MODELING AND PREDICTING MACROMOLECULAR STRUCTURE

Alireza Doustmohammadi

JULAY ٢٠١٩
TARBIAT MODARES UNIVERSITY
Nasr, Jalal al ahmad, Tehran

Abstract:

OSPREY software use to protein design community a unique combination of continuous flexibility modeling, ensemble modeling, and algorithms with provable guarantees. OSPREY ٣،٠ represents a complete refactoring of the code, and presents a simpler and more intuitive interface that makes protein redesign much easier than before. The new, developer-friendly code organization also facilitates adding new features to the free and open-source OSPREY project, both by ourselves and by other contributors. We have introduced a convenient Python scripting interface and added support for GPU acceleration of the bulk of the computation, allowing designs to be completed much more quickly and easily than in previous versions of OSPREY. OSPREY has been used for an impressive number of empirically successful designs, ranging from enzyme design to antibody design to prediction of antibiotic resistance mutations. [١]

Introduction:

OSPREY software use to protein amino acids mutation. For this goal I use ١CC٨.ss protein to check effect of mutation for residues ٢, ٣, ٤, ٣٦ and ٤٢ in chain A (Table one). Osperey software can be parallel and use A* algorithm to find best mutation. This approach cause decrease time complexity.

Table١: Describe Mutations Implement in this project

Mutation Level	Description	
	Residue number	Mutation
١	٢	'ALA' or 'GLY'
	٣	Don’t change or 'VAL'
	٤	Don’t change
٢	٢	'ALA' or 'GLY' or 'VAL' or 'ASN'
	٣	Don’t change or 'VAL'
	٤	Don’t change
٣	٢	'ALA' or 'GLY' or 'VAL' or 'ASN' or 'ARG' or 'PHE' or 'SER' or 'THR'
	٣	Don’tchange or 'VAL' or 'ASP' or 'CYS' or 'GLN' or 'GLU' or 'TRP'
	٤	Don’t change, or 'GLY' or 'HIS' or 'LEU' or 'LYS' or 'MET' or 'TYR'
٤	٢	'ALA', 'GLY','VAL','ASN','ARG','PHE','SER','THR','ASP','CYS','GLN','GLU','TRP','HIS','LEU','LYS','MET','TYR'
	٣	Don’tchange , 'ALA', 'GLY','VAL','ASN','ARG','PHE','SER','THR','ASP','CYS','GLN','GLU','TRP','HIS','LEU','LYS','MET','TYR'
	٤	Don’tchange , 'ALA', 'GLY','VAL','ASN','ARG','PHE','SER','THR','ASP','CYS','GLN','GLU','TRP','HIS','LEU','LYS','MET','TYR'
٥	٢	'ALA', 'GLY','VAL','ASN','ARG','PHE','SER','THR','ASP','CYS','GLN','GLU','TRP','HIS','LEU','LYS','MET','TYR'
	٣٦	Don’tchange , 'ALA', 'GLY','VAL','ASN','ARG','PHE','SER','THR','ASP','CYS','GLN','GLU','TRP','HIS','LEU','LYS','MET','TYR'
	٤٢	Don’tchange , 'ALA', 'GLY','VAL','ASN','ARG','PHE','SER','THR','ASP','CYS','GLN','GLU','TRP','HIS','LEU','LYS','MET','TYR'

Implementation:

OSPREY is developed and maintained by the Donald Lab in the Department of Computer Science at Duke University. To install it I do these steps:

١. **Install prerequisites:**

sudo apt-get install python٢,٧ python-pip openjdk-٨-jre

٢. **Run the install shell script to install Osprey:**

./install.sh

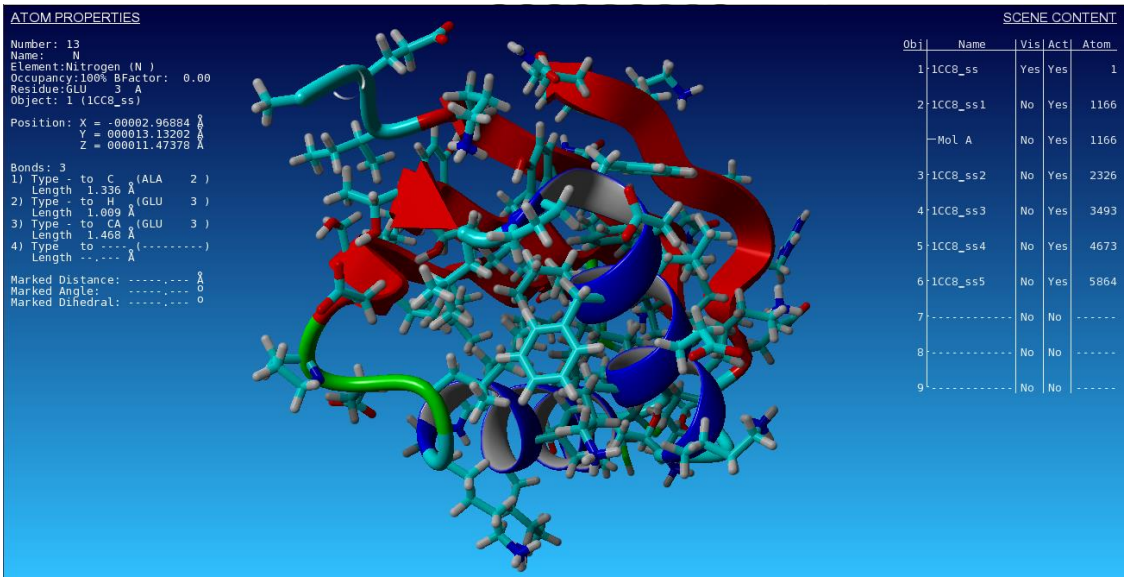
To run Osprey from a Python script we should import Osperey package and start it:

```
import osprey

osprey.start()


# run osprey commands.
```

I mutate ١CC٨.ss protein. Native protein pdb file and protein folding is:



Now I describe a osperey scripts results:

I run puthoon files by:

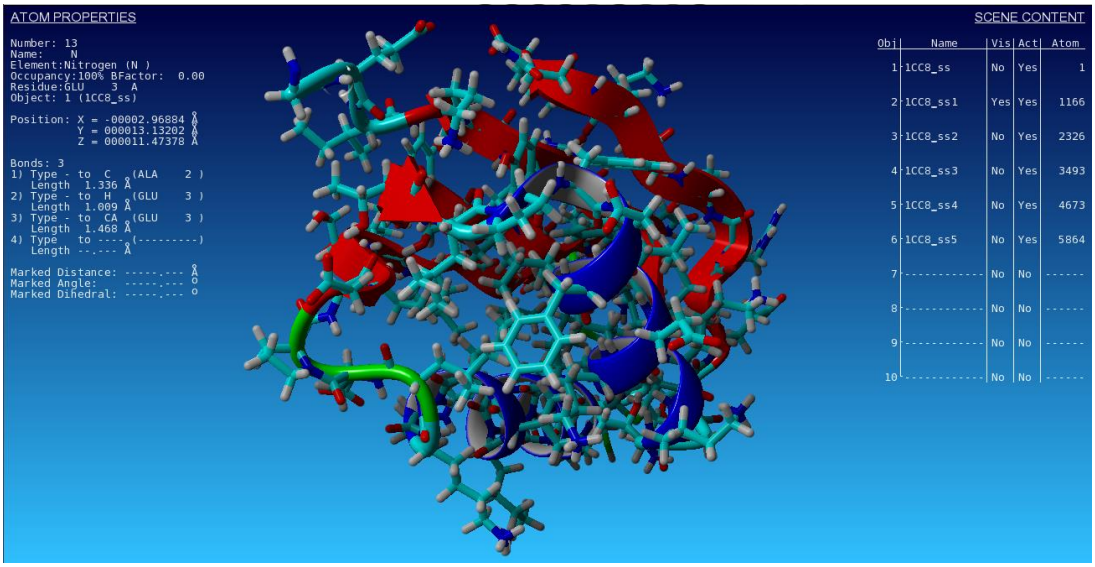
Python2 'filename'

1. Mutation Level one:

Best Energy and Score is:

```
OSPREY 3.0
read PDB file from file: 1CC8.ss.pdb
Calculating energy matrix with 133 entries...
Searching for min score conformation...
        (among 154.0 possibilities)
Found min score conformation in 8.9 ms
Computing energy of min score conf...
Found GMEC!
        Residue Conf Ids       1   3   4
        Residue types        GLY GLU ILE
        Rotamer numbers        L  L3  L4
        Energy               -30.705504
        Score                -30.705504 (gap: 0.000000)
```

