Studying Allosteric Modulation of the Conformational Dynamics of Proteins Using Perturbation Response Scanning (PRS) Method: Literature Review and Research Design

Student

Seyed Damon Sadraie Najafi

Supervisor

Prof. Canan Atilgan
Farzaneh jalalypour

Sabancı Üniversitesi

PURE
PROGRAM FOR UNDERGRADUATE RESEARCH

Abstract:

The dynamic nature of the proteins gives rise to a variety of conformations which are relevant to their biological functions. Conformational changes can be defined as an exchange between main structure states. Studying the structural and biophysical properties of these states are challenging because of the complexity of the information they contain. The Perturbation Response Scanning (PRS) method, could be considered as a promising computational approach to study dynamics of proteins. The goal of this project is to investigate the potential residues/areas of proteins capable of modulating conformational changes (allosteric modulation). To this aim, PRS will be applied to identify key residues of different groups of proteins based on their motions and results will be comparatively evaluated with experimental data. In addition, a summery of other state of the art computational methods for allosteric site identification is provided.

Introduction

Protein motion can be described as a link between protein structure and its function. The role assumed by protein motion varies from transmitting allosteric signals in small scale and conformational changes in large scale for folding and metabolite transport, large assembly formation and cellular movement**.** Protein motions occur on time scales of picoseconds to seconds and on length scales of angstroms locally to nanometers globally. This breadth renders their study with any single experimental or simulation method impossible. Two frequently encountered classes of protein motions include shear and hinge categories. Furthermore, a relative minority of motions exist that cannot be classified as above and demand classes of their own- e.g. partial refolding in, or beta-sheet deformation in TATA-box binding protein -. Another factor to consider when categorizing proteins is the size of motion that is commonly ordered in motion of subunits, domains and fragments (implying bigger than residue movement and smaller than that of the domain).

Allostery is the concept that encompasses other non-orthosteric sites on enzymes, prone to enzyme-ligand interactions, regulating the enzyme activity. Whether positive or negative, the modulation is a product of conformational changes subsequent to the ligand binding that occurs with the propagation of a signal from the allosteric site to the active site. Allosteric site identification in proteins, thus, has become of great pharmaceutical interest. Many allosteric sites have been identified with a variety of biochemical and biophysical methods. A common experimental approach is high-throughput screening followed by X-ray crystallography, phage display and tethering. Other instruments are solid-state and relaxation dispersion nuclear magnetic resonance (NMR), H/D exchange mass spectrometry, patch-clamp fluorometry electro-physiology, disulfide trapping, and fragment-based screening. More recently a number of computational methods have been developed to further aid allosteric site identification. they have been classified as follows:

1. Sequence-based prediction methods; Statistical coupling analysis (SCA):

which utilizes a multiple sequence alignment (MSA) in a given protein family to identify protein sectors –(defined as networks of co-evolving residues). Protein sectors are postulated to the communicational link between allosteric and orthosteric sites in proteins. Hence, identification of surface sites in direct contact with protein sectors can suggest that they can be possible allosteric sites.

2. Structure-based prediction approaches; Allosite:

Utilizing allosteric site database (ASD), a vast collection of allosteric targets and their modulators, Allosite have been developed which uses algorithms such as pocket-based analysis and support vector machine (SVM) classifier to predict the location of allosteric sites in proteins.

3. Normal mode analysis-based prediction approaches; Normal mode analysis:

Normal mode analysis (NMA) provides us with global modes that have functional significance. More than capturing most of the main motions of quaternary structures it also predicts significant regions for mediating allosteric signals. PARS is the most recent advance in normal mode analysis.

PARS:

Protein allosteric and regulatory sites (PARS), is a webserver to predict allosteric sites. This method is based on the alteration of protein flexibility upon ligand binding described by NMA. Possible ligand-binding sites are obtained through the LIGSITEcs program. Then NMA is performed on the *unbound* and ligand-bound *protein*. The difference between the two states is compared and if a significant change is observed, it suggests the presence of a possible allosteric site.

