Methods 209 (2023) 40–47

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Analysis of coordinated NMR chemical shifts to map allosteric regulatory networks in proteins

Erin Skeens, George P. Lisi *

Department of Molecular Biology, Cell Biology, & Biochemistry, Brown University, Providence, RI 02903, United States

ARTICLE INFO

Keywords:
Allostery
Chemical shifts
Reciprocity

ABSTRACT

The exquisite sensitivity of the NMR chemical shift to local environment makes it an ideal probe to assess atomic level perturbations in proteins of all sizes and structural compositions. Recent advances in solution and solid-state NMR spectroscopy of biomolecules have leveraged the chemical shift to report on short- and long-range couplings between individual amino acids to establish "networks" of residues that form the basis of allosteric pathways that transmit chemical signals through the protein matrix to induce functional responses. The simple premise that thermodynamically and functionally coupled regions of a protein (i.e. active and allosteric sites) should be reciprocally sensitive to structural or dynamic perturbations has enabled NMR spectroscopy, the premier method for molecular resolution of protein structural fluctuations, to occupy a place at the forefront of investigations into protein allostery. Here, we detail several key methods of NMR chemical shift analysis to extract mechanistic information about long-range chemical signaling in a protein, focusing on practical methodological aspects and the circumstances under which a given approach would be relevant. We also detail some of the experimental considerations that should be made when applying these methods to specific protein systems.

1. Introduction

Allostery is a ubiquitous regulatory process that functionally couples spatially distinct sites within a protein. In the past decade, significant efforts have been made to map the routes by which proteins transmit biological signals ("allosteric pathways") due to the enhanced spatio-temporal control afforded by modulation of this crosstalk. [1,2] Concurrently, modern experimental and theoretical investigations of allostery have demonstrated that the ability of apo proteins to adopt active conformations is an intrinsic property of a dynamic equilibrium that is then further biased by allosteric activators or inhibitors. [3–7] In effect, these findings suggest that protein structural fluctuations play a large role in allosteric signal propagation. This interpretation, in some ways, diverges from classical paradigms, necessitating the visualization of subtle structural and dynamic changes that modulate protein ensembles to establish mechanisms and leverage allosteric pathways in drug discovery. [8] Of the modern structural methods available to probe allostery, solution nuclear magnetic resonance (NMR) is best suited to report on simultaneous changes in protein structure and multi-timescale dynamics at the atomic level. [9] The beauty of NMR for this application lies in the simplicity of the readout – the chemical shift, which is highly sensitive to local chemical environment and can report on protein folding, dynamics, and structural populations.

A major advantage of employing NMR chemical shift analysis in studies of allostery is that, at the most basic level, residues that are functionally linked within a protein should be structurally sensitive to each other. That is, allosteric couplings should display signatures of reciprocal perturbation in response to mutations, ligands, or protein activation/inhibition. [10–14] Changes in chemical shift frequency can be detected by eye as a reporter of structural perturbation, and when combined with analyses of NMR resonance intensities (reporters of dynamic change), provide a simple molecular readout for a complex biochemical process. A number of formalized methods have been reported to leverage NMR chemical shift information to elucidate and characterize residues that confer functional control of a protein via long-range signaling. Here, we provide a roadmap to implement and adapt these methods based on the system of interest. We explain the optimal circumstances for use and the practical considerations required for several methods, with a focus on expanding usage of simple NMR readouts for specialists and non-specialists alike.

* Corresponding author.
E-mail address: george_lisi@brown.edu (G.P. Lisi).

https://doi.org/10.1016/j.ymeth.2022.12.002
Received 15 September 2022; Received in revised form 7 December 2022; Accepted 14 December 2022
Available online 17 December 2022

2. Mapping allosteric activation on fast timescales

Proteins often undergo conformational transitions that mediate their interactions and functional output. [15] Long-range communication between distal residues (i.e., allosteric signaling) has been shown to play a pivotal role in controlling conformational equilibria that underlie protein function. [16–18] NMR can detect subtle structural and dynamic changes associated with fast-exchanging conformational equilibria, where chemical shifts are population-weighted averages of the conformational states. [19–20] Thus, in the presence of perturbations that modulate the conformational equilibrium, correlated chemical shifts can be leveraged to identify networks of allosteric residues associated with the functional activation of a protein. [21].

A well-established method for characterizing allosteric networks associated with conformational equilibria, such as the transition from an inactive to an active state, is chemical shift covariance analysis (CHESCA). [21–34] Developed and optimized by Melacini and coworkers, CHESCA relies on a set of ligands or mutations that perturb the conformational equilibrium and collectively sample the inactive-to-active conformational spectrum. [21] The basis of CHESCA is to identify residues that exhibit a concerted, linear response to perturbations via their chemical shifts, indicative of sites within a common allosteric network. The implementation of agglomerative clustering and singular value decomposition detects and functionally classifies clusters of residues with highly correlated chemical shifts.

2.1. Implementing chemical shift covariance analysis (CHESCA)

The use of CHESCA to search for novel allosteric sites coupled to activation is optimal for systems with a two-state conformational equilibrium that occurs in the fast-exchange regime ($k_{ex} \gg \Delta\omega$). [21] Additionally, a set of perturbations (≥5 perturbations recommended) must be previously established to sample the conformational equilibrium of activation, including the inactive, intermediate, and active states. These perturbations can include ligand effectors or mutations, as has been demonstrated previously. [21,31,35] Ideally, the perturbations should be spatially localized within a well-defined region of the structure to minimize the perturbation-induced effects to nearby residues, which can introduce nearest-neighbor 'noise' into the analysis and cloud the identification of allosterically-relevant chemical shifts. Melacini and coworkers have reported several iterations of CHESCA to reduce false positives and negatives, and to identify core allosteric residues for further study. [21–22,32] Herein, we highlight complete-linkage CHESCA (CL-CHESCA) as an effective starting point for identifying networks of allosteric residues. [22].

CL-CHESCA begins with the collection of ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectra for all perturbation conditions (fully saturated if using ligands; Fig. 1A). [21–22] Once the ^{1}H and ^{15}N chemical shifts have been assigned for all possible residues, the subsequent analyses can be completed and visualized via the CHESCA-SPARKY plugin in NMRFAM-SPARKY. [28] CHESCA-SPARKY provides a streamlined start-to-finish workflow for all of the analyses that comprise CL-CHESCA, and the semi-automated process simplifies the analyses for researchers who may be unfamiliar with the clustering and dimensionality reduction techniques described below. ^{1}H and ^{15}N chemical shifts (δ_H and δ_N, respectively) for every residue are calculated as a combined chemical shift by $\delta_{NH} = \delta_H + 0.2\,\delta_N$ for the set of perturbations. Pearson correlation coefficients are then computed for all residue pairs X and Y via $= \frac{\sum_{i=1}^{n}(\delta_{Xi}-\bar{\delta}_X)(\delta_{Yi}-\bar{\delta}_Y)}{\sqrt{\sum_{i=1}^{n}(\delta_{Xi}-\bar{\delta}_X)^2\sum_{i=1}^{n}(\delta_{Yi}-\bar{\delta}_Y)^2}}$, where δ_{Xi} and δ_{Yi} are the combined chemical shifts (δ_{NH}) from the set of perturbation conditions for residues X and Y, respectively, $\bar{\delta}_X$ and $\bar{\delta}_Y$ are the averages of δ_X and δ_Y, and n is the number of perturbation conditions tested (Fig. 1A). A chemical shift covariance matrix can subsequently be compiled for residue pairs to highlight residues and regions of the protein with tightly coupled chemical shifts (Fig. 1B).

