Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Structure-based prediction of protein allostery

Joe G Greener and Michael JE Sternberg

Allostery is the functional change at one site on a protein caused by a change at a distant site. In order for the benefits of allostery to be taken advantage of, both for basic understanding of proteins and to develop new classes of drugs, the structure-based prediction of allosteric binding sites, modulators and communication pathways is necessary. Here we review the recently emerging field of allosteric prediction, focusing mainly on computational methods. We also describe the search for cryptic binding pockets and attempts to design allostery into proteins. The development and adoption of such methods is essential or the long-preached potential of allostery will remain elusive.

Address
Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

Corresponding author: Greener, Joe G (joe.greener13@imperial.ac.uk)

Current Opinion in Structural Biology 2018, **50**:1–8

This review comes from a themed issue on **Sequences and topology**

Edited by **Joseph A Marsh and Sarah A Teichmann**

For a complete overview see the Issue and the Editorial

Available online 5th November 2017

https://doi.org/10.1016/j.sbi.2017.10.002

Introduction

Allostery in its broadest sense is the functional change at one site on a protein caused by a change at a distant site. The perturbation at the allosteric site can be non-covalent binding of a molecule (e.g. small molecule, ions, RNA, DNA), covalent binding (e.g. phosphorylation) or light absorption [1]. Changes in structure or dynamics lead to effects such as a reduction or increase in catalytic activity, changes in disordered regions or changes in oligomerisation state.

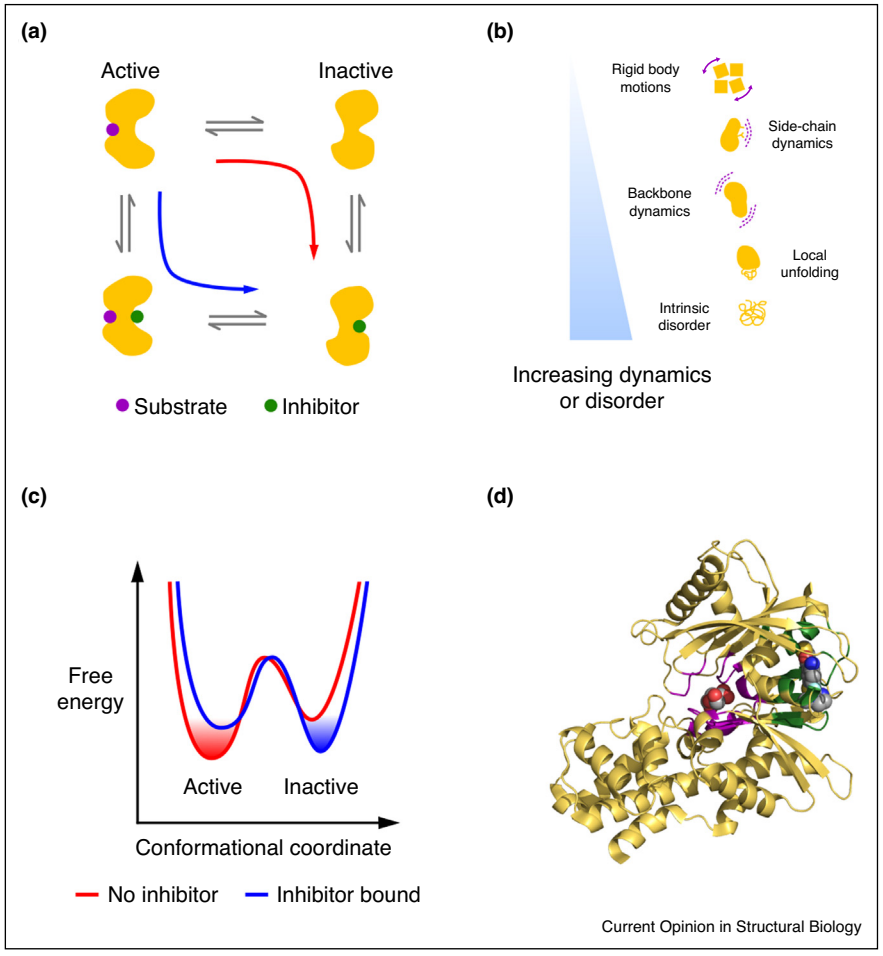
Since the first discovery of allosteric systems more than 50 years ago there have been various models put forward to describe the phenomenon. The dominant proposals for many years were the Monod-Wyman-Changeux (MWC) model, which posited that pre-existing states are subject to an equilibrium shift on modulator binding, and the Koshland–Némethy–Filmer (KNF) model, which advanced the idea that there was an induced fit of a binding site on interaction with a modulator [2••]. The structural view of allostery, which aimed to elucidate the allosteric mechanism by finding structural changes on effector binding, began to fill the gaps left by the phenomenological MWC and KNF descriptions. The discovery that entropic contributions to allostery can be significant predicted the phenomenon of allostery without conformational change, where the allosteric effect is communicated by a change in protein dynamics rather than protein structure [2••].

More recently these views on allostery have been revisited and reconciled in approaches that focus on the ensemble of conformational states that proteins exist in [2••,3]. Figure 1 outlines the current understanding of allostery. A perturbation at any site in the structure leads to a shift in the occupancy of states by the population, so allostery is a property of the conformational ensemble. The effect at the allosteric site is linked to the active site by small conformational changes that transmit the allosteric effect in a wave-like manner along pathways of amino acids in the protein [4]. These pathways may be conserved by evolution. It is also important to consider the effect of allostery on cellular networks and reaction pathways [1], with allosteric effects propagating via protein-protein interactions.

