ELSEVIER

Reaxys®

The Reaxys Medicinal Chemistry Flat File RMCFF

Revision 2023-REV01

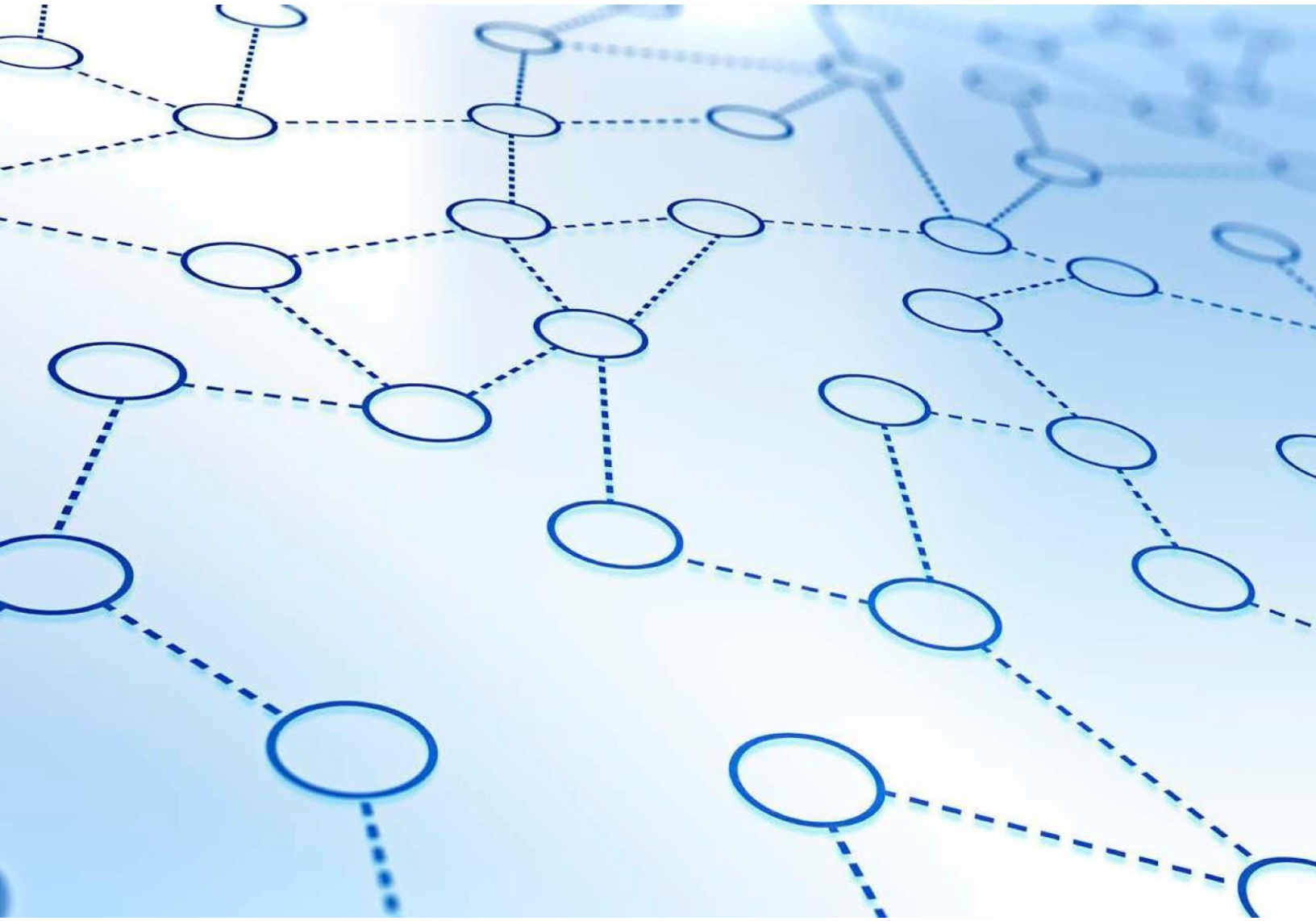


Table of contents

Introduction

The Reaxys Medicinal Chemistry Flat File (FF) delivers information on all substances from Reaxys and Reaxys Medicinal Chemistry, the related fact and reaction availability data, and all bioactivity data from Reaxys Medicinal Chemistry for inhouse usage. This enables organizations to incorporate this rich data set into their data mining or intranet search systems.

The format used to deliver the substances is following Biovia's CTfile Formats description (documentation available on the Dassault Systèmes Website at https://discover.3ds.com/ctfile-documentation-request-form) and comprises structures in Molfile format together with their identification data and a list of available facts and reactions for each compound. The fact availability list includes counts for bioactivity data points and targets. Unstructured substances are included as empty Molfiles. The substances file is segmented into multiple compressed SD files.

An additional document listing the registry numbers of all substances, and a document describing the differences between the current and the previous version of the substances file is provided.

The bioactivity data is delivered as a series of linked data files in XML format, using the Resource Description Framework (RDF), compliant with the Open PHACTS guidelines. The XML files contain information on bioassays, citations, bioactivity data points, substance facts and bioactivity targets.