To identify clusters of residues with correlated chemical shifts, CL-CHESCA employs complete-linkage agglomerative clustering (AC), a hierarchical clustering method. [36] The basis for complete-linkage AC is that residues within a cluster must be correlated with all other residues above a given correlation threshold ($|r| \geq 0.98$ recommended). The resulting clusters represent networks of residues with highly correlated chemical shifts for a set of perturbations. When the chosen perturbations are ligand effectors, it is important to differentiate between clusters of correlated chemical shifts derived from a common allosteric role to those that are due to ligand binding effects. To accomplish this functional delineation, singular value decomposition (SVD) is implemented. [37] SVD can be used as a dimensionality reduction technique to identify principal components from the chemical shift data. The application of SVD to NMR chemical shifts relies on the chemical shifts being calculated relative to a reference state. Given that the chosen ligand effectors modulate the conformational equilibrium that controls activation, and an antagonist will bind the protein without activating it, selecting an antagonist as the reference state will differentiate binding effects (apo vs antagonist-bound) from allosteric activation (agonist-bound vs antagonist-bound). The resulting principal components identified by SVD, and subsequently the AC-derived clusters, can then be assigned a function through analysis of "loading" and "score" plots for the principal components that account for the majority of the total variance (Fig. 1C). The loading plot will indicate which perturbation conditions are the major contributors to the principal components. For example, if the apo-antagonist chemical shifts significantly contribute to

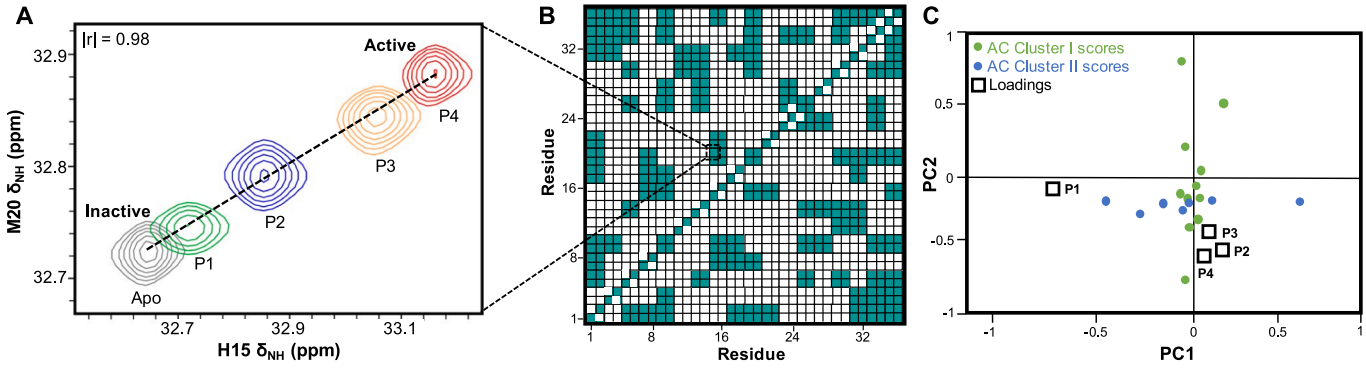

Fig. 1. Chemical shift covariance analysis (CHESCA). **(A)** Chemical shift correlation for residues from example data consisting of the apo protein and four perturbation conditions (P1–P4) that sample the inactive-to-active conformational equilibrium. **(B)** Chemical shift covariance matrix derived from pairwise chemical shift correlations, where residue pairs with $|r| \geq 0.98$ are highlighted in teal. **(C)** Loading and score plots for PC1 and PC2 from SVD analysis for sample data. Blue and green circles denote scores for residues in AC-derived clusters I and II, respectively, and black squares represent loadings (perturbations). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

PC1, then PC1 indicates binding effects. Conversely, if the agonist–antagonist chemical shifts are the major contributor to PC2, then PC2 quantifies allosteric activation. The score plot is comprised of scores calculated for every residue that inform the contribution that each residue has on the principal components and their respective functions. In conjunction with the clusters identified through agglomerative clustering, networks of residues can be functionally assigned as contributing to allosteric activation or binding effects depending on the principal component to which the clusters align.

2.2. Experimental considerations for CHESCA

It is recommended that the perturbation-induced chemical shifts for a given residue have a frequency spread of at least 10 Hz for ^{1}H and 5 Hz for ^{15}N to account for any systematic errors due to referencing or nonspecific binding issues; if it does not meet either of these cutoffs, the residue should be discarded from future analysis. It should also be noted that residues proximal to the sites of perturbation, including ligand binding sites and sites of mutation, should similarly be excluded from analysis due to localized effects caused by the perturbations. The recommended correlation threshold ($|r| \geq 0.98$) can be modified given the system of interest, though several thresholds should be tested to ensure that poorly correlated residue pairs that may give rise to false positives are not included in subsequent analyses. The implementation of a suite of CHESCA-based methods, such as *T*-CHESCA and CLASS-CHESCA, can be helpful for identifying key allosteric residues while reducing the inclusion of false positives and negatives. [32] CHESCA can also be applied to a three-state conformational equilibrium as detailed by Selvaratnam et al. [21].

It should be noted that the CHESCA-SPARKY plugin is written for the analysis of ^{1}H-^{15}N HSQC spectra. However, the plugin allows for modification of the combined chemical shift scaling factor ("*N*-to-H" parameter, defined here as 0.2 for ^{15}N shifts in the calculation of δ_{NH}) and cutoffs that can be tailored to ^{13}C chemical shifts, as has been demonstrated previously for the analysis of methyl spectra. [38–39] Alternatively, complete-linkage AC, AC-derived dendrograms, and SVD can be performed and generated via Cluster 3.0, [40] JAVA TreeView, [41] and Octave, [42] respectively. [22].

In addition to chemical shifts, CHESCA can also be broadly applied to other NMR observables given they satisfy-three important criteria: (1) the observables are linear averages, (2) the linear coefficients are perturbation-dependent, and (3) the observables can be measured at high resolution on a per-residue basis. For instance, a CHESCA-like analysis has been applied to transverse relaxation and dark-state exchange saturation transfer (DEST) rates to identify correlative trends. [43–44].

