ELSEVIER

Recent computational advances in the identification of allosteric sites in proteins

Shaoyong Lu, Wenkang Huang and Jian Zhang

Department of Pathophysiology, Key Laboratory of Cell Differentiation and Apoptosis of Chinese Ministry of Education, Shanghai JiaoTong University, School of Medicine (SJTU-SM), Shanghai 200025, China

Allosteric modulators have the potential to fine-tune protein functional activity. Therefore, the targeting of allosteric sites, as a strategy in drug design, is gaining increasing attention. Currently, it is not trivial to find and characterize new allosteric sites by experimental approaches. Alternatively, computational approaches are useful in helping researchers analyze and select potential allosteric sites for drug discovery. Here, we review state-of-the-art computational approaches directed at predicting putative allosteric sites in proteins, along with examples of successes in identifying allosteric sites utilizing these methods. We also discuss the challenges in developing reliable methods for predicting allosteric sites and tactics to resolve demanding tasks.

Introduction

Allosteric regulation, established almost 50 years ago, is a fundamental process by which distant sites within monomeric proteins or subunits of oligomeric proteins can communicate [1]. Intrinsically, binding of a ligand at an allosteric site topographically distinct from the orthosteric site regulates the functional activity of the protein through alteration of its conformation and/or dynamics. The widespread occurrence of allostery in the cell has motivated the development of two distinct but complementary methods to decipher how allostery works [2]: from a thermodynamic standpoint, allostery works via the population shift of an ensemble of protein conformations from the inactive to the active states, redistributing the conformational states toward the active conformation favored by the ligand [3–5]; from a structural viewpoint, allostery occurs by the propagation of strain energy created at the allosteric site by ligand binding to the functional site, which emphasizes the allosteric coupling (communication) between the allosteric and functional sites. This knowledge of the allosteric mechanism is extremely useful in facilitating allosteric discoveries and applications.

Allosteric modulators do not compete with orthosteric ligands and act primarily by modulating the affinity or efficacy of endogenous ligands [6]. Under certain circumstances, they have the potential to fine-tune protein activity even when the endogenous ligand occupies the orthosteric site on the same target. In contrast to classical approaches that design orthosteric ligands to interact with highly conserved orthosteric sites, targeting allosteric sites can endow allosteric modulators with greater selectivity, better physiochemical properties and fewer side effects [7]. In addition, in the case of class B G-protein-coupled receptors (GPCRs), allosteric modulators can hit their targets, whereas they are intractable to orthosteric manipulation as a result of the natural polypeptide ligands with varying lengths (ranging from 30 to 40 residues) outside classic oral drug-like space [8]. Thus, harnessing the allosteric modulation of protein function is now considered a novel approach in drug discovery [9–17].

Identification of allosteric sites in proteins of pharmaceutical interest is the first step in allosteric drug discovery. Until now, a vast number of allosteric sites deposited into the Allosteric Database (ASD) [18,19] have fortuitously been discovered through biochemical experiments, such as disulfide trapping [20], high-throughput screening [21] and fragment-based screening [22]. However, these experimental approaches face challenges regarding the fast-growing numbers of allosteric drug targets, and the biased chemical libraries that might not detect the potential allosteric sites. Alternatively, *in silico* methods that provide rapid platforms for identifying allosteric sites in proteins have been recognized to be valuable tools. A host of computational

Corresponding author:. Zhang, J. (jian.zhang@sjtu.edu.cn)

1359-6446/06/$ - see front matter http://dx.doi.org/10.1016/j.drudis.2014.07.012

Please cite this article in press as: Lu, S. et al. Recent computational advances in the identification of allosteric sites in proteins, Drug Discov Today (2014), http://dx.doi.org/10.1016/j.drudis.2014.07.012

approaches based on sequence, structure and dynamics have been developed for the prediction of allosteric sites [23–30].

In this review, we survey the recent advances in computational approaches developed for the prediction of allosteric sites. First, the state-of-the art computational approaches are briefly introduced, along with examples that utilize these tools to identify new allosteric sites in proteins. Finally, current challenges facing the computational prediction of allosteric sites and future perspectives are discussed.

