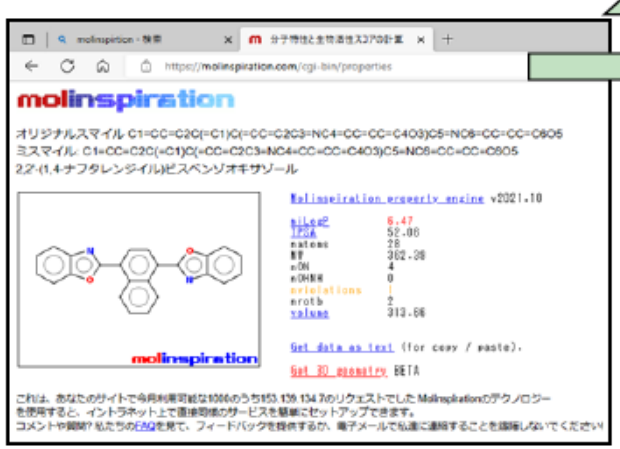
SML6　の予測精度を向上させるデータ・ベースの利用

https://pubchem.ncbi.nlm.nih.gov/
https://www.molinspiration.com/cgi/properties

2023_11_11
2023_11_26　改訂版

Table_03：移行モデルPiringer式とWelle式による拡散係数Dpの違い　Dp値は20℃における値です。

Migrant	CAS.No	Moleqular Weight	Molecular Volume	Dp Welle	Dp Piringer Upper	Dp_Piringer Realistic
アセトアルデヒド	75-07-0	44.05				
ホルムアルデヒド	50-00-0	30.03				
トルエン	108-88-3	92.14				
ベンゼン	71-43-2	78.11				
ビスヒドロキシエチルテレフタル酸	959-36-2					
ジエチレングリコール	111-46-6	106.12				



テクニカル・ノート　SML6.6_No.14　では
より実測値に近い溶出量が得られる
移行モデル_Welle式　を紹介しています。

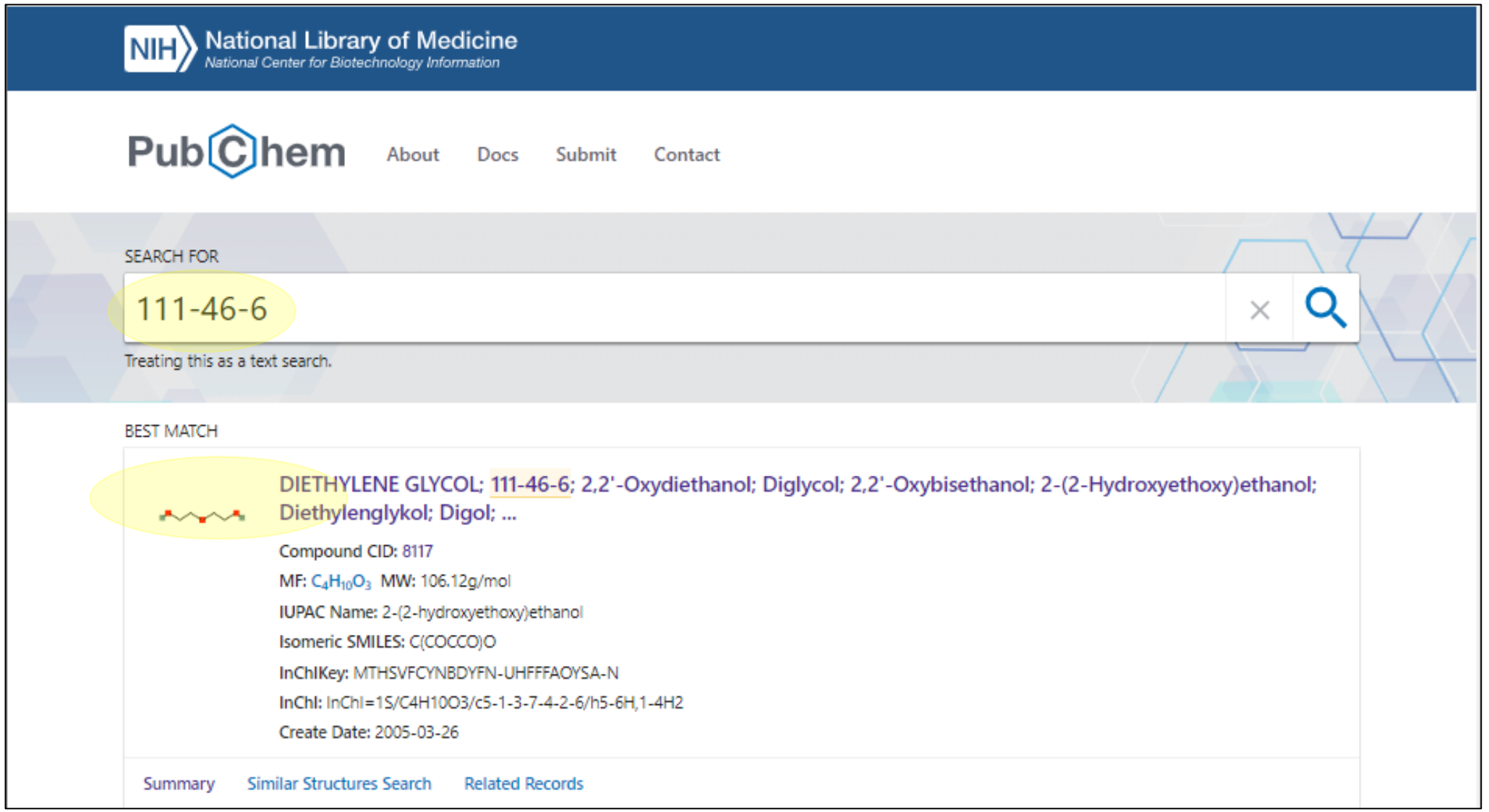
Welle法は拡散係数と予測するために
移行物質(Migrant）のモル体積を使います。

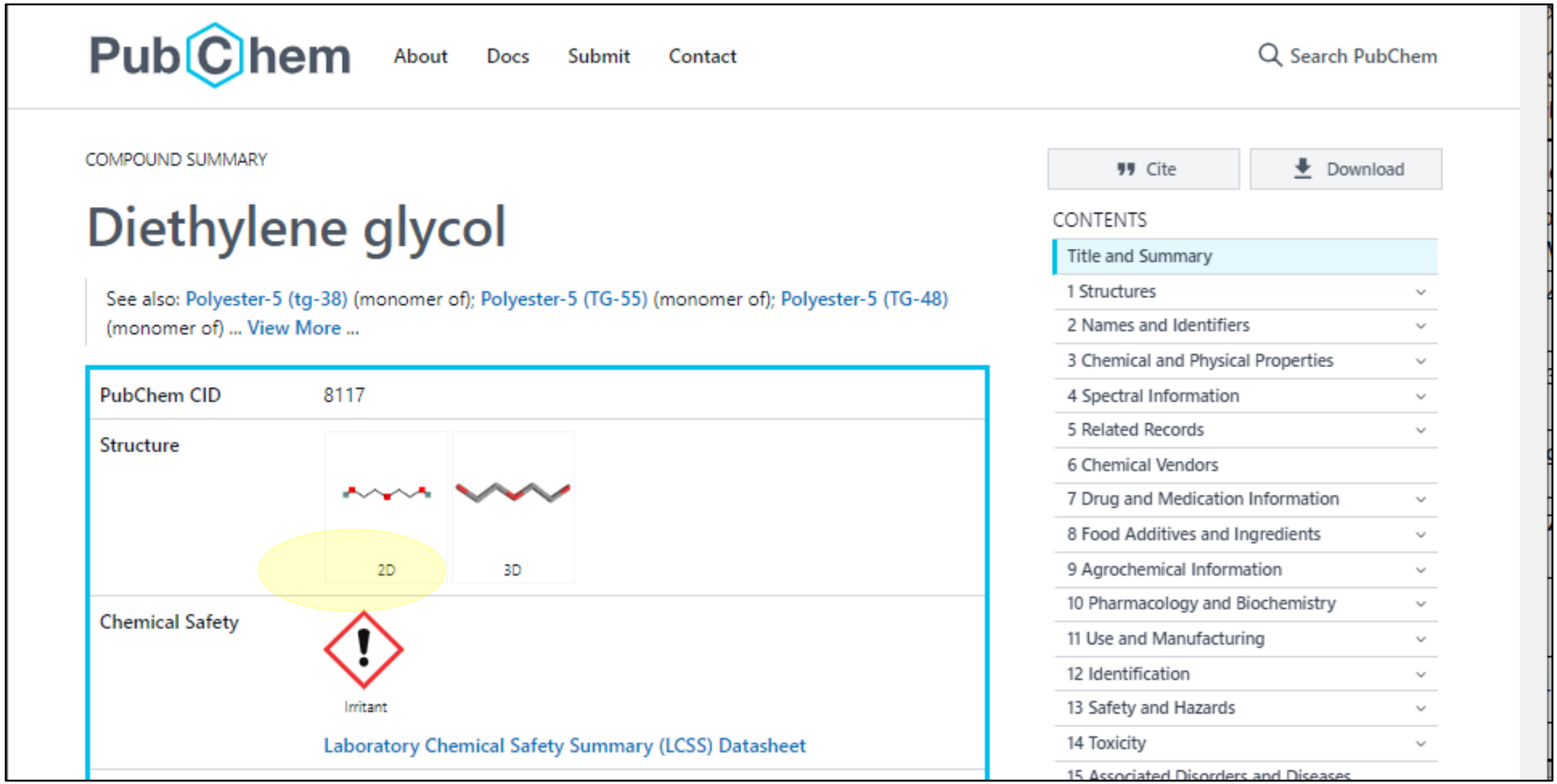
Molecular Volumeを求めるには

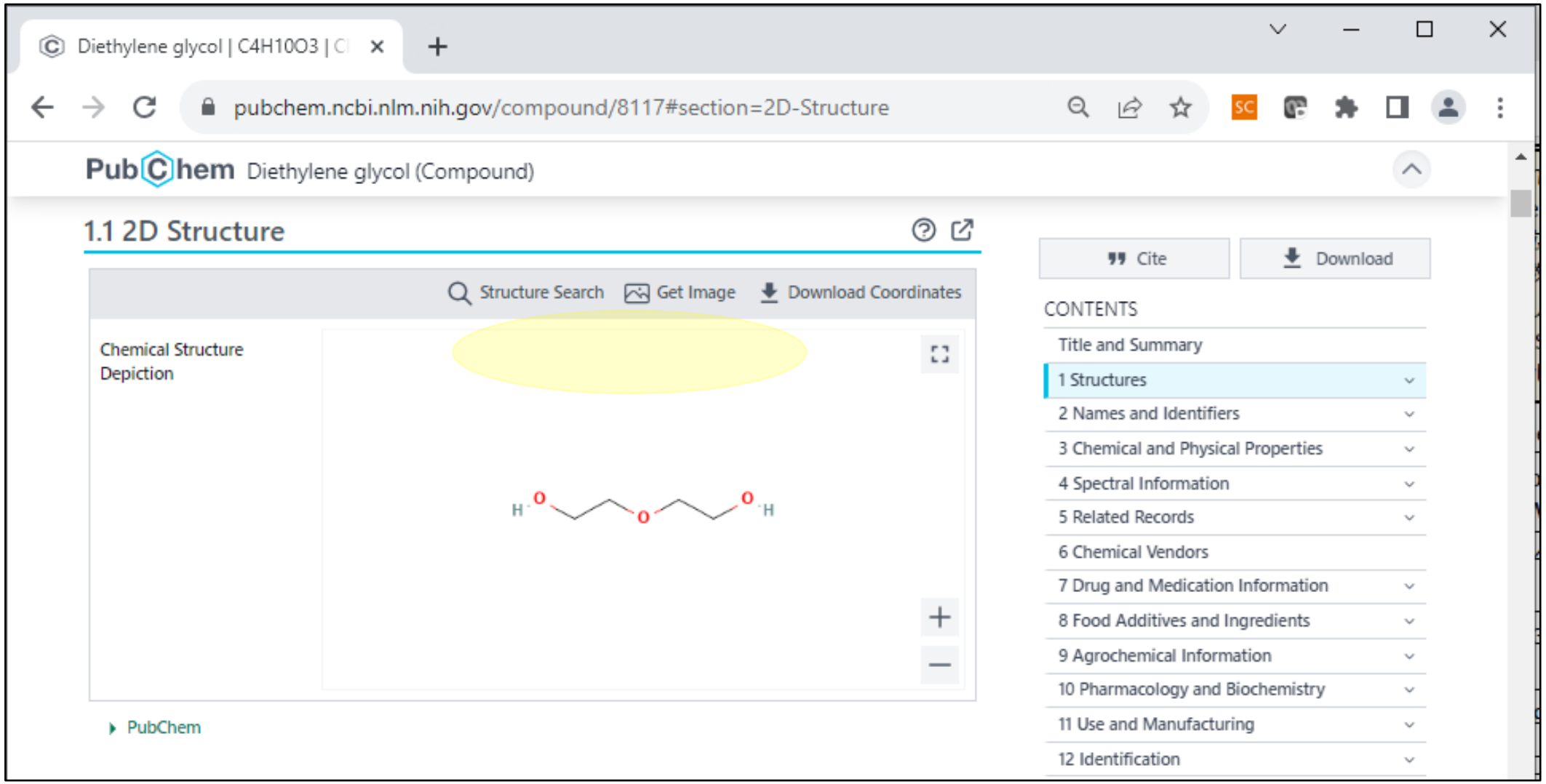
PubChem+**Molinspiration**　の合わせ技
を使うことが一番の近道です。

https://pubchem.ncbi.nlm.nih.gov/

https://www.molinspiration.com/cgi/properties


National Library of Medicine
National Center for Biotechnology Information
PubChem
About
Docs
Submit
Contact
SEARCH FOR
111-46-6
Treating this as a text search.
BEST MATCH
DIETHYLENE GLYCOL; 111-46-6; 2,2'-Oxydiethanol; Diglycol; 2,2'-Oxybisethanol; 2-(2-Hydroxyethoxy)ethanol; Diethylenglykol; Digol; ...
Compound CID: 8117
MF: C4H10O3 MW: 106.12g/mol
IUPAC Name: 2-(2-hydroxyethoxy)ethanol
Isomeric SMILES: C(COCCO)O
InChIKey: MTHSVFCYNBDYFN-UHFFFAOYSA-N
InChI: InChI=1S/C4H10O3/c5-1-3-7-4-2-6/h5-6H,1-4H2
Create Date: 2005-03-26
Summary
Similar Structures Search
Related Records


