Journal of Physical Chemistry & Biophysics

Paola, J Phys Chem Biophys 2014, 4:4
DOI: 10.4172/2161-0398.1000e120

Editorial Open Access

Comment-Protein Contact Networks: A New Framework for Structural Biology

Luisa Di Paola*

Campus Bio-Medico di Roma University, via Alvaro del Portillo 21, 00128 Rome, Italy

Abstract

Protein structure information is not sufficient to predict biological functionality. Among different approach of analysis, protein contact networks provide a previliged point of view through the recognition of proteins as complex networks: the invariant of the networks are related to different properties of protein molecules.

Folding dynamics and allosteric binding find new and detailed description within this framework, pointing at this methodology as one of the most promising in the perspective of big data analysis of whole proteomes.

Introduction

Biological activity relies on the protein structure-function tight relation, which intervenes in any regulatory processes based on protein components [1]. The detailed information of protein structure is, thus, the starting point to analyze its biological activity.

Even though the number of protein structures recorded into devoted repositories (such as PDB) increases exponentially on a yearly basis, the bridge connecting structure and function still lacks, and protein scientists now focus more and more on the interpretation of the protein structures in terms of their activity [2].

Sequence homology allows to describe the protein sequence-structure-function relation with a previsional purpose [3-5]. This methodology relies on the axioma that structure and function are stricly related, but it fails on non-crystalline protein structures with a biological activity: natively unfolded proteins work and fold upon binding recognition, so that only their complexes can be resolved, at least in part [6]. On the other hand, misfolded proteins share exactly the same sequence with their correctly folded counterparts, but their biological activity is drammatically different [7].

With the introduction of methods able to catch the non-crystalline protein (NMR [8], SAXS [9]), the picture has been enlarged, including also misfolded structures in the game. Also functional proteins appear more loosely connected to a given, rigid conformation, acting more effectively upon their structural flexibility [10].

In this framework, it is mandatory to find methods able to identify the structure features in a more general way. Protein structures act as complex systems, since intramolecular interactions act in a nonlinear, cooperative fashion: to correctly model them, we need a simplified description to reduce the information up to a given level, to let common properties of protein structures emerge [11].

In the last years, many researches have called attention on the role of structure topology in protein folding [12] and binding [13,14]: the idea that the geometrical topology is able to describe almost completely the biological functions of proteins is intriguing, since topology has deep rooths in organic chemistry, so that the identification of bonds between atoms is sufficient to derive all molecular properties, without any further information (such as the electronic clouds wave functions, for instance).

We recently reviewed a topology-based approach for the quantitative analysis of three-dimensional protein structures, based on the reduction of information of all atoms position into a contact map, reporting esclusively the intramolecular bonds between residues close to each other within a given range [15].

The choice of the bond length range describes a class of bonds: for instance, in our works [16], we decided to cut off bondw outside the range 4-8 Å, excluding both the stronger (covalent bonds) and looser (scarcely significant noncovalent bond). The idea underlying is that only some noncovalent bonds between residues partecipate into the protein dynamics at the basis of the folding and binding processes. Further, when we analyzed the range contact, we found that the chemico-physical affinity (hydrophobicity) directs long-range bonds whereas short-range contacts form on the basis of residues proximity along sequence [17].

This geometric networks provide useful insight into the folding dynamics [18]: contact order, describing the proneness of nodes to link with residues far in sequence (long-range contacts) drives the folding process, addressing to the contact range as the main determinant of the folding dynamics [19].

The methodology found application altogether in determining and characterizing allosteric hot spots [20,21]: the allosteric mechanism is at the molecular basis of biological regulation, but there is still no agreement on the link between allosterism and conformational change.

In this perspect, the network approach shifted the focus from the conformational change to the modification of topological role upon binding [22-24]: nodes in active sites of allosteric binding proteins change abruptly their topological role upon binding, with respect to active site residues in non-allosteric proteins, which save their role (Figure 1).

The allosterism is altogether strictly linked to the modularity of protein structures: protein domain identification and prediction

***Corresponding author:** Luisa Di Paola, Faculty of Engineering, Campus Bio-Medico di Roma University, via Alvaro del Portillo 21, 00128 Rome, Italy, Tel: 39-06-225419634; Fax: 39-06-225419609; E-mail: l.dipaola@unicampus.it

Received August 26, 2014; **Accepted** August 28, 2014; **Published** August 30, 2014

Citation: Di Paola L (2014) Comment-Protein Contact Networks: A New Framework for Structural Biology. J Phys Chem Biophys 4: e120. doi:10.4172/2161-0398.1000e120

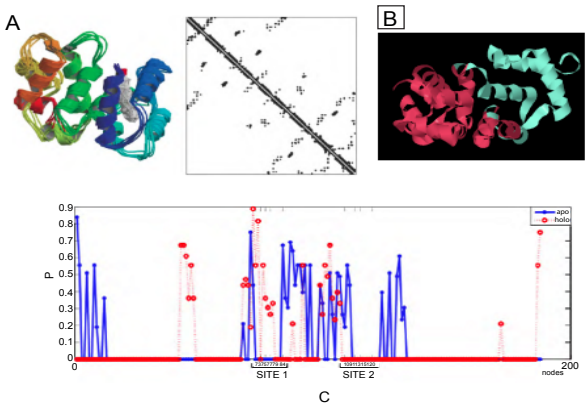

Figure 1: Protein topology and modularitys

represents a pillar of functional structural biology, assigning to each residue a definite role in the protein structure and activity. We developed an unbiased method to identify structural modules on the sole basis of intramolecular interactions network: it succeeds in singling out structural modules satisfactorily matching with protein domains and regions with a biological activity [25].

