Published online 5 May 2016 *Nucleic Acids Research, 2016, Vol. 44, Web Server issue* ***W375–W382***
doi: 10.1093/nar/gkw383

NAPS: Network Analysis of Protein Structures

Broto Chakrabarty and Nita Parekh*

Centre for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500032, India

Received February 15, 2016; Revised April 12, 2016; Accepted April 25, 2016

ABSTRACT

Traditionally, protein structures have been analysed by the secondary structure architecture and fold arrangement. An alternative approach that has shown promise is modelling proteins as a network of non-covalent interactions between amino acid residues. The network representation of proteins provide a systems approach to topological analysis of complex three-dimensional structures irrespective of secondary structure and fold type and provide insights into structure-function relationship. We have developed a web server for network based analysis of protein structures, NAPS, that facilitates quantitative and qualitative (visual) analysis of residue–residue interactions in: single chains, protein complex, modelled protein structures and trajectories (e.g. from molecular dynamics simulations). The user can specify atom type for network construction, distance range (in Å) and minimal amino acid separation along the sequence. NAPS provides users selection of node(s) and its neighbourhood based on centrality measures, physicochemical properties of amino acids or cluster of well-connected residues (k-cliques) for further analysis. Visual analysis of interacting domains and protein chains, and shortest path lengths between pair of residues are additional features that aid in functional analysis. NAPS support various analyses and visualization views for identifying functional residues, provide insight into mechanisms of protein folding, domain-domain and protein–protein interactions for understanding communication within and between proteins. URL:http://bioinf.iiit.ac.in/NAPS/.

INTRODUCTION

Structural stability and function of proteins rely on complex network of inter-residue interactions. Considering amino acid residues as nodes and edges between them drawn based on spatial distance or interaction energy, a network representation of protein, called Protein Contact Network (PCN) or Residue Interaction Network (RIN), is obtained that captures these inter-residue interactions. Over the past decade this alternative approach to protein structure analysis has gained popularity with numerous studies investigating various aspects of network topology to gain insight into structure–function relationship (1–4). These studies have revealed the role of topological analysis in identifying residues crucial for protein stability (5,6) and protein dynamics (7), governing enzymatic activity (8) and allosteric regulation (9,10), signal transduction and protein folding kinetics (11). The modularity analysis of PCNs have been carried out to identify domains (12) and structural repeats in proteins (13). Several specialized online and standalone tools have been developed to assist in visualization and analysis of protein contact maps. For e.g. aminonet provides network based analysis of physico-chemical properties of amino acids (14) while GraProStr (15) allows identification of hubs, cluster of residues, cliques and modularity based analysis. Con-Struct Map (16), CMWeb (17) and CMA (http://ligin.weizmann.ac.il/cma/) allows structural comparison of proteins by analysing the contact maps, and CMA also supports analysis of protein complexes. Standalone programs such as SeqX (18), Protmap2D (19) and CMview (20) allow construction and basic analysis of contact maps. RING provides several methods for the construction of residue interaction network, which can be analysed using Cytoscape (21). RINalyzer (22) and CyToStruct (23) are Cytoscape (24) plugins which can be integrated with other features of Cytoscape for the analysis of protein structures. xPyder (25) and PyInteraph (26) are PyMOL (https://www.pymol.org) plugins integrated with various features of PyMOL for extensive molecular analysis. Major limitation of some of these tools is dependency on other software and specific system requirements. The Cytoscape plugins, designed for a specific version, are not updated with every version, while Pymol plugins, designed with the objective of specific analyses, may not be extended to network analysis. With the tremendous potential of network based analysis in different areas of protein science, there is a need for an integrated platform that would provide a wide variety of options for visualization and analysis of protein structures.

We have developed an online tool for **N**etwork based **A**nalysis of **P**rotein **S**tructures (NAPS). It provides following features for analysis and interactive visualization of

*To whom correspondence should be addressed. Tel: +91 40 6653 1183; Fax: +91 40 6653 1413; Email: nita@iiit.ac.in

PCNs: (i) obtaining PDB (single chain, protein complex or model structure), (ii) selecting network type to be constructed with unweighted or weighted edges, (iii) customizing interactive 3D network view, 2D contact view and 2D distance matrix view with 3D structure view, (iv) performing centrality analysis, (v) highlighting residues based on their physico-chemical properties, (viz., hydrophobic, hydrophilic and charged), (vi) analysis of shortest path between a pair of residues, (vii) visual analysis of interactions in multi-domain proteins, multimeric proteins and protein–protein complexes, (viii) spectral analysis of adjacency and Laplacian matrices, (ix) performing *k*-clique analysis and (x) download options for analyses results in text files and high resolution plots and images. A brief description of the various features in NAPS is discussed below.

NETWORK CONSTRUCTION

First step in network construction is to get the protein structure information in PDB format. The user may either provide a four-letter PDB code (and the PDB file is fetched from the PDB mirror from the backend), or the user may upload the PDB file from local machine. NAPS also supports output of a model structure obtained from any molecular modelling program (in PDB format). In NAPS, construction of five different network representations is provided, summarized in Table 1, both with weighted and unweighted edges. The definition of a node and edge varies in each representation depending on the type of analysis required (5,11–12,27–47). This interaction information is saved in edge list format and can be downloaded.

