Probing conformational landscapes and mechanisms of allosteric communication in the functional states of the ABL kinase domain using multiscale simulations and network-based mutational profiling of allosteric residue potentials



¹ Keck Center for Science and Engineering, Graduate Program in Computational and Data Sciences, Schmid College of Science and Technology, Chapman University, Orange, California 92866, USA

² Department of Biomedical and Pharmaceutical Sciences, Chapman University School of Pharmacy, Irvine, California 92618, USA

³ Department of Chemistry, Center for Research Computing, Center for Drug Discovery, Design, and Delivery (CD4), Southern Methodist University, Dallas, Texas 75205, USA

Note: This paper is part of the JCP Special Topic on New Views of Allostery. ^{a)}Author to whom correspondence should be addressed: verkhivk@chapman.edu. Telephone: 714-516-4586. Fax: 714-532-6048

ABSTRACT

In the current study, multiscale simulation approaches and dynamic network methods are employed to examine the dynamic and energetic details of conformational landscapes and allosteric interactions in the ABL kinase domain that determine the kinase functions. Using a plethora of synergistic computational approaches, we elucidate how conformational transitions between the active and inactive ABL states can employ allosteric regulatory switches to modulate intramolecular communication networks between the ATP site, the substrate binding region, and the allosteric binding pocket. A perturbation-based network approach that implements mutational profiling of allosteric residue propensities and communications in the ABL states is proposed. Consistent with biophysical experiments, the results reveal functionally significant shifts of the allosteric interaction networks in which preferential communication paths between the ATP binding site and substrate regions in the active ABL state become suppressed in the closed inactive ABL form, which in turn features favorable allosteric coupling between the ATP site and the allosteric binding pocket. By integrating the results of atomistic simulations with dimensionality reduction methods and Markov state models, we analyze the mechanistic role of macrostates and characterize kinetic transitions between the ABL conformational states. Using network-based mutational scanning of allosteric residue propensities, this study provides a comprehensive computational analysis of long-range communications in the ABL kinase domain and identifies conserved regulatory hotspots that modulate kinase activity and allosteric crosstalk between the allosteric pocket, ATP binding site, and substrate binding regions.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0133826

INTRODUCTION

Allosteric molecular events involve complex multiscale cascades of thermodynamic and dynamic changes that occur on different spatial and temporal levels.¹⁻⁷ The human protein kinases that orchestrate functional processes in cellular networks are recognized as dynamic regulatory machines that exploit allosteric mechanisms for dynamic switching between the inactive and active kinase forms, thus enabling an efficient control of activity and adaptability during processes of signal transduction and catalysis.⁸ The wealth of structural knowledge about conformational states of the kinase catalytic domain, regulatory assemblies, and ligand-kinase complexes is enormous and has dramatically advanced our understanding of the molecular determinants underlying kinase dynamics, function, and binding.^{11–13} The dynamic equilibrium between inactive and active kinase states can be affected and selectively modulated through activated mutations, by posttranslational modifications, and via ligand and protein binding events. This principle has been successfully exploited in the discovery of small molecule kinase modulators.^{14–16} Conformational transitions between kinase states are orchestrated by three conserved structural motifs in the catalytic domain: the aC-helix, the DFG-Asp motif (DFG-Asp in, active; DFG-Asp out, inactive), and the activation loop (A-loop open, active; A-loop closed, inactive). The conserved His-Arg-Asp (HRD) motif in the catalytic loop and the DFG motif are coupled with the aC-helix to form conserved intramolecular networks termed regulatory spine (R-spine) and catalytic spine (C-spine). A universal mechanism for dynamically driven allosteric conformational transformations and activation of kinases can be mediated by coordinated signal transmission through ordered hydrophobic architecture of the assembled spine networks.¹⁷ Nuclear magnetic resonance (NMR) studies have been instrumental in deciphering the allosteric energy landscapes of protein kinases, showing that the dynamic equilibria between inactive and active states can be regulated through ligand binding to generate positive allosteric cooperativity and long-range communication networks between distal kinase domain regions.^{18,19} These studies suggested that the structure and dynamics of the kinase catalytic core and residues surrounding the active site may have coevolved to optimize the intramolecular allosteric communication. Crystallographic and biochemical studies have provided a molecular framework for understanding mechanisms of ABL regulation by unveiling structural organization of the regulatory complexes.^{20,2}

Despite the established view of kinases as dynamic regulatory machines, atomistic characterization of intrinsic allosteric motions and functionally relevant transient states is often lacking due to large conformational transformations and short-lived kinase intermediates involved in the kinetics of allosteric shifts. A series of pioneering NMR studies has provided a detailed atomistic picture of allosteric regulation in the ABL kinase by showing how interacting signaling modules cooperate with the kinase catalytic domain to form a multilayered regulatory mechanism that exploits various allosteric switches at different sites of the regulatory assemblies.²² A further considerable breakthrough in our understanding of the intrinsic conformational landscape of the ABL kinase domain has been recently achieved in the NMR experiments by Kalodimos and colleagues by revealing atomistic details of the hidden conformational states and determining the thermodynamically stable ground



FIG. 1. The NMR solution structure of the thermodynamically stable fully active ground state of the ABL kinase domain (pdb id 6XR6) is represented by green ribbons with cylindrical helices (a). The NMR structure of ABL in the inactive state I₁ (pdb id 6XR7) is represented by red ribbons (b) and the ABL structure in the closed inactive state I₂ (pdb id 6XR6) is represented by blue ribbons (c). The R-spine residues M309, L320, H380, F401, and D440 in the active state and inactive state I₁ are represented by green and blue spheres, respectively. The R-spine positions in the inactive state I₂ are L309, L320, H380, F401, and D440 and are represented by blue spheres. The structures point to similarities and differences in the key functional regions of the kinase domain exemplified by the α C helix, the A-loop, and the P-loop. In particular, A-loop in the inactive state I₂ (c) adopts a completely different closed conformation. A fully assembled R-spine in the active ABL state (a) becomes partially broken in the inactive state I₁ (b) and is fully disassembled in the closed inactive state I₂ (c).

state of the ABL in the active state [Fig. 1(a)] and two inactive conformations.²³ By using state-of-the-art NMR techniques, this study captured the inactive short-lived ABL conformations I₁ and I₂ [Figs. 1(b) and 1(c)] that occur only 5% of the time and are quite different from each other in the critical regions of the A-loop and the regulatory α C-helix motif. The unprecedented atomistic characterization of the conformational ABL ensemble has shown that these states are intrinsically present in the isolated ligand-free kinase domain and that allosteric transitions between these conformations, and mutations.²³

X-ray crystallography and hydrogen exchange mass spectrometry studies have characterized allosteric myristic pocket in the Cterminal lobe of the kinase domain showing that allosteric inhibitors GNF-2/GNF-5 and Asciminib can bind to a bent conformation of the aI helix that serves as a switch promoting interactions with the SH2 domain in the regulatory complex and stabilization of the closed inhibited conformation.^{24,25} At the same time, ligands that bind to the allosteric pocket but do not cause conformational change in the aI helix have been found to often function as allosteric activators of kinase function.²⁶ Remarkably, NMR spectroscopy studies have recently shown that Imatinib, which inhibits the catalytic site by inducing a specific inactive ABL conformation, can also act as an allosteric activator by binding to the myristic pocket and unexpectedly promoting kinase activity through binding competition between the ATP site an allosteric binding site, which becomes functionally relevant in the Imatinib-resistant ABL variants.²⁷ Recent NMR analyses have suggested an alternative mechanism in which disassembly of the inhibitory ABL-SH2-SH3 state and opening of the regulatory core may be directly caused by Imatinib binding to the ATP site that allosterically enacts breakage of the interdomain interfaces.²⁸⁻³⁰ These crystallographic and NMR studies have revealed functional role of allosteric interaction networks in protein kinases, particularly highlighting the complexity of allosteric transformations and long-range communication networks between the ATP binding site, the substate binding region, and the allosteric site.

Computer simulations have explored the free energy landscape of ABL kinase suggesting that metastable ABL states are often structurally similar to known crystal structures of other kinases in complexes with a variety of inhibitors.^{31–34} Integration of the molecular simulations and dynamic network modeling approaches described conserved allosteric pathways of the ABL and EGFR kinase regulation, providing a mechanistic model of allosteric communications between the ATP binding and regulatory sites.^{35–38} Atomistic modeling of the ABL regulatory complexes bound to allosteric inhibitors and activators has shown that these small molecules can differentially modulate protein kinase activities by altering allosteric communications between the allosteric pocket and the binding regions.^{39,40}