Sequence-based prediction approaches

Statistical coupling analysis

Statistical coupling analysis (SCA) [31] is a sequence-based technique that uses a multiple sequence alignment (MSA) to identify networks of co-evolving residues in a protein family; such networks, also termed as protein sectors, provide a structural basis for allosteric communication between functional and allosteric sites [32–34]. Thus, allosteric sites can be predicted by identifying surface sites that are in direct contact with protein sectors. This assumption finds support in results from the computational prediction of allosteric communication of thermodynamically linked residues in the *Escherichia coli* dihydrofolate reductase [35] and PDZ domains [36], which are responsible for signal propagation within the protein structure, as confirmed through mutational analysis.

Very recently, Novinec *et al.* [23] used SCA to analyze the family of papain-like cysteine peptidases and constructed an MSA of 1239 catalytic domains from this family. Pairwise correlations were calculated between all pairs of residues in the alignment, and protein sector residues were identified through automated sector identification to uncover groups of co-evolving residues, which show a continuous network that surrounds and stretches throughout the molecule when mapped on the structure of cathepsin K, a member of the human peptidases. Alanine-scanning mutagenesis was performed to reveal that 14 out of 15 single-substitution mutants derived from the list of sector residues were involved in allosteric communication. Subsequently, potential ligand-binding sites on the surface of cathepsin K were predicted using AutoLigand, and eight potential allosteric sites were identified by filtering the predicted binding sites using the direct sector contact criterion. High-throughput docking of compound libraries to the eight potential allosteric sites of cathepsin K led to the identification of NSC13345, which was posited to occupy site 6 among the top 0.5% of solutions from the compound library and site 5 among the top 2%. Further X-ray crystallographic investigation of the cathepsin K–NSC13345 complex (PDB ID: 4LEG) determined that NSC13345 was bound to the flexible N terminus of cathepsin K, corresponding to the predicted site 6 distant from the active site (Fig. 1), which was thereby identified as a novel allosteric site in cathepsin K. Overall, these results were proved feasible by exploiting SCA to predict allosteric sites computationally through the identification of allosteric networks within a protein.

Structure-based prediction approaches

Allosite

We have constructed the largest dataset of known allosteric sites deposited into the ASD v2.0 [19]. The dataset contains 907 allosteric sites, of which 218 are unique allosteric sites occupied by 436 diverse chemical modulators. These unique allosteric sites are primarily distributed across several classes of therapeutic targets, including kinases (19.3%), ion channels (5.0%) and transcription factors (4.1%). The topological structures and physiochemical properties of these discovered allosteric sites are conductive for the development of a classification model to differentiate between allosteric and non-allosteric sites.

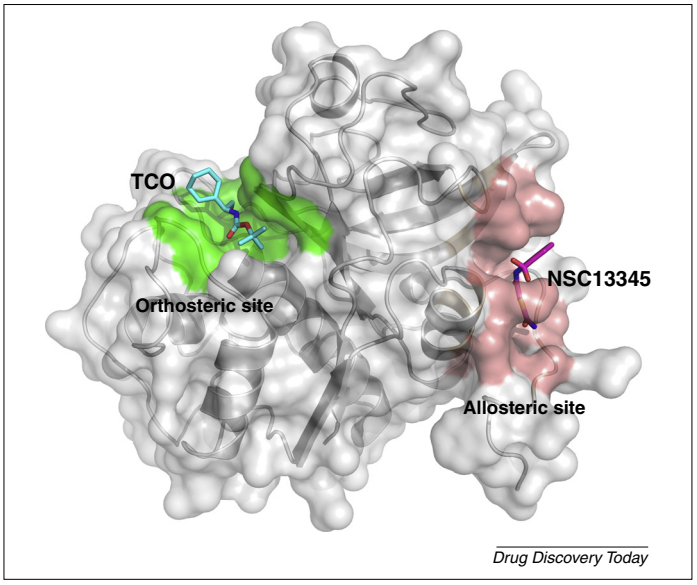

FIGURE 1

View of allosteric and orthosteric sites in cathepsin K (PDB ID: 4LEG). Regions of allosteric and orthosteric sites are highlighted in light pink and green on the surface, respectively, with the allosteric inhibitor NSC13345 (carbon atoms in magenta) and orthosteric inhibitor TCO (carbon atoms in cyan) represented by sticks. TCO is manually docked into the orthosteric site of 4LEG after superimposition with the PDB 1Q6K.