Allosteric drugs have hardly been explored and hold many potential benefits over orthosteric (non-allosteric) drugs: they are highly specific as they do not bind to active sites that are often conserved in protein families; they can activate as well as inhibit a protein; and they can have a ceiling to their effect [5]. Allosteric modulators have been elucidated for targets as diverse as G protein-coupled receptors (GPCRs), protein kinases, the GABA receptor, hepatitis C virus polymerase and RNA. Numerous other allosteric modulators are in various stages of human clinical trails. However, discovery of allosteric drugs presents challenges beyond those encountered in orthosteric drug discovery — see Box 1.

In order to understand and utilise allostery it is necessary to be able to predict allosteric sites, allosteric modulators and residues involved in propagating the allosteric signal. This review outlines advances from the last few years in the structure-based prediction of protein allostery, largely focusing on computational approaches. Previous reviews have covered similar topics [6–9]. The emerging fields of cryptic allosteric site discovery and allosteric site design are described. Challenges faced in the structure-based prediction of allostery and recommended steps for exploring allostery on a protein are also outlined — see Box 1 and Box 2 respectively.

Figure 1



The current conception of allostery. **(a)** A two-state model of allostery where a protein has an active and an inactive conformation. In the presence of the allosteric inhibitor the inactive state is favoured either by the inhibitor binding to the protein when it is in the inactive state (red arrow — conformational selection) or by the inhibitor binding to the active state and causing inactivation (blue arrow — induced fit). **(b)** The variety of motions that can lead to allostery. Larger motions or more disorder are shown further down the vertical axis. Figure based on Figure 2 from [2••]. **(c)** A simplified representation of the change in the energy landscape on binding of an allosteric inhibitor. The shaded regions show the main occupied conformation in each case. On inhibitor binding the relative energies of the active and inactive states are altered. For example, disruption of a hydrogen bond could destabilise the active state and stabilise the inactive state. **(d)** Glucokinase, a well-studied example of allostery [10], is shown as a yellow cartoon. The glucose substrate and the allosteric modulator are shown as spheres coloured by element. The active site and allosteric site are coloured purple and green respectively.

Computational methods

The last few years have seen the emergence of the first general methods that predict allostery based on protein structure. Table 1 summarises these methods, many of which are available as web servers.

Normal mode analysis methods

In normal mode analysis (NMA) the structural fluctuations of a protein around an equilibrium conformation are decomposed into harmonic orthogonal modes. The long-range nature of allosteric communication is often well-described by low-frequency modes that involve the motion of many atoms. The binding leverage approach [21•] predicts how ligand binding couples to the intrinsic motions of a protein. Sites with high binding leverage are predicted to be allosteric. Binding leverage was developed into the web server SPACER [20], and into the general predictor STRESS [22] by a different group. The PARS method [19•,18] calculates normal modes in the presence and absence of a simulated allosteric modulator. If the motions are significantly different the site is predicted as allosteric. The AlloPred method [11] calculates the normal modes of a protein, then holds the springs in the region of a potential allosteric site rigid and measures the effect of this perturbation at the active site. The DynOmics ENM server [15] finds hinge residues that control the two slowest normal modes of a protein, and hence are able to influence its dynamics. NMA is suitable

Box 1 Challenges faced in the structure-based prediction of protein allostery

1 As shown in Figure 1b, an allosteric effect can arise from a variety of different mechanisms. A general predictor would have to account for these in a unified manner. This is particularly challenging when disorder is involved, as approaches based on a defined structure are less applicable. Some approaches to studying disorder and allostery have been proposed [71,72].
2 The conformational changes that cause allostery are often large enough to occur on timescales of microseconds or milliseconds. This makes them too computationally expensive to study using MD without the use of accelerated or targeted MD. NMA is more computationally feasible but the assumption of a harmonic motion around an energy minimum does not correspond well to two distinct states with differing conformations.
3 The properties of active site pockets and small molecules that target the active site have been well-studied, for example Lipinski's rule of five [73]. Allosteric pockets and modulators may have generally different properties that we are not yet fully aware of, so we do not know exactly what to look for [74•,75].
4 The effect of an allosteric modulator is difficult to predict and can range from activation to inhibition, partial or complete. This is in comparison to orthosteric drug discovery, where drug action is presumed to be by competitive inhibition at the active site.
5 The effort of researchers and the protein structural data available is biased towards certain types of protein, such as those relevant in disease. For example, the allostery of GPCRs has been studied in detail [76]. There is a lack of protein structural data for important types of proteins such as membrane proteins and proteins with significant disorder, but these proteins have considerable potential to be allosteric [2••]. There may be different mechanisms or approaches to prediction that are relevant to less-studied protein families. The development of experimental methods such as cryo-electron microscopy should go some way to resolve this discrepancy [77].