C_α network

C_α atom of an amino acid residue is considered as node and an edge is drawn if the C_α–C_α distance between a pair of residues is within a threshold distance, R_c (∼7 Å). It is one of the simplest and most widely analysed protein contact network that captures very well the 3D topology of protein structure.

C_β network

Side-chain C_β atom of an amino acid residue is considered as node with an edge drawn if the C_β–C_β distance (C_α for Gly) between two residues is ≤ threshold R_c (∼7 Å). It captures the 3D topology of the protein fold through side-chain interactions.

Atom pair contact network

Geometric centre of an amino acid residue defines a node and an edge is constructed if the distance between any two atoms of the corresponding residue pair is $\leq R_c$ (∼5 Å). This network representation provides analysis of protein structure at the atomic level by capturing interactions between any pair of atoms.

Centroid network

The centre of mass of an amino acid residue is defined as node with an edge drawn if the distance between the centroids of the two residues is $\leq R_c$ (∼8.5 Å).

Interaction strength network

Geometric centre of the side chain of an amino acid residue is considered as node and an edge is drawn if the interaction strength between two residues, given by $I_{ij} = [n_{ij}/\sqrt{(N_i * N_j)}]*100$, is $\geq I_c$, the threshold interaction strength (default 4%). Here n_{ij} is the number of side chain atom pairs of residues i and j within 4.5 Å; N_i and N_j are the normalization factor for residues types i and j (5,27).

Long range interaction network (LIN)

For each of the above five representations, a network may be constructed by only considering edges between residues that are sequentially separated by 10–12 residues along the protein backbone to capture long range interactions (11). It can also be alternatively constructed by giving a lower distance threshold to neglect edges between spatially very close residues. LINs depict the interactions between secondary structural elements and define the topology of protein structure (28). NAPS also provide options to construct networks with a smaller separation of 2–3 residues to reduce noise in network construction (29).

NETWORK VISUALIZATION

Visualization is an important aspect in the analysis of structure–function relationship of proteins. NAPS provide four different visualization options—3D protein structure and network view, 2D contact map and distance matrix views. When a user selects node(s) of a PCN in the network view/contact map/distance matrix, the corresponding residues in the protein structure view are automatically highlighted in JSmol applet. The biological relevance of network features and their association to the protein structure can thus be evaluated with this powerful interactive visualization feature.

3D structure view

The protein structure is displayed in ribbon format using browser independent open source JavaScript applet JSmol (48). All the functions of JSmol are embedded in NAPS 3D viewer.

3D network view

In the network view, nodes are represented as spheres in 'blue' and edges between them represented by 'grey' lines with the backbone represented in 'dark blue'. To capture the 3D topology of protein structure, nodes are drawn with 3D coordinates of representative atoms in each network type (Table 1) using the information in PDB file. On hovering the mouse over a node in the network view, the node is highlighted in 'magenta' colour and chain id, residue number and type of the amino acid corresponding to the node are displayed. Three highlight options are provided for node selection:

Highlight node. Node(s) can be selected by clicking on it in the network view and the corresponding residue(s) on the JSmol applet is also highlighted in 'red' colour.

Table 1. Five Network representations in NAPS with weighted edges

Network type	Node	R_c	Edge weight	Purpose
C_α	C_α atom	7Å	$w_{ij} = 1/d_{ij}$	Proteins fold analysis (29,35), inter- and intra-molecular communications (36), protein folding kinetics (11), folding of structural repeats (31,32), domain identification (12,33,37).
C_β	C_β atom	7Å	$w_{ij} = 1/d_{ij}$	Protein dynamics (38), identification of key residues (39), side chain clusters (27) and binding cavities (40)
Atom pair contact	Geometric centre of residue	5Å	$w_{ij} = m_{ij}$	Protein folding mechanism (28), identifying repeats and domains (30,41), analyses of allosteric communication (34) and physicochemical properties (42)
Centroid	Centre of mass of residue	8.5Å	$w_{ij} = 1/d_{ij}$	Protein core and exposed residue analysis (43), critical residues for protein function (44)
Interaction strength	Geometric centre of side chain	4%	$w_{ij} = I_{ij}$	Protein thermo-stability (5), DNA binding proteins (45,46), communication pathways in tRNA synthetase (47)

d_{ij}: distance between nodes *i* and *j*; m_{ij}: number of atom pairs of residues *i* and *j* within threshold, R_c; Interaction strength, $I_{ij} = [n_{ij}/\sqrt{(N_i * N_j)}]*100$; $n_{ij}$is the number of side chain atom pairs of residues *i* and *j* within 4.5 Å; N_i, N_j: interaction strength normalization factors for residue types *i* and *j*. For unweighted networks, weight, $w_{ij} = 1$.

Highlight neighbour. On selecting a node with highlight selection as 'neighbour', the node is highlighted in 'red' colour, and its immediate neighbours are highlighted in 'yellow' colour in both the network view and protein structure view, as shown in Figure 1A and B. This is an important feature for ease in visualization of the immediate neighbours of a node in a dense network.

