Journal of
Proteomics & Bioinformatics

Shobana et al., J Proteomics Bioinform 2012, 5:9
DOI: 10.4172/jpb.1000240

Research Article Open Access

Evolutionary Trace Analysis of Azoreductase at the Ligand Binding Site and Enhancing the Active Site through Site Directed Mutagenesis

Shobana S[1]*, Berla Thangam E[2] and Kasmir Raja SV[3]

[1]Department of Bioinformatics, SRM University, Kattankulathur, Chennai- 603203, India
[2]Department of Biotechnology, SRM University, Kattankulathur, Chennai- 603203, India
[3]Dean, Research, SRM University, Kattankulathur, Chennai- 603203, India

Abstract

Azoreductases belong to the family of oxidoreductases and are characterized with detoxification of azo dyes. Azoreductases are involved in bioremediation, to eliminate the color effluents found in waste waters. In this study, the effectiveness of azoreductase in bioremediation was analysed, and the aim was to find a way to increase the efficiency of the enzyme towards toxic compound degradation, more rapidly. To estimate this, the structure function relationship was determined by evolutionary trace (ET) analysis at ligand binding site. The class specific site of ET analysis was mapped with the crystal structure of azoreductase (1NNI). Site directed mutagenesis was performed at the residues, in very close proximity to the active site residues, hydrogen bonded to ligand. Based on the work done, it was found that Glycine at 106th position plays a crucial role in enzyme activity.

Keywords: Evolutionary trace (ET) analysis; Azoreductase; Detoxification; Ligand binding site; Bioremediation

Introduction

The ease of synthesis and chemical stability of azo dyes possessing one or more azo groups helps them to be widely used in textile, printing, cosmetics, pharmaceutical, food and many other industries [1]. Azoreductase catalyses the reductive cleavage of the azo bond, and acts as the key enzyme expressed in all azo dye degrading bacteria [2]. Under anaerobic conditions, the reductive cleavage of azo dyes results in decolorization, which was suggested in many reports [3]. Many bacterial strains contain cytoplasmic enzymes, though unspecific, act as azoreductase and under anaerobic conditions; transfer the electrons to azo dyes via soluble flavins. *Bacillus sp OY12* transforms azo dyes into colorless compounds, and this is mediated by azoreductase activity, in the presence of NADH [4]. *Bacillus subtilis* is considered to be a widely spread bacterium, commonly recovered from water, soil, air and decomposing plant residue. This bacterium produces an endospore, which allows it to overcome extreme conditions of heat and drying up in the environment. *B. subtilis* produces a variety of enzymes that will degrade a variety of natural substrates and contribute to nutrient cycling, and under most conditions but exists in the form of spore [5]. Site-directed mutagenesis is of potential interest for engineering the enzyme to become active so that it is more suitable for biotechnical applications [6]. In this protein (1NNI), four residues namely Asp (33), Tyr (74), His (75), and Lys (83) were identified as catalytic residues, where Tyr (74) and Lys (83) play a crucial role in azobenzene reductase extracted from *Bacillus subtilis* [7]. Olivier Lichtarge in 1996 developed the evolutionary trace (ET) analysis [8]. This method depends on both sequence and structural information in order to analyze functional sites of a protein or group of proteins. It determines the conserved amino acid residues in an alignment, and the information is mapped onto known 3D protein structures. This method gives importance to the residues which are important to the structure or function of a protein and tends to be conserved across species. This method gives ranking of the relative amino acids functional importance in a protein sequence, by correlating their variations during evolution, with divergences in the phylogenetic tree of that sequence family. The best-ranked residues are then clustered spatially in the protein structure and thereby reveal the location of functional sites [9].

Methodology

Dataset

Azoreductase from *Bacillus subtilis* with accession number O07529 was obtained from SWISSPROT database and homologous sequences of the enzyme are retrieved [10]. BLASTP is performed and twenty six protein sequences were more than of 50% identity [11]. All the sequences are aligned together with ClustalW [12], using the Gonnet protein weight matrix [13]. Using the neighbor joining algorithm, a rooted phylogenetic tree was constructed from the multiple sequence alignment and visualized by Phylodraw [14]. The protein crystal structure of azoreductase from *Bacillus subtilis* is complexed with flavo mononucleotide (FMN) taken from the protein data bank (PDB ID1NNI) [15]. The corresponding Swiss Prot ID is O07529 [16].