Recently, Huang *et al.* [24] selected 90 nonredundant allosteric sites from the 218 unique allosteric sites to develop a highly efficient, server-based model called Allosite (http://mdl.shsmu.edu.cn/AST) to predict allosteric sites. The training set consisting of 360 binding sites was classified into two groups: 72 allosteric sites and 288 non-allosteric sites, the latter of which were predicted by FPocket. The support vector machine classifier with 21 site descriptors achieved a sensitivity of >83% and a specificity of >96% in the fivefold cross-validation test; additionally, a prediction accuracy of 96.0% was obtained on an external test set composed of 18 allosteric sites and 231 non-allosteric sites. More importantly, based on the set of allosteric proteins where allosteric sites, as yet, undiscovered, Allosite has been proved to be capable of capturing putative allosteric sites in which several mutations adjacent to these sites and that affect the orthosteric functions of the proteins have been determined by biochemical experiments. Therefore, Allosite is a useful starting point for biologists and medicinal chemists in identifying the location of allosteric sites and, ultimately, allosteric drug design.

Normal-mode-analysis-based prediction approaches

Normal mode analysis (NMA) has the ability to provide global modes that bear functional significance. It not only captures most of the functional motions of quaternary structures but also unearths sites that could play a crucial part in mediating or propagating allosteric signals [37]. Hence, NMA has been used

Please cite this article in press as: Lu, S. et al. Recent computational advances in the identification of allosteric sites in proteins, Drug Discov Today (2014), http://dx.doi.org/10.1016/j.drudis.2014.07.012

for predicting allosteric sites. Recent advances using this method are highlighted below.

PARS

Protein allosteric and regulatory sites (PARS; http://bioinf.uab.cat/pars) [26] is a web server recently developed by Panjkovich and Daura to predict the location of allosteric sites based on the alteration of protein flexibility upon ligand binding described by NMA. Briefly, putative ligand-binding sites in a given protein are first detected using the LIGSITE^{cs} program. NMA is then performed for the *apo* (unbound) protein. As for each potential ligand-binding site, NMA is carried out for the protein–ligand complex, in which small molecules are simplistically represented by a set of dummy atoms. Eventually, the differences in NMA-derived B factors between the *apo* and ligand-bound states of the protein are compared. If a significant change is observed, the cavity where the ligand is positioned is proposed to be an allosteric site. Overall, this simple NMA-based model can predict 44% of known allosteric sites from the benchmark set consisting of 102 allosteric proteins in the first position and 18% in the second position according to the PARS ranking, whereas a total of 73% are observed in the top three positions.

Dynamics-based prediction approaches

Molecular dynamics simulations

Atomistic molecular dynamics (MD) simulations of proteins embedded in an explicit solvent can accurately reproduce local and large-scale conformational changes of proteins [38]. Such simulations can provide a continuous, atomic-level view of the complete process of drug binding [39], without any prior knowledge of the binding site. Thus, this method proves particularly useful in unraveling previously undiscovered allosteric sites.

In the case of the M2 muscarinic acetylcholine receptor (mAChR), a prototypical family A GPCR, Dror *et al.* [27] recently utilized large-scale unbiased MD simulations to locate allosteric sites. In the simulations, a single allosteric modulator was placed in the aqueous region away from the receptor that was embedded in lipids, water and ions, and no artificial forces were added to the simulated systems, thus allowing the allosteric modulator to bind spontaneously to the receptor. The results of microsecond simulations revealed that the allosteric modulators bound to the extracellular vestibule of the receptor, a region approximately 15 Å from the orthosteric site. On the basis of the binding modes of receptor–modulator complexes resulting from MD simulations, radioligand binding experiments on receptor mutants convincingly validated the computationally predicted binding modes between the M_2 mAChR and its allosteric modulators.