The Reaxys Medicinal Chemistry FF is updated weekly together with the Reaxys Medicinal Chemistry database and each update will always comprise the full data set. Customers will be provided with the new version immediately after production for download.

Content Description

Substance Information

The following example shows a typical compound record available in the substances file:

```
R>Mv4.0007261214552D 1   1.00000     0.00000     0
6  6  0  0  0  0  0  0  0  0999 V2000
   0.5292    0.0000    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   1.5095    0.0000    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   0.0000   -0.8593    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   2.0397   -0.8593    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   0.5292   -1.7196    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   1.5095   -1.7196    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
1  2  2  0  0  0  0
1  3  1  0  0  0  0
2  4  1  0  0  0  0
3  5  2  0  0  0  0
4  6  2  0  0  0  0
5  6  1  0  0  0  0
```

```
M  REG 2
M  END
>  <IDE.XRN>
2

>  <IDE.MF>
C6H6

>  <IDE.MW>
78.1136

>  <IDE.INCHI>
UHOVQNZJYSORNB-UHFFFAOYSA-N

>  <IDE.RN>
71-43-2

>  <IDE.CN>
benzene|monobenzene|benzole|Ph-H|PhH
>  <IDE.FA>
YY(1)|PSD(1)|MB(1)|EXTID(2)|CDER(1)|MP(24)|BP(2)|CRYPH(1)|LIQPH(1)|CRT(4)|CRP
(2)|CRD(2)|CRV(1)|MEC(1)|COMP(1)|HFUS(3)|HVAP(14)|HSP(1)|CP(43)|OTHE(7)|OPT(1
)|MSUS(1)|DIC(1)|FINFO(2)|SLB(1)|LVSM(2)|MECM(8)|TRAM(2)|ADSM(3)|ASSM(5)|NMR(
3)|UV(3)|LUM(1)|FLU(2)|PHARM(2)|RX(129)|RCT(111)|PRO(18)|BEH(73)|PRE(40)|DET( 82)|DAT(9)|TRG(3)

>  <COPYRIGHT>
Copyright (C) 2019 Elsevier Life Sciences IP Limited. Reaxys is a trademark of Elsevier
Life Sciences IP Limited, used under license.
$$$$
```

Each record starts with a Molfile, followed by the field IDE.XRN which contains the Reaxys registry number for the given compound, and other fields. A record ends with a sequence of four $ symbols. The complete list of available fields is given in the following:

Field code	Description
IDE.XRN	Reaxys registry number
IDE.RN	CAS registry number (optional)
IDE.CN	Chemical name (optional)
IDE.MF	Molecular formula (optional)
IDE.MW	Molecular weight (optional)
IDE.STYPE	Type of substance (optional; only when containing “markush structure” in order to indicate the difference to “normal” structures)
IDE.INCHI	InChI key (optional)
IDE.FA	Fact availability list
COPYRIGHT	Copyright statement

The field “IDE.FA” contains the list of available facts and reactions for each compound which includes counts for bioactivity data points and targets. The facts are listed in one string separated by vertical lines.

The numbers in brackets denote the occurrence of the given fact, e.g. PRE(40) means that Reaxys has forty preparations available for the given compound, DAT(9) means that Reaxys Medicinal Chemistry has nine bioactivities available for the given compound. A complete list of the possible facts is given in the Appendix.

Bioactivity Data

The following example shows a typical record available in the assays XML file:

```
<rdf:Description rdf:about="Assay/#436483">
       <fct:AFTYPE>Cell behaviour</fct:AFTYPE>
       <fct:ADESC>Effect: mTOR activation
       Bioassay: salivary high metastatic ACC-M cells
       ACC: adenoid cystic carcinoma; mTOR: mammalian target of rapamycin; Western
       blotting</fct:ADESC>
       <fct:CATEG>Toxicity/Safety Pharmacology</fct:CATEG>
       <fct:ANAME>Apoptosis</fct:ANAME>
       <fct:EFFECT>Apoptotic</fct:EFFECT>
       <fct:Copyright> Copyright (C) 2019 Elsevier Life Sciences IP Limited. Reaxys is
       a trademark of Elsevier Life Sciences IP Limited, used under license.
       </fct:Copyright>
</rdf:Description>
```

Each record corresponds to a bioassay and contains the following nodes:

Node	Description
<rdf:Description rdf:about="Assay/#xxxx"> </rdf:Description>	Assays in the file are numbered using the bioassay ID (DAT.AID)
<fct:ANAME></fct:ANAME>	Bioassay name (DAT.ANAME)
<fct:AFTYPE></fct:AFTYPE>	Bioassay full type (DAT.AFTYPE)
<fct:CATEG></fct:CATEG>	Bioassay category (DAT.CATEG)
<fct:MODEL></fct:MODEL>	Animal model (DAT.MODEL)
<fct:PATHO></fct:PATHO>	Bioassay pathology (DAT.PATHO)
<fct:ACTTRG></fct:ACTTRG>	Putative action on target (DAT.ACTTRG)
<fct:ADESC></fct:ADESC>	Bioassay description (DAT.ADESC)
<fct:EFFECT></fct:EFFECT>	Effect (DAT.EFFECT)
<fct:Copyright></fct:Copyright>	Copyright statement