4. Dynamics-based prediction approache:

Atomistic molecular dynamics (MD) is a highly accurate method that inserts proteins in a solvent and simulates local and global conformational alterations of them. MD has provided an atomic-level perspective for the whole process of drug binding. Thus, this method makes a thorough investigation of allosteric sites possible without any previous knowledge of the binding sites.

5. NMA-based prediction coupled with dynamics; SPACER:

SPACER, is a method for identifying allosteric and other biologically relevant sites based on measuring ‘binding leverage’. SPACER has been highly successful in allosteric site discovery in structurally homologous proteins. for the purpose of finding surface binding sites, Monte Carlo simulations are applied. Under deformation, the strain on the ligand-protein contact is acquired by low-frequency normal modes for each site. High strains suggest a high binding leverage, proposing that the site can affect the conformational state of the protein through a population shift mechanism.

Perturbation-Response Scanning (PRS)

MD simulations provide detailed information about the atomic motions. Perturbation Response Scanning (PRS) coupled with all-atom MD simulations is a promising method to explore the dynamics of a protein and allosteric modulation of its main structural states as well as to reveal the corresponding key residues involved in the transition.

PRS is based on linear response theory(LRT) and elastic network models. In this method, target state of protein described by a perturbation of the Hamiltonian of the initial state by applying sequential random forces on each residue. Initial and target states could be open-close or bound-unbound forms of protein. Resulting shift in the coordinates record as a response and approximated by:

$$\Delta R = [\gamma]^{-1} \Delta F$$

PRS needs several steps to set up. The procedure can be described with the flowchart given below.

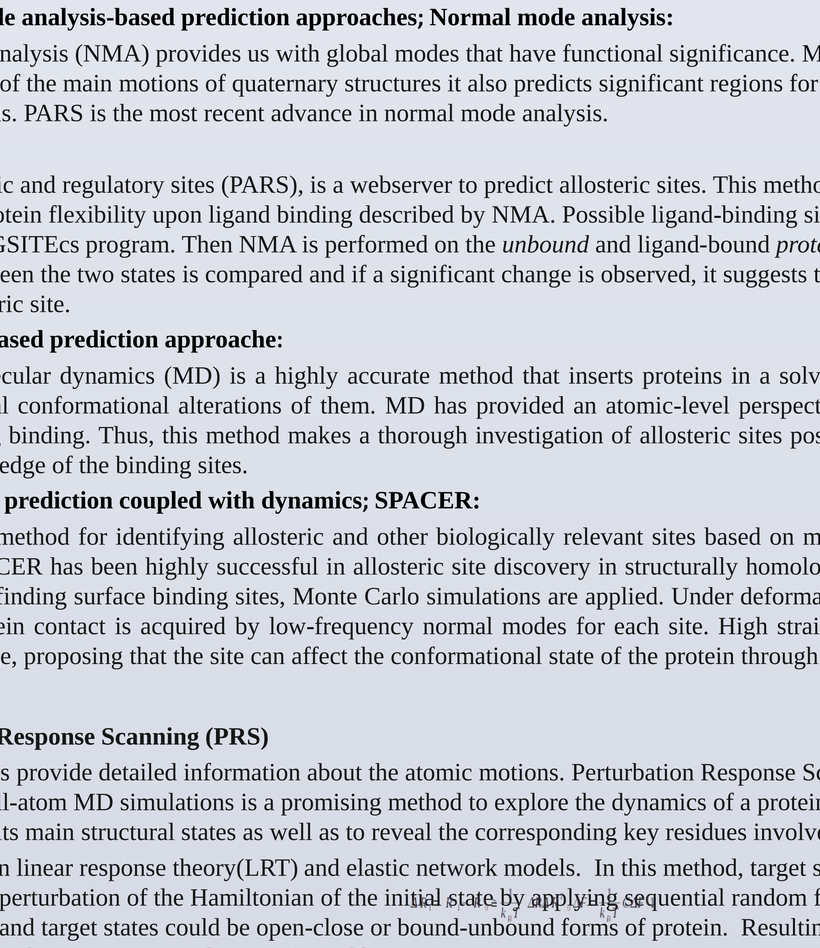

Figure1. Flowchart of the study

1. Database

1.1. Database of Macromolecular Movements (MolMovDB database)

Gerstein and Krebs have developed a database of proteins and nucleic acids movements classified based on their size and packing . Various conformational states of a protein are available in PDB format and a brief explanation have also been included . One can find apo and holo forms of each protein and the type of its movement including shear, hinge ,and rotations. In this project, we have made our groups based on the classifications described in the database (http://molmovdb.org/).