PubChem
About
Docs
Submit
Contact
Search PubChem
COMPOUND SUMMARY
Diethylene glycol
See also: Polyester-5 (tg-38) (monomer of); Polyester-5 (TG-55) (monomer of); Polyester-5 (TG-48) (monomer of) ... View More ...
PubChem CID
8117
Structure
2D
3D
Chemical Safety
Irritant
Laboratory Chemical Safety Summary (LCSS) Datasheet
Cite
Download
CONTENTS
Title and Summary
1 Structures
2 Names and Identifiers
3 Chemical and Physical Properties
4 Spectral Information
5 Related Records
6 Chemical Vendors
7 Drug and Medication Information
8 Food Additives and Ingredients
9 Agrochemical Information
10 Pharmacology and Biochemistry
11 Use and Manufacturing
12 Identification
13 Safety and Hazards
14 Toxicity
15 Associated Disorders and Diseases


Diethylene glycol | C4H10O3 |
pubchem.ncbi.nlm.nih.gov/compound/8117#section=2D-Structure
PubChem Diethylene glycol (Compound)
1.1 2D Structure
Structure Search
Get Image
Download Coordinates
Chemical Structure Depiction
PubChem
Cite
Download
CONTENTS
Title and Summary
1 Structures
2 Names and Identifiers
3 Chemical and Physical Properties
4 Spectral Information
5 Related Records
6 Chemical Vendors
7 Drug and Medication Information
8 Food Additives and Ingredients
9 Agrochemical Information
10 Pharmacology and Biochemistry
11 Use and Manufacturing
12 Identification

AKTS

Palmetrics

https://www.molinspiration.com/cgi/propertiesを使うには

Welle法を使うにはMigrantの分子量(g/mol)ではなく、モル体積が必要です。

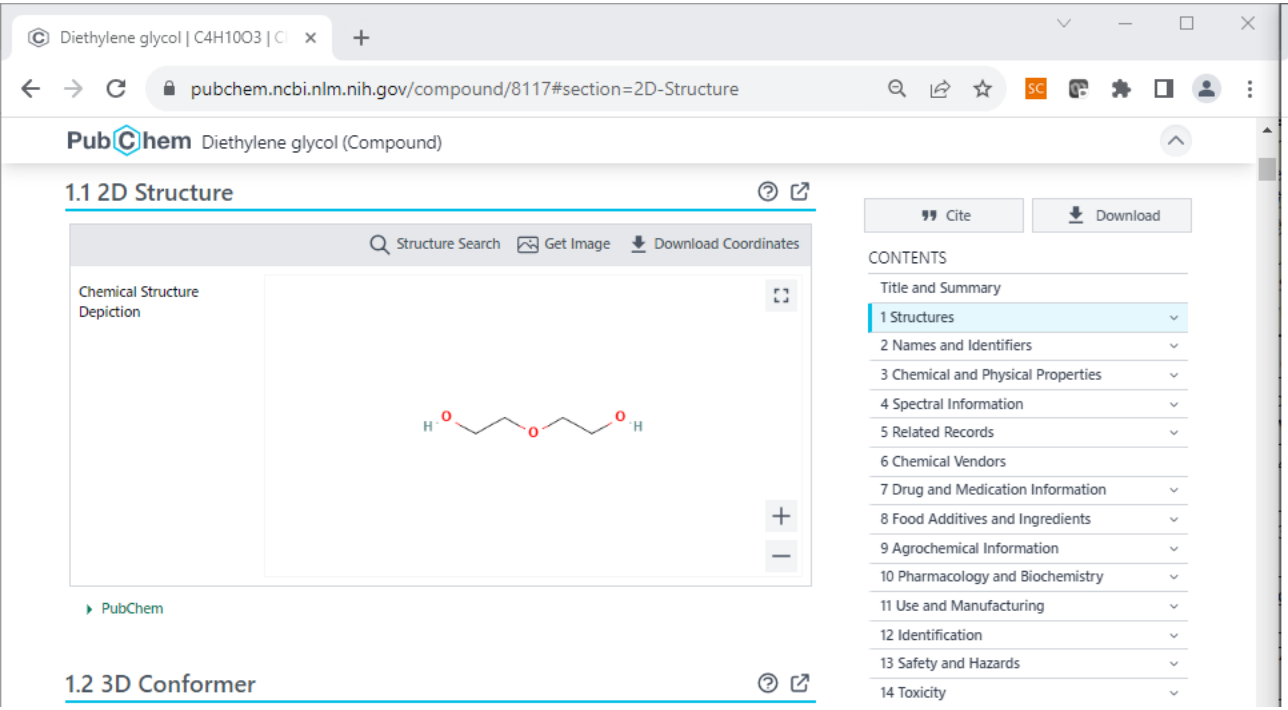
Migrantの化学式をSIMLES表示にして、MolinspirationのSMILES表記を
入力すればモル体積が判明します。（これは2023_07までは可能でした）

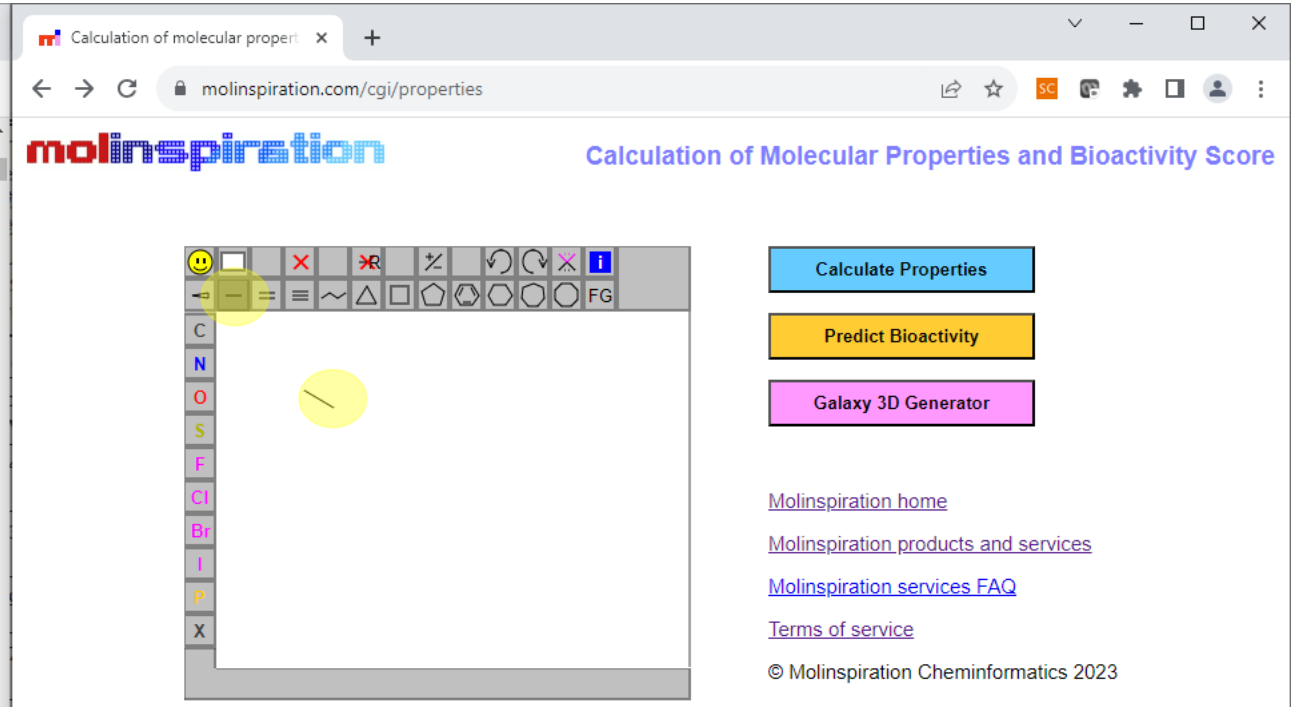
Molinspirationはスロバキアの大学のベンチャーですが、Molinspirationデータベースの評価が
得られた時点で2023_10月からSMILES表記でデータベースを探索することが
できなくなりました。Molinspirationと個別に契約すればSMILES入力は可能です。

ただしMolinspirationの入力欄に化学式を2Dで描写(書き込み)すれば、分子特性が
表示されモル体積情報が得られます。化学式を入力画面で2Dで描写するか、あるいは
SDF file を貼付ける と分子特性情報が得られます。
~~（8～17ページに入力欄に化学式を2Dで書き込みする操作法を紹介します。）~~
18～28ページに入力欄に SDF file を貼付ける操作法を紹介します。
注：ChemDrawの操作に慣れていたとしても、SDFファイルのコピペ操作が簡単です。

2023_11_11
2023_11_26 追加編集

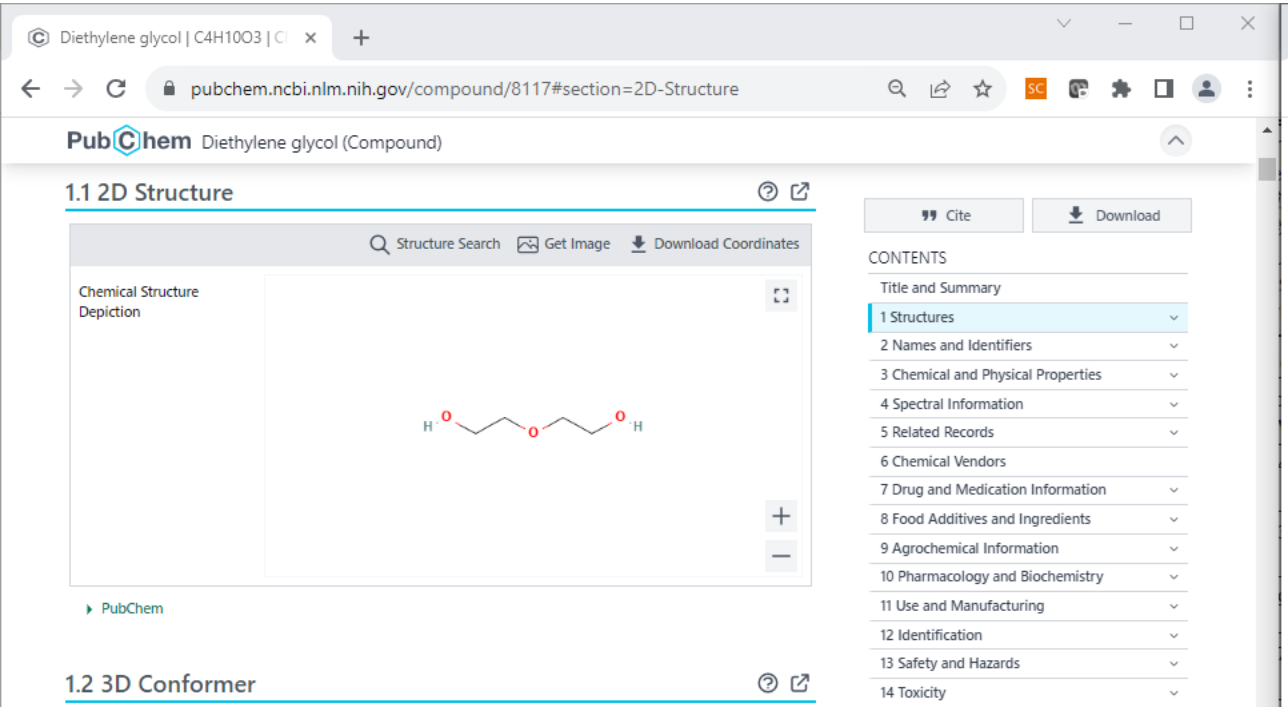
操作法の事例としてジエチレングリコール（CAS.No.111-46-6）で説明しています。
左側がPubChem画面、右がmolinspirationです。

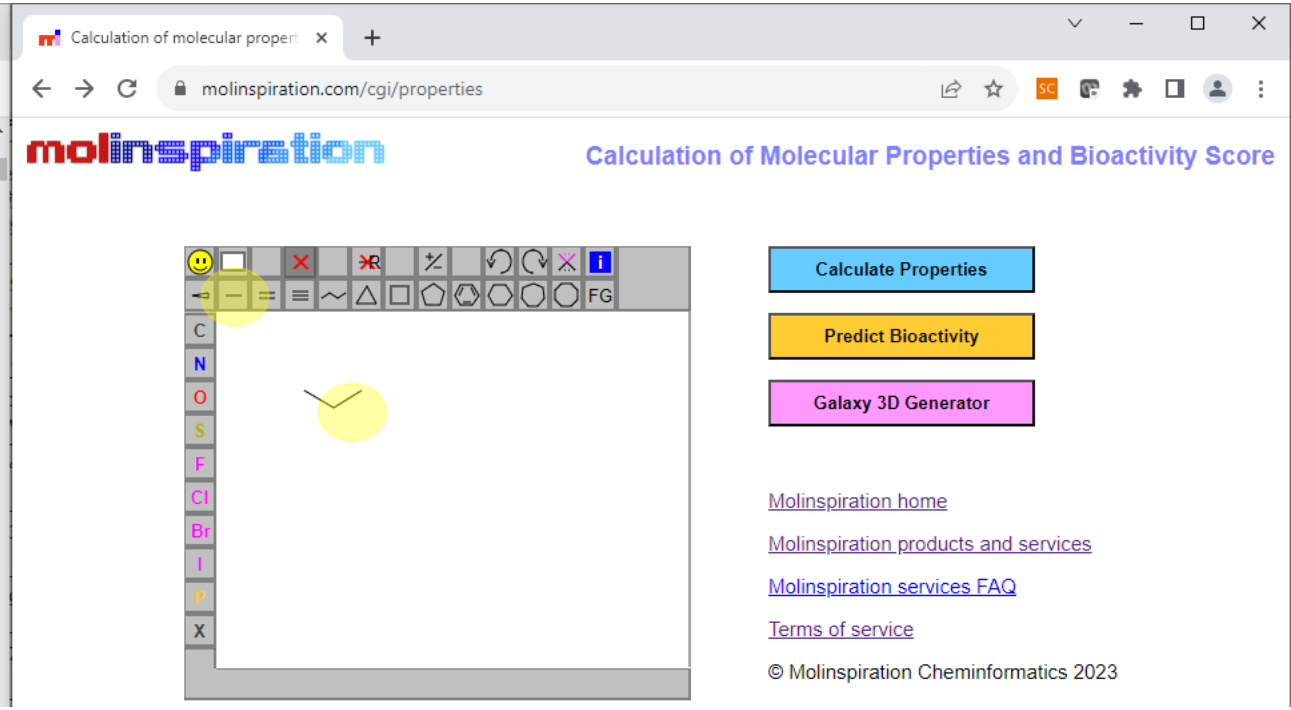




PubChemでジエチレングリコールの”2D_structureを探索します。
次にmolinspirationの入力欄に　2D_structureを１つ１つ書き込みます。

