NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

School of Science

Information Technologies in Medicine and Biology

Direction: *Bioinformatics*

Algorithms in Structural Bioinformatics

Postgraduate Student: *Begetis Nikolaos*

Professor: *Chrysina Evangelia*

Deadline Date: *23/05/2013*

Assignment 4

1. Distances in the plane

1.1

In this fourth assignment we were asked as a first task to find the structural homologues of the enzyme with PDB code 1ctn, which as we saw corresponds to Chitinase and is found in A chain. In addition, we were asked to use either the PDBeFold tool of EMBL-EBI PDB, or the Dali server, or both of them, and write down the procedure we followed and the search parameters we used.

To begin with, in this task we chose to use the Dali server tool, clearly for personal academic reasons, so as to learn one more tool, as in the previous exercise we used PDBeFold. So, as it is shown in Figure 1, we searched for the enzyme 1ctn in the dali server.

Figure 1

The results returned from this search were similar to those we took in the previous assignment from the PDBeFold tool. Figure 2 shows a screenshot of the results. We also saved them in an HyperText format in file res1.htm provided with the submission.

Results

Parseable data
Matches to PDB90

- mol1A

Query: mol1A

MOLECULE: CHITINASE A;

Select neighbours (check boxes) for viewing as multiple structural alignment or 3D superimposition. The list of neighbours is sorted by Z-score. Similarities with a Z-score lower than 2 are spurious. Each neighbour has links to pairwise structural alignment with the query structure, to pre-computed structural neighbours in the Dali Database, and to the PDB format coordinate file where the neighbour is superimposed onto the query structure.

Structural Alignment | Expand gaps | 3D Superimposition (Jmol Applet) | Reset Selection

Summary

No:	Chain	Z	rmsd	lali	nres	%id	PDB	Description
1:	1ctn-A	71.7	0.0	538	538	100	PDB	MOLECULE: CHITINASE A;
2:	2wk2-A	68.6	0.5	538	539	99	PDB	MOLECULE: CHITINASE A;
3:	2wm0-A	68.6	0.6	538	538	99	PDB	MOLECULE: CHITINASE A;
4:	2wlz-A	68.6	0.5	538	538	99	PDB	MOLECULE: CHITINASE A;
5:	1eib-A	68.5	0.4	538	540	99	PDB	MOLECULE: CHITINASE A;
6:	2wly-A	68.4	0.4	538	539	99	PDB	MOLECULE: CHITINASE A;
7:	1ffr-A	68.2	0.4	538	540	99	PDB	MOLECULE: CHITINASE A;
8:	1ffq-A	68.1	0.5	538	540	99	PDB	MOLECULE: CHITINASE A;

Figure 2

From the results we selected the proteins with PDB codes “2WLY” and ”1ITX”, because these are the homologues structures that are found in different molecules.

1.2

Using the JMol tool and the PDB protein databank we concluded that the protein “2WLY” TIM barrel consists of 356 residues in 2 separate fragments. The first one has the Alanine amino-acid in both edges (131,442) and the second has the Alanine in the one edge (518) and the Valine in the other (562).

Similarly to the above, “1ITX” TIM barrel consists of 346 residues and its two fragments are: the first one, Leucine(33) – Proline(336), and the second one, Asparagine(410)-Leucine(451).

Now, in order to check possible domains that do not belong in the TIM barrel we found out, using JMol, that the enzyme “2WLY” consists also of two domains (one with 107 residues (positions 24-130) and the other with 76 residues (positions 443-517)), that are outside the “2WLY” TIM barrel. On the follow, when checking the “1ITX” enzyme we found out that it also has one domain consisting of 76 residues in positions 337-409.

1.3

In this part of the exercise, having done so before in a previous exercise, for the superposing of the enzyme in the same Cartesian system we did as follow:

We used the WinCoot molecular graphics application. At first we tried to use the option File→Open→OpenCoordinates, but the .pdb formatted files for an inquisitive reason could not be opened and so we fetched directly from the database the enzyme proteins we needed. Figure 3 shows how we did that.

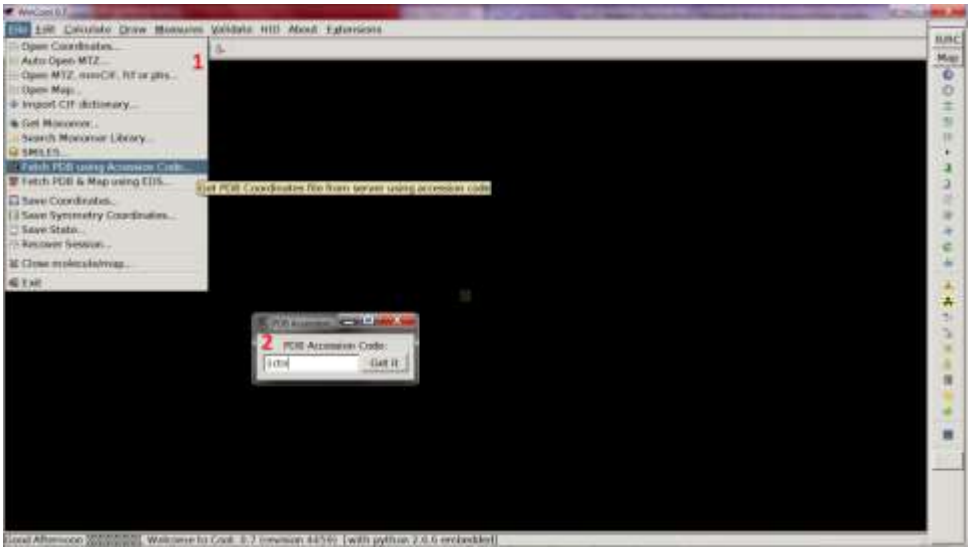

Figure 3

So fetching the pdb file coordinates from server using the accession code of the 1ctn chitinase enzyme protein resulted in Figure's 5 depiction.