3. Quantifying shifts in allosteric activation

Conformational transitions that underlie the functional activation of proteins are allosterically-driven processes that can be modulated by ligands or mutations through the alteration of long-range signaling. In addition to using chemical shift correlations to detect networks of allosteric residues associated with conformational equilibria via CHESCA, as detailed in section 2, chemical shifts can also report on the direction and extent to which perturbations shift allosteric conformational equilibria to impact activation. [45] These characterizations are important for understanding the impacts of ligand- and mutation-induced perturbations on function, and provide insight into the allosteric mechanisms that control conformational activation.

Quantifying shifts in conformational equilibria can be achieved through chemical shift projection analysis (CHESPA), a method formalized by Melacini and coworkers and demonstrated on a number of different systems. [23,27,28,31,34,38,45–47] CHESPA involves the analysis of chemical shifts from three states, including inactive and active states as references, and the perturbation state being interrogated. For a two-state conformational equilibrium in the fast-exchange regime, the magnitude and angle of the chemical shifts observed for the perturbation state relative to the reference states provides a measure of activation or inactivation induced by the perturbation. CHESPA can be implemented to identify sets of ligand effectors or mutations that alter conformational equilibria for analysis of allosteric activation via CHESCA, and detect false positives in CHESCA-derived clusters. CHESPA can also identify deviations from a two-state equilibrium, thus providing insight into new conformational states.

3.1. Implementing chemical shift projection analysis (CHESPA)

CHESPA is ideal for allosteric systems that populate multiple conformations related to activation. [45] Specifically, the analysis assumes a two-state conformational equilibrium in the fast-exchange regime. Experimental conditions that induce the inactive and active conformations must be identified through structural and functional analyses for use as reference states. These reference conditions, as well as the perturbation of interest, can include a combination of mutations and ligands, such as small molecule effectors, interacting proteins, and nucleic acids. A CHESPA-SPARKY plugin is available on NMRFAM-SPARKY to facilitate the analysis of the activation- and perturbation-induced chemical shifts as described below. [28].

Chemical shift projection analysis requires three ^{1}H-^{15}N HSQC spectra comprising the following states: a conformationally inactive reference state, such as the apo wild-type protein, an active reference state, such as when bound to an endogenous ligand, and the perturbation state of interest, such as in the presence of a mutation or ligand effector. The three spectra are then overlaid for analysis of resonances with observed chemical shifts (Fig. 2A). Compounded chemical shifts are calculated as $\Delta\delta = \sqrt{(\Delta\delta_H)^2 + (0.2\Delta\delta_N)^2}$ between the inactive and active states, defining the reference/activation vector ($\vec{R}$), and between the perturbation state and the inactive or active states, defining the perturbation vector ($\vec{P}$). The perturbation vector originating from the inactive state will provide a measure of activation, while conversely, the perturbation vector originating from the active state will provide a measure of inactivation.

Once the vectors have been defined, two important parameters can be calculated to characterize the impact of the perturbation on conformational activation (Fig. 2). Projection of the perturbation vector onto the reference vector quantifies the magnitude of the shift in the conformational equilibrium (i.e., the fractional shift). The fractional shift (X) is calculated as $X = \frac{|\vec{P}|}{|\vec{R}|}\cos(\theta)$, where θ is the angle between the vectors (Fig. 2A and 2B). Cos(θ) informs on the linearity of the reference and perturbation vectors, where vectors for residues that follow a two-state conformational equilibrium are expected to be co-linear such that $|\cos(\theta)| \cong 1$. Deviations from linearity may be due to localized effects caused by proximity to the perturbation site unrelated to conformational activation, or due to altered dynamics induced by the perturbation. Cos(θ), defined as $\cos(\theta) = \frac{\vec{P}\times\vec{R}}{|\vec{P}|\times|\vec{R}|}$, therefore determines the direction of the shift in the conformational equilibrium (Fig. 2A and 2C). When the perturbation vector originates from the inactive state, a positive fractional shift is observed for perturbations that shift the equilibrium to the active state while a negative fractional shift is observed when the equilibrium is shifted towards the inactive state. The reverse is true for a perturbation vector originating from the active state. Residues that exhibit shifts related to a common conformational transition should have similar fractional shift and cos(θ) values. Identification of residue clusters with distinct fractional shift values can indicate the sampling of a third conformational state caused by the perturbation. CHESPA therefore provides information on residues that are allosterically coupled to activation, as well as per-residue effects of allosteric

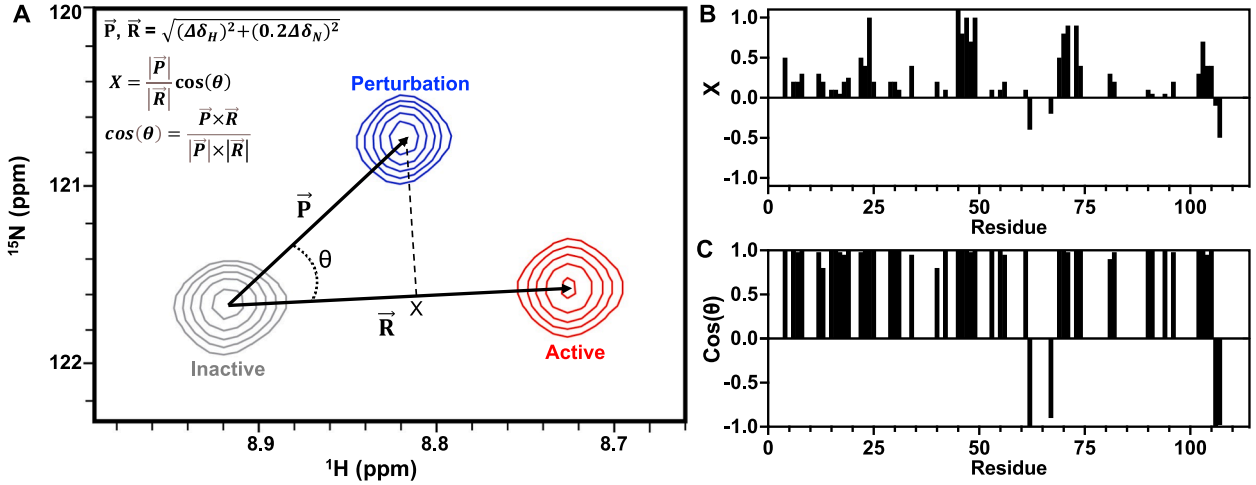

Fig. 2. Chemical shift projection analysis (CHESPA) for quantifying conformational activation of a perturbation state. **(A)** A sample resonance detailing the parameters necessary for the analysis of chemical shifts related to conformational activation. $\vec{R}$ and $\vec{P}$ represent the reference and perturbation vectors between the inactive-to-active state chemical shifts and the inactive-to-perturbation state chemical shifts, respectively. X is the fractional shift of $\vec{P}$ projected onto $\vec{R}$, providing a measure of the perturbation-induced shift in the conformational equilibrium. θ is the angle between $\vec{R}$ and $\vec{P}$, where cos(θ) determines the direction of the shift in the conformational equilibrium. Per-residue plots of X and cos(θ) for the reference and perturbation vectors are shown in **(B)** and **(C)**. Here, consistent with the analysis shown in **(A)**, the positive X values indicate that the perturbation causes a shift in the conformational equilibrium toward the active state.

perturbations.