More recently, Shukla *et al.* [28] conducted massively distributed MD simulations (550 μs) to explore the conformational landscape of the activation pathway in the c-Src tyrosine kinase. Simulations revealed that two intermediate states (I_1 and I_2) existed along the pathway linking the inactive and active states of c-Src kinase. Specially, the intermediate I_2 of c-Src characterized by the broken Glu310–Arg409 ion pair was accommodated by the binding of 8-anilino-1-naphthalene sulphonate (ANS) to the allosteric site formed by the C-helix and β-sheets in the N-terminal lobe of c-Src. Collectively, these results indicate that longer MD simulations can predict the presence of a novel allosteric site in an intermediate state of c-Src that could be potentially used for allosteric drug design. Similarly, a transient allosteric site observed in one of the intermediate states has also been reported in the oncogenic Kirsten-Ras (K-Ras) protein via MD simulations [41].

Two-state Ga model

The coarse-grained (CG) two-state Gā model was recently developed by Qi *et al.* [29] to identify allosteric sites. Briefly, an ensemble of the two functional states of a given protein, such as the *apo* and ligand-bound structures, is first constructed. Subsequently, the energy landscape is tuned to bias one state, according to the population shift mechanism of allostery. Perturbations are added to a potential binding site, and the population distribution of the new ensemble is calculated via CG simulations. If the population distribution has shifted in response to the introduced perturbations, the binding-perturbed site is considered as a potential allosteric site.

Using this methodology, Qi *et al.* [29] predicted two potential allosteric sites, one adjacent to the active site and the nucleotide binding site (site I) and the other proximal to the regulatory domain (site II) in the *E. coli* phosphoglycerate dehydrogenase (PGDH). Molecular-docking-based virtual screening of the SPECS compound library targeted to the site II, together with the subsequent measurement of enzymatic inhibitory activity, identified one of the most potent inhibitors, AF407, with an IC_{50} of 21.6 μM. Further mutagenesis experiments were designed based on the potential binding modes between the PGDH and AF407. The results revealed that three single mutations within site II, namely, K8A, S316A and Y410A, retained their enzymatic activities but markedly eradicated their enzymatic inhibition by AF407. These data confirm that the site II is a novel allosteric site of PGDH that can be used for allosteric drug design. By taking advantage of this technique, more recently, they also discovered allosteric effectors bound to the predicted site I, suggesting that the site I is also a novel allosteric site for PGDH [40].

Combined dynamics- and NMA-based prediction approaches

SPACER

Mitternacht and Berezovsky recently developed a web server called SPACER (http://allostery.bii.a-star.edu.sg/) [25] that is based on the measurement of 'binding leverage' [42] to locate biologically relevant binding sites, including allosteric sites. It first uses Monte Carlo simulations to probe the surface of a protein and produce a list of possible binding sites. For each putative site, the strain on the ligand–protein contacts under the deformations, described by low-frequency normal modes, is evaluated. If the amount of strain measured by the contacts of a ligand with residues that are moving in opposite directions is high, this site is then considered to have a high binding leverage. Ligands binding to such sites can greatly influence the conformational state of the protein through a population shift mechanism. This method has been successful in unraveling latent allosteric sites in structurally homologous proteins.

For example, matrix metalloproteinases (MMPs), a family of zinc-dependent endopeptidases consisting more than 20 structurally homologous members, share high structural homology in their conserved catalytic machinery [43], which represents a key challenge for the design of antagonists targeting the conserved

Please cite this article in press as: Lu, S. et al. Recent computational advances in the identification of allosteric sites in proteins, Drug Discov Today (2014), http://dx.doi.org/10.1016/j.drudis.2014.07.012