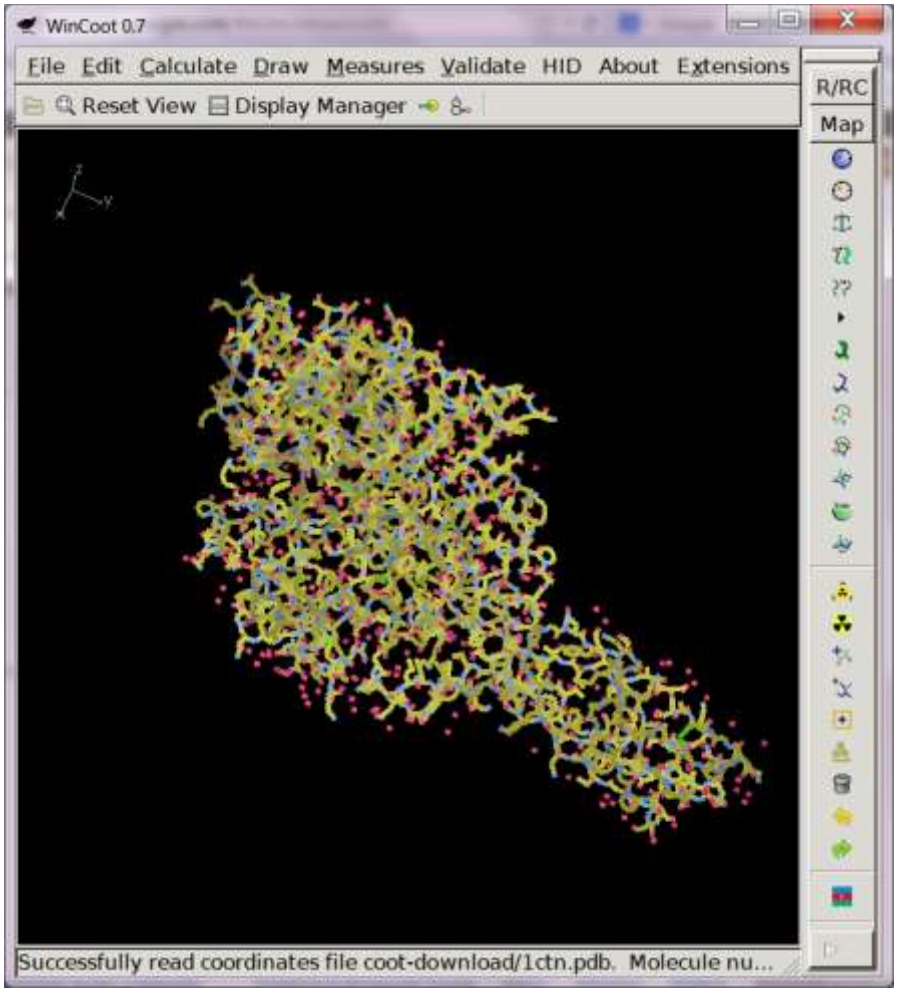

Figure 4

We followed the exact same procedure for the rest two homologue structures highest in score.

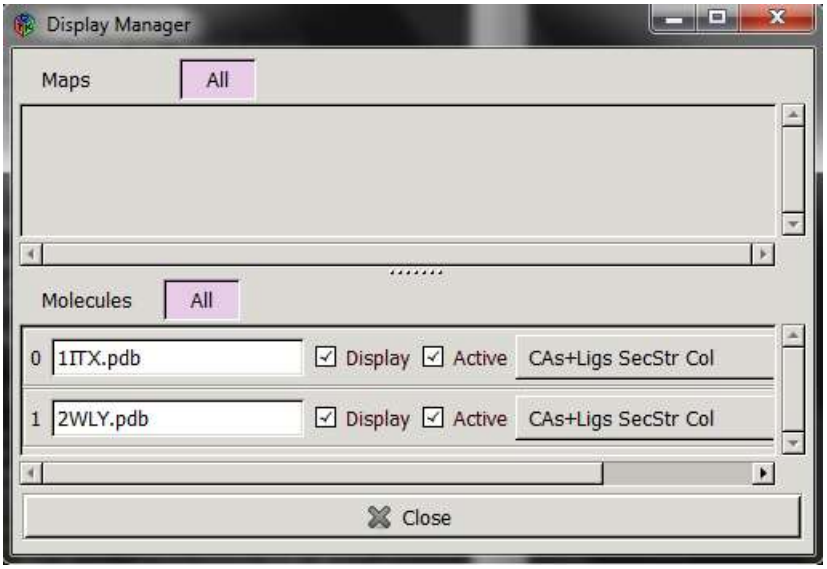

Figure 5

From this part of the 1st exercise and on we had gathered in a colleague's house and we sat together so that we do not make faults in our exercises. So, I do not provide the results and the figures again as they will be the same, having not forestalled to run the same tasks again for a second time (if I lose some points it doesn't matter).

2. Intervor

2.1

Initially, to begin with and state the 3 benefits of the interface model implemented by Intervor against the classical interface model based upon the buried surface area, and the associated dissection of atoms as core-rim, let us give some explanation at first for the classical interface model.

