EurJOC
European Journal of Organic Chemistry

Minireviews
doi.org/10.1002/ejoc.202100506

Chemistry Europe
European Chemical Societies Publishing

VIP *Very Important Paper*

Special Collection

Chemical and Biophysical Approaches to Allosteric Modulation

Monica Civera,[a] Elisabetta Moroni,[b] Luca Sorrentino,[a] Francesca Vasile,[a] and Sara Sattin*[a]

Dedicated to Professor Franco Cozzi *on the occasion of his 70th birthday.*

Allosteric regulation, i.e. the control exerted on an orthosteric site by an effector interacting at a distinct and distant site, represents a prime example of a precise tuning system of several key biological processes like gene transcription, cell adhesion and, most importantly, signal transduction. Since its discovery sixty years ago, the molecular mechanisms underlying this phenomenon have been extensively investigated. The aim of this minireview is to introduce the reader to the topic of protein allostery. In particular, we briefly overview the allosteric models postulated over the years and we describe the most relevant chemical and biophysical approaches reported so far for the identification of putative allosteric sites and/or for the investigation of allosteric signal propagation throughout the protein. An outlook of the main computational and experimental methods is followed by four case studies representative of different proteins classes: enzymes, hub proteins, cell receptors, and intrinsically disordered proteins.

1. Introduction

Allosteric regulation of proteins plays a pivotal role in a variety of biological processes, such as enzymatic activity, cell adhesion, signal transduction and transcriptional regulation. Defined as "the second secret of life" (second only to the genetic code and the central dogma of molecular biology), allostery is the effect observed at the protein functional (orthosteric) site upon binding of an effector at a distant (allosteric), non-overlapping, binding site. Monod[1] and Koshland[2] pioneered the studies on the cooperativity of oxygen binding to oligomeric haemoglobin in the 1960s, setting for decades the benchmark with their symmetric[1] (or MWC, Monod-Wyman-Changeux) and sequential[2] (or KNF, Koshland-Nemethy-Filmer) models, respectively. Both models provide a phenomenological qualitative description of two well-defined end-point structures at different energy levels (i.e. Tensed (T) and Relaxed (R)), tightly linking allostery to conformational change but lacking an explanation on how the information is propagated from one site to the other. An attempt in this direction is represented by the two conceptual models *Conformational Selection* (CS, or population shift) and *Induced Fit* (IF, or reaction front mechanism) (Figure 1).[3] In the case of conformational selection, the protein exists in equilibrium between an active (A) and an inactive (I) state even in the absence of the ligand (L, Figure 1). The ligand binds preferentially to the active state, lowering its energy and driving the population shift towards the A–L complex. On the contrary, in the induced fit model the binding of the ligand to the inactive state induces a perturbation through the protein ultimately leading to the conformational change. This model implies that the active state is not populated in the absence of the ligand. Both mechanisms can take place depending on the ligand concentration and the free energy differences among the two states.[3]

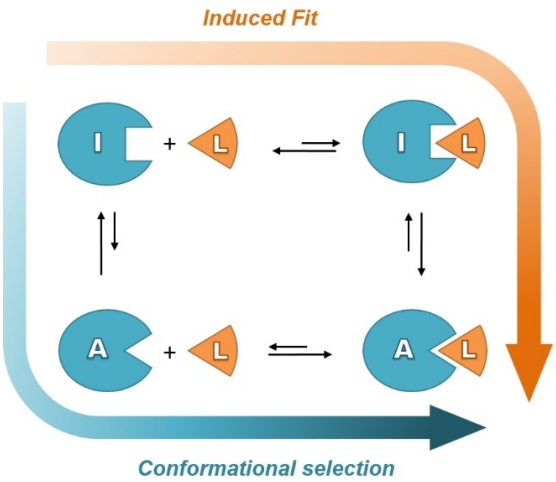


Figure 1. Diagram representing the two (possibly co-existing) conceptual models attempting to explain the transmission of allosteric information between two distinct protein states, i.e. Inactive (I) and Active (A), prompted by an allosteric effector (L) binding. Conformational selection (CS, bottom) involves a pre-existing equilibrium among states I and A, which is perturbed by preferential binding of L to A, effectively shifting the population towards the A–L complex. Induced Fit (IF, top) assumes that L binding to I gradually induces the conformational change leading to the A–L complex.

[a] *Dr. M. Civera, Dr. L. Sorrentino, Prof. Dr. F. Vasile, Prof. Dr. S. Sattin*
Department of Chemistry
Università degli Studi di Milano
via C. Golgi, 19, 20133, Milan, Italy
E-mail: sara.sattin@unimi.it
[b] *Dr. E. Moroni*
Istituto di Scienze e Tecnologie Chimiche Giulio Natta, SCITEC
Via Mario Bianco 9, 20131, Milan, Italy

Special Collection *Part of the "Franco Cozzi's 70th Birthday" Special Collection.*

The advent of structural biology facilitated studies investigating the structural coupling between the orthosteric and the allosteric site, and some of the computational approaches described below are based on this assumption.[4] It is worth mentioning that in the 1980s Cooper et al[5] introduced a model of allostery without conformational change, where variations in thermal fluctuations in response to the binding of an allosteric effector could cooperatively amount to interaction free energies of several $kJ \cdot mol^{-1}$. Almost twenty years later the interchange between ensembles of multiple states replaced the two-state models.[6] Binding of an allosteric ligand has an impact on the energy of the bound ensemble, shifting the distribution among all the possible conformations.[7] The Ensemble Allostery Model (EAM) does not rely on rigid tertiary structures and can therefore be applied to proteins with different degrees of order,[8] allowing us to describe allostery as a dynamic continuum, spanning from rigid body motions to intrinsically disordered proteins (IDPs) (Figure 2).

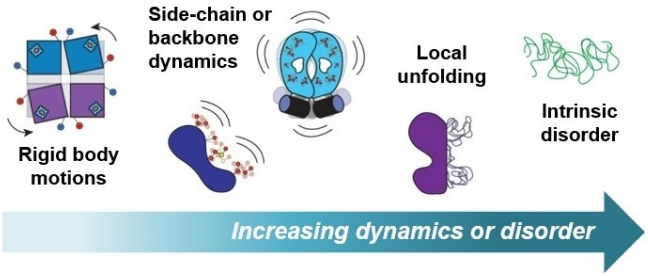


Figure 2. The dynamic continuum of allostery. Allosteric phenomena can be observed in systems with a wide range of dynamic character, spanning from rigid body motions (as in the case of the multimeric enzyme haemoglobin) to intrinsically disordered proteins. Adapted from Ref. [8].

Allosteric protein perturbations can arise from both covalent (e.g. post-translational modifications (PTMs) such as phosphorylation) and non-covalent events (i.e. effector binding, such as lipids, signalling or drug small molecules, other proteins, nucleic acids, or even light absorption). Targeting these processes could provide alternative therapeutic strategies with respect to the more widespread orthosteric drugs.[9] Indeed, allosteric sites are by definition less evolutionary conserved than the functional orthosteric ones,[10] allowing, in principle, for a better selectivity profile. In addition, it has been demonstrated that co-administration of an allosteric modulator restores or even enhances

Monica Civera got her PhD in Chemistry in 2006 at the University of Milan where she is currently an Assistant Professor (RTD-B) in Organic Chemistry. Her research interests are focused on computational chemistry, in particular on the field of drug design, the study of protein-protein interactions and the conformational analysis of biomolecules.

Elisabetta Moroni got her Ph.D. (2006) in Complex Systems for life Sciences at the University of Torino, focusing on MD simulations of DNA-protein complexes. Next, she joined the group of Prof. L. Belvisi at the University of Milano, working on the development of compounds of pharmaceutical interest. After a period as a researcher at the pharma company NMS and IRCCS MultiMedica, in 2019 she was appointed as researcher at SCITEC-CNR in the group of Prof. Colombo. Her research activity focuses on the development of computational methodologies to study biomolecules involved in different pathologies. The acquired knowledge is then applied for the development of potential new drugs.

After completing his PhD in Biomolecular Sciences (Università degli Studi di Milano) in 2015, Luca Sorrentino specialized in drug design projects. From 2016 to 2018, he worked in the Crystallography Unit of the Biophysics Institute (National Research Council) on a project aimed at the design and development of new potential anti-cancer drugs. He is currently a postdoc in the Sattin group on a project focused on the design of anti-microbial compounds targeting bacterial persistence. In 2020, he completed a master course in Pharmaceutical Management.

Francesca Vasile obtained her Ph.D. in Biophysical Sciences and technologies in 2003 from the University of Genoa. In 2017, she became assistant professor and is currently Associate Professor in Organic Chemistry at the University of Milan. Her main research interests include the conformational study of organic molecules, peptides and proteins in solution using Nuclear Magnetic Resonance. Her projects are also focused on the analysis of the intermolecular interactions by applying ligand- and target-based NMR techniques to study the interaction of small molecules with soluble proteins, protein expressed on membrane of living cells and RNA fragments.

Sara Sattin received her Ph.D. in Chemical Sciences in 2009 on antivirals targeting DC-SIGN. After two postdoctoral fellowships at the ICIQ (Spain) and the University of Oxford (UK) she returned to the University of Milan where she is Associate Professor of Organic Chemistry since 2018. Her research interests encompass the design and synthesis of small molecules targeting different biological targets (e.g. lectins and chaperone Hsp90) with a special focus on the bacterial stringent response RSH proteins and bacterial persistence.

the efficacy of an orthosteric drug, while also helping to overcome drug resistance.[11]

In 2004, Cinacalcet (Figure 3), a small molecule targeting a GPCR calcium-sensing receptor, was the first allosteric drug approved by the FDA. Since then, according to the Allosteric Database (ASD),[12] the FDA approved other 18 allosteric drugs, with a growing number of potentially upcoming drugs in preclinical (451), phase I (23), phase II (36) or phase III (9) studies. In ASD, the allosteric modulators (about 82,000 substances divided in activators, inhibitors, and regulators) are paired to their biological target(s) (about 2,000 structures retrieved from the Protein Data Bank) that cover eight main classes: kinases, GPCRs, transcription factors, ion channels, oxidoreductases, lyases, hydrolases, and transferases.

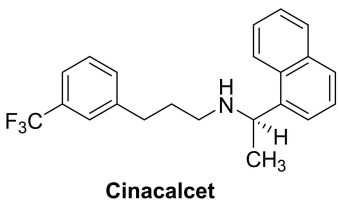
In this minireview, we aim to give an introductory overview of the different chemical and biophysical approaches adopted so far towards the discovery of allosteric sites and effectors, and the analysis of allosteric regulation mechanisms. We will cover the main computational and experimental methods, followed by four paradigmatic case studies covering different protein classes: kinases (enzymes), chaperone Hsp90 (cell hub protein), lectins (cell receptors) and IDPs (disordered proteins).

Computational methods

In the past years, thanks to the increasing number of structural data made available, computational methods have started contributing considerably into the field of allostery. By integrating experimental data into computational workflows, the discovery of allosteric modulators is now moving from an experimental serendipity approach towards a more robust structure-based design.[13] From a purely computational point of view, the complexity of allostery and our still limited understanding of how this phenomenon is elicited within a protein prevent the use of a single technique to model successfully the wide range of possible cases. In the rational design of allosteric modulators, the most challenging task is arguably the identification of allosteric sites, which are often cryptic, i.e. hidden in pockets that become available only when the system adopts a minor or rarely populated conformation. As a viable alternative to experimental methods, many predictive computational tools have been developed over the last two decades.[13a,14] For a comprehensive list of the tools developed so far, see Ref [13a].

Allosteric sites identification

As for orthosteric sites, putative allosteric pockets should be first mapped on the protein surface using a cavity detection tool. In principle, when no experimental information is available, any (binding) pocket located at a distance from the orthosteric site could be a putative allosteric site, making it necessary to discriminate actual allosteric sites from aspecific ones. For example, machine learning approaches[15] are trained on experimental validated datasets of known allosteric modulators. The most commonly used databases are ASD[16] and ASBench.[17] ASD has a web interface for the automatic prediction of allosteric sites that can also perform a virtual screening, while ASBench provides benchmarking sets, retrieved from experimentally determined allosteric sites, to test the prediction of a putative allosteric site and the computational workflow. The Allosite2.0 web server (http://mdl.shsmu.edu.cn/AST) predicts allosteric sites combining a pocket-based analysis and support vector machine (SVM) classifier algorithms trained on ASD. Two allosteric sites of kinase CDK2 predicted by this tool were validated experimentally by mutagenesis studies.[18] The major limitation of these approaches is their range of applicability, restricted to the classes of proteins on which they have been trained.