The following example shows a typical record available in the citations XML file:

```
<rdf:Description rdf:about="Citation/#62510551">
       <fct:ENTRYDATE>2014-02-11</fct:ENTRYDATE>
       <fct:DT>Article</fct:DT>
```

```
        <fct:AU>Vadim T. Ivanov; Tatyana M. Andronova; Mikhail V. Bezrukov;
        Vera A. Rar; Evgenii A. Makarov; Sergei A. Kozmin; Maria V. Astapova; Tamara I.
        Barkova; Vladimir A. Nesmeyanov</fct:AU>
          <fct:JT>Pure and applied chemistry. Chimie pure et appliquee</fct:JT>
        <fct:VL>59</fct:VL>
        <fct:NB>3</fct:NB>
        <fct:PY>1987</fct:PY>
        <fct:PAG>317 - 324</fct:PAG>
        <fct:ISSN>0033-4545</fct:ISSN>
        <fct:Copyright> Copyright (C) 2019 Elsevier Life Sciences IP Limited. Reaxys is
        a trademark of Elsevier Life Sciences IP Limited, used under license.
        </fct:Copyright>
</rdf:Description>
```

Each record corresponds to a citation and contains the following nodes:

Node	Description
<rdf:Description rdf:about="Citation/#xxxx"> </rdf:Description>	Citations in the file are numbered using the citation RN (CNR.CNR)
<fct:ENTRYDATE></fct:ENTRYDATE>	Date document added to Reaxys (CNR.CED)
<fct:DT></fct:DT>	Document type (CIT.DT)
<fct:AU></fct:AU>	Authors (CIT.AU)
<fct:PA></fct:PA>	Patent assignee (CIT.PA)
<fct:PCC></fct:PCC>	Patent country code (CIT.PCC)
<fct:PN></fct:PN>	Patent number (CIT.PN)
<fct:PPY></fct:PPY>	Patent year (CIT.PPY)
<fct:PK></fct:PK>	Patent status (CIT.PK)
<fct:JT></fct:JT>	Journal title (CIT.JT)
<fct:CC></fct:CC>	Country code (CIT.CC)
<fct:LA></fct:LA>	Language code (CIT.LA)
<fct:LO></fct:LO>	Publisher location (CIT.LO)
<fct:VL></fct:VL>	(Series) volume (CIT.VL)
<fct:NB></fct:NB>	Number (CIT.NB)
<fct:PY></fct:PY>	Publication year (CIT.PY)
<fct:PAG></fct:PAG>	Page (CIT.PAG)
<fct:DOI></fct:DOI>	DOI (CIT.DOI)
<fct:ISSN></fct:ISSN>	ISSN (CIT.ISSN)
<fct:Copyright></fct:Copyright>	Copyright statement

The following example shows a typical record available in the datapoints XML file:

```
<rdf:Description rdf:about="Bioactivity/#9837658">
```

```
        <fct:ENTRYDATE>2013-12-16</fct:ENTRYDATE>
        <fct:hasCitation rdf:resource="Citation/#6538965"/>
        <fct:CTYPE>MAIN</fct:CTYPE>
        <fct:hasSubstance rdf:resource="Substance/#10323730"/>
        <fct:MNAME>I</fct:MNAME>
        <fct:MDOSE>3 mM</fct:MDOSE>
        <fct:hasAssay rdf:resource="Assay/#1405052"/>
        <fct:BCELL>CACO2</fct:BCELL>
        <fct:VTYPE>Ki</fct:VTYPE>
        <fct:UNIT>mM</fct:UNIT>
        <fct:VALUE>0.38</fct:VALUE>
        <fct:PAUREUS>3.42</fct:PAUREUS>
        <fct:PAUORIG>pAur = -log(value)</fct:PAUORIG>
        <fct:Copyright> Copyright (C) 2019 Elsevier Life Sciences IP Limited. Reaxys is
        a trademark of Elsevier Life Sciences IP Limited, used under license.
        </fct:Copyright>
</rdf:Description>
```

Each record corresponds to a bioactivity data point and contains the following nodes:

Node	Description
<rdf:Description rdf:about="Bioactivity/#xxxx"> </rdf:Description>	Bioactivity data points in the file are numbered using the bioactivity datapoint ID (DAT.ID)
<fct:ENTRYDATE></fct:ENTRYDATE>	Date bioactivity data point added to Reaxys (DAT.ED)
<fct:hasCitation></fct:hasCitation>	Link to a citation in the citations XML file (DAT.CIT, CNR.CNR)
<fct:hasAssay></fct:hasAssay>	Link to a bioassay in the assays XML file (DAT.AID)
<fct:hasTarget></fct:hasTarget>	Link to a bioactivity target in the targets XML file (DAT.TID, TARGET.ID)
<fct:CTYPE></fct:CTYPE>	Substance type (DAT.CTYPE)
<fct:hasSubstance></fct:hasSubstance>	Link to a substance in the facts XML file, or to a compound record in the substances file (DAT.MRN, IDE.XRN)
<fct:MNAME></fct:MNAME>	Substance name (DAT.MNAME)
<fct:MROUTE></fct:MROUTE >	Substance route of adm. (DAT.MROUTE)
<fct:MREGIM></fct:MREGIM>	Substance dosing regimen (DAT.MREGIM)
<fct:MDOSE></fct:MDOSE>	Substance dose (DAT.MDOSE)
<fct:VTYPE></fct:VTYPE>	Measurement parameter (DAT.VTYPE)
<fct:TEXT></fct:TEXT>	Qualitative results (DAT.TEXT)
<fct:EXACT></fct:EXACT>	Qualitative value (DAT.EXACT)
<fct:VALUE></fct:VALUE>	Quantitative value (DAT.VALUE)
<fct:DEV></fct:DEV>	Deviation (DAT.DEV)

