benzoapyrene
- 10
- 11
- 12
- 13
- ...
- compound1.sdf

A ***record*** is a folder or a zip file containing:

1) All the NMR spectra of a compound (including FID's, acquisition and processing parameters).

2) The .sdf file containing the NMReDATA for each compound assigned for the sample.

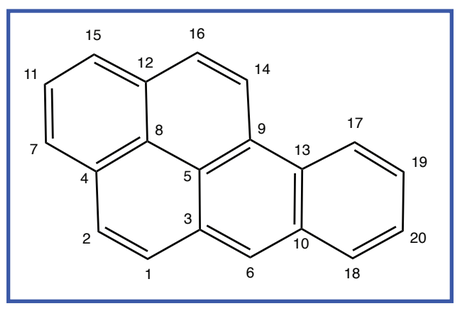
"structure" part of the file

The format is the same as .mol files (compatible with most chemical structure editor)

```
benzo(a)pyrene
demo of .sdf file containing NMReDATA
test_generationV1 1
20 24  0  0  0  0  0  0  0  0999 V2000
   -1.6583    2.2334    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   -0.6633    3.1556    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   -1.3753    0.8191    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    0.7234    2.7603    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    0.0000    0.3848    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
 ....
   -0.5264   -3.3746    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   -3.1687   -2.4653    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   -1.5724   -4.2723    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
   -2.9104   -3.8155    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
 1  2  2  0  0  0
 1  3  1  0  0  0
 2  4  1  0  0  0
 3  5  1  0  0  0
 3  6  2  0  0  0
 4  7  1  0  0  0
 4  8  2  0  0  0
 5  8  1  0  0  0
 5  9  2  0  0  0
 6 10  1  0  0  0
 7 11  2  0  0  0
 8 12  1  0  0  0
 9 13  1  0  0  0
 9 14  1  0  0  0
10 13  2  0  0  0
10 18  1  0  0  0
11 15  1  0  0  0
12 15  2  0  0  0
12 16  1  0  0  0
13 17  1  0  0  0
14 16  2  0  0  0
17 19  2  0  0  0
18 20  2  0  0  0
19 20  1  0  0  0
M  END
```

list of atoms with coordinates (either 2D as in "flat" structures used for drawing or 3D)

list of bonds



Explicit structures (including all hydrogen atoms) should be favored to facilitate the assignment of non-equivalent hydrogens. For aromatic compounds implicit hydrogen is fine.

"NMReDATA" part of the file

```
> <NMREDATA_VERSION>
1.1

> <NMREDATA_ID>
Doi=10.5281/zenodo.1146869\
Record=https://zenodo.org/record/1146869/files/sample1.zip\
Path=compound1.nmredata.sdf\

> <NMREDATA_SOLVENT>
cdcl3

> <NMREDATA_TEMPERATURE>
298.15

> <NMREDATA_ASSIGNMENT>
(1), 128.1712, 1\
(2), 127.9300, 2\
(3), 129.9, 3\
(4), 131.6, 4\
(5), 123.8, 5\
(6), 124.8, 6\
(7), 124.9, 7\
(8), 128.3, 8\
(9), 127.4, 9\
(10), 131.35, 10\
(11), 126.09, 11\
(12), 131.4, 12\
(13), 128.23, 13\
(14), 122.2, 14\
(15), 125.6, 15\
(16), 127.5, 16\
(17), 123.0, 17\
(18), 128.9, 18\
(19), 125.97, 19\
(20), 126.06, 20\
H(1), 8.02, H1\
H(2), 7.94, H2\
H(6), 8.53, H6\
H(7), 8.10, H7\
H(11), 7.98, H11\
H(14), 9.07, H14\
H(15), 8.25, H15\
H(16), 8.34, H16\
H(19), 7.85, H19\
H(20), 7.79, H20\
H(17), 9.06, H17\
H(18), 8.30, H18\
```

Doi are mandatory for NMR Records

The zip file of the record can be accessed directly.

The .sdf file can include "tags". Each tag has a name between "> <" and ">" and data (one or more lines terminating with an empty line. NMReDATA tags start with "NMREDATA_".

This tag associates the labels used in the assignment and the atom(s) of the molecule.

Chemical shifts in ppm (no truncation of the chemical shift before the fourth decimal)

Labels used for the assignment (see right column). In this example, the proton bound to the atom 1 in the structure part (in blue above) was called "H(1)" and the carbon labeled "(1)", but the authors can chose the labels according to IUPAC rules or the requirement of the journal where the data are published.

For compatibility with SDF libraries, the line separator of multiple-lines tags is "\" + CR.

Reference to the atom number in the .mol block (see the structure part in the blue frame). When hydrogen atoms are implicit, the atom number to which the hydrogen atom is bound is given following "H".

Note:
Implicit hydrogen atoms should not be used when non-equivalent hydrogen of CH2's are unambiguously assigned (as in substituded cyclohexane derivatives such as menthol).