Evolutionary trace (ET) analysis

For this purpose, we use multiple sequence alignment and phylogenetic tree available in the dataset [8]. The sequences present on different branches were allowed to form clusters. An early node defines a larger cluster, whereas the later forms a smaller cluster. The cluster defined by the first node to the right of vertical line, is considered to partition the entire tree. Thus, by partition identity cutoffs (PICs) which are the minimum percentage identity within the cluster, generates three different PICs, P1 to P3 based on the phylogram [17]. Sequences are aligned based on the partition identity. The consensus sequence from each cluster was obtained. The consensus residues can be classified as conserved, neutral, group specific residues. Conserved

***Corresponding author:** Dr. Shobana S, Department of Bioinformatics, SRM University, Kattankulathur, Chennai- 603203, India, E-mail: shobana.s@ktr.srmuniv.ac.in

Received August 16, 2012; **Accepted** September 11, 2012; **Published** September 13, 2012

Citation: Shobana S, Berla TE, Kasmir Raja SV (2012) Evolutionary Trace Analysis of Azoreductase at the Ligand Binding Site and Enhancing the Active Site through Site Directed Mutagenesis. J Proteomics Bioinform 5: 222-225. doi:10.4172/jpb.1000240

residues are the amino acids that are conserved in the multiple sequence alignment, whereas neutral residues are those amino acids that are not conserved. Then, the trace residues are obtained which are mapped onto the protein structure of azoreductase (1NNI), which can be done by Jalview [18] and the mapped structure can be visualized by Rasmol [19].

Active site prediction

The crystal structure of azoreductase 1NNI complexed with flavin mononucleotide (FMN) and is visualized in Rasmol for amino acid residues at active site. The amino acids within 5Å distance from FMN are alone considered for the prediction. Amino acids within this distance are referred as active site region.

Site directed mutagenesis

Using PyMOL, the residue can be modified to any other residue of interest, in order to perform site directed mutagenesis [20]. We docked the ligand FMN with the protein 1NNI using "Induced fit docking" module of the Glide. The active residues are generally used for mutational studies. Glide (Grid-based Ligand Docking with Energetics) is a program to perform ligand database screening and high accuracy docking with Glide. Glide is run primarily from the Maestro graphical user interface, but can also be run from the command line. Glide searches for favorable interactions between one or more ligand molecules and a receptor molecule, usually a protein. The glide score or docking score, glide energy, glide emodel before and after performing mutation are analyzed. By default, Schrödinger's proprietary GlideScore multi-ligand scoring function is used to score the poses. If GlideScore was selected as the scoring function, a composite EModel score is then used to rank the poses of each ligand, and to select the poses to be reported to the user. EModel combines GlideScore, the nonbonded interaction energy and for flexible docking, the excess internal energy of the generated ligand conformation. Glide energy is the Modified Coulomb-van der Waals interaction energy [21].

Results and Discussion

Active site residues

The result of binding site analysis when 1NNI is complexed with FMN, predicts 14 active residues as ligand binding site region. The amino acids in (Table 1) are considered as active residues in our analysis. Among these amino acids, 9Thr, 11Arg, 15Arg, 16Thr, 76Ser, 104Val, 105Ala and 106Gly are involved in hydrophilic interactions with the ligand, whereas the remaining residues 73Glu, 74Tyr, 75His and 105Ala exhibit hydrophobic interactions too. 110Gly was discovered to have no interactions with the ligand. The distance between the atoms of ligand and residue, as well as contact surface area and the trace status are mentioned (Table 1).

Phylogenetic tree and ET analysis

The 26 homologous sequences retrieved from O07529 of *Bacillus subtilis* were used to draw the phylogenetic tree. The length of these sequences varies from 171 to 173. *Geobacillus stearothermophilus* and *Bacillus subsp* are closely related to each other and distantly related to seed sequence, *Bacillus subtilis*. After the clustal alignment the rooted tree was drawn, which is further used in the ET algorithm to derive the clusters. Based on the identity P1, P2, P3 are clustered; P1 with 59%, 70%, 75% as shown (Figure 1). Some sequences are withdrawn like Bacillus sp (E5W4C5), as it does not show any relation with the other branches. The trace status shows conserved and neutral residues. 75His and 106Gly form hydrogen bond at the ligand binding site of azoreductase, and it was proved in our analysis that they are conserved.