Cinacalcet

Figure 3. Chemical structure of Cinacalcet, the first allosteric drug approved by the FDA in 2004.

Normal Mode Analysis (NMA) methods[18,19] and Molecular Dynamics (MD) approaches[20] sample instead the conformational space of the protein and evaluate protein motions, allosteric communication or population distribution induced by ligand binding. NMA is a technique for determining the equilibrium modes accessible to a given macromolecule, assuming that the system is stabilized by harmonic potentials. As an example of NMA-based method, the web server AlloPred (http://www.sbg.bio.ic.ac.uk/allopred/home) identifies the allosteric pockets on a protein using a multi-step workflow: the cavity is first detected by the site finder tool Fpocket,[21] while the effect of the perturbation induced by binding of a modulator into the cavity is afterwards evaluated using NMA. Finally, the allosteric sites are ranked according to a combination of the NMA results and Fpocket descriptors using a SVM method. CorrSite[22] uses the Gaussian network model, a simplified version of NMA, to rank the potential sites identified by the CAVITY tool[22] according to a score that quantifies their correlation with the orthosteric site. The web server PARS[19a] analyzes protein flexibility and structural conservation to predict allosteric sites in a protein structure. In the adopted workflow, LIGSITE[23] predicts putative ligand-binding sites in a query protein structure using a simplified ligand representation. Afterwards, NMA is performed for the *apo* structure and for each potential ligand-protein complex (i.e. for each potential ligand binding site): if the difference of the NMA-derived B-factors between the *apo* and ligand-bound states of the protein is significant, the binding site is marked as potentially allosteric. The two-state Gō model[20a] is based instead on an ensemble generated by coarse-grained MD simulations and the populations shift approach: if the perturbation of a binding site leads to a population redistribution among the protein conformations, the perturbed site is considered as a potential allosteric site. Ligand Monte Carlo simulations on the protein surface, combined with a tool for the characterization of the induced protein conformational changes (like NMA), can be also used to detect potential allosteric sites.[24]

Allosteric communication: Monte Carlo methods and MD simulations

Monte Carlo (MC)[25] simulations represent a prime tool to unravel the interaction pathways that could be involved in an allosteric event and to explore the change in the conformational equilibrium of the protein induced by small molecule effectors. Monte Carlo methods consist in exploring the protein energy landscape of a protein by sampling various molecular geometries in a random fashion: starting from an arbitrary configuration of the system, new conformations are randomly generated. The new conformation is accepted based on the Metropolis criterion, which ensures that

conformations lower in energy are always accepted, while conformations higher in energy are accepted or rejected based on an algorithm involving a comparison of the transition probability with a random number. Thanks to this criterion, the search can also proceed "uphill" on the energy landscape and may cross barriers for reaching deeper minima. Thomas et al.[26] used MC simulation for studying three allosteric proteins, CheY, Integrin αL I-domain, and Ras, known to undergo large conformational changes upon binding to effector ligands. They used the Rosetta high-resolution structure prediction methodology, which performs a complete remodelling of the protein structure in specific regions, afterwards refining the whole structure with a Metropolis MC algorithm. The authors were able to predict alternative conformations of the proteins and to reproduce the transition steps between starting and target conformations.

Investigating allostery without large conformational changes

Allosteric signals can be transmitted through the changes in frequencies and amplitudes of macromolecular thermal fluctuations in response to ligand binding. Traditional MD simulations are well suited for studying allosteric signals that are transmitted via fast, local fluctuations (i.e., nanosecond-scale changes in the coordinated motions of adjacent residues), without large-scale conformational transitions. According to this view of allostery, the cumulative perturbation of residue-pair correlated motions creates a network that propagates the dynamics between the substrate and allosteric sites. In many cases, one or several residues play a prominent role in this dynamic network structure, acting as allosteric "hot spots".

In principle, a well-converged MD trajectory contains all the information necessary for describing these processes. However, it can be difficult to filter these signals from the high-dimensional and noisy dynamics that are inherent to MD. A multiplicity of approaches has been proposed for deciphering allosteric events from the analysis of MD trajectories, such as the representation of protein motion through dynamic networks and methods for calculating correlating motions and contacts of atoms/amino acids. Mouawad et al[27] studied the transition between the deoxy T-state and the oxy R-state of human haemoglobin using a combination of MD with the path exploration by distance constraint method. A constraint term pushing the protein to reach a reference conformation is added to the energy potential, while decreasing the mass-weighted root-mean-square deviation (RMSD) between the two conformations through a series of displacements, each consisting of several picoseconds of MD. Thanks to such approach, the authors were able to observe even transient effects not easily detectable using experimental techniques, as well as a sequence of conformational shifts previously predicted to occur in a different order. In order to detect long range couplings, Sharp and Skinner[28] proposed to excite selected atoms/residues with oscillating forces, which propagate to other regions of the protein, and are subsequently probed using Fourier transforms of Cartesian coordinate trajectories. The inter-residue coupling strengths quantify the degree of interaction.

Investigating large conformational changes

Allosteric events in proteins can involve large conformational changes spanning over the ms-μs timescale that are often far beyond what classical MD simulations can manage in terms of time and length scales. Nowadays scalable codes, GPUs, and parallelization algorithms allow simulating systems with millions of atoms for several microseconds. However, although long simulation time is achievable with special supercomputers, it is not enough to perform efficient sampling for most proteins. Indeed, the transition paths of allosteric events involving large conformational changes are stochastic, meaning that random fluctuations push the protein over energy barriers, following many unpredictable routes. In this case, unbiased MD methods are discarded in favor of enhanced sampling techniques, such as biased MD, targeted MD, weighted ensemble methods, accelerated-MD, steered MD, replica-exchange, umbrella sampling and Metadynamics, as exemplified below.

Steered MD simulations allowed observing a large-scale conformational change involving a functional interface of ImGP (imidazole glycerol phosphate) synthase, and a correlated motion of conserved residues during the allosteric transition of the protein.[29] Replica exchange was successfully applied to the study of allostery of IDPs, providing accurate calculation of binding affinities,[30] or suggesting allosteric activation mechanisms upon effector binding.[31] Umbrella sampling MD simulations were employed to compute the reaction path of the conformational transition of the tyrosine phosphatase 1B. The authors were able to estimate the change in the protein-allosteric inhibitor binding free energy due to the conformational change of the protein observed along the simulation time, providing a plausible explanation of the molecular mechanism of allosteric inhibition.[32] Funnel-Metadynamics of Hsp90, integrated with the analysis of internal dynamics of the structural ensembles visited during the simulations, allowed studying the recognition mechanisms of a benzofuran-based allosteric ligand (see case study II) and their consequences on the functional dynamics of the chaperone.[33]

As an alternative to enhanced sampling of the all-atom model, in coarse-grain simulations proteins are represented through an approximation with atom groups. One of the most used simplified representation is the elastic network model, where the vertices of the network represent molecular elements (e.g. atoms, amino acids, segments), while the edges connecting the nodes, represented as springs, model their coupling according to a metric that defines their relationship. Different metrics have been used, such as those based on mutual information, fluctuations in atomic positions or physical interaction. Once the network representation has been determined, several analysis techniques have been developed, in order to identify potential allosteric pathways. This approach has provided in several cases a general formalism of allosteric communication pathways in proteins.[34] As an example, MutInf,[35] a mutual-information metric derived from residue torsion angles, was used to identify statistically significant correlated motions in MD simulations, while VanWart et al[36] built a representative network of residue interactions based on correlated motions of ImGP synthase and identified the shortest path between the allosteric and orthosteric sites. Stetz et al[37] combined atomistic MD simulations and network-based modelling of the Hsp90-Cdc37 chaperone binding to its client kinase Cdk4, revealing the central role of Cdc37 in mediating client recognition and defining allosteric regulation of the chaperone-kinase complex.

According to the allostery model entailing structural coupling between the allosteric and orthosteric sites, an allosteric signal is achieved when highly interconnected residue groups establish long-range interactions. The interconnection degree among residues/residue groups can be determined in different ways. Blacklock and Verkhivker[38] identified groups of long-distance interconnected residues in Hsp90 by analyzing MD trajectories in various functional states. In particular, the authors calculated the force constants for each residue, which are correlated to the energy cost of residue vibration around the equilibrium point. They identified groups of highly connected residues, observing that residues with high force constants moved rigidly, and that the boundary between these protein zones and more flexible ones is characterized by a sudden change of these constants. The method successfully identified the "hot spot" regions, which corresponded to residues known

experimentally to be important hinge or dimerization sites. Using a similar conceptual framework for the evaluation of distant residues coupling, Morra et al used MD-based approaches to identify pathways of allosteric communication between distant residues of some members of the Hsp90 family; this information has then been used to identify molecules able to interfere with the protein activity through an allosteric mechanism.[39] More recently we applied the same approach to the mitochondrial homologue TRAP1.[40] In particular, MD-trajectories are analyzed with respect to the coordination propensity (CP) between any two residues separated by a defined distance as a function of the root mean square fluctuation (RMSF) of their distance. Small values of the CP parameter define amino acids that move in a coordinated way giving the highest contribution in modulating functional motions. The CP profile may change depending on the presence of a ligand, allowing connection between the internal dynamic and the observed ligand effects. The sets of amino acids endowed with potential allosteric control and defining a pocket with stereochemical properties apt to bind drug-like molecules have been used to guide modelling of pharmacophores that recapitulate complementary interactions with the putative site. This model has been used to screen drug databases and allowed the identification of several Hsp90 allosteric binders with interesting anticancer properties (see case study II below).

Experimental methods

From an experimental point of view, the study of protein allostery entails the analysis of local or distal conformational rearrangements and dynamic fluctuations on different timescales and possibly at the atomic level. In this review, we will focus our attention on the most promising approaches used to study such particular regulatory mechanism, like Nuclear Magnetic Resonance spectroscopy, Mass Spectrometry and X-ray crystallography. Solution NMR still represents the main experimental technique, used synergistically with computational methods, to probe allosteric phenomena, with the advantage of exploring diverse timescales. Thanks to recent advancements, mass spectrometry has emerged as a very useful tool to study allostery, enabling the assessment of local conformational dynamics and structural features, and allowing the determination of binding constants. Finally, new approaches in X-ray crystallography allow the study of the same protein in different conformations at atomic resolution, thus leading to the identification of pivotal regulatory mechanisms and of key residues involved.

Nuclear Magnetic Resonance

NMR spectroscopy is one of the main experimental techniques used to study allosteric processes, since it is sensitive to subtle changes in the protein three-dimensional structure. Recently, NMR has been extensively used for describing dynamic processes and quantifying dynamic equilibria on a broad timescale range, i.e. from pico-seconds to seconds. NMR alone is particularly useful when analyzing allosteric events in proteins with MW $<$ 40 kDa, while the combination with computational approaches is well established when characterizing allosteric phenomena in protein complexes with MW $>$ 50 kDa. All NMR spectroscopic observables can be useful for assessing protein conformational changes,[41] but considering that the typical molecular motions occur on the μs-ms timescale, only parameters with high sensitivity, such as chemical shift, residual dipolar coupling and relaxation times, will be considered in this review (Figure 4).

