

The RING 2.0 web server for high quality residue interaction networks

Damiano Piovesan¹, Giovanni Minervini¹ and Silvio C.E. Tosatto^{1,2,*}

¹Department of Biomedical Sciences, University of Padua, Padua 35121, Italy and ²CNR Institute of Neuroscience, Padua 35121, Italy

Received February 10, 2016; Revised April 11, 2016; Accepted April 13, 2016

ABSTRACT

Residue interaction networks (RINs) are an alternative way of representing protein structures where nodes are residues and arcs physico-chemical interactions. RINs have been extensively and successfully used for analysing mutation effects, protein folding, domain-domain communication and catalytic activity. Here we present RING 2.0, a new version of the RING software for the identification of covalent and non-covalent bonds in protein structures, including π - π stacking and π -cation interactions. RING 2.0 is extremely fast and generates both intra and inter-chain interactions including solvent and ligand atoms. The generated networks are very accurate and reliable thanks to a complex empirical re-parameterization of distance thresholds performed on the entire Protein Data Bank. By default, RING output is generated with optimal parameters but the web server provides an exhaustive interface to customize the calculation. The network can be visualized directly in the browser or in Cytoscape. Alternatively, the RING-Viz script for Pymol allows visualizing the interactions at atomic level in the structure. The web server and RING-Viz, together with an extensive help and tutorial, are available from URL: <http://protein.bio.unipd.it/ring>.

INTRODUCTION

Non-covalent interactions in proteins have a wide range of different energies and lengths, making them inherently difficult to characterize (1). While the energy contribution of a single interaction is almost negligible, together they determine the three-dimensional protein structure (2). Describing amino acid properties through continuous functions, although highly informative, requires complex calculations and non-trivial analysis. Some effort to extract valuable information through simplification has been done by applying network theory to protein structures (3–5). Residue in-

teraction networks (RINs) consider single amino acids as nodes and physico-chemical interactions, like covalent and non-covalent bonds, as edges. Representing protein structures as RINs has become common practice to explore the complexity inherent in macromolecular systems (6,7). As a consequence, structure analysis has been simplified, allowing to focus only on a subset of relevant residues. According to the concept of ‘residue centrality’ (8), evolutionary conserved (*central*) residues can be identified by just looking at hyper-connected nodes. RINs have been extensively and successfully used for analysing functional features linked to a broad range of biological processes, the effect of mutations, protein folding, intra protein domain-domain communication and catalytic activity (9–16). On the other hand, software generating RINs still has limitations due to the use of simplified interaction types. RINalyzer (17) for example calculates only hydrogen bond, van der Waals (VDW) and generic contacts based on distance. This limitation can be explained by technical reasons, such as the computational cost of measuring the distance of all possible atom pairs in a protein, in particular for large biopolymers. Another problem is defining distance and angle constraints for certain interactions (e.g. involving π -systems) in large molecules like proteins. For example, the Protein Interaction Calculator (PIC) (18), calculates all types of interactions but lacks atomic resolution for most of them and the distance thresholds are simply based on values reported in the literature which may often be obsolete. The PIC output is provided as separate lists with different formats and hence difficult to import in external network viewers such as Cytoscape (19). The Residue Interaction Network Generator (RING) has been presented to address these limitations (20). RING 2.0 is a new and completely rewritten version of the software based on the Victor library (21). Compared with the previous version, RING-2.0 is available as stand-alone package without the need for third party software, e.g. for the calculation of hydrogen bonds, VDW interactions and secondary structure, increasing the calculation speed by an order of magnitude. RING-2.0 is now also able to return both intra and inter chain interactions as well as contacts involving ‘hetero atoms’ (i.e. ligands, DNA/RNA, cofactors, metal ions and solvent molecules). Moreover, distance

*To whom correspondence should be addressed. Tel: +39 49 827 6269; Fax: +39 49 827 6363; Email: silvio.tosatto@unipd.it

thresholds have been optimized to maximize network reliability on a large-scale by analysing the entire Protein Data Bank (PDB) repository (22). The web server is straightforward to use and allows the user to visualize the network directly in the browser and node attributes as interchangeable layers. RING output is compatible with the RINalyzer (17) and StructureViz (23) plugins for Cytoscape (19). The network can be visualized in the structure at atomic level thanks to the RING-Viz utility for Pymol available on the RING web server.

