Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

HBNG: Graph theory-based visualization of hydrogen bond networks in protein structures

Abhishek Tiwari[1*]& Vivek Tiwari[2]

[1]GVK Biosciences, Informatics Division, Hyderabad-500037, India; [2] CIDSH, Lucknow-226020, India;
Abhishek Tiwari* - Email: abhishek.twr@gmail.com; * Corresponding author

received April 10, 2007; revised May 28, 2007; accepted June 03, 2007; published online July 09, 2007

Abstract:
HBNG is a graph theory-based tool for visualization of Hydrogen Bond Network in 2D. Digraphs generated by HBNG facilitate visualization of cooperativity and anticooperativity chains & rings in Protein Structures. HBNG takes Hydrogen Bonds list files (Output from HBAT, HBexplore, HBPLUS and STRIDE) as input and generates a *DOT* language script and constructs digraphs using freeware AT&T Graphviz tool. HBNG is useful in the enumeration of favorable topologies of hydrogen bond networks in protein structures and determining the effect of cooperativity and anticooperativity on protein stability and folding. HBNG can be applied to protein structure comparison and in the identification of secondary structural regions in protein structures.

Availability: Program is available from the authors for non-commercial purposes.

Keywords: cooperativity; anticooperativity; digraph; *DOT* Script; graphviz

Background:
Many properties, for example, stabilization energy, of '*n*' interconnected hydrogen bonds are not just the sum of those of '*n*' isolated bonds. The total effect either exceeds the simple, additive sum of the individual, isolated effects, or stays lower. This non-additive behavior is explained by two principle mechanisms based on mutual polarization of the groups involved in hydrogen bonding-Cooperativity & Anticooperativity. **[9]** When a hydrogen bond is formed between two groups, the redistribution of electrons changes the ability for further hydrogen bonding. This electron redistribution thus results in cooperativity (e.g. accepting one hydrogen bond encourages the donation of another) and anticooperativity (e.g. accepting one hydrogen bond discourages acceptance of another) in hydrogen bond formation in hydrogen bond networks. Cooperativity may be thought of as synonymous with synergism. σ-bond cooperativity leads to formation of X-H…X-H…X-H chains and rings, in particular for X=O, but also for X=N or S or C. If multiple donors (such as H2O) and/or multiple acceptors are involved, they can interconnect chains and rings to form complex networks of hydrogen bonds. The overall effect of cooperativity is the enhanced stabilization of chains and rings. Hydrogen bonds may not only enhance, but also reduce the strengths of each other due to underlying Anticooperativity mechanism. Hydrogen bonding interactions make non-additive contributions to protein stability, but the non-additive nature depends on whether such interactions are located in the protein interior or on the protein surface. When located on the protein surface, these regular hydrogen-bonding interactions are anticooperative. **[10]** Cooperativity and hydrogen bonding network are widely studied in water clusters. **[8]** Graph theoretical techniques are demonstrated to be of considerable use in the search for stable arrangements of water clusters. The hydrogen-bonding topology was found to have a significant effect on cluster stability. The enumeration of digraphs is a useful method to represent all the possible topology–distinct hydrogen bond patterns of water clusters.

HBNG is a new graph theory based tool for visualization of hydrogen bonding networks in protein structures at different levels with varying definitions of graph nodes (Atoms, Residues or Secondary Structural Units) and edges (Geometrical features of hydrogen bonds). Tools like HERA **[2]** and TOPS **[6]** are also available for providing hydrogen bonding networks visualization. TOPS give topological descriptions (secondary structural topology) of protein structures and used in structure comparison at the topological level. HERA generates hydrogen bonding diagrams of protein structures and optionally helical wheels and helical nets.

Methodology:
The most important feature of the hydrogen bond is that it possesses direction and hence hydrogen bond networks along with cooperativity & anticooperativity can best represented by directed graphs. A *graph* is a mathematical object that captures the notion of connection which may represent the topology of a given network A directed graph (digraph) is a set of vertices (nodes or points) and arcs (arrows or directed edges). For a digraph with n vertices, the directed adjacency matrix A is the nth order square matrix, whose element a_{ij} is equal to 1 for an arc directed from vertex i to vertex j, and 0 otherwise.