A substantial challenge in investigating allosteric mechanisms for large multi-domain kinase assemblies is the inherent difficulty of adapting experimental and computational methodologies to capture the intrinsic flexibility of these structures essential for functionality. Multidisciplinary structural biology studies that exploited synergies between NMR technologies, biophysical approaches, and multiscale computational methods have been particularly fruitful in uncovering the invisible dynamic aspects of diverse protein functions. In particular, this powerful toolkit of structure-centric techniques has been applied for mapping allosteric protein landscapes,⁴¹⁻⁴⁴ dissecting ligand-induced modulation of allosteric activation,45 and characterization of allosteric communication networks that drive protein regulatory responses.^{46–48} A general approach of quantifying mutational effects for multiple molecular phenotypes using a specific implementation of multidimensional deep mutational scanning enabled a comprehensive characterization of allosteric mutations for many proteins.44 Using a combination of triple-resonance NMR and computational network analysis, the allosteric effects of specific kinase mutations and communication paths between regulatory elements and catalytic sites can be characterized.47 Using a combination of NMR spectroscopy, isothermal titration calorimetry (ITC), small-angle x-ray scattering (SAXS), and MD simulations, the intramolecular network of communications in PKA-C were elucidated, suggesting that mutation-induced network changes may compromise highly cooperative regulation process leading to disease progression.⁴⁸ Integration of NMR spectroscopy and surface plasmon resonance revealed dynamic communication networks of residues linking the ligand-binding site to the activation interface in the glucocorticoid receptor and identified a specific motif acting as a ligand- and coregulator-dependent switch for transcriptional activation.⁴⁹ Solution NMR experiments and Gaussian-accelerated molecular dynamics (GaMD) simulations have been performed to examine the structural and dynamic determinants of allosteric signaling within the CRISPR-Cas9 HNH nuclease, advancing our understanding of the allosteric pathway of activation.⁵⁰ Further integration of NMR with multi-microsecond molecular dynamics (MD) simulations and graph-based network modeling probed the effects of mutations on the structure and allosteric communication within the CRISPR-Cas9 system, showing that mutations responsible for increasing the specificity of Cas9 alter the allosteric structure

of the catalytic HNH domain.⁵¹ NMR chemical shift covariance (CHESCA) and projection (CHESPA) analyses⁵²⁻⁵⁵ can identify residue interaction networks that show correlated changes in chemical shifts due to allosteric perturbations caused by ligand binding or mutations designed to modulate allosteric conformational equilibria. Using statistical comparative analyses of NMR chemical shift variations elicited by the selected perturbations, CHESCA approach characterizes perturbation-specific chemical shift patterns serving as signatures of allosteric mechanisms.

In the current study, we propose and integrate several computational approaches to examine principles of the ABL kinase domain allostery where the predicted allosteric signatures can be directly compared with structural and biophysical experiments. Despite unique insights provided by structural approaches in the characterization of conformational landscapes of the ABL kinase, the dynamic and energetic details underlying conformational rearrangements in allosteric networks and the role of allosteric regulatory switches within each of these states are not fully established. MD simulations are combined with the distance fluctuation coupling analysis of ABL conformational ensembles to examine the allosteric role of key functional regions of the kinase domain. We introduce a perturbation-based network approach that implements a "deep" mutational profiling of allosteric residue propensities and interactions in the ABL states. In this method, using conformational ensembles of the ABL structures, we perform systematic modifications of protein residues and evaluate their effect on allosteric interactions and intramolecular long-range communications in the protein structure. By combining MD simulations with dynamic network modeling methods, we show how conformational changes between the active and inactive ABL states can affect allosteric couplings and modulate the intramolecular communication networks between the ATP site, the substrate binding region, and the allosteric myristic pocket. We determine that the preferential communication paths between the ATP binding site and substrate regions in the active state become partly suppressed in the closed inactive state I₂, where, in turn, allosteric communications between the ATP site and the allosteric binding pocket become prevalent. By integrating the results of MD simulations with dimensionality reduction methods and Markov state models (MSMs), we analyze the kinetic transitions between the ABL conformational states and show that the ABL kinase domain can transition following the path from the active state to the inactive state I_1 and then to the inactive state I_2 , while direct transitions from active state to inactive state I₂ are not favorable. Atomistic characterization of dynamic allosteric changes and identification of allosteric regulatory hotspots provide novel insights into the role of allostery in diverse kinase functions, particularly in a mechanism underlying long-range control over communications between the ATP binding site, the substrate binding region, and the allosteric binding site. The results of this study suggest that probing the intramolecular communication networks in the ABL conformations through targeted modifications of vulnerable network links may be useful for engineering and modulating kinase activities.

MATERIALS AND METHODS MD simulations

The crystal structures of the three ABL conformations in the active state, the inactive state I_1 , and inactive state I_2 were taken from

the Protein Data Bank (PDB), (PDB ID 6XR6, 6XR7, and 6XRG, respectively).⁵⁶ The structures were further preprocessed through the Protein Preparation Wizard (Schrödinger, LLC, New York, NY) and included the check of bond order, assignment and adjustment of ionization states, formation of disulphide bonds, removal of crystallographic water molecules and cofactors, capping of the termini, assignment of partial charges, and addition of possible missing atoms and side chains. A box of TIP3P water molecules was used to simulate each system. Assuming normal charge states of ionizable groups corresponding to pH = 7, sodium (Na⁺) and chloride (Cl⁻) counterions were added to achieve charge neutrality and a salt concentration of 0.15M NaCl was maintained. All Na⁺ and Cl⁻ ions were placed at least 8 Å away from protein atoms and from each other. The topology and coordinate files for MD simulations were prepared using tleap⁵⁷ and all-atom MD simulations were performed with the AMBER ff14SB force field.⁵⁸ Energy minimization after addition of solvent and ions was conducted using the steepest descent method for 100 000 steps in each ABL state. In the preparation stage, 100 ps of NVT ensemble Langevin MD simulations were conducted, followed by 200 ps of isothermal-isobaric ensemble (NPT) simulations at 1 atm and 300 K. Three independent 1 μ s NVT MD simulations were conducted for each of the ABL states. All MD simulations were conducted using the graphics processing unit (GPU) version of the OpenMM software environment.⁵⁹ Longrange electrostatic interactions were calculated using the particle mesh Ewald method⁶⁰ with a real space cutoff of 1.0 nm. SHAKE method was used to constrain all bonds associated with hydrogen atoms. Simulations were run using a leap-frog integrator with a 2 fs integration time step.

Distance fluctuations coupling analysis of allosteric residue propensities

The stability and allosteric propensities of protein residues were evaluated using distance fluctuation coupling analysis of the conformational ensembles. The computations and analysis are rooted in the protein mechanics approach^{61,62} in which the fluctuations of the mean distance between a given residue and all other residues in the conformational ensemble are converted into distance fluctuations coupling indices (DFCIs) that measure the energy cost of the residue deformation during MD simulations. Our previous studies^{39,63} showed that the DFCI parameter can correlate with allosteric residue propensities/potentials and the mean square fluctuations between a pair of residues could provide an accurate estimate of the signal commute time. The communication propensity of a pair of residues is inversely related to their commute time CT(i, j) expressed as a function of the variance of the inter-residue distance,

$$CT(i,j) = \left(\left(d_i - \left\langle d_i \right\rangle \right)^2 \right). \tag{1}$$

In our study, MD simulations of the ABL structures are analyzed by computing the fluctuations of the mean distance between each atom within a given residue and the atoms that belong to the remaining residues of the protein. The DFCI index can provide a measure of the allosteric potential for each residue and is calculated by averaging the distances between residues over the simulation trajectory using the following expression:

$$DFCI_i = \frac{3k_BT}{\langle (d_i - \langle d_i \rangle)^2 \rangle},$$
(2)

$$d_i = \left\langle d_{ij} \right\rangle_{i*},\tag{3}$$

where d_{ij} is the instantaneous distance between residue *i* and residue *j*, k_B is the Boltzmann constant, and T = 300 K. (\rangle denotes an average taken over the MD simulation trajectory and $d_i = \langle d_{ij} \rangle_{j*}$ is the average distance from residue *i* to all other atoms *j* in the protein (the sum over j_* implies the exclusion of atoms that belong to the residue *i*). The reported DFCI for each residue is computed as the average of distance fluctuations for all its atoms *i*. It should be noted that in addition, we also computed the mean fluctuations of a given residue by using C_{α} atom positions as well as the reduced representations with a single pseudo-atom per residue^{61,62} and a more refined model in which each amino acid is represented by one pseudo-atoms representing the side chain.⁶⁴ For clarity and consistency of discussion, the reported DFCIs are based on full all-atom protein representation.

Dimensionality reduction analysis of the conformational ensembles

MD equilibrium trajectories were analyzed using the ivis dimensionality reduction method, which is a machine learning model that applies a Siamese neural network architecture composed of three identical neural networks.^{65,66} Each network is composed of three dense layers with sizes of 500, 500, and 2000, followed by a final embedding layer of two neurons. The layers preceding the embedding layer use the selu activation function,

$$\operatorname{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0\\ \alpha e^{x} - \alpha & \text{if } x \le 0. \end{cases}$$
(4)

The values of α and λ are set as the default values of 1.6733 and 1.0507, respectively.⁶⁷ The weights in these layers are initialized randomly using LeCun normal distribution. A linear activation is used in the final embedding layer initialized with Glorot's uniform distribution.

During training, a triplet loss function is calculated as

$$L_{tri}(\theta) = \left[\sum_{a,p,n} D_{a,p} - \min(D_{a,n}, D_{p,n}) + m\right]_+,$$
(5)

where a, p, and n are various kinds of points. a: point of interest, referred to as anchor point; p: positive points, selected through k-nearest neighbors (KNNs); n: negative points, randomly selected from the rest of the samples. A margin (m), the minimum distance gap, was set to the default value of 1. The Euclidean distance (D) between data points is calculated to quantify the similarity,

$$D_{a,b} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}.$$
 (6)

Adam optimizer with default learning rate was applied to train ivis model with GPU acceleration. ivis package (v 2.0.7)

(https://pypi.org/project/ivis/) was used to implement the algorithm.