Box 2 Recommended steps for predicting and rationalising allostery on a protein

1 Assemble available structural data from database searches and homology modelling where appropriate. If possible, obtain different conformations of the same protein, for example with or without an active site inhibitor or known allosteric modulator.
2 Submit structures to available web servers and methods as listed in Table 1. Explore the output to predict allosteric sites and see how far results agree.
3 Carry out further computational studies on the protein of interest, for example MD to investigate conformational changes, changes in dynamics and communication between sites.
4 Validation of the site experimentally. For example, screen molecules against the site using virtual and/or high-throughput screening and test hits using crystallography, NMR and activity studies as appropriate. Site-directed mutagenesis can be used to validate the binding site and suggested allosteric communication mechanisms. Use all available data to propose the mechanism of allostery.

for high-throughput, automated approaches as it can be computationally inexpensive. However whilst NMA-based methods might be expected to reveal perturbations to vibrations, the assumption of harmonic fluctuations around an energetically minimum structure means that other contributing motions to allostery such as local unfolding and rigid body movements [2••] are not taken into account.

Machine learning methods

A few methods have used machine learning to predict allostery. AlloSite [13] uses a support vector machine and features from Fpocket [24] to re-rank pockets in terms of their allosteric character. However the results are often found to be similar to the Fpocket ranking, showing the difficulty of distinguishing pockets that have specific allosteric character from those that are generally suitable for ligand binding. A Random Forest approach [26] uses descriptors for binding sites and associated ligands to assign protein cavities as allosteric, regular or orthosteric.

Molecular dynamics methods

Molecular dynamics (MD) remains the standard computational tool for structural analysis when structures are available. A study on the signalling protein NtrC combined MD simulations and NMR data to explore the free energy landscape and investigate at atomic resolution the transition from active to inactive state [27•]. Perturbation response scanning (PRS), in which the response of the structure to random perturbations at specific positions is examined, is a popular tool for allosteric prediction. For example, allosteric hotspot residues were predicted using PRS for the chaperone Hsp70 [28]. Weinkam *et al.* constructed energy landscapes and explored them with MD [14]. They were able to study the allosteric mechanisms involved in three proteins. The method is available as the AllosMod web server.

Evolutionary methods

Classic work has shown that allosteric communication can be mediated by networks of residues conserved by evolution [29]. One study developed previous work on protein sectors, groups of co-evolving residues physically contiguous in structure, to link sector-connected surface sites to allosteric sites [30••]. A recent approach found that surface and interior critical residues tend to be conserved [22]. The recent discovery that most directly co-evolving residues distant in 3D structure are close in related structures or assemblies [31] brings into question the concept of allosteric and active sites that directly co-evolve. As more structural and conservation information is acquired it will be important to discover to what degree allostery in proteins is a result of selection on specific pathways, and to what degree novel allostery can be discovered on proteins in the absence of previous evolutionary pressure.

Other methods

A recent study [32] constructs an all-atom graph and calculates for each bond the bond propensity, the strength of coupling to the active site through the graph. The method is used to reproduce observed results for three proteins in detail and is also able to predict allosteric sites

Table 1

Computational allosteric prediction methods currently available to run locally or as a web server, ordered alphabetically. In addition there are various pocket prediction methods that aim to predict binding pockets on proteins, but not specifically allosteric pockets [23–25]

Name	Reference(s)	Output(s)	Web server available	Source code available online
AlloPred	[11]	Predicted allosteric pockets	http://www.sbg.bio.ic.ac.uk/allopred/home	Yes, MIT license
AlloSigMA	[12]	Allosteric free energies	http://allosigma.bii.a-star.edu.sg/home	No
AlloSite	[13]	Predicted allosteric pockets	http://mdl.shsmu.edu.cn/AST	No
AllosMod	[14]	Modelled energy landscapes	http://modbase.compbio.ucsf.edu/allosmod	No
ENM method	[15]	Residues coupled to normal modes	http://enm.pitt.edu	Partly as ProDy, MIT license
ExProSE	[16]	Ensemble of protein structures, predicted allosteric pockets	No	Yes, MIT license
MCPath	[17]	Allosteric communication pathways	http://safir.prc.boun.edu.tr/clbet_server	No
PARS	[18,19•]	Predicted allosteric pockets	http://bioinf.uab.cat/pars	No
SPACER	[20,21•]	Predicted allosteric residues, exploration of allosteric communication	http://allostery.bii.a-star.edu.sg	No
STRESS	[22]	Predicted surface-critical and interior-critical residues	No	Yes

in a dataset of 20 allosteric proteins. ExProSE [16] takes two structures of the same protein and generates an ensemble of structures using distance constraints. By adding extra constraints at a possible allosteric site, a perturbed ensemble is generated. By comparing ensembles with and without the allosteric perturbation, allosteric sites can be predicted and the effect of perturbation on structure and dynamics can be explored. This work also includes a quantitative comparison of available allosteric site prediction methods.

Methods not specific to allostery

The identification of binding sites on the protein surface is a problem that has long pre-dated the search for pockets that are specifically allosteric. These methods are however useful in the structure-based prediction of allostery — the identification of a high-affinity binding site distant from a known active site could present an opportunity for allosteric regulation, for example. The FTMap family of web servers [33] predicts ligand-binding hotspots using small organic molecules as probes on the protein surface. By using mixed-solvent MD this principle has been extended to the prediction of allosteric sites in particular, with success on some test cases [34]. Common pocket prediction methods such as LIGSITE^{csc} [23] and Fpocket [24] are able to find pockets on a protein large enough to bind small molecules, and these often correspond to allosteric sites [16].