<fct:UNIT></fct:UNIT>	Unit (DAT.UNIT)
<fct:PAUREUS></fct:PAUREUS>	Measurement pX (DAT.PAUREUS)
<fct:PAUORIG></fct:PAUORIG>	pX origin (DAT.PAUORIG)
<fct:SIGNIF></fct:SIGNIF>	Statistical significance (qualitative) (DAT.SIGNIF)
<fct:PVALUE></fct:PVALUE>	Statistical significance (p-value) (DAT.PVALUE)
<fct:MCOUNT></fct:MCOUNT>	Measurement count (DAT.MCOUNT)
<fct:BNAME></fct:BNAME>	Biological material name (DAT.BNAME)
<fct:BCELL></fct:BCELL>	Cells/cell lines (DAT.BCELL)
<fct:BPART></fct:BPART>	Cell fraction (DAT.BPART)
<fct:BTISSUE></fct:BTISSUE>	Organs/tissues (DAT.BTISSUE)
<fct:BSTRAIN></fct:BSTRAIN>	Strain (DAT.BSTRAIN)
<fct:BSTATE></fct:BSTATE>	Population state (DAT.BSTATE)
<fct:BSPECIE></fct:BSPECIE>	Biological species (DAT.BSPECIE)
<fct:Copyright></fct:Copyright>	Copyright statement

The following example shows a typical record available in the facts XML file:

```
<rdf:Description rdf:about="Substance/#18314799">
       <fct:RN>748106-93-6</fct:RN>
       <fct:CN>5-(acetyloxy)-2-hydroxybenzenesulfonic acid</fct:CN>
       <fct:CN>5-acetoxy-2-hydroxybenzenesulfonic acid</fct:CN>
       <fct:CN>5A-2HBS</fct:CN>
       <fct:MF>C8H8O6S</fct:MF>
       <fct:MW>232.214</fct:MW>
       <fct:INCHI>JFGISRUXKVGGNF-UHFFFAOYSA-N</fct:INCHI>
       <fct:ENTRYDATE>2008/12/05</fct:ENTRYDATE>
       <fct:LOGP>1.282</fct:LOGP>
       <fct:HDONOR>2</fct:HDONOR>
       <fct:HACCOR>4</fct:HACCOR>
       <fct:ROTBND>3</fct:ROTBND>
       <fct:TPSA>109.28</fct:TPSA>
       <fct:LIPINSKI>4</fct:LIPINSKI>
       <fct:VEBER>2</fct:VEBER>
       <fct:SUBSTANCELABEL>
          <rdf:Description>
              <fct:hasCitation rdf:resource="Citation/#101773233"/>
              <fct:LABEL>Amf</fct:LABEL>
          </rdf:Description>
       </fct:SUBSTANCELABEL>
       <fct:SUBSTANCELABEL>
          <rdf:Description>
              <fct:hasCitation rdf:resource="Citation/#8641904"/>
              <fct:hasCitation rdf:resource="Citation/#75353117"/>
              <fct:LABEL>18</fct:LABEL>
```

```
            </rdf:Description>
        </fct:SUBSTANCELABEL>
        <fct:Copyright> Copyright (C) 2019 Elsevier Life Sciences IP Limited. Reaxys is
        a trademark of Elsevier Life Sciences IP Limited, used under license.
        </fct:Copyright>
</rdf:Description>
```

Each record corresponds to a substance and contains the following nodes:

Node	Description
<rdf:Description rdf:about="Substance/#xxxx"> </rdf:Description>	Substances in the file are numbered using the substance RN (IDE.XRN)
<fct:RN></fct:RN>	CAS registry number (IDE.RN)
<fct:CN></fct:CN>	Chemical name (IDE.CN)
<fct:MF></fct:MF>	Molecular formula (IDE.MF)
<fct:MW></fct:MW>	Molecular weight (IDE.MW)
<fct:INCHI></fct:INCHI>	InChI key (IDE.INCHI)
<fct:ENTRYDATE></fct:ENTRYDATE>	Date substance added to Reaxys (IDE.ED)
<fct:LOGP></fct:LOGP>	LogP (CALC.LOGP)
<fct:HDONOR></fct:HDONOR>	H bond donors (CALC.HDONOR)
<fct:HACCOR></fct:HACCOR>	H bond acceptors (CALC.HACCOR)
<fct:ROTBND></fct:ROTBND>	Rotatable bonds (CALC.ROTBND)
<fct:TPSA></fct:TPSA>	Polar surface area (CALC.TPSA)
<fct:LIPINSKI></fct:LIPINSKI>	Lipinski number (CALC.LIPINSKI)
<fct:VEBER></fct:VEBER>	Veber number (CALC.VEBER)
<fct:SUBSTANCELABEL> <rdf:Description> <fct:hasCitation></fct:hasCitation> <fct:LABEL></fct:LABEL> </rdf:Description> </fct:SUBSTANCELABEL>	The SUBSTANCELABEL node contains: Link to a citation in the citations XML file (LB.L, CNR.CNR) Substance label used in citation (LB.LB)
<fct:Copyright></fct:Copyright>	Copyright statement