The chemical shift (δ), one of the principal NMR observables, is influenced by the environment of the observed nuclei. NMR

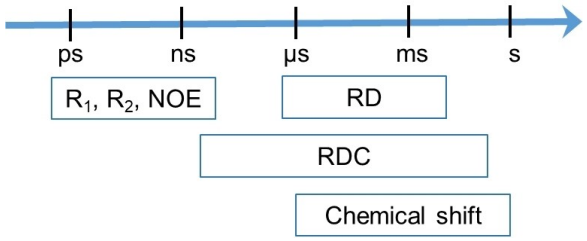


Figure 4. Allosteric phenomena can be investigated by NMR experiments on different timescales. Relaxation rates and NOE effects are useful for fast processes, while Relaxation Dispersion (RD), Residual Dipolar Couplings (RDCs) and Chemical Shifts (CS) allow studying slower phenomena.

chemical shifts are sensitive to even small changes in the protein 3D structure. For instance, a characteristic range of values is observed in random coils,[42] indicative of a lack of a defined structure and therefore of conformational averaging, and values outside the random-coil range point to well-defined structures. An allosteric event (e.g. the interaction of the protein with a binding partner, such as a substrate or an inhibitor) can cause a conformational change that modifies the local chemical environments of the involved amino acids or domains, perturbing NMR chemical shifts. Considering a simple case where two protein conformations (e.g. an inactive and an active state) are in equilibrium, different spectra are observed, depending on the exchange rate between the two. For a system where the two conformations are in slow exchange on the chemical shift time scale (species lifetime is long), the chemical environment of each conformation is sampled individually before exchange, showing separate, well-resolved peaks (Figure 5). In the case of fast-to-intermediate exchange between conformers, the observed chemical shift is the mole fractions weighted average of the δ values of the two states.

Transverse relaxation optimized spectroscopy (TROSY) is an NMR experiment that allows studying large proteins or protein complexes by analyzing the chemical shifts of the backbone ^{1}H-^{15}N nuclei.[43] This technique, in combination with isotope labelling, has opened up new avenues for the study of biomolecules with MW of up to 1000 kDa, as in the case of the slow ligand-mediated allosteric regulation of the 91 kDa TRAP-1 (see case study II).[44] Melacini et al. proposed and validated the chemical shift covariance analysis (CHESCA) that allows to identify correlations between

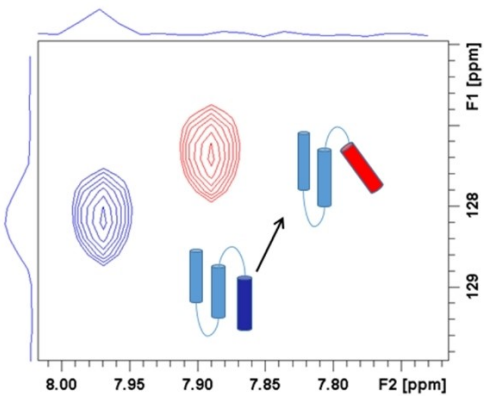


Figure 5. Chemical shift (δ). Schematic representations of the perturbation that can be observed for ^{1}H, ^{15}N or ^{13}C upon a slow conformational change that modifies the local chemical environments of the involved amino acids. δ values reported in ppm.

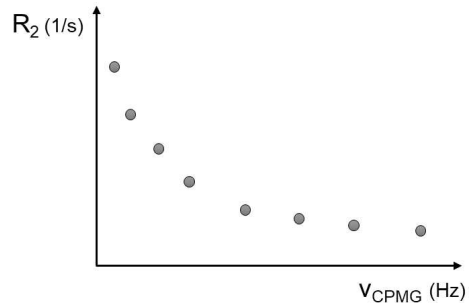


Figure 6. Relaxation dispersion. Schematic representations of a nonflat relaxation dispersion profile that can be detected for the residues experiencing a conformational transition on the micro- to millisecond timescale. In the case of a residue involved in the chemical exchange between two conformations, in an RD experiment a high R2 value is obtained at low field (ν_{CPMG}) and a low value at strong field.

chemical shifts variations and to cluster the protein active and inactive states, based on the assumption that residues that belong to the same allosteric network show a concerted response to the perturbation by an effector.[43]

Residual dipolar couplings (RDCs) are sensitive to small conformational changes in protein structures as well. When the protein molecules align weakly with the magnetic field applied, rather than tumbling freely in solution, the anisotropic dipolar couplings are not averaged anymore, resulting in an alignment tensor frame that provides indications about the relative orientation among internuclei vectors. RDCs are obtained from a comparison of multidimensional NMR spectra in the absence or in the presence of the alignment medium (e.g. phospholipid micelles, PEG ether bilayers, purple membranes, and polyacrylamide gels). In the first case only scalar couplings are observable, with dipolar couplings close to zero, while in the second case both scalar and dipolar couplings are observable. Difference spectra of the two experiments will provide the size as well as the sign of the RDCs. When the vector connecting two nuclei is parallel to the magnetic field, the coupling is at its strongest with a magnitude D_{max}. Observing the RDCs of protein backbone ^{1}H-^{15}N bonds (in two-dimensional ^{1}H-^{15}N HSQC spectra) allows the identification of an extended conformation (negative $^{1}D_{NH}$ RDCs values, corresponding to segments in which the NH vector is largely oriented perpendicular to the polypeptide chain) or α-helical segments (showing positive $^{1}D_{NH}$ values). While the RDCs between amide ^{1}H-^{15}N are widely used, the calculation of RDCs can be applied to different pairs of coupled nuclei. In particular, RDCs are an excellent tool for probing the conformational behavior of IDPs in solution, such as in the case of the determination of long-range and local structural features of α-synuclein[45] (an important IDP involved in human neurodegenerative diseases) and p53[46] (see case study IV).

Heteronuclear NMR relaxation studies (both longitudinal relaxation rate, R1 and transversal relaxation rate, R2) are frequently used to probe protein internal dynamics and conformational changes. Indeed, analysis of ^{13}C relaxation allows studying dynamic processes in the fast-exchange regime, such as methyl-group rotations on the picosecond timescale, and loop (local) motions on the nanosecond timescale. ^{15}N relaxation experiments are useful for the detection of protein backbone dynamics. In particular, ^{15}N R2 rates (R2 measured by Carr-Purcell-Meiboom-Gill (CPMG) sequence and the effective $R2_{eff}$, measured as a function of the delay τ between two successive 180° pulses of the sequence) are important indicators of protein dynamics. Since fast hydrogen exchange of amide protons with solvent protons can significantly affect measured R2 rates, ^{15}N longitudinal relaxation rate in the rotating frame (R1ρ) can be calculated instead. In general, low ^{15}N R1ρ and R2 values are indicative of fast dynamics, while high values can be correlated to slow movements.

The increase of the transverse relaxation rate R2, like that observed in the case of chemical exchange between active and inactive forms in allosteric proteins, causes line broadening of the associated NMR resonances. The measurement of the decay of transverse magnetization (R2) as a function of the strength of the applied radiofrequency field (ν_{CPMG}) is the so-called Relaxation Dispersion (RD, Figure 6).[47] RD measurements allow probing ^{1}H, ^{13}C, or ^{15}N spins providing site-specific information at atomic resolution. In particular, line broadening of the NMR peaks of one allosteric state can be quantified, obtaining information on the protein segment undergoing the induced conformational change.[48] RD techniques can be used even when ^{15}N and $^{13}C^{\alpha}$ chemical shift differences between the two allosteric states are too small to be quantitatively measured. They have already been successfully employed also in the study of allosteric mechanisms, such as in the case of CREB (cAMP-response element-binding protein) binding protein (CBP), where the propagation of allosteric information between two remote binding sites was demonstrated observing the relaxation dispersion data obtained for backbone amide ^{15}N and $^{13}C^{\alpha}$ nuclei.[49] RD experiments have been useful also in the case of galectin-1 (Gal-1), a human cell receptor with a pivotal role in the immune system, where they recently confirmed that there is a conformational entropy gain of the protein upon ligand binding.[50]

Mass spectrometry

Native mass spectrometry (native MS), i.e. mass analysis performed under non-denaturing conditions, provides thermodynamic data of intact protein assemblies[51] with the simultaneous detection of all coexisting states (i.e. species differing for the number of ligands bound). Assuming similar ionization efficiencies, the population of each state is proportional to the peak intensity, allowing the determination of single ligand binding constants. From a technical point of view, electrospray ionization (ESI) is sufficiently mild to preserve the integrity of ligand-protein complexes during the analysis, while Orbitrap and Q-ToF (quadrupole time-of-flight) analyzers have been recently developed to provide the appropriate *m/z* range and resolution. Paradigmatic is the case of the *E. coli* chaperonin GroEL,[52] a 800 kDa double-heptameric ring, containing 14 ATP binding sites, for which it was possible to distinguish ADP *vs* ATP binding, to determine the 14 different ATP binding constants, and to identify the allosteric model involved. It is possible to overcome the limitations of this technique (i.e. the requirement for low salt concentration and the formation of non-specific binding during desolvation in the gas-phase) by using a nanoflow electrospray,[53] more tolerant to high (mM) non-volatile salt concentrations, and by implementing a proposed mathematical method to distinguish between specific and aspecific binding (when the total number of binding sites is known).[54]

Within native MS, ion mobility-MS (IM-MS) is able to provide structural information. In particular, there is a time-resolved separation of ions migrating in a gas chamber, submitted to an electric field, proportional to their size and shape. A higher size therefore corresponds to a lower mobility.[55] Interestingly, this technique allows discriminating between conformational selection and induced fit, since in the first case a number of states would pre-exist in equilibrium in absence of a ligand, while in the second case only one conformational state would be expected before ligand addition. Indeed, conformational selection was identified as the allosteric regulation mechanism for protein kinase A.[56] The

main limitation of this technique is its resolution, since conformational states can be distinguished if they differ by at least 10%.

X-ray crystallography

X-ray crystallography represents the paramount technique for describing static conformations of macromolecules. Indeed, it has been extensively used through the years in combination with computational methods to obtain static pictures of allosteric phenomena and enzyme dynamics, as in the prototypical case of the ubiquitous enzyme dihydrofolate reductase (DHFR).[57] The biological importance of this protein prompted numerous structural studies, leading to more than 100 crystal structures of the enzyme in different states and/or in complex with a wide range of ligands deposited in the PDB.[58] Thanks to such efforts, it was possible to describe the allosteric transitions between a closed and an occluded state occurring in DHFR during the catalytic cycle, due to motions of the Met20 loop.[59]

Recent technical and experimental advancements further expanded the potential of X-ray crystallography in the study of protein flexibility and its direct connection to biological functions.[60] In such context, Multitemperature Multiconformer X-ray crystallography (MMX)[60] represents a particularly interesting new approach based on temperature perturbations and algorithms for alternative conformations modelling.[61]

So far, most protein crystal structures have been determined at cryogenic temperatures below the glass transition range (180–200 K), leading to a depopulation of protein conformational states,[62] hence drastically altering the dynamics of the system. MMX proposes to solve crystal structures at higher temperatures (about 277 K), where protein residues feature broader conformational ensembles, usually pivotal for biological functions.[63] Crystal structures of the same protein solved at different temperatures (both cryogenic and not) will allow to assemble a multiconformer model that includes multiple positions only for specific atoms where heterogeneity is sufficiently justified by collected data. This would result in alternative conformations and/or rotamers of potential key residues, involved in dynamic processes that could be correlated to protein function and/or allosteric regulation.[60] Such conformation ensembles with their relative occupancies are difficult to model manually with sufficient accuracy, due to the typical weakness of the corresponding electron density. An algorithm developed to overcome these limitations, *qFit*, is able to select a narrowed set of conformations that can collectively explain local electron density maps.[64] MMX has been successfully applied to the identification of complex allosteric networks in protein tyrosine phosphatase 1B[65] and could in principle be applied to all proteins yielding diffracting crystal, possibly expanding the perturbation source to pressure or pH.

Compared to traditional cryogenic X-ray diffraction, the multiconformer approach entails the collection of a much larger amount of data, also merged from different data sets. For this reason, data collection must be extremely scrupulous to avoid, e.g., crystal dehydration,[66] cryocooling-related distortions,[63] or radiation damage[67] at non-cryogenic temperatures. In this context, X-ray Free-Electron Lasers (XFELs) could be particularly helpful since they allow, thanks to the ultrafast timescale and extreme brightness of their pulses, not only to avoid substantial radiation damage,[68] but also to collect series of data sets that can provide details about local conformational changes related to specific processes. These structural approaches are therefore promising for the identification of new mechanisms for the transmission of protein allosteric signals.