Allosteric pathway prediction

Allosteric signals can be propagated by multiple communication pathways [4]. Understanding these pathways is necessary in order to predict sites that are able to communicate with the active site [35]. A machine learning approach to predict residues involved in allosteric communication uses a variety of structural and network features and is able to predict these hotspots with reasonable accuracy [36]. A different approach, McPath, uses a Monte Carlo algorithm to define likely allosteric pathways by examining inter-residue interactions in a residue network [17]. A study that added an allosteric domain to a protein analysed residue contact maps to find loops mechanically coupled to the active site [37]. An investigation on the PDZ domain using MD found that allosteric changes are non-linear and occur in a non-local fashion, and are similar in many ways to protein folding [38].

Experimental methods

Experimental studies such as crystallography, NMR and site-directed mutagenesis remain the best tools for exploring allostery in a particular protein. A synthetic azetidine derivative that kills Mycobacterium tuberculosis (Mtb) through allosteric inhibition of tryptophan synthase (TrpAB), a previously untargeted enzyme, was found by a high-throughput screen [39••]. The inhibition is not easily overcome by changes in metabolic environment due to the modulator binding at the TrpAB α-β-subunit interface and affecting multiple steps in the overall reaction of the enzyme. A study on the proteasome [40] crystallised the complex in the presence and absence of an allosteric modulator. Having the active and inactive structures allowed the authors to propose a detailed mechanism of inactivation, which has implications for future allosteric proteasome inhibitors. A study on flavovirus protease [41•] used a virtual screen to select 29 potential allosteric compounds that were tested experimentally. One showed an ability to inhibit the conformational change and also inhibit flavovirus growth. Allosteric pathways in ERG proteins were proposed using fluctuation correlation data and validated by mutating residues in the pathways [42]. However, there are limits to the use of mutational studies to validate allosteric mechanisms. It has been found that mutational data can give evidence for a deliberately poorly conceived allosteric mechanism [43•]. In the future it is to be hoped that experimental screens specifically for allosteric sites

[44–48] become more widespread, opening the path to conventional large-scale screens for allosteric drugs.

Cryptic allosteric sites

The discovery of cryptic binding pockets — pockets that are only available in some conformations of the protein and may not have an associated experimental structure — has the potential to vastly increase the number of druggable sites on proteins [49•] and is directly relevant to allosteric prediction. A recent study [50] showed using NMR data that ligands of the LpxC enzyme access a cryptic site that is invisible to crystallography. One study used Markov state modelling and MD to predict multiple hidden allosteric sites on β -lactamase and tested these using thiol labelling experiments [51], later finding modulators for the sites [52]. The general approach CryptoSite uses machine learning to predict cryptic pockets on proteins using sequence and structural features [25]. However, two problems affect the use of cryptic allosteric pockets over allosteric sites where the pocket is present in most or all conformations. Firstly, the shape of the pocket is not known so rational drug design is difficult. Secondly, there is potentially an energetic cost associated with the protein adopting the conformation required for the cryptic pocket [53]. However, the discovery of ligands with inhibition constants in the low picomolar range in the above study [50] show that these sites are druggable. Further computational and experimental studies are required to explore this promising area.

Design of allosteric sites

The rational design of allosteric sites is a problem closely related to structure-based prediction of allostery. Introducing allosteric sites into existing proteins, or creating fusion proteins to add activity switches, has many potential applications including in biotechnology [54]. A recent study added a PDZ domain into the Cas9 protein at a site that did not disrupt enzyme action [55]. The protein showed modulator-dependent activity in cells, establishing a system for Cas9 activation. Another study created fusion proteins that use conformational entropy to respond to temperature or pH as a switch [56]. Taylor *et al.* engineered *Escherichia coli* LacI to respond to one of four new inducer molecules using computational design and mutagenesis [57]. Dagliyan *et al.* designed a protein with a unique topology, uniRapR, whose conformation is controlled by the binding of a small molecule [58••]. The switching and control ability of uniRapR was confirmed in silico, *in vitro*, and *in vivo*. uniRapR was used as an artificial regulatory domain to control activity of kinases as a proof of concept. The same group built on this and inserted the light-sensitive LOV2 domain into 3 proteins at non-conserved, surface-exposed loops identified computationally using residue contact analysis as being allosterically coupled to active sites [37].

Discussion

It is challenging to compare different methods for allosteric prediction. The different inputs and, more commonly, outputs make systematic comparisons difficult. One quantitative comparison indicated broadly similar performance between four available methods [16]. One of the Critical Assessment of Genome Interpretation challenges in 2015–16 focused on predicting the influence of mutations on the allosteric regulation of human liver pyruvate kinase [59]. However the uptake was limited to four groups and the predictive ability was marginally better than random. In the long run a dedicated community-wide initiative similar to the Critical Assessment of Structure Prediction [60] would be beneficial to the field of allosteric prediction.

One factor holding allosteric prediction back is the lack of a varied and robust set of benchmarks to test methods against. ASBench [61] is a curated set of allosteric proteins, and has been used for example to benchmark AlloPred [11]. It is a subset of the AlloSteric Database (ASD) [62]. ASD v3.0 contains over 1400 proteins and also includes allosteric mechanisms, allosteric networks of proteins and 'allosteromes' of the allostery involved in protein kinases and GPCRs. Improvements in such resources are necessary to prevent the developers of new methods having to assemble their own datasets [19•,21•,32] and to allow systematic comparisons between methods.