Figure 1- Protein topology and modularity: the topological role of residues in allosteric proteins is assigned according to their connectivity role in modules (case of recoverin [26]): A) the protein structure is translated into a network by singling out non-covalent intramolecular interactions; B) according to the network connectivity, the structure is decomposed into modules and residues are characterized by the partecipation coefficient *P*, quantifying their ability to communicate between different modules (for more details, [25]); C) P values for apo and olo forms significantly change, revealing a loss of propensity of active site residues to establish connections between modules.

These results drive to a further application, the prediction of mutation lethality on the basis of topological role of mutating residues: evidences on pathological and physiological mutated forms of hemoglobin highlighted the efficiency of the method to identify single-point mutations with a lethal fate [27].

This methodology is quite young but has already lead to significant results: this approach is endowed by a chemically natural paradigm, the topology of molecular structures, allowing a strong reduction of information against its ability to reveal global properties of the protein structures, on a systems biology perspective [28].

References

1. Hegyi H, Gerstein M (1999) The Relationship Between Protein Structure and Function: a Comprehensive Survey with Application to the Yeast Genome. J Mol Biol 288: 147-164.
2. Orengo CA, Pearl FM, Bray JE, Todd AE, Martin AC, et al. (1999) The CATH Database Provides Insights Into Protein Structure/Function Relationships. Nucleic Acids Res 27: 275-279.
3. Sangar V, Blankenberg DJ, Altman N, Lesk AM (2007) Quantitative sequence-function relationships in proteins based on gene ontology. BMC Bioinformatics 8: 294.
4. Li X, Chen Y, Lu S, Huang Z, Liu X, et al. (2013) Toward an understanding of the sequence and structural basis of allosteric proteins. J Mol Graph Model 40: 30-39.
5. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5: 725-738.
6. Uversky VN (2002) Natively unfolded proteins: a point where biology waits for physics. Protein Sci 11: 739-756.
7. Dobson CM (2003) Protein folding and misfolding. Nature 426: 884-890.
8. Shortle DR (1996) Structural analysis of non-native states of proteins by NMR methods. Curr Opin Struct Biol 6: 24-30.
9. Petoukhov MV, Svergun DI (2013) Applications of small-angle X-ray scattering to biomacromolecular solutions. Int J Biochem Cell Biol 45: 429-437.
10. Teilum K, Olsen JG, Kragelund BB (2011) Protein stability, flexibility and function. Biochim Biophys Acta 1814: 969-976.
11. Di Paola L, Giuliani A (2014) Protein Structures as Complex Systems: a Simplification Conundrum. Advances in Systems Biology 3: 7-9.
12. Shank EA, Cecconi C, Dill JW, Marqusee S, Bustamante C (2010) The folding cooperativity of a protein is controlled by its chain topology. Nature 465: 637-640.
13. Taylor WR, May ACW, Brown NP, Aszódi A (2001) Protein Structure: Geometry, Topology and Classification. Rep Prog Phys 64: 517–590.
14. Oka R, Engkvist O, Chen H (2013) Investigation of the influence of molecular topology on ligand binding. J Mol Graph Model 40: 22-29.
15. Di Paola L, De Ruvo M, Paci P, Santoni D, Giuliani A (2013) Protein contact networks: an emerging paradigm in chemistry. Chem Rev 113: 1598-1613.
16. Paola LD, Paci P, Santoni D, Ruvo MD, Giuliani A (2012) Proteins as sponges: a statistical journey along protein structure organization principles. J Chem Inf Model 52: 474-482.
17. Paci P, Di Paola L, Santoni D, De Ruvo M (2012) Structural and Functional Analysis of Hemoglobin and Serum Albumin Through Protein Long-Range Interaction Networks. Current Proteomics 9: 160-166.
18. Dokholyan NV, Li L, Ding F, Shakhnovich EI (2002) Topological determinants of protein folding. Proc Natl Acad Sci U S A 99: 8637-8641.
19. Plaxco KW, Simons KT, Ruczinski I, Baker D (2000) Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. Biochemistry 39: 11177-11183.
20. Reynolds KA, McLaughlin RN, Ranganathan R (2011) Hot spots for allosteric regulation on protein surfaces. Cell 147: 1564-1575.
21. Nussinov R, Tsai CJ, Csermely P (2011) Allo-network drugs: harnessing allostery in cellular networks. Trends Pharmacol Sci 32: 686-693.
22. De Ruvo M, Giuliani A, Paci P, Santoni D, Di Paola L (2012) Shedding light on protein-ligand binding by graph theory: the topological nature of allostery. Biophys Chem 165-166: 21-9.
23. Amitai G, Shemesh A, Sitbon E, Shklar M, Netanely D, et al. (2004) Network analysis of protein structures identifies functional residues. J Mol Biol 344: 1135-1146.
24. Blacklock K, Verkhivker GM (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.
25. Tasdighian S, Di Paola L, De Ruvo M, Paci P, Santoni D, et al. (2014) Modules identification in protein structures: the topological and geometrical solutions. J Chem Inf Model 54: 159-168.
26. De Ruvo M, Giuliani A, Paci P, Santoni D, Di Paola L (2012) Shedding light on protein-ligand binding by graph theory: the topological nature of allostery. Biophys Chem 165-166: 21-29.
27. Giuliani A, Di Paola L, Setola R (2009) Proteins as Networks: a Mesoscopic Approach Using Haemoglobin Molecule as Case Study. Current Proteomics 6: 235-245.
28. Giuliani A, Filippi S, Bertolaso M (2014) Why network approach can promote a new way of thinking in biology. Front Genet 5: 83.