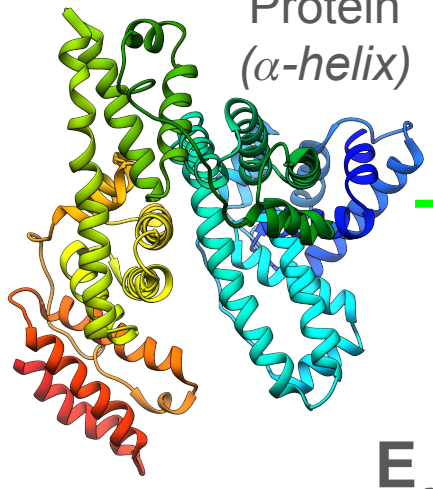
Development and application of parallel CPU/GPU tools to extend the characterization of direct and solvent-mediated H-bond interactions in large structural protein databases

Michele Massuoli (*ISASI-CNR, ICB-CNR*),
Rosa Maria Vitale (*ICB-CNR*), Melania Paturzo (*ISASI-CNR*),
Pietro Amodeo (*ICB-CNR*)

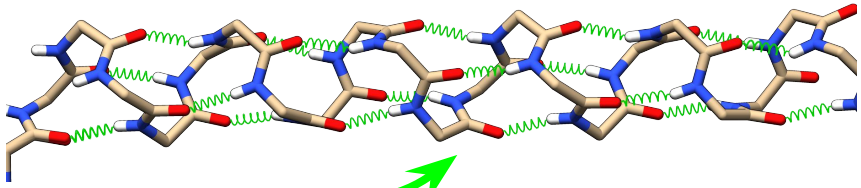
Consiglio Nazionale delle Ricerche

Istituto di Chimica Biomolecolare

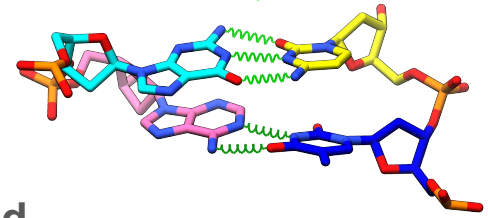
ISASI Institute of Applied Sciences and Intelligent Systems "Eduardo Caianiello"

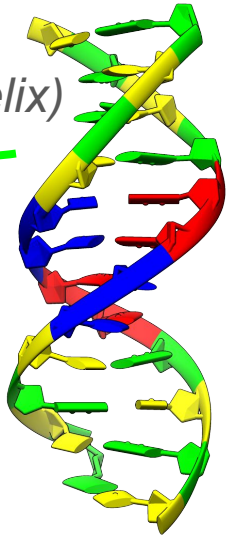


Protein
(α-helix)



Hydrogen bond
(H-bond)



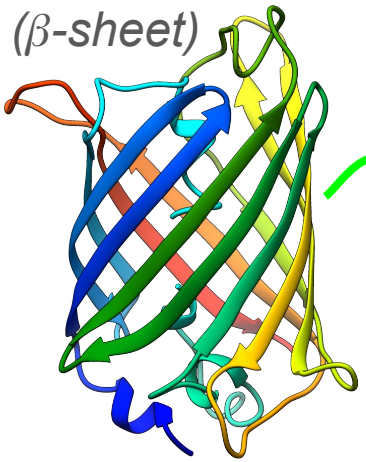


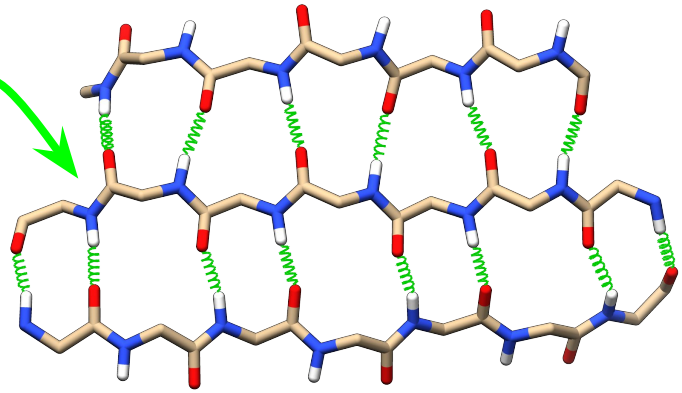
B-DNA
(double helix)

$$E_{\text{covalent bond}} = 10\text{-}50 \times E_{\text{H-bond}}$$

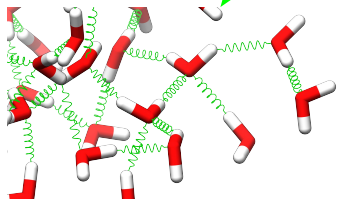
$$E_{\text{H-bond}} = 5\text{-}100 \times E_{\text{van der Waals}}$$

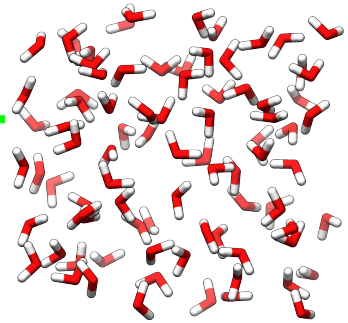
Protein
(β-sheet)