Show residues. The 3D network view can be used for visual analysis of sub-networks based on physicochemical properties of the amino acid residues. A sub-network representation of *hydrophobic*, *hydrophilic* or *charged* residues is created by highlighting the selected residues in 'red', and showing edges only between residues of the chosen type. These residues are simultaneously highlighted in the JSmol applet.

The visual exploration of PCNs often include the study of molecular interactions of active site residues and ligand binding residues. For this purpose, NAPS provide the user options to select a set of residues in the network view. Similarly, the user can also identify interacting residues in the binding interface of two distinct protein domains or between two monomers in multimeric proteins and in complexes of interacting proteins. In protein complexes, the interacting edges are coloured in 'magenta' for highlighting the interface residues involved in protein–protein interactions.

Domain view

The user can provide the information regarding number of domains in the protein and their range, and visualize them highlighted in different colours; in both network and structure views. NAPS supports visualization of both contiguous and non-contiguous domains. The user can easily identify the interface residues and analyse inter-domain interactions by highlighting them in 'red' colour, on selection.

Contact map

Contact map view is a powerful tool for comparison and visualization of structural features and helps in elucidating long range interactions in proteins (30). For example, it can be useful in analysing changes in the contact patterns of the active site and inhibitor binding site residues in inhibitor-bound and unbound states of a protein (30). It is a 2D dot matrix representation of the network: a dot (*i*, *j*) represents an edge between nodes *i* and *j*, and is a symmetric matrix (Figure 1C). The number and extent of secondary structures, helices and strands, can be easily visualized in this representation (49): two helices in protein 1CRN (chain A) can be identified as lines parallel to the diagonal, marked H_1 and H_2 and anti-parallel strands seen as pattern of dots perpendicular to the diagonal, marked S (Figure 1C). The 2D structure representation captures sufficient information to efficiently reconstruct the 3D protein structure, resulting in dimension reduction in storing information of a protein molecule (50). The contact map view in NAPS is interactive, i.e. selecting a dot (*i*,*j*) highlights the residues *i* and *j* forming the edge in the JSmol applet.

Distance matrix

It provides a 2D colour-coded representation of Euclidean distance/interaction strength between pair of nodes in a network (Figure 1D). It clearly highlights the type and strength of interaction between the residues and one can easily visualize strongly interacting cluster of residues, e.g. domains. The contact map is an abstract level extraction of the distance matrix. The distance matrix can guide the user to choose an appropriate distance cutoff for network construction.

NETWORK ANALYSIS

Global properties

On selecting the type of network to be constructed, the global properties depicting the overall topological properties of the network such as number of nodes and edges, diameter, clustering coefficient, average degree and path length are displayed. It is observed that the number of edges vary with the type of network chosen and the threshold, R_c, used for construction of edges. On varying R_c, we may get a connected graph or clusters of disjoint components. The diameter gives information regarding the shape of the protein—is it a compact globular protein or an elongated