The following example shows a typical record available in the targets XML file:

```
<rdf:Description rdf:about="Target/#2866210">
        <fct:ENTRYDATE>2014-02-13</fct:ENTRYDATE>
        <fct:KEY>Acetylcholine receptor subunit alpha [human]:Wild + Nicotinic
        acetylcholine receptor (Delta 1) [human]:Wild + Nicotinic acetylcholine
        receptor (Gamma 1) [human]:Wild + Acetylcholine receptor subunit beta
        [human]:Wild</fct:KEY>
        <fct:SKEY>Acetylcholine receptor subunit alpha [human] + Nicotinic
        acetylcholine receptor (Delta 1) [human] + Nicotinic acetylcholine receptor
        (Gamma 1) [human] + Acetylcholine receptor subunit beta [human]</fct:SKEY>
        <fct:DETAILS>Acetylcholine receptor subunit alpha [human]:Wild + Nicotinic
```

```
        acetylcholine receptor (Delta 1) [human]:Wild + Nicotinic acetylcholine
        receptor (Gamma 1) [human]:Wild + Acetylcholine receptor subunit beta
        [human]:Wild</fct:DETAILS>
        <fct:NATURE>Wild</fct:NATURE>
       <fct:SUBUNIT>
              <rdf:Description>
                      <fct:SUBPROTN>Acetylcholine receptor subunit alpha</fct:SUBPROTN>
                     <fct:SUBSPECIE>human</fct:SUBSPECIE>
                          <fct:SUBNATURE>Wild</fct:SUBNATURE>
                          <fct:SUBSTOICHIO>2</fct:SUBSTOICHIO>
                          <fct:SUBSYNONYM>achra</fct:SUBSYNONYM>
                          <fct:SUBSYNONYM>chnra</fct:SUBSYNONYM>
                          <fct:SUBSYNONYM>chrna1</fct:SUBSYNONYM>
                      <uniprot:TUNIPROT>p02708</uniprot:TUNIPROT>
                     <pdb:TPDB>4zjs</pdb:TPDB>
               </rdf:Description>
    </fct:SUBUNIT>
      <fct:SUBUNIT>
            <rdf:Description>
                    <fct:SUBPROTN>Nicotinic acetylcholine receptor (Delta
                    1)</fct:SUBPROTN>
                    <fct:SUBSPECIE>human</fct:SUBSPECIE>
                    <fct:SUBNATURE>Wild</fct:SUBNATURE>
            </rdf:Description>
       </fct:SUBUNIT>
      <fct:SUBUNIT>
            <rdf:Description>
                    <fct:SUBPROTN>Nicotinic acetylcholine receptor (Gamma
                    1)</fct:SUBPROTN>
                    <fct:SUBSPECIE>human</fct:SUBSPECIE>
                    <fct:SUBNATURE>Wild</fct:SUBNATURE>
            </rdf:Description>
     </fct:SUBUNIT>
      <fct:SUBUNIT>
           <rdf:Description>
                    <fct:SUBPROTN>Acetylcholine receptor subunit beta</fct:SUBPROTN>
                    <fct:SUBSPECIE>human</fct:SUBSPECIE>
                    <fct:SUBNATURE>Wild</fct:SUBNATURE>
                    <fct:SUBSYNONYM>achrb</fct:SUBSYNONYM>
                    <fct:SUBSYNONYM>chrnb</fct:SUBSYNONYM>
                    <fct:SUBSYNONYM>chrnb1</fct:SUBSYNONYM>
                    <uniprot:TUNIPROT>p11230</uniprot:TUNIPROT>
            </rdf:Description>
</fct:SUBUNIT>
       <fct:Copyright> Copyright (C) 2019 Elsevier Life Sciences IP Limited. Reaxys is
       a trademark of Elsevier Life Sciences IP Limited, used under license.
       </fct:Copyright>
</rdf:Description>
```

