BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(14)

Computational Analysis of N-acetyl transferase in Tribolium castaneum

Kailash Singh[1]*, Vineet Kumar Mishra[1], Karabi Nath[2], Naira Rashid[1] & Farzana Parveen[1]

[1]Department of Biotechnology, Faculty of Science, Jamia Hamdard, New Delhi, India; [2]Department of Pharmacy, Faculty of Basic Medical and Pharmaceutical Science, University of Science and Technology Chittagong (USTC), Bangladesh; Kailash Singh – Email: kailash.dna@gmail.com; *Corresponding author

Received July 22, 2013; Accepted July 23, 2013; Published August 07, 2013

Abstract:
N-acetyl transferase (NAT) is responsible to catalyze the transfer of acetyl groups to arylamines from acetyl-CoA. Aralkylamine N-acetyl transferase (AANAT), which belongs to GCN5-related N-acetyl transferase member, is a globular 23-kDa cytosolic protein that forms a reversible regulatory complex with 14-3-3 proteins, AANAT regulates the daily cycle of melatonin biosynthesis in mammals, making it an attractive target for therapeutic control of abnormal melatonin production in mood and sleep disorders. There is no evidence available regarding α and β subunits, active site and their ASA value in Dopamine N-acetyl transferase. Therefore, we describe the development of Dopamine N-acetyl transferase model in Tribolium castaneum. We further document the predicted active sites in the structural model with solvent exposed ASA residues. During this study, the model was built by CPH program and validated through PROCHECK, Verify 3D, ERRAT and ProSA for reliability. The active sites were predicted in the model with further ASA analysis of active site residues. The discussed information thus provides insight to the predicted active site and ASA values of Dopamine N-acetyl transferase model in Tribolium castaneum.

Keywords: Protein selection and validation, Active site and ASA analysis.

Background:
N-acetyl transferase (NAT) was first identified as the genetically controlled step responsible for the inactivation of isoniazid **[1]** in the treatment of Tuberculosis **[2].** NAT is responsible to catalyze the transfer of acetyl groups to arylamines from acetyl-CoA, which have mainly specificity for aromatic amines (e.g. serotonin), and can also catalyze acetyl transfer between arylamines **[3].** GCN5-related N-acetyl transferase (GNAT) is a superfamily of enzymes that are broadly distributed in nature and use acyl-CoAs to acylate their cognate substrates **[4].** All GNAT family members share common structural features associated with acetyl coenzyme A binding; in addition, each member of the family has unique features reflecting substrate specificity for each of a wide range of substrates (e.g. aminoglycosides, diamines, puromycin, histones, and arylalkylamines). Aralkylamine N-acetyl transferase (AANAT), which belongs to GCN5-related N-acetyl transferase member, regulates the daily cycle of melatonin biosynthesis in mammals, making it an attractive target for therapeutic control of abnormal melatonin production in mood and sleep disorders. Melatonin is responsible for time keeping and playing a unique role in vertebrate biology by controlling the rhythmic production of melatonin in the pineal gland. It has been suggested in fish to play a role in osmoregulation, and in salmonids, relate to the timing of adaptive mechanisms during smolting **[5].** AANAT family includes enzyme like dopamine N-acetyl transferase (Dat) **[6].** The function of Dat is not melatonin biosynthesis; rather it is in cuticle sclerotization and neural transmission **[7].** Dat and other AANAT family members have not been found in the same genome **[8].** AANAT is a globular 23-kDa cytosolic protein that forms a reversible regulatory complex with 14-3-3 proteins **[8, 9].** Catalytic core of AANAT forms a cavity encompassing the arylalkylamine and AcCoA binding pockets.

As with other GNAT superfamily members **[10],** AcCoA binds by contacts between the pantethiene moiety and the edge of a rigid sheet directing the thioacetyl group into the center of the enzyme; the adenosine moiety is at the surface Arylalkylamines bind in a funnel-shaped pocket formed by three protein loops, which contact the aromatic ring of substrates via aromatic residues. This position of the protonated amine group of substrates is close to the AcCoA thioacetyl group **[11, 12].** The N-acetyl transferase, found in Tribolium castaneum (red flour beetle), is an AANAT type. The red flour beetle is a tenebrinionid beetle. It's a worldwide pest of stored products such as flour, cereals, pasts, biscuit, nut etc which causes loss and damage [**13**]. Since, there is no adequate information related to the active site amino acids and their nature. So, the present study deals with the identification of active site amino acid and accessible surface area analysis in protein model of Tribolium castaneum.