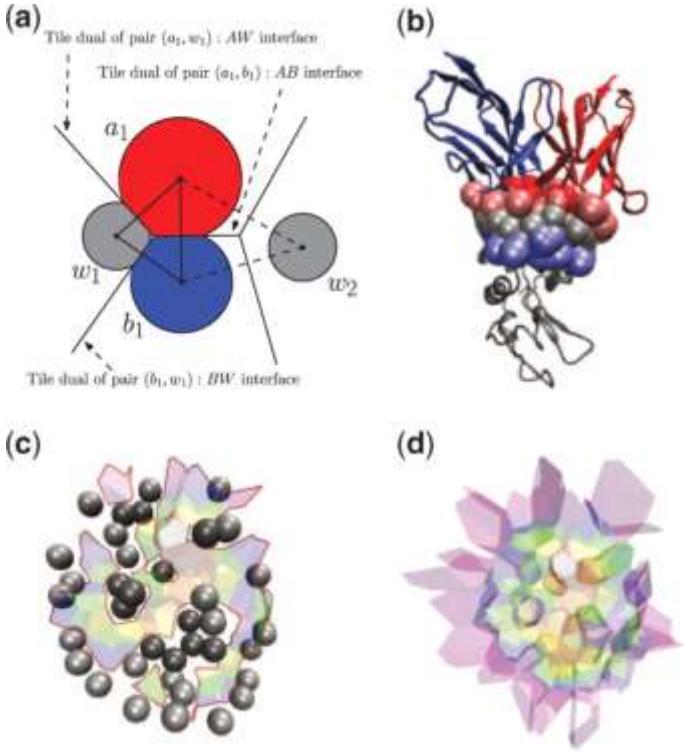

Figure 4.2.1

So, (a) in the classical interface modeling, two interface models are traditionally used. The first is the geometric footprint based model, which consists in considering all pairs of atoms (one on each partner) within some distance threshold; and the second is the solvent accessible model, which selects atoms losing solvent accessibility upon complex formation. While the former tends to overestimate the interface size, the opposite holds for the latter.
This motivates the Voronoi interface model.
In addition to the above, (b) according to the Voronoi interface model, we talked in the previous paragraph, this model functions with direct instead of water-mediated contacts. So, for instance, consider the power diagram of the atomic balls expanded by a water probe which is called as the Voronoi diagram in the sequel. At follow, let a restricted ball or restriction be the intersection of this ball with its Voronoi region, see e.g. the red ball on Figure 4.2.1a, and also, let an interface water molecule be a water molecule such that its restriction has neighboring restrictions of type A and B. Water molecules which are not at

the interface are called bulk water molecules. Now, our interface model consists of pairs of restrictions of type [A; B] or [A; W] or [B; W], with W standing for interface water molecules only. Note that every such pair conveys two pieces of information: (i) the two atoms in contact, which define an interface edge, and (ii) the Voronoi facet, also called tile, dual of this edge(this paragraph will help as to solve the next exercise task too). All atoms involved in at least one pair define the interface atoms, as illustrated on Figure 4.2.1b. Tiles of type AB (AW and BW) define the interface AB (AW and BW, respectively). Tiles of type AB define direct contacts between the partners, while tiles of type AW and BW define contacts between these partners which are mediated by interface water molecules. The union of tiles AW and BW defines the AW−BW interface. The union of the AW−BW and AB interfaces defines the ABW interface, which separates the partners and gives a global overview of the interaction area, regardless of the role played by water molecules. See Figure 4.2.1c and 4.2.1d.
Finally, to close the section of the classical interface model, let us to introduce the shelling of the Voronoi interface. Think of a room whose floor is tiled, the room being delimited by walls. Assign the integer one to all tiles in contact with a wall. Having removed these tagged tiles, move the walls toward the tiles which were in contact with them, and tag these with the integer two. Iterating the process until completion results in a discrete numbering measuring the minimum distance from a tile to the boundary of the room. This process is carried out for the tiles of the ABW interface, starting from its boundary, the resulting number being called the Voronoi Shelling Order or VSO (Fig. 4.2.1d). Since an interface atom contributes at least one tile, the VSO of this atom is defined as the minimum of the values found on the tiles contributed by this atom.

So, having given the above detailed explanation about the classical interface model, it is time to answer the task's demands. The 3 big benefits of the Intervor against the Voronoi interfaces when given two interacting partners, possibly with water molecules squeezed in-between them, are the following:

i) Intervor computes an interface model which identifies the atoms of the partners which are in direct contact and those whose interaction is water mediated, while the classical interface model is not so flexible
ii) Intervor, defines a geometric complex separating the partners, the Voronoi interface, whose geometric and topological descriptions are straightforward (surface area, number of patches, curvature), while the classical interface model deals only with geometric parameters and not topological, and
iii) Intervor, allows the definition of the depth of atoms at the interface, thus going beyond the traditional dissection of an interface into a core and a rim. These features can be used to investigate correlations between structural parameters and key properties such as the conservation of residues, their polarity, the water dynamics at the interface, mutagenesis data, etc. Instead, classical interface models have not equivalent functionality and are more restricted.

2.2

For this second task of this exercise, we downloaded the software of Intervor, described in the paper of S.Loriot and F.Cazals: Modeling macro–molecular interfaces with Intervor, from

the Intervor website: http://cgal.inria.fr/abs/Intervor/. We used both the webserver and the Linux executables, to do our calculations and verify the results of the webserver and the results output from our personal computer.
It is worthy to mention that all the results are provided in the deliverable, accompanied with all the 'sorted by name' screenshot-images taken from both the webserver and the personal computer. To visualize and understand the results we used the VMD tool (PyMol could also be used).
So, the screenshot are also provided in the Miscellaneous part if you want to have a look. The important, thus, that resulted from the outcome is stated in the following paragraphs.

To begin with, the center of all this task is to describe the changes that may be observed in the AB interface and AW-BW interfaces, upon filtering out selected water molecules (as we did, see Miscellaneous chapter) and allowing the filtering by temperature factors.

Initially, we should begin by annotating that the AB interface is the interface specified by edges of type AB, and that this interface describes the contacts between both partners. Also, AW-BW interface, is the interface specified by edges of type AW or BW, which also describes the contacts between the partners and structural water.

Moreover, from the bibliography, we were informed that, in particular, the ABW model is highly relevant for high resolution crystals, with a resolution better than 2 Angstrom, where water molecules are spotted reliably. In respect and in addition to that, if one may wish to retain all such molecules, there should be a relatively low temperature factor in order to have reliable results.

In respect to the above, using the output files from the AB and AW-BW interfaces we compared the relative to each other files and we found out that all atoms which are not inserted into the Delaunay triangulation, are skipped, and also Hydrogens atoms are always discarded. In the AB model, all water molecules (i.e. oxygens) are discarded. In the AW-BW model, water molecules whose temperature factor is beyond a threshold are discarded when the threshold value is set to 80 Angstroms (default). It is worthy to mention that Atoms discarded are also dumped in PDB format.

