BIOINFORMATICS APPLICATIONS NOTE Vol. 27 no. 14 2011, pages 2003–2005
doi:10.1093/bioinformatics/btr191

Structural bioinformatics

Advance Access publication April 14, 2011

RING: networking interacting residues, evolutionary information and energetics in protein structures

Alberto J. M. Martin, Michele Vidotto, Filippo Boscariol, Tomàs Di Domenico, Ian Walsh and Silvio C. E. Tosatto*

Department of Biology, University of Padova, Viale G. Colombo 3, 35131 Padova, Italy

Associate editor: Burkard Rost

ABSTRACT

Motivation: Residue interaction networks (RINs) have been used in the literature to describe the protein 3D structure as a graph where nodes represent residues and edges physico–chemical interactions, e.g. hydrogen bonds or van-der-Waals contacts. Topological network parameters can be calculated over RINs and have been correlated with various aspects of protein structure and function. Here we present a novel web server, RING, to construct physico–chemically valid RINs interactively from PDB files for subsequent visualization in the Cytoscape platform. The additional structure-based parameters secondary structure, solvent accessibility and experimental uncertainty can be combined with information regarding residue conservation, mutual information and residue-based energy scoring functions. Different visualization styles are provided to facilitate visualization and standard plugins can be used to calculate topological parameters in Cytoscape. A sample use case analyzing the active site of glutathione peroxidase is presented.

Availability: The RING server, supplementary methods, examples and tutorials are available for non-commercial use at URL: http://protein.bio.unipd.it/ring/.

Contact: silvio.tosatto@unipd.it

Received on October 9, 2010; revised on March 19, 2011; accepted on April 6, 2011

1 INTRODUCTION

The last 15 years have seen the advent of network representation to tackle the complexity inherent in many problems in biology. By focusing on the network properties of a system it is often possible to gain a new level of insight into apparently unpredictable systems (Albert *et al*., 2000; Barabasi and Albert, 1999). The most studied example in biology is protein interaction networks, where nodes represent proteins and connections between nodes (physical) interactions. Tools have been developed to visualize and analyze such networks, with Cytoscape (Shannon *et al*., 2003) probably being the most widely accepted standard platform due to its open structure and extendibility. More recently, there has been a growing interest in representing protein structures as so-called residue interaction networks (RINs) (Csermely, 2008). RINs consider single amino acids in the protein structure as nodes and connections as physico–chemical interactions, such as covalent bonds and non-covalent contacts (e.g. hydrogen bonds). The intuitive idea is to analyze a structure with the same approach as a protein interaction network in order to investigate whether the same rules apply. Several papers have already used RINs to analyze protein stability and folding, allosteric communication, enzyme catalysis or mutation effect prediction (Buslje *et al*., 2010; Del Sol *et al*., 2007; Dokholyan *et al*., 2002; Soundararajan *et al*., 2010; Suel *et al*., 2003; Swint-Kruse, 2004; Vendruscolo *et al*., 2001). A more complete bibliography is available online.

Here, we present RING as a novel tool to generate RINs for use in Cytoscape. The tool was conceived to yield a simple, intuitive representation that is physico–chemically meaningful and integrates different types of structure-based information with evolutionary information and energy scoring functions.

2 PROGRAM OVERVIEW

RING requires as input a valid PDB identifier or user specified PDB file and provides two user interfaces. A simple user interface with meaningful default values for most parameters is intended for the inexperienced user. A more complex interface where the user can specify exactly how the RIN should be defined is also provided. Three alternative types of network definitions are available, with closest atoms as default. Interactions are defined distinguishing disulfides, salt bridges, hydrogen bonds and aromatic interactions from generic van-der-Waals contacts. Structural features are generated for each node and include secondary structure, solvent accessibility and experimental uncertainty for X-ray structures (i.e. B factor and occupancy). Protein sequence conservation and mutual information (Buslje *et al*., 2009, 2010), determined from PSI-BLAST profiles, and conformational energy preferences determined with FRST and TAP score (Tosatto, 2005; Tosatto and Battistutta, 2007) are to the best of our knowledge unique features of RING. The protocol uses standard programs to derive the data and is described on the website. The output consists in an archive file containing all the necessary network definition files for Cytoscape. Once generated, the RIN can be easily visualized in Cytoscape. Visualization filters are provided in RING to highlight different structural features through the Cytoscape VizMapper feature. Topological network properties can be analyzed with readily available plugins, e.g. NetworkAnalyzer (Assenov *et al*., 2008). Online help pages are provided for the web server together with an extensive documentation on the implemented file format and tutorials.