An issue that requires more study in the field of allosteric prediction is the exact relationship between an allosteric modulator and whether it acts as an activator or inhibitor. It has been shown that under different conditions the same allosteric modulator can have opposite effects [63]. Another viewpoint is the anchor/driver model of allostery, with the concept of a pushing or pulling driver determining which way the ligand acts [64]. An approach to study this would be a quantitative structure activity relationship-like study where a variety of modulators and conditions are explored on the same protein. This would give evidence as to whether small structural differences causing a pushing or pulling effect are enough to reliably switch activator/inhibitor action.

The mechanism of dynamic allostery, where the allosteric effect is transmitted through changes in dynamics and the average structure does not necessarily change, also requires further investigation. While experimental studies [65,66,39••] have found evidence for dynamic allostery, Nussinov and Tsai [67] warn that an apparent lack of conformational change can be an artefact of various factors such as crystal packing, crystallisation conditions, disorder to order transitions, incremental activation, synergy between allosteric sites and changes in oligomeric state. A recent MD study proposes that allostery in the well-studied PDZ domain is driven by changes in

electrostatic effects rather than solely changes in dynamics [68,69]. The role of water in allostery also needs to be further explored as evidence has been found that rearrangement of water molecules is a possible mechanism of allostery [70,32].

Conclusion

For many years papers have pointed to the immense potential of allostery for both understanding and drugging proteins. Yet they regularly contain the qualification that a unified framework of allostery remains 'elusive', and approved allosteric drugs remain rare more than 50 years after the first descriptions of allostery. In order to unlock the dormant potential of allostery, predictive methods need to be as established and robust as those in other areas of bioinformatics. When allosteric prediction is as effective as prediction of secondary structure or disordered regions, the power of allostery will be truly revealed. In an analogous way to allostery itself, it is hoped that the effects of exploring allostery will propagate to all areas of structural biology.

Conflict of interest statement

MJES is a director and shareholder in Equinox Pharma Ltd, which is involved in computer-aided drug discovery.

Acknowledgement

This work was supported by Biotechnology and Biological Sciences Research Council grant BB/J014575/1.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Nussinov R, Tsai CJ: **Allostery in disease and in drug discovery**. *Cell* 2013, **153**:293-305.

2. •• Motlagh HN, Wrabl JO, Li J, Hilser VJ: **The ensemble nature of allostery**. *Nature* 2014, **508**:331-339.
A thorough account of the ensemble view of allostery. Presents a framework to unify the description of allosteric mechanisms from different systems, including the recent discovery of allostery in intrinsically disordered proteins.

3. Cui Q, Karplus M: **Allostery and cooperativity revisited**. *Protein Sci* 2008, **17**:1295-1307.

4. del Sol A, Tsai CJ, Ma B, Nussinov R: **The origin of allosteric functional modulation: multiple pre-existing pathways**. *Structure* 2009, **17**:1042-1050.

5. Wenthur CJ, Gentry PR, Mathews TP, Lindsley CW: **Drugs for allosteric sites on receptors**. *Annu Rev Pharmacol* 2014, **54**:165-184.

6. Schueler-Furman O, Wodak SJ: **Computational approaches to investigating allostery**. *Curr Opin Struct Biol* 2016, **41**:159-171.

7. Wagner JR, Lee CT, Durrant JD, Malmstrom RD, Feher VA, Amaro RE: **Emerging computational methods for the rational discovery of allosteric drugs**. *Chem Rev* 2016, **116**:6370-6390.

8. Guarnera E, Berezovsky IN: **Allosteric sites: remote control in regulation of protein activity**. *Curr Opin Struct Biol* 2016, **37**:1-8.

9. Lu S, Huang W, Zhang J: **Recent computational advances in the identification of allosteric sites in proteins**. *Drug Discov Today* 2014, **19**:1595-1600.

10. Kamata K, Mitsuya M, Nishimura T, Eiki J, Nagata Y: **Structural basis for allosteric regulation of the monomeric allosteric enzyme human glucokinase**. *Structure* 2004, **12**:429-438.

11. Greener JG, Sternberg MJE: **AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis**. *BMC Bioinform* 2015, **16**:1-7.

12. Guarnera E, Tan ZW, Zheng Z, Berezovsky IN: **AlloSigMA: allosteric signaling and mutation analysis server**. *Bioinformatics* 2017. (in press).

13. Huang W, Lu S, Huang Z, Liu X, Mou L, Luo Y, Zhao Y, Liu Y, Chen Z, Hou T, Zhang J: **Allosite: a method for predicting allosteric sites**. *Bioinformatics* 2013, **29**:2357-2359.

14. Weinkam P, Pons J, Sali A: **Structure-based model of allostery predicts coupling between distant sites**. *Proc Natl Acad Sci U S A* 2012, **109**:4875-4880.

15. Wasmuth EV, Lima CD: **The Rrp6 C-terminal domain binds RNA and activates the nuclear RNA exosome**. *Nucleic Acids Res* 2017, **45**:846-860.

16. Greener JG, Filippis I, Sternberg MJE: **Predicting protein dynamics and allostery using multi-protein atomic distance constraints**. *Structure* 2017, **25**:546-558.

17. Kaya C, Armutlulu A, Ekesan S, Haliloglu T: **MCPath: Monte Carlo path generation approach to predict likely allosteric pathways and functional residues**. *Nucleic Acids Res* 2013, **41**: W249-W255.