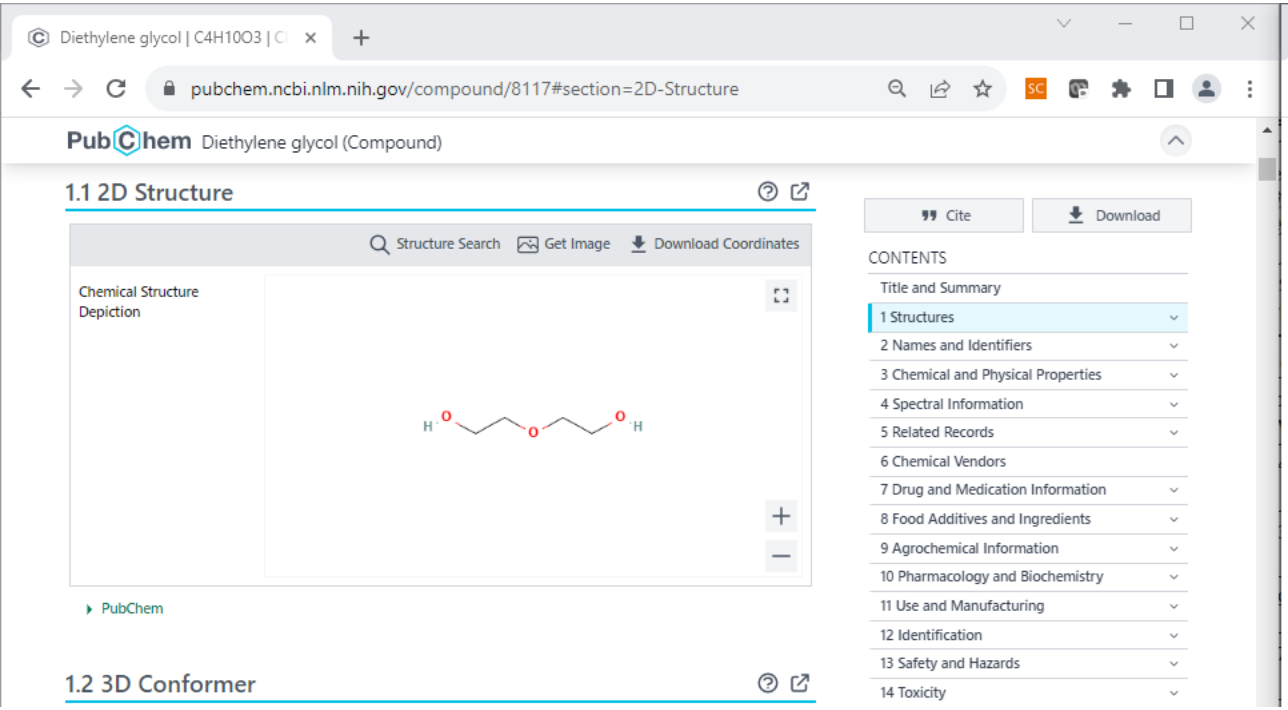
本ページから18ページにスキップしてお読みください。

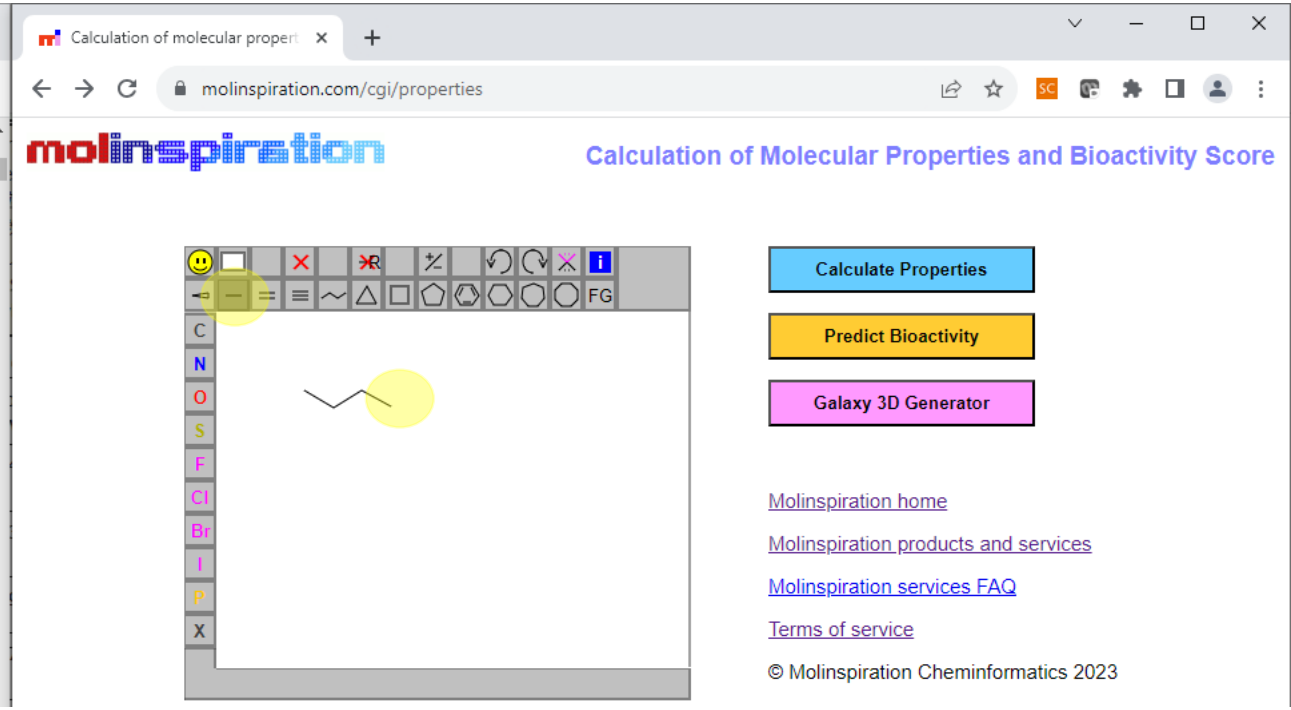




AKTS

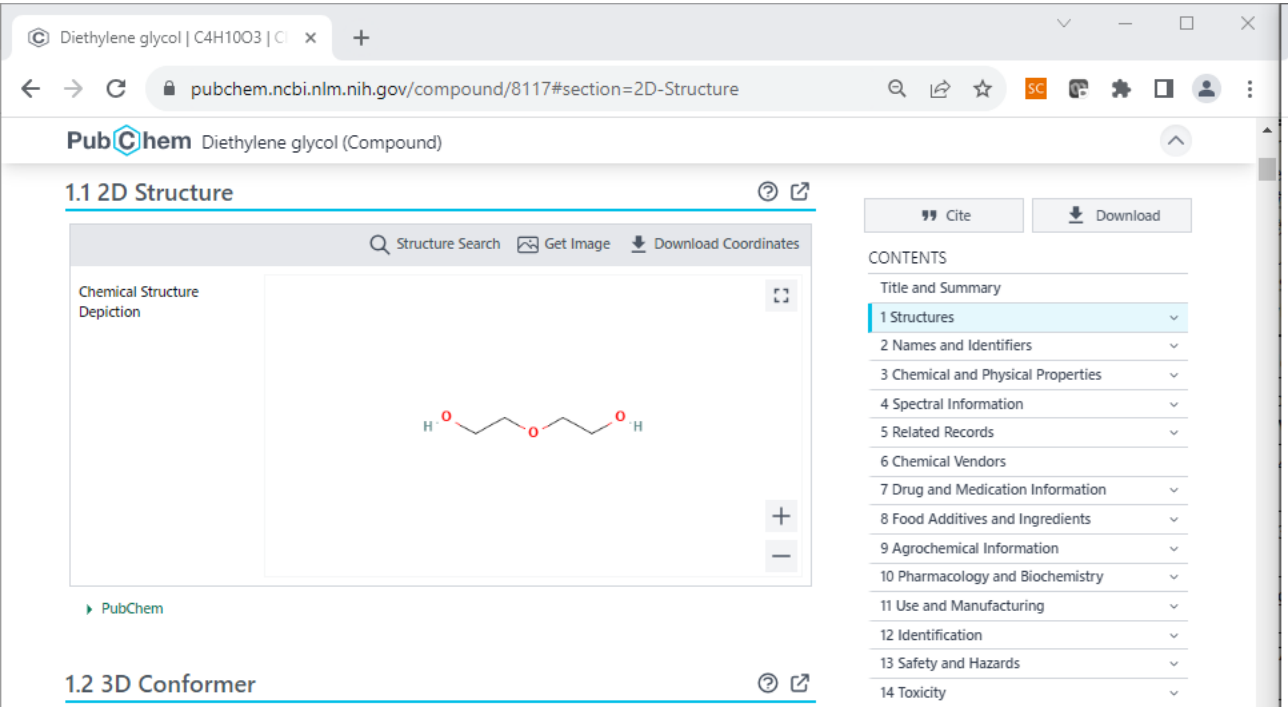
Palmetrics

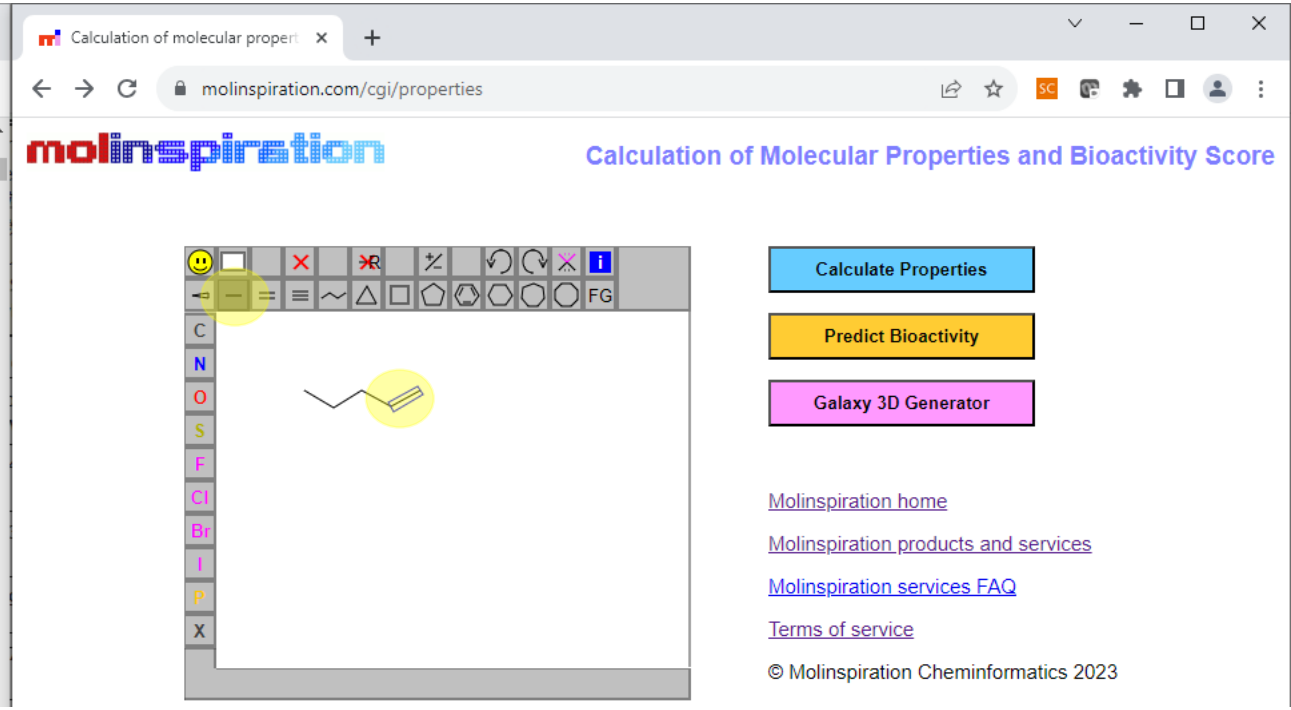




AKTS

Palmetrics





AKTS

Palmetrics

Diethylene glycol | C4H10O3 | C

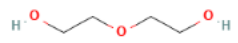
pubchem.ncbi.nlm.nih.gov/compound/8117#section=2D-Structure

PubChem Diethylene glycol (Compound)