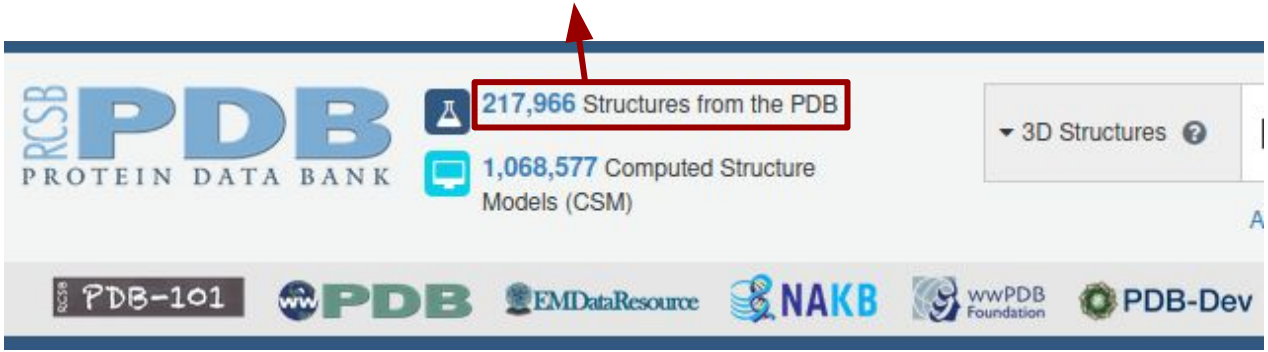
Water
(liquid)