1.2. Protein Data Bank

Protein Data Bank (PDB) is a database collecting three-dimensional structures of biological macromolecules such as proteins and nucleic acids which obtained via experimental methods including NMR and X-Ray crystallography. Each state of a specific molecule has a four-character name or PDB code as well as all related information including sequence, structural information, atomic coordinates- experiment conditions ,and citations (https://www.rcsb.org/)

1.3. Holo / Apo protein set

In this project we used the Holo/Apo protein set selected by Atilgan. et al to compare the results of Network-based PRS with experimental data and MD-based PRS.

1.4. AlloSteric Database (ASD)

ASD is a database of allosteric proteins and their corresponding modulators in three categories of activators, inhibitors and regulators. For the proteins a description of allostery, biological process clinical significance and for the modulators, binding affinity, physicochemical properties and therapeutic areas have been provided. More than 286 allosteric proteins, 565 allosteric diseases and 22008 allosteric modulators have been deposited. In addition, Up-to-date allosteric pathways were manually curated in the database. Subtypes among proteins of the same family can be better defined using ASD and the modulator data can be used by chemists for structure modifications in allosteric drug design.

Proteins under study:

The fourteen selected proteins are known to have functional roles suggesting an allosteric modulation; allosteric key residues for five of them had been identified in the literature and are gathered in the table below. For the remaining proteins no paper was found pointing out to the allosteric sites. hence, PRS predictions for those, would be significant for future wet-lab allosteric identifications.

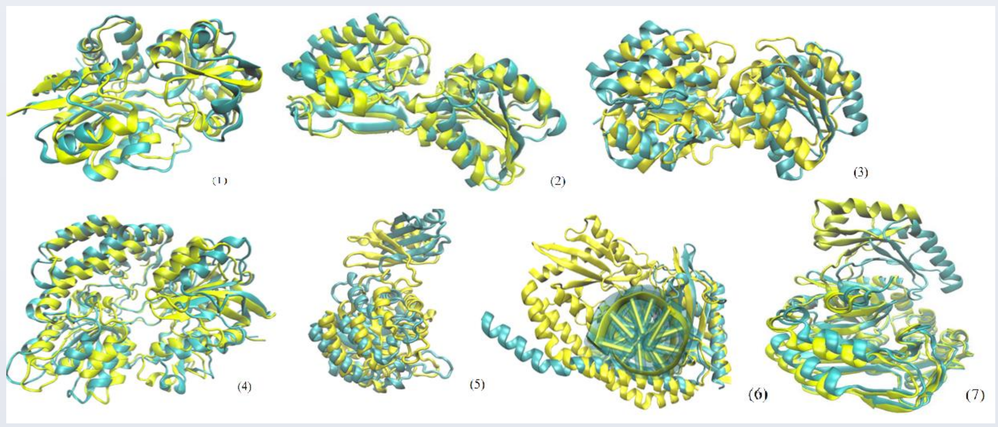

Figure2. All representations are modeled using VMD and the atomic coordinates are taken from PDB database. Cyan color represents the Apo state and yellow color represents the Holo state. Both states are aligned and the disposition between them intuitively suggests a discernible global motion in the protein, making it a qualified PRS candidate. The modeled proteins respectively are; (1) Beta-glucosyltransferase, (2) hexokinase, (4) Osmo protein, (3) D- Allose binding protein, (4) L-leucine-binding protein, (5) Alginate Binding Protein, (6) DAHP synthase and (7) Biotin Carboxylase.

Conclusions and further studies

With this literature review study we have achieved a research framework on which allosteric site identifications can be done using PRS method with the a special attention to the global motion characteristics of the proteins under study.

PDB files are carefully chosen and RMSD calculations have been done for each protein. They are together with the global motions and their known allosteric sites from discussed databases are presented in the table below.