2. Case Studies

2.1. Case study I: type III inhibitors of MEK kinase

Protein kinases are enzymes that catalyze the transfer of a phosphate group from ATP to a protein or peptide substrate. This phosphorylation usually results in a functional change of the substrate and is among the most common mechanisms of cell signalling in both eukaryotes and prokaryotes. Abnormal kinase regulation is implicated in many diseases, particularly in cancer.[69] The classical ATP-competitive approach for kinase inhibitors (i.e. type I and II inhibitors) suffers from cross-reactivity with homologous proteins, leading to unwanted side effects and off-target toxicity. On the other hand, molecules targeting allosteric pockets (i.e. type III and IV, according to their distance from the ATP binding site) are gaining much interest as alternative therapeutic drugs, since they allow overcoming drug resistance and achieving an improved selectivity profile.[70]

Type III inhibitors for the mitogen-activated protein kinases MEK1/2 bind into a pocket adjacent to the ATP site and have been largely investigated. In 2013, the FDA approved the first type III inhibitor, Trametinib[71] (Figure 7 A), boosting the discovery of other type III inhibitors in the field. Currently, there are three FDA approved type III binders (i.e. Trametinib, Cobimetinib[72] and Binimetinib[73]) used for the treatment of melanoma caused by mutations in the kinase BRAF (V600E/K), while other drugs are in development for the treatment of several other cancers, with more than ten additional candidates already in clinical trials.[70a]

Trametinib was discovered through a phenotypic cellular screening assay[74] and was developed following a SAR approach aimed at optimizing pharmaco-kinetic parameters and antiproliferative effect on cancer cells. The identification of MEK1/2 as the Trametinib molecular target was accomplished only after its discovery and optimization, probably explaining the uniqueness of Trametinib's core structure (Figure 7A) compared to other clinical MEK inhibitors. The binding mode of Trametinib into MEK1 allosteric pocket was only recently revealed by crystallography.[75] As expected, the compound occupies the allosteric site adjacent to the ATP pocket, extending further into the KSR (Kinase Suppressor of RAS) interaction interface (Figure 7B). Additionally, drug binding induces a conformational change in the allosteric pocket, enlarging it. The discovery of Cobimetinib relied instead on the X-ray structures of a series of PD0325901 analogues[76] bound to MEK1.[77] In the crystal structures of MEK1-KSR1 complexes with PD0325901 and Cobimetinib no interaction with KSR was observed, instead (Figure 7C and Figure 7D, respectively).

2.2. Case study II: chaperone Hsp90

The 90 kDa heat shock proteins (Hsp90) are homodimeric molecular chaperones playing a pivotal role in maintaining eukaryotic cells homeostasis. As chaperone, Hsp90 promotes, assisted by a cohort of co-chaperones, the correct conforma-

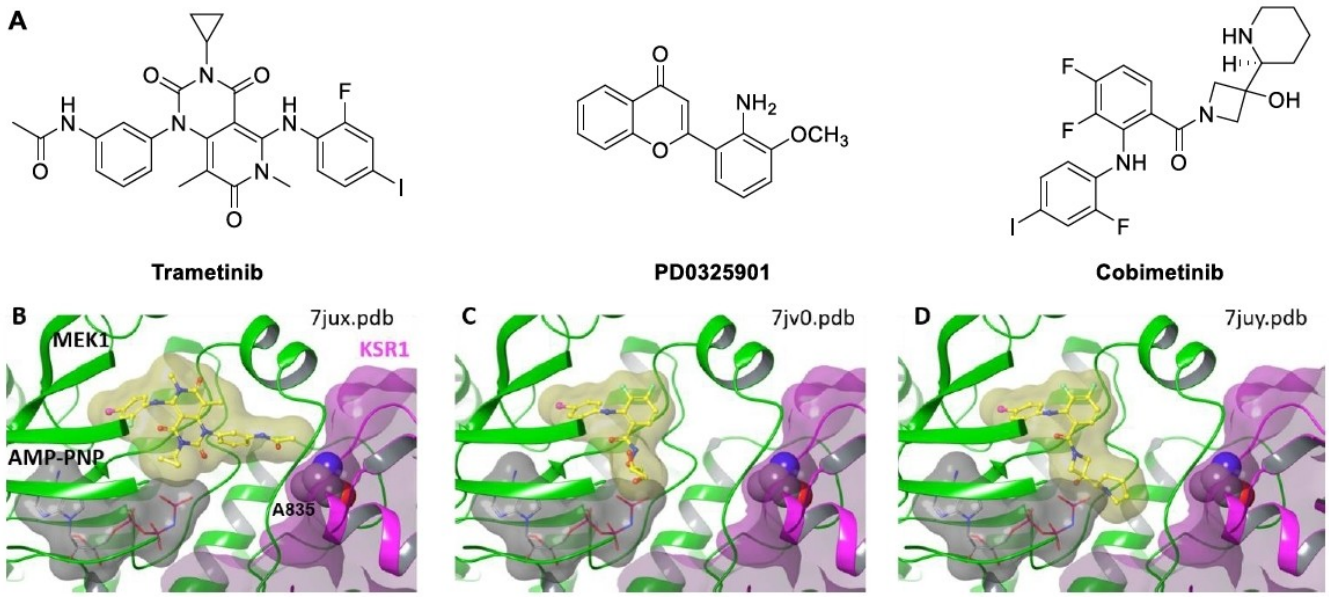


Figure 7. MEK kinase allosteric inhibitors. A) Chemical structure of MEK allosteric inhibitors Trametinib, PD0325901 and Cobimetinib; **B–D)** X-ray crystal structures of each inhibitor bound to the MEK1-KSR1 complex. Only Trametinib (7jux.pdb) is in contact with KSR1 (A835).

tional folding of other proteins, called clients. Having several hundred clients, Hsp90 is considered a hub protein that upholds proteostasis under both physiological and stressful conditions.[78] In humans, there are four functional Hsp90 paralogs: Hsp90α and Hsp90β in the cytoplasm, Grp94 in the endoplasmic reticulum[79] and TRAP1 in mitochondria.[80] These paralogs have 30–40% sequence identity, reflecting the structural similarity of their domains. Hsp90 predominantly forms homodimers, with each monomer consisting of three domains, the highly conserved N-terminal (NTD) and C-terminal (CTD) domains and the middle domain (MD), which is in turn subdivided in large- and small- middle domain.[81] Hsp90 performs its chaperoning function through large conformational changes triggered by ATP binding and hydrolysis at the NTD.[80] Similarly to kinases, the prominent strategy for the development of inhibitors of the enzymatic activity of Hsp90 consists in targeting the N-terminal ATP binding site with small molecules.[82] This approach leads to the concurrent induction of the heat shock response, such as the over-expression of other HSPs (e.g. Hsp70) and the activation of HSF1. Additionally, due to the similarity of the ATP binding site in all Hsp90 family members, these compounds are not paralog specific, blocking indiscriminately all Hsp90-dependent processes.[83] Differences among paralogs in terms of cell localization, expression levels, client and co-chaperone specificity and molecular mechanisms prompted us, and other groups, to search for and target potential allosteric sites.

The Neckers' group experimentally demonstrated the existence of an additional druggable site in the CTD of Hsp90[84] and showed that binding of the coumarin antibiotics novobiocin and chlorobiocin to this site, distal from the nucleotide orthosteric pocket, resulted in impairment of Hsp90 chaperoning function.[85] Importantly, novobiocin and its analogues were shown not to induce the heat shock response.[83] This discovery opened up new avenues for pharmacological targeting of Hsp90 with allosteric modulators. In the absence of an Hsp90-novobiocin complex X-ray crystal structure, design, synthesis, SAR and computational studies have been carried out to develop novel potential ligands.

In this context, the Colombo's group worked on allosteric lead discovery and optimization, ameliorating potencies of these coumarine-based allosteric inhibitors,[86] while we designed and synthetized the first allosteric activators of Hsp90 functional activity.[87] The Colombo's group first identified the putative allosteric binding site in the CTD of Hps90, i.e. residues in the CTD that responded to ATP binding at the orthosteric NTD site, through analysis of the long-range coordination propensity among residue-pairs (CP analysis) derived from molecular dynamics simulations.[39a] Chemical features and conformational properties of this pocket were used as a template for the development of a pharmacophore model, which recapitulated the complementary interactions necessary for a ligand to bind the site. The pharmacophore was then used for a virtual screening followed by in vitro assays from which the natural compound *Eupomatenoid-2* was selected for SAR studies.[39b] Interestingly, the putative binding site identified by CP analysis is in agreement with recent findings based on mutational studies.[88]

Following these studies, a glycorandomization of the *Eupomatenoid-2* aglycon (namely *Eupomatenoid-6*) afforded a first set of 14 glycosylated compounds (Figure 8A).[87a] Combining structural information on the potential allosteric site with docking calculations allowed predicting a putative binding mode for this set, which led to the design and synthesis of a few derivatives aimed at establishing key interactions with specific protein residues. These compounds were able to

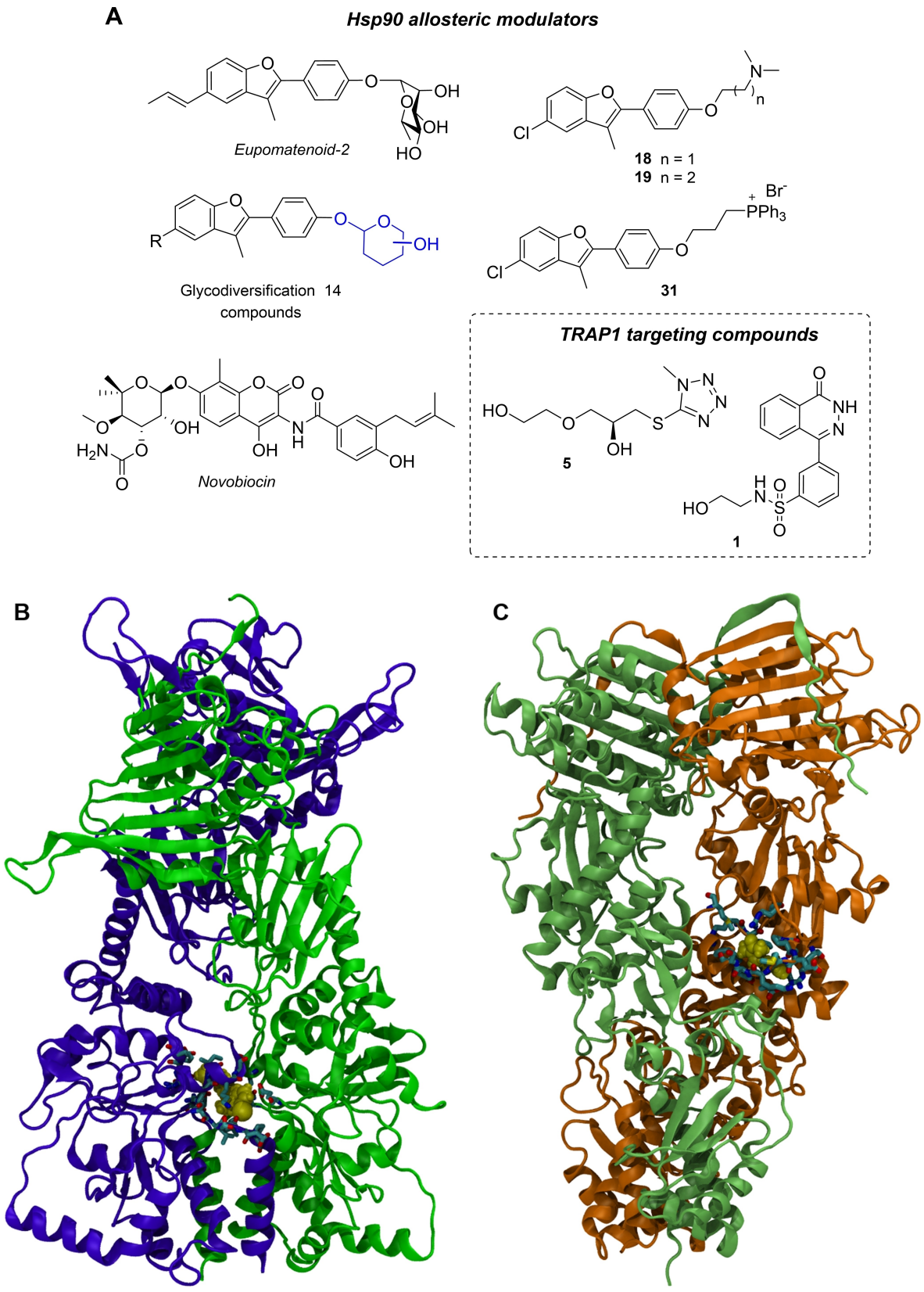


Figure 8. Hsp90 and TRAP1 allosteric inhibitors. A) Chemical structure of some of the benzofuran-based Hsp90 allosteric modulators we have designed and synthesized over last decade (a family of 14 glycodiversified derivatives, Hsp90 activators **18** and **19** and the mitochondrial targeting compound **31**) along with the parent compound *Eupomatednoid-2*, the CTD inhibitor Novobiocin and the most relevant TRAP1 targeting compounds reported, **1** and **5**.[89] **B)** Hsp90 activator **18** (yellow spheres) in complex with Hsp90 in its closed-twisted conformation (protomers in blue and green). **C)** Compound **5** (yellow spheres) in complex with TRAP1 dimer in its closed conformation (protomers in orange and green).

