upgrade
	Reaction Index Upgrade
		Low	DIR-4107

Description	Areas Affected	Risk	User Story
Scanning ROWID conversion table ROWID conversion table scan complete, there are 2 rows Scanning molecule CTABs Molecule at ROWID=AANqImAAPAAAADmAAA cannot find itself with FLEXMATCH Molecule at ROWID=AANqImAAPAAAADmAAA does find itself with FLEXMATCH when ignoring stereochemistry and data Sgroups (IGNORE=STE,DAT) Previous message occurred at ROWID=AANqImAAPAAAADmAAB Molecule CTAB scan has errors SQL> update mdctest set ctab=mol(molfile(ctab)) where rowid = 'AANqImAAPAAAADmAAB'; 1 row updated. SQL> commit; Commit complete.			
Direct FLEXMATCH queries which are meso (have stereocenters but no overall chirality) may now match more targets. Meso stereocenters have always been treated as wildcard stereochemistry. However stereo groups in the query were ignored. Stereo groups are now allowed to participate in the matching, disambiguating parity differences, which can allow more structures to be returned as hits.	Molecule Searching	Low	DIR-1218
Improved documentation for the CHIME operator. The Direct operator CHIME is a synonym for the operator MOLCHIME and is only meant to provide backward compatibility to older Direct versions. Because it is just a synonym it cannot be used within a PL/SQL procedure or function. This is now more clearly documented in the Direct Reference Guide.	Molecule Functions	Low	DIR-3711

Description	Areas Affected	Risk	User Story
BIOVIA Direct domain indexes are now tested and supported with Oracle Materialized Views created with the ' REFRESH ON COMMIT' and 'FAST' options.	Molecule Registration/Indexing	Low	DIR-3725

Added a new indexed similarity search type, in addition to the standard similarity search that uses substructure keys.
The new search type allows the user to create an index using either Accord or Scitegic fingerprints. The structure of the index is identical to that used for the standard substructure key similarity search, the difference being the length of the bit vector and the meaning of each bit.
In CREATE INDEX or ALTER INDEX REBUILD there is a new parameter available that creates a second inverted key index for the user-defined fingerprint:
 	
`FINGERPRINT=type[,size]`
 	
For example:
`create index testix on test (ctab) indextype is c$direct2021.mxixmdl parameters ('fingerprint=accord');`
`create index testix on test (ctab) indextype is c$direct2021.mxixmdl parameters ('fingerprint=ecfp_4,512');`
 	
type is either the keyword ACCORD or one of the standard Scitegic fingerprint names.
ACCORD will compute fingerprint keys very similar to what Accord once did, the difference being that chemical perceptions such as aromaticity may be different from the original Accord data cartridge. Users cannot expect that generated keys will be identical to what was in Accord. The Pipeline Chemistry method that computes Accord fingerprint keys always creates a vector of 384 bits.
	Molecule Index Creation, Molecule Index Maintenance, Molecule Index Upgrade, Molecule Registration/Indexing, Molecule Searching, Molecule Functions	Medium	DIR-3836

Description	Areas Affected	Risk	User Story

The Scitegic fingerprint type is one of the standard name, for example ECFP_6 or FCFP_4. The Scitegic fingerprint keys are folded to the desired *size* (number of bits), which is rounded up to a multiple of 32. If a size is not specified Direct uses 512 bits.
Methods that update pending inversions operate on both the normal substructure/similarity keys and the new fingerprint keys. Methods that rebuild keys will **only** rebuild the substructure/similarity keys. The only way to rebuild just the fingerprint keys is using the FINGERPRINT=type[,size] parameter in an ALTER INDEX REBUILD statement.
To perform a fingerprint similarity search, add the keyword FINGERPRINT to the argument to the MOLSIM or SIMILAR search query. For example:
`select cdbregno, similarity(1) from test where similar(ctab, '/tmp/caffeine.mol', '80 fingerprint', 1) =1 order by similarity(1);`
`select cdbregno from test where molsim(ctab, '/tmp/caffeine.mol', 'normal fingerprint') >= 80;`
The SUB and SUPER keywords also work with fingerprint similarity. An application can force the use of the FINGERPRINT search, even when the *fingerprint* keyword is not present, at either the Oracle session level or for all users of the cartridge:
`execute mdlaux.setflags('ForceFingerprintSearch');`
`execute mdlaux.setproperty('ForceFingerprintSearch', 'True');  -- Execute as user C$DIRECT2021`
There is no way to force use of substructure keys for similarity.			