MATERIALS AND METHODS

RING generates an interaction network in two steps. At first, it identifies a list of residue–residue (residue–ligand or ligand–ligand) pairs eligible for interaction based on all-atom distance measurements. Contacts are then characterized by identifying specific interaction types. Interactions are sorted by position, without repetitions and the index of the source node (first column) being always lower than the target. Multiple interactions occurring between the same residue pair, i.e. involving different atoms are sorted by energy and distance. In this case, the user may choose to receive all interactions, only the first (i.e. most energetic), or only one interaction for each type. Sorting is very helpful for manual inspection of the edges. Algorithm complexity is quadratic depending on the protein size (number of atoms). As a worst case example, for 100 000 atoms distributed over 42 chains (PDB ID: 4V6), the entire computation (ca. 5 billion comparisons) takes 18 min (12 for finding contacts and 6 for sorting/filtering them) on a standard laptop.

Interaction type calculation

Table 1 provides an overview of the RING approach to calculate interaction types. Hydrogen bonds are calculated by applying a donor–hydrogen–acceptor (DHA) angle constraint less or equal to 63° (24), defining a limited set of valid donor/acceptor atoms (25). Generally, only Carbon–Carbon and Carbon–Sulphur pairs are considered valid VDW interactions and evaluated using atom surfaces, i.e. subtracting the atom radius, 1.89 Å for sulphur and 1.77 Å for carbon (26), from the atomic distance. VDW is the only specific interaction type also calculated for ligands, since it is not necessary to know the ligand structure. Special VDW cases involving N and O sidechain atoms of glutamine and asparagine are also considered (27–29). Other types are evaluated based on the distance between pseudo-atoms, i.e. the barycentre of aromatic rings (π – π stacking) or the centre of mass of charged groups (ionic interaction). π –cation interactions are limited to cases where the cation projection over the interacting partner ring lies inside the ring itself.

Network attributes

RING generates attributes for both nodes and edges. Structural features are reported for each node and include secondary structure, vertex degree (the number of directly connected nodes), experimental uncertainty for X-ray structures (i.e. C_α B-factor), conformational energy preferences

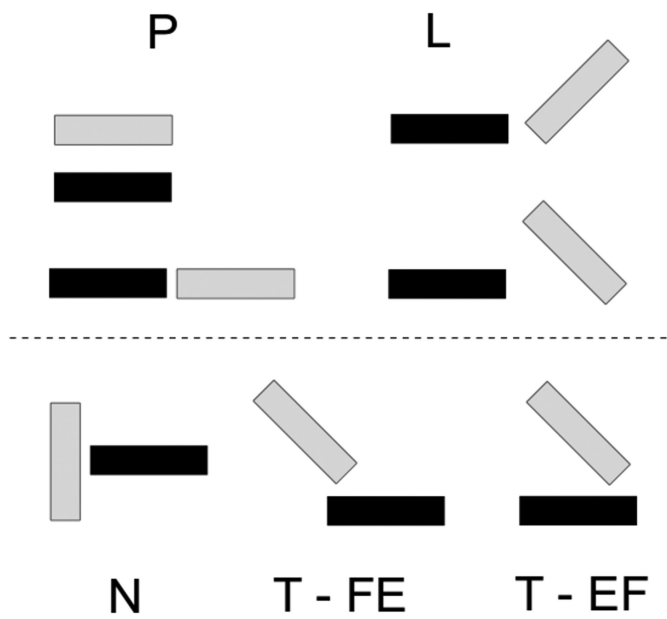


Figure 1. Orientation definition in RING. π – π stacking interactions adopt parallel (P), normal (N), lateral (L), tilted edge to face (T-EF) and tilted face to edge (T-FE) orientations. Orientation in π –cation interactions is provided only for contacts involving arginine and describes guanidine plane positioning relative to the partner ring and it is limited to P, L and N conformations.

determined with FRST (30) and TAP (31), conservation (Shannon entropy) and cumulative mutual information (MI) (32), calculated from PSI-BLAST profiles (33). Interaction energies have been derived from the literature, in particular hydrogen bonds have different energies depending on the donor/acceptor pair (34). Edge attributes include the bond angle (except for VDW), energy and involved atoms. The orientation attribute is calculated only for π – π stacking and π –cation interactions and represents the reciprocal orientation of the two interacting rings and the guanidine group (arginine) positioning relative to the plane of the aromatic ring partner respectively (Figure 1). Moreover, when the sequence profile is calculated RING provides MI and APC corrected MI (32).

Distance threshold optimization

The RING algorithm calculates atomic interactions based on geometrical criteria, without complicated analysis based on force fields, obtaining a reliable interaction network very rapidly. The quality of the interactions strongly depends on geometrical constraints, in particular the distance threshold parameter. An exhaustive analysis has been performed on the entire PDB. The distance distribution of different interaction types has been calculated for the interaction network of all available X-ray and NMR structures (116 568; April 2016). Two different distance thresholds have been chosen for RING to represent strict and permissive parameters (see Figure 2).