Markov state model

Stochastic Markov state models (MSMs)⁶⁸⁻⁷⁰ have become increasingly useful states-and-rates network models with the developed software infrastructure^{71–74} for describing transitions between functional protein states and for modeling allosteric events. In MSM, protein dynamics is modeled as a kinetic process consisting of a series of Markovian transitions between different conformational states at discrete time intervals. A specific time interval, referred to as lag time, needs to be determined to construct transition matrix. MSM depends on the simulation distribution of the projected lowdimensional space obtained by the ivis dimensionality reduction method. First, a MiniBatch k-means clustering method is conducted to assign each simulation frame to a microstate. Macrostates were clustered based on the Perron-cluster cluster analysis (PCCA)⁷⁵ and are considered as kinetically separate equilibrium states. The transition matrix and transition probability were calculated to quantify the transition dynamics among macrostates.

The value of the lag time, as well as the number of macrostates, is selected based on the result of estimated relaxation time scale.⁷⁶ The corresponding transition probability is calculated as

$$P_{ij}(\tau) = Prob(x_{t+\tau} \in S_j | x_t \in S_i).$$
(7)

A proper lag time is required as the MSMs are required to be Markovian at the chosen lag time, which can be assessed through the convergence of implied relaxation time scale. The implied time scale can be calculated using the eigenvalues (λ_i) in the transition matrix as

$$t_i = -\frac{\tau}{\ln|\lambda_i(\tau)|}.$$
(8)

MSMs were constructed and the implied time scales were calculated using the PyEMMA package.⁷⁷ Based on the transition matrix, we obtain implied time scales for transitioning between various regions of the phase space and use this information to determine the number of metastable states. The number of metastable states also defines the resolution of the model by determining how large a barrier must be in order to divide the phase space into multiple states.

Dynamic network analysis

A graph-based representation of protein structures^{78–80} is used to represent residues as network nodes and the inter-residue edges to describe non-covalent residue interactions. We constructed the residue interaction networks using both dynamic correlations^{81,82} and coevolutionary residue couplings.^{83,84} To characterize allosteric couplings of the protein residues and account for a cumulative effect of dynamic and coevolutionary correlations, we employed the generalized correlation coefficient.⁸⁵ The RING program was also employed for the initial generation of residue interaction networks.⁸⁶ The ensemble of shortest paths is determined from matrix of communication distances by the Floyd–Warshall algorithm.⁸⁷ Network graph calculations were performed using the python package NetworkX.⁸⁸ The betweenness of residue *i* is defined as the sum of the fraction of shortest paths between all pairs of residues that pass through residue *i*,

$$B_k = \sum_{s \neq k \neq t}^N \frac{S_{st}(k)}{S_{st}},\tag{9}$$

where S_{st} denotes the number of shortest geodesics paths connecting s and t and $S_{st}(k)$ is the number of shortest paths between residues s and t passing through the node k.

The following Z-score is then calculated:

$$Z_i = \frac{B_k - \langle B \rangle}{\sigma}.$$
 (10)

Through mutation-based perturbations of protein residues, we compute changes in the average short path length (ASPL) averaged over all modifications in a given position. The change of ASPL upon mutational changes of each node is given as

$$\Delta L_i = \left(\left\| \Delta L_i^{node}(j) \right\|^2 \right), \tag{11}$$

where *i* is a given site, *j* is a mutation, and $\langle \cdots \rangle$ denotes averaging over mutations. $\Delta L_i^{node}(j)$ describes the change of ASPL upon mutation *j* in a residue node *i*. ΔL_i is the average change of ASPL triggered by mutational profiling of this position.

The Z-score is then calculated for each node as follows:

$$Z_i = \frac{\Delta L_i - \langle \Delta L \rangle}{\sigma}.$$
 (12)

The ensemble-averaged Z-scores ASPL changes are computed from network analysis of the conformational ensembles using 1000 snapshots of the simulation trajectory for the native protein system.

RESULTS AND DISCUSSION

MD simulations of the ABL kinase allosteric states

We performed multiple independent all-atom MD simulations for the ABL kinase states using the structures of the isolated ABL kinase domain in its active conformational state (pdb id 6XR6) and two inactive conformational states I₁ and I₂ (pdb id 6XR7, 6XRG). All these states have been shown to be intrinsic to the unbound ABL kinase domain and allosteric structural transitions between these major conformational states could drive the kinase activities and response to the substrate and ligand binding (Fig. 1).²³ Structural analysis showed that the inactive ABL states [Figs. 1(b) and 1(c)] are very different, in which the regulatory DFG motif adopts distinct "out" conformations in the I₁ and I₂ states, while the A-loop and the aC helix undergo significant rearrangements. In particular, the aC helix moves from its active "aC-in" position to the intermediate position in the I₁ state [Fig. 1(b)] and to the " α C-out" inactive position in the I_2 state [Fig. 1(c)]. In the active " α C-in" state, a conserved α C helix residue E305 forms an ion pair with K290 in the β 3 strand that coordinates the α and β phosphates of the ATP [Fig. 1(a)]. A substantial deviation from this arrangement is seen in the inactive states, which becomes particularly evident in the "aCout" I_2 state where this hydrogen bond is largely broken [Fig. 1(c)]. The R-spine in ABL consists of M309 from the C-terminal end of the α C-helix, L320 from the β 4-strand, F01 of the DFG motif in the beginning of the A-loop, H380 of the HRD motif in the catalytic loop, and D440 of the α F-helix [Figs. 1(a) and 1(b)]. The R-spine subnetwork is fully assembled in the active ABL kinase [Fig. 1(a)] but becomes partially decoupled in the inactive state I₁ [Fig. 1(b)] and fully disassembled in the inactive state I₂ [Fig. 1(c)]. The C-spine is comprised of hydrophobic residues (V275, A288, L342, C388, L389, V336, S457, and I451) and connects the kinase lobes anchoring catalytically important sites to the C-terminus of the α F-helix.

MD simulations revealed important differences in the conformational dynamics of the ABL kinase core and long-range dynamic coupling between the binding sites that may be relevant for kinase inhibition and activation (Fig. 2). By monitoring the root mean square fluctuations (RMSF) profile, we observed that the active ABL kinase state is more stable than both the inactive forms, displaying smaller thermal fluctuations in the intrinsically more flexible N-terminal lobe and A-loop [Fig. 2(a)]. The dynamics of the inactive state I₁ showed larger fluctuations, but the overall shape of the RMSF profile remained mostly similar [Fig. 2(a)]. More significant differences in the dynamics were seen in the fully inactive state I₂, with large thermal fluctuations observed in the middle of the A-loop that could obstruct the substrate Y412 position and inhibit substrate binding. Of particular importance is the dynamic behavior of the key regulatory elements of the catalytic kinase domain, particularly the 400-DFG-402 motif and the R-spine residues.

The DFG-in motif in the active form flips by adopting the DFGout conformation in both inactive states [Figs. 2(b) and 2(c)]. This DFG movement is local in the I1 state but is much larger in the fully inactive state I₂ where F401 translates by ~10 Å to occupy a distinct hydrophobic pocket [Fig. 2(c)]. Notably, while experiencing pronounced structural rearrangements between the active and inactive ABL states, the DFG motif is stabilized in each of these ABL conformations and is anchored through a local interaction network formed with another indispensable stable element of the kinase core N387 residue. These residues displayed exceedingly small thermal fluctuations in both the active and inactive I_2 state [Fig. 2(a)], indicating that the interactions formed by the DFG motif are important for control of allosteric conformational changes in the ABL kinase. Indeed, the hydrogen bonding network between the N387 and D400 residues was experimentally shown to stabilize the DFG-in active conformation acting as a conformational switch for the DFG flips.⁸⁹ The conformational fluctuations of the regulatory aC-helix (residues 299-311) were markedly smaller in the active ABL state where the stable "aC-in" conformation is coupled to the assembled R-spine required for activation [Fig. 2(a)]. In the disassembled R-spine networks of the inactive states, L320, H380, F401, and D440 residues



FIG. 2. Conformational dynamics profiles of the ABL kinase domain states. (a) The root mean square fluctuation (RMSF) profiles are shown for the active ABL form in green lines, for the inactive I_1 state in red lines, and for the inactive closed I_2 state in blue lines. (b) Structural superposition of the ensemble-averaged conformations for the active state (in green ribbons), for the inactive I_1 state (in red ribbons), and for the inactive I_2 state (in blue ribbons). (c) Structural overlay of the regulatory 400-DFG-402 motif from the MD-averaged conformations of the active state (green sticks), the I_1 state (red sticks), and the I_2 state (blue sticks). A large movement of the F401 residue in the DFG-out conformation of the closed inactive I_2 state can be seen. The covariance maps of dynamic cross correlations between pairs of residues in the ABL active state (d), the inactive state I_1 (e), and the inactive state I_2 (f). Cross correlations of residue-based fluctuations vary between +1 (correlated motion; fluctuation vectors in the same direction, colored in dark red) and -1 (anticorrelated motions; fluctuation vectors in the same direction, colored in dark blue). The values >0.5 are colored in dark red and the lower bound in the color bar indicates the value of the most anticorrelated pairs.

experienced moderate fluctuations, while the larger displacements were seen for the α C-helix spine residue M309 (and L309 respectively in I₂) [Fig. 2(a)]. Interestingly, the thermal fluctuations of the regulatory α I-helix on the C-lobe (residues 504-517) are small in all three states, displaying even greater stability in the inactive ABL states. This helix is a key structural motif involved in the formation of the allosteric binding pocket.^{25–27}

The cross correlation matrices of residue fluctuations along the low frequency modes highlighted broadly distributed and strong positive inter-residue couplings in the ABL active state [Fig. 2(d)]. In this form, long-range couplings were also observed between the kinase lobes, including the allosteric and ATP binding sites. Furthermore, the active ABL state featured strong positive coupling between the aF-helix anchoring the central core of the catalytic domain and the aD-helix and A-loop [Fig. 2(d)]. These correlations may reflect allosteric couplings along the C-spine and R-spine in the active state that link the ATP binding site and the substrate binding regions. These dynamic couplings are weakened in the inactive state I_1 [Fig. 2(e)] but the overall pattern of cross correlations between different functional regions remained intact. A marked reduction of the positive long-range dynamic couplings was observed in the inactive state I₂ [Fig. 2(f)]. In this state, a delocalization of coupled fluctuations in the catalytic core was observed, signaling not only a weakening of the long-range couplings but also a reorganization of the locally coupled structural elements in the kinase domain. In general, the cross correlation maps of residue fluctuations in the catalytic domain were indicative of more cooperative allosteric interactions in the active kinase state as compared to looser and narrower patterns of correlated motions seen in the inactive states.