1.1 2D Structure

Structure Search | Get Image | Download Coordinates

Chemical Structure Depiction



PubChem

1.2 3D Conformer

Cite | Download

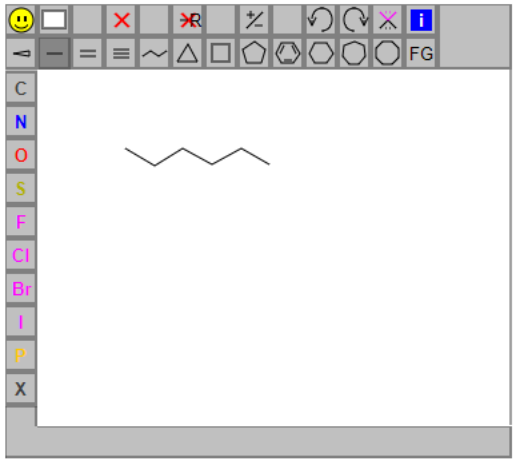
CONTENTS

Calculation of molecular propert

molinspiration.com/cgi/properties

molinspiration

Calculation of Molecular Properties and Bioactivity Score



Calculate Properties

Predict Bioactivity

Galaxy 3D Generator

Molinspiration home

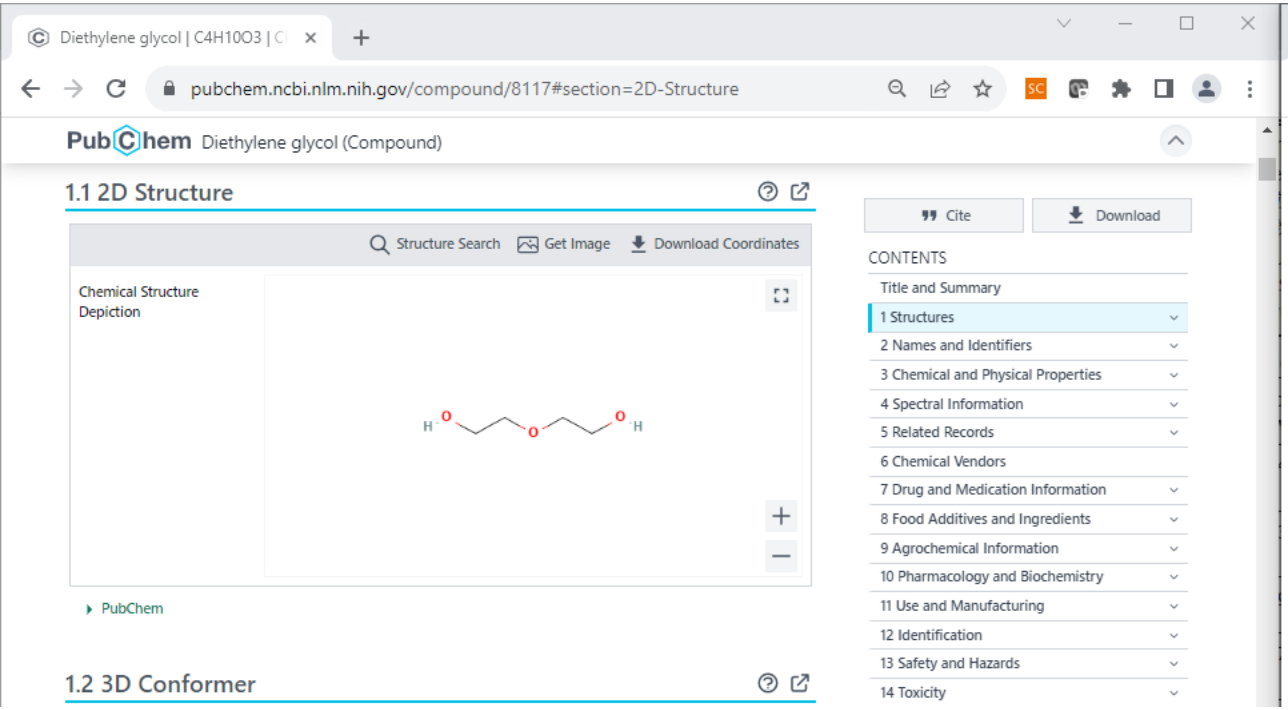
Molinspiration products and services

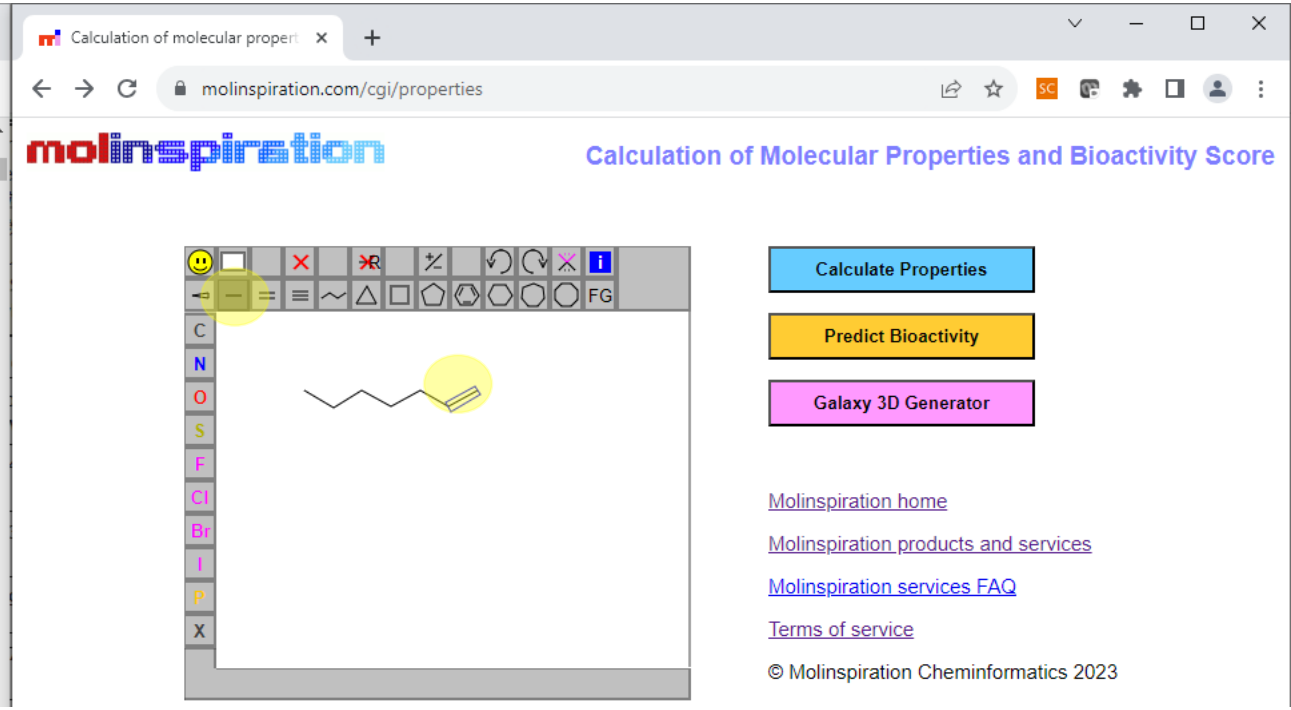
Molinspiration services FAQ

Terms of service

AKTS

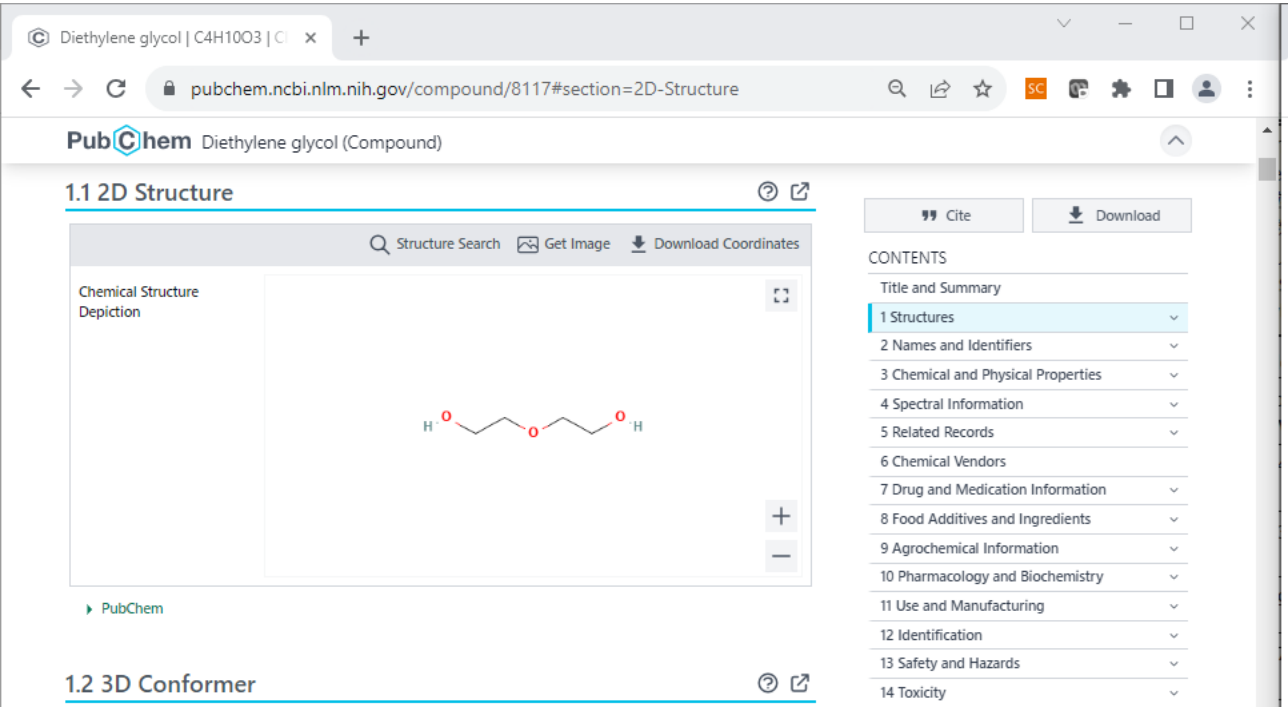
Palmetrics

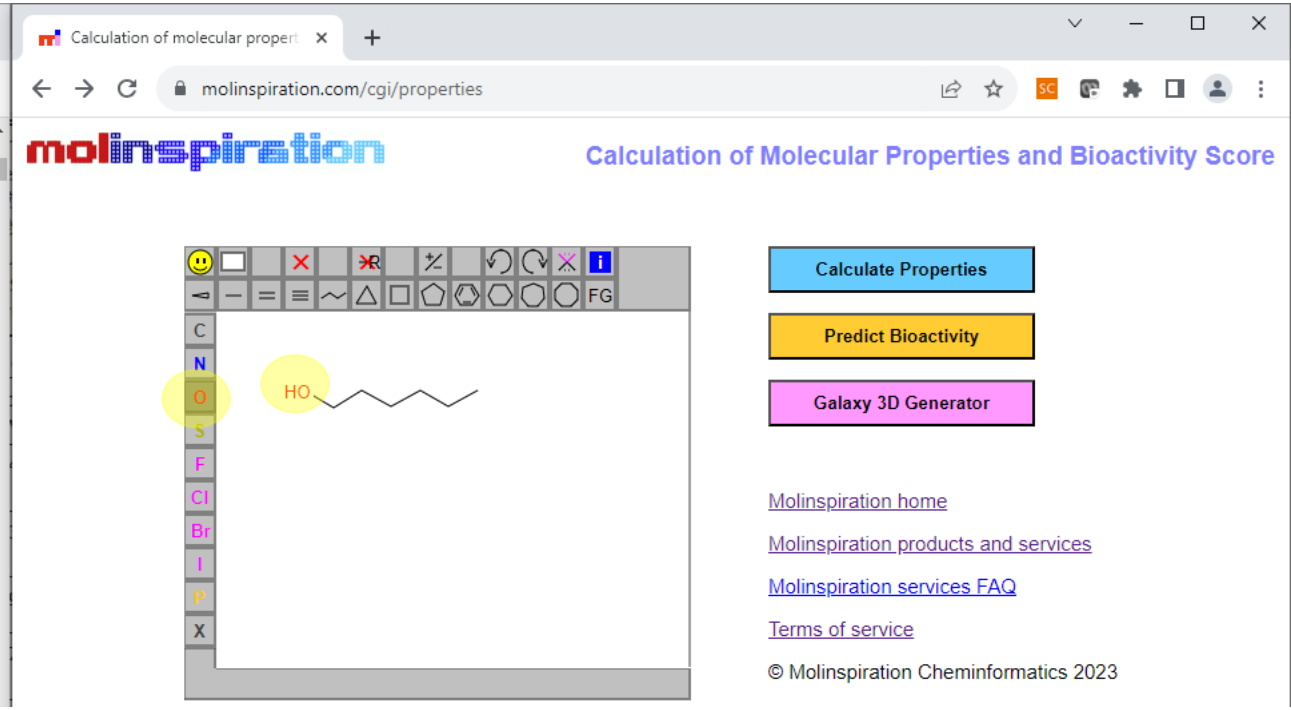




AKTS

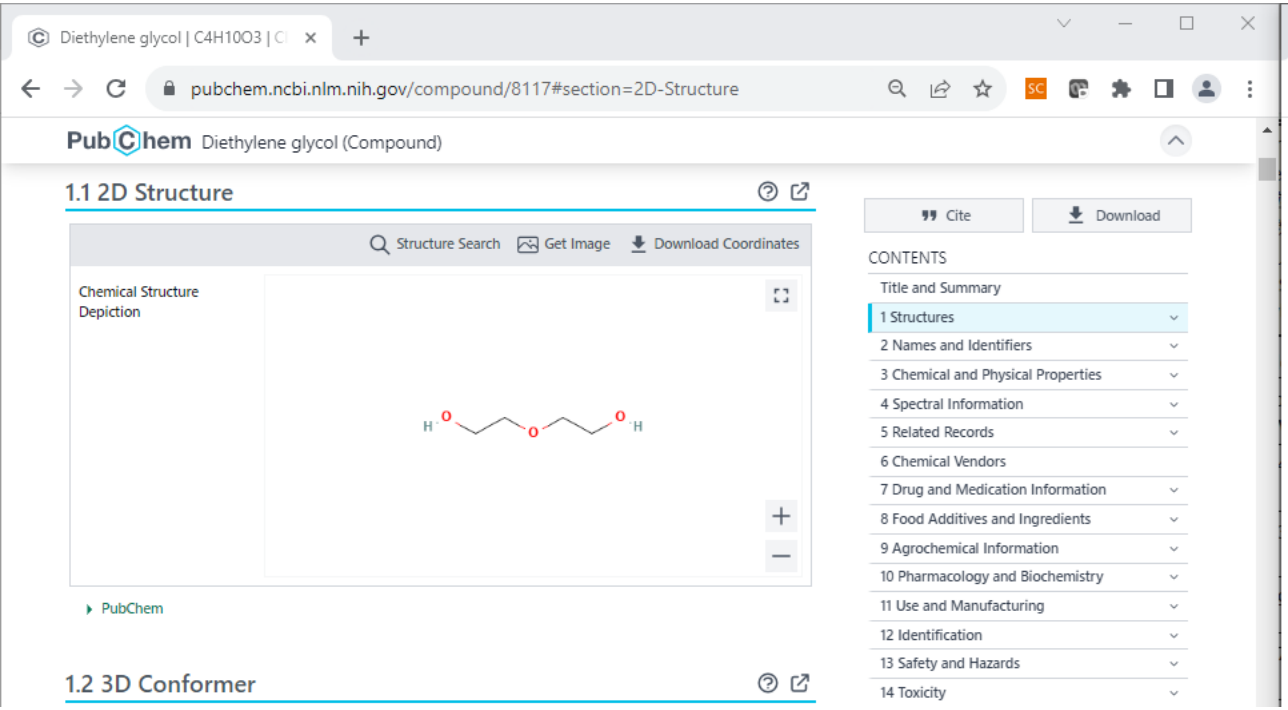
Palmetrics

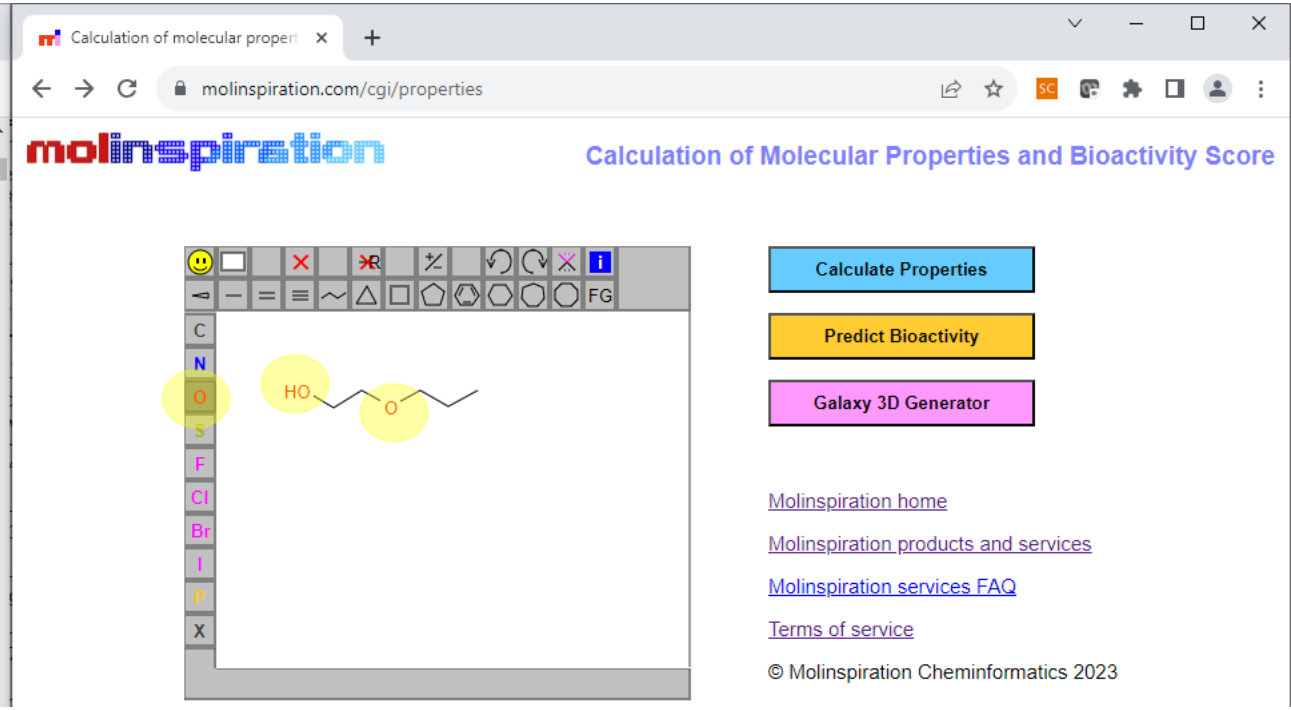




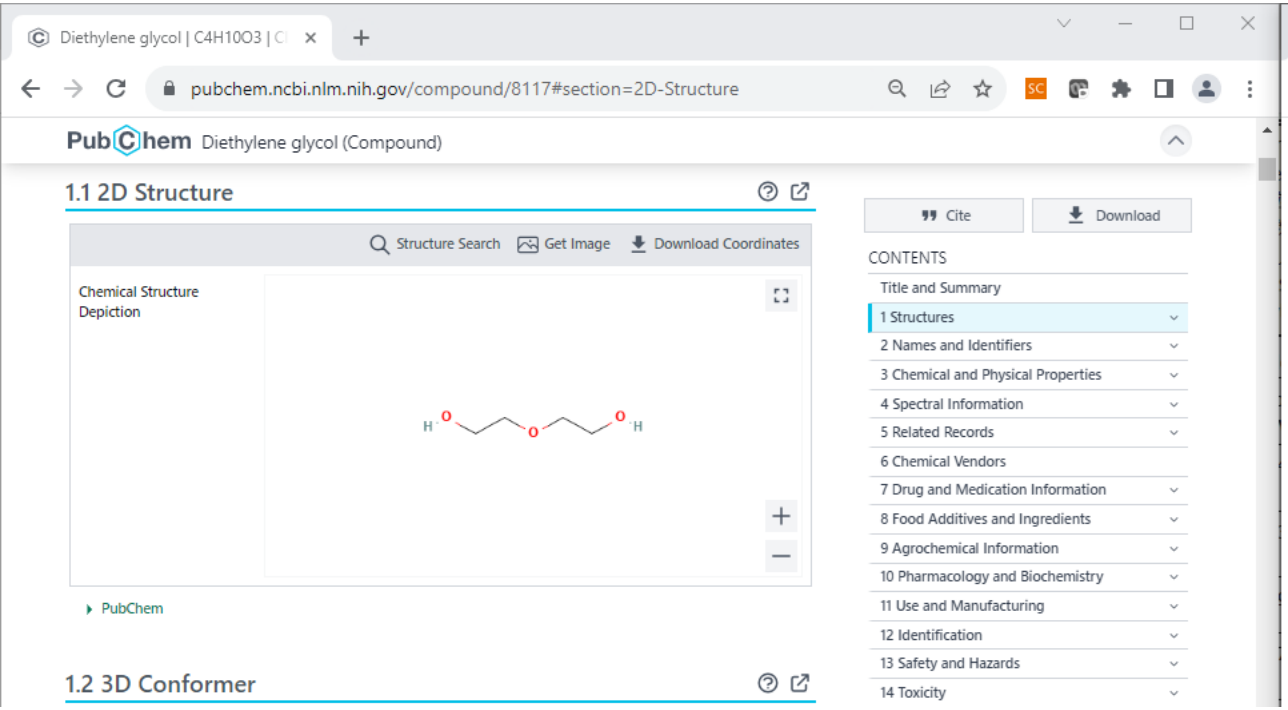
AKTS

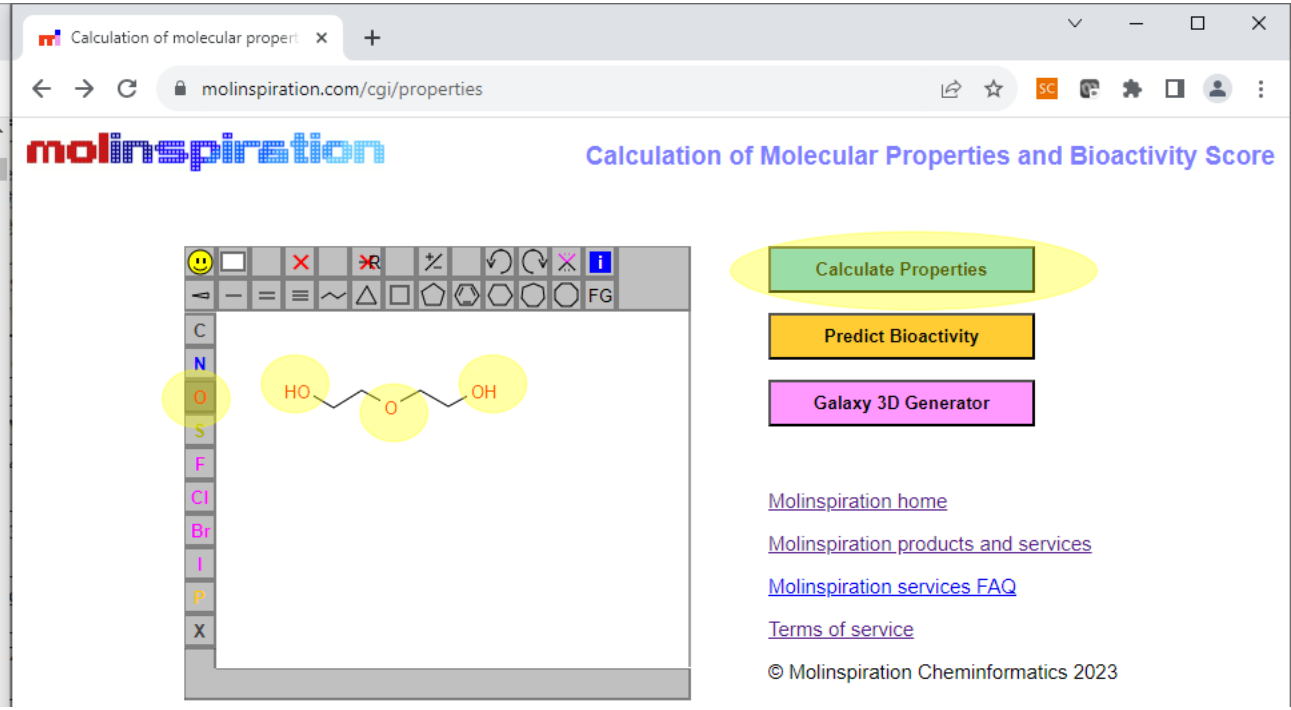
Palmetrics

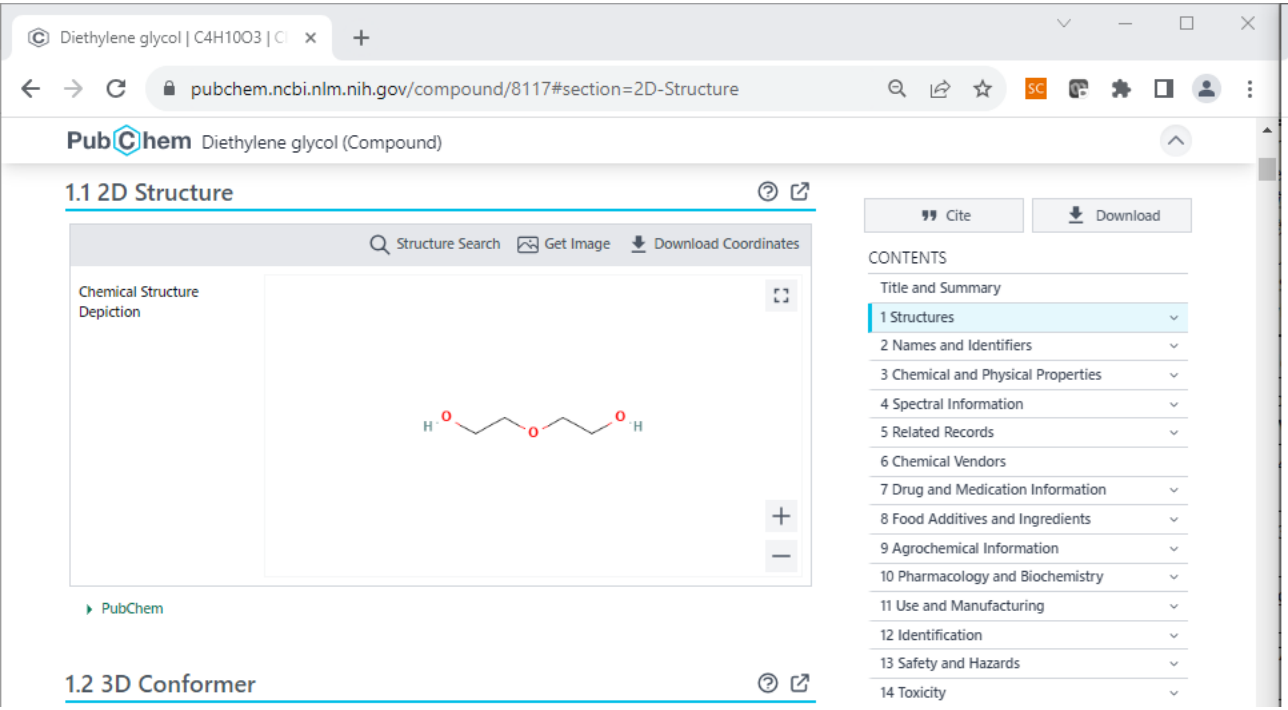


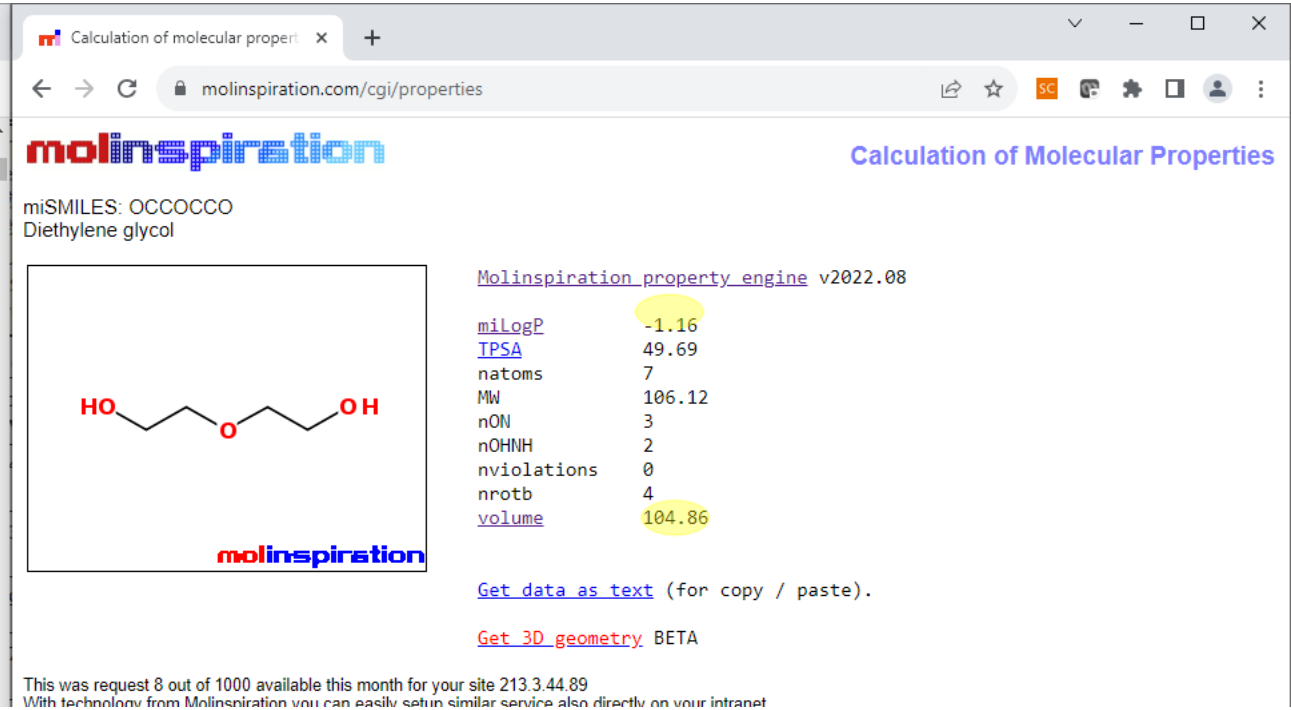


2D_structureの書き込みが完了したら
右欄上端の　Calculate　Properties　をクリックします。



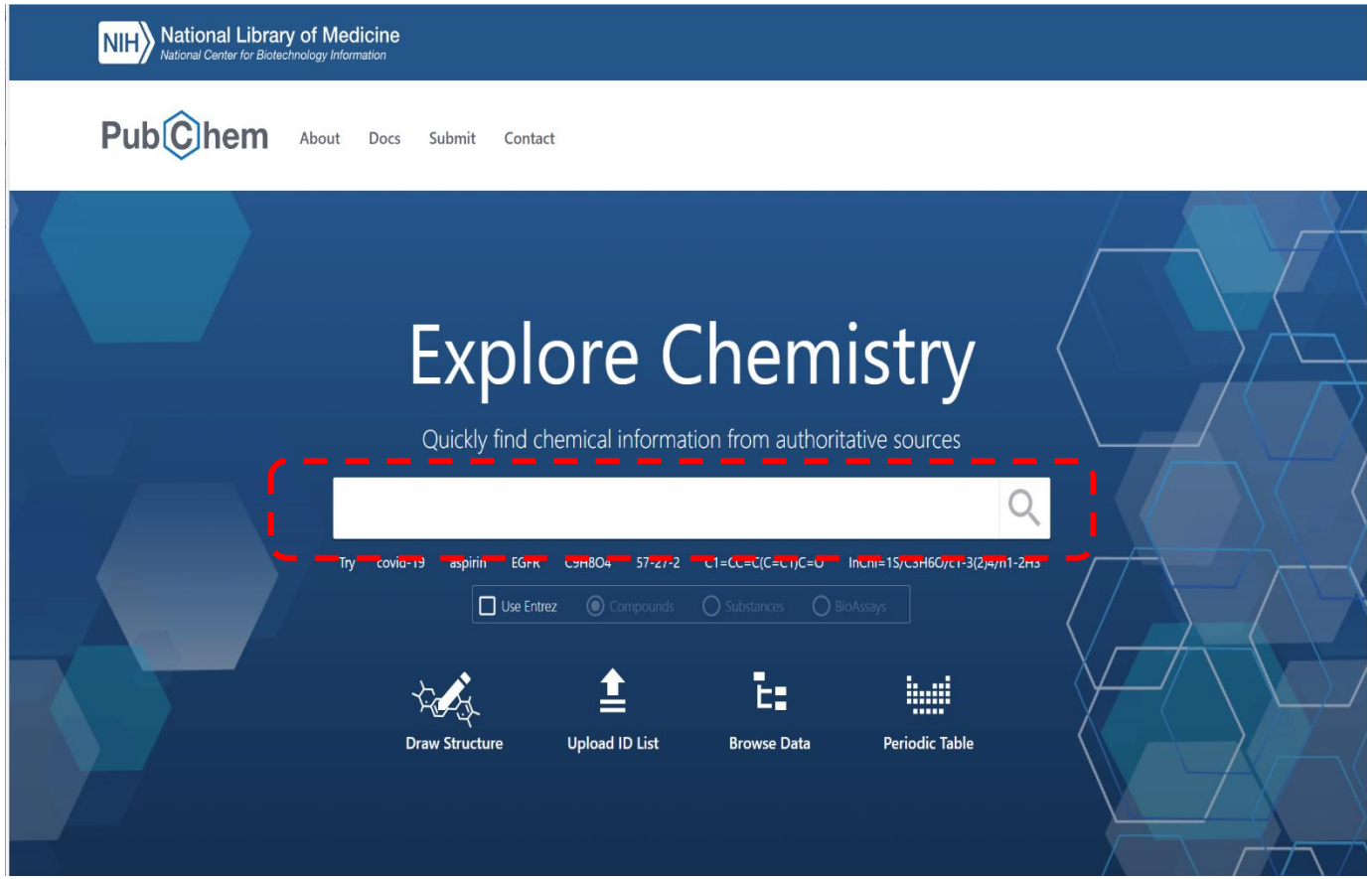






ジエチレン・グリコール　のproperties(2022_08）が　表示されます。
Volume（モル体積）が104.89　mLogP　が-1.16　と表記されています。
