SPORES
Structure PrOtonation and REcognition System
user manual for version 1.28 [1]

T. ten Brink

Tim.tenBrink@uni-konstanz.de

1 running SPORES

In this improved version of the structure recognition and protonation tool SPORES seven running-modes are implemented: complete protonation, reprotonation and adjustment of Sybyl-Types for atoms and bonds with or without checking the connectivity like in the first version. Additionally the new Version is able to produce all combinatorial possible protonation states of the input molecule. Amine groups are switched form R-NH_3^+ to R-NH_2, carboxyl groups are protonated or deprotonated and nitrogens in aromatic rings can be protonated or deprotonated. Planar nitrogen atoms will not be changed. SPORES can be executed by typing *SPORES --mode running-mode yourinputfile youroutputfile* where *running-mode* is one of the options below. *yourinputfile* is the structure file used as input in pdb or mol2 format. *youroutputfile* is the protonated structure file of your molecule in mol2 format.

1.1 running-modes for pdb input files

For processing pdb files SPORES need files from the remediated data base due to its reading routine. The output will be written in mol2 format.

- **completepdb**: structure recognition and protonation
- **reprotpdb**: deletes all hydrogen atoms from a pdb file and reprotonates the input structure
- **splitpdb**: for splitting a downloaded pdb file into protein, ligands and water molecules. This mode will perform a structure recognition and protonation like the "completepdb" mode but the output are written in different files: *protein.mol2* which contains all protein chains from the input, *ligand_ligandname_X.mol2* which contain the ligand molecuels (*ligand_name* is generated from the residue name of the ligand in the original file, and X is a running number in case there are two ligands with identical names) and *water.mol2* which contains all water molecules from the input file.

1.2 running-modes for mol2 input files

- **completemol2**: structure recognition and protonation
- **reprot**: deletes all hydrogen atoms and reprotonates the input structure
- **settypes**: for protonated structures, checks connectivity and assigns Sybyl atom and bond types
- **readbonds**: reads the connectivity from your input file and assigns Sybyl atom and bond types
- **protstates**: conducts structure recognition, protonation and generates all combinatorial possible protonation states of the input molecule. The protonation states will be written in a multi-mol2 file *protstates.mol2*. Additional the number of protonation states and the difference of each state to the standard protonation will be written to the console.
- **stereo**: conducts structure recognition, protonation and generates all combinatorial possible stereoisomers of the input molecule. The stereoisomers are written in a multi-mol2 file *stereo.mol2*.
- **tautomers**: conducts structure recognition, protonation and generates all combinatorial possible tautomers for nitrogen containing planar ring-systems of the input molecule. The tautomers are written in a multi-mol2 file *tautomers.mol2*.
- **ketoenol**: conducts structure recognition, protonation and generates keto-enol-tautomers of the input molecule. The tautomers are written in a multi-mol2 file *ketoenol.mol2*.
- **ringconfs**: conducts structure recognition, protonation and generates ring-conformers of the input molecule. The conformers are written in a multi-mol2 file *ringflips.mol2*. The ring conformers are only generated for 5 and 6 membered rings.
- **script**: A mode for manual correction of errors in structures called by *SPORES --mode running-mode yourinputfile youroutputfile yourscriptfile*. the syntax for the script file is described below.

1.3 script syntax

The script file is is a text file which contains commands to correct most of the common errors in structures.
First line: *file_type pdb* or *mol2* this line must be in the script file.
In the next lines some global definitions can be made:
skip_prot if contained in the script file no protonation is done.
auto_bond if contained in the script file automatic assignment of bond types will be done.
backbone the backbone flag for atoms and bonds is added to the mol2 file.

Some commands for single atoms or bonds:
force_atom_type atomnumber atomtype sets the type of an atom to atomtype.
force_bond_type atomnumber1 atomnumber2 bondtype sets the type of a bond to bondtype.
erase_atom atomnumber deletes the atom from the molecule.
erase_bond atomnumber1 atomnumber2 the same for a bond.
force_prot atomnumber a single nitrogen atom can be protonated.
Some commands for protein residues:
exclude_residue residuenumber the residue is skipped when the molecule is read.
switch_his residuenumber the protonation of the nitrogen atoms in the ring is reversed.

1.4 examples

A pdb structure file *ligand.pdb* is to be protonated so
./SPORES --mode completepdb ligand.pdb ligandprotonated.mol2
will generate a mol2 file with the protonated structure called *ligandprotonated.mol2*

Now the different protonation states can be generated by calling:
./SPORES --mode protstates ligandprotonated.mol2 standartprotonation.mol2
This will generate a multi-mol2 file called *protstates.mol2* and a file called *standardprotonation.mol2* which has the same protonation state as the input file.

If the input file is already in mol2 format you can generate the protonation states directly:
./SPORES --mode protstates ligand.mol2 ligandprotonated.mol2
This will generate the mol2 files called *protstates.mol2* with the different protonation states and the standard protonation state *ligandprotonated.mol2*

References

[1] ten Brink, T.; Exner, T.E. Influence of Protonation, Tautomeric, and Stereoisomeric States on Protein-Ligand Docking Results. *J. Chem. Inf. Model.*, 49:1535–1546, 2009. 1