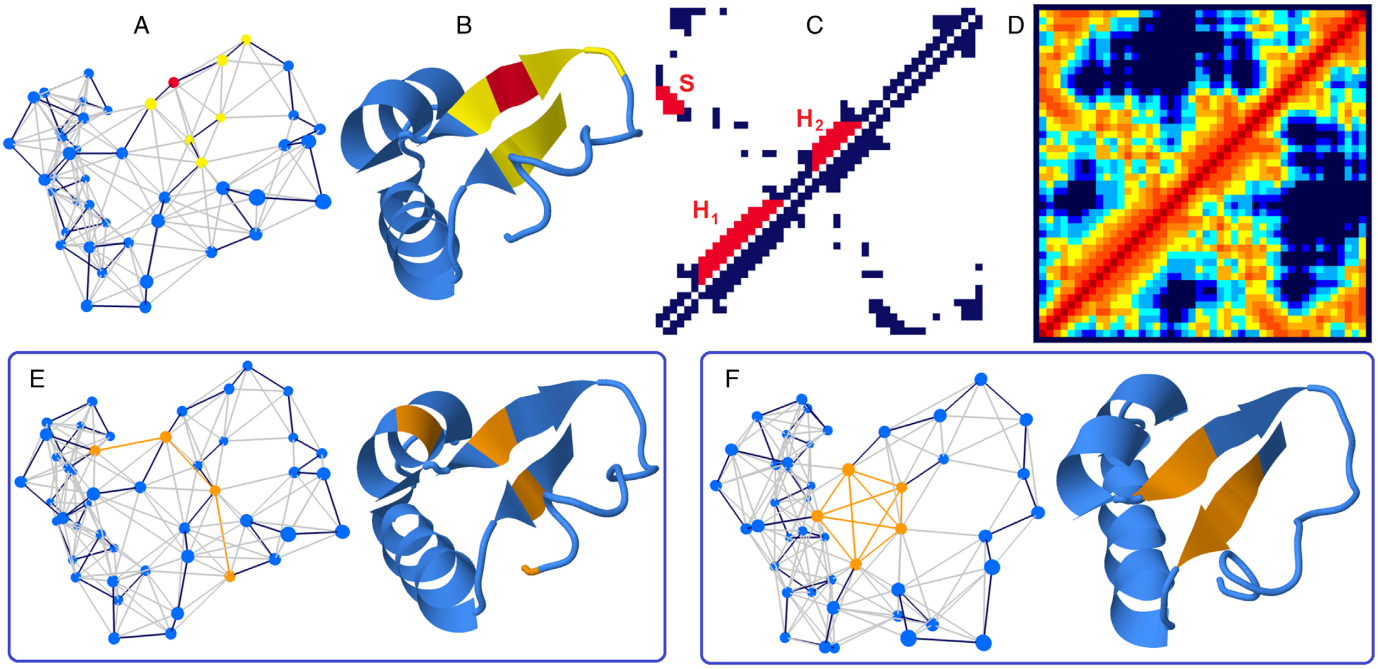

Figure 1. Network analysis of protein structure for 1CRN (chain A). (**A**) Network view showing the 3D network representation with nodes and backbone edges represented in blue and other edges represented by grey. (**B**) JSmol applet showing the 3D protein structure. On clicking a node in network view with neighbour selection option, the node in (A) and the corresponding residue in (B) are highlighted in red, and its immediate neighbouring nodes in (A) and corresponding residues in (B) are highlighted in yellow. (**C**) Contact map view with edges within a helix and strands highlighted in red colour. (**D**) Distance matrix view. (**E**) Shortest path analysis highlighting one of the paths between residues A27 and A46. (**F**) A Clique of size 5 shown.

lamellar protein. Even for the same protein, clustering coefficient of different network types give indication about the average connectivity in the neighbourhood of a node.

Centrality analysis

Centrality identifies the most central, most important, or most significant nodes in a network. Centrality is not defined by a single index, but rather by several indices in correspondence to structural aspects of the interactions that a researcher may intend to focus on. For example, it has been shown that residues crucial for 3D-fold or function are generally high centrality nodes (51), however this association is dependent on the active site shape and mechanism of interaction of the protein (52). The nodes with similar eigenvector profile provide clustering information and have been shown useful in identifying structural motifs/domains (27,31–32). In NAPS the following node centrality measures (both weighted and unweighted) are computed: degree, shortest path, closeness and betweenness, clustering coefficient, eigenvector centrality, eccentricity, average nearest neighbour degree (ANN degree) and edge betweenness centrality (Table 2). NAPS provide various options to examine the results: (i) inspect raw values in a sortable table, (ii) highlight a selected node and its neighbours, or a set of high-centrality nodes and visualize them in the 3D network view and 3D structure view simultaneously, along with highlighting them in the centrality table, (iii) colour code the entire network according to the centrality values, (iv) analyse the node centralities according to hydrophobicity of amino acid residues (Kyte and Doolittle (53)), (v) plot the results in a 2D graph, (vi) save the values in a tab-delimited format and download the network view, structure view and plots as PNG files. Selecting nodes with high-centrality values allow the user to carry out further investigations of their functional and structural characteristics.

Edge betweenness is an important edge centrality and shows the topological importance of edges in the network. It can be correlated to the interaction between two regions in a structure, viz. domain boundaries and interface in multimeric proteins and protein complexes enabling inter-domain and protein–protein interactions.

Graph spectral analysis

The eigenvalues and vector components of the Adjacency and Laplacian matrices associated with a protein contact network provide information on the structure and topology of the graph and analysis of these quantities is known as graph spectral analysis. The graph spectral analysis has yielded valuable results in the identification of clusters in protein structures (54). For the analysis of structural domains and repeats, NAPS provides eigenspectra analysis of: (i) eigenvector components of the principal eigenvalue of the adjacency matrix, and (ii) eigenvector components corresponding to the second smallest eigenvalue of the Laplacian matrix. The spectral analysis of these matrices have been successfully used in the identification of structural repeats and domains (31–33). Typically, residues with similar magnitude of the vector components belong to one cluster and the one with the maximum value corresponds to the centre of the cluster. (54). The NAPS interface for graph spectral analysis is similar to centrality analysis providing selection of a node(s) and its neighbours, sorting of the

Table 2. Centrality measures provided in NAPS for both weighted and unweighted PCNs (computed using igraph (58) and networkx (59) packages respectively)