Future studies will be running PRS on each of the proteins and comparing the yielded allosteric residues to those obtained from different wet lab experiments. Hopefully, we will see a connection between the global motion types and the conformity of PRS results to those obtained through wet lab experiments.

Protein	Residue Count	Apo / Holo	RMSD(Å) (Apo / Holo)	Motion	Allosteric Residues
Beta-glucosyltransferase	351	1JEJ / 1JG6	2.1	Hinge	
Hexokinase	298	2E2N / 2E2O	3.0	Hinge	
Alginate Binding Protein	490	1Y3Q / 1Y3N	4.7	Hinge	
ProX(osmo)	270	1SW5 / 1SW2	5.0	Hinge	
D-allose binding	288	1GUD /1RPJ	4.4	Hinge-bending	
L-leucine-binding protein	345	1USG / 1USI	7.1	Hinge	
Thymidylate Synthase	264	3TMS / 2TSC	0.8	Shear	Asn183, Arg175, Tyr258
Annexin V	316	1ANX / 1AVR	1.7	Hinge	Lys79, His267
myosin	730	1VOM / 2AKA	6.6	Not fully classified	Ser260,Lys265,Arg428 Leu431, Asp590, Pro591 Leu592, Ile617
GroEL	524	1AON/1OEL	12.38	Complex motions	Lys10
DAHP synthase	338	1RZM / 1VR6	10.10	Rotations	
Hemoglobin	141	4HHB / 2HCO	0.73	Allosteric	Arg142, Lys100, Thr138, Leu35,Lys128, Pro96, Thr135,Ala131, Val2, Phe37

Table 1: selected proteins and their relevant properties

Protein	Residue Count	Unbound (APO)	+DNA +Ca ion	+DNA +Mg ion	RMSD(Å) (Apo / Holo)	Motion	Allosteric Residues
SwaI	226	5TGQ	5TGX	5TH3	2.6	Hinge	

Table 2: SwaI and related properties

Protein	Residue Count	Unbound (APO)	+ATP (HOLO)	+ADP +Biotin	RMSD(Å) (Apo / Holo)	Motion	Allosteric Residues
Biotin C.	446	1DV1	1DV2	3G8C	3.6	Hinge	

Table 3: Biotin Carboxylase and its relevant properties

Refrences

Galperin, Michael Y. "The molecular biology database collection: 2005 update." Nucleic acids research 33, no. suppl_1 (2005): D5-D24.

Hilser, Vincent J., James O. Wrabl, and Hesam N. Motlagh. "Structural and energetic basis of allostery." Annual review of biophysics 41 (2012): 585-

Huang, Zhimin, Linkai Mou, Qiancheng Shen, Shaoyong Lu, Chuangang Li, Xinyi Liu, Guanqiao Wang et al. "ASD v2. 0: updated content and novel features focusing on allosteric regulation." Nucleic acids research 42, no. D1 (2013): D510-D516.

Lu, Shaoyong, Wenkang Huang, and Jian Zhang. "Recent computational advances in the identification of allosteric sites in proteins." Drug discovery today 19, no. 10 (2014): 1595-1600.

Huang, Wenkang, Shaoyong Lu, Zhimin Huang, Xinyi Liu, Linkai Mou, Yu Luo, Yanlong Zhao et al. "Allosite: a method for predicting allosteric sites." Bioinformatics 29, no. 18 (2013): 2357-2359.

Bahar, Ivet, Timothy R. Lezon, Ahmet Bakan, and Indira H. Shrivastava. "Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins." Chemical reviews 110, no. 3 (2009): 1463-1497.

Panjkovich, Alejandro, and Xavier Daura. "PARS: a web server for the prediction of protein allosteric and regulatory sites." Bioinformatics 30, no. 9 (2014): 1314-1315.

Mitternacht, Simon, and Igor N. Berezovsky. "Binding leverage as a molecular basis for allosteric regulation." PLoS computational biology 7, no. 9 (2011): e1002148.

Goncearenco, Alexander, Simon Mitternacht, Taipang Yong, Birgit Eisenhaber, Frank Eisenhaber, and Igor N. Berezovsky. "SPACER: server for predicting allosteric communication and effects of regulation." Nucleic acids research 41, no. W1 (2013): W266-W272.)