Protein Data Bank

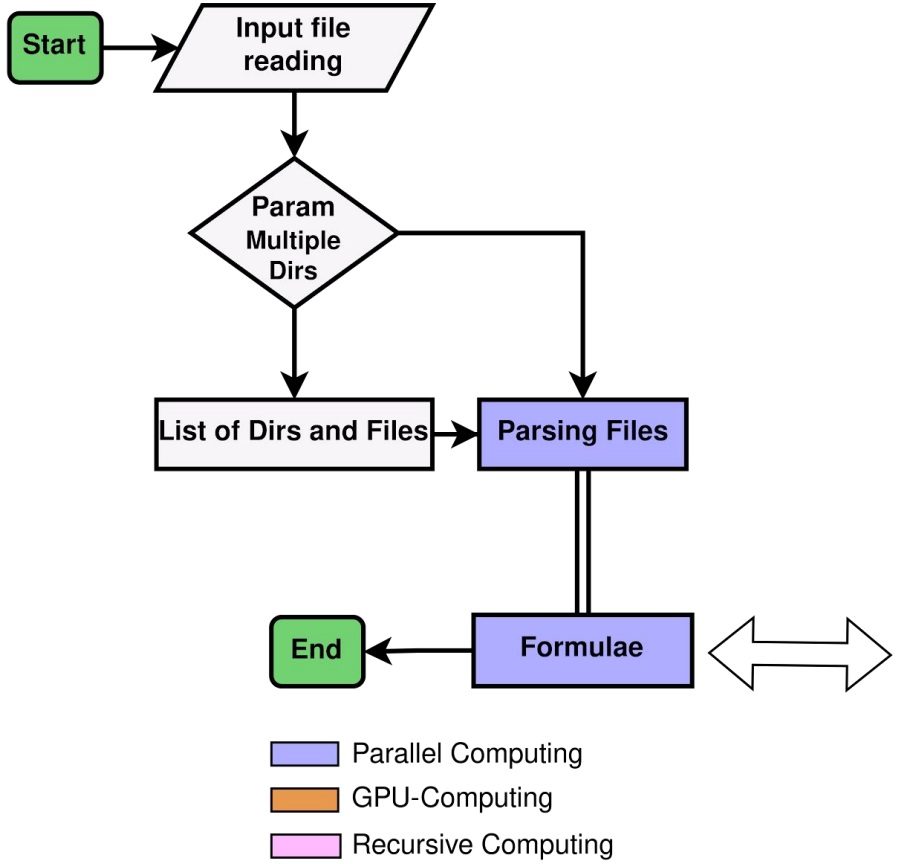
147803 are Protein Complexes

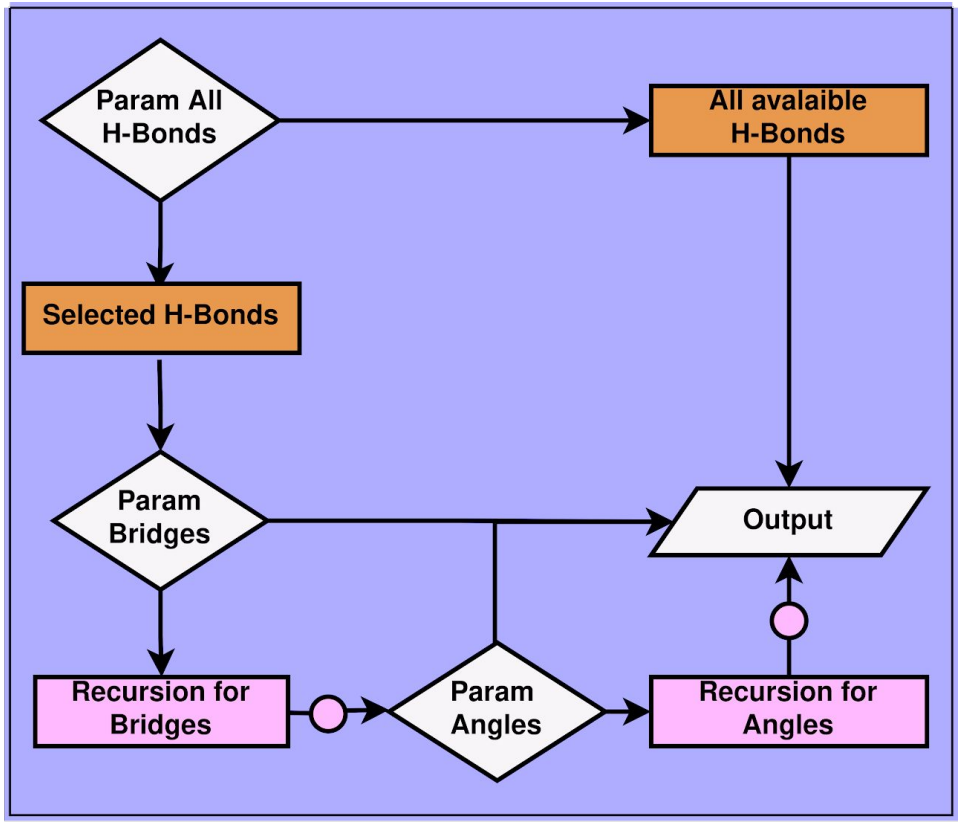


PDB Format Example

```
ATOM      1  N   GLY A   6      21.232  -3.592 -12.215  1.00 10.46           N
ANISOU    1  N   GLY A   6    1537   1300   1136    110   -528    -59       N
ATOM      2  CA  GLY A   6      21.163  -4.550 -11.066  1.00 10.57           C
ANISOU    2  CA  GLY A   6    1558   1396   1063   -267   -343     87       C
ATOM      3  C   GLY A   6      22.447  -4.422 -10.244  1.00  8.93           C
ANISOU    3  C   GLY A   6    1193   1290    908    -97   -117    152       C
ATOM      4  O   GLY A   6      23.428  -3.846 -10.715  1.00 10.86           O
ANISOU    4  O   GLY A   6    1267   1673   1187    -79   -127    346       O
ATOM      5  N   SER A   7      22.415  -4.967  -9.066  1.00  9.67           N
ANISOU    5  N   SER A   7    1214   1683    777   -108    -24     51       N
ATOM      6  CA  SER A   7      23.627  -4.941  -8.213  1.00 10.00           C
ANISOU    6  CA  SER A   7    1162   1784    855     89    -78     87       C
```