Collective dynamics analysis of the ABL conformational states reveals conserved hinge positions controlling modulation of allosteric changes

We characterized collective motions for the ABL states averaged over the ten lowest frequency modes using principal component analysis (PCA) of the MD trajectories (Fig. 3). The local minima along these profiles provide information about potential hinge regions that may control functional movements, while the maxima typically point to the protein regions undergoing global functional movements. It is assumed that functional movements along the preexisting slow modes could drive allosteric transformations between the active and inactive states and provide insight into the role of functional kinase residues in modulating population-shift allosteric changes between ABL states. The overall shape of the essential profiles and the key hinge centers were generally preserved across the three ABL states but showed considerable differences between globally moving regions in the inactive states (Fig. 3). The predicted hinge positions 317-LVQLLGV-323 and 335-EFMTYGN-341 can mediate large-scale movements of the N-terminal lobe [Fig. 3(a)]. This finding is consistent with the experimental evidence,⁹⁰ suggesting that mutations in these regions can increase the hinge flexibility and modulate the kinetics of allosteric changes between kinase states. According to experimental studies, allosteric transitions related to domain movements can be modulated by mutations into the hinge region to generate two mutants with increased flexibility.9

The importance of hinge residues located between the Nand C-lobes of the kinase was also confirmed using NMR analysis, which revealed that mutational disruption of the three molec-



FIG. 3. ollective dynamics profiles of the ABL conformational state. The essential mobility profiles are averaged over the first 10 major low frequency modes. (a) The slow mode profiles of the ABL kinase domain are shown for the active state (in green lines), the inactive I₁ state (in red lines), and the inactive I₂ state (in blue lines). Structural maps of the essential mobility profiles on PCA of the active state (b), the inactive I₁ state (c), and the inactive I₂ state (d). The mobility profiles are projected onto experimentally determined ABL structures represented by ribbons and colored from blue to red according to the rigidity-to-flexibility scale determined by PCA. The R-spine residues are shown in spheres colored according to their level of rigidity/flexibility and annotated.

ular brake residues N549, E565, and K641 in FGFR2K (Q319, E335, K397, respectively, in ABL) activates the enzyme.^{91,92} These studies suggested that a functionally important allosteric pathway between the DFG motif and the molecular brake residues is mediated by I547 (V317 in ABL) located away from the nucleotide and substrate binding pockets. Here, consistent with these studies, we found that dynamic changes in V317 are coupled with the F401 of the DFG motif (Fig. 3). In addition, our analysis revealed another hinge cluster of locally interacting residues that is formed by residues N387, C388, A399 and the regulatory 400-DFG-402 motif [Fig. 3(a)]. According to our findings, the local interaction clusters of hinge residues anchored by V317 and N387 are dynamically coupled to the F401 of the DFG motif. Given the regulatory role of hinge centers in modulating allosteric transitions, we suggest that local structural perturbations of residues in these clusters may induce global rearrangements between the functional ABL states.

The differences between the slow mode profiles of the ABL states are manifested in the distributions of the moving regions. The P-loop region (residues 267-275) in the N-terminal lobe displayed smaller functional displacements in the active state and the inactive I2 state, even though the P-loop adopts a kinked conformation in the active state and a more stretched conformation in the I_2 state [Fig. 3(a)]. Notably, the shape of the slow mode profiles in this region featured a sharp local minimum (local hinge) that is precisely aligned with the Y272 residue. Indeed, the packing interactions between Y272 and F302 and hydrogen bonding of Y472 with E305 are responsible for stabilizing the P-loop and the aC helix-out conformation. At the same time, the P-loop in the inactive I1 state appears to exhibit large functional displacements in slow modes [Fig. 3(a)]. This may be indicative of the "transitional" nature of the inactive I1 state that acquires increased mobility in the functional regions (including the P-loop) to facilitate adaptation of the fully inactive state. The slow mode profiles also pointed to similar and moderate displacements of the A-loop (residues 395-420) in the active state and inactive I_1 states [Fig. 3(a)] as the A-loop conformation is in an open conformation in both of these forms.

In contrast, significantly larger functional displacements can be afforded in the A-loop of the inactive state I2. We argue that functional movements of the P-loop and especially distorted A-loop in the inactive state I₂ could provide driving force for allosteric changes and conformational adjustments of the ABL in the inactive state. Consistent with this assertion, Imatinib binding to the I₂ state may induce additional structural changes in these regions required for optimizing the inhibitor-ABL interactions.²³ Structural mapping of the slow mode profiles onto the ABL structures highlighted subtle but important variations in the distribution of the immobilized and globally flexible kinase segments in slow modes [Figs. 3(b)-3(d)]. In the active state, the stable regions are aligned with the R-spine and C-spine regions linking the kinase lobes as well as the ATP active site and the substrate binding region [Fig. 3(b)]. The immobilized core becomes smaller in the state I_1 [Fig. 3(c)] where a fraction of the Rspine (H380, F401, and D440) is aligned with the hinge regions but the disassembled component in the aC-helix (M309, L320) appeared to be prone to larger displacements. A further readjustment of the moving regions could be seen in the state I_2 [Fig. 3(d)], in which only the F401 residue of the R-spine remains rigid in the low frequency To summarize, the collective dynamics analysis revealed an important integrating role of the hinge cluster that includes locally interacting and strongly coupled stable residues N387, C388, A399, D400, and F401 that collectively may function as a global coordinator of the kinase functional movements. The dynamic couplings between this rigid hinge cluster and flexible functional regions of the α C-helix, P-loop, and A-loop capable of undergoing coordinated movements can serve as a potential mechanism for modulating allosteric transformations between the active and inactive ABL states.

Distance fluctuation coupling analysis of stability and allosteric residue preferences in the ABL functional states

Using the results of MD simulations, we conducted the distance fluctuation coupling analysis and examined how these distributions and positions of the local peaks change in different ABL states (Fig. 4). In this model, dynamically correlated residues whose effective distances fluctuate with low or moderate intensity are expected to communicate with higher efficiency than the residues that experience large fluctuations. By using the DFCI profile of the active state as the reference, we evaluated how dynamic changes in the inactive states can modulate the allocation of stable residues with high allosteric preferences (or allosteric potential) (Fig. 4). A similar shape of the DFCI profile was observed for all ABL states, revealing that specific regions of the catalytic core could form hubs of dynamic couplings. To examine the role of these regions, we first analyzed the distribution obtained for the ABL active state. The multiple peaks distributed in the kinase domain reflected the presence of many spatially distributed allosteric centers mediating a broad network that links the kinase lobes and the binding sites [Fig. 4(a)]. Strikingly, the characteristic peaks of the distribution corresponded to the N387, C388 (C-spine), L389 (C-spine), A399, 400-DFG-402 residues, and W424 from the substrate binding region 24-WTAPE-428 of the C-lobe [Fig. 4(a)]. All these positions have been experimentally recognized as functionally important for kinase function, with the DFG motif being critical for differences between the inactive and active states, the C-spine residues responsible for catalysis, and the substrate binding residues. W424 is the experimentally validated allosteric node in which mutations (Y404A in protein kinase A) can decouple dynamic couplings and shift the dynamic equilibrium.9

In the inactive state I_1 , the DFCI distribution revealed an important change as the indices for N387, A399, D400, and F401 were reduced. According to our model, this may indicate the increased flexibility and partial weakening of the interactions in this key hotspot cluster [Fig. 4(b)]. In the ABL inactive state I_2 , we observed a partial reorganization of the DFCI distribution peaks in which the depletion of some peaks may narrow the spectrum of mediating centers in the allosteric network, whereas the role of locally coupled N387, A399, D400, F401, and L403 residue is amplified as this cluster emerged as a single dominant hotspot of dynamic couplings in the kinase domain [Fig. 4(b)].



FIG. 4. The residue-based DFCI profiles for the ABL conformational states. (a) The DFCI distributions for the active state (in orange bars) and the inactive state I_1 (in maroon-colored bars). (b) The DFCI distributions for the active state (in orange bars) and the inactive state I_2 (in maroon-colored bars). The positions of the R-spine residues (M309/L309, L320, H380, F401, D440) are shown in magenta-colored filled circles. The residues involved in the N387-mediated local contact network (N387, C388, A389, D400, F401) are shown in yellow-colored filled squares. A close-up of the local interaction cluster mediated by N387 in the ensemble-averaged conformation of the active form (c), the inactive I_1 state (d), and the inactive I_2 state (e). The interacting residues are shown in atom-colored sticks and specific interaction contacts are annotated.