18. Panjkovich A, Daura X: **PARS: a web server for the prediction of Protein Allosteric and Regulatory Sites**. *Bioinformatics* 2014, **30**:1314-1315.

19. • Panjkovich A, Daura X: **Exploiting protein flexibility to predict the location of allosteric sites**. *BMC Bioinform* 2012, **13**:1-12.
NMA was carried out on a dataset of allosteric proteins and significant changes in flexibility were found in 70% of cases. Allosteric sites could be selected from predicted pockets by combining changes in normal modes on effector binding and structural conservation information.

20. Goncearenco A, Mitternacht S, Yong T, Eisenhaber B, Eisenhaber F, Berezovsky IN: **SPACER: server for predicting allosteric communication and effects of regulation**. *Nucleic Acids Res* 2013, **41**:W266-W272.

21. • Mitternacht S, Berezovsky IN: **Binding leverage as a molecular basis for allosteric regulation**. *PLoS Comput Biol* 2011, **7**: e1002148.
A quantity called binding leverage, which measures the ability of a binding site to couple to the intrinsic motions of a protein, is introduced. Both catalytic and allosteric sites tend to have high binding leverage. Binding leverage can be calculated on a single crystal structure and used to predict allosteric sites.

22. Clarke D, Sethi A, Li S, Kumar S, Chang RW, Chen J, Gerstein M: **Identifying allosteric hotspots with dynamics: application to inter- and intra-species conservation**. *Structure* 2016, **24**:826-837.

23. Huang B, Schroeder M: **LIGSITE^{csc}: predicting ligand binding sites using the Connolly surface and degree of conservation**. *BMC Struct Biol* 2006, **6**.

24. Le Guilloux V, Schmidtke P, Tuffery P: **Fpocket: an open source platform for ligand pocket detection**. *BMC Bioinform* 2009, **10**.

25. Cimermancic P, Weinkam P, Rettenmaier TJ, Bichmann L, Keedy DA, Woldeyes RA, Schneidman-Duhovny D, Demerdash ON, Mitchell JC, Wells JA, Fraser JS, Sali A: **CryptoSite: expanding the druggable proteome by characterization and prediction of cryptic binding sites**. *J Mol Biol* 2016, **428**:709-719.

26. Chen ASY, Westwood NJ, Brear P, Rogers GW, Mavridis L, Mitchell JBO: **A random forest model for predicting allosteric and functional sites on proteins**. *Mol Inf* 2016, **35**:125-135.

27. • Pontiggia F, Pachov DV, Clarkson MW, Villali J, Hagan MF, Pande VS, Kern D: **Free energy landscape of activation in a signalling protein at atomic resolution**. *Nat Commun* 2015, **6**:7284.

MD simulations and NMR studies are used to explore the free energy landscape of NtrC and examine the active/inactive transition. Functional states are found to be defined in kinetic terms, with the inactive state adopting a variety of conformations.

28. Penkler D, Sensoy O, Atilgan C, Tastan Bishop O: **Perturbation-response scanning reveals key residues for allosteric control in Hsp70**. *J Chem Inf Model* 2017, **57**:1359-1374.

29. Suel GM, Lockless SW, Wall MA, Ranganathan R: **Evolutionarily conserved networks of residues mediate allosteric communication in proteins**. *Nat Struct Biol* 2003, **10**:59-69.

30. •• Reynolds KA, McLaughlin RN, Ranganathan R: **Hot spots for allosteric regulation on protein surfaces**. *Cell* 2011, **147**:1564-1575.

Allosteric sites are found to be surface sites of protein sectors — physically connected and co-evolving groups of residues. This is shown for DHFR and the PDZ domain using a domain insertion scan where the light-sensitive LOV2 domain is inserted at surface-exposed residues and used to probe for allostery.

31. Anishchenko I, Ovchinnikov S, Kamisetty H, Baker D: **Origins of coevolution between residues distant in protein 3D structures**. *Proc Natl Acad Sci U S A* 2017, **114**:9122-9127.

32. Amor BR, Schaub MT, Yaliraki SN, Barahona M: **Prediction of allosteric sites and mediating interactions through bond-to-bond propensities**. *Nat Commun* 2016, **7**:12477.

33. Kozakov D, Grove LE, Hall DR, Bohnuud T, Mottarella SE, Luo L, Xia B, Beglov D, Vajda S: **The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins**. *Nat Protoc* 2015, **10**:733-755.

34. Ghanakota P, Carlson HA: **Moving beyond active-site detection: MixMD applied to allosteric systems**. *J Phys Chem B* 2016, **120**:8685-8695.

35. Dokholyan NV: **Controlling allosteric networks in proteins**. *Chem Rev* 2016, **116**:6463-6487.

36. Demerdash ONA, Daily MD, Mitchell JC: **Structure-based predictive models for allosteric hot spots**. *PLoS Comput Biol* 2009, **5**:e1000531.

37. Dagliyan O, Tarnawski M, Chu PH, Shirvanyants D, Schlichting I, Dokholyan NV, Hahn KM: **Engineering extrinsic disorder to control protein activity in living cells**. *Science* 2016, **354**:1441-1444.

38. Buchenberg S, Sittel F, Stock G: **Time-resolved observation of protein allosteric communication**. *Proc Natl Acad Sci U S A* 2017, **114**:E6804-E6811.