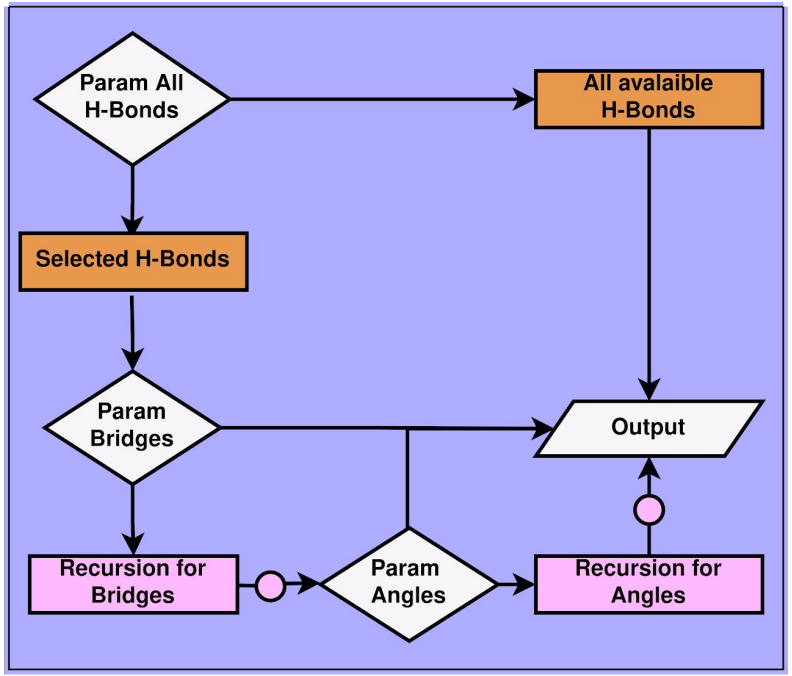
Algorithm flowchart





Formulae Block

Formulae and Input



Parallel Computing
GPU-Computing
Recursive Computing

Input.txt example

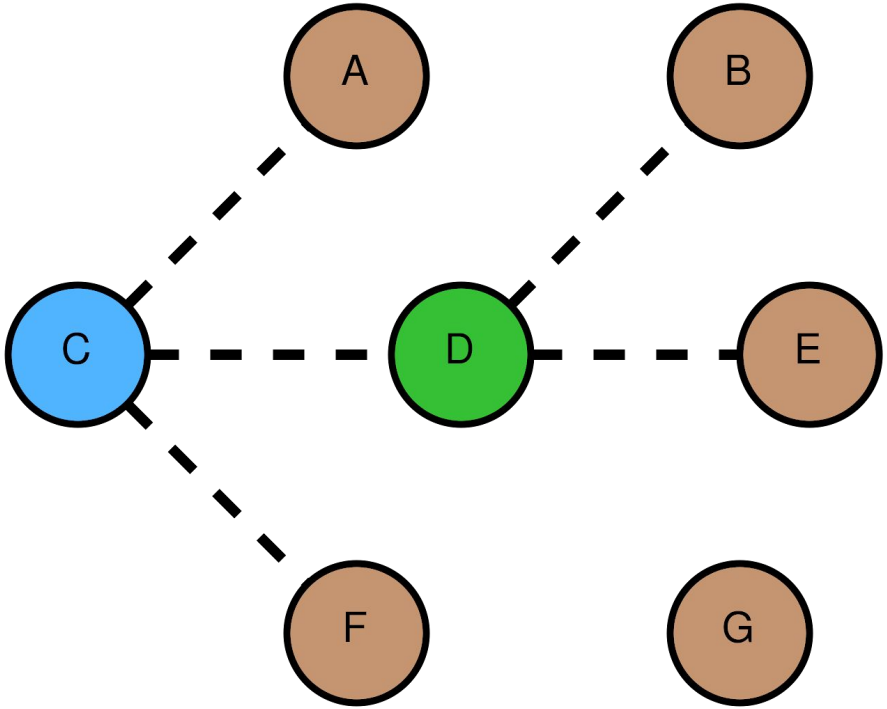
```
output_directory=           /home/user/Desktop
output_file_name=           IBiSCo_ws
output_extension=           all
input_resolution=           1.5
input_directory=            /home/user/Dataset
input_multi_dir=            True
input_multi_name=           DB_CIF_
input_multi_numb=           9
molecule_standard_res=      ARG(NE)
molecule_define_res_1=      HOH, O: don-acceptors
calculate_direct_hbonds=    ARG
calculate_bridges_hbond=    True
calculate_bridges_set=      3
calculate_angles=           True
calculate_angles_define_1= HOH(O) | ARG(NE), acc
calculate_angles_define_2= HOH(O) | ARG(NE), don
calculate_angles_define_3= ARG(NE) | acc, HOH(O)
calculate_angles_define_4= ARG(NE) | acc
```

Descriptors highlight

```
molecule_standard_res=          ARG(NE)
molecule_define_res_1=          HOH, O: don-acceptors
```

Bridging H-Bonds

```
C  molecule_standard_res=        ARG(NE)
D  molecule_define_res_1=        HOH, O: don-acceptors
```