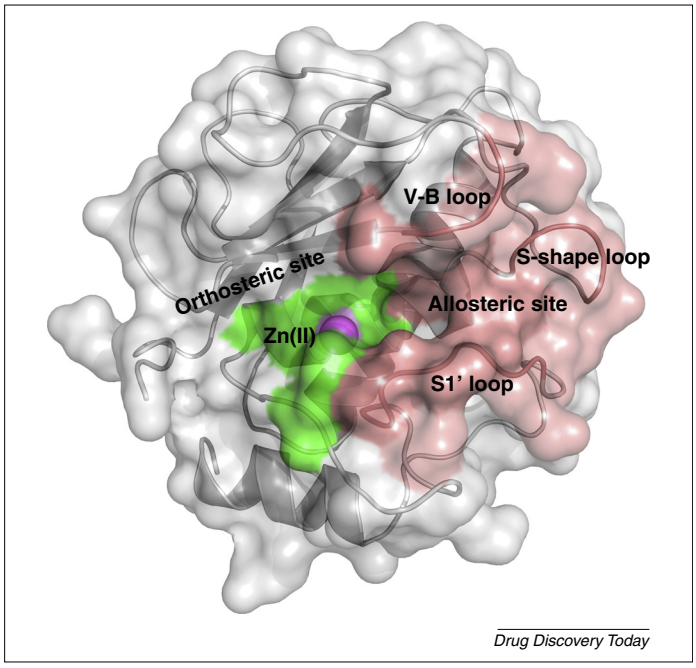

FIGURE 2

View of allosteric and orthosteric sites in matrix metalloproteinase (MMP)-12 (PDB ID: 4IJ0). Regions of allosteric and orthosteric sites are highlighted in light pink and green on the surface, respectively. Zn(II) in the active site is depicted by a violet sphere.

MMP catalytic zinc sites. Recently, Udi *et al.* [44] used the binding leverage computational method to predict the existence of hidden allosteric regulatory sites in MMP-12 outside the enzyme catalytic pocket. Zwitterionic copolymer YU158 was designed to detect this putative allosteric site. Structural analyses of MMP-12–YU158 complexes via the evaluation of chemical shift perturbations determined by solution NMR and the electron density by X-ray crystallography confirmed that YU158 bound to MMP-12 directly at the cavity formed between the S-shape, V-B and S1□ loops, which is a hidden allosteric regulatory site of MMP-12 (Fig. 2).

Concluding remarks: current challenges and future directions

Computational approaches for identifying allosteric sites provide useful empirical knowledge that complements existing experimental techniques and contributes to pharmaceutical R&D. Although impressive progress has been achieved, the predictive methods developed thus far still fall short of fully predicting experimental observations. In Box 1 we discuss some challenges faced in the development of sufficiently accurate approaches to predict allosteric sites in proteins computationally.

Despite the considerable challenges, the rapid development in experimental determination of allosteric sites, especially in the repertoire of GPCRs, fuels allosteric drug discovery. For instance, GPCRs in complex with their drug-like allosteric modulators, such as the M2 mAChR–LY2119620 [45], CCR5 chemokine receptor–maraviroc [46], metabotropic glutamate receptor 1–FITM [47] and corticotrophin-releasing factor type 1–CP376395 complexes [48], have recently been solved. In addition, the physicochemical and structural features of allosteric proteins [49,50] and allosteric modulators [51,52] have

BOX 1

Challenges faced in computational prediction of allosteric sites

- The shape and pharmacophoric properties of allosteric sites are not necessarily similar to those of highly conserved orthosteric sites (i.e. allosteric sites have evolved along different evolutionary pathways compared with orthosteric sites) [50]. In many cases, if not all, allosteric sites cannot be easily predicted from the *apo* structure, thus emphasizing that most of the allosteric sites are remarkably adaptable [56]. In addition, some of allosteric sites exist exclusively in intermediate states pertaining to activation or deactivation processes of proteins [28]. Together, these scenarios suggest an inability to predict allosteric sites from crystallographic or NMR structures sufficiently.
- Structure- or topology-based prediction models frequently suffer from a biased application. This is because the allosteric sites used to build the predictive model are selected mainly from protein kinases and enzymes, and no allosteric sites from G-protein-coupled receptors (GPCRs) are used. Therefore, these models can render the prediction of allosteric sites in GPCRs ineffective.
- Normal mode analysis (NMA)-based models are based on coarser approximations, in which the protein is represented by a subset of C_α atoms interconnected by a network of elastic springs. Although NMA provides an understanding of the global movements of biomolecular systems, it loses accuracy and specificity at the local scale [37]. Colombo *et al.* previously demonstrated the large energetic effects of hydration on hemoglobin conformation and thus protein allostery [57]. Recent studies also pinpointed that water molecules might modulate an allosteric coupling between allosteric and functional sites and act as allosteric modulators to induce a population shift of a conformational ensemble [58,59]. However, current NMA-based models are developed using the proteins in vacuum and do not account for water hydration effects, which could have a profound effect on the accurate prediction of allosteric sites.
- Rigorous molecular dynamics (MD) simulations can be used for the identification of allosteric sites. However, although potentially more accurate, long-MD simulations are computationally expensive, and the timescales that are suitable for effective sampling are frequently inaccessible for large molecular systems. In addition, the point charge model force-field used for amino acid residues in the MD simulations might not completely account for the impact of local environments on a given residue, which can possibly affect the prediction of the naturally occurring allosteric communication between allosteric and functional sites.
- In principle, most of computational approaches can predict multiple potential allosteric sites on the surface of a given protein. Yet, the current utility of suitable and efficient tools for determining the allosteric sites to which ligands bind preferentially is insufficient.
- Generally, allosteric regulation is a noncompetitive mechanism. Although computational approaches can predict the location of allosteric sites and screen potential allosteric modulators, experimentally distinguishing between allostery and competition among modulating ligands is currently challenging [60].