The role of the predicted functional residues in mediating the dynamic inter-residue couplings in the ABL catalytic domain can be better understood by examining the local interaction clusters formed by these key sites in different ABL states. In the ABL active state, the carbonyl oxygen of N387 is hydrogen bound to NH of A399, and the amide group of N387 makes several hydrogen bonds with the side chain and carbonyl oxygen of D400 [Fig. 4(c)]. A strong interaction cluster formed by N387 with A399 and D400 in the active state provides additional support to the fully assembled R-spine and contributes to stabilization of the DFG-in active conformation. This is consistent with a critical role of N387 residue acting as a conformational switch of the ABL dynamics found in the functional studies showing that mutations of N387 that weaken the local interactions in this cluster can allosterically promote the increased global flexibility of ABL and facilitate faster Imatinib dissociation.⁸⁹ Interestingly, a recent computational study showed that Imatinib can dissociate from the wild-type ABL and ABL-N387S mutant through two distinct pathways where the predominant Imatinib unbinding pathway through the kinase hinge region in the wild-type ABL is altered in the ABL-N387 mutant and occurs via the αC-helix region.94

In the inactive I₁ state, the interactions mediated by the N387 residue are weakened due to the loss of hydrogen bonding seen in the active state that could not be compensated by suboptimal N387 side chain contacts with the carbonyl oxygen of F401 [Fig. 4(d)]. On the other hand, in the inactive I₂ state, the hydrogen bonding

mediated by the side chain of N387 with the backbone of A399 and side chain of D400 can be partly restored [Fig. 4(e)]. In addition, F401 of the DFG-out motif is rigidified in its inactive position through favorable contacts with F336 and L403 residues. Our results suggest that the structural changes in the local interactions mediated by N387 can be coupled with the potential allosteric role of these residues. This may partly explain why F401V mutation that enhances stability of the local interactions in this region could preferentially stabilize the inactive I₂ state but not the I₁ inactive form.²³ In addition, the N387-mediated contact network in the I2 state can rigidify the H380-F401-D440 fragment of the distorted R-spine, thereby disfavoring movements required to restore connections with M309/L320 N-terminal component of the R-spine [Fig. 4(e)]. The observed reorganization of the intramolecular contacts between key stabilization centers could therefore promote a complete disintegration of the R-spine and alter the intramolecular allosteric network. Structural mapping of the major distribution peaks associated with the N387 interaction network that includes A399, D400, and F401 residues illustrated the coupling of this local cluster with the Rspine residues (Fig. 5). In the active ABL state, the N387 cluster is tightly coupled to the fully assembled R-spine and together with Y412, R405, and R381 of the 380-HRD-382 motif form an interconnected stabilizing network connecting the kinase lobes and linking the binding sites [Fig. 5(a)]. The "grip" of the N387 mediating interactions on the partially disassembled R-spine in the inactive



ABL Active State (6XR6) ABL Inactive State I1 (6XR7) ABL Inactive State I2 (6XRG

FIG. 5. Structural mapping of the N387-medited local interaction cluster and the R-spine residues in the active state (a), the inactive I₁ form (b), and the inactive I₂ conformation (c). The residues involved in the N387-mediated local contact network (N387, A389, D400, F401) are shown in atom-colored sticks. The R-spine residues are shown in red spheres and annotated. The ABL structures are shown in light-pink colored ribbons.

state I_1 is reduced [Fig. 5(b)]. On the other hand, the N387 cluster could strengthen the disassembled R-spine in the inactive state I_2 [Fig. 5(c)].

To summarize, the DFCI analysis revealed a key role of locally interacting functional residues N387, A399, D400, and F401 that feature a high allosteric potential. Our results indicated that these allosteric centers together with the R-spine residues could modulate dynamic couplings and the intramolecular communication networks in the ABL states.

Network-based mutational profiling of allostery: Conformation-specific reorganization of allosteric interaction networks modulates the intramolecular communications

The residue interaction networks in the ABL conformational states were built using a graph-based representation of protein structures^{78–80} in which residue nodes are interconnected through dynamic^{81,82} and coevolutionary correlations.^{83,84} In this method, we perform systematic modifications of protein residues and evaluate their effect on the dynamic inter-residue couplings and the residue interaction network organization as was initially proposed in our previous studies.95 By computing mutation-induced changes in the topological network parameters such as ASPL that characterize the efficiency and robustness of allosteric communications in the ABL states, we identify the allosteric regulatory hotspots as positions in which the average effect of mutations can cause a significant ASPL change, leading to network alterations that may compromise the efficiency of allosteric communications in the ABL states. The mutational profiling distribution in the active state featured a considerable number of peaks that are broadly distributed in the kinase domain [Figs. 6(a) and 6(b)]. The key sites that affect allosteric communication in the active ABL are precisely aligned with the R-spine residues L320, H380, F401, and D440 [Fig. 6(a)]. In addition, the local interaction cluster anchored by N387 is also involved in mediating the allosteric network in the active form. The distribution highlighted a noticeable peak in the substrate binding motif 424-WTAPE-428 of the C-lobe, which is consistent with the results of functional studies, showing that W424 can act a conserved allosteric node linked to the R-spine residues.93 Structural mapping of the 20 residues featuring the highest allosteric potentials showed a broad and dense interaction network connecting the substrate site with the ATP site [Fig. 6(b)]. This network is determined by connectivity afforded through dynamic couplings between the aC-helix-in, the fully assembled R-spine, and the substrate binding site in the active state. The predicted topological map of allosteric communications in the active ABL is consistent with the experimentally observed patterns of the intramolecular cooperativity in other kinases. For example, biochemical experiments identified a similar dynamically coupled allosteric network between the ATP- and substrate binding sites in Src kinase, showing that robust intramolecular communications and long-range cooperativity in the active state underlie the enzymatic kinase function.⁹⁶ Our findings are also consistent with the NMR-based CHESCA analysis of correlation maps between functional network communities in the active state of PKA enzyme showing strong allosteric communications between the ATP binding site, the aC-helix, and the A-loop.48

We also found that modifications of residues from the allosteric binding pocket (R351, A356, L359, L360, A363, L448, L451, T453, Y454, M456 G482, C483, V487, F512, V525, L529) could induce only minor changes in the allosteric communications of the active form and thus have a minor effect on the efficiency of the allosteric networks [Figs. 6(a) and 6(b)]. It appeared from this analysis that the long-range couplings between the allosteric site and the ATP site may be partially depleted in the active kinase form. These observations are consistent with recent biochemical and structural studies showing a lack of cooperation between ATP-competitive inhibitors that stabilize the α C-helix-in kinase conformation and allosteric inhibitors of ABL.97-100 At the same time, these studies suggested that conformation-selective ATP-site inhibitors that favor the inactive ABL conformations can afford a stronger binding cooperation of the allosteric site with the ATP site. An alternative view based on NMR and isothermal titration calorimetry studies suggested that Asciminib can bind ABL concomitantly with both type-1 and type-2 ATP-competitive inhibitors to form ternary complexes.9

Mutational scanning of allosteric propensities in the inactive state I1 featured a similar distribution shape but revealed reduced peaks associated with the R-spine positions M320, H380, and F401, while the peak corresponding to D440 completely vanished [Figs. 6(c) and 6(d)]. In the I₁ state, the α C helix remains in the "in" conformation with the catalytic bridge K290-E305 intact as in the active state, but the DFG motif swings into inactive "out" position, rendering this state catalytically inactive.²³ From structural mapping of the top allosterically sensitive positions, it is evident that the allosteric network becomes weaker and involves only a portion of the R-spine. Nonetheless, our results suggested that allosteric communications between the ATP site and the substrate binding site in the inactive I1 state can be still functional but less robust than in the active form [Fig. 6(d)]. Interestingly, despite a partially disconnected R-spine, allosteric communication flow between the N-lobe and C-lobe of ABL can still be preserved owing to a large number of mediating centers in the functional regions. These observations are consistent with NMR and network-based studies of allostery in p38y kinases,⁴⁷ showing that a physically connected R-spine in the inactive state may be "dynamically" sufficient to enable allosteric signal propagation between the binding sites along the conserved architecture, while a completely assembled R-spine is a necessary



FIG. 6. Mutational profiling of allosteric residue propensities in the ABL states. The residue-based Z-score profile estimates the average mutation-induced changes in the ASPL network parameter for the active state (a), the inactive I₁ state (c), and the inactive I₂ state (e). The profiles are shown in brown-colored lines. The positions of the R-spine residues on the distribution are highlighted in magenta-filled circles. The residues in the N387-anchored mediating local cluster (N387, C388, A399, D400, F401) are shown in orange-colored circles. The sites targeted by Imatinib-resistant mutations G269E, 71H, Y272H, E274V, T334I, F378V, and H415V are indicated by yellow-colored filles squares. Structural mapping of the top 20 kinase residues with the highest allosteric mediating potential in the active state (b), the inactive I₁ state (d), and the inactive I₂ state (f). The highlighted centers (shown in red spheres) are determined by Z-score evaluation of mutational effects on the intramolecular communication paths. The mapped allosteric mediating sites form intramolecular communication routes connecting the ATP site, the substrate binding region, and the allosteric pocket. The observed shifts in preferential communication paths between the active site and the inactive state I₂ are highlighted.

prerequisite for phospho-transferase activity. Consistent with this study, we similarly found that stronger dynamic residue correlations and a broader interaction network in the active ABL form ensure efficient allosteric communication mediated by the fully assembled R-spine.