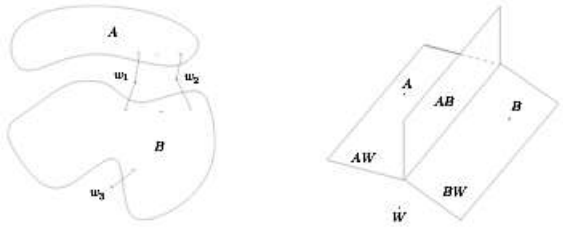

Figure 4.2.2: Left: Water molecules centered at w_1 and w_2 are interface water molecules. The water molecule w_3 is not. Right: the boundary of the union of the AB and AW-BW interfaces may not be one manifold.

What is more, from the tests we implemented and the related bibliography, we can infer that when considering an interface, an interesting question is the role played by structural water molecules. As these water molecules are described from the crystal as protein atoms are, we also expand their radius by a measured quantity. Notice again that this expansion aims at mimicking an implicit continuous layer of solvent molecules on the atoms found in the crystal be they protein atoms of water molecules. So, when having three molecular types A, B and W there can be defined three types of bicolor interfaces. But since we

primarily care for the AB interface, contact of type AW and BW are of interest only when located near the AB interface, see Fig. 4.2.2.
Taking a deeper sight in this problem, the AW-BW interface is the collection of Voronoi facets dual of edges of type AW or BW. A connected component of AW-BW interface is a collection of edge-connected Voronoi facets dual of interface edges of type AW or BW. So, to study the AW or the BW or the AW-BW interface, one should observe that edge connected Voronoi facets of types AW and BW are defined from bicolor, tricolor or quadricolor tetrahedra. Let us analyze the last two cases. In such a tetrahedron, we identify the labels A and B, since we do not report facets dual of such edges, so that the congurations found are those of bicolor tetrahedra. More precisely, a tetrahedron of type AABW where AB edges are omitted yields the same topological congurations as a bicolor (3, 1) tetrahedron for any bicolor interface. A tricolor tetrahedron of type ABWW is similar to a bicolor (2,2) tetrahedron for any bicolor interface. Finally, a ABWX tetrahedron is equivalent to a (2,1,1) tetrahedron for any bicolor interface.

Therefore, finally, the AW, the BW, and the AW-BW interfaces have the same topological properties as the AB interface i.e. are surfaces with possibly pinched vertices.

2.3

As described in the first task of this exercise, two of the benefits of the interface model implemented by Intervor over the classical interface model based upon the buried surface area, and the associated dissection of the atoms as core-rim are that

i) Intervor, defines a geometric complex separating the partners, the Voronoi interface, whose geometric and topological descriptions are straightforward (surface area, number of patches, **curvature**), while the classical interface model deals only with geometric parameters and not topological, and

ii) Intervor, allows the definition of the depth of atoms at the interface, thus going beyond the traditional dissection of an interface into a core and a rim. These features can be used to investigate correlations between structural parameters and key properties such as the conservation of residues, their polarity, the water dynamics at the interface, mutagenesis data, etc. Instead, classical interface models have not equivalent functionality and are more restricted.

Taking the above under consideration, at first we will provide a figure (Figure 4.2.3 (a)) from the paper “Geometric, topological and contact analysis of interfaces in macro-molecular complexes: from the atomic scale to the complex scale using Intervor”[F.Cazals]

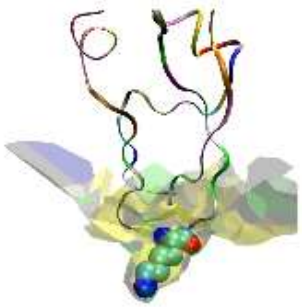

Figure 9: Interface of complex Kallikrein - pancreatic trypsin inhibitor (2kai) in the AB model. The inhibitor is drawn as a ribbon, and its Lys 15 drawn in VdW mode. The concave active site of the protease accommodates Lys 15 of the inhibitor. The cylindrical geometry of the active site is reflected by a high *discrete mean curvature*. Equivalently, an atom of the inhibitor has more than twice more neighbors than an atom of the protease.

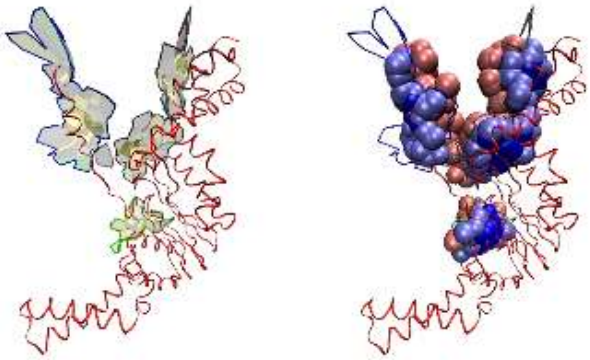

Figure 10: Interface patches of complex E. Coli EF-Tu/Ts (1efu) in the AB model. EF-Ts is shown as a ribbon (a)The interface splits into three main connected components whose boundaries are depicted in green, blue, brown (b)Interface atoms for each connected component. The blue/red atoms respectively correspond to EF-Tu and EF-Ts.

Figure 4.2.3 (a)

For instance from the above explanations given in these figures captions one can infer that even that with the classical interface model it was not possible to accommodate alternate locations of atoms, due to the Intervor, this became possible.

The function for accommodating alternate locations of atoms, if any in the PDB file is not provided from the webserver, so we had to use this option from our personal computer. Below Figure 4.2.3 (b) shows the option *–A [--alternate] +argument* provided in the help option of the Intervor, which alternates the coordinates to be used. After that Figure 4.2.3 (c) shows an example execution with keeping atoms with alternate location 2.