Please cite this article in press as: Lu, S. et al. Recent computational advances in the identification of allosteric sites in proteins, Drug Discov Today (2014), http://dx.doi.org/10.1016/j.drudis.2014.07.012

also recently been elucidated. Therefore, the aforementioned contributions can be used to improve the prediction accuracy of structure- or topology-based models. In the dynamics-based methods, the recent development of improved and/or polarized force fields [53] can be used in MD simulations to enhance the accuracy of conformational sampling, and appropriate analytical methodologies are needed to extract unique snapshots from simulated trajectories [54]. With respect to the CG NMA-based predictive methods, the recent development of the electrostatic force and the modified spring-force constant [55] should be integrated into the model to underscore water hydration effects at ionized interfacial surfaces. In light of soaring computing power, all-atom NMA of large biological systems with an explicit solvent will be developed to predict allosteric sites in the near future. Nevertheless, as the accuracy and predictive power of computational approaches in the identification of allosteric sites improve, they are expected to provide valuable avenues for allosteric drug discovery.

Conflict of interest

The authors declare that they do not have any conflicts of interest related to this manuscript.

Acknowledgements

This work was supported by National Natural Science Foundation of China (81322046 and 81302698), the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning and the Program for New Century Excellent Talents in University (NCET-12-0355).