これらの情報はSML6のWelle式には不可欠な特性です。

Fig_03:PubChemデータベースを立ち上げた時の初期画面
bikoku



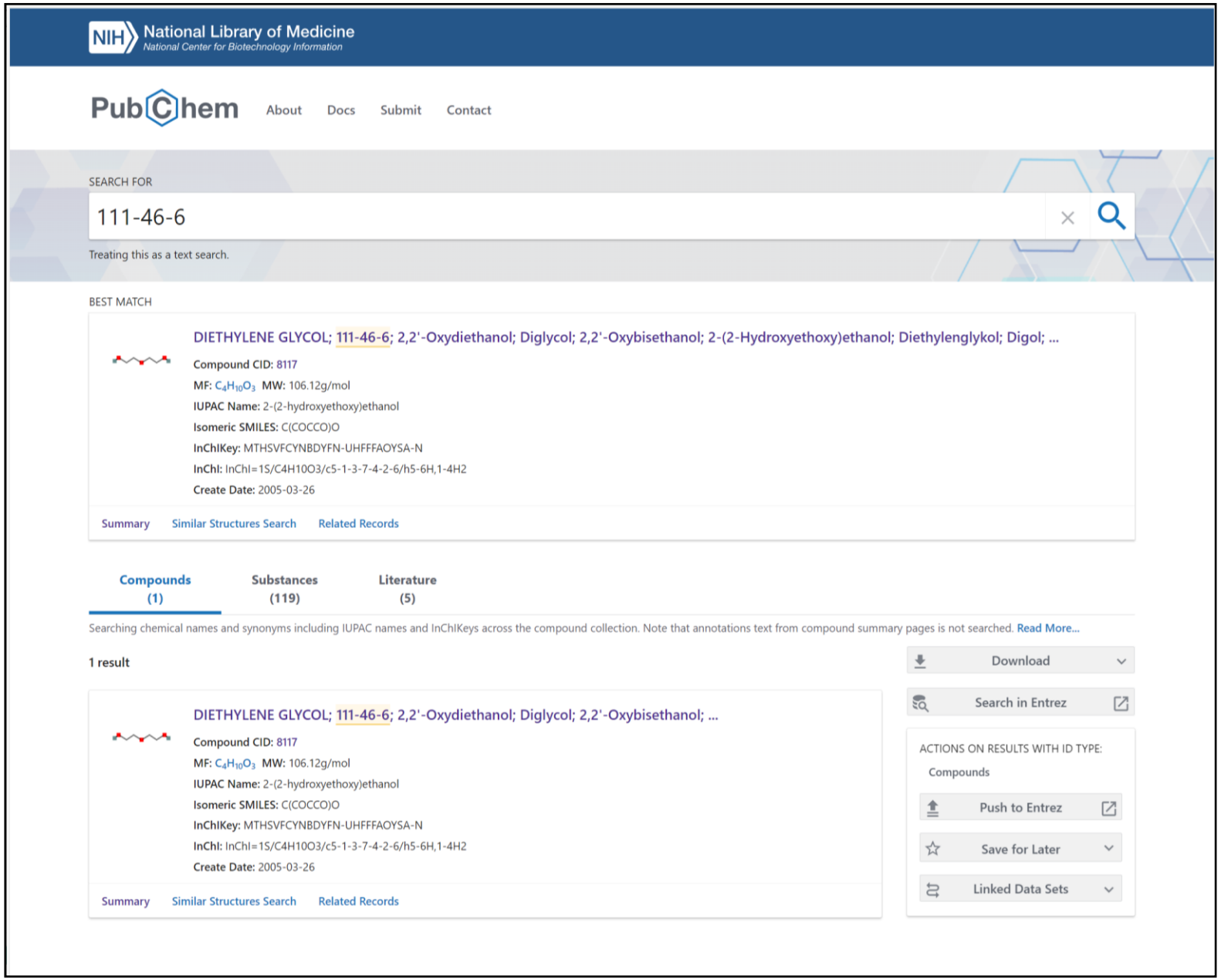
移行物質のジエチレン・グリコールを事例とします。
CAS.No は 111-46-6 です。

モル体積 と Log Powの値を探索します。

最初にPubChem から ジエチレン・グリコール の
SDFファイルをダウンロードする操作の説明をします。

① PubChemを起動するには
https://pubchem.ncbi.nlm.nih.gov を
立ち上げます。
左がPubChemの初期画面です。

② 赤破線枠に**CAS No.の “111-46-6”**
を入力します。


National Library of Medicine
National Center for Biotechnology Information
PubChem
About
Docs
Submit
Contact
SEARCH FOR
111-46-6
Treating this as a text search.
BEST MATCH
DIETHYLENE GLYCOL; 111-46-6; 2,2'-Oxydiethanol; Diglycol; 2,2'-Oxybisethanol; 2-(2-Hydroxyethoxy)ethanol; Diethylenglykol; Digol; ...
Compound CID: 8117
MF: C4H10O3 MW: 106.12g/mol
IUPAC Name: 2-(2-hydroxyethoxy)ethanol
Isomeric SMILES: C(COCCO)O
InChIKey: MTHSVFCYNBDYFN-UHFFFAOYSA-N
InChI: InChI=1S/C4H10O3/c5-1-3-7-4-2-6/h5-6H,1-4H2
Create Date: 2005-03-26
Summary
Similar Structures Search
Related Records
Compounds (1)
Substances (119)
Literature (5)
Searching chemical names and synonyms including IUPAC names and InChIKeys across the compound collection. Note that annotations text from compound summary pages is not searched. Read More...
1 result
Download