Van der Waals. Most VDW interactions involve C–C pairs with a distance of [0.71, 0.74] Å. C–S pairs are mainly found in the [0.19, 0.22] interval (data not shown). The number of

Table 1. RING-2.0 interaction types

Interaction Type	Residue or ligand	Atom		Strict (Å)	Relaxed (Å)	Energy (Kj/mol)
Hydrogen bond	All	Acceptor C (main chain)	Donor NH (main chain)			
	Arg		NE, NH1, NH2			
	Asn	OD1	ND2			
	Asp	OD1, OD2				115.0 ($d \leq 1.5 \text{ \AA}$)
	Cys		SG			
	Gln	OE1	NE2	3.5	5.5	40.0 ($1.5 < d \leq 2.2 \text{ \AA}$)
	Glu	OE1, OE2				
	His	ND1	NE2, ND1			
	Lys		NZ			
	Met	SD				17.0 ($d > 2.2 \text{ \AA}$)
	Ser	OG	OG			
	Thr	OG1	OG1			
	Trp		NE1			
Tyr	OH	OH				
Ionic	Asp	Positive	Negative Charged group mass centre			
	Glu		Charged group mass centre	4.0	5.0	20.0
	Arg	CZ				
	Lys	NZ				
His	Charged group mass centre					
Disulphide	Cys	SG	2.5	3.0	167.0	
Van-der-Waals	All	Any C, any S				
	Gln	NE1, OE1		0.5	0.8	6.0
	Asn	ND2, OD1				
π - cation	Arg	Cation CZ	π -partner			
	Lys	NZ		5.0	7.0	9.6
	Tyr		Closest ring atom			
	Trp		Closest ring atom			
	Phe		Closest ring atom			
π - π stacking	His	Ring barycentre				
	Tyr	Ring barycentre		6.5	7.0	7.0
	Trp	Ring barycentre				
	Phe	Ring barycentre				

For directed interaction types (H bonds, π -cation and Ionic), the atom column is split in source/target pairs. Atom names are reported according to the PDB standard. For VDW, the distance is calculated considering atom surface instead of atom centre. Strict and relaxed correspond to the optimized thresholds available in the web server. For hydrogen bonds, energy is provided as a function of the distance (d).

contacts under the ideal threshold of 0.5 Å is ~74 million and 119 million under a more relaxed cutoff of 0.8. Interestingly, a lot of atom pairs (ca. 11 million) clash, and are shown as negative values in the figure. These come from low quality structures. Also interesting, two secondary peaks at [0.74, 0.78] and [1.20, 1.23] Å correspond to false positive VDW interactions that are found inside α -helices and between close strands in β -sheets respectively.

Hydrogen bonds. The hydrogen bond distribution has been split into main chain (MC) and side chain (SC) interactions. The peak at [2.84, 2.87] Å corresponds to interactions that stabilize the packing of different secondary structure elements, i.e. bridges between alpha-helices or turns. A similar peak for the MC bonds at [2.94, 2.98] corresponds to interactions between adjacent strands in β -sheets, whereas the second MC peak [5.01, 5.04] Å identifies bonds in α -helices separated by a turn (four residues). Over 5.6 Å only spurious interactions are identified.

Salt bridges. Salt bridges (ionic interactions) are defined between a positively charged amino acid (Arg, Lys, His) and a negatively charged residue (Asp or Glu). Both arginine and lysine have a weak preference for interacting with glutamic acid, on the contrary, histidine slightly prefers as-

partic acid (data not shown). The main peak for arginine at [3.63, 3.66] Å represents residues interacting with a planar or orthogonal orientation relative to the ring plane. The second peak at [4.38, 4.41] Å corresponds either to gauche conformations or situations where the interaction is altered by neighbouring forces. The lysine peak at [3.75, 3.78] Å corresponds to interactions occurring when the lysine C_ϵ - N_ζ axis lies on the plane of the carboxylic group of the interaction partner. Above 4 Å this orientation is lost and in most cases false interactions are predicted. Histidine has a peak in the range [4.59, 4.62] Å. Above this distance interactions are spurious.

π - π stacking. π - π stacking interactions involve aromatic side chain rings. According to the reciprocal orientation (Figure 1), it is possible to identify four different categories with different distributions. The orthogonal (N, normal) conformation is found at [5.36, 5.4] Å. Both the parallel (P) and tilted edge-to-face (T-EF) have a maximum at [5.94, 5.99] Å, whereas the L conformation (resembling the letter "L") has a peak at [6.03, 6.08] Å. In general, beyond 7 Å, a π - π interaction is unreliable, since the straight line connecting the rings either passes through other atoms or the side chains point in opposite directions.