39. •• Wellington S, Nag PP, Michalska K, Johnston SE, Jedrzejczak RP, Kaushik VK, Clatworthy AE, Siddiqi N, McCarren P, Bajrami B, Maltseva NI, Combs S, Fisher SL, Joachimiak A, Schreiber SL, Hung DT: **A small-molecule allosteric inhibitor of Mycobacterium tuberculosis tryptophan synthase**. *Nat Chem Biol* 2017, **13**:943-950.

A synthetic azetidine derivative is described that kills Mtb through allosteric inhibition of TrpAB. Binding occurs at the TrpAB α-β-subunit interface and affects multiple steps in the overall enzyme reaction. This demonstrates the effectiveness of targeting naturally highly dynamic proteins with allostery.

40. Haselbach D, Schrader J, Lambrecht F, Henneberg F, Chari A, Stark H: **Long-range allosteric regulation of the human 26S proteasome by 20S proteasome-targeting cancer drugs**. *Nat Commun* 2017, **8**:15578.

41. • Brecher M, Li Z, Liu B, Zhang J, Koetzner CA, Alifarag A, Jones SA, Lin Q, Kramer LD, Li H: **A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease**. *PLoS Pathog* 2017, **13**:e1006411.

A conformational switch assay that can monitor the conformational transition of a flavovirus protease component was developed and used to characterise allosteric inhibitors. The most potent hit inhibited protease activity with an IC_{50} value of 1.8 μM, and also inhibited the growth of multiple flavoviruses.

42. Ye W, Qian T, Liu H, Luo R, Chen HF: **Allosteric autoinhibition pathway in transcription factor ERG: dynamics network and mutant experimental evaluations**. *J Chem Inf Model* 2017, **57**:1153-1165.

43. • Tang Q, Alontaga AY, Holyoak T, Fenton AW: **Exploring the limits of the usefulness of mutagenesis in studies of allosteric mechanisms**. *Hum Mutat* 2017, **38**:1144-1154.

A deliberately problematic allosteric mechanism is proposed for a pyruvate kinase and mutagenesis studies were carried out to find evidence for the mechanism. Much of the data supports the conclusion, highlighting the limits of mutagenesis experiments in supporting proposed allosteric mechanisms.

44. Martin MP, Alam R, Betzi S, Ingles DJ, Zhu JY, Schonbrunn E: **A novel approach to the discovery of small-molecule ligands of CDK2**. *ChemBioChem* 2012, **13**:2128-2136.

45. Jayakar SS, Ang G, Chiara DC, Hamouda AK: **Photoaffinity labeling of pentameric ligand-gated ion channels: a proteomic approach to identify allosteric modulator binding sites**. *Methods Mol Biol* 2017, **1598**:157-197.

46. Pellerano M, Tcherniuk S, Perals C, Ngoc Van TN, Garcin E, Mahuteau-Betzer F, Teulade-Fichou MP, Morris MC: **Targeting conformational activation of CDK2 kinase**. *Biotechnol J* 2017, **12**:1600531.

47. Pisco JP, de Chiara C, Pacholarz KJ, Garza-Garcia A, Ogrodowicz RW, Walker PA, Barran PE, Smerdon SJ, de Carvalho LPS: **Uncoupling conformational states from activity in an allosteric enzyme**. *Nat Commun* 2017, **8**:1-10.

48. Raman S, Taylor N, Genuth N, Fields S, Church GM: **Engineering allostery**. *Trends Genet* 2014, **30**:521-528.

49. • Boehr DD, Nussinov R, Wright PE: **The role of dynamic conformational ensembles in biomolecular recognition**. *Nat Chem Biol* 2009, **5**:789-796.

Biomolecular recognition is examined via the perspective of conformational selection and induced fit. The view that a primary conformational selection event is followed by induced fit motions to optimise the interactions is described.

50. Lee CJ, Liang X, Wu Q, Najeeb J, Zhao J, Gopalaswamy R, Titecat M, Sebbane F, Lemaitre N, Toone EJ, Zhou P: **Drug design from the cryptic inhibitor envelope**. *Nat Commun* 2016, **7**:10638.

51. Bowman GR, Bolin ER, Hart KM, Maguire BC, Marqusee S: **Discovery of multiple hidden allosteric sites by combining Markov state models and experiments**. *Proc Natl Acad Sci U S A* 2015, **112**:2734-2739.

52. Hart KM, Moeder KE, Ho CMW, Zimmerman MI, Frederick TE, Bowman GR: **Designing small molecules to target cryptic pockets yields both positive and negative allosteric modulators**. *PLOS ONE* 2017, **12**:e0178678.

53. Oleinikovas V, Saladino G, Cossins BP, Gervasio FL: **Understanding cryptic pocket formation in protein targets by enhanced sampling simulations**. *J Am Chem Soc* 2016, **138**:14257-14263.

54. Makhlynets OV, Raymond EA, Korendovych IV: **Design of allosterically regulated protein catalysts**. *Biochemistry* 2015, **54**:1444-1456.

55. Oakes BL, Nadler DC, Flamholz A, Fellmann C, Staahl BT, Doudna JA, Savage DF: **Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch**. *Nat Biotechnol* 2016, **34**:646-651.

56. Choi JH, Laurent AH, Hilser VJ, Ostermeier M: **Design of protein switches based on an ensemble model of allostery**. *Nat Commun* 2015, **6**:1-9.