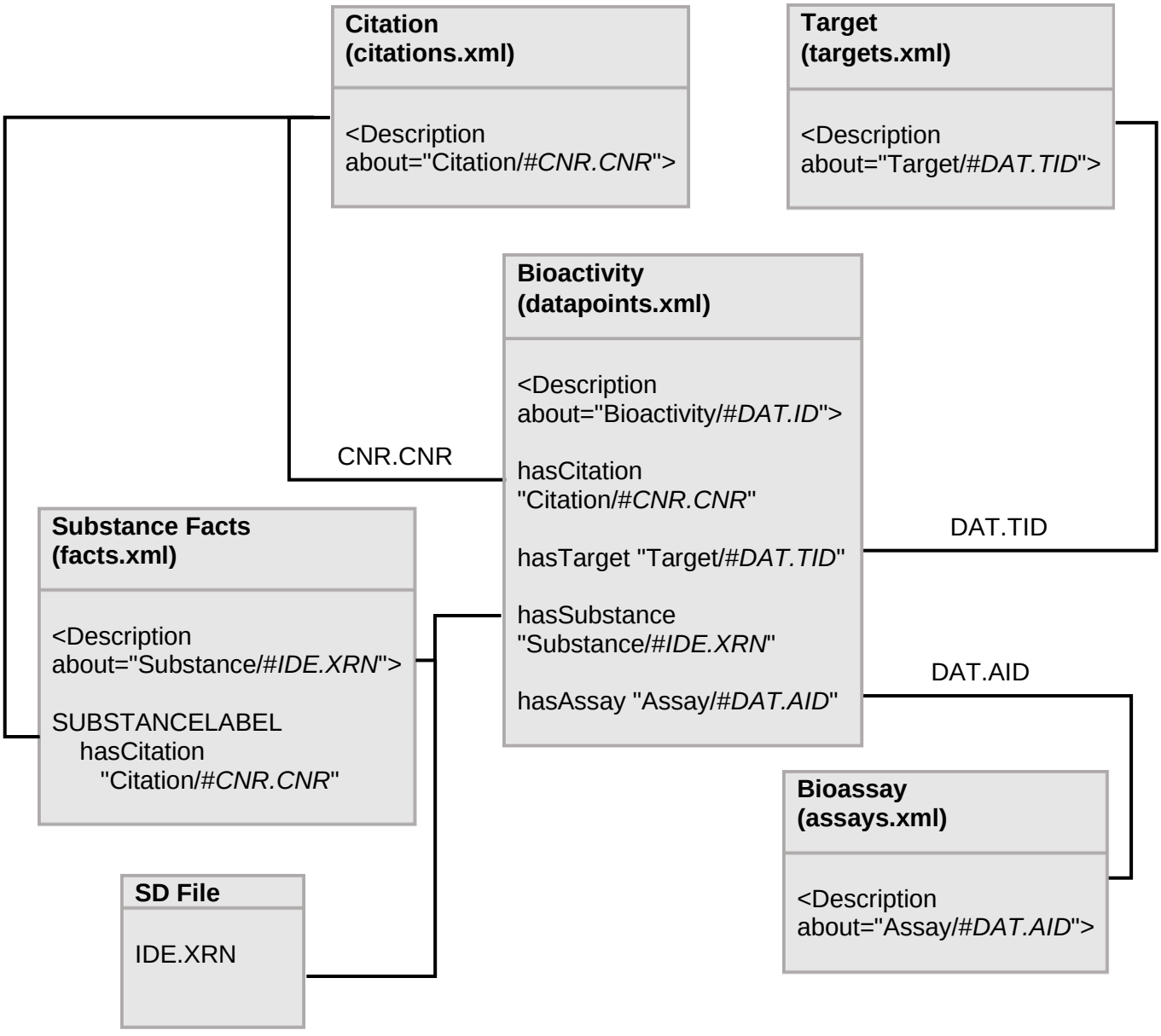
Each record corresponds to a bioactivity target and contains the following nodes:

Node	Description
<rdf:Description rdf:about="Target/#xxxx"> </rdf:Description>	Bioactivity targets in the file are numbered using the target ID (TARGET.ID)
<fct:ENTRYDATE></fct:ENTRYDATE>	Date target added to Reaxys (TARGET.ED)

<fct:SITE></fct:SITE>	Compound action site (TARGET.SITE)
<fct:ROLE></fct:ROLE>	Target role (TARGET.ROLE)
<fct:NATURE></fct:NATURE>	Target nature (TARGET.NATURE)
<fct:DETAILS></fct:DETAILS>	Mutant/chimera details (TARGET.DETAILS)
<fct:KEY></fct:KEY>	Target key (TARGET.KEY)
<fct:SKEY></fct:SKEY>	Target short key (TARGET.SKEY)
<fct:SUBUNIT> <rdf:Description> <fct:SUBPROTN></fct:SUBPROTN> <fct:SUBSPECIE></fct:SUBSPECIE> <fct:SUBNATURE></fct:SUBNATURE> <fct:SUBTRANSFECT></fct:SUBTRANSFECT> <fct:SUBSTOICHIO></fct:SUBSTOICHIO> <fct:SUBACTSITE></fct:SUBACTSITE> <fct:PMID></fct:PMID> <fct:PMPROT></fct:PMPROT> <fct:PMSPECIE></fct:PMSPECIE> <fct:SUBSYNONYM></fct:SUBSYNONYM> <fct:SUBUID></fct:SUBUID> <uniprot:TUNIPROT></uniprot:TUNIPROT > <pdb:TPDB></pdb:TPDB> <interpro:TINTERPRO></interpro:TINTERPRO> <go:TGOID></go:TGOID> </rdf:Description> </fct:SUBUNIT>	The “subunit” node contains subnodes for: - Subunit protein (SUBUNIT.PROTN) - Subunit species (SUBUNIT.SPECIE) - Subunit nature (SUBUNIT.NATURE) - Transfection (SUBUNIT.TRANSFECT) - Stoichiometry (SUBUNIT.STOICHIO) - Active site indicator (SUBUNIT.ACTSITE) - Promoter identifier (SUBUNIT.PMID) - Promoter protein (SUBUNIT.PMPROT) - Promoter species (SUBUNIT.PMSPECIE) - Subunit synonym (SUBUNIT.SYNONYM) - RMC taxonomy concept UID (SUBUNIT.UID) - Uniprot ID (SUBUNIT.UNIPROT) - PDB ID (SUBUNIT.PDB) - InterPro ID (SUBUNIT.INTERPRO) - GO ID (SUBUNIT.GO)
<fct:Copyright></fct:Copyright>	Copyright statement