Pdb result with mutation as shown as below:



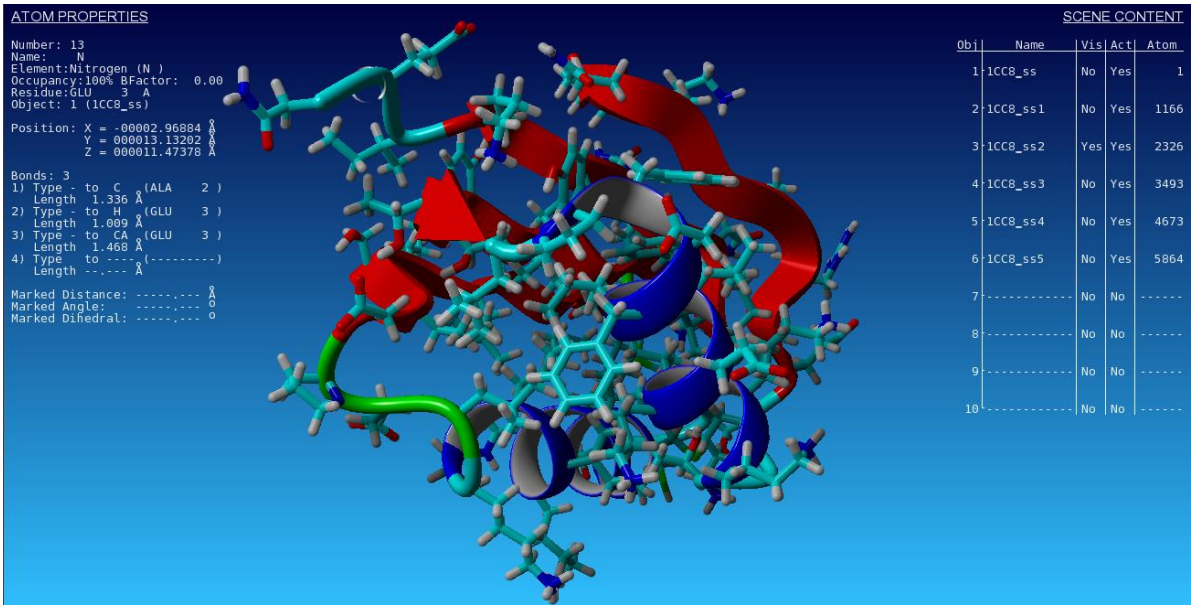
٢. Mutation Level two:

Best Energy and Score is:

```
Progress:  100.0%   ETA: 0 ns
Finished in 210.9 ms
checked 9 conformations

Found GMEC!
        Residue numbers      A2    A3    A4
        Residue types        ASN   GLU   ILE
        Rotamer numbers      L5    L3    L4
        Residue Conf Ids     10    3     4
        Energy               -39.225390
        Score                -39.225390 (gap: 0.000000)
Found 9 total conformations in energy window
```

Pdb result with mutation as shown as below:

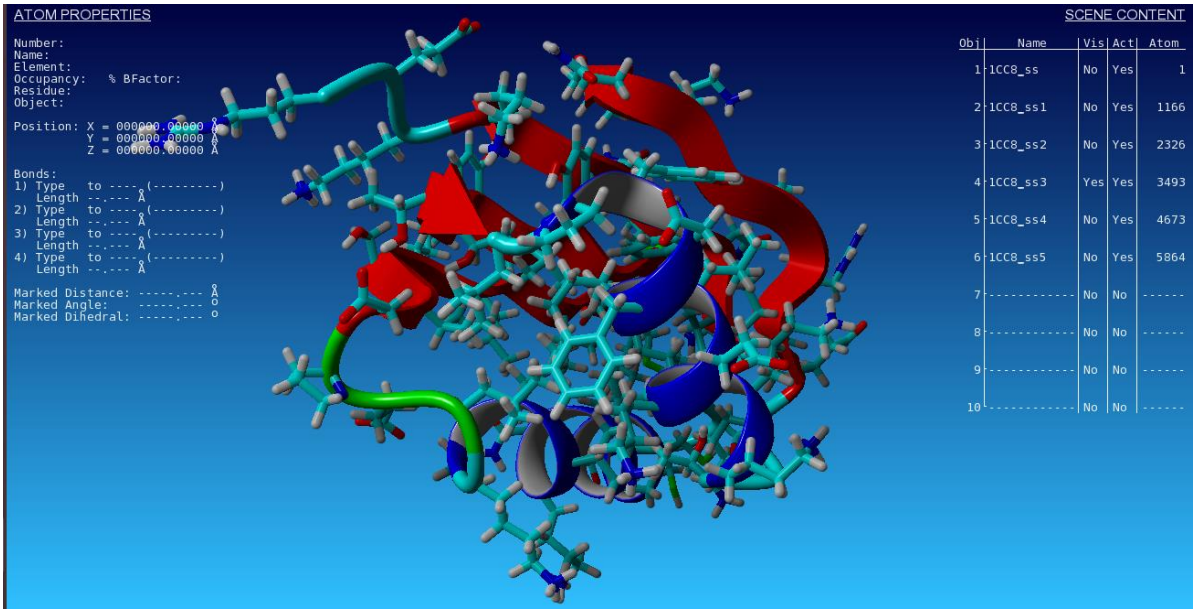


٣. Mutation Level three:

Best Energy and Score is:

```
Found min score conformation in 347.4 ms
Computing energy of min score conf...
Found GMEC! (it's actually the min score conformation too!)
        Residue numbers      A2    A3    A4
        Residue types        ARG   GLU   LYS
        Rotamer numbers      L24   L3    L19
        Residue Conf Ids     36    3     40
        Energy               -60.432884
        Score                -60.432884 (gap: 0.000000)
```

Pdb result with mutation as shown as below:



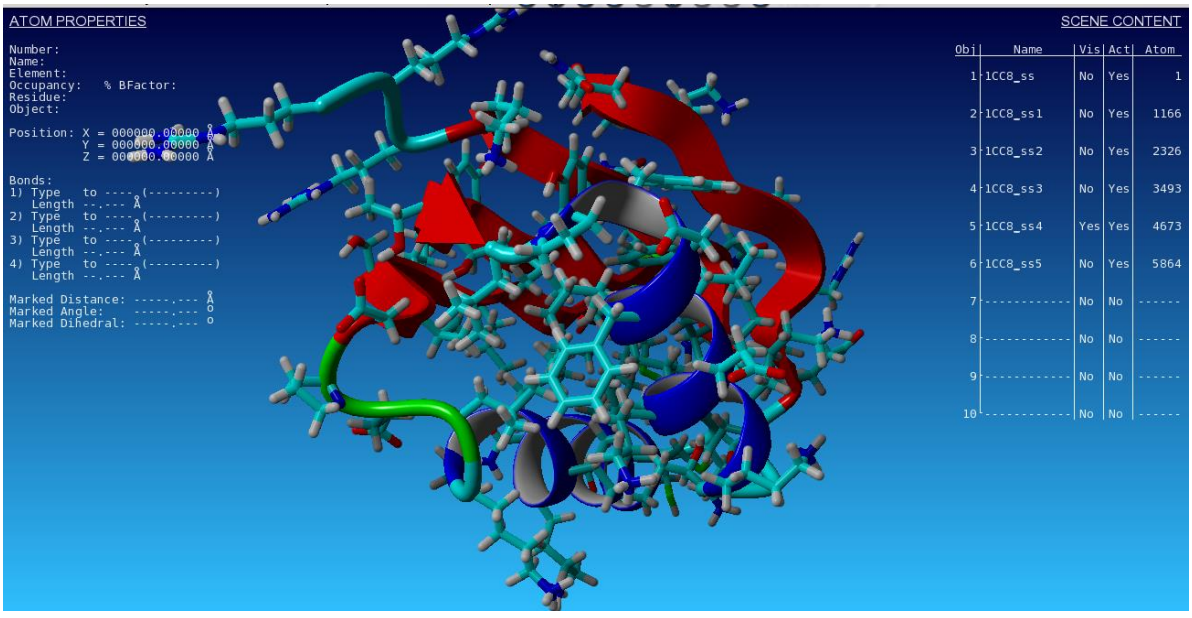
٤. Mutation Level four:

Best Energy and Score is:

```
Found min score conformation in 891.2 ms
Computing energy of min score conf...
Found GMEC! (it's actually the min score conformation too!)
        Residue numbers      A2    A3    A4
        Residue types        ARG   ARG   ARG
        Rotamer numbers      L24   L15   L24
        Residue Conf Ids     36    35    43
        Energy               -85.895743
        Score                -85.895743 (gap: 0.000000)


Residues           A2    A3    A4
Sequence           ARG   ARG   ARG
Rotamers           L24   L15   L24
Residue Conf IDs   36    35    43
Energy             -85.895743
Score              -85.895743   (gap: 0.000000)
```

Pdb result with mutation as shown as below:



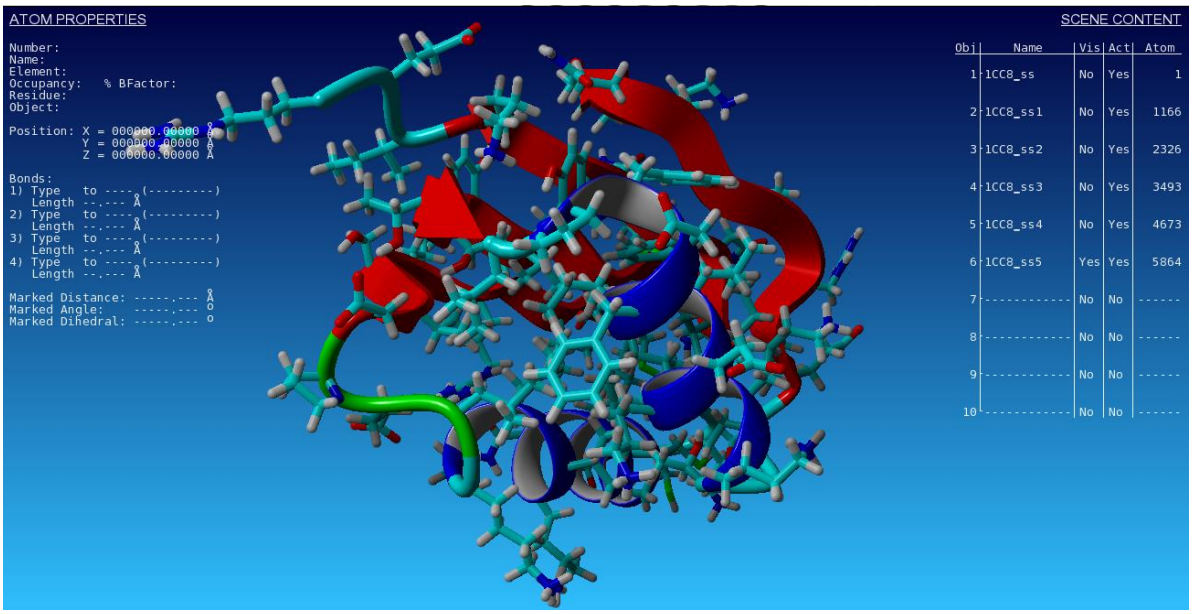
٥. Mutation Level five:

Best Energy and Score is:

```
Computing energies for 28 conformations...
Progress:  100.0%   ETA: 3.0 ms
Finished in 770.7 ms
checked 29 conformations

Found GMEC!
        Residue numbers      A2    A36   A42
        Residue types        ARG   ARG   ARG
        Rotamer numbers      L24   L0    L27
        Residue Conf Ids     36    19    66
        Energy               -96.991462
        Score                -96.991462 (gap: 0.000000)
Found 29 total conformations in energy window
```

Pdb result with mutation as shown as below:



References

[١] Mark A. Hallen,Jeffrey W. Martin,Adegoke Ojewole, Jonathan D. Jou,Anna U. Lowegard,Marcel S. Frenkel,Pablo Gainza,Hunter M. Nisonoff,Aditya Mukund,Siyu Wang,Graham T. Holt,David Zhou,Elizabeth Dowd,Bruce R. Donald, "OSPREY ٣,٠: Open-Source Protein Redesign for You, with Powerful New Features," *Journal of Computational Chemistry,* ٢٠١٨.