Methodology:

Isolation of Template selection and confirmation of absent N-acetyl transferase: The Dopamine N-acetyl transferase in complex with acetyl-COA from Drosophila Melanogaster was isolated from RSCP PDB **[14]** as template and compared against a complete Non-Redundant Protein Database using The NCBI PSI-BLAST to ensure that those protein matching eukaryotic protein exclusively in the first step truly had no Dopamine N-acetyl transferase in Tribolium castaneum. Visualization and Model Validation: The protein structure was predicted by homology modeling with different server CPH model (http://www.cbs.dtu.dk/services/CPHmodels) and SAVES (http: //nihserver.mbi.ucla.edu/SAVES). The program CPH models 3.0 servers **[15]** was used to build the model as PDB file of N-acetyl transferase in Tribolium castaneum according to the homology modeling method. The PDB of N-acetyltransferase in Tribolium castaneum was visualized by using PyMOL version 1.3 (http: //pymol. sourceforge.net/) **[16].** The structure which was obtained from homology modeling was validated by SAVES. The steriochemical quality of the model was verified with program PROCHECK **[17]** in order to select best model.3D-profiling of the residue was done by VERIFY 3D Structure evaluation Server **[18, 19].** ERRAT was used for verifying protein structure for evaluating the progress of crystallographic model building and refinement **[20]**. The model was submitted to the ProSA **[21]** to obtain the Z-Score. Protein Structure analysis: Secondary structure of the protein was predicted through SOPMA program (http://npsapbil.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_sopma.html), the program determined individual role of amino acid for building the secondary structure with their positions **[22].**

Active site and accessible surface area analysis: One of the keenest areas in the bioinformatics is the active sites of the protein and its residual identification and characterization. The structure was used to find out the ligand binding efficiency with the predicted model and to detect whether the protein was better or weak with other reported sources. CASTp (http://sts.bioengr.uic.edu/castp/calculation.php/) was used for identification of active site from the protein structure **[23].** POPS (http://mathbio.nimr.mrc.ac.uk/~ffranca/POPS) is a new method based on an empirically parameterisable analytical formula to calculate solvent accessible surface areas of protein **[24].**

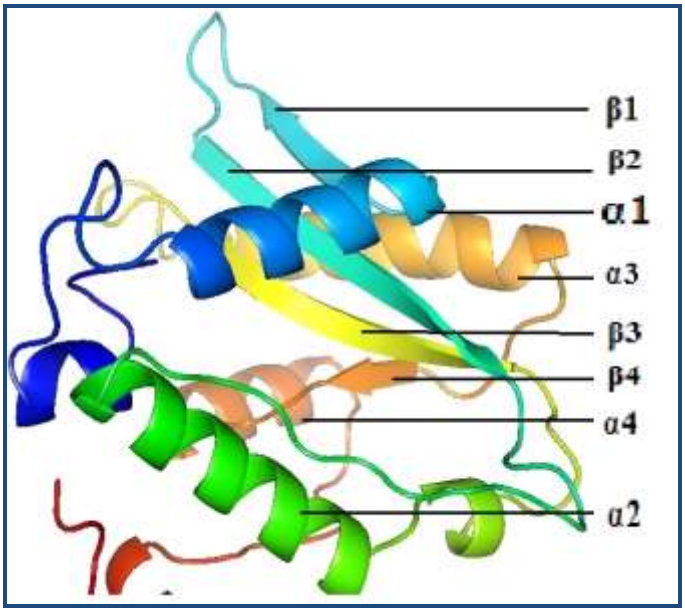

Figure 1: 3-D structure of the predicated model by PyMOL

Figure 2: Active site in predicated model computed with CASTp. Green color represents active site with largest area and volume and other colors represent the remaining active site with different areas and volumes

Discussion:

Isolation of Template selection and confirmation of absent Dopamine N-acetyl transferase: The Dopamine N-acetyl transferase in complex with acetyl-COA from Drosophila Melanogaster (accession ID: 3TE4) was retrived from RSCP PBD and was compared against complete Non-Redundant Protein Database using The NCBI PSI-BLAST and N-acetyl transferase absent in PBD used as subject in Tribolium castaneum (Accession ID: ACN43328.1). Visualization and Model Validation: The three dimensional structure of Dopamine N-acetyl transferase was generated by using PyMOL program **(Figure 1).** To verify the predicted structure, validation was carried out with PROCHECK program. Ramchandran plot of non-glycine and non-proline residue in the structure showed that 89.7% of the total amino acids were presented in most favored regions and the other 10.3% of amino acids were presented in allowed regions including disallowed region with 0.0%. VERIFY_3D shows 87.43% of the residues had an averaged 3D-1D score greater than 0.2 indicates that the environment profile of the model is good. ERRAT2 shows 92.208 overall quality factors indicating good resolution structure. Moreover, quality of the model can be

compared to reference structure of high resolution obtained from X-Ray crystallography analysis through Z score and "0" is the average Z score for good model. The Z score of N-acetyl transferase is -5.74 showing the possibility to be a better model. Figure 1: 3-D structure of the predicated model by PyMOL