3.2. Experimental considerations for CHESPA

To account for experimental errors, a chemical shift minimum should be established to ensure that the projection analyses are completed on shifts that reliably inform on conformational changes at a given residue. A minimum of 0.05 ppm is recommended for chemical shifts calculated for the reference and perturbation vectors (|R|, |P| greater than 0.05 ppm). [45] For the assessment of linearity, a threshold should be established to exclude residues with cos(θ) values that deviate from that expected for a two-state conformational equilibrium ($|\cos(\theta)| \cong 1$). Melacini and coworkers recommend a threshold of $|\cos(\theta)| \geq 0.90–0.95$. However, these thresholds can be tailored to the system of interest through testing of different cutoff values to limit false positives and negatives due to a variety of system-specific factors. As in CHESCA-SPARKY, the CHESPA-SPARKY plugin is written for ^{1}H-^{15}N chemical shifts, though the scaling factor ("*N*-to-H") for calculating compounded chemical shifts ($\Delta\delta$) can be modified by the user to analyze other nuclei of interest. Alternatively, $\Delta\delta$, cos(θ), and X can be calculated using the equations detailed in section 2.1 and Fig. 2A, with the scaling factor modified for $\Delta\delta$.

4. Mapping allosteric coupling on slow timescales

NMR is an established tool for detecting protein–ligand interactions, where chemical shifts are used to map ligand binding sites and quantify binding affinities. [48–49] Given its sensitivity, NMR can also detect residues distal to a binding site that exhibit ligand-induced chemical shift perturbations via long-range allosteric communication. [50] In some cases, protein function is regulated by the allosteric coupling of two ligand binding events, where the binding of one ligand promotes or precludes binding of a second ligand to spatially and temporally control downstream responses. [51] Elucidation of the residues responsible for propagating the allosteric signals between ligand binding sites is critical for dissecting the mechanisms of distinct, yet coupled, binding events. These allosteric sites can further be leveraged for the design of ligand effectors for experimental and therapeutic purposes.

McDermott and coworkers recently reported a method for identifying allosteric residues involved in the coupling of two ligand binding sites, called chemical shift detection for allostery participants (CAP). [51] CAP relies on systems where conformational changes associated with binding are in slow exchange ($k_{ex} \ll \Delta\omega$), allowing for visualization of all possible binding states, including apo, ligand A bound, ligand B bound, and ligands A and B bound, via unique chemical shift behavior (Fig. 3A). The goal of CAP is to identify residues that exhibit distinct chemical shift changes when both ligands are bound that are not observed for the apo protein or when one ligand is bound. This is a strong indicator that the residues are coupled to a conformational change that is dependent on the occupation of the two ligand binding sites and likely plays a role in mediating the allosteric cooperativity between the binding event (i.e., allostery participants). Chemical shift correlations of population changes are then used to identify networks of residues that are allosterically coupled to a common binding event.

4.1. Implementing chemical shift detection of allostery participants (CAP)

Chemical shift detection of allostery via CAP is ideal for systems with two distal binding sites occupied by distinct ligands that are known to be functionally and allosterically coupled. Additionally, the conformational changes associated with ligand binding should occur via slow exchange to allow for detection and quantitation of chemical shifts that report on the bound and unbound states, where changes in resonance intensity or volume (i.e., the population of each state) are indicative of ligand-dependent conformational states. [51].

The first phase of CAP involves determining which residues respond to the binding of one ligand, two ligands, or are insensitive to either ligand. To begin, four two-dimensional NMR spectra consisting of the apo protein, ligand A-bound, ligand B-bound, and ligands A- and B-bound are collected under limiting ligand concentrations, as necessary, and resonances are assigned. CAP assumes that when a residue is sensitive to a conformational change associated with a single ligand binding event, there will be two chemical shifts observed for that residue in slow exchange under limiting conditions, which represent the bound and unbound states. As ligand concentrations are increased, the resonance corresponding to the ligand-bound state will increase in intensity while the opposite trend is observed for the unbound resonance, denoting a population change dependent on ligand concentration. Therefore, peak intensities serve as a reporter of ligand occupancy. These residues can be readily identified through comparative analysis of the apo and ligand-