Name	Description	Definition
Degree	Number of direct neighbours of a node	$C_d(u) = \sum_{v \in V} A_{uv}$ A: adjacency matrix; V: set of all nodes, *u*, *v*: nodes
Closeness	Inverse of the shortest path distance of the node to all other nodes in the network.	$C_{cl}(u) = (n-1)/\sum_{v \in V} dist(u,v)$ *n*: number of nodes in the network dist(*u*,*v*)[b]: shortest path distance between nodes *u* and *v*
Betweenness	Ratio of all the shortest paths passing through a node and the total number of shortest paths in the network.	$C_b(u) = \sum_{s \neq u \in V} \sum_{t \neq u \in V} \sigma_{st}(u)/\sigma_{st}$ σ_{st}[b]: shortest path between *s* and *t* $\sigma_{st}(u)$[b]: shortest path between *s* and *t* passing through *u*.
Clustering coefficient	Ratio of number of connected neighbours of a node to total number of connections possible between the neighbours.	$C_{cc}(u) = \lambda(u)/\tau(u)$ $\tau(u) = C_d(u)(C_d(u)-1)/2$ λ(u): neighbours of *u* connected by an edge.
Eigenvector centrality	Eigenvector component corresponding to largest eigenvalue of adjacency matrix.	$x_i = \frac{1}{\lambda}\sum_{j=1}^{N} A_{ij}x_j$ A_{ij}: ij^{th} element of Adjacency matrix; λ: largest eigen value of A x_i: eigenvector centrality of node *i*
Eccentricity	Shortest path distance of the node to the farthest node in the network.	$C_e(u) = \max(dist(u,v))$
Average nearest neighbour degree	Average of the degree of its immediate neighbours.	$C_{an}(u) = \sum_{v \in N(u)} C_d(u)/N(u)$ N(u): neighbours of u.
Strength[a]	Weighted degree represented by cumulative weights of all the edges connected to a node.	$C_s(u) = \sum_{v \in N(u)} w_{uv}$ w_{uv}: weight of the edge joining nodes *u* and *v*
Edge betweenness	Ratio of all the shortest paths passing through an edge and the total number of shortest paths in the network.	$C_{eb}(e) = \sum_{u \in V} \sum_{v \in V} \sigma_{uv}(e)/\sigma_{uv}$ σ_{uv}: shortest path between u and *v* σ_{uv}(e): shortest path between *u* and *v* passing through edge *e*.

[a]In weighted network.
[b]Weighted shortest path distance is considered in weighted networks.

residues according to spectral components and for visual analysis, the colour coding of network and structure view according to eigenvector components and hydrophobicity values.

Shortest path analysis

Shortest path distance between two nodes is the minimum number of nodes traversed to reach from one node to the other and displays the path of long range interaction in the protein molecule. The residues in the shortest path have been experimentally found to take part in allosteric communication where inter-molecular signal propagates from one functional site of the protein to the other (34). For a pair of selected nodes, NAPS lists all possible shortest paths, and selecting a path highlights the nodes and edges along that path in orange colour in both the network view and JSmol applet (see Figure 1E).

k-clique analysis

A *k*-clique is a sub-network of *k* nodes where all the *k* nodes are connected to each other. In protein structures, *k*-cliques represent compact structural regions with large number of interactions among the residues, indicating structurally stable regions of the protein. The conformational changes in the protein have been analysed by the change in the distribution of *k*-cliques during allosteric regulation in tRNA synthetase (55). The percolation of *k*-cliques in a network has been used in understanding protein folding (56). On selecting a *k* value, all the *k*-cliques with the residues involved are listed and any clique can be highlighted by selecting it from the list (see Figure 1F).

ANALYSIS OF PROTEIN COMPLEX

All the features discussed above for the analysis of a single protein can also be performed on a protein complex formed by two interacting protein chains (of same or different proteins). Network parameters and various features are computed by considering all the nodes in the two pro-

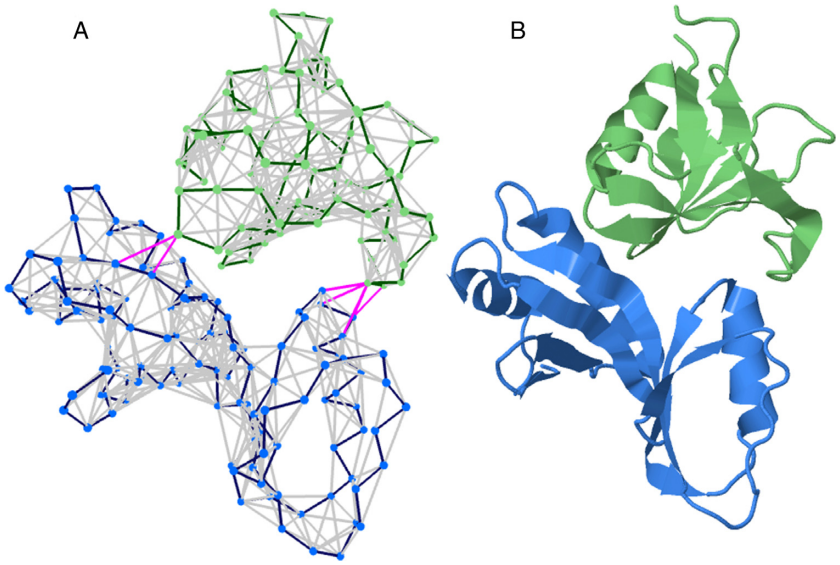

Figure 2. Protein complex 3FPN (chain A and B). (**A**) Network view showing 3D network representation with chain A and B coloured in blue and green colours respectively. The edges joining nodes of chain A and B are highlighted in magenta colour. (**B**) 3D protein structure in JSmol applet.

tein chains as one single network. However, for visualization purposes, the two chains are shown in different colours in both the network view and JSmol applet (see Figure 2A and B). To highlight the interacting residues, the inter-chain edges connecting nodes of two different chains are represented in 'magenta' colour while all non-backbone intra-chain edges are shown in 'grey'. Statistically significant high betweenness residues have been observed to have high correlation with experimentally proven hot spots of interaction (57). The residues at the interface of the complex have high betweenness values as the shortest paths between nodes of the two protein chains pass through these nodes. NAPS provides a platform for extensive network analysis of protein complexes which can be used for better understanding of the protein–protein interactions.