Bioinformation by Biomedical Informatics Publishing Group *open access*

www.bioinformation.net **Hypothesis**

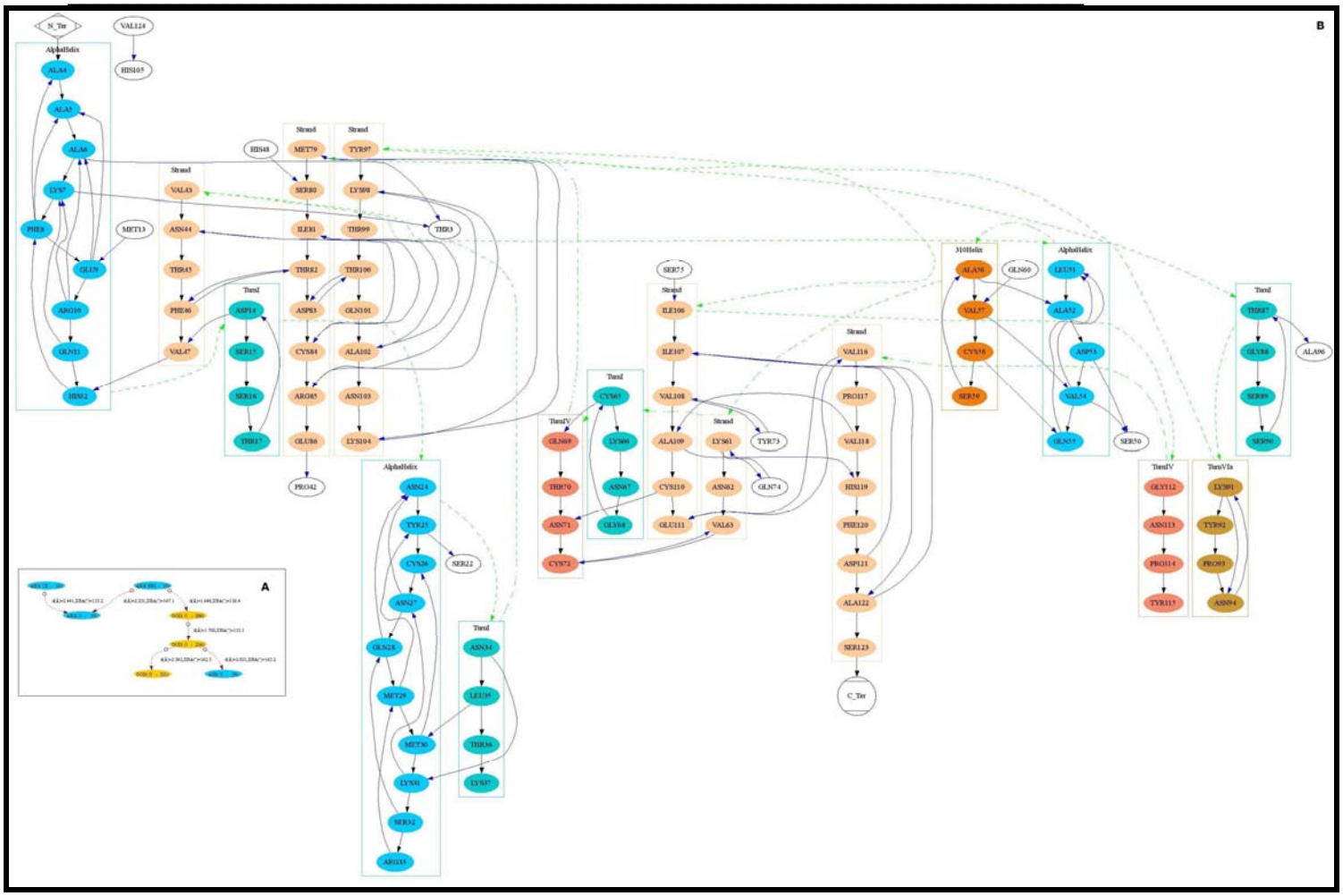

Figure 1: HBNG for PDB entry 6rsa. (A) Nodes are labeled with residue name, residue no and atom type. Edges are labeled with geometrical features of hydrogen bond like θ (XHA bond angle) and d (H…A distance). Node ARG(NH1- 10) represent donor atom NH1 which forms N-H…O type hydrogen bonds with O atoms of DOD(O -296) and ARG(O- 33) nodes. (B) SSEs level hydrogen bonding network where amino acid residues are grouped and colored according to their secondary structural units.