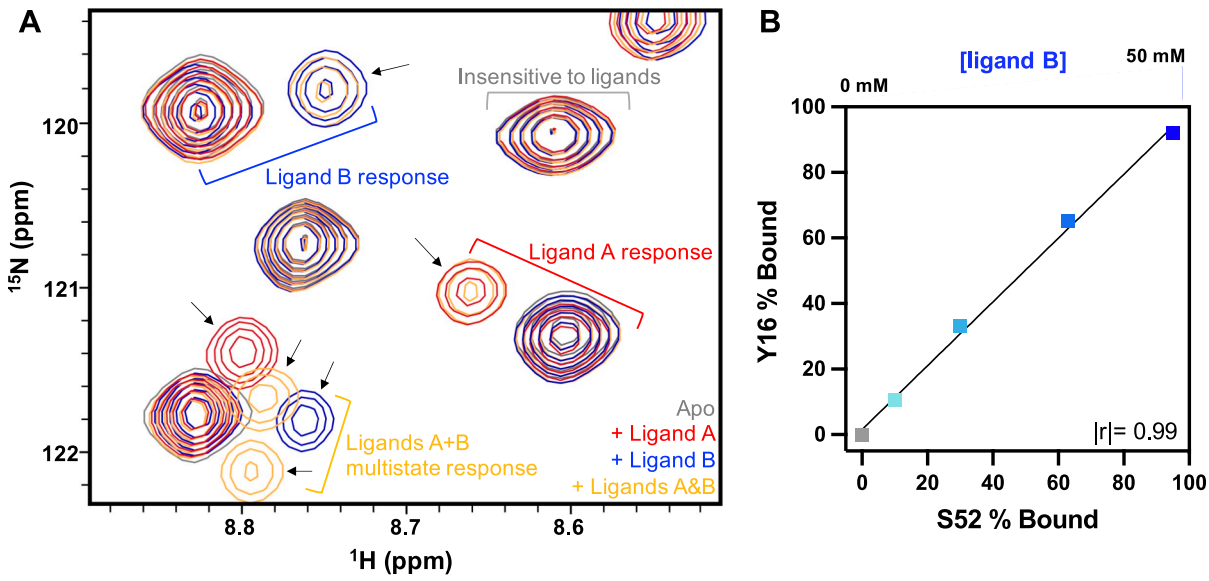

Fig. 3. Chemical shift detection of allostery participants (CAP). **(A)** A ^{1}H-^{15}N spectral overlay of selected resonances from a test protein in the apo (gray), ligand A-bound (red), ligand B-bound (blue), and ligands A- and B-bound (yellow) states. Slow-exchanging minor peaks (i.e., ligand bound conformation) are indicated with an arrow. The spectral overlay highlights residues insensitive to the ligands, residues coupled to ligand A or ligand B binding events, and residues that are allosteric participants in the coupled ligand A/ligand B binding events and exhibit a multistate response only when ligands A and B are present. **(B)** Chemical shift correlation for two residues, where the percent bound is determined by quantifying population shifts over a ligand titration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bound spectra (Fig. 3A). Residues distal from the binding sites that exhibit ligand-induced chemical shifts are suggested to play an allosteric role in the respective binding event. Conversely, when two ligands are bound, residues that are coupled to both binding events will exhibit chemical shift behaviors that diverge from what is observed when a single ligand is bound. Xu et al. demonstrated that these multistate chemical shift behaviors can present as more than two chemical shifts, indicative of a distinct conformation that requires both ligands to be bound, or NMR line broadening, revealing altered dynamics that may contribute to the propagation of signals between binding sites. Residues exhibiting multistate behaviors in the presence of two ligands are likely participants in the allosteric coupling of the binding events. Collectively, these analyses will identify residues that are allosterically responsive to one or two ligand binding events.

The second phase of CAP is identifying residues that exhibit correlated chemical shift perturbations in the presence of one or two ligands, suggestive of a coordinated allosteric role in the binding events. Although conceptually similar to the covariance analysis of CHESCA, CAP leverages population changes of the slow exchange chemical shifts to elucidate correlated responses to ligand binding. To accomplish this, spectra are collected at varying ligand concentrations, where population changes of the chemical shifts identified in the first phase of CAP are quantified through peak integration and denote the percentage of ligand bound (Fig. 3B). Pearson correlation coefficients can then be determined for the population shifts of residue pairs X and Y by $r = \frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}$, where X_i and Y_i represent population percentages of the ligand-bound state of residues X and Y, respectively, for each ligand concentration tested, $\overline{X}$ and $\overline{Y}$ are the average population percentages, and n is the number of ligand concentrations tested. Analysis of the covariance matrices of population shifts in the presence of one or two ligands highlights residue pairs that exhibit a concerted response to ligand binding through correlated population changes established across varied ligand concentrations (Fig. 3B). Importantly, resonances that exhibit correlated population changes in the presence of two ligands can pinpoint residues that may contribute to the route of allosteric communication between binding sites to regulate function.

4.2. Experimental considerations for CAP

For some systems with coupled binding events, the binding of one ligand may promote or attenuate binding of the other ligand. Consequently, residues that are allosterically coupled to one ligand binding event may appear to exhibit a response to the other ligand when ligand-induced binding or release occurs, thus clouding efforts to identify residues that participate in the allosteric process. Therefore, it may be necessary to identify conditions where ligand binding can be more tightly controlled to avoid such ambiguity, such as modulating pH, ligand concentration, etc. [51] If the ligand-bound conformations of a system with coupled binding events are occurring in fast exchange, such that peaks indicative of both conformations are not visible for quantifying population changes, the CAP method is still relevant to elucidate residues involved in a single binding event versus two binding events. However, the principles of CAP must be applied in conjunction with a method that relies on fast exchanging conformations in the analysis, such as CHESCA. [21].

5. Allosteric reciprocity

Since allostery describes the exchange of chemical information between spatially distinct sites, an inherent property of allostery is the reciprocity of signals that propagate from one site to another. This bidirectional transfer of information (i.e. crosstalk) implies that a perturbation at a one site will impact the other. Thus, the reciprocal nature of allostery can be leveraged to experimentally elucidate novel allosteric sites, employing the sensitivity of NMR to detect bidirectional chemical shift changes in response to perturbations. [52].

As demonstrated by Loria and coworkers, the reciprocity of allosteric signaling can be experimentally tested through mutational analysis of a known active site (Fig. 4). [52] When residues at the active site are mutated, chemical shift perturbations observed for residues distal to the active site suggests the existence of allosteric coupling. Therefore, if allosteric signals are being propagated between them, mutations made at the distal residues should now result in chemical shift perturbations at the active site, thus exhibiting a bidirectional response. This method can be extended to other functionally relevant sites in a protein as well, such

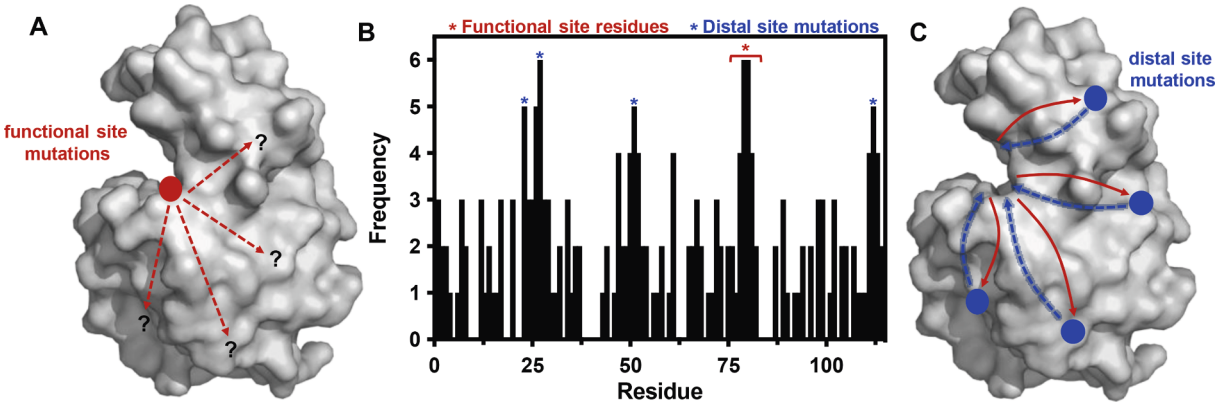

Fig. 4. Reciprocal chemical shift perturbations via mutational analysis. **(A)** Single-point mutations of residues that comprise a functional site are used to detect residues that may be allosterically coupled to the functional site through analysis of NMR chemical shift perturbations. **(B)** The frequency at which a given residue exhibits a chemical shift perturbation above a significance threshold is plotted on a histogram to identify residues distal from the functional site that are consistently perturbed across the set of mutations. **(C)** Distal site residues identified in (B) are then mutated to determine if chemical shift perturbations are reciprocally observed at the functional site, thus leveraging the inherent reciprocity of allosteric communication between spatially distinct sites.

as a generic ligand binding site.