enhance Hsp90 ATPase turnover by up to 7-fold (**18** and **19**, Figure 8A).[87b] The benzofuran derivatives showed the ability to stimulate Hsp90 ATPase in a structure-dependent manner, acting synergistically with the activating cochaperone Aha1 and the ability of modulating the direct interaction of Hsp90 with its cochaperone Sba1 (p23).[87b] The Colombo's group performed also extensive Funnel Metadynamics calculations of HSP90 in complex with compound **18**, shedding light on the allosteric mechanism exerted by the compound at atomic level, and allowing the identification of structural elements of the receptor that guide the entrance of the ligand into the allosteric cavity.[33] Conjugation of the best activator with a mitochondrial penetrating group resulted in cancer cells cytotoxicity in the order of 500 nM (**31**, Figure 8A). These compounds showed antiproliferative activity also against drug-resistant cancer cell lines, suggesting the idea that acceleration of conformational dynamics could represent a new way of perturbing the chaperoning mechanisms that underlie cell viability.[87c] Generated ligand-protein complexes were used to investigate the effect of the allosteric ligands on Hsp90 conformational dynamics, suggesting a protein population shift towards the catalytically active asymmetric state upon ligand binding (Figure 8B).[90] Finally, some of these compounds (**18** and **19**) were able, via boosting of Hsp90 chaperoning activity, to rescue the activity of ΔF508-CFTR, the temperature-sensitive ΔF508 mutant of the cystic fibrosis transmembrane receptor.[91]

TRAP1,[80] the Hsp90 mitochondrial paralog, is even more complex to decipher from a functional point of view compared to cytosolic Hsp90. Indeed, both under- and overexpression of TRAP1 has been observed in cancer cells compared to healthy cells. Therefore, the development of TRAP1 selective inhibitors/modulators would help shed light on its role in tumor progression.[92] ^{15}N and ^{13}C labelled TRAP1 was analyzed by TROSY NMR experiments[92] in its inactive and active (bound to tryptophan) states. The evaluation of the line broadening and of the chemical shift perturbations in both *apo* and *holo* forms demonstrated that tryptophan plays a key role in TRAP modulation by altering its RNA-binding affinity highlighting the dynamic behavior of the protein.

The crystal structure of TRAP1 bound to different ATP analogues showed that one of the two monomers is buckled, resulting in a peculiar asymmetric conformation where the maximal differences occur at the interface between the middle and the C-terminal domain, in a broadly conserved region responsible for client binding.[80] Experimental results showed that asymmetry is also present in solution, and it determines differential hydrolysis for each protomer, with the buckled conformation favoring ATP hydrolysis. In this model, asymmetry plays a key role in the mechanism that drives chaperone function.[93]

The nucleotide-dependent internal dynamics of TRAP1 was investigated, establishing a mechanistic link between the nucleotide state in the N-terminal domain and the asymmetric modulation of the dynamic and structural properties of the client-binding region in the Middle domain.[40] Specifically, using the CP analysis of atomistic MD simulations of TRAP1 in different bound states in the N-terminal domain it was demonstrated that the asymmetric substructures at the interface between the distal Middle and C-terminal domains of the protein undergo the most intense dynamic modulation upon nucleotide exchange, depending on the specific protomer to which the nucleotide is bound. In particular, CP analysis indicated that the presence of ATP in the N-terminal domains generates a conformational allosteric signal that is relayed throughout the protein via specific secondary-structure motifs. These signals at either catalytic sites were directed toward the same protein region in the middle domain of TRAP1, consistent with the experimental findings reported by Schulze *et al.*[94] This result discloses a specific cross-talk at atomic level between the nucleotide binding site involved in ATP hydrolysis and the client-recognition protein region, and supports the observation that structural asymmetry plays a role in functional regulation of the chaperone, within the closed active state of TRAP1.

Building on this result, it was demonstrated that this asymmetric substructure was a promising starting point for specifically targeting TRAP1 through small molecule inhibitors/modulators (Figure 8A and Figure 8C).[89] Indeed, even if the sequence of the client-binding region is broadly conserved, this structural asymmetry has not been observed in other Hsp90 homologues. A druggable pocket was identified in the most populated structural ensembles obtained from MD simulations. The pharmacophore, generated as reported above for Hsp90, was then used for a virtual screening and the selected small molecules, endowed with optimal chemical and structural characteristics to establish productive interactions with the client-binding region, were tested in biochemical, cellular, and in vivo experiments. The results showed that they selectively perturb TRAP1 ATPase activity, they affect the enzymatic activity of the TRAP1 client succinate dehydrogenase, both in cancer cells and in zebrafish larvae, and they abolish cell tumourigenicity. Using this approach, a structure- and dynamics-based model of allosteric perturbation was also generated. According to this analysis, allosteric inhibitors disrupt the dynamic coupling mechanisms between the M-domain and ATP binding/hydrolysis, favoring an ensemble of states where the functional motion of the chaperone is hampered, which ultimately reverberates in the observed functional perturbation. These results demonstrated that for these systems the analysis of the protein conformational dynamics on a short time scale could be an effective way for identifying selective allosteric modulators and for clarifying their perturbation mechanism of the protein activity.

Summing up, the strengths of this approach are the exploitation of protein dynamics to identify potential allosteric pockets, the generation of pharamcophore models that take into account the dynamic features of the cavity, and the capability to study and characterize the effects of targeting unexplored allosteric sites at the atomic level. We expect that the integration of structural dynamics studies with the development of new compounds will contribute to disentangle the specific functions of each Hsp90 paralog, which is fundamental to comprehend how pathways dysfunctions regulated by this protein family influence the pathogenesis of various diseases,

ranging from cancer to neurodegeneration and inflammation, and thus to conceive new therapeutic intervention strategies.

2.3. Case study III: lectin cell receptors

A third class of proteins relevant in the context of allosteric regulation are cell receptors. Sensing external stimuli and transforming them into biochemical signals across the cell membrane, G-protein-coupled receptors (GPCRs) are arguably the most important family of this class. These proteins are intrinsically allosteric and their regulation mechanisms have already been recently reviewed.[95]

Here we would like to discuss the allosteric regulation of lectins, a less known but perhaps equally interesting class of receptors. Lectins, i.e. proteins that specifically recognize carbohydrates, can be found across most life kingdoms performing a myriad of biological functions. On one hand, pathogens such as viruses, bacteria and fungi are able to start their infection process by exposing lectins able to adhere to the cell surface of host organisms. On the other hand, mammalian host cells have developed their own set of lectins. In particular, within the innate immune system, the so-called pattern recognition receptors (PRRs) are able to discriminate between self and non-self, triggering the appropriate immune response against pathogens, while lectins expressed on different cell types mediate cell-recognition processes such as leukocytes rolling and extravasation, and inflammation.[96] Targeting lectins with glycoconjugates and glycomimetics, both on the microbial or the host side, is indeed an innovative strategy to prevent the first step of microbial infection, i.e. pathogen adhesion to the host cell, avoiding the rapid onset of resistance.[97]

As an example of bacterial lectin, FimH mediates the adhesion of Gram-negative uropathogenic *Escherichia coli* (UPEC) to urothelial cells (i.e. cells covering the bladder and the urinary tract). This is the key step in urinary tract infections (UTIs), where bacterial adhesion is mediated by surface organelles called pili and fimbriae, constituted by oligomeric proteins. Adhesin FimH is located on the tip of fimbriae and binds to mannosylated glycoconjugates (α-D-linked mannoses of N-linked glycans of the receptor uroplakin 1a) displayed by the bladder endothelium, with a carbohydrate recognition domain (CRD) that is highly conserved in all *E. coli* clinical isolates. UTIs are particularly difficult to eradicate also due to a clever adaptation mechanism called catch-bond,[98] i.e. the ligand-receptor interaction is reinforced by mechanical stress, such as the shear force produced by urine flow. UPEC cells can therefore adapt to dynamically varying conditions, allowing for colonization of the host in absence of tensile forces and for solid anchoring under stressful conditions.

From a structural point of view, the FimH lectin domain (H_L, red, Figure 9) interacts with the pilin domain (H_P, yellow, Figure 9), determining the affinity state for the mannosylated ligand. In a resting state, the interaction of these two domains stabilizes H_L in a low-affinity state characterized by a shallow binding pocket (encounter complex, Figure 9). The mannosylated ligand binding induces a conformational change leading to a medium-affinity state characterized by micromolar affinities and fast off-rates. The shear force produced by urine flow pulls apart H_L and H_P, locking a high-affinity conformation (up to

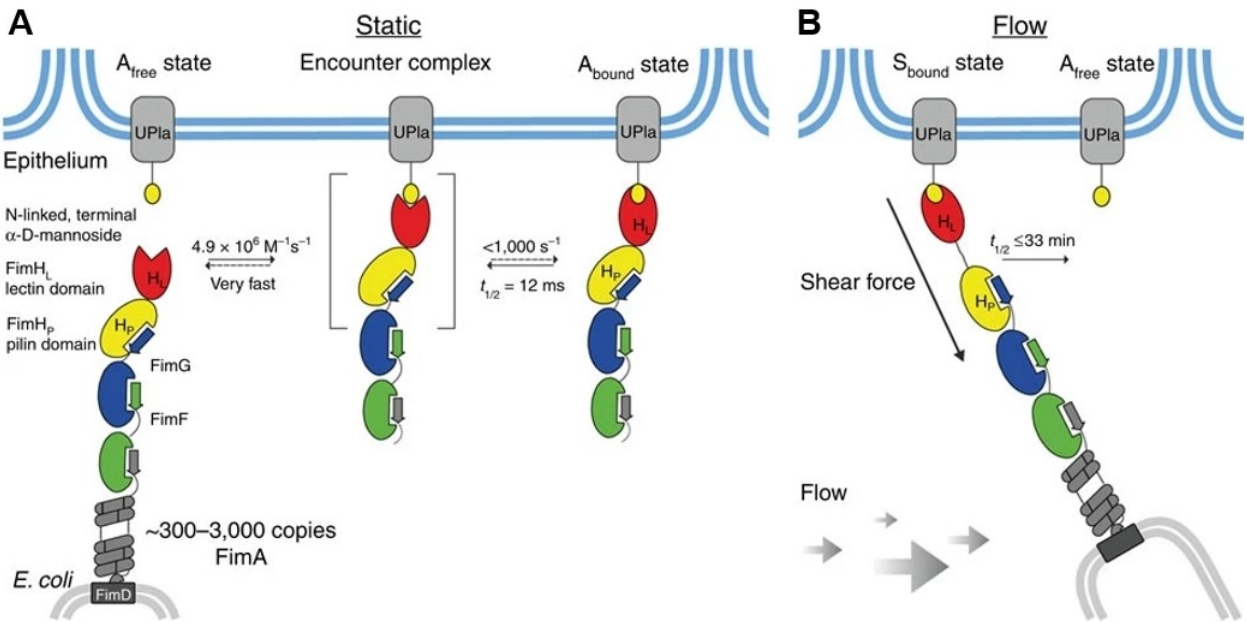


Figure 9. Allosteric regulation of FimH-mediated UPEC cell adhesion. A) In the absence of shear force, the FimH-Uroplakin 1a (UPIa) complex is loose and dynamic, enabling bacterial motility on the bladder epithelium. The pilin domain acts as a negative allosteric regulator, promoting the dissociation of the receptor from the FimH lectin domain. **B**) Shear force increases the population of the high affinity $\mathbf{S_{bound}}$ state, in which the pilin and lectin domains are separated. Dissociation of the FimH-UPIa complex is greatly disfavored, promoting the firm anchorage of the bacterium to the host epithelium. Reproduced under the CC BY 4.0.[98]