References

[1] Changeux, J.-P. (2013) The concept of allosteric modulation: an overview. *Drug Discov. Today: Technol.* 10, e223–e228
[2] Tsai, C.-J. and Nussinov, R. (2014) A unified view of ''How allostery works''. *PLoS Comput. Biol.* 10, e1003394
[3] Hilser, V.J. (2010) An ensemble view of allostery. *Science* 327, 653–654
[4] Bai, F. *et al.* (2010) Conformational spread as a mechanism for cooperativity in the bacterial flagellar switch. *Science* 327, 685–689
[5] Popovych, N. *et al.* (2006) Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* 13, 831–838
[6] Nussinov, R. and Tsai, C.-J. (2012) The different ways through which specificity works in orthosteric and allosteric drugs. *Curr. Pharm. Des.* 18, 1311–1316
[7] Nussinov, R. and Tsai, C.-J. (2013) Allostery in disease and in drug discovery. *Cell* 153, 293–305
[8] Hoare, S.R.J. (2005) Mechanisms of peptide and nonpeptide ligand binding to class B G-protein-coupled receptors. *Drug Discov. Today* 10, 417–427
[9] Gao, Z.-G. and Jacobson, K.A. (2013) Allosteric modulation and functional selectivity of G protein-coupled receptors. *Drug Discov. Today: Technol.* 10, e237–e243
[10] Kimatrai-Salvador, M. *et al.* (2013) Allosteric modulation of A1-adenosine receptor: a review. *Drug Discov. Today: Technol.* 10, e285–e296
[11] Engers, D. and Lindsley, C.W. (2013) Allosteric modulation of class C GPCRs: a novel approach for the treatment of CNS disorders. *Drug Discov. Today: Technol.* 10, e269–e276
[12] Nemeth, E.F. (2013) Allosteric modulators of the extracellular calcium receptor. *Drug Discov. Today: Technol.* 10, e277–e284
[13] Desimmie, B.A. *et al.* (2013) Rational design of LEDGINs as first allosteric integrase inhibitors for the treatment of HIV infection. *Drug Discov. Today: Technol.* 10, e517–e522
[14] Schann, S. *et al.* (2013) Technology combination to address GPCR allosteric modulators drug-discovery pitfalls. *Drug Discov. Today: Technol.* 10, e261–e267
[15] Palmieri, L. and Rastelli, G. (2013) αC helix displacement as a general approach for allosteric modulator of protein kinases. *Drug Discov. Today* 18, 407–414
[16] Lu, S. *et al.* (2014) Harnessing allostery: a novel approach to drug discovery. *Med. Res. Rev.* http://dx.doi.org/10.1002/med.21317
[17] Wootten, D. *et al.* (2013) Emerging paradigms in GPCR allostery: implications for drug discovery. *Nat. Rev. Drug Discov.* 12, 630–644
[18] Huang, Z. *et al.* (2011) ASD: a comprehensive database of allosteric proteins and modulators. *Nucleic Acids Res.* 39, D663–D669
[19] Huang, Z. *et al.* (2014) ASD v2.0: updated content and novel features focusing on allosteric regulation. *Nucleic Acids Res.* 42, D510–D516
[20] Ostrem, J.M. *et al.* (2013) K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* 503, 548–551
[21] Feldman, T. *et al.* (2012) A class of allosteric caspase inhibitors identified by high-throughput screening. *Mol. Cell* 47, 585–595
[22] Jahnke, W. *et al.* (2010) Allosteric non-bisphosphonate FPPS inhibitors identified by fragment-based discovery. *Nat. Chem. Biol.* 6, 660–666
[23] Novinec, M. *et al.* (2014) A novel allosteric mechanism in the cysteine peptidase cathepsin K discovered by computational methods. *Nat. Commun.* 5, 3287
[24] Huang, W. *et al.* (2013) Allosite: a method for predicting allosteric sites. *Bioinformatics* 29, 2357–2359
[25] Goncearenco, A. *et al.* (2013) SPACER: server for predicting allosteric communication and effects of regulation. *Nucleic Acids Res.* 41, W266–W272
[26] Panjkovich, A. and Daura, X. (2014) PARS: a web server for the prediction of protein allosteric and regulatory sites. *Bioinformatics* 30, 1314–1315
[27] Dror, R.O. *et al.* (2013) Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature* 503, 295–299
[28] Shukla, D. *et al.* (2014) Activation pathway of Src kinase reveals intermediate states as targets for drug design. *Nat. Commun.* 5, 3397
[29] Qi, Y. *et al.* (2012) Identifying allosteric binding sites in proteins with a two-state Gā model for novel allosteric effector discovery. *J. Chem. Theory Comput.* 8, 2962–2971
[30] Collier, G. and Ortiz, V. (2013) Emerging computational approaches for the study of protein allostery. *Arch. Biochem. Biophys.* 538, 6–15
[31] Lockless, S.W. and Ranganathan, R. (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* 286, 295–299
[32] Gasper, P.M. *et al.* (2012) Allosteric networks in thrombin distinguish procoagulant vs. anticoagulant activities. *Proc. Natl. Acad. Sci. U. S. A.* 109, 21216–21222
[33] Sethi, A. *et al.* (2009) Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* 106, 6620–6625
[34] van Wart, A.T. *et al.* (2014) Weighted implementation of suboptimal paths (WISP): an optimized algorithm and tool for dynamical network analysis. *J. Chem. Theory Comput.* 10, 511–517
[35] Reynolds, K.A. *et al.* (2011) Hot spots for allosteric regulation on protein surfaces. *Cell* 147, 1564–1575
[36] McLaughlin, R.N., Jr *et al.* (2012) The spatial architecture of protein function and adaptation. *Nature* 491, 138–142
[37] Bahar, I. *et al.* (2010) Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chem. Rev.* 110, 1463–1497
[38] Sahún-Roncero, M. *et al.* (2013) The mechanism of allosteric coupling in choline kinase α1 revealed by the action of a rationally designed inhibitor. *Angew. Chem. Int. Ed.* 52, 4852–4856
[39] Shan, Y. *et al.* (2011) How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* 133, 9181–9183
[40] Hocker, H.J. *et al.* (2013) Andrographolide derivatives inhibit guanine nucleotide exchange and abrogate oncogenic Ras function. *Proc. Natl. Acad. Sci. U. S. A.* 110, 10201–10206
[41] Wang, Q. *et al.* (2014) Discovery of novel allosteric effectors based on predicted allosteric sites for *Escherichia coli* D-3-phosphoglycerate dehydrogenase. *PLoS ONE* 9, e94829
[42] Mitternacht, S. and Berezovsky, I.N. (2011) Binding leverage as a molecular basis for allosteric regulation. *PLoS Comput. Biol.* 7, e1002148
[43] Nagase, H. and Woessner, J.F. (1999) Matrix metalloproteinases. *J. Biol. Chem.* 274, 21491–21494
[44] Udi, Y. *et al.* (2013) Unraveling hidden regulatory sites in structurally homologous metalloproteases. *J. Mol. Biol.* 425, 2330–2346