5.1. *Implementing reciprocal chemical shift perturbations via mutational analysis*

Identification of allosteric sites through analysis of reciprocal chemical shift perturbations is ideal for proteins with one known functional site, such as an enzyme active site or ligand binding site. An advantage of this method is that it requires very little prior knowledge of the system beyond the residues that compose the functional site. However, a biochemical assay for the site in question is useful for subsequent characterizations of the newly identified allosteric residues to understand their role in modulating protein function.

The method requires engineering single-point mutations at some, or all, residues that comprise the functional (active) site (Fig. 4A). ^{1}H-^{15}N HSQC spectra are then collected on the wild-type protein and the set of functional site variants. Chemical shift perturbations are calculated from composite chemical shifts via $\Delta\delta = \sqrt{\frac{1}{2}(\delta_H^2 + \frac{\delta_N^2}{25})}$, where δ_H and δ_N are changes in the amide proton and nitrogen chemical shifts, respectively, relative to the wild-type spectrum. A significance cutoff, for example two standard deviations (2σ) above the mean of the data, is used to identify residues that exhibit the largest chemical shift perturbations due to the mutations. The frequency at which a residue is deemed significantly perturbed for the set of functional site mutations can then be quantified and visualized in a histogram (Fig. 4B). Residues that frequently exhibit significant chemical shift perturbations due to mutations, but are distal to the functional site, likely experience long-range crosstalk and thus become candidates for further analysis.

To determine whether the communication between the functional and distal sites follows the inherent reciprocity that defines allostery, the distal residues identified by the frequency histogram are similarly mutated (Fig. 4C). HSQC spectra are then collected, with composite chemical shift perturbations calculated as described above. Analysis of chemical shift behavior is now focused on residues at the functional site. Distal residues that cause reciprocal chemical shift perturbations at the functional site confirm a bidirectional response to perturbation at the spatially distinct sites. The distal site variants can then be assessed through biochemical assays to determine the functional impact of the potential allosteric site.

5.2. *Experimental considerations for allosteric reciprocity*

The set of functional site mutations tested should be large enough to elucidate residues that exhibit the strongest coupling to the functional site based on a high frequency of observed chemical shift perturbations. For instance, Cui et al. tested 13 active site mutations, though the number is dependent on the system, and functional site, being studied. [52] While not explicitly confined to any exchange regime, this method works best in the fast chemical-exchange regime, with defined changes in chemical shifts quantified for the analysis. However, line broadening or the presence of slow-exchanging major and minor peaks could be taken into consideration as an observable change in chemical shift behavior, indicative of altered structural or dynamic properties of a given residue caused by mutations at the functional site. The cutoff for determining significant chemical shifts to be included in the frequency histogram can also be modified given the system.

6. General experimental considerations for allosteric characterizations via chemical shift analysis

An important consideration for the study of allostery by NMR is the molecular weight of the system of interest. As a general metric, proteins with molecular weights greater than 25 kDa tend to exhibit diminished signal and spectral crowding that can complicate or preclude analysis. There are several techniques to mitigate these issues for larger systems, including the use of high and ultra-high field magnets, deuterium incorporation, transverse relaxation optimized spectroscopy (TROSY) pulse sequences, sparse isotopic labeling, and high-resolution magic angle spinning (HR-MAS) NMR. [53] However, some systems may remain intractable to NMR studies due to molecular weight limitations. Additionally, a chemical shift standard, such as sodium trimethylsilylpropanesulfonate (DSS) or ^{15}N-acetylglycine, should be used to ensure the accurate and consistent measurement of chemical shifts for all samples. [21] For systems where mutations are used in the analysis to probe conformational activation or allosteric reciprocity, initial assessments of protein stability and conformation should be conducted to limit significant mutation-induced changes to structure and dynamics that may cloud comparative analyses of chemical shifts.

The methods described above rely on straightforward two-dimensional NMR spectra for analysis of chemical shifts. Though many focus on the use of ^{1}H-^{15}N HSQC spectra, other two-dimensional NMR methods can be similarly applied as well, with combined and composite chemical shift equations modified to the nuclei of interest as necessary. For example, implementation of CHESPA and CHESCA with ^{1}H-^{13}CH$_3$ heteronuclear multiple quantum coherence (HMQC) spectra has been demonstrated by Liptak et al. and Latham and coworkers, respectively, for allosteric characterizations of methyl labeled systems, [35,38] while Xu et al. used ^{13}C–^{13}C dipolar-assisted rotation resonance (DARR) solid-state NMR experiments to study ligand binding of a large

transmembrane protein in proteoliposomes via CAP. [51].

7. Conclusions

The ubiquitous use and inherent simplicity of the NMR chemical shift as a metric of protein structure makes it an ideal probe of allosterically controlled systems. While not discussed in detail here, it should be noted that NMR spin relaxation experiments focusing on changes in resonance intensity, and indirectly reporting on chemical shift differences, are also well-suited for resolving allosteric networks under many of the same conditions described above. [19–20,23,27] Thus, NMR data of either type, in conjunction with crystallographic, biophysical, or molecular dynamics-based studies, are tremendously powerful for the visualization of biological mechanisms driven by conformational equilibria.