MDLAUX.SMILESTOMOLFILE now converts an argument that is NULL or an empty string into a molecule with zero atoms and bonds (a "no-structure"). It is still an error to provide a NULL or	Molecule Functions	Low	DIR-3886

Description	Areas Affected	Risk	User Story
empty string molecule argument to other operators such as MOLFMLA, SSS, or FLEXMATCH.			
Direct can now compute and return InChI AuxInfo for molecules and reactions. Additionally, the function that converts a reaction InChI string to a reaction now accepts an optional argument containing the reaction AuxInfo to assist in the conversion. There are two new operators and two new functions to return the AuxInfo data: inchiauxinfo(ctab [, options]) - returns CLOB rinchiauxinfo(rctab [, options]) - returns CLOB mdlaux.inchiauxinfo(any-molecule [, options]) - returns CLOB mdlaux.rinchiauxinfo(any-reaction [, options]) - returns CLOB The function mdlaux.rinchitorxnfile now takes RAuxInfo as an optional second argument. If present is it used to assist in the conversion. Note that unlike the corresponding molecule conversion function the RInChI string is always required for the conversion. The RAuxInfo string provides some additional information but is not in itself complete.	Molecule Functions Reaction Functions	Low	DIR-4019
Improved the documentation for the MDLAUX.PREPAREINDEXEXPORT function. The function creates a table that must explicitly be included in exported Oracle dump file as it is not a member of the secondary Oracle tables of the domain index. This is now described more clearly in the Direct Administration Guide.	Other	Low	DIR-4025
Improved the BIOVIA Direct Reference Guide to describe how molecules and reactions used as query in the molsim, similar, rxnsim, and rxnmolsim operators must not include query features.	Other	Low	DIR-4043
Added a new function and operator to return the HELM version 2 (HELM2) string for a biological sequence molecule. The primary difference between HELM and HELM2 in Direct is that annotations	Molecule Functions	Low	DIR-4044

Description	Areas Affected	Risk	User Story
are stored at the end of the string in HELM and adjacent to the relevant monomer in HELM2. The new operator is: HELM2(ctab-column-name) The new function is: HELM2(table-or-index-name, any-molecule) Both return a statement-duration temporary CLOB containing the HELM version 2 string, or NULL if the input is not a biological sequence molecule.			
Enhanced the Direct Administration Guide to document that a change in the perception of coordinate bonds and hydrogen bonds in the count of ring bonds requires the recreation of molecular keys when updating existing domain indexes to Direct version 2021.	Molecule Index Upgrade Reaction Index Upgrade	Low	DIR-4066

Fixed Defects

This release of BIOVIA Direct includes the following fixed defects.

Severity	Description	Areas Affected	Risk	Defect
Major	Fixed an issue where an SSS query that includes a 6-member ring, one C-C aromatic bond (fused benzene ring), one or more A query atoms, and five single bonds would miss hits with an aromatic bond.	Molecule Searching	Low	DIR-4006

Severity	Description	Areas Affected	Risk	Defect
Major	Added an additional restriction when the HYD flexmatch switch is present, to prevent erroneous matching of two unequal structures. This restriction was preventing some tautomers from matching one another when the both the TAU and HYD flags were present. This problem is fixed. The additional HYD restriction is only added with an exact match when the flexmatch switches include BON, CHA, and RAD.	Molecule Searching	Low	DIR-4076
Minor	Fixed a memory issue where the MDLAUX.SCANINDEX function with a 'CTAB' option was consuming ~600 bytes of memory for each row scanned.	Cartridge Functions	Low	DIR-4037
Minor	Renamed and revised the appendix “Direct 8 to Direct 9 Chemistry Update” in the Direct Administration Guide to reflect recent releases. The appendix is renamed to “Direct 8 to Direct 2021 Chemistry Update”.	Other	Low	DIR-4053
Minor	Direct 2020 added support for returning the InChI key of a molecule with zero atoms (a "no-structure"). The computed key was incorrect, it should not have included the InChI prefix. This problem has been fixed in Direct 2021—the string returned by INCHIKEY and MDLAUX.INCHIKEY for a molecule with zero atoms is now correct.	Molecule Functions	Low	DIR-4064
Minor	Fixed an issue where some molecules with multiple fragments may have triggered a "MDL-2010: Warning: Removed invalid Chiral flag" warning in the execution of the mdlaux.scanindex command.	Molecule Index maintenance Reaction Index maintenance	Low	DIR-4079

Severity	Description	Areas Affected	Risk	Defect
Minor	Updated the BIOVIA Chemical Representation Guide to document a fix in Direct 2020 that changed the perception of the hydrogen count at an atom in a 5-member heteroaromatic ring. Accordingly, the descriptions of the HYD flexmatch switch and the Exact Match/As Drawn set of flexmatch switches are updated.	Other	Low	DIR-4085

Known Issues

This release of BIOVIA Direct has the following known issues.

Severity	Description	Defect
Minor	The third-party OpenEye library in BIOVIA Direct 2021 does not generate the correct IUPAC name for symmetric cyclobutanes with attached stereo bonds. BIOVIA has reported this problem to OpenEye.	DIR-3074