Entity Relationships

The information on targets, bioactivities, and substances are distributed over six individual data sets. Therefore, this individual data have to be assembled to reconstruct a complete view. For this purpose, each entity has a unique identifier assigned, which links the individual data sets (see diagram below).



Download

Customers will be provided with the new version immediately after production for download from the Amazon S3. The Amazon S3 (Simple Storage Service) is an online file storage web service offered by Amazon Web Services (AWS). For downloading the Reaxys Medicinal Chemistry FF, the AWS Command Line Interface (CLI) must be obtained from the Amazon web site at https://aws.amazon.com/cli/. The AWS CLI is available for Microsoft Windows or Linux operating systems.

Upon starting the AWS CLI for the first time, the client needs to be configured by typing the following command into the Windows Command Prompt or the Linux Console:

```
aws configure
```

The user is required to enter their Access Key ID, Secret Access Key, default region name and default output format. Access Key ID and Secret Access Key are customer specific and will be provided to the customer by email. As the default region name, "**us-east-1"** should be entered. As the default output format, the default should be selected, by pressing the "Enter" key.

The configuration can be found at any time in the following path:

Linux: `/home/USER/.aws/config`

Windows: `C:\Users\USER\.aws\config`

The name of the accessible directories on the Amazon S3 can be checked by typing the following command into the Windows Command Prompt or the Linux Console:

```
aws s3 ls s3://rmc-ff-download
```

Each new version of the Reaxys Medicinal Chemistry FF will be available from a separate directory.

Single files can be copied from a directory on the Amazon S3 to a local directory by typing the following command into the Windows Command Prompt or the Linux Console:

Linux:
```
aws s3 cp s3://rmc-ff-download /DIRECTORY/FILE.zip
/home/USER/PATH/
```

Windows: `aws s3 cp s3://rmc-ff-download/DIRECTORY/FILE.zip C:\Users\USER\PATH\`

Multiple files or directories on the Amazon S3 can be synchronized with local files or a local directory by typing the following command into the Windows Command Prompt or the Linux Console:

Linux: `aws s3 sync s3://rmc-ff-download/ /home/USER/PATH/`

Windows: `aws s3 sync s3://rmc-ff-download/ C:\Users\USER\PATH\`

Appendix

A complete list of facts is given in the following table. In the “available from” column, a description of the license is given that is required to retrieve the data. Possible licenses are: RX (Reaxys only license), RMC (Reaxys Medicinal Chemistry only license), and combined (both Reaxys and Reaxys Medicinal Chemistry license).

Field code	Field name	Available from *
ADSM	Adsorption (MCS)	RX, combined
ALLOY	Alloy composition	RX, combined
ASSM	Association (MCS)	RX, combined
AUTI	Autoignition	RX, combined
AZE	Azeotropes (MCS)	RX, combined
BEH	Chemical behavior presence	RX, combined
BIO	Bioaccumulation, biomagnification and biomonitoring	RX, combined
BIOD	Biodegradation	RX, combined
BP	Boiling point	RX, combined
BSPM	Boundary surface phenomena (MCS)	RX, combined
BV	Bulk viscosity	RX, combined
CALC	Druglikeness	RMC, combined
CAT	Catalyst investigation	RX, combined
CDEN	Density of the crystal	RX, combined
CDER	Derivative	RX, combined
CDIC	Circular dichroism	RX, combined
CHROMAT	Chromatographic data	RX, combined
CIP	Electron binding	RX, combined
CMC	Critical micelle concentration (MCS)	RX, combined
CNF	Conformation	RX, combined
COMP	Compressibility	RX, combined
CP	Heat capacity Cp	RX, combined
CP0	Heat capacity Cp0	RX, combined