Residue	Dist	Surf	HB	Arom	Phob	DC	Trace status
9 THR*	3.0	12.0	+	-	-	-	Conserved
11 ARG*	2.7	58.9	+	-	-	+	Neutral
14 GLY*	2.8	24.6	-	-	-	-	Conserved
15 ARG*	2.5	50.3	+	-	-	-	Neutral
16 THR*	2.5	44.0	+	-	-	-	Conserved
72 PRO*	3.1	12.1	-	-	-	+	Neutral
73 GLU*	2.9	41.1	+	-	+	+	Conserved
74 TYR*	2.7	68.4	+	+	+	-	Conserved
75 HIS*	2.9	58.6	+	+	+	+	Conserved
76 SER*	2.9	27.4	+	-	-	-	Neutral
104 VAL*	3.0	37.1	+	-	-	+	Neutral
105 ALA*	3.3	33.6	+	-	+	+	Neutral
106 GLY*	3.1	30.4	+	-	-	-	Neutral
110 GLY*	3.5	12.4	-	-	-	-	Conserved

Table 1: Active residue shown in 1NNI showing their trace status
Dist- nearest distance (A°) between atoms of the ligand and the residue
Surf- contact surface area (A°) between the ligand and the residue
HB- hydrophobic-hydrophilic contact (hydrogen bond)
Arom- aromatic-aromatic contact
Phob- hydrophobic-hydrophobic contact
DC- hydrophobic-hydrophilic contact (destabilizing contact)
+/- - indicates presence/absence of a specific contacts
* - indicates residues contacting ligand by their side chain (including CA atoms)

Trace residues at the ligand binding

We analyzed the amino acid sequence alignment of all azoreductases at the binding site, to search for possible amino acid positions that could be manipulated by site-directed mutagenesis experiments, to design laccase with broader substrate spectrum. The trace residues which are neutral at the ligand binding site, are both polar and non-polar in nature. The non-polar residues 104 Val, 105 Ala, 106 Gly are non-polar amino acids, which can be shown in the alignment (Figure 2). These are the good sites for mutations. And it is really true that 106Gly was aligned with much more smaller non-polar amino acids like valine and alanine. This suggests us that non-polar property is more important than the aliphatic property of glycine at this position. This can be observed by performing induced fit docking.

Modification of Gly at 106 position with Val, Ala by "Induced fit docking"

It was considered that 106Gly can be a good site for mutagenesis and can be replaced with valine or alanine, which are small non-polar amino acid residues. Schrodinger is used for induced fit docking. The Figure 1 shows the induced fit docking before mutation. The "Red" color coding denotes a decrease in the value with respect to the reference value (i.e. value obtained before mutation) and the "Yellow" color coding represents the increase in value with respect to the value obtained before mutation.

Before mutation (wild)

Entry residue	Title	Docking score	Glide EModel	Glide energy
106Gly	azoreductase	-5.917	-51.534	-33.316

After mutation with valine

Entry residue	Title	Docking score	Glide EModel	Glide energy
106 Val	Azoreductase	-6.147	-67.701	-31.295

After mutation with alanine

Entry residue	Title	Docking score	Glide EModel	Glide energy
106 Ala	Azoreductase	-6.060	-64.086	-36.585

Discussion

The Evolutionary trace analysis at the ligand binding site suggests that glycine at 106th position to be the effective candidate for the site directed mutagenesis. Substitution of the amino acid with other smaller amino acid might result in paving way for the entry of larger substrates for degradation. The ligand binding site of azoreductase from *Bacillus subtilis* is different from the catalytic or active site of the enzyme, hence site directed mutagenesis carried out in the ligand binding site, is not going to affect the catalytic activity of the enzyme azoreductase. The reason for selecting this position for mutational studies is that it lies in the active site. Valine and alanine are the amino acids substituted in 106 position of glycine. The results obtained after substitution of valine (Figure 3 and 4) possess a low score of glide model and docking score, whereas the glide energy increased slightly suggests that, the activity of the protein is high though it inhibits the protein. The substitution of alanine (Figure 5), determines all our three parameters simultaneously (Docking score, Glide emodel, Glide Energy) to reach an overall low score (as usual by comparing with the results obtained before mutation). Since a better binding at the active site would ensure, quicker degradation of the toxic compounds bounded to azoreductase,