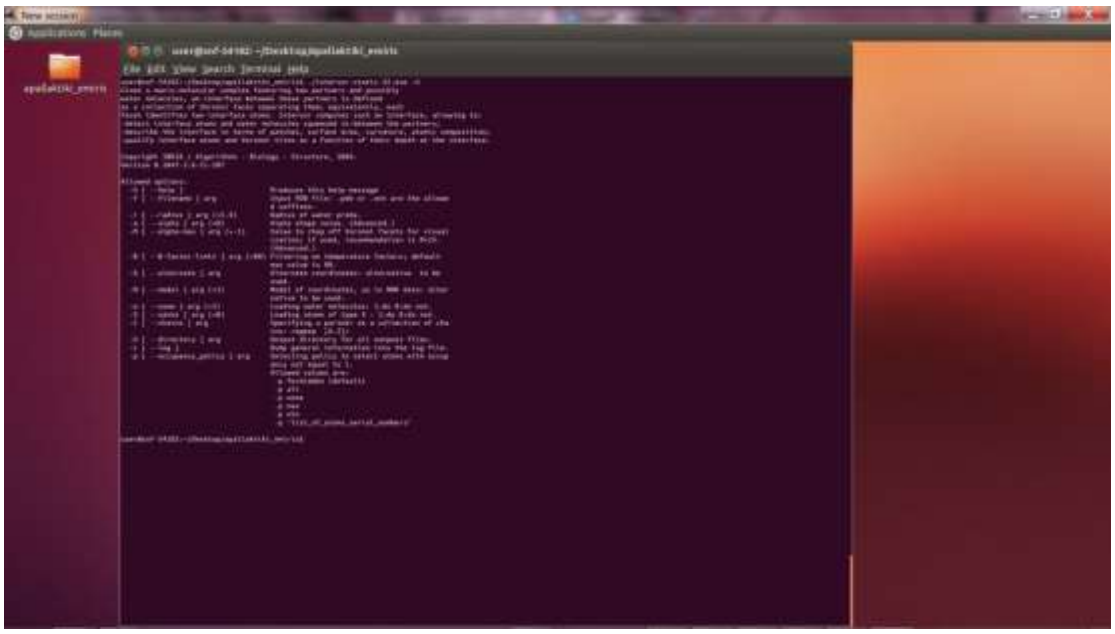

Figure 4.2.3 (b)

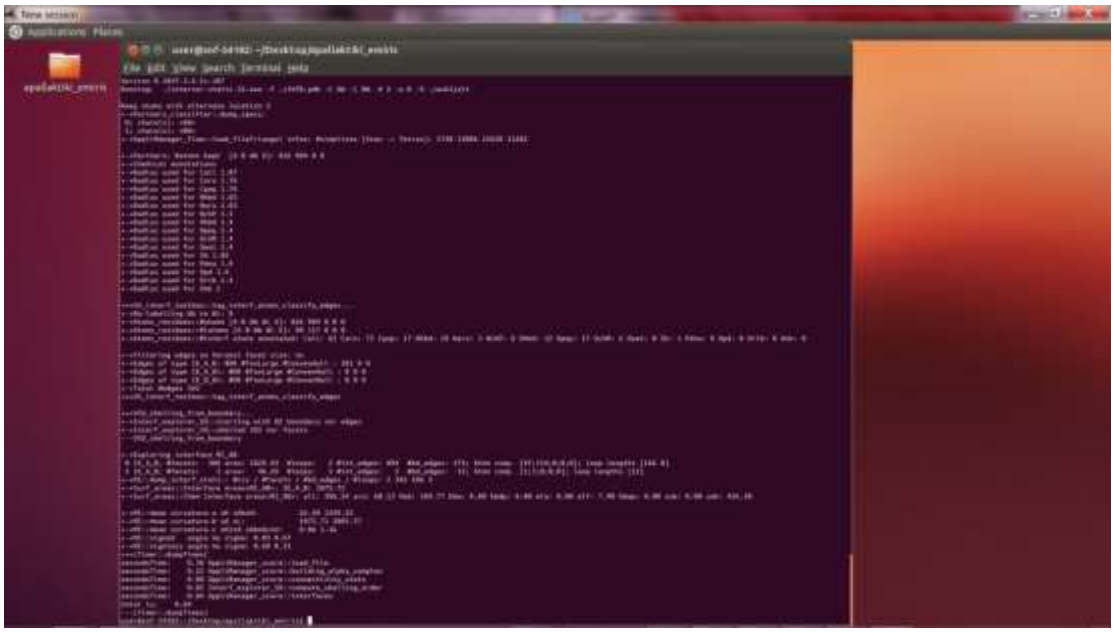

Figure 4.2.3 (c)

Now we can go to the new output folder named “alt” as shown in the above figure’s query and then run VMD or PyMol to see possible differencies.

3. Miscellaneous

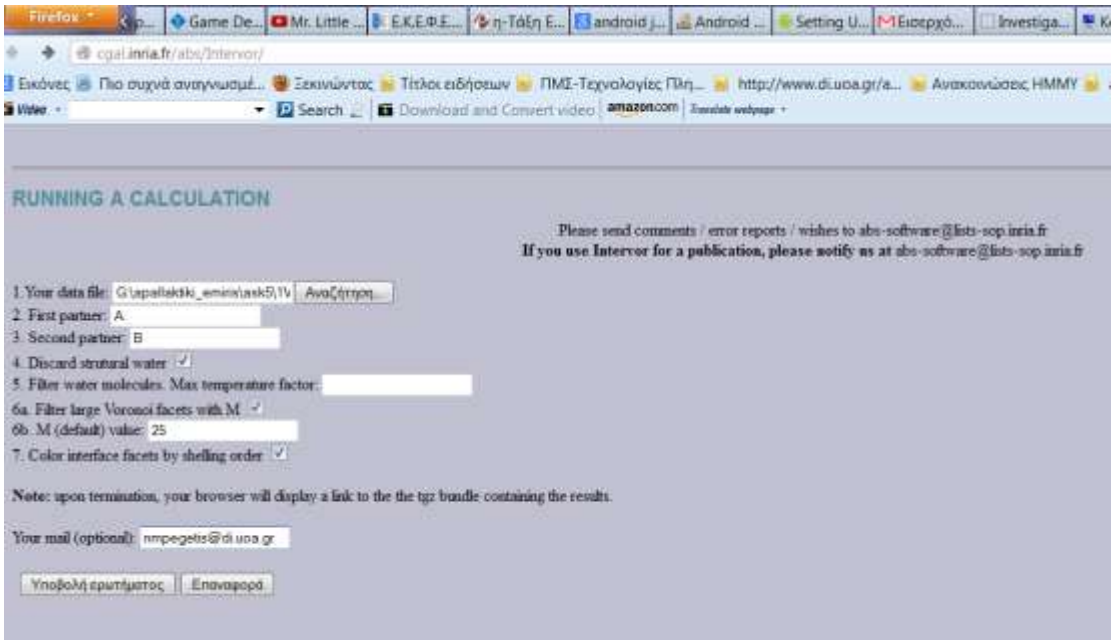

Figure 3.3: Querying AB interface, waters are discarded and temperature factor is not set.

