Hydrogen Bond Analysis

G. Matthias Ullmann

[1] Computational Biochemistry, University of Bayreuth, Universitätsstr. 30, NW I, 95447 Bayreuth, Germany

March 30, 2022

Contents

1 Hydrogen Bonding – Theoretical Background

Hydrogen bonds are essential in many molecules. Usually they are defined by geometric criteria. The geometric criteria used in our program are listed in Figure 1. Often, just the distance d_{AD} are used as parameter. However, this distance may not be ideally suited, since sulfer or selenium have a larger radius than oxygen or nitrogen. Therefore, we use the Van-der-Waals radii of the donor and acceptor atoms, which are either read from a PQR-file or from a the PSE file.

In principle, one should differentiate between sp3, sp2 and sp1. However this is probably not required because of the relatively rough definition of the hydrogen bond in this context.

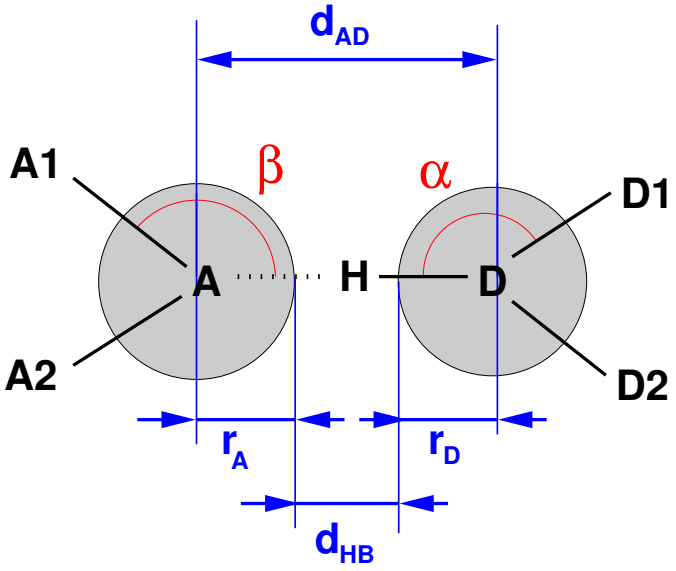

Figure 1: Hydrogen Bond

1.1 Program: `hb-search`

The program searches for hydrogen bonds in a molecular structure file. A hydrogen bond is defined as as shown in Figure 1.

Hydrogen atoms are not considered

Also connections to the solvent are searched. If a donor or an acceptor atom is at the surface of the molecule, an hydrogen bond to the solvent is assigned.

If the molecule is a membrane protein, it is also possible to assigned the side of the membrane. If the z-coordinate is less then 0 and the donor/acceptor is at the surface of the molecule, a hydrogen bond to NEG is assigned. If the z-coordinate is larger then 0 and the donor/acceptor is at the surface of the molecule, a hydrogen bond to POS is assigned.

To Do:

- add connection to solvent
- differentiate between cytoplasma and extra cellular

- get pathways
- get coordination shells

```
hb-search [-blab blab] [-hb file] [-solv key] [-pse file] [-con 0/1] mol

Default values
          hb-distance <= R_VdW(1) + R_VdW(2) +  DIST_EPS
            80.0 <= alpha <=  150.0
            80.0 <= beta  <=  150.0
Solv Key:
          NONE -- no solvent HBonds
          HOMO -- homogeneous solvent
          MEMB -- membrane environment

          if con is set to 1, special connection (CONNECT)
          are taken into account
```

One of the important input file is a file defining the hydrogen bond parameter and the donor and acceptor atoms.

The default name of this files is `hb-define.txt`. It has the following structure

```
#
# ACCE -- HB-Aceptor
# DONO -- HB-Donor
# DOAC -- Both HB-Aceptor and HB-Donor as for instance -OH
#

DIST_EPS   1.0
LALPHA    80.0
RALPHA   150.0
LBETA     80.0
RBETA    150.0
DENSIT     255
PROBE      1.4
TessType     0

#-------------------------------------------------
# Backbone

ACCE  ALA:O   ALA:C
DONO  ALA:N   ALA:CA

ACCE  ARG:O   ARG:C
DONO  ARG:N   ARG:CA
.
.
.
```

```
#-------------------------------------------------
# aminoacid side chains
#
DOAC  ARG:NE  ARG:CD
DOAC  ARG:NH1 ARG:CZ
DOAC  ARG:NH2 ARG:CZ

DOAC  ASN:OD1 ASN:CG
DOAC  ASN:ND2 ASN:CG
.
.
```

There are the following parameter in the input file.

DIST_EPS (default: 1.0):

LALPHA (default: 80.0):

RALPHA (default: 150.0):

LBETA (default: 80.0):

RBETA (default: 150.0):

DENSIT (default: 255):

PROBE (default: 1.4):

DOAC : Define atom and its neighbor of atoms that can act as donor or acceptor.

ACCE : Define atom and its neighbor of atoms that can act as acceptor

DONO : Define atom and its neighbor of atoms that can act as donor

The donor/acceptor atom are defined by as `RESNAME:RESID`. If there is no neighboring atom, (for instance for `HOH`) one writes `---`.

The output is written to `stdout`. It has a structure that allows an easy use of the program `grep`. The file has the following structure:

The hydrogen bond network generated by this software can be used in other programs such as for instance `hb-network`. If some atoms are not available to hydrogen bonding, the output file can be modified either "by hand" (i.e. by an ASCII editor) or by software such as `awk`, `sed`, or `grep`.

2 Program: `hb-network`

3 Program: `get-water-cluster`

4 How to deal with a membrane protein?

1. generate a dummy membrane using `dot-surface` using a radius of 2.4 Å (which corresponds to the radius of methane). Add some layers around the molecule
2. fill cavities with water
3. check if any cavities are left and repeat if required
4. add dummy layer and make clefts to dummy cavities; fill these clefts with water
5. check if any clefts or cavities are left and repeat if required
6. generate hydrogen bonded network and find paths to `SOLV` or `CP/EC`

This program writes the HB network to stdout. This output can be edited and for instance atoms not available for HB bond can be deleted. Then the file is read by other programs and analyzed:

- find connected network
- shortest path search
- find clusters with betweenness

There is a difference between hydrogen bond networks and proton transfer networks. Proton transfer networks contain atoms that can accept and donate protons, while in hydrogen bond networks, one may have just hydrogen bond (such as in protein backbones).

Thus, each proton transfer network is a hydrogen bond network, but not vice versa. Often proton transfer networks are subsets of the whole hydrogen bond network.