Hydrogen bond networks can be represented by digraphs, where vertices correspond to donor and acceptor group, and arcs correspond to hydrogen bonds from proton-donor to proton-acceptor and can be labeled with D, d or θ. In case of hydrogen bond networks one can call this directed adjacency matrix as Hydrogen Bond Network Adjacency (HBNA) Matrix.

Programs like HBAT **[7]**, HBPLUS **[5]**, HBexplore **[4]** and STRIDE **[1]** can produce hydrogen bond list (HB list file) for a given PDB file of protein macromolecules. HBNG parses the HB list file to generate an HBNA Matrix. The HBNA Matrix is employed by HBNG to generate a *DOT* language script. DOT language describes three kinds of objects: graphs, nodes, and edges and was originally developed by AT&T labs to support Graphviz. Graphviz **[3]** is open source graph visualization software. Graphviz has many useful features for concrete diagrams, such as options for colors, fonts, tabular node layouts, line styles, hyperlinks, and custom shapes. Graphviz includes *dot* (draws directed graphs as hierarchies) and *dotty* (a graph editor for the X-Window System) and both accept a DOT script file as input. HBNG is written in PERL and the installation for Windows operating system is available with the Graphviz components. HBNG provides the option for filtering weak hydrogen bonds based on angle and distance cutoff to avoid highly cross linked hydrogen bond networks which means network topology will depend on the cutoff values. One can assign different colors (red, blue, green etc) and shapes (circle, rectangle, ovals, double circle) to vertices according residue or molecule type (amino acid, nucleic acid, water, ligands) along with labeling of vertices and arcs.

Results & application:

The directive power of hydrogen bonds is apparently the major factor for the uniqueness and specificity of biopolymer structures. Directed graphs generated by HBNG facilitate visualization of

cooperativity and anticooperativity chains & rings. HBNG is useful in the identification of clusters of amino acid residues that stabilize the protein structure and protein-protein interfaces. It will give better insights into protein structure and stability by elucidating the role of the amino acids in maintaining the unique topology of protein structures and will be helpful in studies of protein folding, stability and design. Compared to HERA and TOPS, HBNG can include water molecules, ligands and water-mediated interaction (see Figure 1.a) in its hydrogen bond network graph. It also includes directed edges or arcs labeling which display available geometrical information about hydrogen bond. In HBNG hydrogen bonding networks can be visualized at many levels like at atom level (atoms participating in hydrogen bond formation make nodes for the graph, see Figure 1.a), residue level (residues participating in inter residue hydrogen bond formation make nodes for the graph) and SSEs level (SSEs i.e., Secondary Structure Elements participating in inter SSEs hydrogen bond formation make nodes for the graph, see Figure 1.b). For SSEs level visualization HBNG utilizes STRIDE **[1]** for identification of secondary structural elements. HBNG can be applied to comparing proteins at different levels and correlating protein structure with attributes of individual residues.

Conclusion:

HBNG is a very useful tool for the visualization of Hydrogen Bond Networks in 2-D and facilitate in analysis of the role of hydrogen bond networks and cooperativity in the stability of biopolymer structures.

References:

[01] D. Frishman, *et al., Proteins*, 23:566 (1995) [PMID: 8749853]
[02] E. G. Hutchinson, *et al., Proteins,* 8:203(1990) [PMID: 2281084]
[03] E. R. Gansner, *et al., Soft Prac and Experience,* 30:1203(1999)
[04] K. Lindauer, *et al., Comput Appl Biosci,* 12:281(1996) [PMID: 8902345]
[05] I. K. McDonald, *et al., J Mol Biol*, 238:777(1994) [PMID: 8182748]
[06] I. Michalopoulos, *et al., Nucleic Acid Res,* 32:D252 (2004) [PMID: 14681405]
[07] S. K. Panigrahi, *et al., Proteins,* 67:128(2007) [PMID: 17206656]
[08] S. S. Xantheas, *Chemical Physics,* 258:225(2000)
[09] T. Steiner, *Angew Chem Int Ed,* 41:48(2002)
[10] X. Li, *et al., Proteins,* 60:46(2005) [PMID: 15849756]

Edited by R. Sowdhamini

Citation: Tiwari & Tiwari, Bioinformation 2(1): 28-30 (2007)