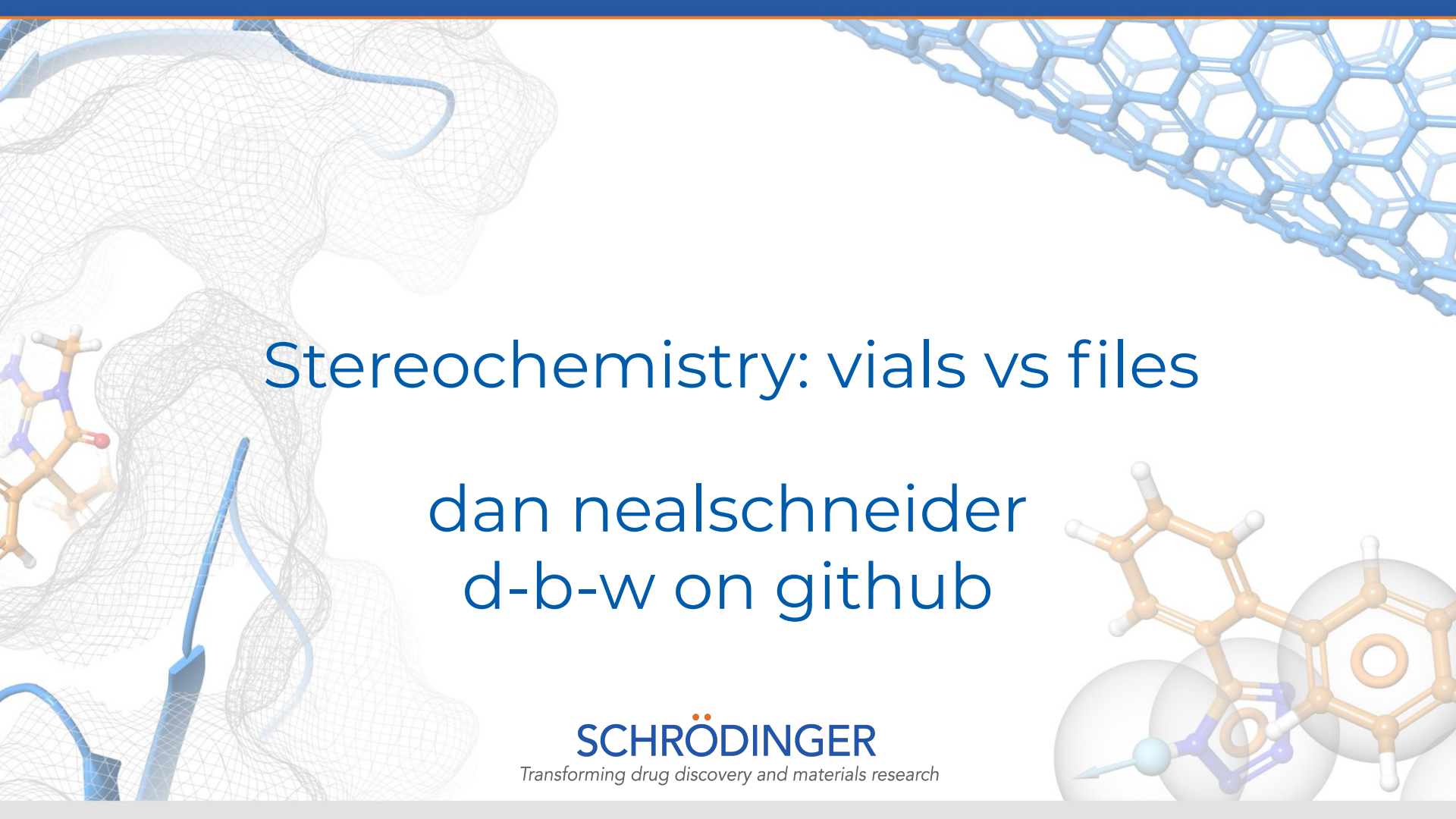
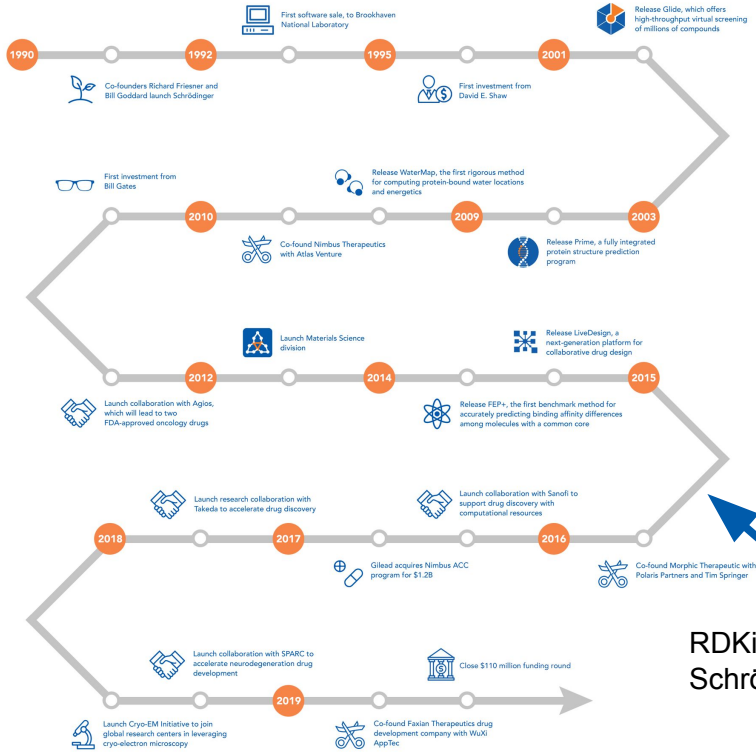
# Stereochemistry: vials vs files