Paths

Length 5
A - C - D - B - E
F - C - D - B - E

Length 4
A - C - D - B
A - C - D - E
F - C - D - B
F - C - D - E

Length 3
A - C - D
A - C - F
F - C - D
C - D - B
B - D - E
C - D - E

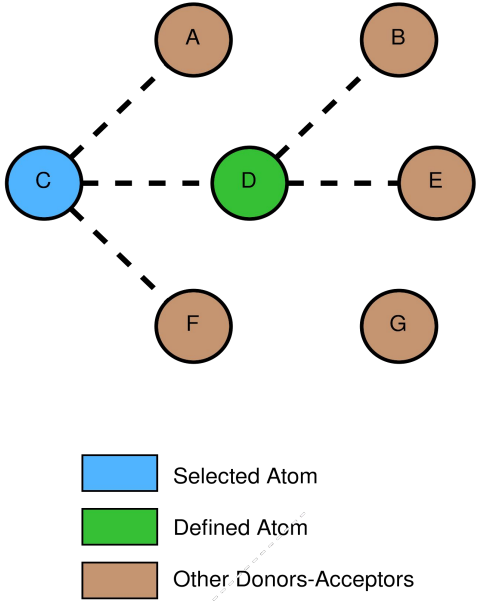
Selected Atom

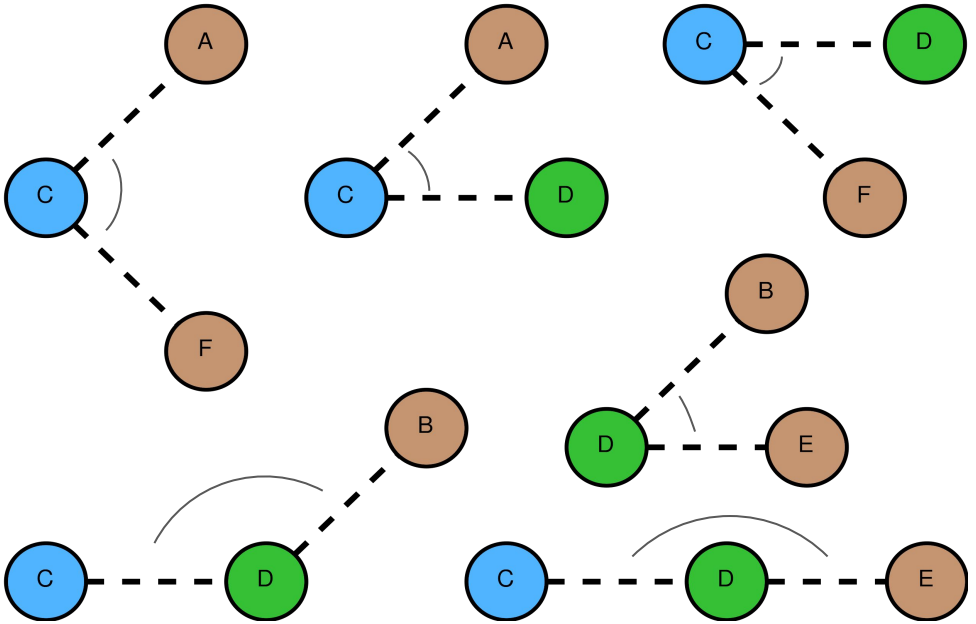
Defined Atom

Other Donors-Acceptors

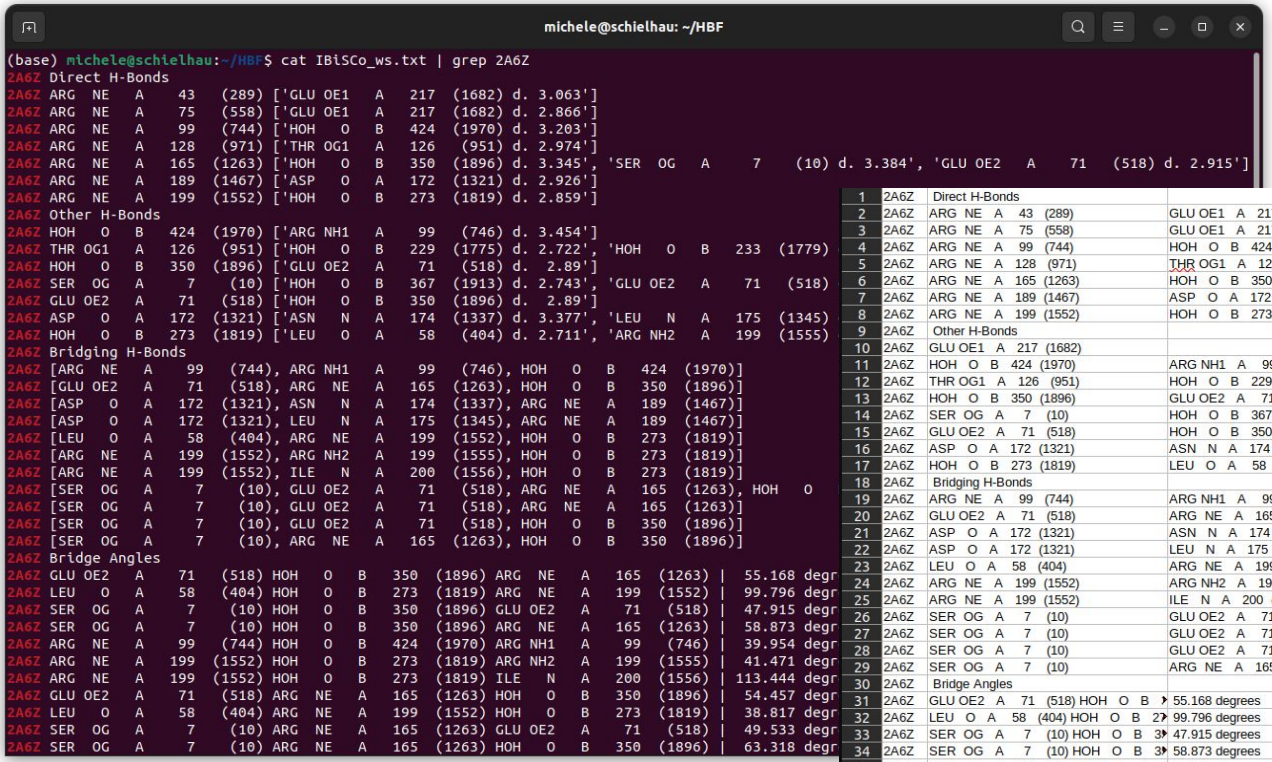
Bridge Angles

```
calculate_angles_define_1= HOH(O) | ARG(NE), acc
calculate_angles_define_2= HOH(O) | ARG(NE), don
calculate_angles_define_3= ARG(NE) | acc, HOH(O)
calculate_angles_define_4= ARG(NE) | acc
```





