

SESA: a program for analytic computations of solvent-accessible surface area and solvent-excluded surface area

Lincong Wang*

Abstract

The surface area of a molecule, an inherent geometric property of its structure, plays important roles in its solvation and functioning. Here we present an accurate and robust program, SESA, for the analytic computations of both solvent-accessible surface (SAS) area and solvent-excluded surface (SES) area. The accuracy and robustness are achieved through the analytic computations of all the SAS regions for a surface atom and of probe-probe intersections. The detailed comparisons of the areas for a large set of protein structures by SESA and MSMS, a de-facto standard for analytic SAS and SES computations, confirm SESA's accuracy to a good extent and in the same time reveal significant differences between them. The unprecedented accuracy and robustness of SESA make it possible to analyze in great detail the surface areas of any molecules in general and biomolecules in particular.

1 Compilation

SESA is written in C++, developed under Ubuntu 20.04 LTS and compiled with windows10 with minGW.

2 Installation

The installation is very simple since no Qt is required. SESA could be installed as follows.

1. DOWNLOAD the file "sesaV001-windows.zip"
2. UNZIP the file
3. GO TO "sesA/bin"

The program, *sesa*, will be in *bin* directory.

The directory, *bioParam*, contains the charmm force field parameter files.

The directory, *lib*, contains libraries required by *sesa*. To run *sesa* one needs to edit the path as follows: search for "edit the system environment variables", then Click "Environment Variables", then Edit Path by adding the absolute path of the *lib* folder. Alternatively one could place the libraries from the *lib* folder into the *bin* folder. The former is preferred.

3 The usage

By CLICK on *sesa* directly, it computes SES areas for "2FR3FH.pdb" which is in the same folder as the program *sesa* itself.

3.1 Terminal (Windows PowerShell) command

```
.\sesA ..\1b18FH.pdb
```

*Corresponding author: Lincong Wang, Email: wlincong@gmail.com.

3.2 Terminal output

```
***** PROcess 1bl8FH.pdb *****
HEADER      MEMBRANE PROTEIN                      23-JUL-98    1BL8
TITLE       POTASSIUM CHANNEL (KCSA) FROM STREPTOMYCES LIVIDANS
REMARK      2 RESOLUTION.      3.20 ANGSTROMS.
HETNAM      K POTASSIUM ION
```

```
The largest compound has heavy atoms = 4
      chainid=B      chainid=D      chainid=C      chainid=A
MORE than ONE chain
```

```
-----
1bl8FH Chains:
has NO N-terminus NHs !
      has NO C-terminal!
      Has 57 missing atoms:
( B 1412)_atom = 57, 4.04%
has NO N-terminus NHs !
      has NO C-terminal!
      Has 57 missing atoms:
( D 1412)_atom = 57, 4.04%
has NO N-terminus NHs !
      has NO C-terminal!
      Has 57 missing atoms:
( C 1412)_atom = 57, 4.04%
has NO N-terminus NHs !
      has NO C-terminal!
      Has 57 missing atoms:
( A 1412)_atom = 57, 4.04%
```

```
***** SES areas by Atom *****
Atom_Probe_CPUtime(s)_clockTime(s): 5648 6720 1.861 1.861
```

The terminal output includes some basic information about the structure and the time for SES area computation.

3.3 The SES area file

The atomic SES areas computed by *sesa* are saved to a file named as “pdbidFH_protSES.txt” in the same directory. The four rows of “1bl8FH_protSES.txt” are shown as follows. The first row is the header. The columns are, respectively, atom indices, SAS area (as a part of ses area), toroidal area, probe area, ses area, atom identifier, SAS area and SAS regions. The SAS area in column 2 differs from the SAS area in column 7 in that the latter is computed by adding probe radius (1.4Å) to atom radius. The latter is the typical SAS area used in many applications. These areas are for the exterior surface (e-surface) only.

atIndex	sas	toroidal	probe	ses	atomID	SAS	SASregions
0	11.712991	6.57483	2.14069	20.42851	B:A23_N	42.42760	
1	0.439732	1.74494	0.64285	2.82752	B:A23_CA	1.46222	0 0 0 0.075952 0.232487 0.131293
1157	0.028927	0.55302	2.74808	3.33003	B:S102_O	0.10675	0 10 0.028927 0.042037

the last column “SASregions” is empty for atom 0: meaning that atom 0 has only one SAS region and that region is on the e-surface of 1bl8. The “SASregions” for atom 1 is “0 0 0 0.075952 0.232487 0.131293”: meaning that it has three SAS regions, all of them are on the e-surface and their areas are, respectively, 0.075952, 0.232487 and 0.131293. The “SASregions” for atom 1157 is “0 10 0.028927 0.042037”: meaning it has two SAS regions, the first one is on the e-surface while the second one is on interior surface (i-surface) 10 and their areas are, respectively, 0.028927 and 0.042037.

The rows of atomic areas is followed by three rows listing some statistics about the protein and the total SES areas.

```
Atoms= 5648 surfAtom= 2813 buriedAtom= 2835
Total: sasArea= 4064.5881 torusArea= 6001.8320 probeArea= 5936.7197
Average: sasArea= 1.4449 torusArea= 2.1336 probeArea= 2.1105
```

The total number of atoms (with proton added), the total number of surface atoms (including those of i-surfaces), and the total number of buried atoms are listed in the first row. The second row lists the total SAS area (as part of ses area), toroidal area, and probe area. The last row lists

their average values.

The last parts of the file (starting with “— internal Cavities —”) list the atoms for every i-surface. The last row lists the total SAS area (as a part of ses ara) of all the i-surfaces of 1bl8: “sasAreaOfAllCavities= 26.7479”.