dan nealschneider
d-b-w on github



SCHRÖDINGER
*Transforming drug discovery and materials research*

# Schrodinger and the RDKit

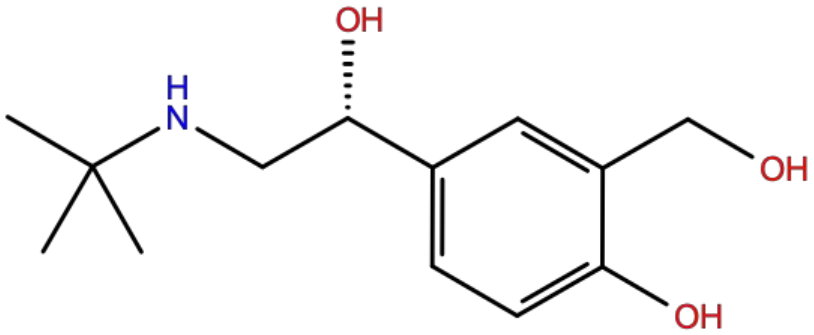



RDKit first bundled with Schrödinger suite



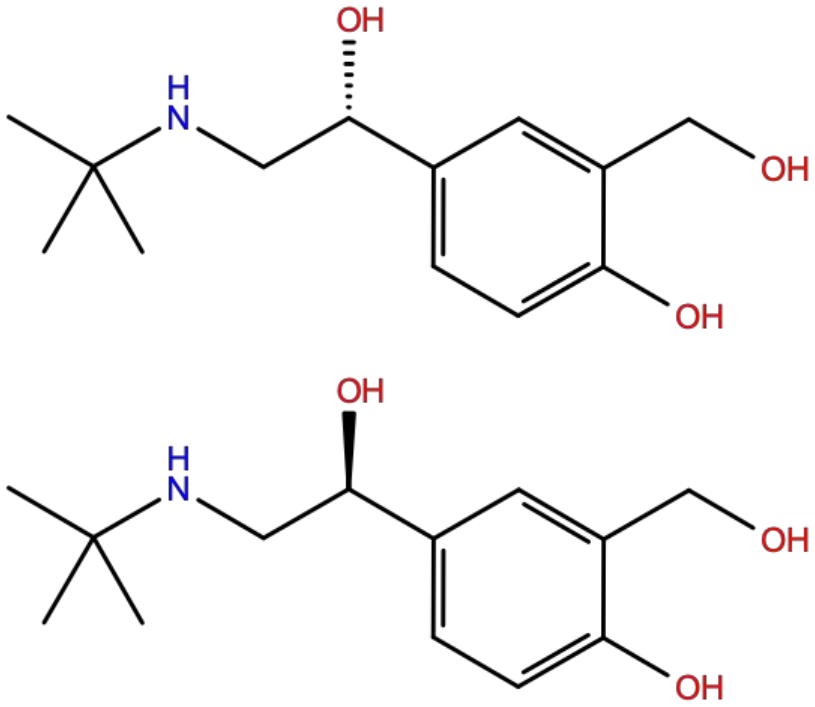
SCHRÖDINGER