Investigating Interfaces with Intervor

clientfile: 1VFB.pdb **partnerA**: A **partnerB**: B **water_01**: on **MFilter_yn**: on **MFilter_value**: 25 **color_facets_by_so**: on **yourmail**: mpegetis@di.uoa.gr

Attention: directory 1VFB is there. I clean it up.
Calling intervor as: intervor.exe -f data/1VFB/1VFB.pdb -C A -C B -w 0 -D data/1VFB --log

Results here.

Figure 3.2: Result of Figure's 3.1. query

Figure 3.3: Querying AW-BW interface, waters are discarded and temperature factor is not set.

Investigating Interfaces with Intervor

clientfile: 1VFB.pdb **partnerA**: AW **partnerB**: BW **water_01**: on **MFilter_yn**: on **MFilter_value**: 25 **color_facets_by_so**: on **yourmail**: mpegetis@di.uoa.gr

Attention: directory 1VFB is there. I clean it up.
Calling intervor as: intervor.exe -f data/1VFB/1VFB.pdb -C AW -C BW -w 0 -D data/1VFB --log

Results here.

Figure 3.4: Result of Figure's 3.3 query

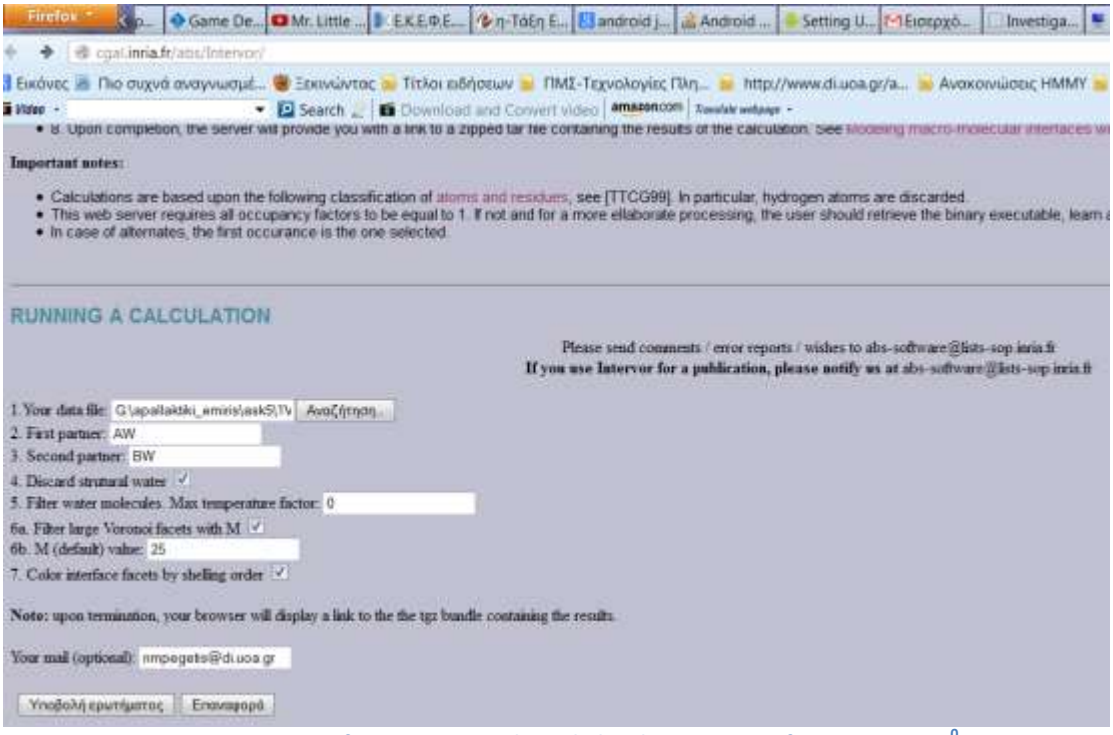

Figure 3.5: Querying AW-BW interface, waters are discarded and temperature factor is set to 0 ^{0}C.

Figure 3.6: Result of Figure's 3.5 query.