300-fold enhancement, Figure 9B). Point mutation experiments allowed identifying the allosteric regulatory regions,[99] while computational studies pinpointed a *population shift* model for the allosteric regulation mechanism.[100] More recently, a minimalistic allosteric system was engineered at the lectin domain level alone, showing that a single mutation or the introduction of a disulphide bridge in the H_L–H_P interdomain region can convert the intrinsically high-affinity monomeric H_L into a low affinity state.[101]

Langerin is a mammalian lectin belonging to the C-type subfamily, which is characterized by a common protein fold where the carbohydrate recognition event is predominantly mediated by a Ca^{2+} ion.[102] C-type lectins are increasingly recognized for their function in innate and adaptive antimicrobial immune responses and, more recently, for their important role in autoimmune diseases.[103] In particular, langerin is mainly found on Langheran cells, i.e. antigen-presenting dendritic cells (DCs) that reside in epithelia,[104] and is able to detect both self (e.g. blood group antigens[105]) and non-self-epitopes (e.g. several viruses, including HIV[106] and fungi). Interestingly, allosteric regulation of langerin's function is twofold.The affinity of the Ca^{2+} ion, necessary for ligand recognition, is fine-tuned by a pH-dependent allosteric network where a histidine residue acts as pH sensor that indirectly regulates the site accessibility via coupled motion of two neighbouring loops.[107] These conformational changes therefore play a role also in cargo uptake and release.

Lectins binding sites are generally broad, hydrophilic and solvent exposed, making the design of drug-like ligands rather difficult. The search for more druggable secondary sites (either cryptic and/or allosteric) is therefore a sought-out approach when targeting this class of proteins. In this framework, langerin represents the first example of an allosteric inhibition of a mammalian lectin with small molecules.[108] Fragment screening followed by SAR studies on a focused analogues library led to the identification of thiazolopyrimidines as potent (low micromolar range) allosteric inhibitors of murine langerin, showing also selectivity over the human homologue and DC-SIGN.[108] Titration of the ^{15}N-labelled langerin CRD with mannose (the orthosteric site ligand) and a thiazolopyrimidine derivative induced chemical shift perturbations of different ^{1}H-^{15}N cross-peaks, strongly supporting the hypothesis of different binding sites. This is an important feature since ligands targeting the primary carbohydrate-binding sites are generally promiscuous for C-type lectin with rare exceptions.[109]

2.4. Case study IV: Intrinsically Disordered Proteins (IDPs)

The long established dogma that a protein's function depends on its well defined and folded three-dimensional structure has been challenged in the mid-1990s by the discovery that disordered regions are very common in eukaryotic proteins.[110] Indeed, Intrinsically Disordered Proteins (IDPs), despite being fully functional, do not fold spontaneously into a stable and defined arrangement, because they generally lack the bulky hydrophobic amino acids required to constitute the hydrophobic core at the base of structured domains. Instead, IDPs rapidly sample a range of conformations with varying degrees of disorder, participate in many cell processes, especially regulation and signalling,[110] and frequently undergo PTMs, further increasing their number of functional states.[111]

Among the models described above, the Ensemble Allosteric Model (EAM)[8] is the only suitable one for describing allostery in IDPs, where a definite structure is not necessarily present and where function is not a direct consequence of a folded structure. In the EAM, binding of an allosteric effector is treated as an energetic perturbation of an ensemble of pre-existing protein conformations. The perturbation redistributes the ensemble population, affecting the binding of a third partner (i.e. the protein function).[112]

The allosteric regulatory mechanisms in which IDPs play a role are diverse.[113] As an example, allostery is driven by thermodynamic coupling in the human Glucocorticoid Receptor (GR),[114] a nuclear transcription factor involved in development, metabolism and inflammatory response. Of the three inherently dynamic functional domains composing GR, the N-terminal intrinsically disordered (N-term ID) domain is of particular interest and can be divided into a regulatory (R) and a functional (F) subdomain. Thermodynamic coupling between the R and F subdomains, regulated by the N-term ID domain length observed in different translational isoforms, was reported to impact significantly on the receptor physiological distribution and functional specificity.[115] Further studies also revealed an additional regulatory level where coupling occurs between the N-term ID domain and the downstream structured DNA-binding domain.[116]

Another important example of proteins with intrinsic disorder is provided by p53, a pivotal transcription factor that regulates cell cycle progression, DNA repair and apoptosis, mainly functioning as tumor suppressor, hence named "guardian of the genome".[117] Under normal conditions, p53 intracellular levels are usually low thanks to the proteasome degradation system, mainly by interaction with the Mdm2 ubiquitin E3 ligase. On the contrary, stressful conditions (e.g. DNA damage, hypoxia, heat shock) induce p53 stabilization and activation. From a structural point of view, p53 is a homotetramer characterized by largely folded core domains (i.e. the tetramerization (TD) and DNA-binding (DBD) domains) that are linked together and flanked by intrinsically disordered domains at the N- (transactivation domain, TAD) and C- (regulatory domain, RD) termini, with 37% of the protein structure being intrinsically disordered. Even though it is mostly folded, the DBD of p53 shows a relatively low thermodynamic stability and a high conformational plasticity.[118] The structures of individual (folded) core domains were first solved by using X-ray crystallography and NMR, while the abovementioned residual dipolar couplings (RDCs) technique, in combination with SAXS and MD simulations, was particularly useful for the elucidation of the N-terminal TAD.[46] In particular, a single helix turn, populated to ca. 30%, was identified as the nascent helix that folds completely upon interaction with Mdm2.[46] Further studies showed that increasing the residual helicity of this region by mutation of specific

proline residues in the TAD led to a tenfold increase in Mdm2 affinity, although impairing p53 downstream cell signalling and disrupting its protective DNA repair function.[119] Indeed, p53 allosteric regulation is particularly intricate, depending not only on the intrinsic dynamicity of the protein but also on several types of PTMs (e.g. acetylation and multisite phosphorylation) and oligomerization.[120]

Mutations in the *TP53* gene are the most common in cancer cells and are generally localized in the DBD, thus altering its global conformation. Interestingly, the conformational regulation of p53 relies on a complex chaperone machine composed of Hsp70, Hsp90 (described in case study II), and their co-chaperones. The two chaperones seem to execute opposed and complementary effects: while both chaperones support p53 DNA binding, Hsp70 can sequester p53 in the cytoplasm negatively regulating its stability. On the other hand, Hsp90 promotes the stability of p53 by interacting with its DBD and maintaining its folding, and support its nuclear translocation. In particular, Hsp90 associates much more stably with mutated variants of p53 than with the wild type protein, therefore protecting it from ubiquitination by Mdm2.[121]

3. Conclusions

Considerable progress has been attained in the last decades in understanding the principles governing allosteric phenomena, and in the enlargement of the toolbox available to chemists and biophysicists for their investigation.

From a computational point of view, software for allosteric site prediction have been developed, along with accessible databases of known allosteric modulators, while MD trajectories can be perused from different angles to detect signal propagation pathways. Experimentally, NMR sensitivity to subtle changes in protein three-dimensional structures makes it a powerful technique to investigate allosteric events on a wide time-scale range. Remarkable technical advancements in Mass Spectrometry currently allow the measurement of binding constants and the detection of all coexisting states, therefore helping shedding light on the allosteric mechanism involved. Finally, an innovative X-ray crystallography technique (MMX) allows generating multiconformer models accounting for a certain degree of conformational flexibility of the protein residues.

Despite such remarkable progresses, allostery is far from being an open book, with its role in evolution and adaptation just starting to be unveiled.[122] Certainly, the advantages showed by allosteric over orthosteric drugs in terms of resistance and selectivity profile are strongly stimulating research on this topic, and we posit this might become, whenever possible, the preferred approach for the development of future therapeutic agents.

Acknowledgements

E.M. would like to thank Prof. Giorgio Colombo for his direction and support in the development of the Hsp90 and TRAP1 computational work mentioned in this minireview.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: Allostery · Molecular dynamics · Native mass spectrometry · NMR spectroscopy · X-ray crystallography

[1] J. Monod, J. Wyman, J. P. Changeux, *J. Mol. Biol.* **1965**, *12*, 88–118.
[2] D. E. Koshland, Jr., G. Nemethy, D. Filmer, *Biochemistry* **1966**, *5*, 365–385.
[3] G. G. Hammes, Y.-C. Chang, T. G. Oas, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13737.
[4] D. Thirumalai, C. Hyeon, P. I. Zhuravlev, G. H. Lorimer, *Chem. Rev.* **2019**, *119*, 6788–6821.
[5] A. Cooper, D. T. F. Dryden, *Eur. Biophys. J.* **1984**, *11*, 103–109.
[6] C.-J. Tsai, S. Kumar, B. Ma, R. Nussinov, *Protein Sci.* **1999**, *8*, 1181–1190.
[7] C. J. Tsai, R. Nussinov, *PLoS Comput. Biol.* **2014**, *10*, e1003394.
[8] H. N. Motlagh, J. O. Wrabl, J. Li, V. J. Hilser, *Nature* **2014**, *508*, 331–339.
[9] R. Nussinov, C. J. Tsai, *Cell* **2013**, *153*, 293–305.
[10] J. S. Yang, S. W. Seo, S. Jang, G. Y. Jung, S. Kim, *PLoS Comput. Biol.* **2012**, *8*, e1002612.
[11] D. Ni, Y. Li, Y. Qiu, J. Pu, S. Lu, J. Zhang, *Trends Pharmacol. Sci.* **2020**, *41*, 336–348.
[12] X. Liu, S. Lu, K. Song, Q. Shen, D. Ni, Q. Li, X. He, H. Zhang, Q. Wang, Y. Chen, X. Li, J. Wu, C. Sheng, G. Chen, Y. Liu, X. Lu, J. Zhang, *Nucleic Acids Res.* **2020**, *48*, D394–D401.
[13] a) S. Lu, X. He, D. Ni, J. Zhang, *J. Med. Chem.* **2019**, *62*, 6405–6421; b) S. Lu, Q. Shen, J. Zhang, *Acc. Chem. Res.* **2019**, *52*, 492–500; c) G. M. Verkhivker, S. Agajanian, G. Hu, P. Tao, *Front. Mol. Biosci.* **2020**, *7*, 136.
[14] a) O. Sheik Amamuddy, W. Veldman, C. Manyumwa, A. Khairallah, S. Agajanian, O. Oluyemi, G. Verkhivker, O. Tastan Bishop, *Int. J. Mol. Sci.* **2020**, *21*; b) J. G. Greener, M. J. Sternberg, *Curr. Opin. Struct. Biol.* **2018**, *50*, 1–8.
[15] a) W. Huang, S. Lu, Z. Huang, X. Liu, L. Mou, Y. Luo, Y. Zhao, Y. Liu, Z. Chen, T. Hou, J. Zhang, *Bioinformatics* **2013**, *29*, 2357–2359; b) A. S. Chen, N. J. Westwood, P. Brear, G. W. Rogers, L. Mavridis, J. B. Mitchell, *Mol. Inf.* **2016**, *35*, 125–135; c) R. Akbar, V. Helms, *Chem. Biol. Drug Des.* **2018**, *91*, 845–853.
[16] Q. Shen, G. Wang, S. Li, X. Liu, S. Lu, Z. Chen, K. Song, J. Yan, L. Geng, Z. Huang, W. Huang, G. Chen, J. Zhang, *Nucleic Acids Res.* **2016**, *44*, D527–D535.
[17] W. Huang, G. Wang, Q. Shen, X. Liu, S. Lu, L. Geng, Z. Huang, J. Zhang, *Bioinformatics* **2015**, *31*, 2598–2600.
[18] K. Song, X. Liu, W. Huang, S. Lu, Q. Shen, L. Zhang, J. Zhang, *J. Chem. Inf. Model.* **2017**, *57*, 2358–2363.
[19] a) A. Panjkovich, X. Daura, *Bioinformatics* **2014**, *30*, 1314–1315; b) J. G. Greener, M. J. Sternberg, *BMC Bioinf.* **2015**, *16*, 335.
[20] a) Y. Qi, Q. Wang, B. Tang, L. Lai, *J. Chem. Theory Comput.* **2012**, *8*, 2962–2971; b) P. Weinkam, J. Pons, A. Sali, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4875–4880.
[21] P. Schmidtke, V. Le Guilloux, J. Maupetit, P. Tufféry, *Nucleic Acids Res.* **2010**, *38*, W582–W589.
[22] a) X. Ma, H. Meng, L. Lai, *J. Chem. Inf. Model.* **2016**, *56*, 1725–1733; b) Y. Xu, S. Wang, Q. Hu, S. Gao, X. Ma, W. Zhang, Y. Shen, F. Chen, L. Lai, J. Pei, *Nucleic Acids Res.* **2018**, *46*, W374–W379.
[23] B. Huang, M. Schroeder, *BMC Struct. Biol.* **2006**, *6*, 19.
[24] S. Mitternacht, I. N. Berezovsky, *PLoS Comput. Biol.* **2011**, *7*, e1002148.
[25] D. Clarke, A. Sethi, S. Li, S. Kumar, R. W. F. Chang, J. Chen, M. Gerstein, *Structure* **2016**, *24*, 826–837.
[26] B. A. Kidd, D. Baker, W. E. Thomas, *PLoS Comput. Biol.* **2009**, *5*, e1000484.