Search in Entrez
ACTIONS ON RESULTS WITH ID TYPE:
Compounds
Push to Entrez
Save for Later
Linked Data Sets
DIETHYLENE GLYCOL; 111-46-6; 2,2'-Oxydiethanol; Diglycol; 2,2'-Oxybisethanol; ...
Compound CID: 8117
MF: C4H10O3 MW: 106.12g/mol
IUPAC Name: 2-(2-hydroxyethoxy)ethanol
Isomeric SMILES: C(COCCO)O
InChIKey: MTHSVFCYNBDYFN-UHFFFAOYSA-N
InChI: InChI=1S/C4H10O3/c5-1-3-7-4-2-6/h5-6H,1-4H2
Create Date: 2005-03-26
Summary
Similar Structures Search
Related Records

Compounds (1) | Substances (119) | Literature (5)

Searching chemical names and synonyms including IUPAC names and InChIKeys across the compound collection. Note that annotations text from compound summary pages is not searched. Read More...

1 result

Download

Search in Entrez

DIETHYLENE GLYCOL; 111-46-6; 2,2'-Oxydiethanol; Diglycol; 2,2'-Oxybisethanol; ...

Compound CID: 8117

MF: $C_4H_{10}O_3$ **MW:** 106.12g/mol

IUPAC Name: 2-(2-hydroxyethoxy)ethanol

Isomeric SMILES: C(COCCO)O

InChIKey: MTHSVFCYNBDYFN-UHFFFAOYSA-N

InChI: InChI=1S/C4H10O3/c5-1-3-7-4-2-6/h5-6H,1-4H2

Create Date: 2005-03-26

Summary | Similar Structures Search | Related Records

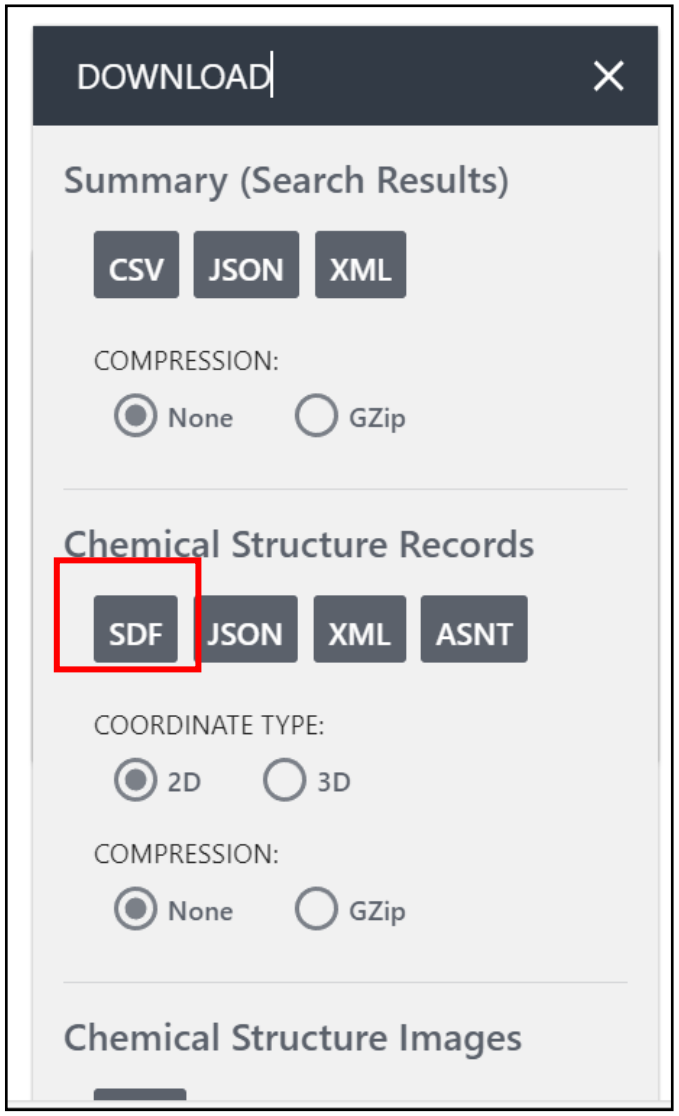
ACTIONS ON RESULTS WITH ID TYPE:

Compounds

Push to Entrez

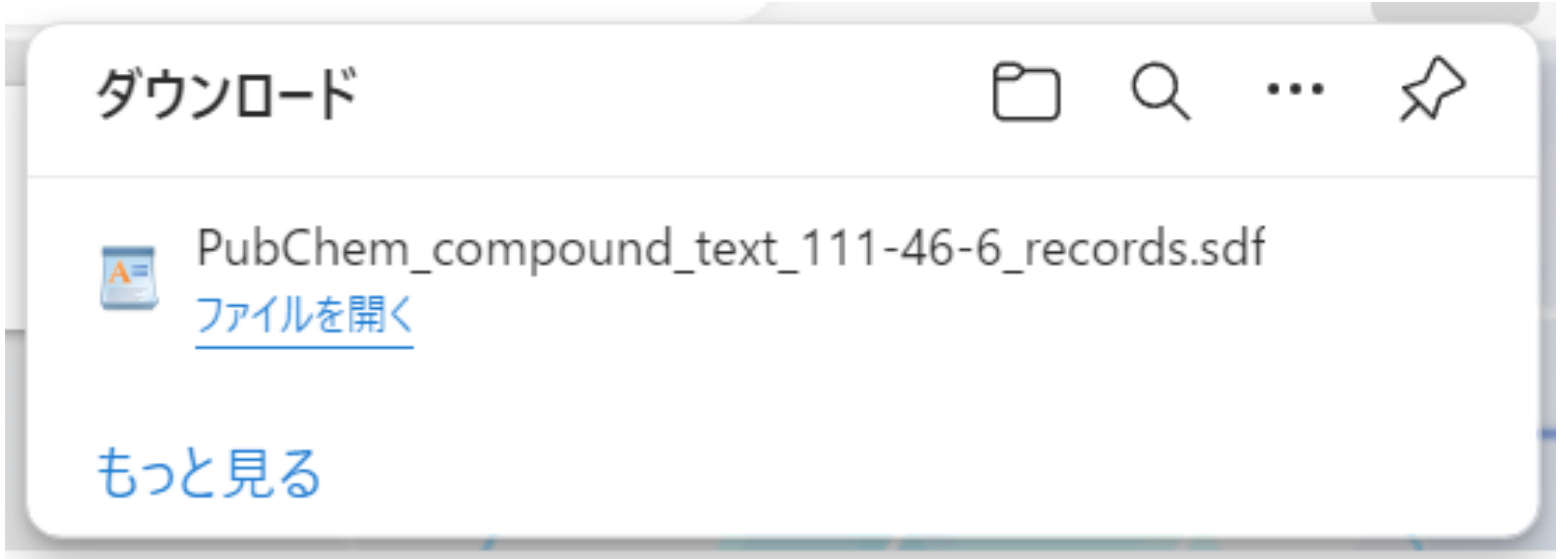
Save for Later

Linked Data Sets



④ 目標とするDownloadはSDF(Structure Design File)です。
赤線枠のSDF をクリックします。

⑤ Download が開始されます。
PubChemデータベースからSDFファイルを入手することができました。
なおファイルの属性はTXTですので、TXTファイルとして取り込みます。



PubChem_compound_text_111-46-6_records - メモ帳

ファイル(F) 編集(E) 書式(O) 表示(V) ヘルプ(H)

```
8117
  -OEChem-11152320422D

 17 16  0     0  0  0  0  0  0  0999 V2000
    5.1350   -0.2500    0.0000 O   0  0  0  0  0  0  0  0  0  0  0  0
    7.7331    0.2500    0.0000 O   0  0  0  0  0  0  0  0  0  0  0  0
    2.5369    0.2500    0.0000 O   0  0  0  0  0  0  0  0  0  0  0  0
    6.0010    0.2500    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    4.2690    0.2500    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    6.8671   -0.2500    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    3.4030   -0.2500    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    6.3996    0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    5.6025    0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    4.6675    0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    3.8705    0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    6.4685   -0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    7.2656   -0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    3.0044   -0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    3.8015   -0.7249    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    8.2700   -0.0600    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
    2.0000   -0.0600    0.0000 H   0  0  0  0  0  0  0  0  0  0  0  0
  1  4  1  0  0  0  0
  1  5  1  0  0  0  0
  2  6  1  0  0  0  0
  2 16  1  0  0  0  0
  3  7  1  0  0  0  0
  3 17  1  0  0  0  0
  4  6  1  0  0  0  0
  4  8  1  0  0  0  0
  4  9  1  0  0  0  0
  5  7  1  0  0  0  0
  5 10  1  0  0  0  0
  5 11  1  0  0  0  0
  6 12  1  0  0  0  0
  6 13  1  0  0  0  0
  7 14  1  0  0  0  0
  7 15  1  0  0  0  0
M  END
> <PUBCHEM_COMPOUND_CID>
8117
```

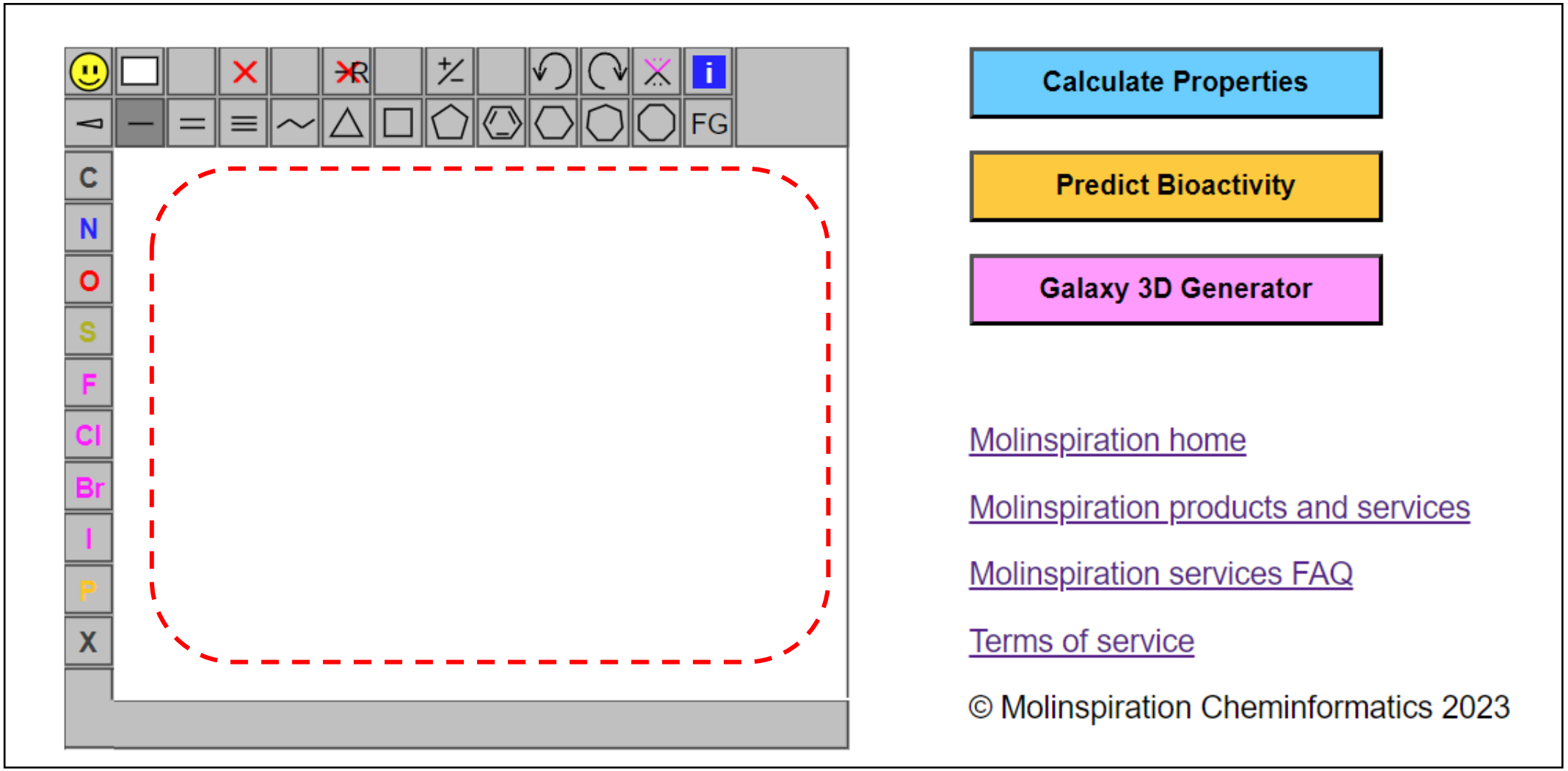
⑥ DownloadされたファイルはTXTファイルです。
ワードパッドあるいはメモ帳で表示させます。
左図はメモ帳の表示例です。
ここで表示させる意味はないのですが、111-46-6の
化学構造がテキストファイルの内容を確認しています。

SDFファイル名はCAS.Noが含まれており識別が容易です。

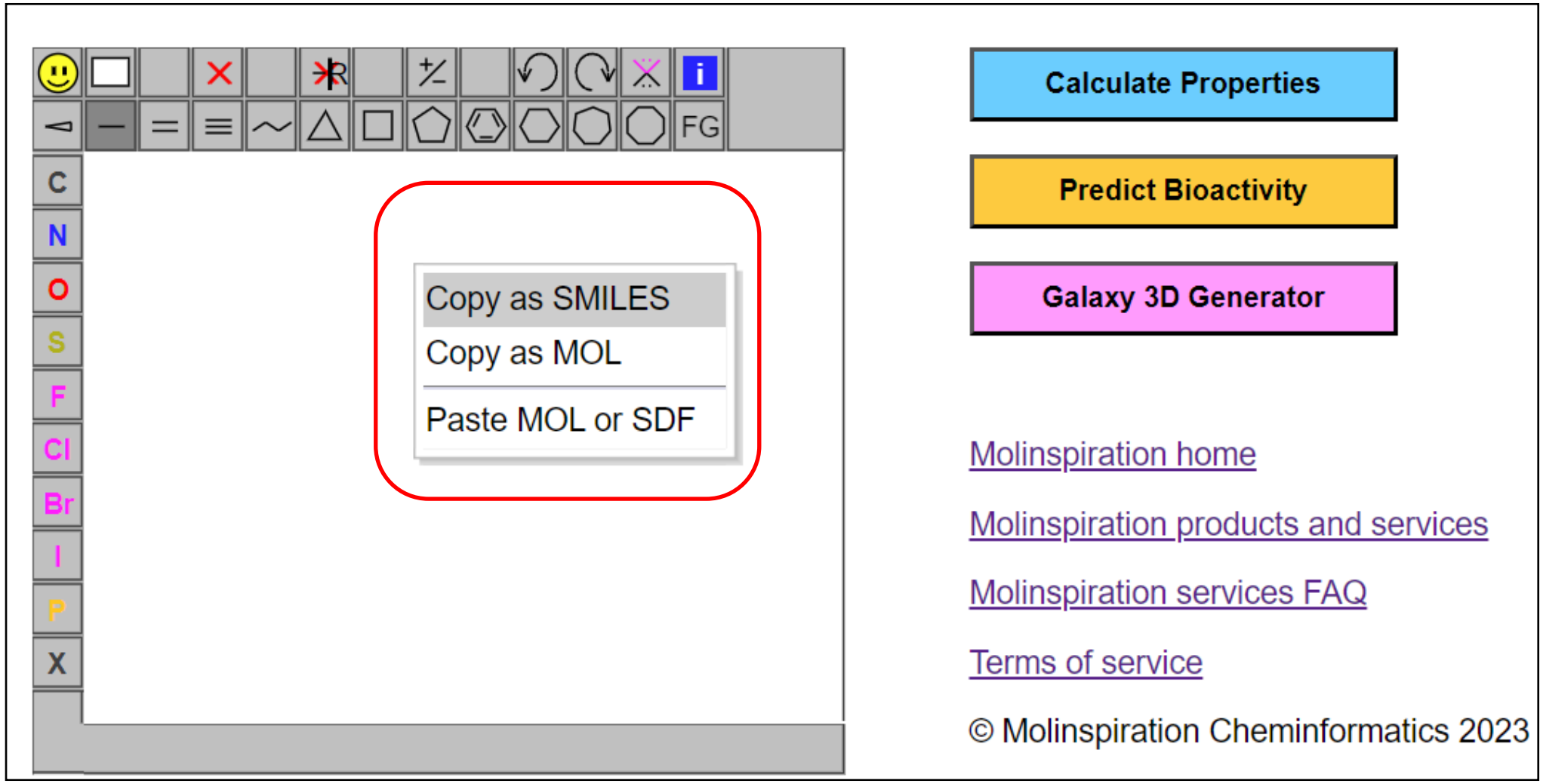
PubChem_compound_text_111-46-6_records.sdf

⑦ 次にmolinspirationを起動します。
molinspirationのURLは
https://molinspiration.com/cgi-bin/properties

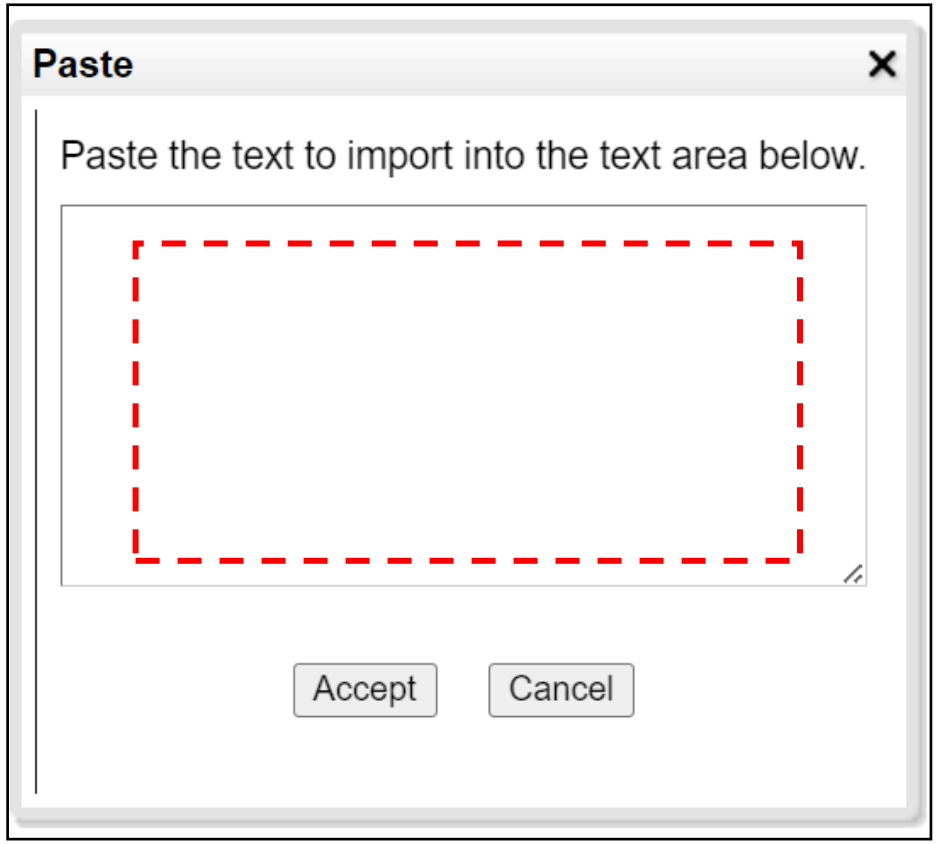
下図は**molinspiration**の起動後の画面です。
⑧ここで赤破線枠内にカーソルを置いて**Right_Click**　します。