CPD	Crystal property description	RX, combined
CPEM	Complex phase equilibria (MCS)	RX, combined
CPTP	Transition point(s) of crystalline modification(s)	RX, combined
CRD	Critical density	RX, combined
CRP	Critical pressure	RX, combined
CRT	Critical temperature	RX, combined
CRV	Critical volume	RX, combined
CRYPH	Crystal phase	RX, combined
CSG	Space group	RX, combined
CSYS	Crystal system	RX, combined
CV	Heat capacity Cv	RX, combined
DAT	Bioactivity data points	RMC, combined
DE	Dissociation exponent	RX, combined
DEN	Density of the liquid	RX, combined
DET	Detailed reaction presence	RX, combined
DFM	Molecular deformation	RX, combined
DIC	Dielectric constant	RX, combined
DP	Decomposition	RX, combined
DV	Dynamic viscosity	RX, combined
EBC	Energy barriers	RX, combined
ECA	Exposure assessment	RX, combined
ECC	Concentration in the environment	RX, combined
ECDH	Abiotic degradation, hydrolysis	RX, combined
ECDP	Abiotic degradation, photolysis	RX, combined
ECS	Stability in soil	RX, combined
ECT	Ecotoxicology	RX
ECTD	Transport and distribution	RX, combined
EDIS	Dissociation energy	RX, combined
EDM	Electrical data (MCS)	RX, combined
ELCB	Electrochemical behavior	RX, combined
ELCH	Electrochemistry data	RX, combined
ELE	Electrical data	RX, combined
ELP	Electrical polarizability	RX, combined
ELYC	Electrolytic conductivity	RX, combined
EM	Electrical moment	RX, combined
ENEM	Energy data (MCS)	RX, combined
EOD	Oxygen demand	RX, combined
ESR	ESR spectroscopy	RX, combined
EXPL	Explosion limits	RX, combined
EXTID	External identifiers	RX, combined

FINFO	Further information	RX, combined
FLAP	Flash point	RX, combined
FLU	Fluorescence spectroscopy	RX, combined
GP	Gas phase	RX, combined
HCOM	Enthalpy of combustion	RX, combined
HEN	Henry constant (MCS)	RX, combined
HFOR	Enthalpy of formation	RX, combined
HFUS	Enthalpy of fusion	RX, combined
HHDG	Enthalpy of hydrogenation	RX, combined
HPTP	Enthalpies of other phase transitions	RX, combined
HSP	Enthalpy of sublimation	RX, combined
HVAP	Enthalpy of vaporization	RX, combined
IDA	Interatomic distances and angles	RX, combined
IEP	Isoelectric point pH	RX, combined
INP	Isolation from natural product	RX, combined
IP	Ionization potential	RX, combined
IR	IR spectroscopy	RX, combined
KV	Kinematic Viscosity	RX, combined
LB	Substance label	RX, combined
LIGM	Multi-center ligands	RX, combined
LIGO	One-center ligands	RX, combined
LIQPH	Liquid phase	RX, combined
LLSM	Liquid/liquid systems (MCS)	RX, combined
LPTP	Transition point(s) of liquid modification(s)	RX, combined
LSSM	Liquid/solid systems (MCS)	RX, combined
LUM	Luminescence spectroscopy	RX, combined
LVSM	Liquid/vapour systems (MCS)	RX, combined
MAG	Magnetic data	RX, combined
MB	Molecular BIN	RX, combined
MEC	Mechanical properties	RX, combined
MECM	Mechanical and physical properties (MCS)	RX, combined
MP	Melting point	RX, combined
MS	Mass spectrometry	RX, combined
MSUS	Magnetic susceptibility	RX, combined
MUT	Mutarotation	RX, combined
NMR	NMR spectroscopy	RX, combined
NQR	NQR spectroscopy	RX, combined
ODM	Optical data (MCS)	RX, combined
OPT	Optics	RX, combined
ORD	Optical rotatory dispersion	RX, combined

ORP	Optical rotatory power	RX, combined
OSM	Other spectroscopic methods	RX, combined
OTHE	Other thermochemical data	RX, combined
PHARM	Pharmacological data	RX
PHO	Phosphorescence spectroscopy	RX, combined
POT	Electrochemical characteristics	RX, combined
POW	Partition octan-1-ol/water (MCS)	RX, combined
PRE	Preparation presence	RX, combined
PRO	Presence as product	RX, combined
PSD	Patent-specific data	RX, combined
PUR	Purification	RX, combined
QUAN	Quantum chemical calculations	RX, combined
RAMAN	Raman spectroscopy	RX, combined
RCT	Presence as reactant	RX, combined
RI	Refractive index	RX, combined
ROT	Rotational spectroscopy	RX, combined
RSTR	Related structure	RX, combined
RX	Reaction	RX, combined
SDIC	Static dielectric constant	RX, combined
SDIF	Self-diffusion	RX, combined
SEQ	Sequence	RX, combined
SLB	Solubility (MCS)	RX, combined
SLBP	Solubility product (MCS)	RX, combined
SOLM	Solution behavior (MCS)	RX, combined
SOUND	Sound properties	RX, combined
SP	Sublimation	RX, combined
ST	Surface tension	RX, combined
TD	Transport data	RX, combined
TEXP	Thermal expansion	RX, combined
TP	Triple point	RX, combined
TRG	Bioactivity targets	RMC, combined
TRAM	Transport phenomena (MCS)	RX, combined
USE	Use	RX, combined
UV	UV/VIS spectroscopy	RX, combined
VP	Vapour pressure	RX, combined
XREF	Crossfile reference	RX, combined
XS	Cross-Sections	RX, combined
YY	Structure	RX, combined

*) Description of the license that is required to retrieve data. Possible licenses are: RX (Reaxys only license), RMC (Reaxys Medicinal Chemistry only license), and combined (both Reaxys and Reaxys Medicinal Chemistry license).