[27] L. Mouawad, D. Perahia, C. H. Robert, C. Guilbert, *Biophys. J.* **2002**, *82*, 3224–3245.
[28] K. Sharp, J. J. Skinner, *Proteins* **2006**, *65*, 347–361.
[29] R. E. Amaro, A. Sethi, R. S. Myers, V. J. Davisson, Z. A. Luthey-Schulten, *Biochemistry* **2007**, *46*, 2156–2173.
[30] Y. Wang, C. L. Brooks, *J. Phys. Chem. Lett.* **2019**, *10*, 5963–5968.
[31] X. Pang, H.-X. Zhou, *Biophys. J.* **2015**, *109*, 1706–1715.
[32] W. Cui, Y.-H. Cheng, L.-L. Geng, D.-S. Liang, T.-J. Hou, M.-J. Ji, *J. Chem. Inf. Model.* **2013**, *53*, 1157–1167.
[33] I. D'Annessa, S. Raniolo, V. Limongelli, D. Di Marino, G. Colombo, *J. Chem. Theory Comput.* **2019**, *15*, 6368–6381.
[34] a) W. Zheng, B. R. Brooks, D. Thirumalai, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7664; b) C. Chennubhotla, I. Bahar, *PLoS Comput. Biol.* **2007**, *3*, e172; c) D. Shukla, C. X. Hernández, J. K. Weber, V. S. Pande, *Acc. Chem. Res.* **2015**, *48*, 414–422; d) Z. Yang, P. Majek, I. Bahar, *PLoS Comput. Biol.* **2009**, *5*, e1000360; e) D. Shukla, Y. Meng, B. Roux, V. S. Pande, *Nat. Commun.* **2014**, *5*, 3397.
[35] a) C. L. McClendon, G. Friedland, D. L. Mobley, H. Amirkhani, M. P. Jacobson, *J. Chem. Theory Comput.* **2009**, *5*, 2486–2502; b) C. L. McClendon, L. Hua, G. Barreiro, M. P. Jacobson, *J. Chem. Theory Comput.* **2012**, *8*, 2115–2126.
[36] A. T. Van Wart, J. Eargle, Z. Luthey-Schulten, R. E. Amaro, *J. Chem. Theory Comput.* **2012**, *8*, 2949–2961.
[37] G. Stetz, G. M. Verkhivker, *J. Chem. Inf. Model.* **2016**, *56*, 1490–1517.
[38] K. Blacklock, G. M. Verkhivker, *PLoS Comput. Biol.* **2014**, *10*, e1003679.
[39] a) G. Morra, G. Verkhivker, G. Colombo, *PLoS Comput. Biol.* **2009**, *5*, e1000323; b) G. Morra, M. A. C. Neves, C. J. Plescia, S. Tsustsumi, L. Neckers, G. Verkhivker, D. C. Altieri, G. Colombo, *J. Chem. Theory Comput.* **2010**, *6*, 2978–2989; c) G. Morra, R. Potestio, C. Micheletti, G. Colombo, *PLoS Comput. Biol.* **2012**, *8*, e1002433.
[40] E. Moroni, D. A. Agard, G. Colombo, *J. Chem. Theory Comput.* **2018**, *14*, 1033–1044.
[41] a) K. W. East, E. Skeens, J. Y. Cui, H. B. Belato, B. Mitchell, R. Hsu, V. S. Batista, G. Palermo, G. P. Lisi, *Biophys. Rev. Lett.* **2020**, *12*, 155–174; b) P. J. Farber, A. Mittermaier, *Biophys. Rev. Lett.* **2015**, *7*, 191–200; c) S. Grutsch, S. Bruschweiler, M. Tollinger, *PLoS Comput. Biol.* **2016**, *12*, e1004620.
[42] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **1995**, *5*, 67–81.
[43] R. Selvaratnam, S. Chowdhury, B. VanSchouwen, G. Melacini, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6133–6138.
[44] C. McElroy, A. Manfredo, A. Wendt, P. Gollnick, M. Foster, *J. Mol. Biol.* **2002**, *323*, 463–473.
[45] L. Salmon, G. Nodet, V. Ozenne, G. Yin, M. R. Jensen, M. Zweckstetter, M. Blackledge, *J. Am. Chem. Soc.* **2010**, *132*, 8407–8418.
[46] M. Wells, H. Tidow, T. J. Rutherford, P. Markwick, M. R. Jensen, E. Mylonas, D. I. Svergun, M. Blackledge, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5762–5767.
[47] E. Walinda, D. Morimoto, K. Sugase, *Curr. Protoc. Protein Sci.* **2018**, *92*, e57.
[48] P. J. Farber, J. Slager, A. K. Mittermaier, *J. Phys. Chem. B* **2012**, *116*, 10317–10329.
[49] S. Bruschweiler, P. Schanda, K. Kloiber, B. Brutscher, G. Kontaxis, R. Konrat, M. Tollinger, *J. Am. Chem. Soc.* **2009**, *131*, 3063–3068.
[50] S. Bertuzzi, A. Gimeno, R. Núñez-Franco, G. Bernardo-Seisdedos, S. Delgado, G. Jiménez-Osés, O. Millet, J. Jiménez-Barbero, A. Ardá, *Chemistry – A European Journal* **2020**, *26*, 15643–15653.
[51] a) M. Sharon, A. Horovitz, *Curr. Opin. Struct. Biol.* **2015**, *34*, 7–16; b) R. Gruber, A. Horovitz, *Philos. Trans. R. Soc. London Ser. B* **2018**, *373*.
[52] A. Dyachenko, R. Gruber, L. Shimon, A. Horovitz, M. Sharon, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7235–7239.
[53] H. Hernández, C. V. Robinson, *Nat. Protoc.* **2007**, *2*, 715–726.
[54] a) L. Shimon, M. Sharon, A. Horovitz, *Biophys. J.* **2010**, *99*, 1645–1649; b) T. Daubenfeld, A. P. Bouin, G. van der Rest, *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1239–1248.
[55] F. Lanucara, S. W. Holman, C. J. Gray, C. E. Eyers, *Nat. Chem.* **2014**, *6*, 281–294.
[56] S. Badireddy, G. Yunfeng, M. Ritchie, P. Akamine, J. Wu, C. W. Kim, S. S. Taylor, L. Qingsong, K. Swaminathan, G. S. Anand, *Mol. Cell. Proteomics* **2011**, *10*, M110 004390.
[57] J. R. Schnell, H. J. Dyson, P. E. Wright, *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 119–140.
[58] M. Goldstein, N. M. Goodey, in *Allostery: Methods and Protocols* (Eds.: L. Di Paola, A. Giuliani), Springer US, New York, NY, **2021**, pp. 185–219.
[59] J. Chen, R. I. Dima, D. Thirumalai, *J. Mol. Biol.* **2007**, *374*, 250–266.
[60] D. A. Keedy, *Acta Crystallogr. Sect. D* **2019**, *75*, 123–137.
[61] a) P. T. Lang, H. L. Ng, J. S. Fraser, J. E. Corn, N. Echols, M. Sales, J. M. Holton, T. Alber, *Protein Sci.* **2010**, *19*, 1420–1431; b) G. C. P. van Zundert, B. M. Hudson, S. H. P. de Oliveira, D. A. Keedy, R. Fonseca, A. Heliou, P. Suresh, K. Borrelli, T. Day, J. S. Fraser, H. van den Bedem, *J. Med. Chem.* **2018**, *61*, 11183–11198; c) N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane, F. von Delft, *Nat. Commun.* **2017**, *8*, 15123.
[62] D. Ringe, G. A. Petsko, *Biophys. Chem.* **2003**, *105*, 667–680.
[63] D. A. Keedy, H. van den Bedem, D. A. Sivak, G. A. Petsko, D. Ringe, M. A. Wilson, J. S. Fraser, *Structure* **2014**, *22*, 899–910.
[64] H. van den Bedem, A. Dhanik, J. C. Latombe, A. M. Deacon, *Acta Crystallogr. Sect. D* **2009**, *65*, 1107–1117.
[65] D. A. Keedy, Z. B. Hill, J. T. Biel, E. Kang, T. J. Rettenmaier, J. Brandao-Neto, N. M. Pearce, F. von Delft, J. A. Wells, J. S. Fraser, *eLife* **2018**, *7*.
[66] C. Farley, G. Burks, T. Siegert, D. H. Juers, *Acta Crystallogr. Sect. D* **2014**, *70*, 2111–2124.
[67] M. Warkentin, J. B. Hopkins, R. Badeau, A. M. Mulichak, L. J. Keefe, R. E. Thorne, *J. Synchrotron Radiat.* **2013**, *20*, 7–13.
[68] J. L. Thomaston, R. A. Woldeyes, T. Nakane, A. Yamashita, T. Tanaka, K. Koiwai, A. S. Brewster, B. A. Barad, Y. Chen, T. Lemmin, M. Uervirojnangkoorn, T. Arima, J. Kobayashi, T. Masuda, M. Suzuki, M. Sugahara, N. K. Sauter, R. Tanaka, O. Nureki, K. Tono, Y. Joti, E. Nango, S. Iwata, F. Yumoto, J. S. Fraser, W. F. De Grado, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13357–13362.
[69] K. S. Bhullar, N. O. Lagarón, E. M. McGowan, I. Parmar, A. Jha, B. P. Hubbard, H. P. V. Rupasinghe, *Mol. Cancer* **2018**, *17*, 48.
[70] a) X. Lu, J. B. Smaill, K. Ding, *Angew. Chem. Int. Ed. Engl.* **2020**, *59*, 13764–13776; b) R. Roskoski, *Pharmacol. Res.* **2016**, *103*, 26–48; c) L. K. Gavrin, E. Saiah, *MedChemComm* **2013**, *4*, 41–51; d) R. Roskoski, Jr., *Pharmacol. Res.* **2020**, *152*, 104609.
[71] C. J. M. Wright, P. L. McCormack, *Drugs* **2013**, *73*, 1245–1254.
[72] K. P. Garnock-Jones, *Drugs* **2015**, *75*, 1823–1830.
[73] A. A. N. Rose, *Drugs Today (Barc)* **2019**, *55*, 247–264.
[74] T. Yoshida, J. Kakegawa, T. Yamaguchi, Y. Hantani, N. Okajima, T. Sakai, Y. Watanabe, M. Nakamura, *Oncotarget* **2012**, *3*, 1533–1545.
[75] Z. M. Khan, A. M. Real, W. M. Marsiglia, A. Chow, M. E. Duffy, J. R. Yerabolu, A. P. Scopton, A. C. Dar, *Nature* **2020**, *588*, 509–514.
[76] D. T. Dudley, L. Pang, S. J. Decker, A. J. Bridges, A. R. Saltiel, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7686.
[77] K. D. Rice, N. Aay, N. K. Anand, C. M. Blazey, O. J. Bowles, J. Bussenius, S. Costanzo, J. K. Curtis, S. C. Defina, L. Dubenko, S. Engst, A. A. Joshi, A. R. Kennedy, A. I. Kim, E. S. Koltun, J. C. Lougheed, J.-C. L. Manalo, J.-F. Martini, J. M. Nuss, C. J. Peto, T. H. Tsang, P. Yu, S. Johnston, *ACS Med. Chem. Lett.* **2012**, *3*, 416–421.
[78] F. H. Schopf, M. M. Biebl, J. Buchner, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 345–360.
[79] M. Marzec, D. Eletto, Y. Argon, *Biochim. Biophys. Acta* **2012**, *1823*, 774–787.
[80] L. A. Lavery, J. R. Partridge, T. A. Ramelot, D. Elnatan, M. A. Kennedy, D. A. Agard, *Mol. Cell* **2014**, *53*, 330–343.
[81] M. M. U. Ali, S. M. Roe, C. K. Vaughan, P. Meyer, B. Panaretou, P. W. Piper, C. Prodromou, L. H. Pearl, *Nature* **2006**, *440*, 1013–1017.
[82] a) J. Travers, S. Sharp, P. Workman, *Drug Discovery Today* **2012**, *17*, 242–252; b) L. Neckers, P. Workman, *Clin. Cancer Res.* **2012**, *18*, 64.
[83] T. Kijima, T. L. Prince, M. L. Tigue, K. H. Yim, H. Schwartz, K. Beebe, S. Lee, M. A. Budzynski, H. Williams, J. B. Trepel, L. Sistonen, S. Calderwood, L. Neckers, *Sci. Rep.* **2018**, *8*, 6976.
[84] M. G. Marcu, A. Chadli, I. Bouhouche, M. Catelli, L. M. Neckers, *J. Biol. Chem.* **2000**, *275*, 37181–37186.
[85] M. G. Marcu, T. W. Schulte, L. Neckers, *J. Natl. Cancer Inst.* **2000**, *92*, 242–248.
[86] a) H. Zhao, E. Moroni, B. Yan, G. Colombo, B. S. Blagg, *ACS Med. Chem. Lett.* **2012**, *4*, 57–62; b) H. Zhao, E. Moroni, G. Colombo, B. S. J. Blagg, *ACS Med. Chem. Lett.* **2014**, *5*, 84–88; c) E. Moroni, H. Zhao, B. S. J. Blagg, G. Colombo, *J. Chem. Inf. Model.* **2014**, *54*, 195–208.
[87] a) L. Morelli, A. Bernardi, S. Sattin, *Carbohydr. Res.* **2014**, *390 C*, 33–41; b) S. Sattin, J. Tao, G. Vettoretti, E. Moroni, M. Pennati, A. Lopergolo, L. Morelli, A. Bugatti, A. Zuehlke, M. Moses, T. Prince, T. Kijima, K. Beebe, M. Rusnati, L. Neckers, N. Zaffaroni, D. A. Agard, A. Bernardi, G. Colombo, *Chem. Eur. J.* **2015**, *21*, 13598–13608; c) I. D'Annessa, S. Sattin, J. Tao, M. Pennati, C. Sanchez-Martin, E. Moroni, A. Rasola, N. Zaffaroni, D. A. Agard, A. Bernardi, G. Colombo, *Chem. Eur. J.* **2017**, *23*, 5188–5192; d) S. Sattin, M. Panza, F. Vasile, F. Berni, G. Goti, J. H. Tao, E.

