BIOINFORMATION
Discovery at the interface of physical and biological sciences

Open access

BIOMEDICAL INFORMATICS

www.bioinformation.net

Volume 14(4)

Software

ASBAAC: Automated Salt-Bridge and Aromatic-Aromatic Calculator

Chittran Roy & Saumen Datta*

[1]Structural Biology and Bioinformatics Division, Council of Scientific and Industrial Research – Indian Institute of Chemical Biology, 4 Raja SC Mullick Road, Jadavpur, Kolkata-700032, West Bengal, India; Saumen Datta – E-mail: saumen_datta@iicb.res.in; Telephone Number: +913324995896 Fax: +913324735197, +913324723967; *Corresponding Author

Received March 19, 2018; Revised April 5, 2018, Accepted April 5, 2018, Published April 30, 2018

doi:10.6026/97320630014164

Abstract:
Biological systems are made of complex networks non-covalent interactions observed among protein-protein, protein-DNA, protein-lipid complexes using hydrogen-bonds, salt-bridges, aromatic-aromatic, van der Waals (vdW), hydrophobic-interactions and several others using distance criteria. Hence, large-scale data analysis is required to understand the principles of biological complex formation. Therefore, it is of interest to analyze non-covalent interaction namely, salt-bridge and aromatic-aromatic contacts in known and modeled protein complex structures. Here, we describe ASBAAC for automatic calculation of salt-bridges and aromatic-aromatic contacts in protein complexes. This software tool is fast, robust and user-friendly for large-scale analysis of inter-chain salt bridges and aromatic-aromatic contact in protein complexes.

Availability: ASBAAC is available for free at https://sourceforge.net/projects/asbaac

Abbreviation: S-B **=** Salt-Bridge, A-A = Aromatic-Aromatic

Keywords: salt-bridges, aromatic-aromatic contacts, protein, program, software

Background:
Interaction between two or more proteins play a crucial role in maintaining cellular systems through non-covalent interactions such as hydrogen bonds, salt bridges and aromatic-aromatic (A-A) contacts. A particular region where two or more proteins interact with each other is called the interface **[1].** Knowledge of amino acids, which are involved in the formation of salt-bridges and A-A contacts with other interacting interfaces between various proteins, is important for the better understanding of protein-protein binding. Salt-bridges most often arise from the anionic amino acids (aspartic acid or glutamic acid) and the cationic amino acids (lysine, arginine or histidine) **[2].** The interaction with these functional groups plays an important role in the formation of the structure and function of proteins. Formation of Salt Bridge between two residues occurs at a distance of 4A° **[3].** There are numerous salt-bridge interaction strategies such as simple or complex, isolated or networked **[4],** intra- helical, coiled, strand or inter-helical, coiled and strand **[5].** Mutagenesis studies and nuclear resonance technique reveal that the contribution of salt-bridge plays an important role in the overall stability of proteins.

Apart from salt-bridge, the aromatic residue interactions also play crucial role in protein stabilization, protein-protein recognition, ligand binding and protein folding **[6].** On average 60% of aromatic side chains of protein involves in the formation of aromatic pair and 80% of which involves in the formation of a network of three or more interacting aromatic side chains **[7].** Interaction occurs when pi rings range is between 4.5 to 7.5 A° assisting in the formation of pi-pi stacking **[8].**

There is about fifty thousand (until February 2018) high-resolution (≤ 2.5 A°) multi-chain protein structure available at the protein data bank (PDB) used for the analysis of salt-bridges and aromatic-aromatic contacts. There has been extensive research on hydrogen bonds. However, data on salt-bridges and aromatic-

ISSN 0973-2063 (online) 0973-8894 (print)

BIOINFORMATION
Discovery at the interface of physical and biological sciences

Open access

aromatic interactions is limited. Therefore, there is a need to develop computer software tools to analyze S-B and A-A interaction of multi-chain proteins using known structure complexes. We describe ASBAAC for the large-scale analysis of S-B and A-A. ASBAAC is fast, robust, simple and user-friendly. ASBAAC is freely available for academic use.

Program input:
ASBAAC uses .pdb file format as input. This program scans the PDB text file and calculates the inter-chain salt bridge interaction and aromatic-aromatic contacts. ASBAAC should be run from UNIX prompt in directory containing PDB files for large-scale analysis.

Program output:
Outputs are generated in the same working directory. A detail flow-chart for ASBAAC is shwon in **Figure 1.** It checks for chain continuity and then constructs a sequence while scanning the topology of the structure .pdb file. ASBAAC then calculates the distance between amino acids identifying salt-bridges, and aromatic-aromatic contacts in all possible chain combinations. The program runs at three modes. First mode checks the salt-bridge interactions, second mode extracts aromatic interactions and third mode combines all possible results to produce output in text format within the same working directory.