Output Examples



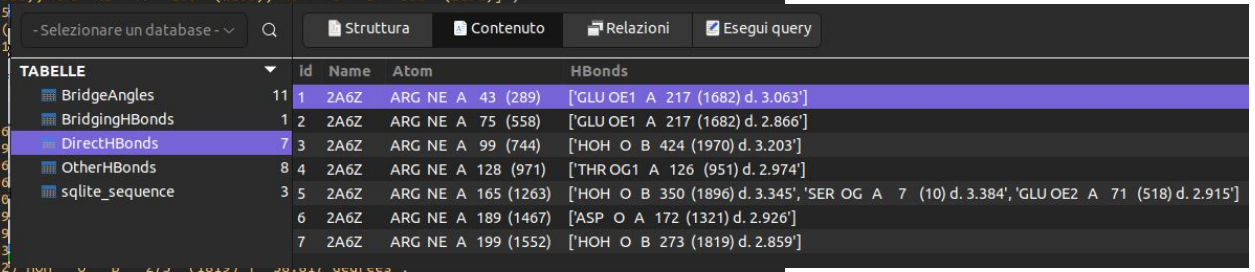
1	2A6Z	Direct H-Bonds			
2	2A6Z	ARG NE A 43 (289)	GLU OE1 A 217 (1682) d. 3.063		
3	2A6Z	ARG NE A 75 (558)	GLU OE1 A 217 (1682) d. 2.866		
4	2A6Z	ARG NE A 99 (744)	HOH O B 424 (1970) d. 3.203		
5	2A6Z	ARG NE A 128 (971)	THR OG1 A 126 (951) d. 2.974		
6	2A6Z	ARG NE A 165 (1263)	HOH O B 350 (1896) d. 3.345	SER OG A 7 (10) d. 3.384	GLU OE2 A 71 (518) d. 2.915
7	2A6Z	ARG NE A 189 (1467)	ASP O A 172 (1321) d. 2.926		
8	2A6Z	ARG NE A 199 (1552)	HOH O B 273 (1819) d. 2.859		
9	2A6Z	Other H-Bonds			
10	2A6Z	GLU OE1 A 217 (1682)			
11	2A6Z	HOH O B 424 (1970)	ARG NH1 A 99 (746) d. 3.454		
12	2A6Z	THR OG1 A 126 (951)	HOH O B 229 (1775) d. 2.722	HOH O B 233 (1779) d. 3.086	
13	2A6Z	HOH O B 350 (1896)	GLU OE2 A 71 (518) d. 2.89		
14	2A6Z	SER OG A 7 (10)	HOH O B 367 (1913) d. 2.743	GLU OE2 A 71 (518) d. 2.67	ARG NH2 A 165 (1266) d. 3.085
15	2A6Z	GLU OE2 A 71 (518)	HOH O B 350 (1896) d. 2.89		
16	2A6Z	ASP O A 172 (1321)	ASN N A 174 (1337) d. 3.377	LEU N A 175 (1345) d. 3.47	
17	2A6Z	HOH O B 273 (1819)	LEU O A 58 (404) d. 2.711	ARG NH2 A 199 (1555) d. 3.3	ILE N A 200 (1556) d. 3.127
18	2A6Z	Bridging H-Bonds			
19	2A6Z	ARG NE A 99 (744)	ARG NH1 A 99 (746)	HOH O B 424 (1970)	
20	2A6Z	GLU OE2 A 71 (518)	ARG NE A 165 (1263)	HOH O B 350 (1896)	
21	2A6Z	ASP O A 172 (1321)	ASN N A 174 (1337)	ARG NE A 189 (1467)	
22	2A6Z	ASP O A 172 (1321)	LEU N A 175 (1345)	ARG NE A 189 (1467)	
23	2A6Z	LEU O A 58 (404)	ARG NE A 199 (1552)	HOH O B 273 (1819)	
24	2A6Z	ARG NE A 199 (1552)	ARG NH2 A 199 (1555)	HOH O B 273 (1819)	
25	2A6Z	ARG NE A 199 (1552)	ILE N A 200 (1556)	HOH O B 273 (1819)	
26	2A6Z	SER OG A 7 (10)	GLU OE2 A 71 (518)	ARG NE A 165 (1263)	HOH O B 350 (1896)
27	2A6Z	SER OG A 7 (10)	GLU OE2 A 71 (518)	ARG NE A 165 (1263)	
28	2A6Z	SER OG A 7 (10)	GLU OE2 A 71 (518)	HOH O B 350 (1896)	
29	2A6Z	SER OG A 7 (10)	ARG NE A 165 (1263)	HOH O B 350 (1896)	
30	2A6Z	Bridge Angles			
31	2A6Z	GLU OE2 A 71 (518) HOH O B	55.168 degrees		
32	2A6Z	LEU O A 58 (404) HOH O B 27	99.796 degrees		
33	2A6Z	SER OG A 7 (10) HOH O B 3	47.915 degrees		
34	2A6Z	SER OG A 7 (10) HOH O B 3	58.873 degrees		
35	2A6Z	ARG NE A 99 (744) HOH O B 4	39.954 degrees		
36	2A6Z	ARG NE A 199 (1552) HOH O B	41.471 degrees		
37	2A6Z	ARG NE A 199 (1552) HOH O B	113.444 degrees		
38	2A6Z	GLU OE2 A 71 (518) ARG NE A	54.457 degrees		
39	2A6Z	LEU O A 58 (404) ARG NE A 1	38.817 degrees		
40	2A6Z	SER OG A 7 (10) ARG NE A 3	49.533 degrees		
41	2A6Z	SER OG A 7 (10) ARG NE A 3	63.318 degrees		