Moroni, D. Agard, G. Colombo, A. Bernardi, *Eur. J. Org. Chem.* **2016**, *2016*, 3349–3364.
[88] a) B. K. Zierer, M. Rübbelke, F. Tippel, T. Madl, F. H. Schopf, D. A. Rutz, K. Richter, M. Sattler, J. Buchner, *Nat. Struct. Mol. Biol.* **2016**, *23*, 1020–1028; b) A. Rehn, E. Moroni, B. K. Zierer, F. Tippel, G. Morra, C. John, K. Richter, G. Colombo, J. Buchner, *J. Mol. Biol.* **2016**, *428*, 4559–4571.
[89] C. Sanchez-Martin, E. Moroni, M. Ferraro, C. Laquatra, G. Cannino, I. Masgras, A. Negro, P. Quadrelli, A. Rasola, G. Colombo, *Cell Rep.* **2020**, *31*, 107531.
[90] G. Vettoretti, E. Moroni, S. Sattin, J. Tao, D. A. Agard, A. Bernardi, G. Colombo, *Sci. Rep.* **2016**, *6*, 23830.
[91] M. Bagdany, G. Veit, R. Fukuda, R. G. Avramescu, T. Okiyoneda, I. Baaklini, J. Singh, G. Sovak, H. Xu, P. M. Apaja, S. Sattin, L. K. Beitel, A. Roldan, G. Colombo, W. Balch, J. C. Young, G. L. Lukacs, *Nat. Commun.* **2017**, *8*, 398.
[92] A. Rasola, L. Neckers, D. Picard, *Trends Cell Biol.* **2014**, *24*, 455–463.
[93] J. R. Partridge, L. A. Lavery, D. Elnatan, N. Naber, R. Cooke, D. A. Agard, *eLife* **2014**, *3*, e03487.
[94] A. Schulze, G. Beliu, D. A. Helmerich, J. Schubert, L. H. Pearl, C. Prodromou, H. Neuweiler, *Nat. Chem. Biol.* **2016**, *12*, 628–635.
[95] D. M. Thal, A. Glukhova, P. M. Sexton, A. Christopoulos, *Nature* **2018**, *559*, 45–53.
[96] J. Meiers, E. Siebs, E. Zahorska, A. Titz, *Curr. Opin. Chem. Biol.* **2019**, *53*, 51–67.
[97] a) S. Sattin, A. Bernardi, *Trends Biotechnol.* **2016**, *34*, 483–495; b) A. E. Clatworthy, E. Pierson, D. T. Hung, *Nat. Chem. Biol.* **2007**, *3*, 541–548.
[98] M. M. Sauer, R. P. Jakob, J. Eras, S. Baday, D. Eris, G. Navarra, S. Berneche, B. Ernst, T. Maier, R. Glockshuber, *Nat. Commun.* **2016**, *7*, 10738.
[99] V. B. Rodriguez, B. A. Kidd, G. Interlandi, V. Tchesnokova, E. V. Sokurenko, W. E. Thomas, *J. Biol. Chem.* **2013**, *288*, 24128–24139.
[100] G. Interlandi, W. E. Thomas, *Proteins* **2016**, *84*, 990–1008.
[101] S. Rabbani, B. Fiege, D. Eris, M. Silbermann, R. P. Jakob, G. Navarra, T. Maier, B. Ernst, *J. Biol. Chem.* **2018**, *293*, 1835–1849.
[102] a) K. Drickamer, M. E. Taylor, *Curr. Opin. Struct. Biol.* **2015**, *34*, 26–34; b) P. Valverde, J. D. Martinez, F. J. Canada, A. Arda, J. Jimenez-Barbero, *ChemBioChem* **2020**, *21*, 2999–3025.
[103] G. D. Brown, J. A. Willment, L. Whitehead, *Nat. Rev. Immunol.* **2018**, *18*, 374–389.
[104] J. Valladeau, O. Ravel, C. Dezutter-Dambuyant, K. Moore, M. Kleijmeer, Y. Liu, V. Duvert-Frances, C. Vincent, D. Schmitt, J. Davoust, C. Caux, S. Lebecque, S. Saeland, *Immunity* **2000**, *12*, 71–81.
[105] T.-L. Hsu, S.-C. Cheng, W.-B. Yang, S.-W. Chin, B.-H. Chen, M.-T. Huang, S.-L. Hsieh, C.-H. Wong, *J. Biol. Chem.* **2009**, *284*, 34479–34489.
[106] L. de Witte, A. Nabatov, M. Pion, D. Fluitsma, M. A. de Jong, T. de Gruijl, V. Piguet, Y. van Kooyk, T. B. Geijtenbeek, *Nat. Med.* **2007**, *13*, 367–371.
[107] J. Hanske, S. Aleksić, M. Ballaschk, M. Jurk, E. Shanina, M. Beerbaum, P. Schmieder, B. G. Keller, C. Rademacher, *J. Am. Chem. Soc.* **2016**, *138*, 12176–12186.
[108] J. Aretz, U. R. Anumala, F. F. Fuchsberger, N. Molavi, N. Ziebart, H. Zhang, M. Nazare, C. Rademacher, *J. Am. Chem. Soc.* **2018**, *140*, 14915–14925.
[109] a) M. Thépaut, C. Guzzi, I. Sutkeviciute, S. Sattin, R. Ribeiro-Viana, N. Varga, E. Chabrol, J. Rojo, A. Bernardi, J. Angulo, P. M. Nieto, F. Fieschi, *J. Am. Chem. Soc.* **2013**, *135*, 2518–2529; b) L. Medve, S. Achilli, J. Guzman-Caldentey, M. Thepaut, L. Senaldi, A. Le Roy, S. Sattin, C. Ebel, C. Vives, S. Martin-Santamaria, A. Bernardi, F. Fieschi, *Chem. Eur. J.* **2019**, *25*, 14659–14668.
[110] P. E. Wright, H. J. Dyson, *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29.
[111] R. van der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright, M. M. Babu, *Chem. Rev.* **2014**, *114*, 6589–6631.
[112] a) J. Li, V. J. Hilser, *Methods Enzymol.* **2018**, *611*, 531–557; b) V. J. Hilser, E. B. Thompson, *Proc Natl Acad Sci USA* **2007**, *104*, 8311–8315.
[113] R. B. Berlow, H. J. Dyson, P. E. Wright, *J. Mol. Biol.* **2018**, *430*, 2309–2320.
[114] J. T. White, J. Li, E. Grasso, J. O. Wrabl, V. J. Hilser, *Philos. Trans. R. Soc. London Ser. B* **2018**, *373*.
[115] J. Li, H. N. Motlagh, C. Chakuroff, E. B. Thompson, V. J. Hilser, *J. Biol. Chem.* **2012**, *287*, 26777–26787.
[116] J. Li, J. T. White, H. Saavedra, J. O. Wrabl, H. N. Motlagh, K. Liu, J. Sowers, T. A. Schroer, E. B. Thompson, V. J. Hilser, *eLife* **2017**, *6*.
[117] B. Vogelstein, D. Lane, A. J. Levine, *Nature* **2000**, *408*, 307–310.
[118] P. Wang, M. Reed, Y. Wang, G. Mayr, J. E. Stenger, M. E. Anderson, J. F. Schwedes, P. Tegtmeyer, *Mol. Cell. Biol.* **1994**, *14*, 5182–5191.
[119] W. Borcherds, F. X. Theillet, A. Katzer, A. Finzel, K. M. Mishall, A. T. Powell, H. Wu, W. Manieri, C. Dieterich, P. Selenko, A. Loewer, G. W. Daughdrill, *Nat. Chem. Biol.* **2014**, *10*, 1000–1002.
[120] P. Muller, J. M. Chan, O. Simoncik, M. Fojta, D. P. Lane, T. Hupp, B. Vojtesek, *Protein Sci.* **2018**, *27*, 523–530.
[121] M. Boysen, R. Kityk, M. P. Mayer, *Mol. Cell* **2019**, *74*, 831–843, e834.
[122] H. G. Saavedra, J. O. Wrabl, J. A. Anderson, J. Li, V. J. Hilser, *Nature* **2018**, *558*, 324–328.

Manuscript received: April 23, 2021
Revised manuscript received: May 31, 2021
Accepted manuscript online: June 1, 2021