Caveats and future development:
ASBAAC software is written in AWK programming language, which can be successfully run from C shell Unix prompt in 32-bit CYGWIN OS. It can also be made run from B shell LINUX and WINDOWS environment. We are developing web interface to integrate ASBAAC in future development.

Methodology:
Distance calculation:
Inter-atomic SB and A-A distance in inter-chain protein is calculated. Inter-atomic distance between charged amino acids of two chains is calculated using equation 1 given below.

$$dist = \sqrt{\left(X_a - X_b\right)^2 + \left(Y_a - Y_b\right)^2 + \left(Z_a - Z_b\right)^2} \qquad (1)$$

Here, *a* and *b* are specified side chain atoms of acidic and basic residues respectively. *X*, *Y* and *Z* are atomic coordinates.

This rule is also applicable for aromatic-aromatic interaction. In case of tyrosine, phenylalanine, histidine and tryptophane (along with its indole ring), the distance between any of these residues is calculated by taking the distance between the centroids of aromatic rings.

Salt-bridge and aromatic-aromatic bond calculation:
Positively charged atoms are Lys NZ1, Arg NH1, Arg NH2 and His NE2. Negatively charged atoms are Asp OD1, Asp OD2, Glu OE1, Glu OE2. A salt bridge is defined if two oppositely charged atoms lie within 4A° across the interface. A pi-pi interaction is defined between two aromatic amino acids if the distance calculated from their centroid is less than 7.5A°.

Analysis of NMR solved structure:
ASBAAC analyzes inter-chain salt-bridge and aromatic-aromatic interactions at the interface of homo-meric or hetero-meric subunits. However, the program does not work for NMR structures with variable number of conformers containing homo-meric or hetero-meric subunit structures. Analysis should be completed after separating the conformers in such cases **(Figure 2).**

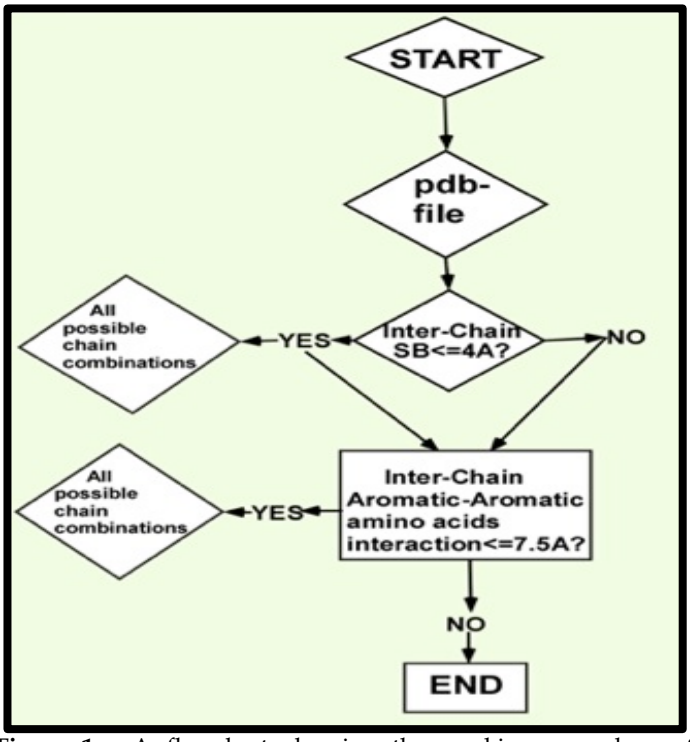

Figure 1: A flowchart showing the working procedure of ASBAAC. The first function is to scan the pdb file. If distance between atomic SB is within 4A° (i.e. YES), ASBAAC calculates all possible chain combination and goes to next step. If distance between atomic SB not within 4A° (i.e. NO), it comes out and goes to next level. Then the program enters the next lap to calculate pi-pi interaction. If the distance is within 7.5 A° (i.e. YES) it calculates all possible chain combinations until end of program.