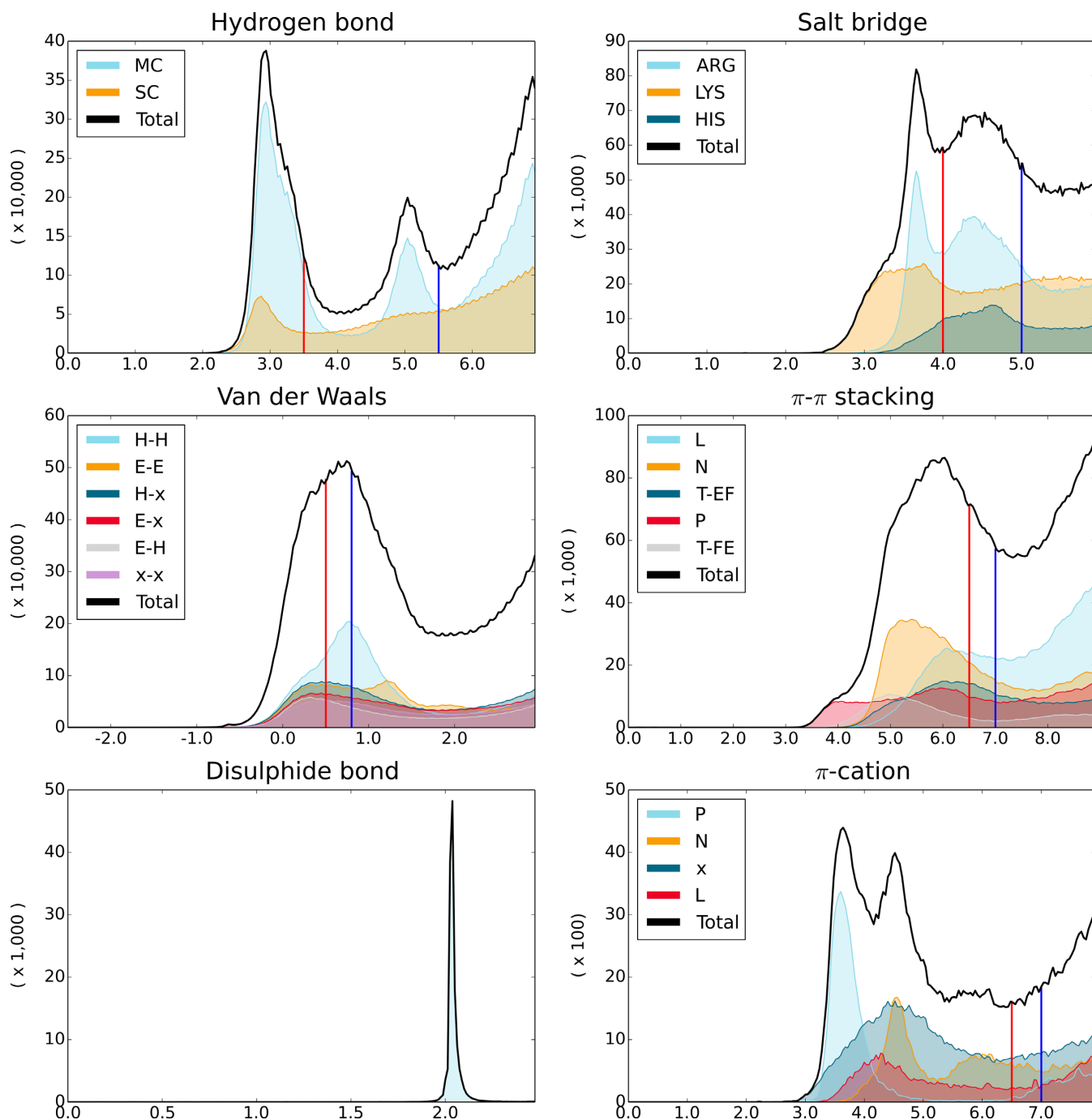


Figure 2. Distance distribution for the six different interaction types. Hydrogen bonds are split into side chain (SC) and main chain (MC). Ionic interactions are characterized by the positively charged residue. Van der Waals interaction by secondary structure of the pair (E = sheet, H = helix, x = undefined). π - π stacking and π -cation interactions are separated by orientation type (see 'Materials and Methods' section). Red and blue vertical lines correspond respectively to the 'strict' and 'relaxed' thresholds in the web server.