Protein Structure analysis: The secondary analysis indicates whether a given amino acid is located in helix strand or coil. The result obtained from SOPMA described that about 47.70% of amino acids presented in helix containing 4 α-subunits, 24.14% of amino acids in random coil, 21.26% of amino acids in extended strand as 4 β subunits and rest of all amino acids in bridge and turn. All the subunit is shown in (**Figure 1)**. Active site and accessible surface area analysis: A total of 10 active sites were evaluated in the structure through CASTp software with ideal parameters shown in table 6. All 10 pockets were characterized to find out its residues around probe radius of 1.4Å and among them; largest and second largest active sites have an area of 2014.5 Å and 97.6 Å respectively and volume of 3378.1 Å and 91.1 Å respectively. The green color (figure 2) shows the largest active site position in the build protein which lies between amino acid 1 and 161. Accessible surface area (ASA) analysis of the predicted model showed the amino acids with low ASA value are buried inside the catalytic cleft and with high ASA values are on the surface of the cleft. Some of the ligand binding residues and metal ion binding residues were found to have high ASA values (LYS17, ASN42, LYS44, TYR58, GLU60, GLU63, GLU64, ARG113,) and some others were found to have low ASA values (CYS20, SER27, GLY47,CYS49,THR123,ALA127, ALA145). The active site amino acids, which are hydrophobic, are CYS20, ALA127 and ALA145 with low ASA values. While those active site amino acids, which is hydrophilic, are LYS17, GLU63 and GLU64 with high values. Most of the active site amino acids such as SER10, SER33, THR55, TYR78, SER86, SER105, SER108, SER141, THR144, TYR157, SER158, TYR161 and THR173, which are oxidized mostly, show average ASA values with hydrophilic nature indicating that, are **(Figure 2**): Active site in predicated model computed with CASTp. Green color represents active site with largest area and volume and other colors represent the remaining active site with different areas and volumes located in the middle of the ceft. Except SER10, SER33, SER108, and THR173, all the above active site amino acids are the part of large active site in the predicted model.

Conclusion:

In this study, we proposed a valid and stable 3D model of Dopamine N-acetyl transferase in Tribolium castaneum whose structure is not present in PDB (Protein Data Bank). Further analysis provides information about α and β subunits, active site and their ASA value analysis in the predicted model. On the basis of the findings, it could be concluded that further characterization of Dopamine N-acetyl transferase from Tribolium castaneum will be important in regulation of the melanin sysnthesis. Such models can be effectively used to design and support the drug design project. Knowledge gained from this study will be used in broad screening on inhibitors of the protein and can be further implemented in future drug designing.

Reference:

[1] Peter JH *et al. J Pharmacol Exp Ther*. 1965 **150**: 298 [PMID: 5855885]
[2] Tam CM *et al. Int J Tuberculosis Lung Dis*. 2000 **4:** 262 [PMID: 10751074]
[3] Evans DA, *Pharmacol*. 1989 **42**: 157 [PMID: 2664821]
[4] Matthew *et al. Arch Biochem Biophys*. 2005 **433**: 212 [PMID: 15581578]
[5] Sekeris CE & Karlson P, *Pharmacol Rev*. 1966 **18**: 89 [PMID: 5904192]
[6] Caniato *et al. Adv Exp Med Biol*. 2003 **527:** 593 [PMID: 15206778]
[7] Lopez-Patino MA *et al. J Exp Biol*. 2011 **214**: 928 [PMID: 21346120]
[8] Coon SL *et al Science*. 1995 **270**: 1681 [PMID: 7502081]
[9] Klein DC *et al*. *Recent Prog Horm Res*. 1997 **52**: 307 [PMID: 9238858]
[10] Dyda F *et al. Annu Rev Biophys Biomol Struct*. 2000 **29**: 81 [PMID: 10940244]
[11] Wolf E *et al J Mol Biol*. 2002 **317:** 215 [PMID: 11902838]
[12] Hickman AB *et al. Mol Cell*. 1999 **3:** 23 [PMID: 10024876]
[13] Peters JM *et al. J Pharmacol Exp Ther*. 1965 **150**: 298 [PMID: 5855885]
[14] Berman HM *et al. Nucleic Acids Res*. 2000 **28**: 235 [PMID: 10592235]
[15] Nielsen M *et al. Nucleic Acids Res*. 2010 **38:** W576 [PMID: 20542909]
[16] Lorenz IC *et al. Viruses*. 2010 **2**: 1635 [PMID: 21994698]
[17] Laskowski *et al. J Biomol NMR*. 1996 **8**: 477 [PMID: 9008363]
[18] Bowie JU *et al*. *Science*. 1991 **253:** 164 [PMID: 1853201]
[19] Luthy R *et al. Nature*. 1992 **356**: 83 [PMID: 1538787]
[20] Colovos C & Yeates TO, *Protein Sci*. 1993 **2**: 1511 [PMID: 8401235]
[21] Benkert P *et al. Bioinformatics*. 2011 **27:** 343 [PMID: 21134891]
[22] Geourjon C & Deleaqe G, *Comput Appl Biosci*. 1995 **11**: 681 [PMID: 8808585]
[23] Dundas J *et al. Nucleic Acids Res*. 2006 **34**: W116 [PMID: 16844972]
[24] Kleninjung J & Fraternali F, *Nucleic Acids Res*. 2005 1: W342 [PMID: 15980485]

Edited by P Kangueane
Citation: Singh *et al.* Bioinformation 9(14): 715-717 (2013)