# *(R)*-Salbutamol aka albuterol

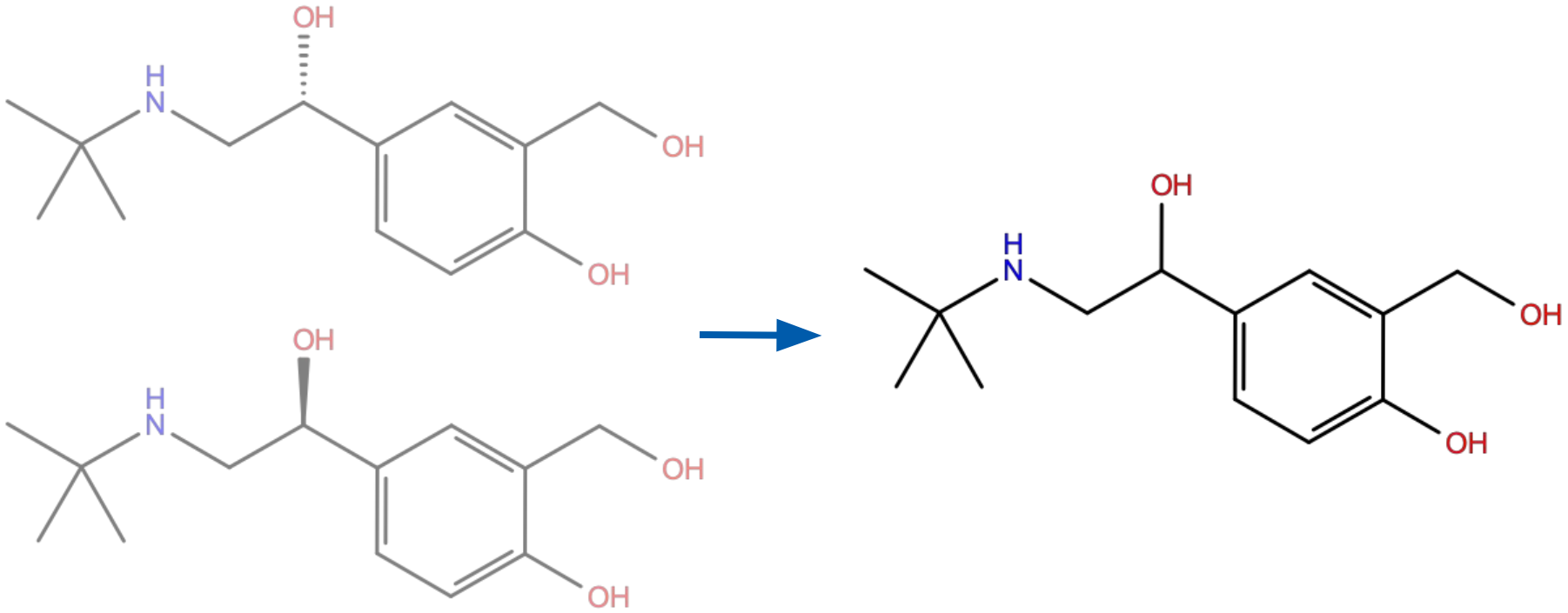


# albuterol

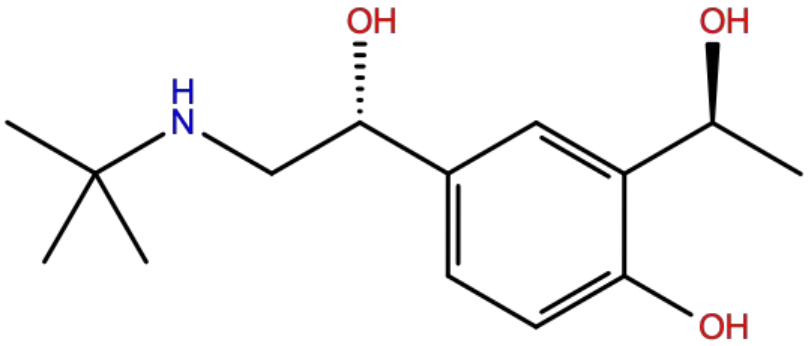
# Salbutamol aka albuterol



# Salbutamol aka albuterol



# A bit more complicated...



# Use Enhanced Stereochemistry

Also called:

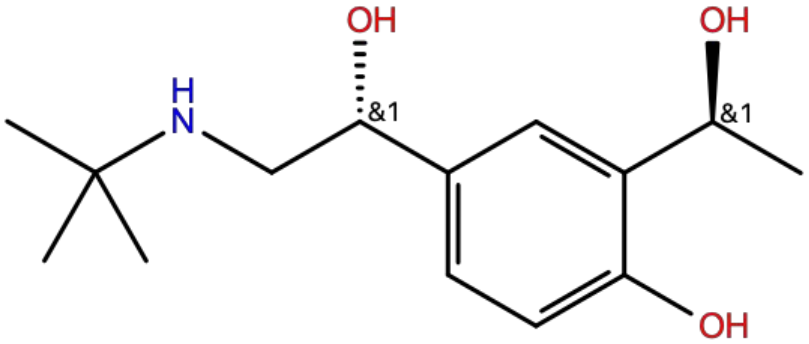
- Extended stereochemistry
- Relative stereochemistry
- Stereo groups

Specified in these file formats

- CXSMILES
- SDF v3000

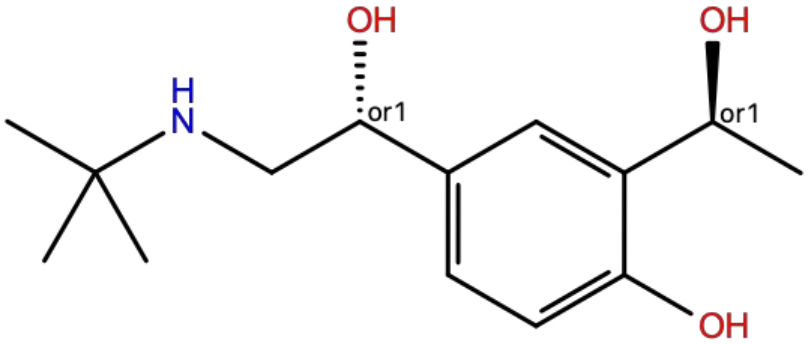
We added support in the RDKit

# AND: A mixture of diastereomers



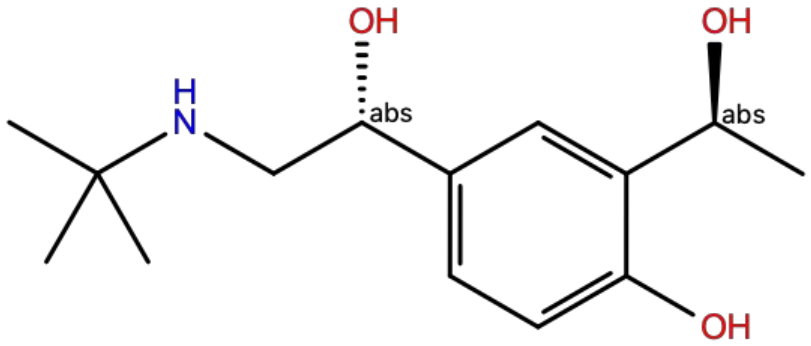

C[C@H](O)c1cc([C@@H](O)CNC(C)(C)C)ccc1O **|&1:1,6|**

# Or - a distinct but unknown stereoisomer

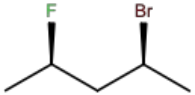
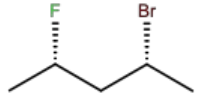
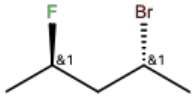
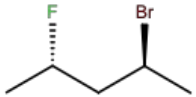
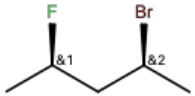
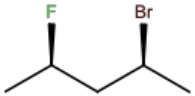
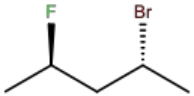
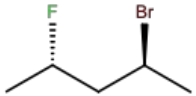
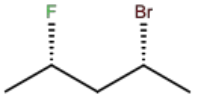
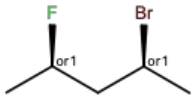
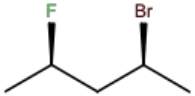
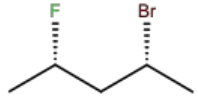
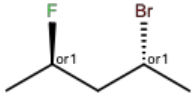
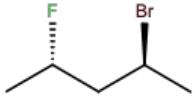
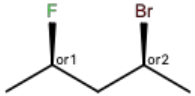
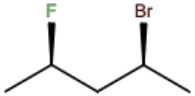
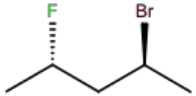
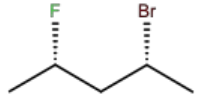