The structural and dynamic changes in the inactive I2 state appeared to incur more significant rearrangements in the interaction network in which contributions of the R-spine residues with the exception F401 are reduced [Fig. 6(e)]. The invariant cluster of allosteric hotspots includes N387, C388, L389, A399, D400, and F401 residues [Figs. 6(e) and 6(f)]. However, in contrast to other ABL states, we observed the emergence of additional mediating centers in residues L406, Y408, Y412, L448, I451, T453, Y454 as well as residues from the allosteric binding pocket [Fig. 6(f)]. Notably, the mutation-sensitive allosteric centers also include C-spine residues C388, L389, and I451. NMR analysis showed a significant role of Y408 in increasing the population of the I2 state, and the interactions of the displaced Y412 A-loop residue with the L403/L406/M407 positions are important in allosteric changes of the inactive I2 form.²³ The structural map of top 20 allosteric centers revealed a narrower and more localized pattern as compared to the active state and I_1 [Fig. 6(f)]. Strikingly, the predicted mediating centers in the inactive I2 state form an allosteric route between the ATP binding site

and the allosteric binding pocket, while the long-range communications between the ATP and substrate binding sites become partially depleted.

Our results revealed that the allosteric communications of the ATP site with the substrate binding region and allosteric pockets may coexist and are not mutually exclusive. However, the presented findings indicated a potential shift in the allosteric communication ensemble of the I₂ state that favors long-range coupling between the ATP site and allosteric binding pocket [Figs. 6(e) and 6(f)].

These findings are consistent with recent NMR studies showing lack of synergy and "weak antagonism" between allosteric inhibitor Asciminib bound in the allosteric site and ATP-competitive inhibitors that stabilize the open α C-helix ABL conformation, whereas in contrast, an ATP-competitive inhibitor of ABL that stabilizes the α C-helix-out, closed conformation showed increased synergy with Asciminib binding.⁹⁷ According to these experiments, the mechanism by which Asciminib binding inhibits the ABL catalytic function may be related to the synergy between the allosteric site and the ATP site in the inactive ABL conformations.

To summarize, the central finding of the network-based mutational scanning analysis is conformation-specific modulation of the allosteric interactions that may control shifts in the preferential communication routes connecting the ATP site, the substrate binding site, and the allosteric binding pocket. Notably, our results and analysis are performed using the experimentally determined conformational states of the unbound ABL catalytic domain and did not consider the complete regulatory ABL-SH2-SH3 construct where the presence of the interacting SH2 and SH3 domains in the assembled state can influence the dynamic properties of the allosteric pocket. Nonetheless, the presented data provided evidence that allosteric crosstalk between the binding site and regulatory site can be differentially modulated in the active and inactive states through "activation" and "suppression" of the regulatory hotspots and preferential stabilization of conformation-specific communication routes.

Dimensionality reduction analysis and Markov state model of allosteric transitions

While the analysis of MD simulations and functional movements provided important insights into the underlying conformational landscape of ABL, the high dimensionality of the datasets produced by simulations often hinders salient dynamic signatures associated with the mechanisms of allosteric transitions. Here, to facilitate the conformational landscape analysis we employed a deep learning-based dimensionality reduction method to project the results of MD simulations into low-dimensional space.⁶ The loss function used for the training process in ivis is a triplet loss function that calculates the Euclidean distance among data points and simultaneously minimizes the distances between data of the same labels while maximizing the distances between data of different labels, thus allowing to preserve both local and global structures in a low-dimensional space. The pairwise Ca distances are often used to characterize protein conformations and movements. In ABL kinase, there are 287 residues, leading to 41 041 pairs of Ca distances. To obtain a low-dimensional space projection of the MD ensembles, the ivis dimensionality reduction method was applied to the produced distance dataset. We observed the active state covers a wider region in comparison with the other two states [Fig. 7(a)].

Moreover, there is no overlap between the active state and the inactive state I_2 but instead a big gap as shown in the ivis 2D surface [Fig. 7(a)]. In addition, we noticed that the low-dimensionality

projection of the MD ensemble highlighted a topological similarity between the active state and the inactive I_1 state while also demonstrating a larger conformational space available for the open active ABL conformation [Fig. 7(a)]. A different low-dimensional signature was obtained for the inactive I_2 conformation, reflecting considerable structural and dynamic rearrangements in this state, which are associated with the disassembly and partitioning of intramolecular interaction networks [Fig. 7(a)].

Importantly, compared with dimensionality reduction methods, ivis method was shown to be more robust for constructing MSMs and maintaining high similarity between high and low dimensions with the least information loss. Using the conformational ensembles of the ABL states and ivis dimensionality reduction, we constructed MSM and performed analysis of the macrostates to obtain insight into the kinetic aspects of allosteric transformations in ABL. The projected low-dimensional space produced by the ivis method [Fig. 7(a)] was applied to the MSM construction. In the ivis 2D space, a MiniBatch k-means clustering method was applied to partition the 2D data into 500 microstates. The implied time scales were calculated with lag times ranging from 1 to 2500. The top 15 time scales are shown in Fig. 7(b). The trend of implied relaxation time scale revealed that the estimated time scale converged after ~2500 steps, which was chosen as the lag time in the construction of MSM. Based on the gap of time scales, the number of macrostates was set to eight and the distribution of macrostates with their transition probabilities enabled analysis of allosteric changes between the active and inactive ABL states (Fig. 8). After partition of the MSM microstates to macrostates, the stationary distribution and transition probabilities were calculated based on the constructed MSM. According to the MSM partition of states, macrostates 1-3 belong to the active kinase conformation, macrostates 4 and 6 are associated with the inactive structure I₁, and macrostates 5-8 belong to the inactive structure I₂ (Fig. 8). For each macrostate, the RMSD between its representative structure and the corresponding ABL conformation was calculated. Based on the root mean square deviation (RMSD) values, macrostates 1, 6, and 7 were assigned as the active state, inactive state I, and inactive state I₂, respectively, while other five macrostates were designated as intermediate states (Fig. 8).

The structural signatures of the obtained ABL macrostates are consistent with the MSM study of ABL kinase that revealed a total of



FIG. 7. 2D ivis dimensionality reduction map and the corresponding implied time scales under different lag times. (a) The distribution of three protein conformations on the reduced 2D space. (b) The estimated relaxation time scale of MSMs under different lag times. The number of steps in MD simulations were used as lag times, ranging from 1 to 2500. In each lag time, an MSM was constructed, and the estimated relaxation time scale was calculated from the transition matrix. The implied time scale converged with a lag time of 2000 steps, which was chosen for MSM construction in further analysis.

J. Chem. Phys. **157**, 245101 (2022); doi: 10.1063/5.0133826 Published under an exclusive license by AIP Publishing



FIG. 8. MSM analysis of the ABL conformational landscape. Macrostates 1–3 belong to the active kinase conformation, macrostates 4 and 6 are associated with the inactive structure I, while macrostates 5–8 belong to the inactive structure I₂. Based on the RMSDs between representative structures in macrostates and the corresponding native structures, macrostate 1, 6, and 7 were considered as the active state, inactive state I₁, and inactive state I₂, respectively. Other macrostates were treated as intermediate states.

16 important macrostates differing in the functional conformations adopted by the α C-helix (in, out), the DFG motif (in, out), the P-loop (kinked, extended), and the A-loop (open closed).³¹ We observed that each macrostate is metastable with high probabilities to remain in the same macrostate.

Moreover, the macrostates can more easily interconvert within each of the allosteric ABL states but have higher kinetic barriers for transition to the other states (Fig. 8). For example, for the inactive state I_1 (macrostates 4 and 6), there is an overall probability of 6.9% to shift between two macrostates, but only 2.3% to transition to the active state and 0.6% to the inactive state I_2 (Fig. 8). The transition probabilities revealed that ABL kinase can transition following the path from the active state to the inactive state I1 to the inactive state I2 while direct transitions from the active state to the inactive state I₂ are kinetically unfavorable. Furthermore, there is an overall 3% probability to transition from the active state to inactive state I1 (2% from macrostate 3 to 6 and 1% from macrostate 2 to 6) but only 0.6% probability to transition from the inactive state I1 to inactive state I2 (from macrostate 4 to 8). These results are consistent with the experimental findings of a linear equilibrium between the ground active state G and inactive states $G \leftrightarrow I_1 \leftrightarrow I_2$ where the kinetic exchange between the active and inactive state I2 is slow due to DFG flip and dramatic rearrangements in the A-loop.²³

Structural superposition of microstates 1–3 with the structure of the ABL active conformation highlighted the variability of the P-loop and open A-loop [Fig. 9(a)]. The P-loop can directly affect ligand binding and experiences transitions from the "kinked" conformation in microstate 1 to an "extended" conformation seen in macrostates 2 and 3 [Fig. 9(a)]. Interestingly, the extended P-loop conformation may be a preferable form as it was observed in a large number of protein kinase crystal structures,³¹ while a less common partially folded "kinked" conformation is seen in the active ABL



FIG. 9. Structural analysis of the ABL macrostates. (a) Structural superposition of the conformations representing macrostates 1(in green ribbons), 2 (in red ribbons), and 3 (in blue ribbons) that belong to the active kinase conformation. The experimental active state conformation is shown in light-pink ribbons. (b) Structural superposition of the conformations representing macrostates 4 (in green ribbons) and 6 (in red ribbons) that belong to the inactive I₁ state. The experimental I₁ conformation is shown in light-pink ribbons. (c) Structural superposition of the conformations representing macrostates 5 (in green ribbons), and 8 (in blue ribbons) that belong to the inactive I₂ state (shown in light-pink ribbons).

form²³ and is important for selective binding of Imatinib for ABL kinase over c-SRC kinase.^{101,102} A substantial conformational variability of the open A-loop in the macrostates associated with the active ABL form was also noticeable [Fig. 9(a)]. Importantly, in these macrostates, the α C-helix maintains active "in" conformation, which is the key signature of the active kinase state [Fig. 9(a)]. By aligning the microstates 4 and 6 with the structure of the inactive I₁ state, we observed appreciable variations in the main functional regions α C-helix, the P-loop, and the A-loop [Fig. 9(b)]. While the A-loop in these macrostates undergoes moderate conformational changes, more significant displacements were observed for the α C-helix that experienced movements between the α C helix-in and the inactive α C helix-out positions [Fig. 9(b)].