TXT - CSV

Output Examples

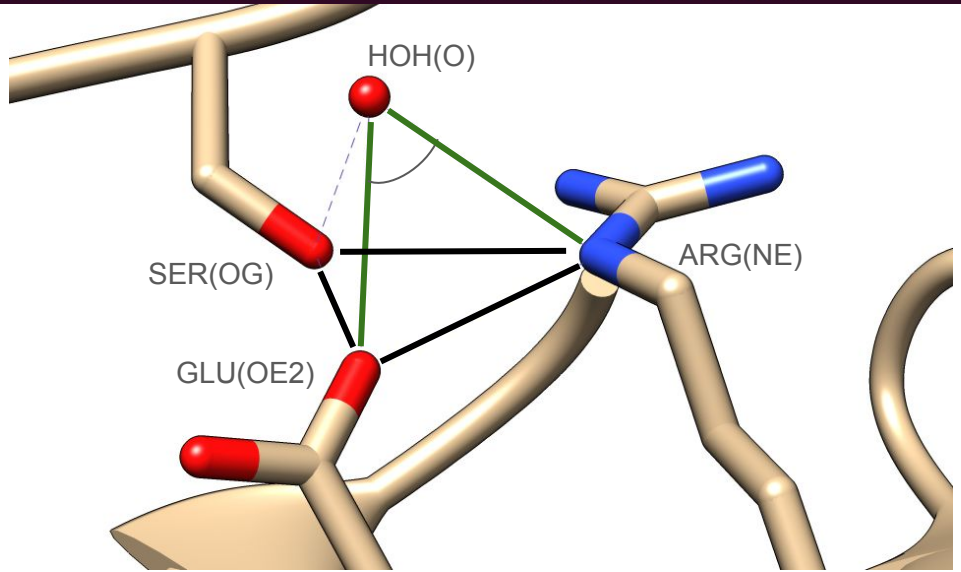

```
{
    "2A6Z":{
        "Direct H-Bonds":{
            "ARG  NE   A    43   (289)": "['GLU OE1   A   217  (1682) d. 3.063']",
            "ARG  NE   A    75   (558)": "['GLU OE1   A   217  (1682) d. 2.866']",
            "ARG  NE   A    99   (744)": "['HOH   O   B   424  (1970) d. 3.203']",
            "ARG  NE   A   128   (971)": "['THR OG1   A   126   (951) d. 2.974']",
            "ARG  NE   A   165  (1263)": "['HOH   O   B   350  (1896) d. 3.345', 'SER  OG   A     7    (10) d. 3.384', 'GLU OE2   A    71   (518) d. 2.915']",
            "ARG  NE   A   189  (1467)": "['ASP   O   A   172  (1321) d. 2.926']",
            "ARG  NE   A   199  (1552)": "['HOH   O   B   273  (1819) d. 2.859']"
        }
    },
    "2A6Z":{
        "Other H-Bonds":{
            "GLU OE1   A   217  (1682)": "[]",
            "HOH   O   B   424  (1970)": "['ARG NH1   A    99   (746) d. 3.454']",
            "THR OG1   A   126   (951)": "['HOH   O   B   229  (1775) d. 2.722', 'HOH   O   B   233  (1779) d. 3.086']",
            "HOH   O   B   350  (1896)": "['GLU OE2   A    71   (518) d.  2.89']",
            "SER  OG   A     7    (10)": "['HOH   O   B   367  (1913) d. 2.743', 'GLU OE2   A    71   (518) d. 2.673', 'ARG NH2   A   165  (1266) d. 3.085']",
            "GLU OE2   A    71   (518)": "['HOH   O   B   350  (1896) d.  2.89']",
            "ASP   O   A   172  (1321)": "['ASN   N   A   174  (1337) d. 3.377', 'LEU   N   A   175  (1345) d.  3.47']",
            "HOH   O   B   273  (1819)": "['LEU   O   A    58   (404) d. 2.711', 'ARG NH2   A   199  (1555) d. 3.395', 'ILE   N   A   200  (1556) d. 3.127']"
        }
    },
    "2A6Z":{
        "Bridging H-Bonds":{
            "1": "[ARG  NE   A    99   (744), ARG NH1   A    99   (746), HOH   O   B   424  (1970)]",
            "2": "[GLU OE2   A    71   (518), ARG  NE   A   165  (1263), HOH   O   B   350  (1896)]",
            "3": "[ASP   O   A   172  (1321), ASN   N   A   174  (1337), ARG  NE   A   189  (1467)]",
            "4": "[ASP   O   A   172  (1321), LEU   N   A   175  (1345), ARG  NE   A   189  (1467)]",
            "5": "[LEU   O   A    58   (404), ARG  NE   A   199  (1552), HOH   O   B   273  (1819)]",
            "6": "[ARG  NE   A   199  (1552), ARG NH2   A   199  (1555), HOH   O   B   273  (1819)]",
            "7": "[ARG  NE   A   199  (1552), ILE   N   A   200  (1556), HOH   O   B   273  (1819)]",
            "8": "[SER  OG   A     7    (10), GLU OE2   A    71   (518), ARG  NE   A   165  (1263), HOH   O   B   350  (1896)]",
            "9": "[SER  OG   A     7    (10), GLU OE2   A    71   (5",
            "10": "[SER  OG   A     7    (10), GLU OE2   A    71   (",
            "11": "[SER  OG   A     7    (10), ARG  NE   A   165  (1
        }
    },
    "2A6Z":{
        "Bridge Angles":{
            "1": "GLU OE2   A    71   (518) HOH   O   B   350  (1896
            "2": "LEU   O   A    58   (404) HOH   O   B   273  (1819
            "3": "SER  OG   A     7    (10) HOH   O   B   350  (1896
            "4": "SER  OG   A     7    (10) HOH   O   B   350  (1896
            "5": "ARG  NE   A    99   (744) HOH   O   B   424  (1970
            "6": "ARG  NE   A   199  (1552) HOH   O   B   273  (1819
            "7": "ARG  NE   A   199  (1552) HOH   O   B   273  (1819
            "8": "GLU OE2   A    71   (518) ARG  NE   A   165  (1263
            "9": "LEU   O   A    58   (404) ARG  NE   A   199  (1552) HOH   O   B   273  (1819) |  38.817 degrees",
            "10": "SER  OG   A     7    (10) ARG  NE   A   165  (1263) GLU OE2   A    71   (518) |  49.533 degrees",
            "11": "SER  OG   A     7    (10) ARG  NE   A   165  (1263) HOH   O   B   350  (1896) |  63.318 degrees"
        }
    },
```