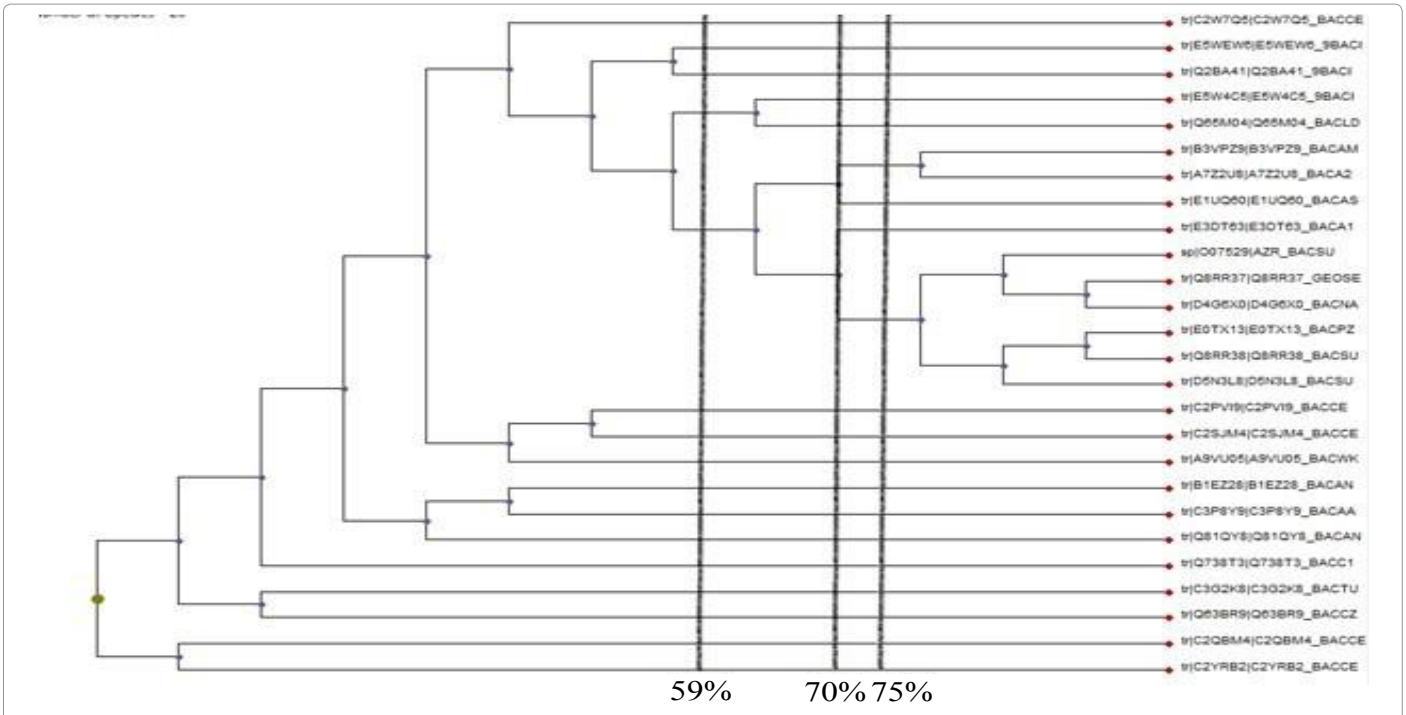

Figure 1: Phylogenetic tree of all azoreductases.