IMPLEMENTATION

NAPS takes a protein structure (experimental or modelled) in PDB format as input for network analysis. It can support single chain as well as multimeric proteins and protein complexes. Depending on the type of analysis to be performed, the user needs to select chain(s), network type, weighted or unweighted, threshold for network construction and residue separation along the backbone chain. The default threshold for each network type has been obtained from the literature and is auto filled on choosing a network type. The backend computation for network construction is carried out using python scripts, and network analyses are performed by C and python programs using well established graph libraries igraph (58) and Networkx (59). The input parameters and the results are displayed using frontend HTML and PHP pages. Once a network is constructed, the various analyses can be performed by navigating through the top menu bar.

Visualization of large protein structures and their PCNs is a challenging task for web portals. NAPS addresses this issue by minimizing the server to client site data exchange. For any analysis, the backend programs are executed on the server side and the results are sent to the client site browser. The visualization of protein structure and network structure are carried out through interactive browser independent JavaScript applets using WebGL API which utilize the system resources at user's end. Thus, limitation on the size of protein structure that can be analysed using NAPS is governed by the system configuration of the user. Interactive applets with inter-applet communication allow the user to select nodes/contacts in various network and visual analyses and the cenrtality tables. The computationally efficient and easy to use applets enhance the user experience and provide quick and efficient way of protein structure analysis.

Download options

The portal is build with an objective to provide a single interface for end-to-end network based analysis of protein structures and provide high resolution images which can be used for publication. The edgelist of the network can be downloaded in text format, which can be used for further analysis by the user. The global parameters, node/edge centrality values, all shortest paths between a pair of nodes and all k-cliques of size k can be downloaded in text files. One can generate the plot of various centrality measures which can be downloaded in PNG format. The JavaScript applets for protein structure (JSmol), network 3D structure (network view), contact map view and distance matrix view provide an option to download the snapshot of the applet in PNG format.

CONCLUSION

For structural biologists, an explosion of protein structure data and exponential increase in computing power has prompted them to look for alternative, computationally efficient approaches for structural analysis of proteins. Although the principles of graph theory have been known

for over a century, exciting applications of network based approach has contributed substantially to our understanding of protein structure, folding, stability, function and dynamics. In NAPS we provide an interactive platform to analyse the relevance of a node/edge based on centrality measures, all possible shortest paths between a pair of nodes, and cluster of fully connected nodes, *k*-cliques, along with interactive network and structure views. It has been shown that residues crucial for 3D-fold or function generally are high degree/betweenness nodes, while nodes with similar eignevector profile provide clustering information and useful in identifying structural motifs/domains. Shortest path analysis can help in analysing allosteric interaction between residue pairs, and structural rigidity/plasticity with the analysis of *k*-cliques in protein–ligand complexes. The network analysis of protein complex can be useful in predicting coevolving residues in interacting proteins. Network view is useful in identifying long-range connections between residues, contact map highlights the secondary structure arrangement, while distance matrix view provides insight into local and long-range interactions and, thus can be useful in cluster analysis. It supports various analyses and visualization views for identifying functional residues including potential targets for therapeutic intervention, provide insight into the mechanisms of protein folding and protein–protein interactions, and for understanding the communication within and between proteins. NAPS provides various features for protein structure analysis on a single platform and has scope for expansion, such as identification of structural motifs, repeats, domains and structure–structure comparison.

FUNDING

DBT, Government of India for Research Fellowship [BT/PR-14715/PBD/16/903/2010 to B.C.]. The open access publication charge for this paper has been waived by Oxford University Press - *NAR*.
Conflict of interest statement. None declared.