JSON - DATABASE



Viewing Results

```
2A6Z ARG  NE   A   165  (1263) ['HOH   O   B   350  (1896) d. 3.345', 'SER  OG   A     7    (10) d. 3.384', 'GLU OE2   A    71   (518) d. 2.915']
2A6Z HOH   O   B   350  (1896) ['GLU OE2   A    71   (518) d.  2.89']
2A6Z SER  OG   A     7    (10) ['HOH   O   B   367  (1913) d. 2.743', 'GLU OE2   A    71   (518) d. 2.673', 'ARG NH2   A   165  (1266) d. 3.085']
2A6Z GLU OE2   A    71   (518) ['HOH   O   B   350  (1896) d.  2.89']
2A6Z [GLU OE2   A    71   (518), ARG  NE   A   165  (1263), HOH   O   B   350  (1896)]
2A6Z [SER  OG   A     7    (10), GLU OE2   A    71   (518), ARG  NE   A   165  (1263), HOH   O   B   350  (1896)]
2A6Z [SER  OG   A     7    (10), GLU OE2   A    71   (518), ARG  NE   A   165  (1263)]
2A6Z [SER  OG   A     7    (10), GLU OE2   A    71   (518), HOH   O   B   350  (1896)]
2A6Z [SER  OG   A     7    (10), ARG  NE   A   165  (1263), HOH   O   B   350  (1896)]
2A6Z GLU OE2   A    71   (518) HOH   O   B   350  (1896) ARG  NE   A   165  (1263) |  55.168 degrees
2A6Z SER  OG   A     7    (10) HOH   O   B   350  (1896) GLU OE2   A    71   (518) |  47.915 degrees
2A6Z SER  OG   A     7    (10) HOH   O   B   350  (1896) ARG  NE   A   165  (1263) |  58.873 degrees
2A6Z GLU OE2   A    71   (518) ARG  NE   A   165  (1263) HOH   O   B   350  (1896) |  54.457 degrees
2A6Z SER  OG   A     7    (10) ARG  NE   A   165  (1263) GLU OE2   A    71   (518) |  49.533 degrees
2A6Z SER  OG   A     7    (10) ARG  NE   A   165  (1263) HOH   O   B   350  (1896) |  63.318 degrees
```



2A6Z.pdb

Future Developments:

HBF Platform

- Multi-platform stand-alone program, downloadable, for calculations on local dataset
- Online service with textual input
- Online service with GUI
- Free access to pre-computed community datasets
- Development of further analyses

Thank you for your attention!

Consiglio Nazionale delle Ricerche
Istituto di Chimica Biomolecolare
ISASI Institute of Applied Sciences and Intelligent Systems "Eduardo Caianiello"