In the I₁ structure, the kinase adopts the DFG-out/ α C-in conformation and becomes catalytically inactive. The analysis of I₁ macrostates showed that dynamically this inactive state may be susceptible to large movements of the α C-helix that would facilitate conformational transformation to the spectrum of inactive ABL conformations. We also observed large displacements of the "extended" P-loop conformations that are coupled with reciprocal movements of the α C-helix [Fig. 9(b)]. Overall, these findings pointed to globally increased functional movements of the key functional regions in the I₁ macrostates.

Structural variations among microstates 5, 7, and 8 that belong to the inactive I₂ state are even more significant particularly in the N-lobe and A-loop [Fig. 9(c)]. We noticed appreciable displacements of the P-loop that can adopt various "extended" stretched conformations that are necessary to accommodate diverse closed conformations of the A-loop. At the same time, the α C-helix movements are moderate and similar to the α C-helix-out position in the inactive I₂ structure [Fig. 9(c)]. The macrostates revealed a remarkable diversity of the closed A-loop conformations that may be available in the I₂ state. It was proposed in the original structural study that Imatinib could bind preferentially to the ABL I₂ state but induces additional changes to optimize the binding energy.²³ MSM analysis indicated that Imatinib binding can induce a less intrinsically favorable "kinked" conformation of the P-loop and significantly reduce mobility of the closed A-loop. As a result, Imatinib binding may significantly reduce conformational mobility and impair favorable entropy intrinsically present in the I₂ state that could only be partly compensated by the ligand–protein interactions and result in the observed moderate binding affinity.²³ Structural analysis of macrostates associated with low-populated forms of the ABL kinase can expand our understanding of "hidden" allosteric states and invisible aspects of protein kinase functionalities. While computationally predicted inactive macrostates represent shortlived intermediates, their dynamic proximity to the experimentally determined inactive conformations may be leveraged for the design of selective inhibitors with reduced sensitivity to drug-resistant mutations.

CONCLUSIONS

In this work, we introduced and synergistically applied several computational approaches to examine at atomistic detail conformational landscapes and mechanisms of allosteric regulation in the distinct functional states of the ABL kinase domain. By using MD simulations and ensemble-based distance fluctuations analysis, we characterized a critical role of dynamic coupling between N387 residue, the regulatory 400-DFG-402 motif, and the R-spine residues in mediating structural stability and allosteric communications in the ABL states. Our results revealed a previously unappreciated role of N387 as a regulatory anchor that couples functional kinase regions to modulate allostery in the kinase domain. These results are consistent with the recent biophysical experiments and explain the importance of N387 in modulating long-range functional effects in the ABL kinase.

We proposed a network-based mutational profiling approach for probing allosteric residue propensities and communications that revealed a group of conserved allosteric regulatory hotspots in the ABL states. The results of mutational disclosed profiling modulation of the allosteric interaction networks between the active and inactive states where the long-range communication pathways between the ATP binding site and substrate regions are dominant in the active state but are partially suppressed in the closed inactive state I2 and replaced with the stronger allosteric couplings are between the ATP site and the allosteric binding pocket. The important finding of this study is evidence of conformation-specific modulation of allosteric interaction networks and communications that can alter preferential communication routes and crosstalk between the ATP site, the substrate binding site, and allosteric site. These results agree with the latest NMR experiments suggesting a stronger cooperativity between the ATP binding site and allosteric pocket in the inactive closed conformation.97

By integrating the results of MD simulations with MSM-based identification of macrostates, we analyzed the kinetic transitions between the ABL functional conformations that follow the path from the active state to the inactive state I_1 and then to the inactive state I_2 , while direct transitions from active state to inactive state I_2 are not favorable. Through structural analysis of the ABL macrostates, we showed that conformational variability and long-range cooperativity between the functional regions of α C-helix, the P-loop, and the A-loop can provide a concerted mechanism for allosteric transitions underlying regulation of the kinase activity. We argue that probing and exploiting diverse ensemble of the intramolecular communication networks in the ABL conformations through targeted modifications of vulnerable network link and inter-community bridges may be useful for engineering and modulating kinase activities. Together with NMR studies of kinase dynamics and advances in deep mutational scanning of allostery,^{103–105} the proposed computational approach may open up new venues for probing kinase-centric signaling processes and engineering of allosteric functions.

ACKNOWLEDGMENTS

G.M.V. acknowledges support by the Kay Family Foundation Grant No. A20-0032.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Keerthi Krishnan: Data curation (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal). Hao Tian: Data curation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal). Peng Tao: Data curation (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Writing – original draft (equal). Gennady M. Verkhivker: Conceptualization (lead); Data curation (equal); Formal analysis (equal); Funding acquisition (lead); Investigation (lead); Methodology (equal); Project administration (lead); Resources (equal); Software (equal); Supervision (lead); Validation (equal); Visualization (equal); Writing – original draft (lead); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹S. J. Wodak *et al.*, Structure **27**, 566 (2019).
- ²C.-J. Tsai and R. Nussinov, PLoS Comput. Biol. 10, e1003394 (2014).
- ³H. N. Motlagh *et al.*, Nature **508**, 331 (2014).

- ⁵M. Leander et al., Proc. Natl. Acad. Sci. U. S. A. 117, 25445 (2020).
- ⁶O. Bozovic et al., Proc. Natl. Acad. Sci. U. S. A. 117, 26031 (2020).

⁷G. Stock and P. Hamm, Philos. Trans. R. Soc. London, Ser. B 373, 20170187 (2018).

- ⁸S. S. Taylor et al., Philos. Trans. R. Soc. London, Ser. B 367, 2517 (2012).
- ⁹S. S. Taylor et al., Nat. Rev. Mol. Cell Biol. 13, 646 (2012).

¹⁰K. Oruganty and N. Kannan, Philos. Trans. R. Soc. London, Ser. B 367, 2529 (2012).

- ¹¹H. S. Meharena *et al.*, PLoS Biol. **11**, e1001680 (2013).
- ¹² A. P. Kornev and S. S. Taylor, Trends Biochem. Sci. 40, 628 (2015).
- ¹³S. S. Taylor *et al.*, J. Biol. Chem. **296**, 100746 (2021).
- ¹⁴D. Fabbro, Mol. Pharmacol. 87, 766 (2015).

⁴N. V. Dokholyan, Chem. Rev. 116, 6463 (2016).

- ¹⁵R. Roskoski, Jr., Pharmacol. Res. **100**, 1 (2015).
- ¹⁶G. M. Verkhivker, Curr. Med. Chem. **24**, 4838 (2017).
- ¹⁷J. Kim *et al.*, Sci. Adv. **3**, e1600663 (2017).
- ¹⁸Y. Wang et al., Nat. Commun. 10, 799 (2019).

¹⁹L. R. Masterson, A. Cembran, L. Shi, and G. Veglia, Adv. Protein Chem. Struct. Biol. 87, 363 (2012).