1D spectra

```
> <NMREDATA_1D_1H>
;note that integrals (E) were not measured in the spectrum but set
to 1\
Spectrum_Location=file:./benzoapyrene/10/pdata/1/\
Larmor=500.13\
9.0600, L=H(17), S=d, J=8.44(H(19)), E=1.00\
9.0700, L=H(14), S=d, J=9.13(H(16)), E=1.03\
8.5300, L=H(6), S=s, E=1.03\
8.3400, L=H(16), S=d, J=9.13(H(14)), E=1.05\
8.3000, L=H(18), S=d, J=8.06(H(20)), E=1.02\
8.2500, L=H(15), S=d, J=7.65(H(11)), E=1.00\
8.1000, L=H(7), S=d, J=7.31(H(11)), E=1.06\
8.0200, L=H(1), S=d, J=9.07(H(2)), E=1.02\
7.9800, L=H(11), S=dd, J=7.31(H(7)), 7.65(H(15)), E=1.02\
7.9400, L=H(2), S=d, J=9.07(H(1)), E=1.01\
7.8500, L=H(19), S=dd, J=6.77(H(20)), 8.44(H(17)), E=1.05\
7.7900, L=H(20), S=dd, J=6.77(H(19)), 8.06(H(18)), E=0.98\
```

Chemical shift (or range in case of overlap or complex multiplet)

location of the spectrum in the folder/zip files

Label (see left)

Multiplicity ("d" for doublet, etc.)

Coupling constants. They can be assigned as here: $J_{H(11),H(7)}$ = 7.31 Hz

Integral (crude exp. values). The rounded number of atom can be given following N=)

For a list of all 1D signal attributes, see: http://nmredata.org/wiki/1D_attributes

```
> <NMREDATA_1D_13C>
;note that intensisites were not measured in the spectrum 110 for CH
100 for Cq\
Spectrum_Location=file:./benzoapyrene/13/pdata/1/\
Larmor=125.0\
Decoupled=1H\
131.6000, L=(4), I=100.00\
131.4000, L=(12), I=100.00\
131.35, L=(10), I=100.00\
129.9, L=(3), I=100.00\
128.9, L=(18), I=110.00\
128.17, L=(1), I=110.00\
128.23, L=(13), I=100.00\
127.9, L=(2), I=110.00\
127.4, L=(9), I=100.00\
127.5, L=(16), I=110.00\
126.09, L=(11), I=110.00\
126.06, L=(20), I=110.00\
128.3, L=(8), I=100.00\
125.97, L=(19), I=110.00\
125.6, L=(15), I=110.00\
124.8, L=(6), I=110.00\
124.9, L=(7), I=110.00\
123.8, L=(5), I=100.00\
123.0, L=(17), I=110.00\
122.2, L=(14), I=110.00\
```

For a list of all spectra properties, see: http://nmredata.org/wiki/NMReDATA_tag_format

Chemical shift in ppm

Label (see the NMREDATA_ASSIGNMENT tag)

Signal intensity (usefull to distinguish signals in DEPT spectra, to determine the sign of the peak and quantidy NOE. Here they were set to 100 and 110 but the experimental values should be given here.)

2D spectra

Isotope in F1 / type of mixing / isotope in F2

```
> <NMREDATA_2D_1H_NJ_1H>
Larmor=500.13\
CorType=COSY\
Spectrum_Location =file:./benzoapyrene/12/1/pdata/1/\
H(1)/H(2)\
H(7)/H(11)\
H(11)/H(15)\
H(14)/H(16)\
H(19)/H(20)\
H(19)/H(17)\
H(20)/H(18)\
```

"N": multiple-bond, "J": scalar coupling mixing

Correlation of "H(1)" with "H(2)".

13C isotope in F1, 1J mixing, 1H isotope in hdimension

```
> <NMREDATA_2D_13C_1J_1H>
Larmor=500.13\
CorType=HSQC\
Spectrum_Location =file:./benzoapyrene/13/pdata/1/\
(1)/H(1)\
(2)/H(2)\
(6)/H(6)\
(7)/H(7)\
(11)/H(11)\
(14)/H(14)\
(15)/H(15)\
(16)/H(16)\
(17)/H(17)\
(18)/H(18)\
(19)/H(19)\
(20)/H(20)\
```

Link to the spectra (the data must include the fid, and the processed data). This points either to a folder where the .sdf file is located or to a permanent record in an open-access database.

Correlation of carbon "(1)" with hydrogen "H(1)".

For a list of all 2D signal attributes, see: http://nmredata.org/wiki/2D_attributes

```
...
```

HMBC, NOESY data would follow with the same format

```
> <NMREDATA_LITERATURE>
Source=Journal
DOI=....
```

Reference to the main publication associated to the record.

```
> <...>
```

Other data such as the origin of the sample, the sample preparation and other analytics data could (and should) be added to .sdf file ***outside*** the NMReDATA initiative.

```
$$$$
```

End of file code