BIOINFORMATION
Discovery at the interface of physical and biological sciences

Open access

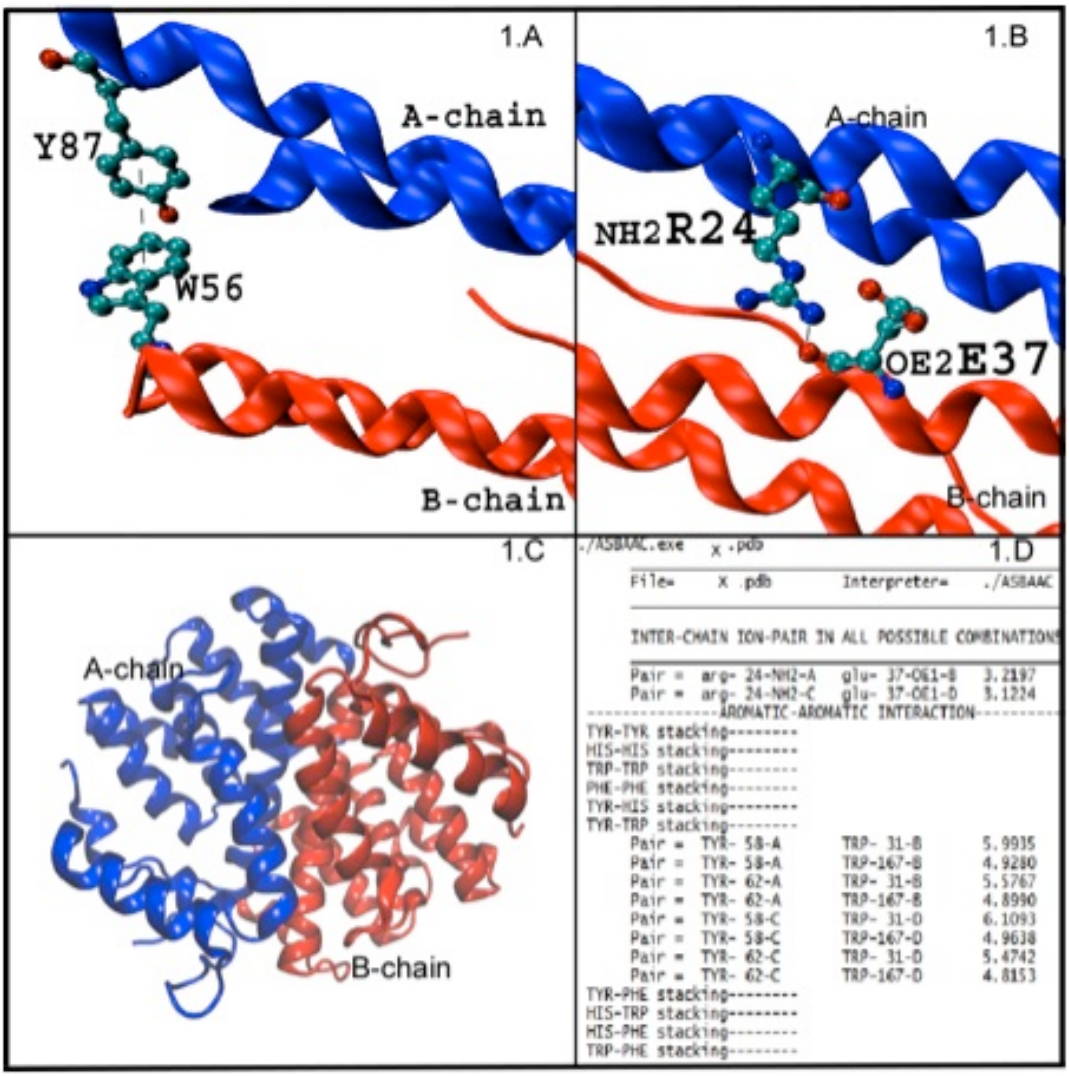

Figure 2: Details of ASBAAC extracted aromatic-aromatic (A) and salt-bridge (B) of multi-chain protein (C) analysis (D). X denotes any complex/multi-chain protein PDB ID.

Acknowledgement:
We acknowledge the computational facility at the structural biology and bioinformatics department, Indian Institute of Chemical Biology, Kolkata.

References:

[1] Keskin O *et al*. Methods Mol Biol. 2008, 484 [PMID: 18592198]
[2] Bosshard HR *et al*. J Mol. Recognition 2004, **17:**1 [PMID: 14872533]
[3] Kumar S *et al*. ChemBioChem. 2002, **3:**7 [PMID: 12324994]
[4] Sarakatsannis JN *et al*. Proteins. 2005, **60:**4 [PMID: 16021620]
[5] Gupta PSS *et al*. Bioinformation. 2015, **11:**1 [PMID: 25780279]
[6] Burley SK *et al*. Science. 1985, **229:**4708 [PMID: 3892686]
[7] Makwana KM et al. Protein Sci. 2015, **24:**12 [PMID: 26402741]
[8] Anjana R *et al*. Bioinformation. 2012, **8:**24 [PMID: 23275723]

Edited by P Kangueane
Citation: **Roy & Datta *et al.*** Bioinformation 14(4): 164-166 (2018)