- ²⁰S. Panjarian *et al.*, J. Biol. Chem. **288**, 5443 (2013).
- ²¹ A. J. Lamontanara *et al.*, Nat. Commun. 5, 5470 (2014).
- ²²T. Saleh, P. Rossi, and C. G. Kalodimos, Nat. Struct. Mol. Biol. 24, 893 (2017).
- ²³T. Xie, T. Saleh, P. Rossi, and C. G. Kalodimos, Science **370**, eabc2754 (2020).
- ²⁴F. J. Adrian *et al.*, Nat. Chem. Biol. **2**, 95 (2006).
- ²⁵J. Zhang *et al.*, Nature **463**, 501 (2010).
- ²⁶J. Yang *et al.*, Chem. Biol. **18**, 177 (2011).
- ²⁷T. Xie *et al.*, J. Mol. Biol. **434**, 167349 (2022).
- ²⁸L. Skora *et al.*, Proc. Natl. Acad. Sci. U. S. A. **110**, E4437 (2013).
- ²⁹ R. Sonti *et al.*, J. Am. Chem. Soc. **140**, 1863 (2018).
- ³⁰S. Grzesiek, J. Paladini, J. Habazettl, and R. Sonti, Magn. Reson. 3, 91 (2022).
- ³¹Y. Meng et al., J. Chem. Theory Comput. 14, 2721 (2018).
- ³²Y. Meng, M. P. Pond, and B. Roux, Acc. Chem. Res. **50**, 1193 (2017).
- ³³F. Paul, Y. Meng, and B. Roux, J. Chem. Theory Comput. 16, 1896 (2020).
- ³⁴F. Paul, T. Thomas, and B. Roux, J. Chem. Theory Comput. **16**, 7852 (2020).
- ³⁵A. Dixit and G. M. Verkhivker, PLoS Comput. Biol. 7, e1002179 (2011).
- ³⁶A. Tse and G. M. Verkhivker, J. Chem. Inf. Model. 55, 1645 (2015).
- ³⁷K. A. James and G. M. Verkhivker, PLoS One 9, e113488 (2014).
- ³⁸A. Tse and G. M. Verkhivker, PLoS One **10**, e0130203 (2015).
- ³⁹L. Astl and G. M. Verkhivker, J. Chem. Theory Comput. 15, 3362 (2019).
- ⁴⁰L. Astl, A. Tse, and G. M. Verkhivker, Adv. Exp. Med. Biol. 1163, 187 (2019).
 ⁴¹M. ElGamacy *et al.*, Structure 27, 853 (2019).
- ⁴²M. J. Bostock, A. S. Solt, and D. Nietlispach, Curr. Opin. Struct. Biol. 57, 145 (2019).
- ⁴³E. T. Usher and S. A. Showalter, J. Biol. Chem. 295, 17411 (2020).
- ⁴⁴A. J. Faure *et al.*, Nature **604**, 175 (2022).
- ⁴⁵A. Zorba *et al.*, Proc. Natl. Acad. Sci. U. S. A. **116**, 13937 (2019).
- ⁴⁶S. Grutsch, S. Brüschweiler, and M. Tollinger, PLoS Comput. Biol. 12, e1004620 (2016).
- ⁴⁷P. C. Aoto, B. T. Martin, and P. E. Wright, Sci. Rep. 6, 28655 (2016).
- ⁴⁸C. Olivieri *et al.*, Commun Biol **4**, 321 (2021).
- ⁴⁹C. Köhler *et al.*, Sci. Adv. **6**, eabb5277 (2020).
- ⁵⁰K. W. East *et al.*, J. Am. Chem. Soc. **142**, 1348 (2020).
- ⁵¹L. Nierzwicki *et al.*, <u>eLife</u> **10**, e73601 (2021).
- ⁵²R. Selvaratnam, S. Chowdhury, B. VanSchouwen, and G. Melacini, Proc. Natl. Acad. Sci. U. S. A. **108**, 6133 (2011).
- ⁵³S. Boulton and G. Melacini, Chem. Rev. **116**, 6267 (2016).
- ⁵⁴S. Boulton, R. Selvaratnam, R. Ahmed, and G. Melacini, <u>Methods Mol. Biol.</u> 1688, 391 (2018).
- ⁵⁵M. Akimoto *et al.*, Biochem. J. **479**, 825 (2022).
- ⁵⁶P. W. Rose *et al.*, Nucleic Acids Res. **45**, D271 (2017).
- ⁵⁷D. A. Case *et al.*, J. Comput. Chem. **26**, 1668 (2005).
- ⁵⁸J. A. Maier et al., J. Chem. Theory Comput. 11, 3696 (2015).
- ⁵⁹ P. Eastman *et al.*, PLoS Comput. Biol. **13**, e1005659 (2017).
- ⁶⁰ M. Di Pierro, R. Elber, and B. Leimkuhler, J. Chem. Theory Comput. **11**, 5624 (2015).
- ⁶¹S. Sacquin-Mora and R. Lavery, Biophys. J. 90, 2706 (2006).
- ⁶²S. Sacquin-Mora, O. Delalande, and M. Baaden, Biophys. J. 99, 3412 (2010).

- ⁶³G. Stetz, L. Astl, and G. M. Verkhivker, J. Chem. Theory Comput. 16, 4706 (2020).
- ⁶⁴M. Zacharias, Protein Sci. 12, 1271 (2003).
- 65 B. Szubert, J. E. Cole, C. Monaco, and I. Drozdov, Sci. Rep. 9, 8914 (2019).
- ⁶⁶H. Tian and P. Tao, J. Chem. Inf. Model. **60**, 4569 (2020).
- ⁶⁷C. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, *Advances in Neural Information Processing Systems Long Beach California USA Decembe* 4–9, 2017 (Red Hook, NY).
- ⁶⁸D. Shukla, C. X. Hernández, J. K. Weber, and V. S. Pande, Acc. Chem. Res. 48, 414 (2015).
- ⁶⁹E. Suárez, J. L. Adelman, and D. M. Zuckerman, J. Chem. Theory Comput. 12, 3473 (2016).
- ⁷⁰G. R. Bowman et al., Proc. Natl. Acad. Sci. U. S. A. 112, 2734 (2015).
- ⁷¹ M. P. Harrigan *et al.*, Biophys. J. **112**, 10 (2017).
- ⁷²B. Cronkite-Ratcliff and V. Pande, Bioinformatics **29**, 950 (2013).
- ⁷³G. R. Bowman and F. Noe, Adv. Exp. Med. Biol. **797**, 139 (2014).
- ⁷⁴G. R. Bowman, Methods Mol. Biol. 1084, 141 (2014).
- ⁷⁵P. Deuflhard and M. Weber, Linear Algebra Appl. **398**, 161 (2005).
- ⁷⁶G. R. Bowman, X. Huang, and V. S. Pande, Methods **49**, 197 (2009).
- ⁷⁷M. K. Scherer *et al.*, J. Chem. Theory Comput. **11**, 5525 (2015).
- ⁷⁸K. V. Brinda and S. Vishveshwara, Biophys. J. 89, 4159 (2005).
- ⁷⁹M. S. Vijayabaskar and S. Vishveshwara, Biophys. J. 99, 3704 (2010).
- ⁸⁰P. Hansia, A. Ghosh, and S. Vishveshwara, Mol. BioSyst. 5, 1860 (2009).
- ⁸¹ A. Sethi, J. Eargle, A. A. Black, and Z. Luthey-Schulten, Proc. Natl. Acad. Sci. U. S. A. **106**, 6620 (2009).
- ⁸² M. C. R. Melo, R. C. Bernardi, C. de la Fuente-Nunez, and Z. Luthey-Schulten, J. Chem. Phys. 153, 134104 (2020).
- ⁸³F. L. Simonetti, E. Teppa, A. Chernomoretz, M. Nielsen, and C. Marino Buslje, Nucleic Acids Res. 41, W8 (2013).
- ⁸⁴G. Stetz and G. M. Verkhivker, PLoS Comput. Biol. 13, e1005299 (2017).
- ⁸⁵O. F. Lange and H. Grubmüller, Proteins **62**, 1053 (2006).
- ⁸⁶D. Clementel *et al.*, Nucleic Acids Res. **50**, W651 (2022).
- ⁸⁷R. W. Floyd, Commun. ACM 5, 345 (1962).

⁸⁸A. A. Hagberg, D. A. Schult, and P. J. Swart, in *Proceedings of the 7th Python in Science Conference*, edited by G. Varoquaux, T. Vaught, and J. Millman (Pasadena, 2008), pp. 11–15.

- ⁸⁹A. Lyczek et al., Proc. Natl. Acad. Sci. U. S. A. 118, e2111451118 (2021).
- ⁹⁰ F. Li *et al.*, Nat. Commun. **13**, 3541 (2022).
- ⁹¹ H. Chen *et al.*, Mol. Cell **27**, 717 (2007).

⁹²W. M. Marsiglia, J. Katigbak, S. Zheng, M. Mohammadi, Y. Zhang, and N. J. Traaseth, Structure 27, 1308 (2019).

- 93 L. G. Ahuja et al., Proc. Natl. Acad. Sci. U. S. A. 114, E931 (2017).
- ⁹⁴M. Shekhar, Z. Smith, M. A. Seeliger, and P. Tiwary, Angew. Chem., Int. Ed. Engl. **61**, e202200983 (2022).

⁹⁵G. M. Verkhivker, S. Agajanian, D. Y. Oztas, and G. Gupta, J. Chem. Inf. Model. 61, 5172 (2021).

⁹⁶Z. H. Foda et al., Nat. Commun. 6, 5939 (2015).

- ⁹⁷T. K. Johnson *et al.*, Angew. Chem., Int. Ed. Engl. **60**, 20196 (2021).
- ⁹⁸W. Jahnke *et al.*, Angew. Chem., Int. Ed. Engl. **61**, e202117276 (2022).
- ⁹⁹T. K. Johnson *et al.*, Angew. Chem., Int. Ed. Engl. **61**, e202209518 (2022).
- ¹⁰⁰B. Oruganti *et al.*, J. Biol. Chem. **298**, 102238 (2022).

¹⁰¹M. A. Seeliger *et al.*, Cancer Res. **69**, 2384 (2009).

¹⁰²M. A. Seeliger *et al.*, Structure **15**, 299 (2007).

- ¹⁰³K. K. Nishikawa *et al.*, Nat. Commun. **12**, 5562 (2021).
- ¹⁰⁴M. Leander, Z. Liu, Q. Cui, and S. Raman, eLife 11, e79932 (2022).
- ¹⁰⁵J. W. McCormick *et al.*, Elife **10**, e68346 (2021).