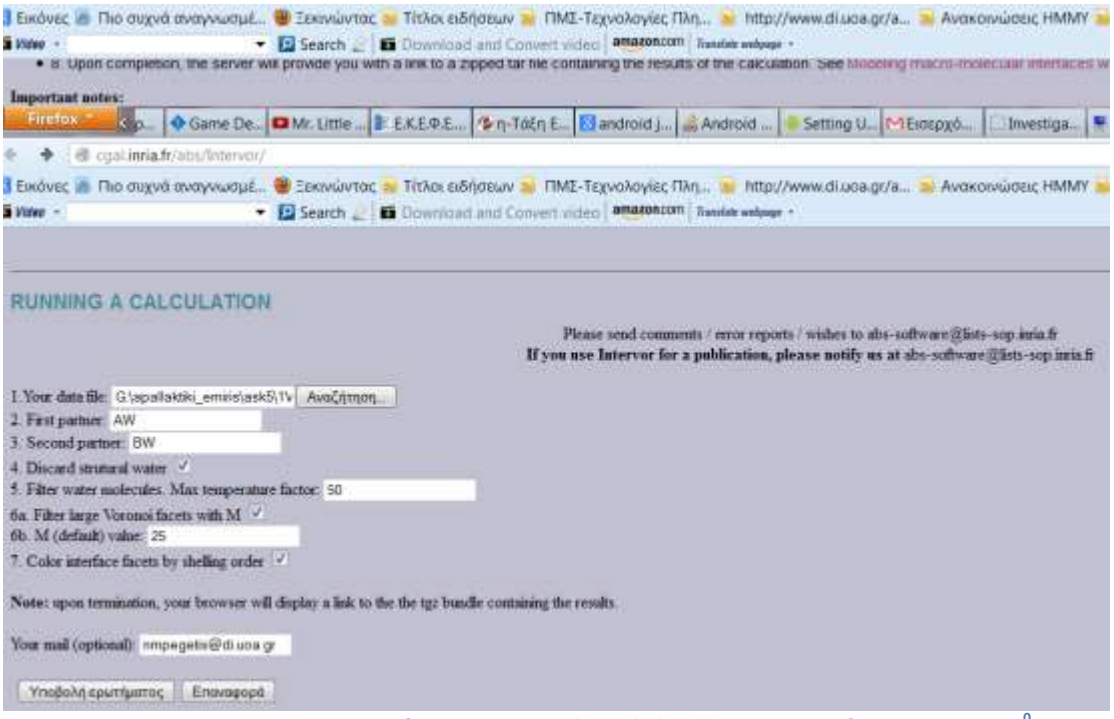

Figure 3.7: Querying AW-BW interface, waters are discarded and temperature factor is set to 50 ^{0}C.

Figure 3.8: Result of Figure's 3.7 query.

Figure 3.9: Querying AW-BW interface, waters are discarded and temperature factor is set to 100 ^{0}C.

Investigating Interfaces with Intervor

cleanfile: 1VFB.pdb **partnerA**: AW **partnerB**: BW **water_01**: on **TempFact_value**: 100 **MFilter_yn**: on **MFilter_value**: 25 **color_facets_by_so**: on **yourmail**: mpegetis@di.uoa.gr

Attention: directory 1VFB is there. I clean it up.
Calling intervor as: intervor.exe -f data/1VFB/1VFB.pdb -C AW -C BW --w 0 -D data/1VFB --log

Results here

Figure 3.10: Result of Figure's 3.9 query.

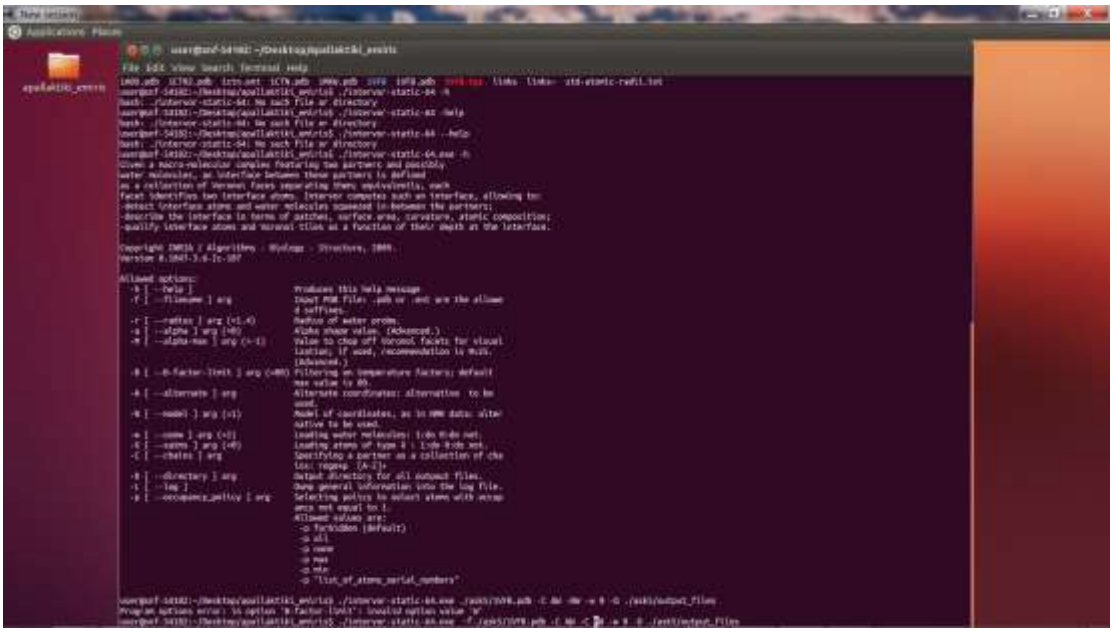

Figure 3.11: Querying AW-BW interface, waters are discarded and temperature factor is not set using personal computer. On the top the executable's help is provided and at the bottom the query is shown.