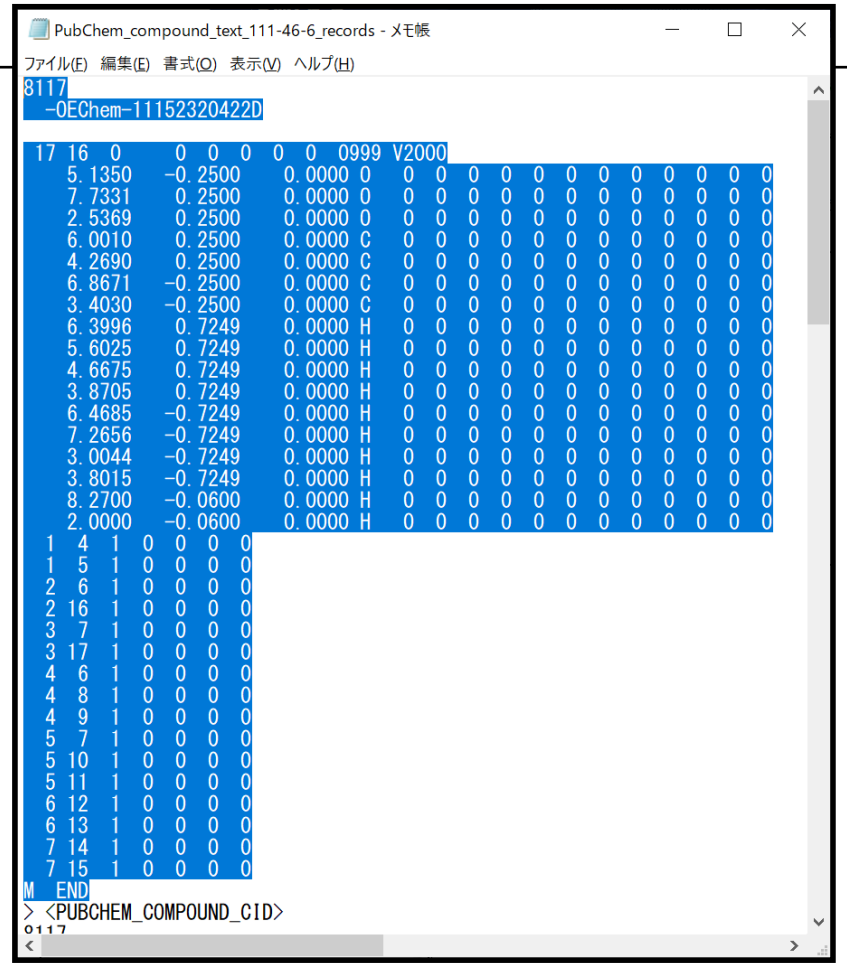


⑧ここで赤破線枠内にカーソルを置いて**Right_Click** すると 赤線枠が表示されます。
⑨ Paste MOL or SDF を選択します。

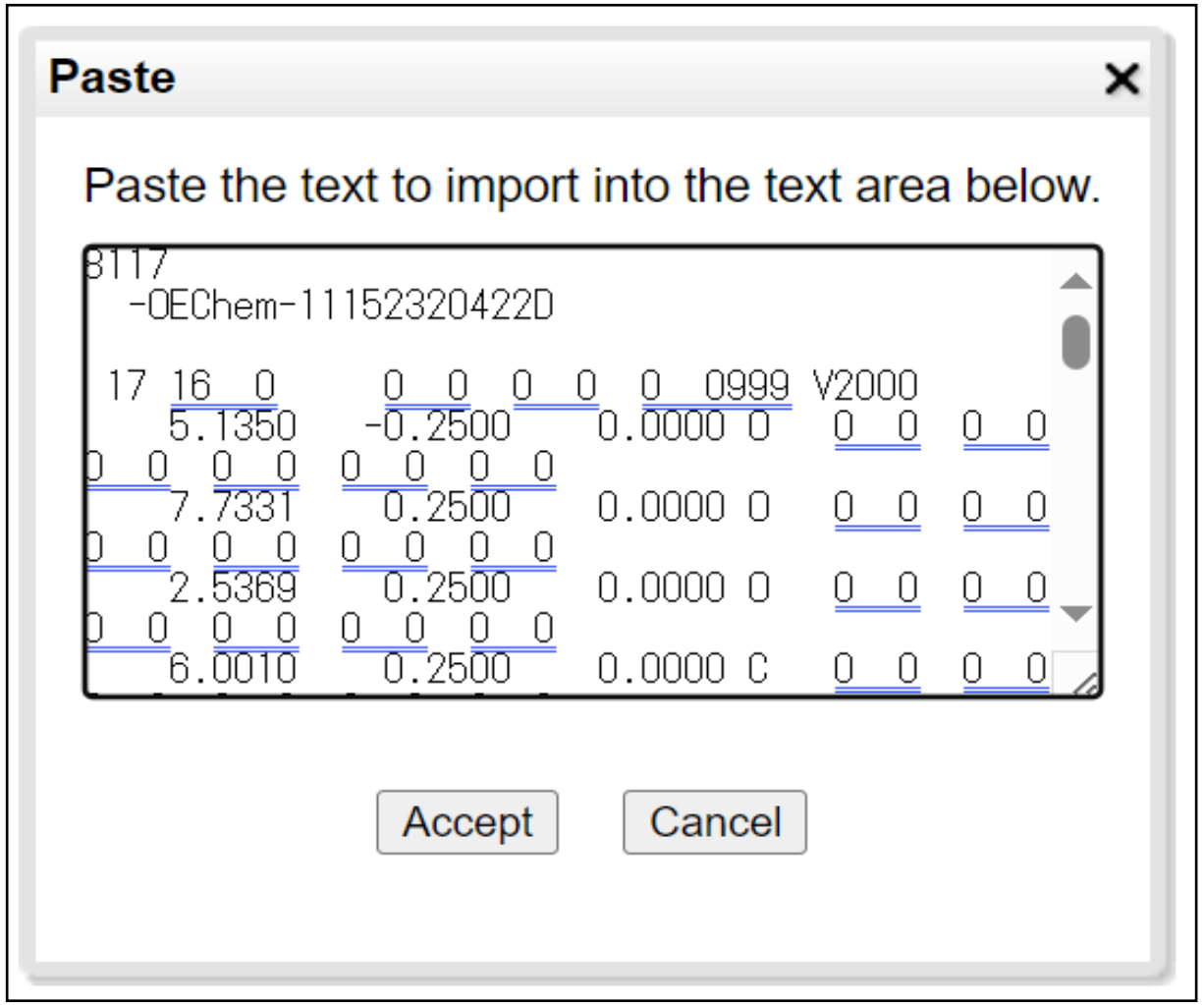


⑩ Paste MOL or SDF を Clickします。左図が表示されるので赤破線枠内にSDFファイルをPasteします。
ただしPasteするときはFig_05のテキスト表示した状態でテキストファイルをペーストすることが必要です。
SDFファイルをコピーしてからペーストするのではなく、テキストをベタコピーしてからペーストすることが必要です。
Copy＆Pasteしたら、 左・下図のAccept をclickします。

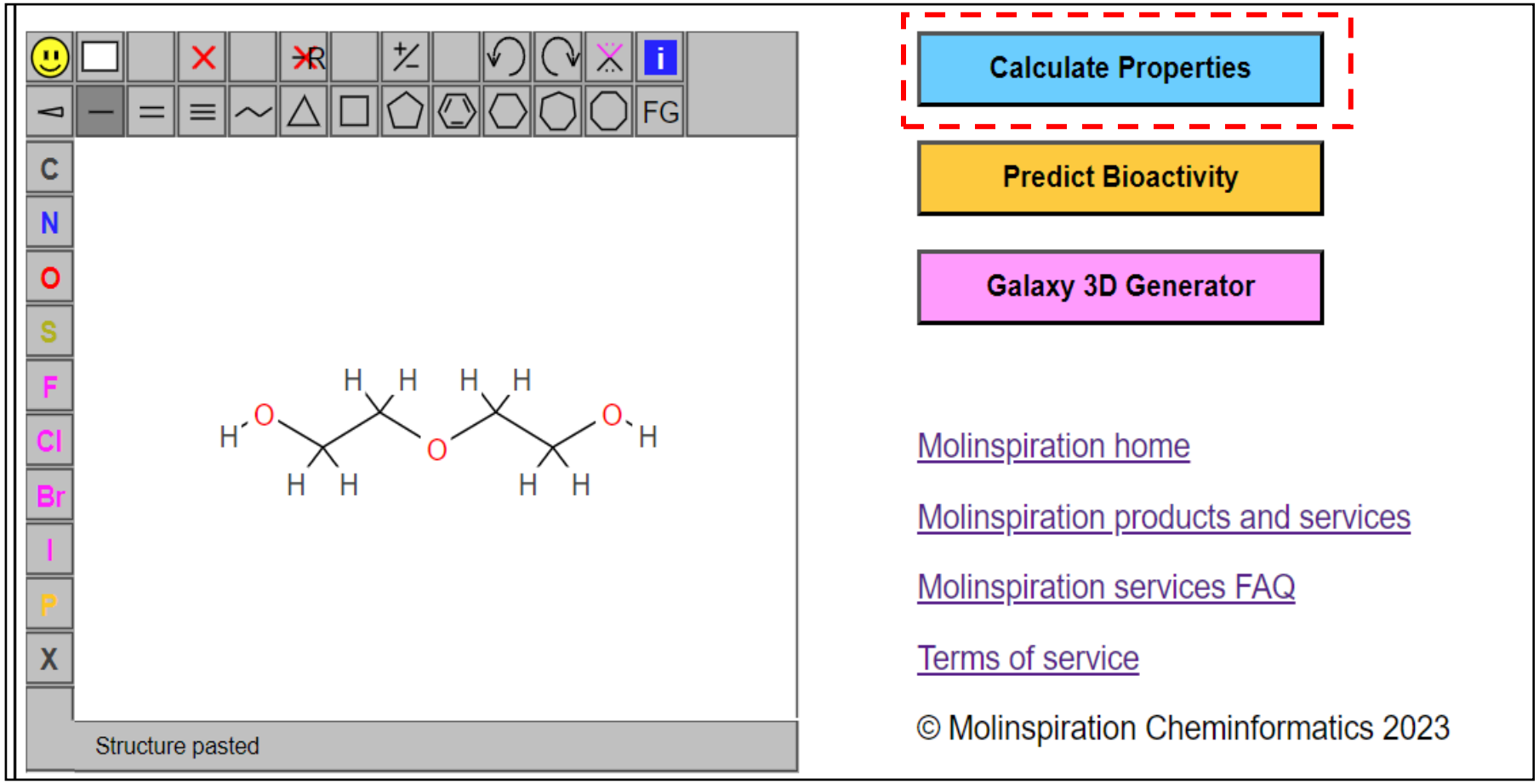




⑪　Copy＆Pasteしたら、Accept　をclickします。

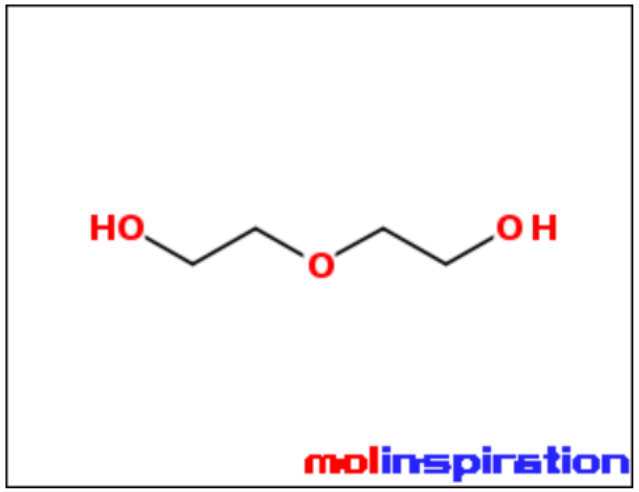


SDFファイルをテキストファイルで表示してから　Copy & Paste して、Acceptをクリックすると操作完了です。
ジエチレン・グリコールの化学構造を　ChemDrawのように操作をしなくても、2Dno化学構造図が入力完了！
化学特性を探索するための入力作業が完了しました。
最後に　⑫赤破線枠のCalculate　Properties　をクリックします。



molinspiration

miSMILES: [H]OC([H])([H])C([H])([H])OC([H])([H])C([H])([H])O[H]
Diethylene glycol



Molinspiration property engine v2022.08

miLogP	-1.16
TPSA	49.69
natoms	7
MW	106.12
nON	3
nOHNH	2
nviolations	0
nrotb	4
volume	104.86

Get data as text (for copy / paste).

Get 3D geometry BETA

This was request 2 out of 1000 available this month for your site 123.217.197.242
With technology from Molinspiration you can easily setup similar service also directly on your intranet.
Comments or questions ? See our FAQ and do not hesitate to provide feedback or contact us by email !

New molecule Predict bioactivity About properties Molinspiration home

SML6のWelle式アプローチでは
下記2項目が最も必要とする物性値です。

Log_Pow： -1.16
モル体積： 104.86