π -cation. π -cation interactions involving arginine are characterized by different orientations of the charged group relative to the ring plane of the partner (Figure 1). Parallel, lateral and normal orientations have a peak at [3.64, 3.68] Å, [4.28, 4.32] Å and [4.56, 4.60] Å respectively. The orthogonal (N) conformation presents another relevant peak around [6.12, 6.16] Å, corresponding to a situation where the charged group is found opposite to the interacting ring.

Interactions involving lysine are more difficult to interpret, as they have a peak around [4.40, 4.44] Å and above ca. 5 Å only spurious interactions are found.

Visualizing RINs in PDB structures

RINs generated by RING can be visualized directly in the structure thanks to the RING-Viz script for Pymol. The

script is invoked from the command line, taking the RING network and corresponding PDB structure file as input. RING-Viz works out of the box in both Windows and Linux systems requiring only Pymol as dependency. The script also accepts other parameters to customize edge rendering or to filter interactions by type, distance, orientation and node identifier. Once the script finishes loading, nodes and edges appear as new objects corresponding to different interaction types. In this way, it is possible to customize the node and edge view transparently as for normal atom selections.

Server implementation

The RING web server is implemented in Node.js (<https://nodejs.org>) using the REST (Representational State Transfer) architecture and can be accessed through the web interface or programmatically exploiting the REST functionality. The web interface is built using the Angular.js (<https://angularjs.org>) framework and Bootstrap CSS style (<http://getbootstrap.com>). The network layout in the result page is calculated on the client-side exploiting a force-directed algorithm provided by the D3.js library (<http://d3js.org>). Unlike the previous version, RING-2.0 is now also available as stand-alone package. Hydrogen bonds and VDW interactions as well as secondary structure are calculated without the need for external tools. RING-2.0 is now also able to return both intra- and inter-chain interactions and contacts involving hetero atoms. Moreover, newly optimized distance thresholds are available as built-in defaults in strict and relaxed versions. The RING output network is provided both in text and GraphML (XML) formats, improving compatibility and integration with external tools.

SERVER DESCRIPTION

Input

The RING web interface is straightforward to use. The main page features an input box, which accepts either a PDB identifier or a structure file. By default, RING processes the first chain, alternatively the user can select the chain manually or decide to perform the calculation on all chains, obtaining both intra and inter-chain connections. RING compares residues that are not adjacent (i.e. separated) in the sequence. By default, the distance is set to 3, i.e. it compares position i and $i + 3$, but the user can vary the threshold to further filter local interactions. Two important options are related to the edge cardinality and distance thresholds. RING can return one, multiple or all possible interactions between a node pair. By default, it provides multiple interactions, but only one for each type. Distance thresholds are set automatically, but the user can choose between a stringent and relaxed definition to provide an easy way to generate both inclusive and very reliable networks. The two sets have been defined through large scale analysis, as described in the Methods section. Mutual information and residue conservation (entropy) are calculated on demand, since they require a time consuming PSI-BLAST profile. However, the server is designed to be always very responsive. The output network is generated immediately and

missing attributes are added transparently when the calculation finishes.

Output

RING provides the network as an interactive graph on the results page (see Figure 3). Node positions are updated dynamically thanks to a force-directed algorithm that tries to optimize the layout. The layout can also be adjusted manually by modifying the force parameters or dragging nodes with the mouse. Nodes can be coloured to highlight different aspects, like residue chemical propensity, vertex degree, secondary structure, mutual information and conservation (when available). Additional details are shown on a tooltip when the mouse hovers over a node or edge element. Multiple connections between nodes are shown as curved lines and ‘hetero’ molecules are grey circles with a black outline. RING output is also provided as different files, including the network in both GraphML (XML) and text format, the processed PDB structure with hydrogen atoms and the vector image (SVG) of the graph. The network can be loaded in Cytoscape (<http://www.cytoscape.org>) and visualized in the structure by running the RING-Viz program (see ‘Materials and Methods’ section), which is able to draw atomic level connections in Pymol (<https://www.pymol.org>). The XML network file can also be used by the RINalyzer/StructureViz (35) plugin to synchronize residue selection in Cytoscape with the 3D visualization in Chimera (23). Detailed instructions and examples are available in the tutorial and information about output formats in the help section of the website.