Please cite this article in press as: Lu, S. et al. Recent computational advances in the identification of allosteric sites in proteins, Drug Discov Today (2014), http://dx.doi.org/10.1016/j.drudis.2014.07.012

[45] Kruse, A.C. *et al.* (2013) Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* 504, 101–106
[46] Tan, Q. *et al.* (2013) Structure of the CCR5 chemokine receptor-HIV entry inhibitor maravirov complex. *Science* 341, 1387–1390
[47] Wu, H. *et al.* (2014) Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* 344, 58–64
[48] Hollenstein, K. *et al.* (2013) Structure of class B GPCR corticotrophin-releasing factor receptor 1. *Nature* 499, 438–443
[49] Li, X. *et al.* (2013) Toward an understanding of the sequence and structural basis of allosteric proteins. *J. Mol. Graph. Model.* 40, 30–39
[50] Yang, J.S. *et al.* (2012) Rational engineering of enzyme allosteric regulation through sequence evolution analysis. *PLoS Comput. Biol.* 8, e1002612
[51] Wang, Q. *et al.* (2012) Toward understanding the molecular basis for chemical allosteric modulator design. *J. Mol. Graph. Model.* 38, 324–333
[52] van Westen, G.J.P. *et al.* (2014) Chemical, target, and bioactive properties of allosteric modulation. *PLoS Comput. Biol.* 10, e1003559
[53] Song, J. *et al.* (2013) The critical effect of polarization on the dynamical structure of guanine quadruplex DNA. *Phys. Chem. Chem. Phys.* 15, 3846–3854
[54] Feher, V.A. *et al.* (2014) Computational approaches to mapping allosteric pathways. *Curr. Opin. Struct. Biol.* 25, 98–103
[55] Martin, D.R. and Matyushov, D.V. (2012) Solvated dissipative electro-elastic network model of hydrated proteins. *J. Chem. Phys.* 137, e165101
[56] Hardy, J.A. and Wells, J.A. (2004) Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* 14, 706–715
[57] Colombo, M.F. *et al.* (1992) Protein solvation in allosteric regulation: a water effect on hemoglobin. *Science* 256, 655–659
[58] Prakash, P. *et al.* (2012) The role of conserved waters in conformational transitions of Q61H K-ras. *PLoS Comput. Biol.* 8, e1002394
[59] Reichow, S.L. *et al.* (2013) Allosteric mechanism of water-channel gating by Ca^{2+}-calmodulin. *Nat. Struct. Mol. Biol.* 20, 1085–1092
[60] Yang, P.S. *et al.* (2014) Allostery in Ca^{2+} channel modulation by calcium-binding proteins. *Nat. Chem. Biol.* 10, 231–238

Please cite this article in press as: Lu, S. et al. Recent computational advances in the identification of allosteric sites in proteins, Drug Discov Today (2014), http://dx.doi.org/10.1016/j.drudis.2014.07.012