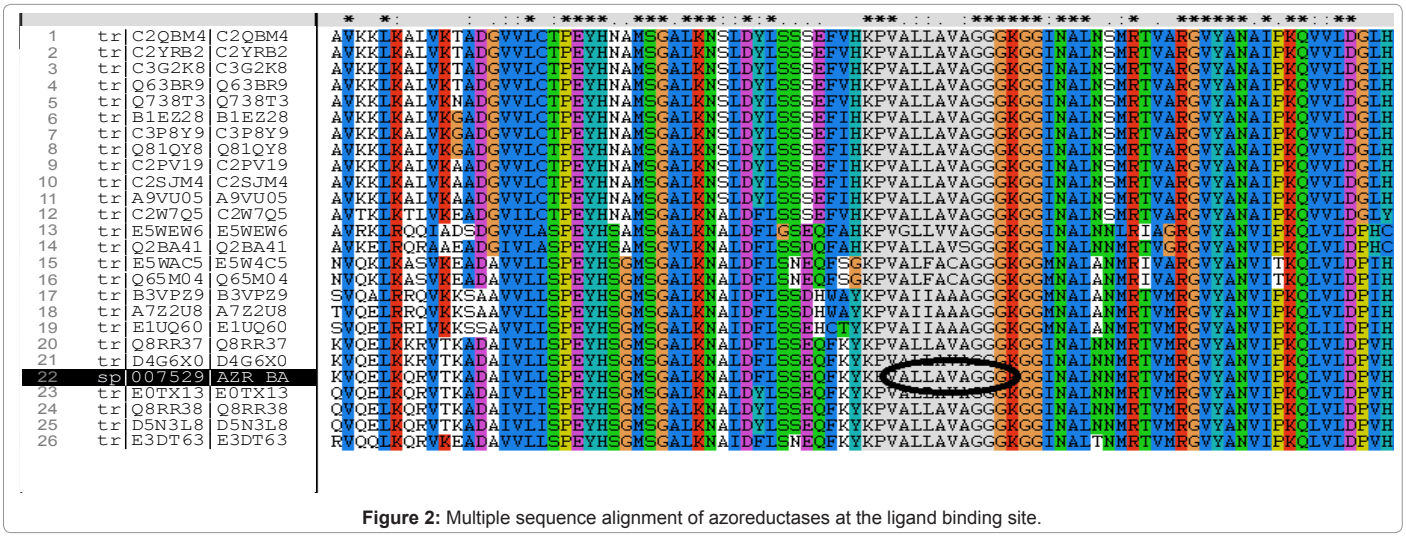

Figure 2: Multiple sequence alignment of azoreductases at the ligand binding site.

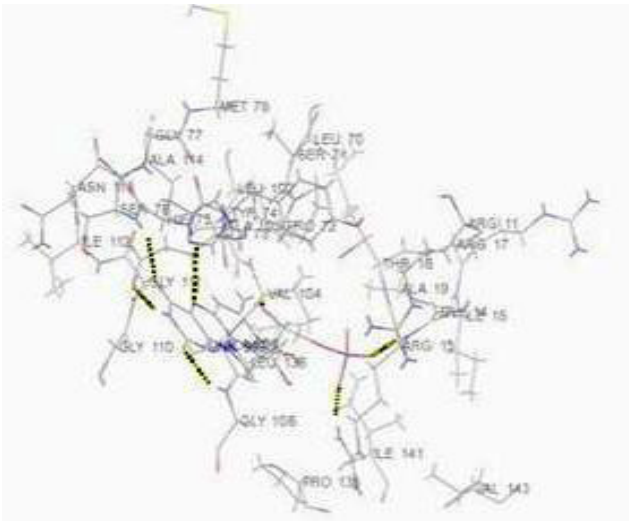

Figure 3: Induced fit docking before mutation.

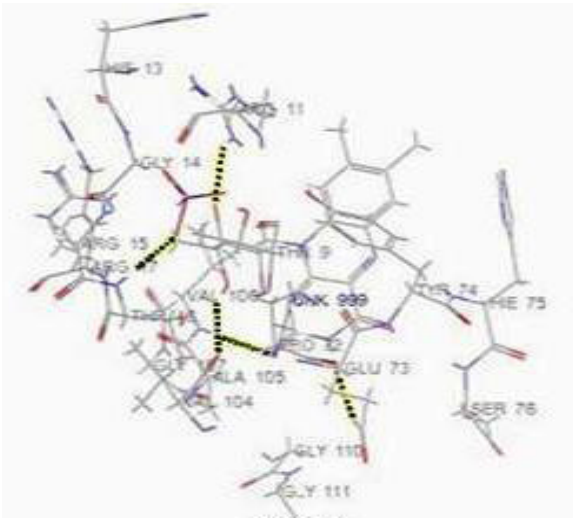

Figure 4: Induced fit docking with valine.