Usage example

Cyclin-dependent kinase (CDK) inhibitors play a central role in the regulation of eukaryotic cell cycles (16,36). A common event during malignant cancer progression is the deregulation of cell cycle phase transitions due to mutations frequently inactivating the kinase inhibitor activity. p27^{Kip}, a member of the Kip/Cip protein family, is known to act as a tumour suppressor protein (37). It is an intrinsically disordered protein (38) lacking a hydrophobic core characterized by consecutive secondary structure elements not interacting with each other. This extended conformation is used to form a relatively large contacting surface allowing multiple types of interactions with its binding partners (39). Here, we used RING 2.0 to analyse the crystal structure of p27^{Kip} bound to the cyclinA-Cdk2 complex (PDB code: 1JSU) and show how the web server can be used to easily retrieve information from a crystal structure. The generated RIN, covering three different chains, has a total of 510 nodes and 722 edges (see Figure 3). In the first panel (top-left corner), nodes are coloured by chain and both π - π and ionic interaction lines are thicker. Visual inspection revealed both intra- and inter-chain clusters of π - π interactions. One cluster (blue circle in the figure) represents the residues connecting the p27^{Kip}-Cdk2 chains. The inhibitor uses a β -strand to clamp around a β -sheet of Cdk2. This specific interaction induces a structural change which contributes to kinase inactivation. A similar cluster representation was also generated for the p27^{Kip} LGF binding motif. It lies in a shallow groove of cyclinA