Funding

This work was supported by NSF grant MCB 2143760 and NIH grant R01 GM144451 to GPL.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] G.P. Lisi, J.P. Loria, Allostery in enzyme catalysis, Curr. Opin. Struct. Biol. 47 (2017) 123–130.
[2] K.W. East, E. Skeens, J.Y. Cui, H.B. Belato, B. Mitchell, R. Hsu, V.S. Batista, G. Palermo, G.P. Lisi, NMR and computational methods for molecular resolution of allosteric pathways in enzyme complexes, Biophys. Rev. 12 (2020) 155–174.
[3] S. Boulton, G. Melacini, Advances in NMR methods to map allosteric sites: from models to translation, Chem. Rev. 116 (11) (2016) 6267–6304.
[4] D. Ming, M.E. Wall, Allostery in a coarse-grained model of protein dynamics, Phys. Rev. Lett. 95 (19) (2005) 198103–198106, https://doi.org/10.1103/PhysRevLett.95.198103.
[5] A.L. Lee, Contrasting roles of dynamics in protein allostery: NMR and structural studies of CheY and the third PDZ domain from PSD-95, Biophys. Rev. 7 (2015) 217–226.
[6] N. Popovych, S. Sun, R.H. Ebright, C.G. Kalodimos, Dynamically driven protein allostery, Nat. Struct. Mol. Biol. 13 (9) (2006) 831–838, https://doi.org/10.1038/nsmb1132.
[7] H.N. Motlagh, J.O. Wrabl, J. Li, V.J. Hilser, The ensemble nature of allostery, Nature 508 (7496) (2014) 331–339, https://doi.org/10.1038/nature13001.
[8] R. Nussinov, C.J. Tsai, Allostery in disease and in drug discovery, Cell 153 (293–305) (2013) 293–305, https://doi.org/10.1016/j.cell.2013.03.034.
[9] G.P. Lisi, J.P. Loria, Solution NMR spectroscopy for the study of enzyme allostery, Chem. Rev. 116 (11) (2016) 6323–6369, https://doi.org/10.1021/acs.chemrev.5b00541.
[10] J.-P. Changeux, 50 Years of allosteric interactions: the twists and turns of the models, Nat. Rev. Mol. Cell Biol. 14 (12) (2013) 819–829, https://doi.org/10.1038/nrm3695.
[11] J.-P. Changeux, S.J. Edelstein, Allosteric mechanisms of signal transduction, Science 308 (5727) (2005) 1424–1428, https://doi.org/10.1126/science.1108595.
[12] G.P. Lisi, K.W. East, V.S. Batista, J.P. Loria, Altering the allosteric pathway in IGPS suppresses millisecond motions and catalytic activity, PNAS 114 (2017) E3414–E3423.
[13] A. Velyvis, H.K. Schachman, L.E. Kay, Application of methyl-TROSY NMR to test allosteric models describing effects of nucleotide binding to aspartate transcarbamoylase, J. Mol. Biol. 387 (3) (2009) 540–547, https://doi.org/10.1016/j.jmb.2009.01.066.
[14] S.R. Tzeng, C.G. Kalodimos, Dynamic activation of an allosteric regulatory protein, Nature 462 (7271) (2009) 368–372, https://doi.org/10.1038/nature08560.
[15] B.J. Grant, A.A. Gorfe, J.A. McCammon, Large conformational changes in proteins: signaling and other functions, Curr. Opin. Struct. Biol. 20 (2) (2010) 142–147, https://doi.org/10.1016/j.sbi.2009.12.004. From NLM Medline.
[16] J. Guo, H.X. Zhou, Protein allostery and conformational dynamics, Chem. Rev. 116 (11) (2016) 6503–6515, https://doi.org/10.1021/acs.chemrev.5b00590. From NLM Medline.
[17] D. Kern, E.R. Zuiderweg, The role of dynamics in allosteric regulation, Curr. Opin. Struct. Biol. 13 (6) (2003) 748–757, https://doi.org/10.1016/j.sbi.2003.10.008. From NLM Medline.
[18] K.W. East, E. Skeens, J.Y. Cui, H.B. Belato, B. Mitchell, R. Hsu, V.S. Batista, G. Palermo, G.P. Lisi, NMR and computational methods for molecular resolution of allosteric pathways in enzyme complexes, Biophys. Rev. 12 (1) (2020) 155–174, https://doi.org/10.1007/s12551-019-00609-z.
[19] R. Selvaratnam, S. Chowdhury, B. VanSchouwen, G. Melacini, Mapping allostery through the covariance analysis of NMR chemical shifts, Proc. Natl. Acad. Sci. U.S. A. 108 (15) (2011) 6133–6138, https://doi.org/10.1073/pnas.1017311108. From NLM Medline.
[20] S. Boulton, M. Akimoto, R. Selvaratnam, A. Bashiri, G. Melacini, A tool set to map allosteric networks through the NMR chemical shift covariance analysis, Sci. Rep. 4 (2014) 7306, https://doi.org/10.1038/srep07306. From NLM Medline.
[21] B. VanSchouwen, R. Ahmed, J. Milojevic, G. Melacini, Functional dynamics in cyclic nucleotide signaling and amyloid inhibition, Biochim. Biophys. Acta (BBA) – Proteins Proteomics 1865 (11) (2017) 1529–1543.
[22] H. Chen, W.M. Marsiglia, M.K. Cho, Z. Huang, J. Deng, S.P. Blais, W. Gai, S. Bhattacharya, T.A. Neubert, N.J. Traaseth, et al., Elucidation of a four-site allosteric network in fibroblast growth factor receptor tyrosine kinases, Elife 6 (2017), https://doi.org/10.7554/eLife.21137. From NLM Medline.
[23] S. Boulton, R. Selvaratnam, R. Ahmed, G. Melacini, Implementation of the NMR CHEmical shift covariance analysis (CHESCA): a chemical biologist's approach to allostery, Methods Mol. Biol. 1688 (2018) 391–405, https://doi.org/10.1007/978-1-4939-7386-6_18. From NLM Medline.
[24] J.A. Byun, G. Melacini, NMR methods to dissect the molecular mechanisms of disease-related mutations (DRMs): Understanding how DRMs remodel functional free energy landscapes, Methods 148 (2018) 19–27, https://doi.org/10.1016/j.ymeth.2018.05.018. From NLM Medline.
[25] H. Shao, S. Boulton, C. Olivieri, H. Mohamed, M. Akimoto, M.V. Subrahmanian, G. Veglia, J.L. Markley, G. Melacini, W. Lee, P. Yann, CHESPA/CHESCA-SPARKY: automated NMR data analysis plugins for SPARKY to map protein allostery, Bioinformatics 37 (8) (2021) 1176–1177.
[26] D.K. Weber, U.V. Reddy, S. Wang, E.K. Larsen, T. Gopinath, M.B. Gustavsson, R. L. Cornea, D.D. Thomas, A. De Simone, G. Veglia, Structural basis for allosteric control of the SERCA-Phospholamban membrane complex by Ca(2+) and phosphorylation, Elife (2021) 10, https://doi.org/10.7554/eLife.66226. From NLM Medline.
[27] J.M. Axe, E.M. Yezdimer, K.F. O'Rourke, N.E. Kerstetter, W. You, C.E. Chang, D. D. Boehr, Amino acid networks in a (beta/alpha)(8) barrel enzyme change during catalytic turnover, J. Am. Chem. Soc. 136 (19) (2014) 6818–6821, https://doi.org/10.1021/ja501602t. From NLM Medline.
[28] J.M. Axe, D.D. Boehr, Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling, J. Mol. Biol. 425 (9) (2013) 1527–1545, https://doi.org/10.1016/j.jmb.2013.01.030. From NLM Medline.
[29] H. Mohamed, U. Baryar, A. Bashiri, R. Selvaratnam, B. VanSchouwen, G. Melacini, Identification of core allosteric sites through temperature- and nucleus-invariant chemical shift covariance, Biophys. J . 121 (11) (2022) 2035–2045, https://doi.org/10.1016/j.bpj.2022.05.004. From NLM Medline.
[30] R. Selvaratnam, M. Akimoto, B. VanSchouwen, G. Melacini, cAMP-dependent allostery and dynamics in Epac: an NMR view, Biochem. Soc. Trans. 40 (1) (2012) 219–223, https://doi.org/10.1042/BST20110628. From NLM Medline.
[31] R. Selvaratnam, M.T. Mazhab-Jafari, R. Das, G. Melacini, A. Hofmann, The auto-inhibitory role of the EPAC hinge helix as mapped by NMR, PLoS One 7 (11) (2012), https://doi.org/10.1371/journal.pone.0048707. From NLM Medline e48707.
[32] S. Rahman, M. Beikzadeh, M.P. Latham, Biochemical and structural characterization of analogs of MRE11 breast cancer-associated mutant F237C, Sci. Rep. 11 (1) (2021) 7089, https://doi.org/10.1038/s41598-021-86552-0. From NLM Medline.
[33] E.K. Tokuda, C.H. Comin, L.D. Costa, Revisiting agglomerative clustering, Phys. A (2022) 585, https://doi.org/10.1016/j.physa.2021.126433.
[34] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Cluster analysis and display of genome-wide expression patterns, PNAS 95 (25) (1998) 14863–14868, https://doi.org/10.1073/pnas.95.25.14863. From NLM Medline.
[35] C. Liptak, M.M. Mahmoud, B.E. Eckenroth, M.V. Moreno, K. East, K.S. Alnajjar, J. Huang, J.B. Towle-Weicksel, S. Doublie, J.P. Loria, et al., I260Q DNA polymerase beta highlights precatalytic conformational rearrangements critical for fidelity, Nucleic Acids Res. 46 (20) (2018) 10740–10756, https://doi.org/10.1093/nar/gky825. From NLM Medline.
[36] M.P. Williamson, Using chemical shift perturbation to characterise ligand binding, Prog. Nucl. Magn. Reson. Spectrosc. 73 (2013) 1–16, https://doi.org/10.1016/j.pnmrs.2013.02.001. From NLM Medline.
[37] M.J. de Hoon, S. Imoto, J. Nolan, S. Miyano, Open source clustering software, Bioinformatics 20 (9) (2004) 1453–1454, https://doi.org/10.1093/bioinformatics/bth078. From NLM Medline.
[38] A.J. Saldanha, Java Treeview–extensible visualization of microarray data, Bioinformatics 20 (17) (2004) 3246–3248, https://doi.org/10.1093/bioinformatics/bth349. From NLM Medline.
[39] John W. Eaton, D. B., Søren Hauberg, Rik Wehbring. *GNU Octave version 7.3.0 manual: a high-level interactive language for numerical computations*; 2019.
[40] R. Ahmed, J. Huang, M. Akimoto, T. Shi, G. Melacini, Atomic resolution map of hierarchical self-assembly for an amyloidogenic protein probed through thermal (15)N-R(2) correlation matrices, J. Am. Chem. Soc. 143 (12) (2021) 4668–4679, https://doi.org/10.1021/jacs.0c13289. From NLM Medline.