C[C@H](O)c1cc([C@@H](O)CNC(C)(C)C)ccc1O **|o1:1,6|**

# ABSOLUTE: Exactly what it says in the structure




C[C@H](O)c1cc([C@@H](O)CNC(C)(C)C)ccc1O **|a:1,6|**

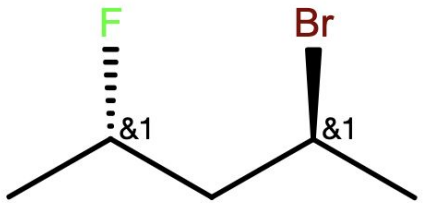
| | | |
|---|---|---|
|  | *Mixture* |   |
|  | *Mixture* |   |
|  | *Mixture* |     |
|  | *Single* |   |
|  | *Single* |   |
|  | *Single* |     |

# CXSMILES - Read and Write

In [1]:
```
from rdkit import Chem
import dan_rendering
```

In [2]:
```
m = Chem.MolFromSmiles('C[C@H](F)C[C@@H](Br)C |&1:1,4|')
m
```

Out[2]:




In [3]:
```
Chem.MolToCXSmiles(m)
```

Out[3]:
```
'C[C@H](F)C[C@H](C)Br |&1:1,4|'
```

# SD files - Read and Write

```
In [4]: print(Chem.MolToMolBlock(m))
```

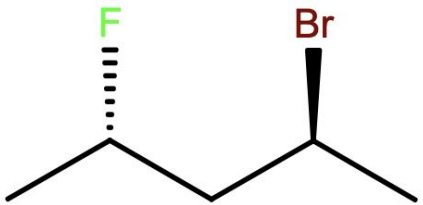
```
     RDKit          2D

  0  0  0  0  0  0  0  0  0  0999 V3000
M  V30 BEGIN CTAB
M  V30 COUNTS 7 6 0 0 0
M  V30 BEGIN ATOM
M  V30 1 C -2.519 -0.1968 0 0
M  V30 2 C -1.6518 0.3014 0 0
M  V30 3 F -1.6498 1.3014 0 0
M  V30 4 C -0.7868 -0.2002 0 0
M  V30 5 C 0.0802 0.298 0 0
M  V30 6 Br 0.0822 1.298 0 0
M  V30 7 C 0.9452 -0.2038 0 0
M  V30 END ATOM
M  V30 BEGIN BOND
M  V30 1 1 2 1 CFG=3
M  V30 2 1 2 3
M  V30 3 1 2 4
M  V30 4 1 4 5
M  V30 5 1 5 6
M  V30 6 1 5 7 CFG=1
M  V30 END BOND
M  V30 BEGIN COLLECTION
M  V30 MDLV30/STERAC1 ATOMS=(2 2 5)
M  V30 END COLLECTION
M  V30 END CTAB
M  END
```

# Add or remove stereo groups (C++ or Python)

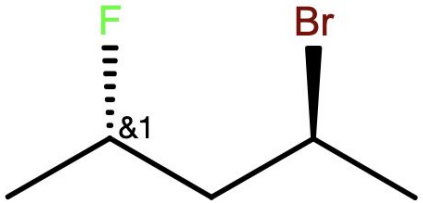
In [5]:
```
editmol = Chem.RWMol(m)
# Clear enhanced stereochemistry groups
editmol.SetStereoGroups([])
Chem.Mol(editmol)
```

Out[5]:




In [6]:
```
from rdkit.Chem import rdchem
# Add a new enhanced stereochemistry group
group1 = rdchem.CreateStereoGroup(rdchem.StereoGroupType.STEREO_AND, editmol, [1])
editmol.SetStereoGroups([group1])
Chem.Mol(editmol)
```

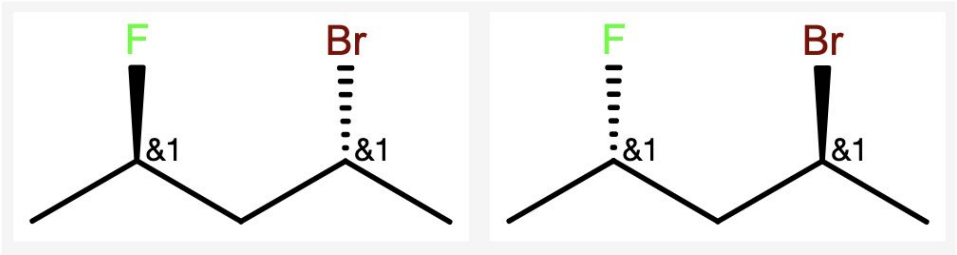
Out[6]:



# Enumerate the stereoisomers (Python only)

In [7]:
```
from rdkit.Chem import EnumerateStereoisomers
# Enumerate the stereoisomers
stereoIsomers = list(EnumerateStereoisomers.EnumerateStereoisomers(m))
dan_rendering.renderMols(stereoIsomers)
```

Out[7]:



# Match any stereoisomer with a MolBundle

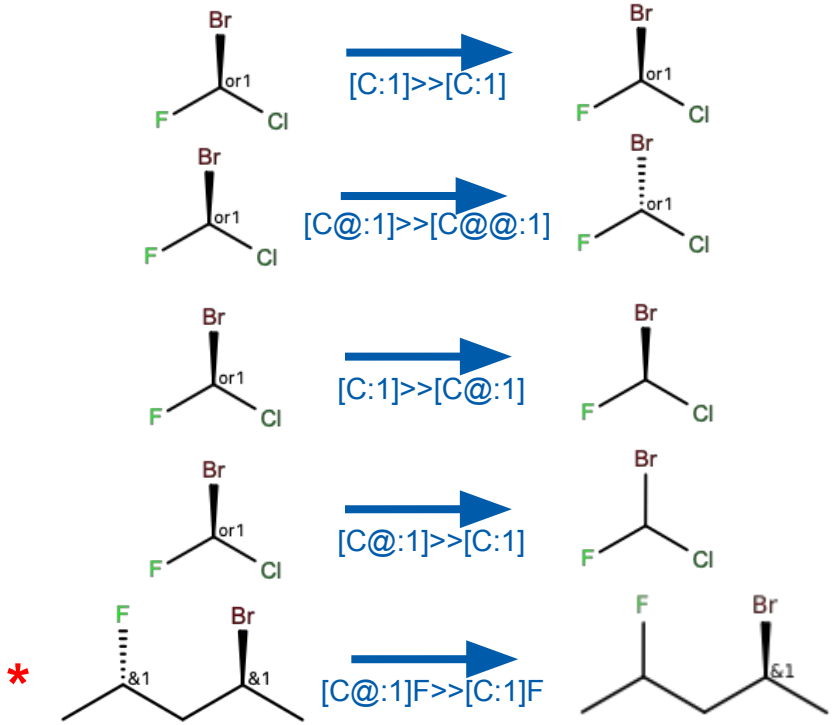
In [8]:
```
Chem.MolToCXSmiles(m)
```

Out[8]:
```
'C[C@H](F)C[C@H](C)Br |&1:1,4|'
```

In [9]:
```
stereoIsomerBundle = Chem.MolBundle()
for isomer in stereoIsomers:
    stereoIsomerBundle.AddMol(isomer)
params = Chem.SubstructMatchParameters()
params.useChirality = True
assert Chem.MolFromSmiles('C[C@H](F)C[C@@H](Br)C').HasSubstructMatch(stereoIsomerBundle, params)
assert Chem.MolFromSmiles('C[C@@H](F)C[C@H](Br)C').HasSubstructMatch(stereoIsomerBundle, params)
assert not Chem.MolFromSmiles('C[C@H](F)C[C@H](Br)C').HasSubstructMatch(stereoIsomerBundle, params)
```

# Reactants with enhanced stereochemistry

Sites are preserved unless chirality is created or destroyed



# Future work

- Depiction
- Canonicalization of SMILES including stereo groups
  - may be complicated for “or” stereo groups
- Substructure searching with stereo groups

# Conclusions

- Enhanced stereo labels allow representation of complex stereochemical relationships
- We expect that the groups will be preserved through RDKit operations
- Matching and uniquification are not seamless
- Other tools may not recognize stereo groups

# Thanks

- Ricardo Rodríguez-Schmidt *dev*
- Greg Landrum *RDKit*
- Marybeth Grimes *ddag*
- Shawn Watts *ddag*
- Amy Rask *ddag*

SCHRÖDINGER®

*Transforming drug discovery and materials research*