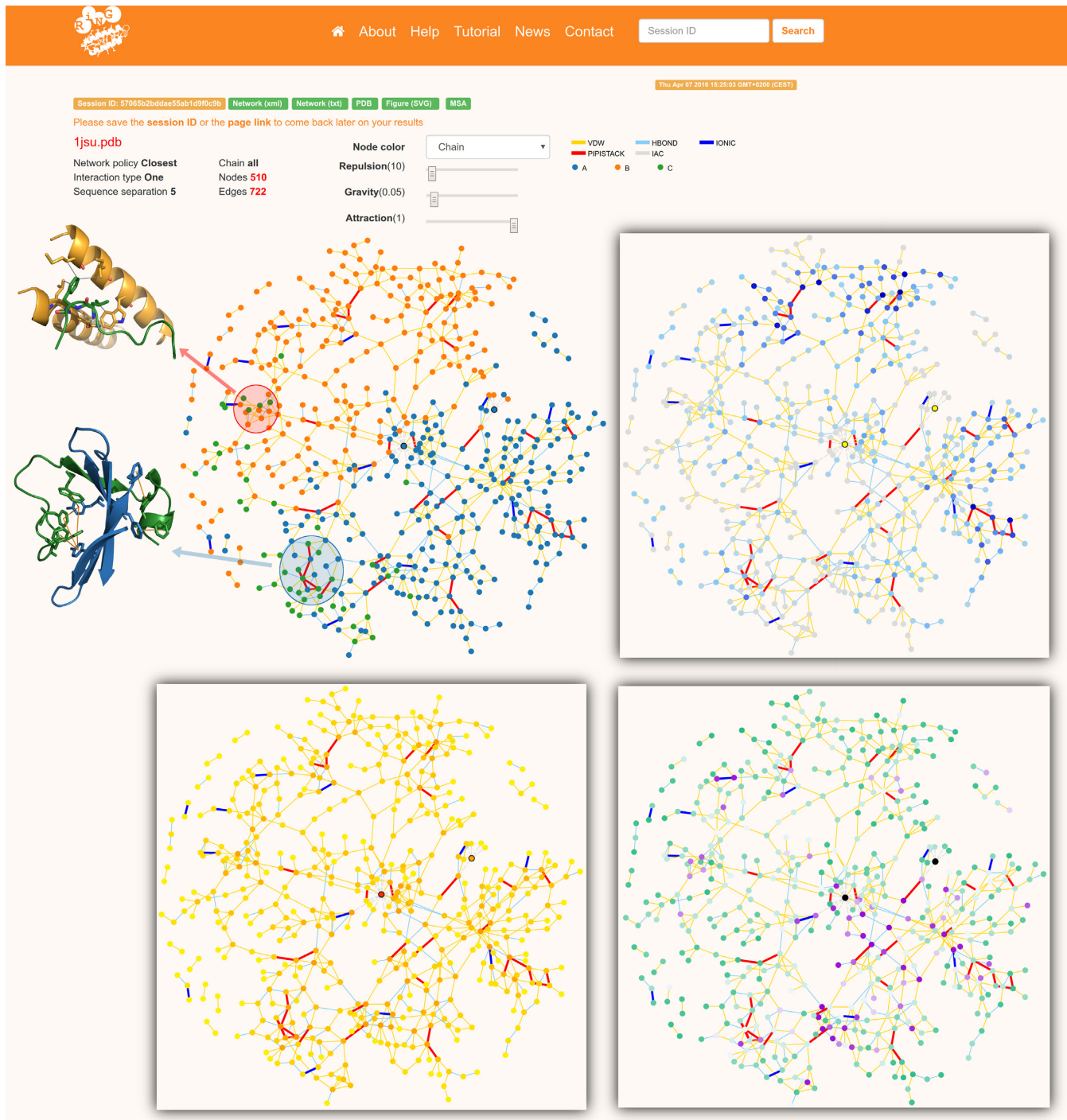


Figure 3. RING result page for the human p27^{Kip} kinase inhibitory domain bound to the phosphorylated cyclinA-Cdk2 complex (PDB code: 1JSU). The top-left graph shows the RIN with nodes and edges coloured according to the legend in the top right part. Highlighted interactions are shown in the lateral panels (structures in cartoon representation, interacting residues as sticks) and have been generated using the RING-Viz script (see 'Materials and Methods' section). The three inserts represent the same network graph with different colouring schemes. Clockwise from the top-right corner, the highlighted node attributes are: mutual information, conservation and node degree.

formed by the $\alpha 1$, $\alpha 3$ and $\alpha 4$ helices of the cyclin-box repeat (red circle in the figure). Ring 2.0 highlights important interactions at a glance, allowing a fast and useful recognition of functional residues. The three inserts in Figure 3 highlight other graph representations. Mutual information, conservation (entropy) and node degree are shown clockwise from the top-right corner. In general, node degree correlates with conservation and both provide indications on key residues. Mutual information is calculated only for intra-chain connections and highlights residues relevant for structural stability. It is interesting to note that chain C (green nodes) lacks residues with high mutual information values (pale-blue nodes in the top-right panel). This is not surprising as chain C lacks a hydrophobic core. Intra-chain contacts in elongated, disordered structures are less important and therefore less sensitive to correlated mutations.

CONCLUSIONS

We have presented RING 2.0, a new version of the RING software, for identification of both covalent and non-covalent bonds in protein structures. RING 2.0 is extremely fast and generates both intra- and inter-chain interactions while also considering ‘hetero atoms’, i.e. solvent, ligand and DNA/RNA atoms. A new empirical reparameterization of distance thresholds was performed on the entire PDB repository, ensuring a more reliable detection of real interactions. By default, RING output is generated with optimal parameters, but the web server provides an exhaustive interface to customize calculations. The network can be visualized directly on the web server or in Cytoscape. Alternatively, the RING-Viz script for Pymol allows visualizing atomic level interactions in the structure.

ACKNOWLEDGEMENT

The authors are grateful to members of the BioComputing UP group for insightful discussions.

FUNDING

University of Padova [CPDR123473 to S.T., in part]; Fondazione Italiana per la Ricerca sul Cancro [16621 to D.P.]; Associazione Italiana per la Ricerca sul Cancro [MFAG12740 to S.T.; IG17753 to S.T.]. Funding for open access charge: University of Padova.

Conflict of interest statement. None declared.