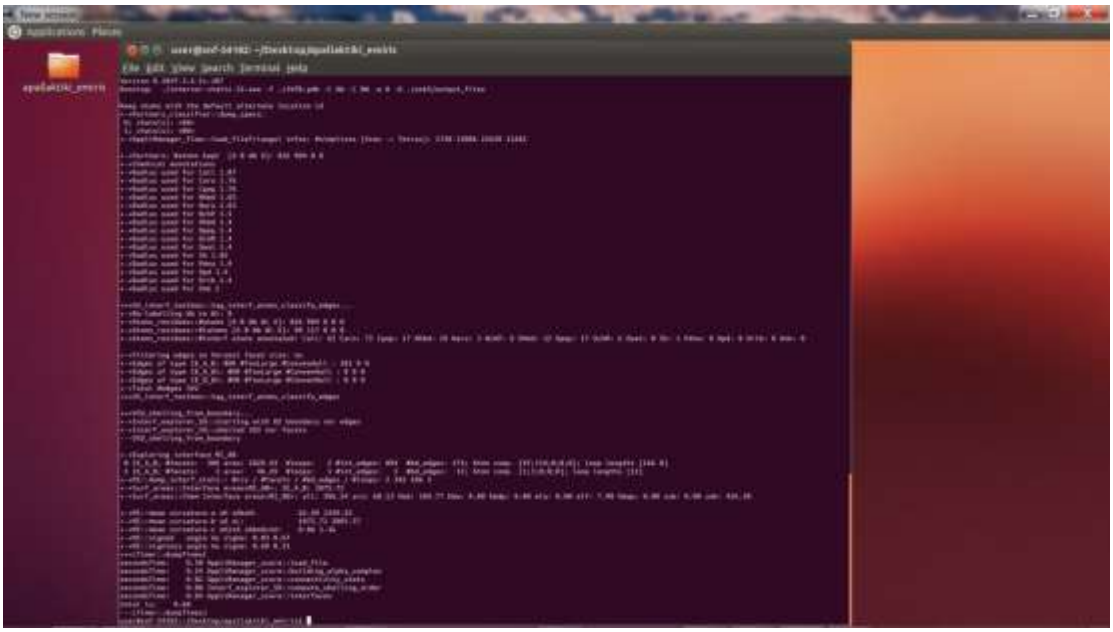

Figure 3.12: Results of the Figure's 3.13 query.

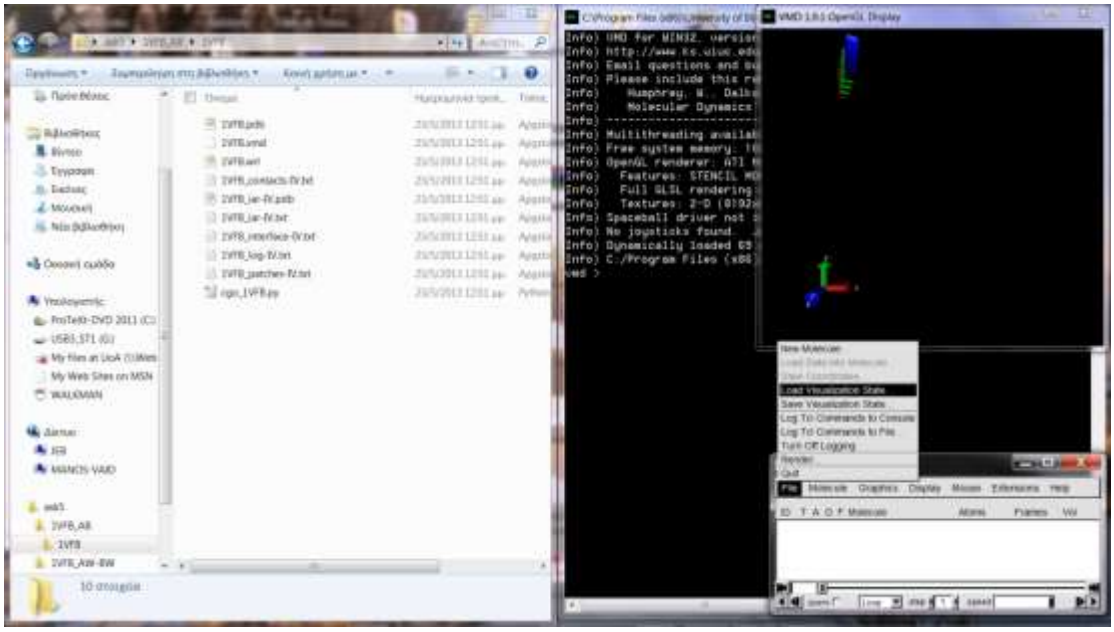

Figure 3.13: Loading in VMD visualization tool the results provided from each of the above queries.

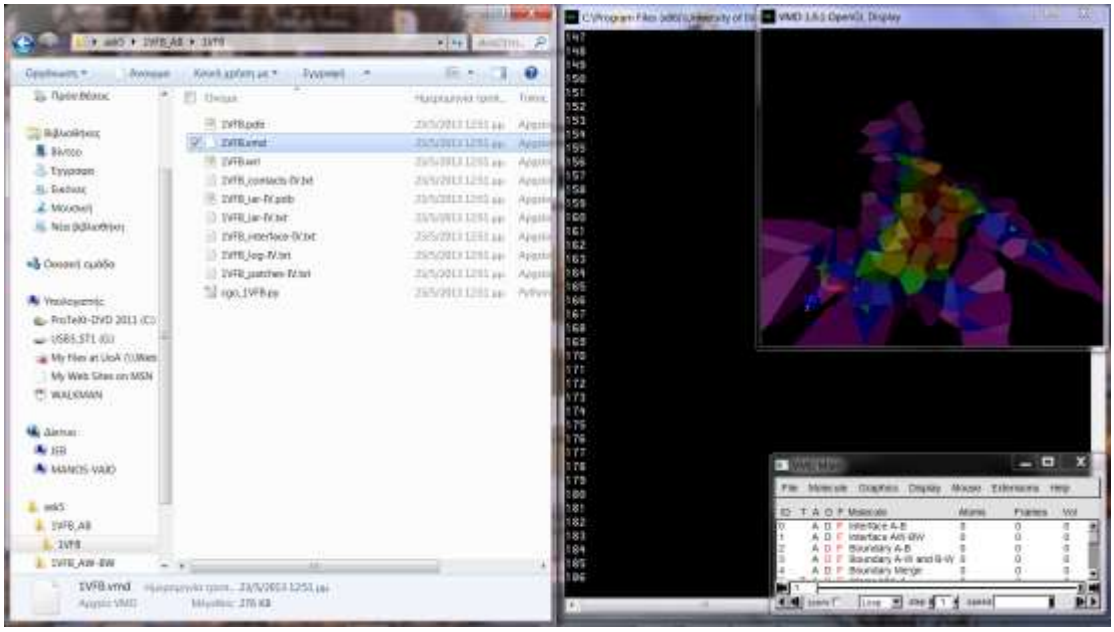

Figure 3.14: Example output visualization. All functions can be found in the top menu of VMD main, while in the 3D modeling there can be easily seen the different interfaces, colored in different colors