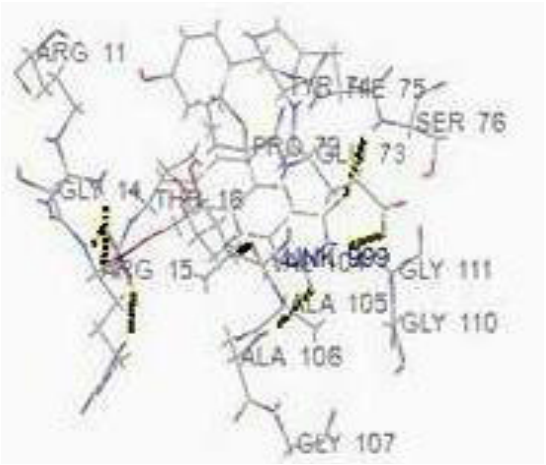

Figure 5: Induced fit docking with alanine.

we hope that in vitro lab experiments would get along with our hypothesis of "increasing the efficiency of azoreductase through site directed mutagenesis" (the reason is that these Insilco trials are carried out using Glide, which mostly gives a result close to the experimental observations). Hence, we could conclude by saying that the active site of azoreductase is modified to degrade toxic compounds more efficiently. However, the cavity enclosing FMN is rather wide, thus allowing the accommodation of substrates of various sizes to get into this cavity and get degraded after all, the substitution further provides more space for bigger toxic ligands to get in.

References

1. Rafii F, Hall JD, Cerniglia CE (1997) Mutagenicity of azo dyes used in foods, drugs and cosmetics before and after reduction by *Clostridium* species from the human intestinal tract. Food Chem Toxicol 35: 897–901.
2. Gurulakshmi M, Sudarmani DNP, Venba R (2008) Biodegradation of leather acid dye by Bacillus subtilis. Advanced Biotech 12-18.
3. Yeh MS, Chang JS (2004) Bacterial decolorization of an azo dye with a natural isolate of *Pseudomonas luteola* and genetically modified *Escherichia coli*. J Chem Technol Biotechnol 79: 1354–1360.
4. Puvaneswari N, Muthukrishnan J, Gunasekaran P (2006) Toxicity assessment and microbial degradation of azo dyes. Indian J Exp Biol 44: 618-626.
5. Alexander M (1977) Introduction to Soil Microbiology. John Wiley and Sons, Inc., New York. L. J. Jönsson, Markku Saloheimo, Merja Penttilä
6. Jönsson LJ, Saloheimo M, Penttilä M (1997) Laccase from the white-rot fungus Trametes versicolor: cDNA cloning of lcc1 and expression in *Pichia pastoris*. Curr Genet 32: 425-430.
7. Ramanathan K, Shanthi V, Sethumadhavan R (2009) In silico identification of catalytic residues in azobenzene reductase from *Bacillus subtilis* and its docking studies with azo dyes. Interdiscip Sci 1: 290-297.
8. Lichtarge O, Bourne HR, Cohen FE (1996) An evolutionary trace method defines binding surfaces common to protein families. J Mol Biol 29: 342-358.
9. Mohamad SB, Ong AL, Ripen AM (2008) Evolutionary trace analysis at the ligand binding site of laccase. BioInformation 18: 369-372.
10. http://www.expasy.org/
11. http://www.ncbi.nlm.nih.gov/blast/
12. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.
13. Gonnet GH, Cohen MA, Benner SA (1992) Exhaustive matching of the entire protein sequence database. Science 256: 1443-1445.
14. Choi JH, Jung HY, Kim HS, Cho HG (2000) PhyloDraw: a phylogenetic tree drawing system. Bioinformatics 16: 1056-1058.
15. Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res 29: 221-222.
16. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, et al. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31: 365-370.
17. Lichtarge O, Bourne HR, Cohen FE (1996) An evolutionary trace method defines binding surfaces common to protein families. J Mol Biol 257: 342-358.
18. Clamp M, Cuff J, Searle SM, Barton GJ (2004) The jalview java alignment editor. Bioinformatics 20: 426-427.
19. Sayle RA, Milner-White EJ (1995) RASMOL: biomolecular graphics for all. Trends Biochem Sci 20: 374.
20. http://www.pymol.org/
21. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, et al. (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 47: 1750-1759.