REFERENCES

1. Grewal,R.K. and Roy,S. (2015) Modeling proteins as residue interaction networks. *Protein Pept. Lett.*, **22**, 923–933.
2. Yan,W., Zhou,J., Sun,M., Chen,J., Hu,G. and Shen,B. (2014) The construction of an amino acid network for understanding protein structure and function. *Amino Acids*, **46**, 1419–1439.
3. Di Paola,L., De Ruvo,M., Paci,P., Santoni,D. and Giuliani,A. (2013) Protein contact networks: an emerging paradigm in chemistry. *Chem. Rev.*, **113**, 1598–1613.
4. Greene,L.H. (2012) Protein structure networks. *Brief. Funct. Genomics*, **11**, 469–478.
5. Brinda,K.V. and Vishveshwara,S. (2005) A network representation of protein structures: implications for protein stability. *Biophys. J.*, **89**, 4159–4170.
6. Taylor,N.R. (2013) Small world network strategies for studying protein structures and binding. *Comput. Struct. Biotechnol. J.*, **5**, e201302006.
7. Böde,C., Kovács,I.A., Szalay,M.S., Palotai,R., Korcsmáros,T. and Csermely,P. (2007) Network analysis of protein dynamics. *FEBS Lett.*, **581**, 2776–2782.
8. Amitai,G., Shemesh,A., Sitbon,E., Shklar,M., Netanely,D., Venger,I. and Pietrokovski,S. (2004) Network analysis of protein structures identifies functional residues. *J. Mol. Biol.*, **344**, 1135–1146.
9. Blacklock,K. and Verkhivker,G.M. (2014) Computational modeling of allosteric regulation in the Hsp90 chaperones: a statistical ensemble analysis of protein structure networks and allosteric communications. *PLoS Comput. Biol.*, **10**, e1003679.
10. Tse,A. and Verkhivker,G.M. (2015) Molecular dynamics simulations and structural network analysis of c-Abl and c-Src kinase core proteins: capturing allosteric mechanisms and communication pathways from residue centrality. *J Chem. Inf. Model.*, **55**, 1645–1662.
11. Bagler,G. and Sinha,S. (2007) Assortative mixing in Protein Contact Networks and protein folding kinetics. *Bioinformatics*, **23**, 1760–1767.
12. Sistla,R.K., K V,B. and Vishveshwara,S. (2005) Identification of domains and domain interface residues in multidomain proteins from graph spectral method. *Proteins*, **59**, 616–626.
13. Chakrabarty,B. and Parekh,N. (2012) Analysis of graph centrality measures for identifying Ankyrin repeats. In: *2012 World Congress on Information and Communication Technologies (WICT)*. IEEE, Trivandrum, pp. 156–161.
14. Aftabuddin,M. and Kundu,S. (2010) AMINONET- a tool to construct and visualize amino acid networks, and to calculate topological parameters. *J. Appl. Cryst*, **43**, 367–369.
15. Vijayabaskar,MS, Niranjan,V. and Vishveshwara,S. (2011) GraProStr—graphs of protein structures: a tool for constructing the graphs and generating graph parameters for protein structures. *Open Bioinform. J.*, **5**, 53–58.
16. Chung,J.-L., Beaver,J.E., Scheeff,E.D. and Bourne,P.E. (2007) Con-Struct Map: a comparative contact map analysis tool. *Bioinformatics*, **23**, 2491–2492.
17. Kozma,D., Simon,I. and Tusnády,G.E. (2012) CMWeb: an interactive on-line tool for analysing residue-residue contacts and contact prediction methods. *Nucleic Acids Res.*, **40**, W329–W333.
18. Biro,J.C. and Fördös,G. (2005) SeqX: a tool to detect, analyze and visualize residue co-locations in protein and nucleic acid structures. *BMC Bioinformatics*, **6**, 170.
19. Pietal,M.J., Tuszynska,I. and Bujnicki,J.M. (2007) PROTMAP2D: visualization, comparison and analysis of 2D maps of protein structure. *Bioinformatics*, **23**, 1429–1430.
20. Vehlow,C., Stehr,H., Winkelmann,M., Duarte,J.M., Petzold,L., Dinse,J. and Lappe,M. (2011) CMView: interactive contact map visualization and analysis. *Bioinformatics*, **27**, 1573–1574.
21. Martin,A.J.M., Vidotto,M., Boscariol,F., Di Domenico,T., Walsh,I. and Tosatto,S.C.E. (2011) RING: networking interacting residues, evolutionary information and energetics in protein structures. *Bioinformatics*, **27**, 2003–2005.
22. Doncheva,N.T., Klein,K., Domingues,F.S. and Albrecht,M. (2011) Analyzing and visualizing residue networks of protein structures. *Trends Biochem. Sci.*, **36**, 179–182.
23. Nepomnyachiy,S., Ben-Tal,N. and Kolodny,R. (2015) CyToStruct: augmenting the network visualization of cytoscape with the power of molecular viewers. *Structure*, **23**, 941–948.
24. Shannon,P., Markiel,A., Ozier,O., Baliga,N.S., Wang,J.T., Ramage,D., Amin,N., Schwikowski,B. and Ideker,T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
25. Pasi,M., Tiberti,M., Arrigoni,A. and Papaleo,E. (2012) xPyder: a PyMOL plugin to analyze coupled residues and their networks in protein structures. *J. Chem. Inf. Model.*, **52**, 1865–1874.
26. Tiberti,M., Invernizzi,G., Lambrughi,M., Inbar,Y., Schreiber,G. and Papaleo,E. (2014) PyInteraph: a framework for the analysis of interaction networks in structural ensembles of proteins. *J. Chem. Inf. Model.*, **54**, 1537–1551.
27. Kannan,N. and Vishveshwara,S. (1999) Identification of side-chain clusters in protein structures by a graph spectral method. *J. Mol. Biol.*, **292**, 441–464.
28. Greene,L.H. and Higman,V.A. (2003) Uncovering network systems within protein structures. *J. Mol. Biol.*, **334**, 781–791.
29. Patra,S.M. and Vishveshwara,S. (2000) Backbone cluster identification in proteins by a graph theoretical method. *Biophys. Chem.*, **84**, 13–25.
30. Hrabe,T. and Godzik,A. (2014) ConSole: using modularity of contact maps to locate solenoid domains in protein structures. *BMC Bioinformatics*, **15**, 119.
31. Chakrabarty,B. and Parekh,N. (2014) Identifying tandem Ankyrin repeats in protein structures. *BMC Bioinformatics*, **15**, 6599.
32. Chakrabarty,B. and Parekh,N. (2014) PRIGSA: protein repeat identification by graph spectral analysis. *J. Bioinform. Comput. Biol.*, **12**, 1442009.