REFERENCES

- Cockroft, S.L. and Hunter, C.A. (2007) Chemical double-mutant cycles: dissecting non-covalent interactions. *Chem. Soc. Rev.*, **36**, 172–188.
- Dill, K.A., Ozkan, S.B., Shell, M.S. and Weikl, T.R. (2008) The protein folding problem. *Annu. Rev. Biophys.*, **37**, 289–316.
- Vishveshwara, S., Ghosh, A. and Hansia, P. (2009) Intra and inter-molecular communications through protein structure network. *Curr. Protein Pept. Sci.*, **10**, 146–160.
- Csermely, P. (2008) Creative elements: network-based predictions of active centres in proteins and cellular and social networks. *Trends Biochem. Sci.*, **33**, 569–576.
- Yan, W., Sun, M., Hu, G., Zhou, J., Zhang, W., Chen, J., Chen, B. and Shen, B. (2014) Amino acid contact energy networks impact protein structure and evolution. *J. Theor. Biol.*, **355**, 95–104.
- Albert, R., Jeong, H. and Barabasi, A.-L. (2000) Error and attack tolerance of complex networks. *Nature*, **406**, 378–382.
- Barabasi, A.-L. and Albert, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.
- 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. Publ. Protein Soc.*, **15**, 2120–2128.
- 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.
- Dokholyan, N.V., Li, L., Ding, F. and Shakhnovich, E.I. (2002) Topological determinants of protein folding. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 8637–8641.
- Soundararajan, V., Raman, R., Raguram, S., Sasisekharan, V. and Sasisekharan, R. (2010) Atomic interaction networks in the core of protein domains and their native folds. *PLoS One*, **5**, e9391.
- Swint-Kruse, L. (2004) Using networks to identify fine structural differences between functionally distinct protein states. *Biochemistry*, **43**, 10886–10895.
- Vendruscolo, M., Paci, E., Dobson, C.M. and Karplus, M. (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature*, **409**, 641–645.
- Hu, G., Yan, W., Zhou, J. and Shen, B. (2014) Residue interaction network analysis of Dronpa and a DNA clamp. *J. Theor. Biol.*, **348**, 55–64.
- Boehr, D.D., Schnell, J.R., McElheny, D., Bae, S.-H., Duggan, B.M., Benkovic, S.J., Dyson, H.J. and Wright, P.E. (2013) A distal mutation perturbs dynamic amino acid networks in dihydrofolate reductase. *Biochemistry*, **52**, 4605–4619.
- Scaini, M.C., Minervini, G., Elefanti, L., Ghiorzo, P., Pastorino, L., Tognazzo, S., Agata, S., Quaggio, M., Zullato, D., Bianchi-Scarrà, G. et al. (2014) CDKN2A unclassified variants in familial malignant melanoma: combining functional and computational approaches for their assessment. *Hum. Mutat.*, **35**, 828–840.
- Doncheva, N.T., Assenov, Y., Domingues, F.S. and Albrecht, M. (2012) Topological analysis and interactive visualization of biological networks and protein structures. *Nat. Protoc.*, **7**, 670–685.
- Tina, K.G., Bhadra, R. and Srinivasan, N. (2007) PIC: Protein Interactions Calculator. *Nucleic Acids Res.*, **35**, W473–W476.
- 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.
- 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.
- Hirsh, L., Piovesan, D., Giollo, M., Ferrari, C. and Tosatto, S.C.E. (2015) The Victor C++ library for protein representation and advanced manipulation. *Bioinformatics*, **31**, 1138–1140.
- Berman, H., Henrick, K., Nakamura, H. and Markley, J.L. (2007) The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res.*, **35**, D301–D303.
- Morris, J.H., Huang, C.C., Babbitt, P.C. and Ferrin, T.E. (2007) structureViz: linking Cytoscape and UCSF Chimera. *Bioinformatics*, **23**, 2345–2347.
- Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
- Torshin, I.Y., Weber, I.T. and Harrison, R.W. (2002) Geometric criteria of hydrogen bonds in proteins and identification of “bifurcated” hydrogen bonds. *Protein Eng.*, **15**, 359–363.
- Alvarez, S. (2013) A cartography of the van der Waals territories. *Dalton Trans.*, **42**, 8617–8636.
- Word, J.M., Lovell, S.C., Richardson, J.S. and Richardson, D.C. (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.*, **285**, 1735–1747.
- Pace, C.N., Scholtz, J.M. and Grimsley, G.R. (2014) Forces Stabilizing Proteins. *FEBS Lett.*, **588**, 2177–2184.
- Lovell, S.C., Word, J.M., Richardson, J.S. and Richardson, D.C. (1999) Asparagine and glutamine rotamers: B-factor cutoff and correction

- of amide flips yield distinct clustering. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 400–405.
30. Tosatto, S.C.E. (2005) The victor/FRST function for model quality estimation. *J. Comput. Biol.*, **12**, 1316–1327.
31. Tosatto, S.C.E. and Battistutta, R. (2007) TAP score: torsion angle propensity normalization applied to local protein structure evaluation. *BMC Bioinformatics*, **8**, 155.
32. Dunn, S.D., Wahl, L.M. and Gloor, G.B. (2008) Mutual information without the influence of phylogeny or entropy dramatically improves residue contact prediction. *Bioinformatics*, **24**, 333–340.
33. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
34. Emsley, J. (1980) Very strong hydrogen bonding. *Chem. Soc. Rev.*, **9**, 91–124.
35. 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.
36. Abbas, T. and Dutta, A. (2009) p21 in cancer: intricate networks and multiple activities. *Nat. Rev. Cancer*, **9**, 400–414.
37. Polyak, K., Lee, M.H., Erdjument-Bromage, H., Koff, A., Roberts, J.M., Tempst, P. and Massagué, J. (1994) Cloning of p27Kip1, a cyclin-dependent kinase inhibitor and a potential mediator of extracellular antimitogenic signals. *Cell*, **78**, 59–66.
38. Potenza, E., Domenico, T.D., Walsh, I. and Tosatto, S.C.E. (2014) MobiDB 2.0: an improved database of intrinsically disordered and mobile proteins. *Nucleic Acids Res.*, **43**, D315–D320.
39. Russo, A.A., Jeffrey, P.D., Patten, A.K., Massagué, J. and Pavletich, N.P. (1996) Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature*, **382**, 325–331.