A special feature of RING is to generate meaningful sub-networks. It has been suggested that RIN complexity can be focused on essential interactions by limiting analysis to buried residues

*To whom correspondence should be addressed.

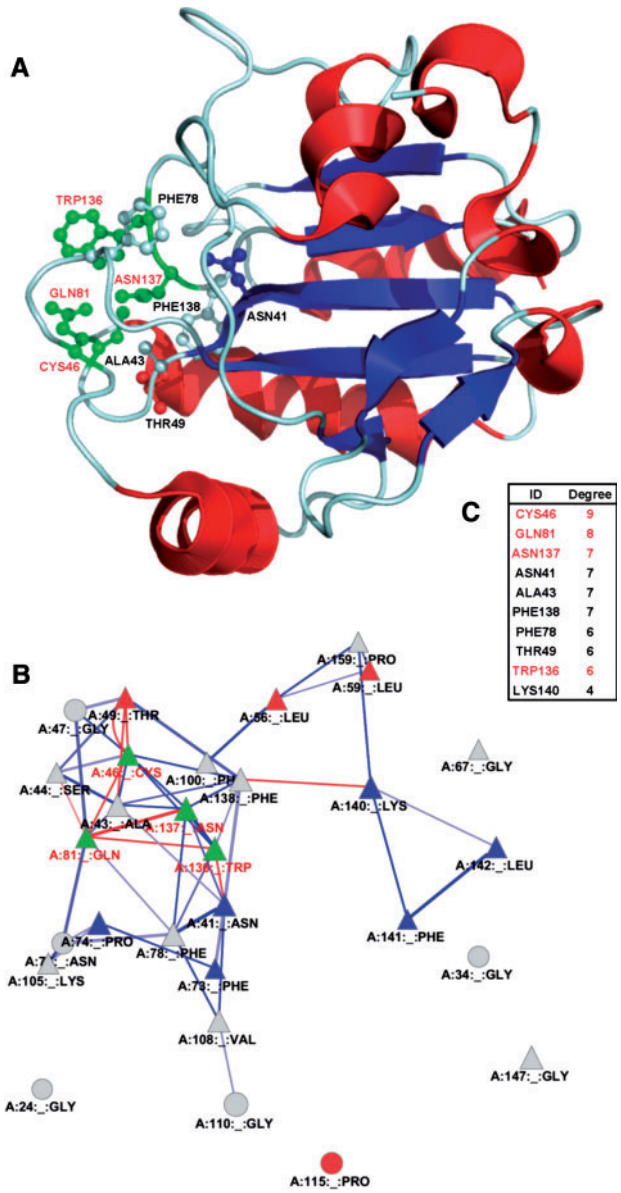

Fig. 1. Protein structure of human glutathione peroxidase 4 (PDB identifier 2obi). The 3D structure is shown as cartoons (**A**) colored by secondary structure. The active site is shown in balls and sticks and labeled. Lighter labels are used for the catalytic tetrad. The corresponding RING residue interaction network in Cytoscape is shown for positions with at least 80% sequence conservation (**B**) with circles representing exposed and diamonds buried residues. Nodes are colored by secondary structure. Edges are colored according to interaction type and width proportional to mutual information. Notice how the most highly connected nodes in the sub-network correspond to the active site and its immediate surroundings. The top 10 most connected nodes of the network are shown in inset (**C**).

(Soundararajan *et al*., 2010). RING allows the user to limit the generated network to buried residues, conserved residues or both. Intuitively, limiting RIN analysis to a network of conserved residues will help to focus on the essential interactions for a protein active site as shown in the following case study.