33. Yalamanchili,H.K. and Parekh,N. (2009) Graph spectral approach for identifying protein domains. *Bioinformatics and Computational Biology*. Springer, Berlin Heidelberg, Vol. **5462**, pp. 437–448.
34. del Sol,A., Fujihashi,H., Amoros,D. and Nussinov,R. (2006) Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol. Syst. Biol.*, **2**, 2006.0019.
35. Paola,L.D., Paci,P., Santoni,D., Ruvo,M.D. and Giuliani,A. (2012) Proteins as sponges: a statistical journey along protein structure organization principles. *J. Chem. Inf. Model.*, **52**, 474–482.
36. Vishveshwara,S., Ghosh,A. and Hansia,P. (2009) Intra and inter-molecular communications through protein structure network. *Curr. Protein Pept. Sci.*, **10**, 146–160.
37. Ryslik,G.A., Cheng,Y., Cheung,K.-H., Modis,Y. and Zhao,H. (2014) A graph theoretic approach to utilizing protein structure to identify non-random somatic mutations. *BMC Bioinformatics*, **15**, 86.
38. Atilgan,A.R., Akan,P. and Baysal,C. (2004) Small-world communication of residues and significance for protein dynamics. *Biophys. J.*, **86**, 85–91.
39. Atilgan,A.R., Turgut,D. and Atilgan,C. (2007) Screened nonbonded interactions in native proteins manipulate optimal paths for robust residue communication. *Biophys. J.*, **92**, 3052–3062.
40. Estrada,E. (2010) Universality in protein residue networks. *Biophys. J.*, **98**, 890–900.
41. Xu,Y., Xu,D., Gabow,H.N. and Gabow,H. (2000) Protein domain decomposition using a graph-theoretic approach. *Bioinformatics*, **16**, 1091–1104.
42. Aftabuddin,M. and Kundu,S. (2007) Hydrophobic, hydrophilic, and charged amino acid networks within protein. *Biophys. J.*, **93**, 225–231.
43. Alves,N.A. and Martinez,A.S. (2007) Inferring topological features of proteins from amino acid residue networks. *Physica A*, **375**, 336–344.
44. Cusack,M.P., Thibert,B., Bredesen,D.E. and Del Rio,G. (2007) Efficient identification of critical residues based only on protein structure by network analysis. *PLoS One*, **2**, e421.
45. Sathyapriya,R., Brinda,K.V. and Vishveshwara,S. (2006) Correlation of the side-chain hubs with the functional residues in DNA binding protein structures. *J. Chem. Inf. Model.*, **46**, 123–129.
46. Sathyapriya,R., Vijayabaskar,M.S. and Vishveshwara,S. (2008) Insights into protein-DNA interactions through structure network analysis. *PLoS Comput. Biol.*, **4**, e1000170.
47. Ghosh,A. and Vishveshwara,S. (2007) A study of communication pathways in methionyl- tRNA synthetase by molecular dynamics simulations and structure network analysis. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 15711–15716.
48. Hanson,R.M., Prilusky,J., Renjian,Z., Nakane,T. and Sussman,J.L. (2013) JSmol and the Next-Generation Web-Based Representation of 3D Molecular Structure as Applied to Proteopedia. *Isr. J. Chem.*, **53**, 207–216.
49. Barah,P. and Sinha,S. (2009) Analysis of protein folds using protein contact networks. *Pramana*, **71**, 369–378.
50. Duarte,J.M., Sathyapriya,R., Stehr,H., Filippis,I. and Lappe,M. (2010) Optimal contact definition for reconstruction of contact maps. *BMC Bioinformatics*, **11**, 283.
51. Thibert,B., Bredesen,D.E. and del Rio,G. (2005) Improved prediction of critical residues for protein function based on network and phylogenetic analyses. *BMC Bioinformatics*, **6**, 213.
52. del Sol,A., Fujihashi,H., Amoros,D. and Nussinov,R. (2006) Residue centrality, functionally important residues, and active site shape: analysis of enzyme and non-enzyme families. *Protein Sci.*, **15**, 2120–2128.
53. Kyte,J. and Doolittle,R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105–132.
54. Vishveshwara,S., Brinda,K.V. and Kannan,N. (2002) Protein structure: insights from graph theory. *J. Theor. Comput. Chem.*, **1**, 187–211.
55. Ghosh,A. and Vishveshwara,S. (2008) Variations in clique and community patterns in protein structures during allosteric communication: investigation of dynamically equilibrated structures of methionyl tRNA synthetase complexes. *Biochemistry*, **47**, 11398–11407.
56. Deb,D., Vishveshwara,S. and Vishveshwara,S. (2009) Understanding protein structure from a percolation perspective. *Biophys. J.*, **97**, 1787–1794.
57. del Sol,A. and O'Meara,P. (2005) Small-world network approach to identify key residues in protein-protein interaction. *Proteins*, **58**, 672–682.
58. Csardi,G. and Tamas,N. (2006) The igraph software package for complex network research. *Inter J. Complex Syst.*, **2006**, 1695.
59. Hagberg,A., Schult,D. and Swart,P. (2008) Exploring network structure, dynamics, and function using NetworkX. In: *7th Python in Science conference (SciPy 2008)*. Pasadena, pp. 11–15.