[41] R. Ahmed, M. Akcan, A. Khondker, M.C. Rheinstädter, J.C. Bozelli, R.M. Epand, V. Huynh, R.G. Wylie, S. Boulton, J. Huang, C.P. Verschoor, G. Melacini, Atomic resolution map of the soluble amyloid beta assembly toxic surfaces, Chem. Sci. 10 (24) (2019) 6072–6082.
[42] R. Selvaratnam, B. VanSchouwen, F. Fogolari, M.T. Mazhab-Jafari, R. Das, G. Melacini, The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants, Biophys. J. 102 (3) (2012) 630–639, https://doi.org/10.1016/j.bpj.2011.12.030. From NLM Medline.
[43] D. Gagne, C. Narayanan, N. Doucet, Network of long-range concerted chemical shift displacements upon ligand binding to human angiogenin, Protein Sci. 24 (4) (2015) 525–533, https://doi.org/10.1002/pro.2613. From NLM Medline.
[44] C. Narayanan, D.N. Bernard, K. Bafna, D. Gagne, P.K. Agarwal, N. Doucet, Ligand-induced variations in structural and dynamical properties within an enzyme superfamily, Front. Mol. Biosci. 5 (2018) 54, https://doi.org/10.3389/fmolb.2018.00054. From NLM PubMed-not-MEDLINE.
[45] M. Arai, J.C. Ferreon, P.E. Wright, Quantitative analysis of multisite protein-ligand interactions by NMR: binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP, J. Am. Chem. Soc. 134 (8) (2012) 3792–3803, https://doi.org/10.1021/ja209936u. From NLM Medline.
[46] M. Goldflam, T. Tarrago, M. Gairi, E. Giralt, NMR studies of protein-ligand interactions, Methods Mol. Biol. 831 (2012) 233–259, https://doi.org/10.1007/978-1-61779-480-3_14. From NLM Medline.
[47] G. Manley, J.P. Loria, NMR insights into protein allostery, Arch. Biochem. Biophys. 519 (2) (2012) 223–231, https://doi.org/10.1016/j.abb.2011.10.023. From NLM Medline.
[48] Y. Xu, D. Zhang, R. Rogawski, C.M. Nimigean, A.E. McDermott, Identifying coupled clusters of allostery participants through chemical shift perturbations, PNAS 116 (6) (2019) 2078–2085, https://doi.org/10.1073/pnas.1811168116. From NLM Medline.
[49] D.S. Cui, V. Beaumont, P.S. Ginther, J.M. Lipchock, J.P. Loria, Leveraging reciprocity to identify and characterize unknown allosteric sites in protein tyrosine phosphatases, J. Mol. Biol. 429 (15) (2017) 2360–2372, https://doi.org/10.1016/j.jmb.2017.06.009.
[50] R. Puthenveetil, O. Vinogradova, Solution NMR: A powerful tool for structural and functional studies of membrane proteins in reconstituted environments, J. Biol. Chem. 294 (44) (2019) 15914–15931, https://doi.org/10.1074/jbc.REV119.009178. From NLM Medline.
[51] F. Ziarelli, L. Peng, C.C. Zhang, S. Viel, High resolution magic angle spinning NMR to investigate ligand-receptor binding events for mass-limited samples in liquids, J. Pharm. Biomed. Anal. 59 (2012) 13–17, https://doi.org/10.1016/j.jpba.2011.10.006. From NLM Medline.
[52] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J. L. Markley, B.D. Sykes, 1H, 13C and 15N chemical shift referencing in biomolecular NMR, J. Biomol. NMR 6 (2) (1995) 135–140, https://doi.org/10.1007/BF00211777. From NLM Medline.
[53] Z.K. Boswell, S. Rahman, M.D. Canny, M.P. Latham, A dynamic allosteric pathway underlies Rad50 ABC ATPase function in DNA repair, Sci. Rep. 8 (1) (2018) 1639, https://doi.org/10.1038/s41598-018-19908-8. From NLM Medline.