3 CASE STUDY

The glutathione peroxidase (GPx) enzymes are an evolutionarily conserved family catalyzing the reduction of hydroperoxides to alcohols and the concomitant oxidation of thiols to disulfides (Toppo *et al*., 2008). Long thought to function by means of a catalytic triad, a detailed structural analysis of sequence conservation has recently suggested a fourth residue to be catalytically active which was experimentally confirmed (Tosatto *et al*., 2008).

Figure 1 shows the GPx structure and corresponding RING sub-network of conserved residues. The GPx active site can be easily identified as the nodes with the highest number of connections in the RIN. The combination of evolutionary conservation and basic network topology thus provides an easy way to replicate the work that led to the recent characterization of the GPx active site (Tosatto *et al*., 2008). Moreover, mutual information provided by RING can serve to investigate cases of co-evolution between residues. An online tutorial explains the necessary steps to reproduce the steps to generate the RIN visualization. Several further example are available as part of the online tutorial.

In summary, RING is a novel web tool for use with Cytoscape designed for the visualization and analysis of protein structures in terms of physico–chemical interactions, evolutionary information and energetics while taking advantage of the powerful network paradigm. We anticipate this novel combination to facilitate the functional annotation of proteins.

ACKNOWLEDGEMENTS

The authors are grateful to Mario Albrecht, Francisco Domingues and to members of the BioComputing UP lab for insightful discussions.

Funding: University of Padova (grants CPDA098382, CPDR097328 to S.T.); FIRB Futuro in Ricerca (grant RBFR08ZSXY to S.T.).

Conflict of Interest: none declared.

REFERENCES

Albert,R. *et al.* (2000) Error and attack tolerance of complex networks. *Nature*, **406**, 378–382.

Assenov,Y. *et al.* (2008) Computing topological parameters of biological networks. *Bioinformatics*, **24**, 282–284.

Barabasi,A.L. and Albert,R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.

Buslje,C.M. *et al.* (2009) Correction for phylogeny, small number of observations and data redundancy improves the identification of coevolving amino acid pairs using mutual information. *Bioinformatics*, **25**, 1125-1131.

Buslje,C.M. *et al.* (2010) Networks of high mutual information define the structural proximity of catalytic sites: implications for catalytic residue identification. *PLoS Comput Biol*, **6**, e1000978.

Csermely,P. (2008) Creative elements: network-based predictions of active centres in proteins and cellular and social networks. *Trends Biochem. Sci.*, **33**, 569–576.

Del Sol,A. *et al.* (2007) Modular architecture of protein structures and allosteric communications: potential implications for signaling proteins and regulatory linkages. *Genome Biol.*, **8**, R92.

Dokholyan,N.V. *et al.* (2002) Topological determinants of protein folding. *Proc. Natl Acad. Sci. USA*, **99**, 8637–8641.

Shannon,P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.

Soundararajan,V. *et al.* (2010) Atomic interaction networks in the core of protein domains and their native folds. *PLoS One*, **5**, e9391.

Suel,G.M. *et al.* (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.*, **10**, 59–69.

Swint-Kruse,L. (2004) Using networks to identify fine structural differences between functionally distinct protein states. *Biochemistry*, **43**, 10886–10895.

Toppo,S. *et al.* (2008) Evolutionary and structural insights into the multifaceted glutathione peroxidase (Gpx) superfamily. *Antioxid Redox Sign.*, **10**, 1501–1514.

Tosatto,S.C. (2005) The victor/FRST function for model quality estimation. *J. Comput. Biol.*, **12**, 1316–1327.

Tosatto,S.C. and Battistutta,R. (2007) TAP score: torsion angle propensity normalization applied to local protein structure evaluation. *BMC Bioinformatics*, **8**, 155.

Tosatto,S.C. *et al.* (2008) The catalytic site of glutathione peroxidases. *Antioxid Redox Sign.*, **10**, 1515–1526.

Vendruscolo,M. *et al.* (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature*, **409**, 641–645.