57. Taylor ND, Garruss AS, Moretti R, Chan S, Arbing MA, Cascio D, Rogers JK, Isaacs FJ, Kosuri S, Baker D, Fields S, Church GM, Raman S: **Engineering an allosteric transcription factor to respond to new ligands**. *Nat Methods* 2016, **13**:177-183.

58. •• Dagliyan O, Shirvanyants D, Karginov AV, Ding F, Fee L, Chandrasekaran SN, Freisinger CM, Smolen GA, Huttenlocher A, Hahn KM, Dokholyan NV: **Rational design of a ligand-controlled protein conformational switch**. *Proc Natl Acad Sci U S A* 2013, **110**:6800-6804.
A protein with a unique topology whose conformation is controlled by the binding of a small molecule is rationally designed. The ability of the protein to act as a switch and control pathways is demonstrated in Src kinase leading to two phenotypes *in vivo*.

59. Xu Q, Tang Q, Katsonis P, Lichtarge O, Jones D, Bovo S, Babbi G, Martelli PL, Casadio R, Lee GR, Seok C, Fenton AW, Dunbrack RL: **Benchmarking predictions of allostery in liver pyruvate kinase in CAGI4**. *Hum Mutat* 2017, **38**:1123-1131.

60. Moult J, Fidelis K, Kryshtafovych A, Schwede T, Tramontano A: **Critical assessment of methods of protein structure prediction: progress and new directions in round XI**. *Proteins* 2016, **84(Suppl 1)**:4-14.

61. Huang W, Wang G, Shen Q, Liu X, Lu S, Geng L, Huang Z, Zhang J: **ASBench: benchmarking sets for allosteric discovery**. *Bioinformatics* 2015, **31**:2598-2600.

62. Shen Q, Wang G, Li S, Liu X, Lu S, Chen Z, Song K, Yan J, Geng L, Huang Z, Huang W, Chen G, Zhang J: **ASD v3.0: unraveling allosteric regulation with structural mechanisms and biological networks**. *Nucleic Acids Res* 2016, **44**:D527-D535.

63. Motlagh HN, Hilser VJ: **Agonism/antagonism switching in allosteric ensembles**. *Proc Natl Acad Sci U S A* 2012, **109**:4134-4139.

64. Nussinov R, Tsai CJ: **Unraveling structural mechanisms of allosteric drug action**. *Trends Pharmacol Sci* 2014, **35**:256-264.

65. Popovych N, Sun S, Ebright RH, Kalodimos CG: **Dynamically driven protein allostery**. *Nat Struct Mol Biol* 2006, **13**:831-838.

66. Capdevila DA, Braymer JJ, Edmonds KA, Wu H, Giedroc DP: **Entropy redistribution controls allostery in a metalloregulatory protein**. *Proc Natl Acad Sci U S A* 2017, **114**:4424-4429.

67. Nussinov R, Tsai CJ: **Allostery without a conformational change? Revisiting the paradigm**. *Curr Opin Struc Biol* 2015, **30**:17-24.

68. Kumawat A, Chakrabarty S: **Hidden electrostatic basis of dynamic allostery in a PDZ domain**. *Proc Natl Acad Sci U S A* 2017, **114**:E5825-E5834.

69. Liu J, Nussinov R: **Energetic redistribution in allostery to execute protein function**. *Proc Natl Acad Sci U S A* 2017, **114**:7480-7482.

70. Buchli B, Waldauer SA, Walser R, Donten ML, Pfister R, Blochliger N, Steiner S, Caflisch A, Zerbe O, Hamm P: **Kinetic response of a photoperturbed allosteric protein**. *Proc Natl Acad Sci U S A* 2013, **110**:11725-11730.

71. Singh S, Bowman GR: **Quantifying Allosteric Communication via Both Concerted Structural Changes and Conformational Disorder with CARDS**. *J Chem Theory Comput* 2017, **13**:1509-1517.

72. Wang J, Custer G, Beckett D, Matysiak S: **Long distance modulation of disorder-to-order transitions in protein allostery**. *Biochemistry* 2017, **56**:4478-4488.

73. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ: **Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings**. *Adv Drug Deliv Rev* 2001, **46**:3-26.

74. • van Westen GJ, Gaulton A, Overington JP: **Chemical target and bioactive properties of allosteric modulation**. *PLoS Comput Biol* 2014, **10**:e1003559.
A large set of allosteric and non-allosteric ligands from the ChEMBL database of bioactive molecules is analysed to explore differences in physicochemical and structural features. In general allosteric modulators are found to be relatively smaller, more lipophilic and more rigid.

75. Wang Q, Zheng M, Huang Z, Liu X, Zhou H, Chen Y, Shi T, Zhang J: **Toward understanding the molecular basis for chemical allosteric modulator design**. *J Mol Graph Model* 2012, **38**:324-333.

76. Wootten D, Christopoulos A, Sexton PM: **Emerging paradigms in GPCR allostery: implications for drug discovery**. *Nat Rev Drug Discov* 2013, **12**:630-644.

77. Ozorowski G, Pallesen J, de Val N, Lyumkis D, Cottrell CA, Torres JL, Copps J, Stanfield RL, Cupo A, Pugach P, Moore JP, Wilson IA, Ward AB: **Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike**. *Nature* 2017, **547**:360-363.