Atilgan, Ali Rana, Ayse Ozlem Aykut, and Canan Atilgan. "Subtle p H differences trigger single residue motions for moderating conformations of calmodulin." The Journal of chemical physics 135, no. 15 (2011): 10B613.

Abdizadeh, Haleh, Gokce Guven, Ali Rana Atilgan, and Canan Atilgan. "Perturbation response scanning specifies key regions in subtilisin serine protease for both function and stability." Journal of enzyme inhibition and medicinal chemistry 30, no. 6 (2015): 867-873.

Gerstein, Mark, and Werner Krebs. "A database of macromolecular motions." Nucleic acids research 26, no. 18 (1998): 4280-4290.

Deshpande, Nita, Kenneth J. Addess, Wolfgang F. Bluhm, Jeffrey C. Merino-Ott, Wayne Townsend-Merino, Qing Zhang, Charlie Knezevich et al. "The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema." Nucleic acids research 33, no. suppl_1 (2005): D233-D237.

Atilgan, C., Z. N. Gerek, S. B. Ozkan, and A. R. Atilgan. "Manipulation of conformational change in proteins by single-residue perturbations." Biophysical Journal 99, no. 3 (2010): 933-943.

Huang, Zhimin, Liang Zhu, Yan Cao, Geng Wu, Xinyi Liu, Yingyi Chen, Qi Wang et al. "ASD: a comprehensive database of allosteric proteins and modulators." Nucleic acids research39, no. suppl_1 (2010): D663-D669

Humphrey, William, Andrew Dalke, and Klaus Schulten. "VMD: visual molecular dynamics." Journal of molecular graphics 14, no. 1 (1996): 33-38.

Lovelace, Leslie L., Lydia M. Gibson, and Lukasz Lebioda. "Cooperative inhibition of human thymidylate synthase by mixtures of active site binding and allosteric inhibitors." Biochemistry 46, no. 10 (2007): 2823-2830.

Tönsing, Katja, Sergej Kakorin, Eberhard Neumann, S. Liemann, and R. Huber. "Annexin V and vesicle membrane electroporation." European biophysics journal 26, no. 4 (1997): 307-318.

Preller, Matthias, Krishna Chinthalapudi, René Martin, Hans-Joachim Knölker, and Dietmar J. Manstein. "Inhibition of Myosin ATPase Activity by Halogenated Pseudilins: A Structure–Activity Study." Journal of medicinal chemistry 54, no. 11 (2011): 3675-3685.

Roseman, Alan M., Shaoxia Chen, Helen White, Kerstin Braig, and Helen R. Saibil. "The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL." Cell 87, no. 2 (1996): 241-251.

Inobe, Tomonao, and Kunihiro Kuwajima. "Φ Value Analysis of an Allosteric Transition of GroEL based on a Single-pathway Model." Journal of molecular biology 339, no. 1 (2004): 199-205.

Abraham, Donald J., Fred C. Wireko, Ramnarayan S. Randad, Claude Poyart, Jean Kister, Brigitte Bohn, Jean Francois Liard, and Mary Pat Kunert. "Allosteric modifiers of hemoglobin: 2-[4-[[(3, 5-disubstituted anilino) carbonyl] methyl] phenoxy]-2-methylpropionic acid derivatives that lower the oxygen affinity of
-hemoglobin in red cell suspensions, in whole blood, and in vivo in rats." Biochemistry 31, no. 38 (1992): 9141-9149.

Boyiri, Telih, Martin K. Safo, Richmond E. Danso-Danquah, Jean Kister, Claude Poyart, and Donald J. Abraham. "Bisaldehyde allosteric effectors as molecular ratchets and probes." Biochemistry 34, no. 46 (1995): 15021-15036.

Yokoyama, Takeshi, Saburo Neya, Antonio Tsuneshige, Takashi Yonetani, Sam-Yong Park, and Jeremy RH Tame. "R-state haemoglobin with low oxygen affinity: crystal structures of deoxy human and carbonmonoxy horse haemoglobin bound to the effector molecule L35." Journal of molecular biology 356, no. 3 (2006): 790-801.